International Journal of
Interactive Multimedia and **Artificial Intelligence**

December 2017, Vol IV, Number 6, ISSN: 1989-1660

"The real question is, when will we draft an artificial intelligence bill of rights? What will that consist of? And who will get to decide that?"
Gray Scott

http://ijimai.unir.net

INTERNATIONAL JOURNAL OF INTERACTIVE MULTIMEDIA AND ARTIFICIAL INTELLIGENCE
ISSN: 1989-1660 –VOL. 4, NUMBER 6

Editor's Note

The International Journal of Interactive Multimedia and Artificial Intelligence provides an interdisciplinary forum in which scientists and professionals can share their research results and report new advances on Artificial Intelligence and Interactive Multimedia techniques. The research works presented in this regular issue are based on various topics of interest, among which are included: nature inspired optimization algorithms, multi-agent systems, fast motion estimation, handwritten recognition, supervised and unsupervised machine learning methods, or web mining.

The fields of application are diverse: e-commerce, computer security, 2.0 enterprises, decision making in business, online banking systems, video compression, user experience evaluation, accessibility for Deaf people, or wireless networks, which is the domain of the first article of this issue. This first article is authored by Kaur and Arora [1], who compare the performance of nature inspired optimization algorithms when solving the problem of localization of sensor nodes in wireless sensor networks. They show the behavior of the algorithms considering values of relevant parameters not previously analyzed. Specifically, they apply Flower Pollination Algorithm (FPA), Firefly Algorithm (FA), Particle Swarm Optimization (PSO) and Grey Wolf Optimization (GWO) algorithm, finding that FPA shows better localization accuracy.

Idrus et al. [2] present a framework for decision making applied to the construction domain problem, whose resolution involves the proposal of solutions, negotiation and conflict resolution. They focus on the first part of the process, proposing an algorithm for software agents to generate solutions and rank them to facilitate subsequent negotiation. They apply the algorithm to a validation scenario and show that the algorithm ranks solutions according to each stakeholder position, but conflicts between the different stakeholders exist, which makes necessary to develop a conflict resolution algorithm for the last phase of the process.

With the aim of facilitating industrial site selection for decision-makers, Taibi and Atmani [3] also provide a ranking model through the combination of Fuzzy Analytic Hierarchy Process (FAHP), Geographic Information System (GIS) and decision rules. The model provides a rank for each zone that is an index that allows to optimize the use of zones in the future.

Arora et al. [4] propose a hybrid technique for fast motion estimation, a key process in video compression. Simulation results are compared with results obtained with other existing techniques. The comparison shows that the proposed solution increases video quality and search efficiency and reduces the computation required to estimate the motion vectors.

A dissertation on handwriting recognition is presented by Boulid et al. in next article [5]. The recognition task bases on the multi-agent systems paradigm and it is inspired by the mechanisms the human reader applies while reading. Authors focus on Arabic handwritten documents, obtaining promising results. Continuing with this line, Souhar et al. [6] explore methods based on the watershed transform technique to segment text lines in handwritten Arabic documents. They compare the methods and discuss about the possible reasons that justify the results.

In order to prevent malicious mobile agents attacking a mobile agent platform, Bagga et al. [7] propose the use of machine learning algorithms to detect those unknown malicious agents. They consider an n-gram representation of the agent, which is used as feature for the classification process undertaken by commonly used algorithms, such as Naïve Bayesian or J48 Decision Tree. The different algorithms are compared via extensive experiments done with a benchmark dataset involving malware and benign traces, probing the suitability of supervised machine learning for the detection of malicious agents.

Related also to the field of computer security, Harish and Kumar [8] present a network anomaly detection method based on fuzzy clustering, which, as unsupervised learning method, has the advantage of better detecting unknown attacks when compared with supervised ones. They apply different techniques to mitigate disadvantages of unsupervised methods such as the higher false alarm rate. They compare the performance of this method with other clustering methods finding that it outperforms the other ones.

In the field of semantic web, Anoop and Asharaf [9] propose a method guided by topic modeling for extracting concepts and relationships from unstructured e-commerce product descriptions. The described evaluation shows that their method outperforms other existing techniques, some guided also on topic modelling. Besides, authors explore the possibility of creating a product knowledge base, which can potentially serve for product discovery experience for customers.

The article authored by Bader et al. [10] analyses the state of research on heuristics used for the evaluation of user experience, contributing to reduce the existing research gap in that topic. From the analysis done, they propose a mapping of heuristics with the different dimensions of the User Experience Questionnaire, a widely used questionnaire to measure the user experience with products. Moreover, they also provide a set of quality criteria for heuristics. The article concludes with some recommendations about the convenient heuristics when it is intended to apply heuristic evaluation for measuring user experience.

With the same focus on user experience, the article of Schrepp et al. [11] describes the design and evaluation of a short version of the above mentioned User Experience Questionnaire. The aim is to propose an alternative short questionnaire for those scenarios in which the use of the longer one is not practical.

Reguieg and Taghezout [12] apply an enterprise 2.0 project to help organizations improve business processes, based on multi-agent systems. They implement a collaborative environment, a social network, for employers to share diagnosis and fault repair procedures. Besides, they propose a coordination protocol that manages interaction between agents.

In addition, in order to also improve processes in companies, specifically the processes in bank operation centers, Serengil and Ozpinar [13] propose a hybrid multi stage approach for workforce planning. The approach is based on supervised and unsupervised learning algorithms. Expected workload is predicted via a neural network while employees are grouped via k-means clustering.

Next article, authored by Farhane et al. [14] describes a robust adaptive fuzzy neural network sliding mode (AFNNSM) control design for a variable speed wind turbine. A fuzzy neural network is used to approximate the model function, which provides a better description of the plant. In order to optimize the learning rate of backpropagation algorithm used to train the neural network, a particle swarm optimization algorithm is also used. Through simulations, the authors show the effectiveness of the method, with the trajectory tracking errors converging to zero.

Last article corresponds to an interdisciplinary research done by Pérez et al. [15] whose main result is a system that translates video subtitles in oral language to SignWriting, a method of writing Sign

DOI: 10.9781/ijimai.2017.09.003

Language. This system complements a platform that automatically provides accessible web content for Deaf people and that is prepared to be extended to satisfy the needs of other people. A first prototype of the video translator has been tested and results in usability and accessibility tests show that this tool can enhance the accessibility of video content available on the Web for Deaf people.

Dr. Elena Verdú

References

[1] R. Kaur and S. Arora, "Nature Inspired Range Based Wireless Sensor Node Localization Algorithms", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 7-17, Dec. 2017.

[2] A. Idrus, M. A. Mahmoud, M. S. Ahmad, A. Yahya, and H. Husen, "A Solution Generator Algorithm for Decision Making based Automated Negotiation in the Construction Domain", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 18-23, Dec. 2017.

[3] A. Taibi and B. Atmani, "Combining Fuzzy AHP with GIS and Decision Rules for Industrial Site Selection", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 60-69, Dec. 2017.

[4] S. M. Arora, K. Khanna, and N. Rajpal, "A Novel Hybrid Approach for Fast Block Based Motion Estimation", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 24-30, Dec. 2017.

[5] Y. Boulid, A. Souhar, and M.E. Elkettani, "Multi-agent Systems for Arabic Handwriting Recognition", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 31-32, Dec. 2017.

[6] A. Souhar, Y. Boulid, E. Ameur, and M. M. Ouagague, "Segmentation of Arabic Handwritten Documents into Text Lines using Watershed Transform", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 96-102, Dec. 2017.

[7] P. Bagga, R. Hans, and V. Sharma, "N-grams Based Supervised Machine Learning Model for Mobile Agent Platform Protection against Unknown Malicious Mobile Agents", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 33-39, Dec. 2017.

[8] B. S. Harish and S.V.A. Kumar, "Anomaly based Intrusion Detection using Modified Fuzzy Clustering", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 54-59, Dec. 2017.

[9] V. S. Anoop and S. Asharaf, "A Topic Modeling Guided Approach for Semantic Knowledge Discovery in e-Commerce", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 40-47, Dec. 2017.

[10] F. Bader, E.-M. Schön, and J. Thomaschewski, "Heuristics Considering UX and Quality Criteria for Heuristics", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 48-53, Dec. 2017.

[11] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S)", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 103-108, Dec. 2017.

[12] S. Reguieg and N. Taghezout, "Supporting Multi-agent Coordination and Computational Collective Intelligence in Enterprise 2.0 Platform", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 70-80, Dec. 2017.

[13] S. I. Serengil and A. Ozpinar, "Workforce Optimization for Bank Operation Centers: A Machine Learning Approach", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 81-87, Dec. 2017.

[14] N. Farhane, I. Boumhidi, and J. Boumhidi, "Smart Algorithms to Control a Variable Speed Wind Turbine", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 88-95, Dec. 2017.

[15] E. V. Pérez, B. C. P. G-Bustelo, M. A. M. Sánchez, and R. G. Crespo, "A System to Generate SignWriting for Video Tracks Enhancing Accessibility of Deaf People", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 6, pp. 109-115, Dec. 2017.

TABLE OF CONTENTS

OPEN ACCESS JOURNAL

ISSN: 1989-1660

Nature Inspired Range Based Wireless Sensor Node Localization Algorithms

Ranjit Kaur*, Sankalap Arora

DAV University, Jalandhar, Punjab (India)

Received 1 April 2016 | Accepted 10 January 2017 | Published 30 March 2017

unir
LA UNIVERSIDAD EN INTERNET

Abstract

Localization is one of the most important factors highly desirable for the performance of Wireless Sensor Network (WSN). Localization can be stated as the estimation of the location of the sensor nodes in sensor network. In the applications of WSN, the data gathered at sink node will be meaningless without localization information of the nodes. Due to size and complexity factors of the localization problem, it can be formulated as an optimization problem and thus can be approached with optimization algorithms. In this paper, the nature inspired algorithms are used and analyzed for an optimal estimation of the location of sensor nodes. The performance of the nature inspired algorithms viz. Flower pollination algorithm (FPA), Firefly algorithm (FA), Grey Wolf Optimization (GWO) and Particle Swarm Optimization (PSO) for localization in WSN is analyzed in terms of localization accuracy, number of localized nodes and computing time. The comparative analysis has shown that FPA is more proficient in determining the coordinates of nodes by minimizing the localization error as compared to FA, PSO and GWO.

Keywords

Wireless Sensor Network, Localization, Flower Pollination Algorithm, Particle Swarm Optimization, Firefly Algorithm, Grey Wolf Optimization.

DOI: 10.9781/ijimai.2017.03.009

I. Introduction

Wireless Sensor Network (WSN) consists of independent similar or diverse types of nodes that monitor the environment. These tiny, autonomous, portable and economical nodes sense the data and pass it to designated location or the base station by using wireless ad-hoc network technique. Each sensor node has a transducer, power supply, microcomputer and transceiver to record and process the sensory data [1]. WSN is being used for many applications like industrial process monitoring, battlefield, forest fire detectors, natural disaster prevention, traffic monitoring, etc. [2]. Most of the applications need location information of the sensor nodes for tracking and monitoring [3].

In WSN, sensor nodes sense and report the events of interest which can be examined when the position of target nodes reporting the event is known. The information gathered at sink node will be in vain without localization information of sensor nodes. The estimation of the position of sensor nodes is one of the important issues of the WSN and is known as localization problem [4]. Localization can be defined as determining the coordinates of unknown nodes called as target nodes using the position of known nodes called as anchor nodes or beacons [5] based on the measurements such as Time of arrival (TOA), Time difference of arrival (TDOA), Angle of arrival measurement (AOA), etc. [6]. The localization issue can be resolved by deploying each node with Global Positioning System (GPS) but this is not preferred due to size, cost and power factors [7]. Moreover, GPS has restricted functionality as it cannot work indoor and underwater. So, an efficient alternative is required for the localization. The non-GPS based localization algorithms can be used which is categorized into range based and range free algorithms. Range based localization algorithms uses point to point distance estimation or angle based estimation between sensor nodes whereas range free localization algorithms do not require range information between target node and anchor node but depends on the topological information. The former provides more accuracy as compared to range free localization algorithms.

The range based localization of nodes has two phases – ranging phase and position estimation phase. In the ranging phase, each target node measures its distance from the anchor nodes using the strength of received signal or the signal propagation time. Accurate measurement of distance is not possible due to noise. There is a noisy range measurement irrespective of the ranging method used whereas in position estimation phase, the information acquired from ranging phase is used to determine position of target node. It can also be estimated using geometric approach or by using an optimization algorithm.

The optimization algorithms are really effective in solving NP-Hard problems like Traveling salesman problem, decision subset sum problem, localization, etc. Localization problem is considered as an optimization problem due to size and complexity factors. The analytical methods of optimization like linear programming takes more computation time for solving optimization problems and increase the complexity as the size of problem increases [8]. This propelled to use nature inspired optimization algorithms for WSN as these are robust and effective [9]. These algorithms became popular from the last decade as they can easily adjust to frequently changing environment and have high efficiency [10]. The various algorithms like Particle Swarm Optimization (PSO) [11], Firefly Algorithm (FA) [12] [13], Genetic Algorithm (GA) [14], Grey Wolf Optimization (GWO) [15], Flower pollination Algorithm [16], etc. have been used to determine positions of target nodes. The objective of the various optimization algorithms in WSN localization is to minimize the position estimation

* Corresponding author.
E-mail address: kaurranjit2212@gmail.com

error. These nature inspired algorithms really performed well on benchmark functions and localization problem [17]. The work on FPA based node localization algorithm [18] does not provide comparison of two or more localization algorithms with FPA. The distance between target node and an estimated node for each target node has not been reported. In order to provide a fair comparison, it is necessary to show the behavior of the aforementioned techniques considering other values of the parameters involved in the considered algorithms which has not been done yet.

In this paper, the application of FPA, PSO, FA and GWO for the optimal location estimation of sensor nodes is analysed. All the aforementioned factors are considered for the comparison and analysis. These range based localization algorithms are compared with each other to determine the proficient algorithm which performs better to solve localization issue. Factors like transmission range, number of anchor nodes and number of iterations affecting the localization error are considered and analysed for each localization algorithm and compared graphically. All the localization algorithms are analysed in 36 trail runs by changing number of anchor nodes and target nodes in sensor field in terms of localized nodes, localization accuracy and computation time.

The organization of the rest of the paper is as follows: Section II includes literature survey on WSN. Section III describes the meta-heuristic optimization algorithms. Section IV present the localization based on optimization algorithms. In section V, simulation results and comparative study are given and discussed. In section 6, a conclusion along with the future work is presented.

II. Related Work

Localization has become an active research topic in WSN in recent years as exact localization information is really desirable for the performance of WSN. This problem is approached with different methods by the researchers. The localization with Ad-hoc Positioning System (APS)-distributed method in an ad-hoc network is developed by Niculescu [19]. The principle of APS is similar to the GPS but it extends the potential of GPS to non-GPS sensor nodes as position information of anchor nodes is passed to all sensor nodes in the wireless network. For each target node, minimum three anchor nodes are necessary to execute the triangulation or trilateration to determine its location. Another algorithm was proposed by Savarese [20] to improve Niculescu's method which consists of two phases-hop terrain and refinement. Hop terrain phase is similar to the APS. The refinement phase uses an iterative procedure for the accuracy of location of each node which is enhanced by estimating the least square distances from the neighbouring sensor nodes [21]. The sensor nodes in the sensor field work collectively to determine the location but the error is accumulated in the network. To avoid the accumulation, the kalman filter along with the least square estimation method is used to estimate the co-ordinates of location simultaneously as proposed by Savvides [22]. Localization problem is approached with the convex optimization which is based on semi-definite programming [23]. The technique of semi-definite program is extended further to non- convex inequality constraints by Prtaik [24]. Tzu-Chen Liang extends the Pratik's method to apply gradient search method [25]. It uses technique of data analysis called multidimensional scaling (MDS) [26]. The range based anchor less localization algorithms are discussed in [27-29].

The localization problem can be considered as an optimization problem. The optimization methods use population based stochastic method to evaluate the location of nodes by minimizing the mean square error [30]. Simulated annealing method was proposed to locate the position of nodes but this doesn't provide good results. An improved simulated annealing method was proposed which has two phases. In first phase, the location of target node is estimated whereas in second phase, the optimization is executed on those nodes that may have flip ambiguity problem [31]. The localization algorithms based on GA are proposed in [14], [32], [33]. To minimize the localization error, localization algorithm based on PSO is proposed [30]. The Biogeography Based optimization (BBO) and H-best particle swarm optimization (HPSO) performs better in terms of accuracy and localized nodes as described in [34]. FPA based localization is also proposed to solve the localization issue [18].

III. Nature Inspired Algorithms

Nature inspired algorithms mimics' nature to solve various hard and complex problems as nature exhibits flexible, robust, diverse and dynamic phenomenon. The nature inspired becomes popular as it can easily adjust to frequently changed environment and the conventional or traditional methods were inefficient. These optimization algorithms really perform well to solve the optimization problems like localization issue, traveling salesman problem, etc. The various algorithms like PSO, FA, FPA, GWO and GA. have been applied to solve the localization problem in WSN. The algorithms which are used for the analysis and comparison are:

A. Flower Pollination Algorithm

FPA is a nature inspired algorithm proposed by Xin-She Yang in 2012 [35]. It is inspired from the natural process 'pollination of flowers'. This metaheuristic algorithm has evolutionary characteristics and its convergence rate is relatively high as compared to other nature inspired algorithms [36].

Pollination is an intriguing process in which pollen grains are transferred from the anther to the stigma with the help of pollinators for reproduction. It has two major types: biotic and abiotic pollination. Biotic requires help of some living organisms like insects, bats etc. to transfer pollen grains whereas the latter are dependent on wind and water for pollination. There is a process known as flower constancy in which some pollinators visit specific species of flowers bypassing others.

Pollination can be attained by two ways i.e. self-pollination and cross-pollination. Self-pollination is defined as the reproduction of flower by the transfer of pollen grains from the same or different flower of the same plant species whereas in cross-pollination, pollinators like birds, bees, bats etc. travel a long distance for pollination and follows Levy flight behaviour in which the step length obeys Levy distribution.

There are four rules which have been derived for FPA based on the characteristics of pollination which are:

1. Global pollination process is attained by considering biotic and cross pollination because various pollinators performs levy flights.
2. Local pollination is attained through abiotic and self- pollination.
3. Flower constancy is defined as the probability of reproduction which directly depends on the similarity of two flowers which are involved.
4. Switch probability $p \in [0, 1]$ helps in controlling local pollination (exploitation) and global pollination (exploration). In overall pollination activities, local pollination can have a value of p in significant fraction due to the various factors like physical proximity, wind etc.

These rules are formulated into mathematical equations to have position updating formulas. The two most important steps of FPA are local pollination and global pollination. In global pollination, pollinators carry the pollen grains to the different flowers by traveling a long distance as birds, insects [37]. Through this, exploration can be attained and reproduction of the fittest is ensured.

The first and third rule can be mathematically represented in Eq. (1).

$$x_i^{t+1} = x_i^t + L(x_i^t - g^*) \quad (1)$$

Here x_i^t is pollen i at iteration t or solution vector and g^* is the global best value in every generation. L is the most important parameter for pollination as the various insects use Levy flight to move over a long distance for pollination [38]. Thus, we draw $L > 0$ from levy distribution which is valid for large steps i.e. $s > 0$ [39] and it can be represented as:

$$L \sim \frac{\lambda \Gamma(\lambda) \sin(\frac{\pi\lambda}{2})}{\pi} \frac{1}{s^{(1+\lambda)}}, \quad (s > s_0 > 0) \quad (2)$$

The local pollination takes place by the transfer of pollen grains with the help of abiotic pollinators from one flower to another. The second rule and flower constancy can be represented mathematically as:

$$x_i^{t+1} = x_i^t + \epsilon (x_j^t - x_k^t) \quad (3)$$

where x_i^t and x_k^t are pollens from distinct flowers of the same plant species at iteration t. It imitates the flower constancy in restricted neighbourhood which helps to attain convergence quickly. If x_k^t and x_j^t are from same population and ϵ is drawn from a uniform distribution [0, 1], then it becomes a local random walk. Switch probability p helps to switch from local to global pollination and vice-versa. The pseudo code of the FPA is given in Fig. 1.

```
Objective f(x_i), i = (1,2,3,...d)^t
  Initialize the flower population with n flowers
  Find the best flower g* in the population.
  Set switch probability p ϵ [0,1]
   while (t < Max number of generations)
    for i =1:n
      if rand < p,
        Draw Levy flight using Eq.(1)
        Do global pollination using Eq.(2)
      else
        Draw ϵ from uniform distribution in [0,1]
        Do local pollination using Eq.(3)
      end if
       Update the fitter solutions in existing population
    end for
      Find the best solution in the population
  end while
```

Fig. 1. Pseudo code of FPA.

B. Firefly Algorithm

FA was proposed by Xin she Yang in 2009 and it was mainly inspired from the flashing qualities or characteristics of fireflies [40]. The fireflies get attracted to each other despite of their sex. Attractiveness between fireflies depends directly on the brightness due to which less luminous fireflies will be attracted towards brighter firefly. But attractiveness decreases with an increase in distance between them [41]. If there is no brighter firefly available, then firefly will move randomly in the search space. The brightness of the firefly is evaluated by the objective function which is to be optimized [42]. The brightness determines the attractiveness between fireflies corresponding to the objective function. FA is mainly based on two important factors i.e. change in light intensity and formulation of attractiveness [43]. The intensity of light I changes with distance r having a constant light absorption coefficient which is described mathematically as follows:

$$I = I_o e^{-\gamma r} \quad (4)$$

Here I_o represents light intensity at initial value. The attractiveness between fireflies is relative which is seen or judged by fireflies and it varies with distance r between two fireflies. Thus, the attractiveness β is defined as follows which is corresponding to the light intensity judged or seen by fireflies.

$$\beta = \beta_o e^{-\gamma r^2} \quad (5)$$

Here the β_o represents attractiveness between fireflies when distance r is 0. The Cartesian distance method is used to calculate distance between two fireflies. The firefly i move towards the brighter firefly j which is determined with the help of Eq. (6).

$$x_i = x_i + \beta_o e^{-\gamma r_{ij}^2}(x_j - x_i) + \alpha \epsilon_i \quad (6)$$

The second part of the equation represents the attractiveness and the third part is for randomization where the parameter α lies in the range of [0, 1] which helps in randomization and ϵ_i is a random variable whose value is drawn from Gaussian distribution. The pseudo code for this algorithm is given in Fig. 2.

```
Objective f(x_i), i = (1,2,3,...d)^t
Initialize the population of fireflies x_i (i=1,2,3...n)
Set light absorption coefficient γ
While (t < Max number of generations)
 for i =1: n fireflies
  for j =1: n fireflies
   Light intensity I_iat x_i is determined by f(x_i)
   if (I_i>I_j)
    Move firefly i towards j in d dimensions.
   end if
   Attractiveness changes with distance r using exp[-γr]
   Evaluate new solutions and update the light intensity
  end for j
 end for i
Rank the fireflies according to light intensity and
determine the best firefly
end while
```

Fig. 2. Pseudo code of FA.

C. Particle Swarm Optimization

The PSO algorithm was proposed by Kennedy and Eberhart [44]. This algorithm is mainly inspired from the behaviour of birds flocking in nature. The flock of birds communicate with each other while migrating to the destination and find the bird at best position. Each bird in the flock moves towards best position with velocity dependent on the current position of the bird and then, they explore the search space from their new positions. The process is reiterated until they reach their destination [45].

In this algorithm, the social interaction and intelligence of birds are involved. The birds grasp knowledge from their own experience as well as the experience of other birds [46]. Each particle or bird possesses three values i.e. the present position (X_i), the local best value

(P_i) and their velocity (V_i). The objective function helps to find out the best particle's position (P_g). Every particle in the flock updates their velocity with respect to best particle using following mathematical formula given in Eq. (7).

$$New\ V_i = \omega * V_i + c_1 * rand(\ ) * (P_i - X_i) + c_2 * Rand(\ ) * (P_g - X_i) \quad (7)$$

Here rand ( ) and Rand ( ) are two random functions whose value lies in the range [0, 1]. The c_1 and c_2 parameters are acceleration constants or learning factors and w represents inertia weight which is mainly used to control the impact of previous velocities of particle on current velocity. The second part of the equation is used to compare the particle's current position to its own local best position whereas third part compares the particle's position to the global best particle. The position of a particle using new velocity is updated with the help of Eq. (8).

$$New\ X_i = current\ X_i + New\ V_i \quad (8)$$

The value of V_i lies in the range of user specified values of Vmax, i.e., $V_{max} \geq V_i \geq -V_{max}$ to control the effect of change in particle's velocity. The pseudo code of PSO algorithm is given in Fig. 3.

```
Objective f(x_i), i = (1,2,3,...d)^t
Initialize the population of particles
Initialize the value of inertia weight w
While (t < Max number of generations)
  for i: n particles
    find local best (pbest) of all particles
  end for
find global best (gbest) as the best fitness of all particles
  for i: n particles
    Calculate the velocity of particle using Eq. (7)
    Update the position of particle using Eq. (8)
  end for
end While
```

Fig. 3. Pseudo code of PSO.

D. Grey Wolf Optimization

Grey Wolf Optimizer (GWO) is a nature inspired algorithm proposed by Mirjalili et al. in 2014 that focuses on social behaviour of grey wolves [47]. This algorithm is inspired from grey wolves that belong to canidae family. It simulates the leadership quality and the hunting behaviour of grey wolves in three steps as tracking, encircling and attacking. Grey wolves consists of 5-12 wolves. Grey wolves live in pack that contains 5-12 wolves. α, β, δ and ω are four types of grey wolves following a strict social hierarchy. α is the dominant wolf among the other grey wolves that makes different decisions which are followed by other submissive grey wolves. β grey wolf is second in the hierarchy after α grey wolf. β grey wolf help the dominant leader α to make decisions about sleeping etc.

Approaching and encircling the prey are behaviour of team hunting followed by the grey wolves' pack which is mathematically modelled in Eq. (9) and Eq. (10).

$$\vec{D} = \left|\vec{C}.\vec{X_p}(t) - \vec{X}(t)\right| \quad (9)$$

$$\vec{X}(t+1) = \vec{X_p}(t) - \vec{A}.\vec{D} \quad (10)$$

Here $\vec{X}$ and $\vec{X_p}$ represents the position vector of grey wolf and prey. Vectors $\vec{A}$ and $\vec{C}$ are depicted with the help of following equations:

$$\vec{A} = 2\vec{a}.\vec{r_1} - \vec{a} \quad (11)$$

$$\vec{C} = 2.\vec{r_2} \quad (12)$$

Here r_1 Here and r_2 are random vectors in range [0, 1] and parameter $\vec{a}$ is linearly decreased from 2 to 0. The best three solutions are saved and further the candidate solutions i.e., grey wolves update their positions accordingly. Social behavior of hunting mechanism is mathematically derived using Eq. (13), (14) and Eq. (15).

$$\begin{aligned} \overrightarrow{D_\alpha} &= \left|\vec{C}.\overrightarrow{X_\alpha} - \vec{X}\right|, \\ \overrightarrow{D_\beta} &= \left|\vec{C}.\overrightarrow{X_\beta} - \vec{X}\right|, \\ \overrightarrow{D_\delta} &= \left|\vec{C}.\overrightarrow{X_\delta} - \vec{X}\right| \end{aligned} \quad (13)$$

$$\begin{aligned} \overrightarrow{X_1} &= \overrightarrow{X_\alpha}(t) - \overrightarrow{A_1}.\overrightarrow{(D_\alpha)}, \\ \overrightarrow{X_2} &= \overrightarrow{X_\beta}(t) - \overrightarrow{A_2}.\overrightarrow{(D_\beta)}, \\ \overrightarrow{X_3} &= \overrightarrow{X_\delta}(t) - \overrightarrow{A_3}.\overrightarrow{(D_\delta)} \end{aligned} \quad (14)$$

$$\vec{X}(t+1) = \frac{\overrightarrow{X_1}+\overrightarrow{X_2}+\overrightarrow{X_3}}{3} \quad (15)$$

At the end, when the last criterion specified will be satisfied, GWO algorithm will get terminated and the best position of α wolf will be considered as the outcome. All the steps are presented in pseudo code which is given in Fig. 4.

```
Initialize the population of grey wolves x_i , (i=1,2,3....n)
Initialize a, A and C
Calculate the fitness of each search agent or wolf
X_α=the best search agent
X_β=the second best search agent
X_δ=the third best search agent
 while (t < Max number of iterations)
    for each search agent
     Update the position of current search agent using Eq.(15)
     end for
   Update a, A and  C
   Calculate the fitness of all search agents.
   Update X_α, X_β, X_δ
 end while
 return X_α
```

Fig. 4. Pseudo code of GWO.

IV. Localization Using Nature Inspired Algorithms

The objective of sensor node localization using nature inspired algorithms is to evaluate the position of the maximum number of target nodes using analytical information about the position of anchor nodes. The localization problem can be formulated as an objective function which is to be minimized using nature inspired algorithm. The overall flowchart of range based distributed localization of sensor nodes using

nature inspired algorithm is shown in Fig. (5).

The following steps are followed to perform node localization of each target node in WSN:

1. M target nodes and N beacons or anchor nodes are deployed randomly in the sensor field. Each anchor node has transmission range R. The localized nodes act as beacon in next iteration to remove flip ambiguity problem.

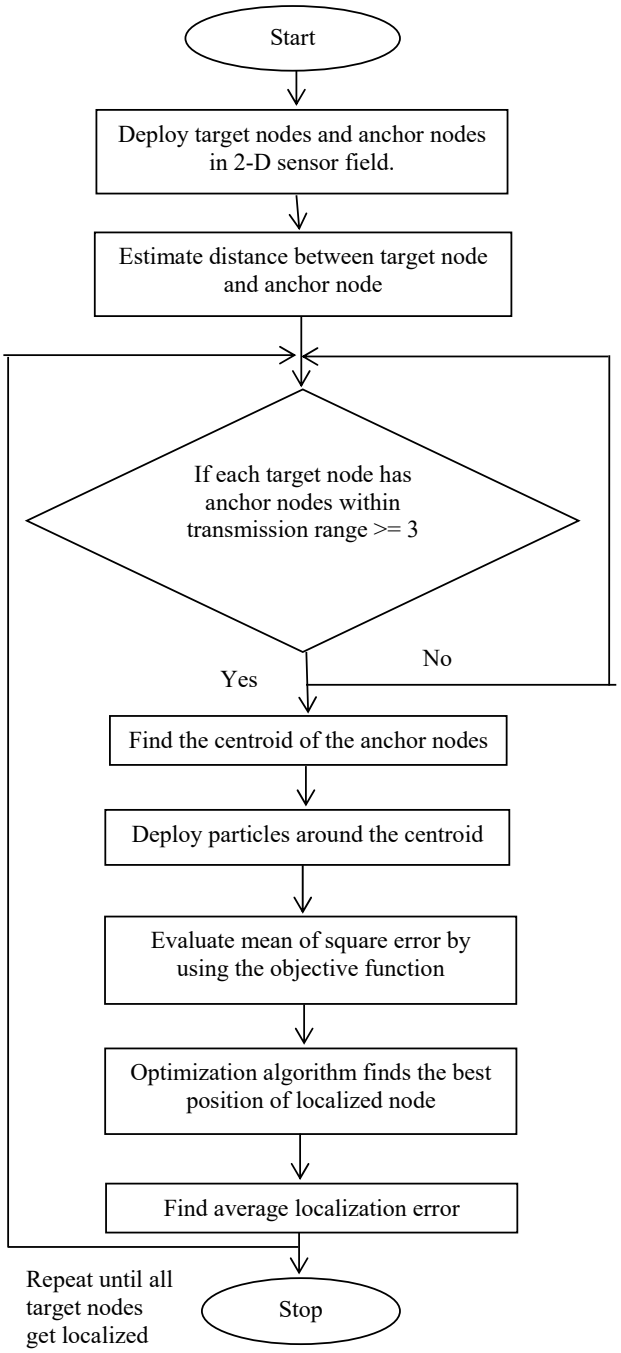


Fig. 5. Flow chart of Sensor node localization using nature inspired algorithm.

2. The distance between each target node and anchor node is measured. There is an additive white Gaussian noise which blurs the measurement. Each target node's distance from every anchor node is measured as $\widehat{d}_i = d_i + n_i$ where d_i is the actual distance and n_i is the measurement noise which has gaussian distributed random value in the range $d_i \pm d_i(\frac{P_n}{100})$ which really affects the performance of the localization algorithm as the localization error increases with the increase in noise. The actual distance can be measured by using the Eq. (16):

$$d_i = \sqrt{(x - x_i)^2 + (y - y_i)^2} \tag{16}$$

Where (x, y) and (x_i, y_i) are the coordinates of target node and anchor node respectively.

3. The target node that has anchor node ≥ 3 within its transmission range is known as localizable node. This target node is localized using an optimization algorithm.
4. For each target node, an optimization algorithm is executed independently. Initialize the particles of the algorithm with the centroid of the anchor nodes (within the transmission range) by using the Eq. (17) as follows:

$$(\mathrm{x_c}, \mathrm{y_c}) = \left(\frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}} \mathrm{x_i}, \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}} \mathrm{y_i}\right) \tag{17}$$

where N is the number of the anchor nodes within transmission range of each target node.

5. The algorithm minimizes the objective function or error for the localization problem in WSN which is given mathematically in Eq. (18).

$$f(x, y) = \frac{1}{N}\left(\sum_{i=1}^{N} \sqrt{(x - x_i)^2 + (y - y_i)^2} - \widehat{d}_i\right)^2 \tag{18}$$

where $N \geq 3$ are the number of anchor nodes within transmission range, (x_i, y_i) is the anchor node within the transmission range and (x,y) is the position of the particles.

6. FPA finds the optimal value (X, Y) after the maximum number of iterations.
7. When the position of all localized nodes get estimated, compute the average localization error to find the localization accuracy by using the Eq. (19) as follows:

$$E_l = \frac{1}{N_L}\sum_{i=1}^{N} \sqrt{(x_i - X_i)^2 + (y_i - Y_i)^2} \tag{19}$$

where (x_i, y_i) are the coordinates of actual node, (X_i, Y_i) are the coordinates of estimated position and N_L is the total number of localized nodes.

8. Repeat steps 2-7 until all the target nodes get localized or no more target nodes can be localized. The localization algorithm's performance is based on the estimated average localization error and the number of un-localized nodes N_{N_L} where $N_{N_L} = \mathrm{M} - N_L$. Localization accuracy is better if value of E_l and N_{N_L} is less.

The number of localized nodes increases as the iteration increments. A node that has been localized can be used as a beacon for the next node. This decreases the problem of flip ambiguity as more references are available for the localized node. However, this increases the computation time.

V. Simulation Results and Discussion

The simulation of WSN localization is conducted using PSO, FA, GWO and FPA in QT Creator 2.4.0. In 2-D sensor field, target nodes and anchor nodes are deployed in the region of 100*100 square units. The transmission range of beacons or anchor nodes is set as 30 units. The performance of each localization algorithm is analysed considering other values of the parameters involved in the considered

algorithms in terms of localized node (N_L), localization error (E_L) and computation time (T_L). The parameter values that provide better localization accuracy are considered for the localization algorithms. The strategic settings and parameter values of PSO, FA, GWO and FPA are discussed below:

A. PSO Based Localization

The performance of PSO based localization with different parameters is analysed and summarized in Table 1. The parameter values that result in less localization error are considered. The inertia weight ω is an important parameter to control the effect of the previous velocities on the existing velocity and the parameters c_1 and c_2 is the learning factors. The number of particles n helps to localize a target node by updating the position according to optimization algorithm. The number of iterations represents the number of times the position is updated to find an optimal solution. The parameter values that works best for localization is considered for PSO based node localization algorithm which are as follows.

- No. of particles or population (n) = 30
- No. of iterations = 100
- Cognitive or social scaling parameter $c_1 = c_2$ = 1.494
- Inertia weight (w) = 0.7

TABLE I. Performance of PSO with Different Tuning Parameters

S. No.	Parameters		N_L	E_L	T_L
1.	w	0.7	100	0.584384	2.287
		0.4	100	**0.650246**	2.206
2.	$c_1 = c_2$	1.494	100	**0.690518**	2.259
		2.0	100	0.694749	2.45
3.	n	20	100	0.721967	2.21
		30	100	**0.717235**	1.997
4.	iteration	100	100	0.698956	2.591
		200	100	**0.567573**	4.335

The PSO based localization algorithm is conducted using above parameters for specified number of iterations to find the localization information. The localized nodes, 50 target nodes and 15 anchor nodes are depicted in Fig. 6.

B. FA Based Localization

Each localized node runs FA to estimate the position of sensor nodes in the sensor field. The performance of FA based localization is evaluated by tuning parameters and summarized in Table 2. The parameter γ is the light absorption coefficient which is really important for the convergence of the algorithm. The value of randomization factor α lies between [0, 1] and β is the initial attractiveness when the distance between two fireflies is 0. The value of n shows the number of fireflies to be deployed in the sensor field to attain the localization information by running the algorithm for specified number of iterations.

The parameters which shows less localization error are considered for localization which are:

- No. of fireflies (n) = 30
- No. of iterations = 100
- Randomization parameter (α) = 0.25
- Absorption coefficient (γ) = 1.0
- Initial attractiveness (β) = 1.0

FA is run for each target node till the specified number of iterations to localize target nodes. FA based localization for 50 target nodes is represented in Fig.7

TABLE II. Performance of FA with Different Tuning Parameters

S. No.	Parameters		N_L	E_L	T_L
1.	α	0.25	100	**0.903181**	6.503
		0.50	100	1.01864	8.737
2.	β	1.0	100	0.528133	7.892
		0.2	100	**0.378845**	9.133
3.	γ	1.0	100	**0.410389**	7.618
		0.7	100	0.430746	7.743
4.	n	20	100	**0.360542**	5.229
		30	100	0.366548	6.763
5.	Iteration	100	100	**0.305937**	8.681
		200	100	0.576672	9.315

C. GWO Based Node Localization

The performance of GWO based node localization algorithm considering the parameters is summarized in Table 3. The vector $\vec{a}$ is a controlling parameter for exploration and exploitation in an algorithm which linearly decreases from 2 to 0. The numbers of grey wolves or particles n helps to determine the localization information of target node by updating their position.

The following parameters are considered for the localization of target nodes using GWO:

- No. of particles (n) = 30
- No. of iterations = 100
- $\vec{a}$ vector = 2 to 0

TABLE III. Performance of GWO with Different Tuning Parameters

S. No.	Parameters		N_L	E_L	T_L
1.	$\vec{a}$	2 to 0	100	0.382065	2.894
2.	n	20	100	0.442893	2.531
		30	100	**0.362075**	2.734
3.	iteration	100	100	0.445948	2.531
		200	100	**0.314762**	2.406

GWO based node localization of 50 target nodes with 15 anchor nodes is illustrated in Fig. 8 which shows target nodes, localized nodes and anchor nodes.

D. FPA Based Node Localization

The performance of the FPA based localization algorithm is analysed and summarized by considering parameters in Table 4.

Table IV. Performance of FPA with Different Tuning Parameters

S. No.	Parameters		N_L	E_L	T_L
1.	p	0.7	100	**0.203275**	2.854
		0.8	100	0.609618	2.445
2.	λ	1	100	0.261968	1.873
		1.5	100	**0.201976**	2.138
3.	n	20	100	0.766729	1.954
		30	100	**0.265837**	3.098
4.	iteration	100	100	**0.244462**	2.225
		200	100	0.685683	3.771

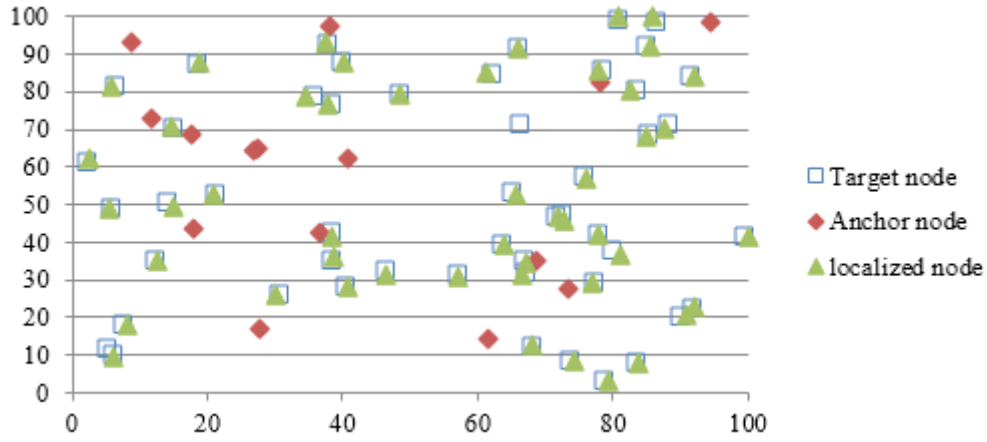


Fig. 6. Node localization using PSO.

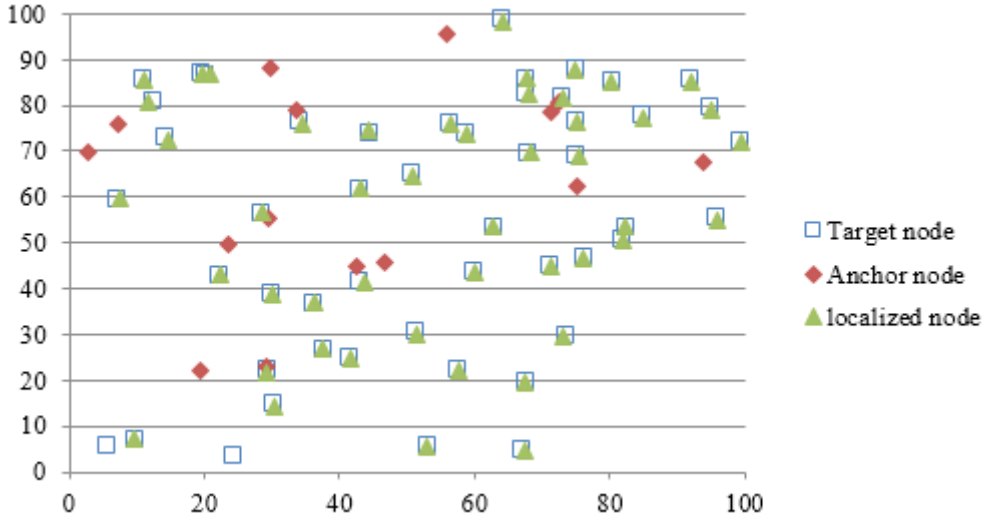


Fig. 7. Node localization using FA.

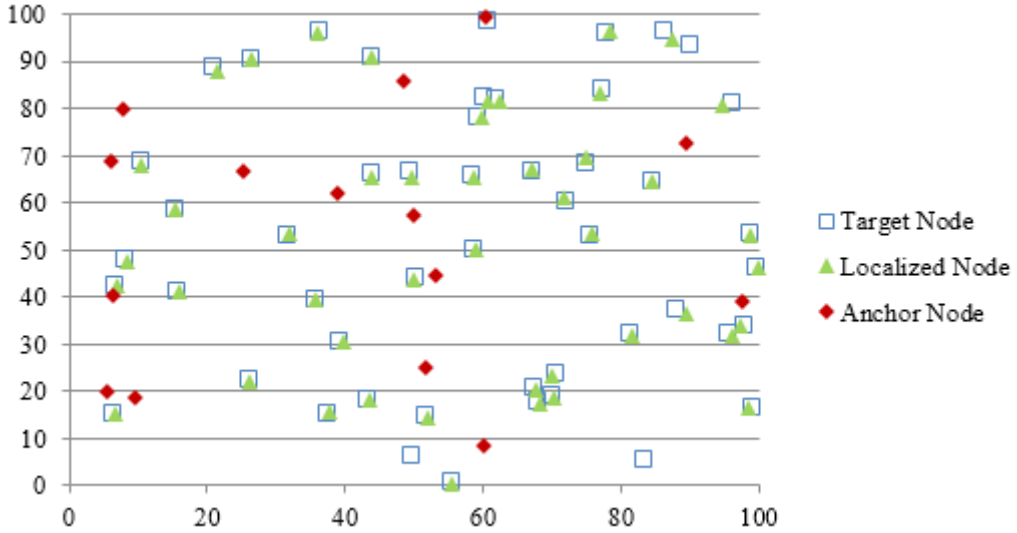


Fig. 8. Node localization using GWO.

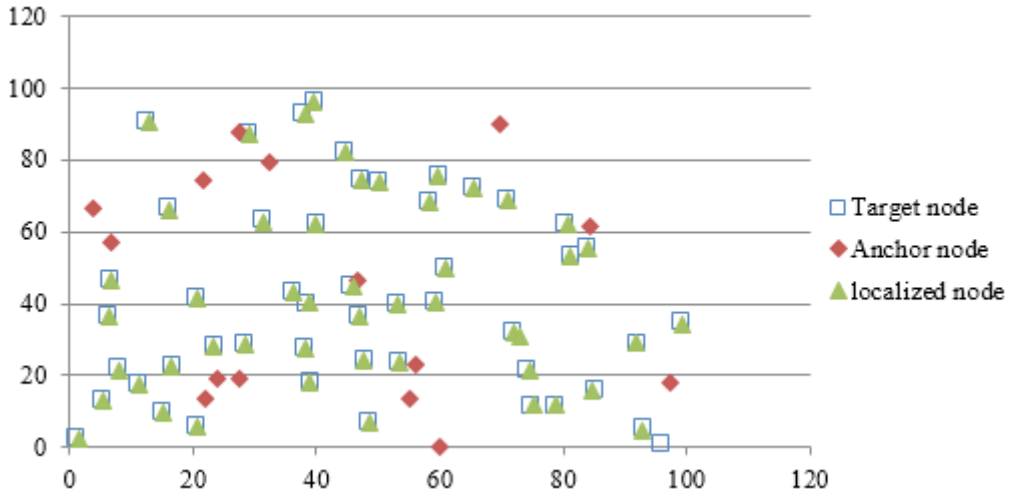


Fig. 9. Node localization using FPA.

The parameter switch probability p is mainly used to control the global pollination and local pollination of the algorithm. The value of λ is really crucial for the Levy flight which plays an important role for global pollination. The number of pollens n is deployed around the centroid of the anchor nodes which update the position with the help of position updating formulas for specified number of times. The following parameters are considered for FPA based localization algorithm:

- No of pollen/flowers (n) = 30
- No. of iterations = 100
- Switch probability (p) = 0.7
- Lambda (λ) = 1.5

50 target nodes estimated by FPA based node localization with the help of 15 anchor nodes is depicted in Fig.9.

The nature inspired optimization algorithms are stochastic in nature. So, same results are not produced in all runs or experiments. Due to this, the results of 30 trial experiments are averaged by using $P_n = 2$, M = 50 and N = 15. The results are summarized in Table 5 which shows that FPA performs better with respect of localization error and un-localized node. The computation time taken by PSO is less than other algorithms. GWO performs less among other algorithms in terms of localization error.

TABLE V. SUMMARY OF RESULTS OF 30 TRIAL RUNS

Algorithms	Mean N_{N_L}	Mean E_L	Computing time (s)
FPA	**4.5**	**0.28374**	0.767
PSO	7.1	0.584231	**0.743**
FA	5.4	0.725323	3.891
GWO	5.0	0.802848	0.832

The initial deployment of sensor nodes is random due to which the localization accuracy, the number of un-localized node and the total computing time will be different for every run of the localization algorithm. The beacons, target node and the position estimated by the algorithms like PSO, FA, GWO and FPA are shown in Fig. 6, 7, 8 and 9 respectively

The transmission range, additive Gaussian noise and the number of anchor nodes are the important parameters to determine the localization error. The performance of localization algorithms is influenced by these parameters. Dependency of the percentage of the localized node on the number of anchor nodes for FA, PSO, GWO and FA based localization algorithms is shown in Fig. 10. The performance of the localization algorithm depends on the density of anchor nodes. It is difficult to locate position of nodes if sufficient number of anchor nodes (N >= 3) are not available. The less number of anchor nodes localize very few number of target nodes. Location estimation accuracy and the percentage of localized nodes increase with the increase in anchor node density. The increase in transmission range of anchor nodes helps in improving the performance as the number of anchor nodes within transmission range of each target node will be more. This will also increase the number of localized nodes. The dependency of localized nodes on the transmission range for FA, PSO, GWO and FPA is shown in Fig. 11. Smaller transmission range localizes very less number of sensor nodes.

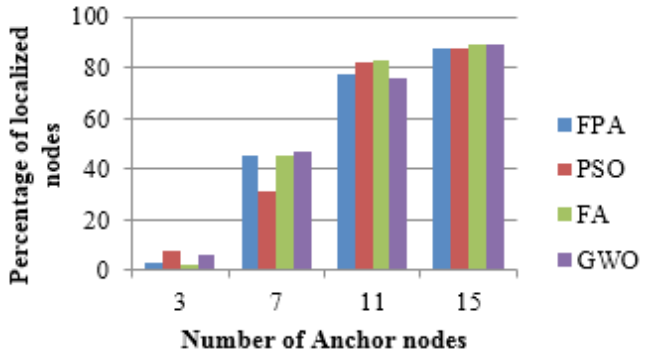


Fig. 10. Percentage of localized nodes depending on the number of anchor nodes

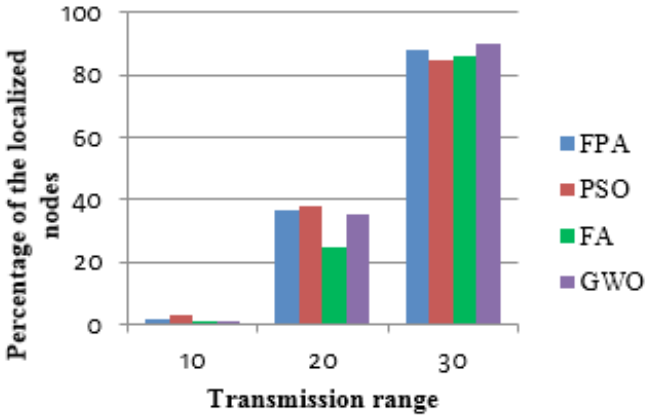


Fig. 11. Percentage of localized nodes depending on the transmission range.

Gaussian additive noise is also an important parameter that has a great impact on localization accuracy. As noise in distance measurement increases, E_l increases which leads to decrease in localization accuracy. Due to this, all the experiments are conducted by considering Gaussian noise $P_n = 2$.

The localization accuracy also improves with the increase in number of iterations as shown in Fig. 12. As the number of iteration progresses, the localization error declines. FPA shows more decline in the error as compared to other two algorithms.

The distance between target node and estimated node for varying number of target nodes is shown in Fig. 13. In this estimated distance for each target node using optimization algorithms is compared which shows FPA based node localization performs better for target nodes which shows its proficiency. Node localization based on FA, PSO, GWO and FPA algorithms by varying number of anchor and target nodes is summarized in Table 6. All the experiments are conducted with different configurations in 36 trails. All the optimization algorithms performed well to estimate the location of nodes in WSN. The FPA based localization algorithms provides less localization error for each number of target nodes whereas PSO estimate the position in less computing time but it has high localization error. It shows better localization accuracy to estimate the position than FA, PSO and GWO localization algorithms in terms of mean square error. But the computing time required by FPA in locating the sensor nodes is more than other optimization algorithms. However, the performance of GWO is comparative less to other algorithms like FA, PSO and FPA in terms of localization accuracy and computation time.

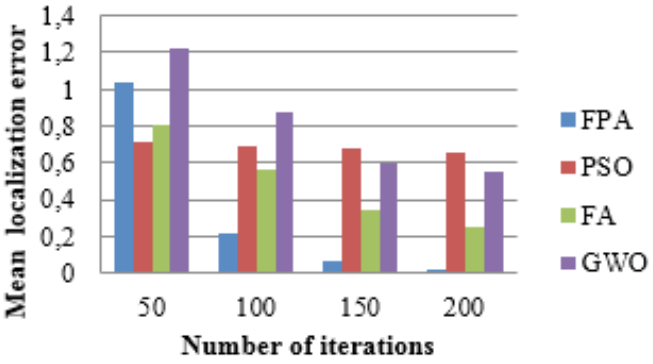


Fig. 12. Error V/S iterations.

TABLE VI. Summary of Results of FPA, FA, PSO and GWO Based Node Localization

Target Node	Anchor Node	Trial	PSO			FA			FPA			GWO		
			N_L	E_l	T_l	N_L	E_l	T_l	N_L	E_l	T_l	N_L	E_l	T_l
25	20	1	22	0.807158	0.36	19	0.335551	1.44	21	**0.133565**	0.563	19	0.65320	0.460
		2	17	0.728214	0.39	20	0.264623	1.44	18	**0.20504**	0.473	23	0.425891	0.480
		3	18	0.79765	0.40	21	0.296398	1.70	20	**0.197211**	0.454	24	0.477140	0.570
50	15	1	48	0.578797	0.74	50	0.505511	2.50	49	**0.325364**	0.668	46	0.421750	0.890
		2	50	0.753254	0.85	49	0.32698	4.42	50	**0.22299**	0.685	44	0.519238	0.970
		3	47	0.587004	0.75	49	0.254824	1.63	48	**0.367064**	0.941	48	0.525000	0.830
75	20	1	75	0.67414	1.31	74	0.703964	2.97	74	**0.169253**	2.046	75	0.471716	1.370
		2	75	0.720123	1.35	75	0.291862	2.73	75	**0.278071**	2.451	75	0.448495	1.380
		3	73	0.771325	1.30	72	0.279126	3.84	73	**0.189639**	1.008	75	0.428363	1.370
100	25	1	100	0.668227	2.49	100	0.779716	5.66	100	**0.171925**	1.506	100	0.3623	2.592
		2	100	0.614843	2.10	100	0.299194	6.33	100	**0.181158**	2.156	100	0.55558	2.590
		3	100	0.608155	2.20	100	0.385758	6.55	100	**0.151736**	2.853	100	0.51877	2.590

N_L = number of localized nodes E_l = localization error T_l = computing time (in seconds)

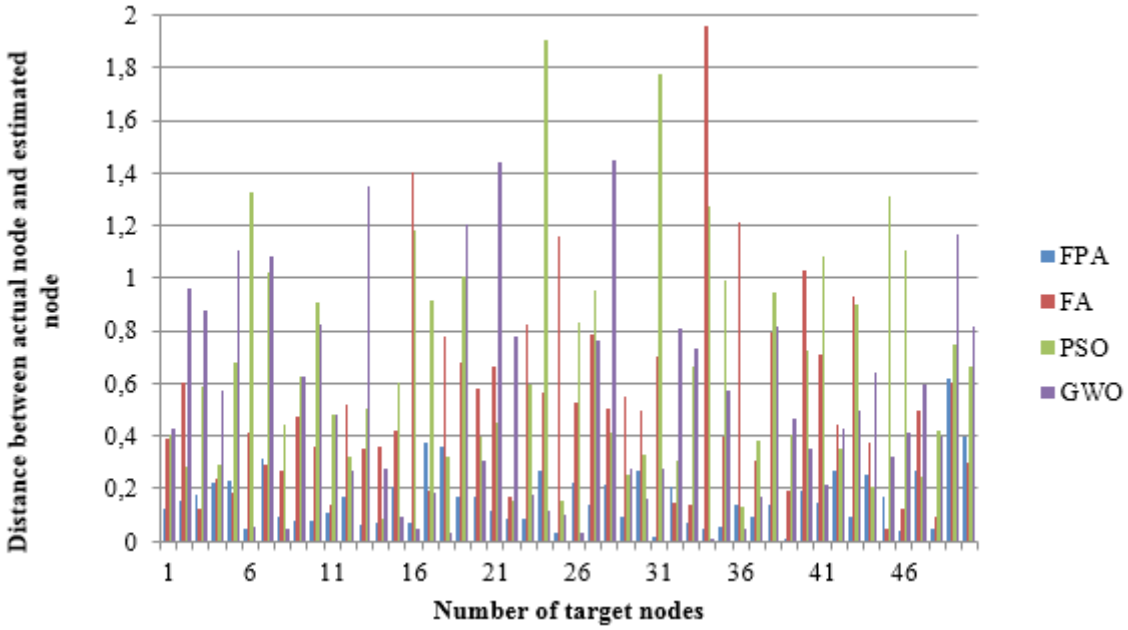


Fig. 13. Distance between actual node and estimated node with different optimization algorithms.

VI. Conclusion

Localization of sensor nodes is really important for the performance of WSN as many applications of WSN require localization information. The main objective of this optimization problem is to minimize the localization error with the help of nature-inspired optimization algorithms. In this paper, node localization using meta-heuristic optimization algorithm like Firefly Algorithm (FA), Particle Swarm Optimization (PSO) and Grey Wolf Optimization (GWO) algorithm is conducted to determine the position of the sensor nodes in WSN. This paper has analysed the localization problem and solved it with different optimization algorithms and provides the summary of results by comparing the algorithm with the each other in terms of localization error, localized nodes and computing time. The FPA based node localization algorithm shows better localization accuracy in estimating the position than other algorithms. GWO performs comparative less with respect to localization error and computation time among the optimization algorithms. These distributed localization algorithms are better than centralized algorithms as the number of transmissions to the sink node is reduced which helps to conserve the energy of sensor nodes. These localization algorithms further can be implemented for centralized method and can be compared with distributed method for analysis. FPA can be hybridized with other optimization algorithm to further minimize the location estimation error. These optimization algorithms can be implemented in 3D scenario to find localization accuracy. The parameter values can be varied to improve the optimization algorithm to improve convergence rate, localization accuracy and computation time.

References

[1] R. V. Kulkarni, A. Förster, G. K. Venayagamoorthy, Computational intelligence in wireless sensor networks: a survey, Communications Surveys & Tutorials, IEEE 13 (1) (2011) 68–96.

[2] I. F. Akyildiz,W. Su, Y. Sankarasubramaniam, E. Cayirci, Wireless sensor networks: a survey, Computer networks 38 (4) (2002) 393–422.

[3] J. Yick, B. Mukherjee, D. Ghosal, Wireless sensor network survey, Computer networks 52 (12) (2008) 2292–2330.

[4] R. V. Kulkarni, G. K. Venayagamoorthy, Particle swarm optimization in wireless-sensor networks: A brief survey, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 41 (2) (2011) 262–267.

[5] J.Wang, R. K. Ghosh, S. K. Das, A survey on sensor localization, Journal of Control Theory and Applications 8 (1) (2010) 2–11.

[6] G. Mao, B. Fidan, B. D. Anderson, Wireless sensor network localization techniques, Computer networks 51 (10) (2007) 2529–2553.

[7] R. G. Crespo, G. G. Fernandez, O. S. Martínez, V. García-Díaz, L. J. Aguilar, E. T. Franco, In premises positioning–fuzzy logic, in: International Work-Conference on Artificial Neural Networks, Springer, 2009, pp. 284–291.

[8] R. V. Kulkarni, G. K. Venayagamoorthy, Bio-inspired algorithms for autonomous deployment and localization of sensor nodes, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 40 (6) (2010) 663–675.

[9] R. V. Kulkarni, G. K. Venayagamoorthy, M. X. Cheng, Bioinspired node localization in wireless sensor networks, in: Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on, IEEE, 2009, pp. 205–210.

[10] J. Meza, H. Espitia, C. Montenegro, R. G. Crespo, Statistical analysis of a multi-objective optimization algorithm based on a model of particles with vorticity behavior, Soft Computing (2015) 1–16.

[11] A. Gopakumar, L. Jacob, Localization in wireless sensor networks using particle swarm optimization, in: Wireless, Mobile and Multimedia Networks, 2008. IET International Conference on, IET, 2008, pp. 227–230.

[12] R. Harikrishnan, V. J. S. Kumar, P. S. Ponmalar, Firefly algorithm approach for localization in wireless sensor networks, in: Proceedings of 3rd International Conference on Advanced Computing, Networking and Informatics, Springer, 2016, pp. 209–214.

[13] S. Arora, S. Singh, A conceptual comparison of firefly algorithm, bat algorithm and cuckoo search, in: Control Computing Communication & Materials (ICCCCM), 2013 International Conference on, IEEE, 2013, pp. 1–4.

[14] Q. Zhang, J. Wang, C. Jin, J. Ye, C. Ma, W. Zhang, Genetic algorithm based wireless sensor network localization, in: Natural Computation, 2008. ICNC'08. Fourth International Conference on, Vol. 1, IEEE, 2008, pp. 608–613.

[15] M. M. Fouad, A. I. Hafez, A. E. Hassanien, V. Snasel, Grey wolves optimizer-based localization approach in wsns, in: 2015 11th International Computer Engineering Conference (ICENCO), IEEE, 2015, pp. 256–260.

[16] M. Sharawi, E. Emary, I. A. Saroit, H. El-Mahdy, Flower pollination optimization algorithm for wireless sensor network lifetime global optimization, International Journal of Soft Computing and Engineering 4 (3) (2014) 54–59.

[17] X.-S. Yang, Nature-inspired metaheuristic algorithms, Luniver press, 2010.

[18] S. Goyal, M. S. Patterh, Flower pollination algorithm based localization f wireless sensor network, in: 2015 2nd International Conference on Recent Advances in Engineering & Computational Sciences (RAECS), IEEE, 2015, pp. 1–5.

[19] D. Niculescu, B. Nath, Ad hoc positioning system (aps), in: Global Telecommunications Conference, 2001. GLOBECOM'01. IEEE, Vol. 5, IEEE, 2001, pp. 2926–2931.

[20] C. S. J. Rabaey, K. Langendoen, Robust positioning algorithms for distributed ad-hoc wireless sensor networks, in: USENIX technical annual conference, 2002, pp. 317–327.

[21] N. Bulusu, D. Estrin, L. Girod, J. Heidemann, Scalable coordination for wireless sensor networks: self-configuring localization systems, in: International Symposium on Communication Theory and Applications (ISCTA 2001), Ambleside, UK, 2001.

[22] A. Savvides, H. Park, M. B. Srivastava, The bits and flops of the n-hop multilateration primitive for node localization problems, in: Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications, ACM, 2002, pp.112–121.

[23] L. Doherty, K. S. Pister, L. El Ghaoui, Convex position estimation in wireless sensor networks, in: INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, Vol. 3, IEEE, 2001, pp. 1655–1663.

[24] P. Biswas, Y. Ye, Semidefinite programming for ad hoc wireless sensor network localization, in: Proceedings of the 3rd international symposium on Information processing in sensor networks, ACM, 2004, pp. 46–54.

[25] T.-C. Liang, T.-C. Wang, Y. Ye, A gradient search method to round the semidefinite programming relaxation solution for ad hoc wireless sensor network localization, Sanford University, formal report 5.

[26] Y. Shang,W. Rum, Improved mds-based localization, in: INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies, Vol. 4, IEEE, 2004, pp. 2640–2651.

[27] S. C˘ apkun, M. Hamdi, J.-P. Hubaux, GPS-free positioning in mobile ad hoc networks, Cluster Computing 5 (2) (2002) 157–167.

[28] U. Ferner, H. Wymeersch, M. Z. Win, Cooperative anchor-less localization for large dynamic networks, in: Ultra-Wideband, 2008. ICUWB 2008. IEEE International Conference on, Vol. 2, IEEE, 2008, pp. 181–185.

[29] A. Youssef, A. Agrawala, M. Younis, Accurate anchor-free node localization in wireless sensor networks, in: Performance, Computing, and Communications Conference, 2005. IPCCC 2005. 24th IEEE International, IEEE, 2005, pp. 465–470.

[30] S. Yun, J. Lee,W. Chung, E. Kim, S. Kim, A soft computing approach to localization in wireless sensor networks, Expert Systems with Applications 36 (4) (2009) 7552–7561.

[31] A. A. Kannan, G. Mao, B. Vucetic, Simulated annealing based wireless sensor network localization with flip ambiguity mitigation, in: Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd, Vol. 2, IEEE, 2006, pp. 1022–1026.

[32] Q. Zhang, J. Huang, J. Wang, C. Jin, J. Ye, W. Zhang, A new centralized localization algorithm for wireless sensor network, in: Communications and Networking in China, 2008. China-Com 2008. Third International Conference on, IEEE, 2008, pp. 625–629.

[33] Q. Zhang, J. Wang, C. Jin, Q. Zeng, Localization algorithm for wireless sensor network based on genetic simulated annealing algorithm, in: Wireless Communications, Networking and Mobile Computing, 2008. WiCOM'08. 4th International Conference on, IEEE, 2008, pp. 1–5.

[34] A. Kumar, A. Khosla, J. S. Saini, S. Singh, Meta-heuristic range based node localization algorithm for wireless sensor networks, in: Localization and GNSS (ICL-GNSS), 2012 International Conference on, IEEE, 2012, pp. 1–7.

[35] X.-S. Yang, Flower pollination algorithm for global optimization, in: Unconventional computation and natural computation, Springer, 2012, pp. 240–249.

[36] X.-S. Yang, M. Karamanoglu, X. He, Flower pollination algorithm: a novel approach for multiobjective optimization, Engineering Optimization 46 (9) (2014) 1222–1237.

[37] X.-S. Yang, Nature-inspired optimization algorithms, Elsevier, 2014.

[38] S. Arora, S. Singh, Butterfly algorithm with levy flights for global optimization, in: Signal Processing, Computing and Control (ISPCC), 2015 International Conference on, IEEE, 2015, pp. 220–224.

[39] I. Pavlyukevich, L´evy flights, non-local search and simulated annealing, Journal of Computational Physics 226 (2) (2007) 1830–1844.

[40] X.-S. Yang, Firefly algorithm, stochastic test functions and design optimisation, International Journal of Bio-Inspired Computation 2 (2) (2010) 78–84.

[41] S. Arora, S. Singh, S. Singh, B. Sharma, Mutated firefly algorithm, in: Parallel, Distributed and Grid Computing (PDGC), International Conference on, IEEE, 2014, pp. 33–38.

[42] S. Arora, S. Singh, Performance research on firefly optimization algorithm with mutation, in: International Conference, Computing & Systems, 2014.

[43] S. Gupta, S. Arora, A hybrid firefly algorithm and social spider algorithm for multimodal function, in: Intelligent Systems Technologies and Applications, Springer, 2016, pp. 17–30.

[44] R. C. Eberhart, J. Kennedy, et al., A new optimizer using particle swarm theory, in: Proceedings of the sixth international symposium on micro machine and human science, Vol. 1, New York, NY, 1995, pp. 39–43.

[45] J. Meza, H. Espitia, C. Montenegro, E. Giménez, R. González- Crespo, Movpso: Vortex multi-objective particle swarm optimization, Applied Soft Computing.

[46] Y. Shi, R. C. Eberhart, Empirical study of particle swarm, in: Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on, Vol. 3, IEEE, 1999.

[47] S. Mirjalili, S. M. Mirjalili, A. Lewis, Grey wolf optimizer, Advances in Engineering Software 69 (2014) 46–61.

Ranjit Kaur

Ranjit Kaur received M.Tech in computer science and engineering from DAV University, Jalandhar, Punjab (India) and B.Tech in Computer Science Engineering from Punjab Technical University, Jalandhar, Punjab (India). Her current research interests includes optimization, nature inspired algorithms and wireless sensor networks.

Sankalap Arora

Sankalap Arora was born on Nov 3, 1988. He received his Bachelor's degree (B.Tech.) and Master's degree (M.Tech.) from Lovely Professional University, Phagwara, Punjab (India) with specialization in Computer Science & Engineering. He is working as an Assistant Professor in Department of Computer Science & Engineering in DAV University, Janadhar, Punjab (India). He has more than 5 years' research and teaching experience. His fields of special interest include Nature Inspired Algorithms, Engineering Design problems and Wireless Sensor Networks. He has published nearly 30 research papers in reputed International Journals and Conferences.

A Solution Generator Algorithm for Decision Making based Automated Negotiation in the Construction Domain

Arazi Idrus[1], Moamin A. Mahmoud[2], Mohd Sharifuddin Ahmad[2], Azani Yahya[1], Hapsa Husen[1] *

[1]National Defence University of Malaysia, Kuala Lumpur (Malaysia)
[2]Universiti Tenaga Nasional, Selangor (Malaysia)

Received 1 December 2016 | Accepted 20 February 2017 | Published 30 March 2017

Abstract

In this paper, we present our work-in-progress of a proposed framework for automated negotiation in the construction domain. The proposed framework enables software agents to conduct negotiations and autonomously make value-based decisions. The framework consists of three main components which are, solution generator algorithm, negotiation algorithm, and conflict resolution algorithm. This paper extends the discussion on the solution generator algorithm that enables software agents to generate solutions and rank them from 1st to nth solution for the negotiation stage of the operation. The solution generator algorithm consists of three steps which are, review solutions, rank solutions, and form ranked solutions. For validation purpose, we present a scenario that utilizes the proposed algorithm to rank solutions. The validation shows that the algorithm is promising, however, it also highlights the conflict between different parties that needs further negotiation action.

Keywords

Intelligent Software Agent, Multi-agent Systems, Agent and Negotiation, Automated Negotiation, Value Management.

DOI: 10.9781/ijimai.2017.03.010

I. Introduction

In build engineering and construction domains, deciding on a new project depends upon a company's strategy. If the strategy is based on a decision by a stakeholder, then it takes a shorter time to decide. However, such decision has no significance in terms of value management, because the decision-making process does not include other experienced stakeholders that have different backgrounds.

A project manager usually cares more about the cost and schedule of a project than the function while a design manager is more concerned about the function than the cost. Thus, for any decision to be made regarding a new project, stakeholders must propose an optimal solution. However, a problem may arise when stakeholders propose many solutions. In such a situation, stakeholders need to negotiate on the proposed solutions and agree on an optimal solution. But the negotiation may not be easy and smooth because when stakeholders possess different backgrounds, often their views about an optimal solution for a particular project are different. Such differences cause conflicts in arriving at a decision. In addition, stakeholders may work at different branches throughout the country or other parts of the world which make a meeting for decision more difficult and costly. While applying Value Management on decision making in the construction domain is useful, it faces communication difficulties between stockholders and conflicting issues that require negotiation.

* Corresponding author.
E-mail addresses: arazi@upnm.edu.my (Arazi Idrus), moamin@uniten.edu.my (Moamin A. Mahmoud), sharif@uniten.edu.my (Mohd Sharifuddin Ahmad), azani@upnm.edu.my (Azani Yahya), hapsa@upnm.edu.my (Hapsa Husen).

We attempt to overcome these difficulties by developing a framework for automated multi-agent negotiation for decision making based on value management in the construction domain. This paper is an extension to our work in the concepts of automated multi-agent negotiation [1, 2, 29].

The framework in [1, 2, 29] enables software agents to conduct negotiations and autonomously make value-based decisions. The framework consists of three main components which are, solution generator algorithm, negotiation algorithm, and finally conflict resolution algorithm. This paper focuses and extends the discussion on the solution generator algorithm that enables software agents to generate solutions and rank them from 1st to nth solution for the next stage of the negotiation operation. The solution generator algorithm consists of three steps which are, review solutions, rank solutions, and form ranked solutions. For validation purpose, we present a scenario that utilizes the proposed algorithm to rank solutions. The validation shows that the algorithm is promising, however, it also highlights the conflict between different parties that needs further negotiation action.

While this work is inspired by the work of Utomo [3], his study is only in conceptual level and lacks a complete negotiation process that aids an agent to interact and negotiate with other agents and respond to its environment and eventually influences its autonomy level in decision making.

II. Related Work

In this section, we discuss two prominent topics of this research which are value management and application of negotiation in multi-agent systems.

Value Management (VM) is defined as "a structured, organized team approach to identify the functions of a project, product, or service that will recognize techniques and provide the necessary functions to meet the required performance at the lowest overall cost" [4]. Utomo et al. [3] defined VM as one of the decision methodologies that include a multi-disciplinary, team-oriented approach to problem solving [5]. Therefore, negotiation plays an important role on VM using a value-based group decision [3]. VM is based on a data collection method from reliable resources and functional requirements to fulfill the needs, wants and desires of customers [3].

The application of VM in decision making has been reported by many researchers [3, 6, 7]. One of the techniques that is relevant to VM is weighting and scoring in which a decision needs to be made in selecting an option from a number of competing options, and the best option is not immediately identifiable [3, 8, 9].

Intelligent software agents have been widely used in distributed artificial intelligence and due to their autonomous, self-interested, rational abilities [11, 12, 13, 14, 15, 16], and social abilities [17, 18, 19, 20], agents are well-suited for automated negotiation on behalf of humans [10]. According to Kexing [10], automated negotiation is a system that applies artificial intelligence and information and communication technology to negotiation strategies, utilizing agents and decision theories.

Numerous research have discussed negotiation on multi-agent systems in various domains [21, 22, 23, 24, 25]. Coutinho et al. [26] proposed a negotiation framework to serve collaboration in enterprise networks to improve the sustainability of interoperability within enterprise information systems. Utomo [3] presented a conceptual model of automated negotiation that consists of a negotiation methodology and an agent-based negotiation. Dzeng and Lin [27] presented an agent-based system to support a negotiation between constructors and suppliers via the Internet. Anumba et al. [28] proposed a collaborative design of light industrial buildings based on multi-agent systems to automate the interaction and negotiation between the design members. Ren et al. [22] developed a multi-agent system representing participants, who negotiate with each other to resolve construction claims.

III. A Conceptual Framework for Value Decision Making based Automated Negotiation

A decision made by an agent goes through several processes. These processes work by gradually reducing candidate solutions of a project until a single solution is reached. Consequently, in this work, the process of nominating a single solution from a set of solutions is called decision-making.

There are three main processes in decision-making for a specific project, which are propose solutions, negotiate solutions and handling conflicting outcomes (conflict resolution).

Figure 1 shows the flowchart of the decision-making process as described above. The process starts when agents receive a new project. The agents first propose solutions in ranked order. They then negotiate these solutions. If they agree upon a single solution, then the decision is made, otherwise, the conflict resolution process takes over to drop the weak and risky solutions. If the outcome of the conflict resolution process is a single solution then the decision is made. Otherwise, the agents negotiate the outcome of the conflict resolution process. Ultimately, one coalition's solution is accepted.

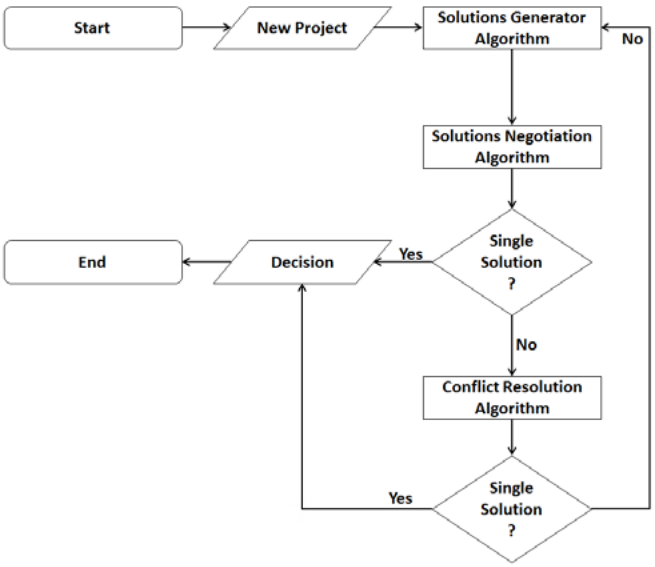


Fig. 1. Decision Making Process.

IV. Solutions Generator

A. Overview and Definitions

In this section, we discuss the preliminary issues in proposing an algorithm for software agents to generate solutions and rank them from 1st to nth solution for the negotiation stage of the operation.

The proposed solution generator algorithm (A_{SG}) is based on two parameters, which are Cost and Function. In real situations, various stakeholders have different level of interest about the cost and function parameters based on their positions and values they uphold. Thus, those stakeholders appraise their solutions based on their interest level on these parameters. For example, in the construction domain, a Design Manager cares more about the function in contrast with a Project Manager who cares more about the cost, while a Facility Manager's interest is in between the Design and Project Managers' interests. Therefore, the Design Manager normally attempts to find a solution that provides high function, whereas the Project Manager normally attempts to find a solution that provides low cost. The Facility Manager attempts to find a moderate solution that provides acceptable cost and function.

Definition 1: A Stakeholder, S, is a person that has an ability and authority to propose solutions for a particular issue and contribute in the decision for that issue's solution. In this work, a software agent represents a stakeholder. If α_S is a stakeholder agent, D is a decision, L is a solution, then,

$$\forall\, \alpha_S\, \forall\, D\, \exists\, L\, (\text{generate}\, (\alpha_S, L) \wedge \text{contribute}\, (\alpha_S, D))$$

which means that for all stakeholder agents and for all decisions, there exist solutions, which stakeholder agents generate and contribute in the decision.

Definition 2: A Single Solution, L_S, is a result of agents' negotiation to solve a particular problem. Initially, each agent ranks the proposed solutions from 1st to nth solution, where n is any natural number. After negotiation, the agents produce a single solution. The agents rank the solutions based on the parameters of cost and function.

Definition 3: A Cost, C, is the price of completing a specific item of a project. A Cost is ranked from low to high based on an original total amount of a project. The minimum (lowest) cost contributes to the highest rank in the solution and vice versa. If C_{MAX} is the maximum

1st (Optimal)	2nd	3rd	4th
C_{MIN} , F_{MAX}	C_{MIN} , F_{MAX-n}	C_{MIN} , F_{MAX-2n}	C_{MIN} , F_{MIN}
5th	6th	7th	8th
C_{MIN+n}, F_{MAX}	C_{MIN+n}, F_{MAX-n}	C_{MIN+n}, F_{MAX-2n} ...	C_{MIN+n}, F_{MIN}
.			
.			
.			
9th	10th	11th	nth
C_{MAX}, F_{MAX}	C_{MAX}, F_{MAX-n}	C_{MAX}, F_{MAX-2n}	C_{MAX}, F_{MIN}

Fig 2. Complete track of cost route.

1st (Optimal)	2nd	3rd	4th
F_{MAX}, C_{MIN}	F_{MAX}, C_{MIN+n}	F_{MAX}, C_{MIN+2n}	F_{MAX}, C_{MAX}
5th	6th	7th	8th
F_{MAX-n}, C_{MIN}	F_{MAX-n}, C_{MIN+n}	F_{MAX-n}, C_{MIN+2n} ...	F_{MAX-n}, C_{MAX}
.			
.			
.			
9th	10th	11th	nth
F_{MIN}, C_{MIN}	F_{MIN}, C_{MIN+n}	F_{MIN}, C_{MIN+2n}	F_{MIN}, C_{MAX}

Fig 3. Complete track of function route.

(highest) cost, C_{MIN} is the minimum (lowest) cost, n is any natural number, then,

$$C_{MIN}^{MAX} = C_{MIN} \quad C_{MIN+n} \quad C_{MIN+2n}, \ldots\ldots, \quad C_{MAX} \tag{1}$$

Definition 4: A Function, F, is the measure of usefulness of an entity in fulfilling its purpose. It is a solution's quality measurement for a specific project. A Function is ranked from high to low based on its useful effect on a solution. The maximum function contributes highly to the solution rank and vice versa. If F_{MAX} is the maximum function, F_{MIN} is the minimum function, then,

$$F_{MAX}^{MIN} = F_{MAX}, \; F_{MAX-n}, \; F_{MAX-2n}, \ldots\ldots, \; F_{MIN} \tag{2}$$

B. Solution Routes Method (SRM)

We identify three routes to generate ranked solutions, as follows:

- **Cost Route (R_C):** This route is preferred by agents (stakeholders) who care more about the cost, e.g. Project Manager. From Eq. 1 and 2 since this route emphasizes more on the cost, then the first rank solution starts with the minimum cost, C_{MIN}, and maximum function, F_{MAX}, as an optimal solution. However, the optimal solution (C_{MIN}, F_{MAX}) is the same for all agents' types e.g. Design Manager, Project Manager and so on, but it depends if such solution exists. If this solution does not exists, then the next optimal solution attempts to keep the cost low with reduced function (C_{MIN}, F_{MAX-n}). This route progresses with decreasing function until the minimum acceptable function is found (C_{MIN}, F_{MIN}). An alternative cost route increases cost to C_{MIN+N} while maintaining the maximum function F_{MAX}, (C_{MIN+N}, F_{MAX}). If no optimal solution is found, the cost cycle is repeated with reduced function. The completed track of cost route is as shown in Figure 2.
- **Function Route (R_F):** This route is preferred by agents that care more about the function, e.g. Design Manager. Since this route emphasizes more on the function, then the first rank solution starts with the maximum function, F_{MAX}, and minimum cost, C_{MIN} as an optimal solution. If the optimal solution is not found, the next optimal solution attempts to maintain the maximum function with increased cost (F_{MAX}, C_{MIN+n}). This route progressively increases the cost until an optimal solution is found or the maximum acceptable cost (F_{MAX}, C_{MAX}) is reached. An alternative function route progressively decreases function to F_{MAX-n} while maintaining the minimum cost C_{MIN}, (C_{MIN+n}, F_{MAX}). If no solution is found, the function cycle is repeated with reduced cost. The complete track of the function route is as shown in Figure 3.
- **Mixed Route (R_M):** This route is preferred by agents that moderately care about both the cost and function, e.g. Facility Manager. This route accepts both earlier mentioned routes' solutions. For example, the two solutions (C_{MIN}, F_{MAX-N}; F_{MAX}, C_{MIN+N}) are acceptable, an agent from this type selects the highest weightage solution (see (3)), which we shall discuss in the next section.

$$R_M = \left(R_C \vee R_F \right) \tag{3}$$

Figure 5 shows the solution routes where the blue line represents the function route, red line represents the cost route, and green line represents the mixed route. The function route takes the horizontal direction, the solutions' rank of which is represented by 1st, 2nd, 3rd in blue color. The cost route takes the vertical direction, the solutions' rank of which is represented by 1st, 2nd, 3rd in red color. The mixed route takes the oblique direction, the solutions' rank of which is represented by 1st, 2nd, 3rd in red and blue colors.

C. The Solutions Generator Algorithm (A_{SG})

The A_{SG} algorithm consists of three steps as shown in Figure 4. We present these steps as follows:

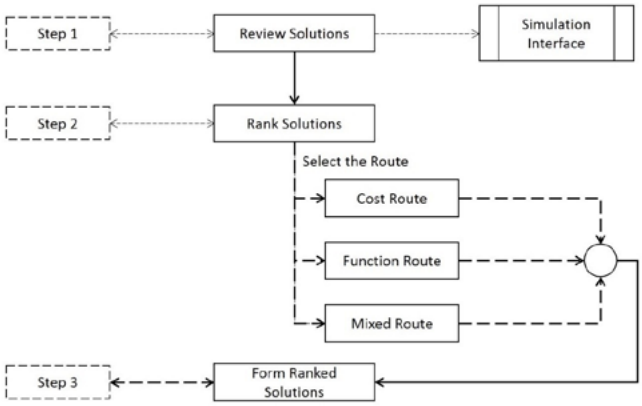


Fig. 4. The Solutions Generator Algorithm (A_{SG}) Process.

Step1: Review Solutions, agents first review solutions.

Step 2: Rank Solutions, each agent selects a suitable route by using the Solution Route Method (SRM) to rank solutions from 1st to nth.

Step 3: Form Ranked Solutions, after each agent has selected the suitable route according to its type and has ranked them, it forms the outcome as follows:

```
agentName (solutionNumber, solutionRank)
```

Algorithm 1 shows the process of Solution Generator Algorithm.

V. Scenario Validation

In this section, we present a scenario to clarify and validate the proposed Solution Generator algorithm. As the proposed algorithm is for the construction domain, we assume the task is building the Roof System. From the literature [3, 4], there are five possible solutions

```
1.  Begin
2.     ∃ α contributes in Decision Do
3.     exploit(α,A_SG),
4.     begin
5.        review(α, solutions)
6.           returns (∃ Solution, (C, F, T))
7.        rank(α, solutions)
8.           if care (α, C)
9.              grade (R_C , L_1^N ),
10.          else
11.          if care (α, F)
12.             grade (R_F , L_1^N ),
13.          else
14.          if care (α, (C,F))
15.             grade (R_M , L_1^N ),
16.          end if
17.    form(α, ranked solutions)
18.       agentName (solutionNumber, solutionRank)
19.    end
20.End
```

Algorithm 1. The Solution Generator Algorithm (A_{SG})

which are, Steel Structure, Pre-cast Structure, Timber System, Cast in Situ Reinforced Concrete, and Space Frame. In addition, the three characteristics, as defined in this paper, are Cost, Function, and Time and each characteristic has a value from low to high, 1 to 5. Three types of agents are simulated in this scenario, which are Design Manager, Facility Manager, and Project Manager. Figure 6 shows the set up specifications of the scenario.

Table 1 shows the assumed values of cost, function, and time for each solution,

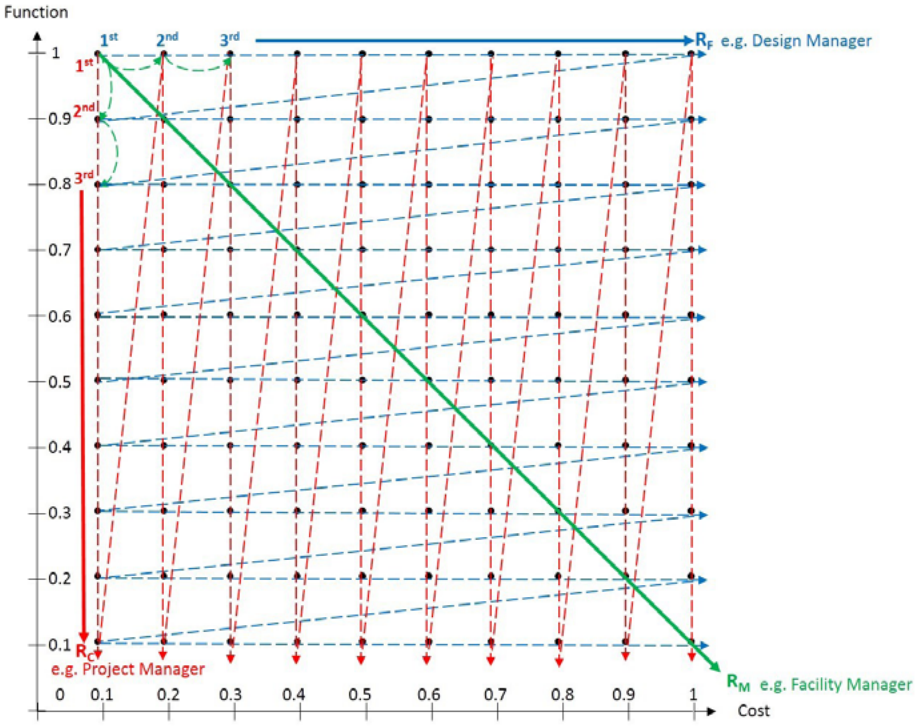


Fig. 5. Grid for Solution Algorithm.

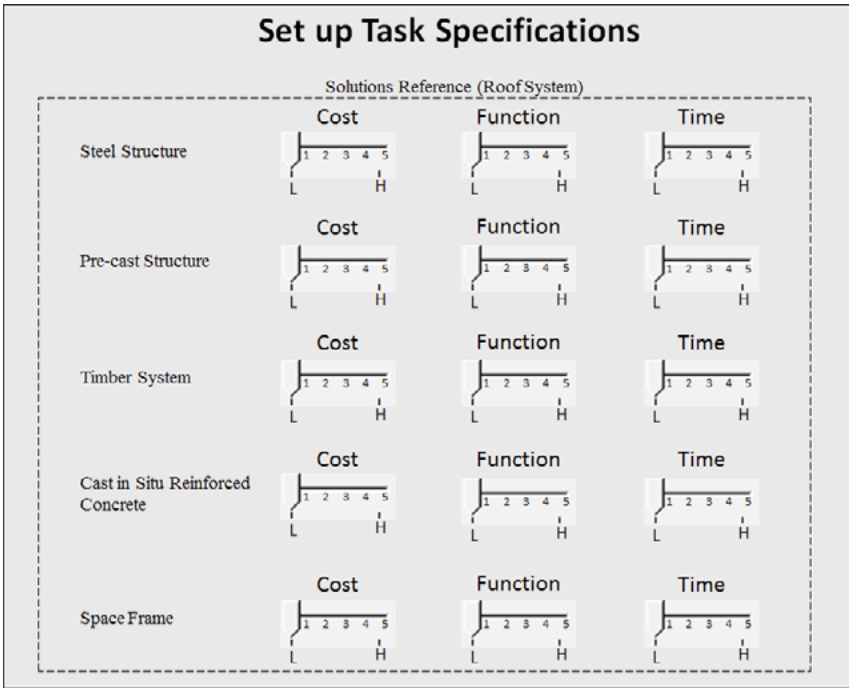


Fig 6. Set up task specifications.

TABLE I. The Assumed Values of Cost, Function, and Time for each Solution, the Scale from 1 represents Very Low to scale 5 represents Very High.

No	Solution	Cost	Function	Time
1	Steel Structure	4	5	5
2	Pre-cast Structure	3	2	4
3	Timber System	3	2	3
4	Cast in Situ Reinforced Concrete	5	3	4
5	Space Frame	4	4	3

According to [3, 4], a Project Manager cares more about the cost, and time, then function. While a Design Manager cares more about the Function then the cost and time. Facility manager has no conflict and can adapt with both Project and Design Managers. According to algorithm 1 and assumed values in Table 1, Table 2 shows the solutions ranking of each stockholder (agent).

Table 2 shows that the algorithm is able to rank the solutions according to each stakeholder/agent characteristic/ position. It also highlights a clear conflict between different stakeholders that needs further action.

VI. Conclusion and Future Work

This paper presents a solution generator algorithm to enable software agents to generate solutions and rank them from 1st to nth solution as preparation for the subsequent negotiation process. The developed algorithm consists of three steps which are, review solutions, rank solutions, and form ranked solutions. The algorithm works on ranking through three identified factors which are, the cost, function, and time and a stakeholder characteristic that inspired by the position e.g. Design Manager, Facility Manager, and Project Manager. We also present a validation scenario that deploys the developed algorithm to rank solutions. The validation shows that the algorithm is able to rank solutions according to each stakeholder/agent characteristic/position.

However, the scenario also highlights a clear conflict between different stakeholders/agents. Consequently, in our future work, we shall develop a negotiation algorithm and conflict resolution algorithm for agents to negotiate solutions and solve the conflict if any.

References

[1] Mahmoud, M. A., Ahmad, M. S., Yuso, M. Z. M., & Idrus, A. (2015). An Automated Negotiation-based Framework via Multi-Agent System for the Construction Domain. International Journal of Artificial Intelligence and Interactive Multimedia, 3(5), 23-27.

[2] Mahmoud, M. A., Ahmad, M. S., Yusoff, M. Z. M., & Idrus, A. (2015). Automated Multi-agent Negotiation Framework for the Construction Domain. In Distributed Computing and Artificial Intelligence, 12th International Conference (pp. 203-210). Springer International Publishing.

[3] Utomo C., Development of a negotiation Support Model for Value Management in Construction, PhD Thesis, University Teknologi PETRONAS, December 2009.

[4] SAVE International, value methodology standards, 2001.

[5] Kelly J. and Male S., Value Management in Decision and Construction, The Economic Management of Projects. Spon Press, London.

[6] Jaapar, A., Endut, I.R., Bari, N.A.A. and Takim, R. The impact of value management implementation in Malaysia. Journal of Sustainable Development 2 (2), 2009.

[7] Shen, Q., Chung, J.K.H., Li, H. and Shen, L.. A Group Support System for improving value management studies in construction. Automation in Construction, 13 (2004): 209–224, 2004.

[8] Cariaga, I, El-Diraby, T and Osman, H., Integrating Value Analysis and Quality Function Deployment for Evaluating Design Alternatives. Construction Engineering and Management, 133(10), 761-770, 2007.

[9] Qing Y. and Wanhua Q., 2007, Value Engineering Analysis and Evaluation. For the Second Beijing Capital Airport. Value World, Spring, SAVE International.

[10] Kexing L.. A survey of agent based automated negotiation. In Network Computing and Information Security (NCIS), 2011 International Conference on, vol. 2, pp. 24–27 (IEEE, 2011).

[11] Ahmed M., Ahmad M S, Yusoff M Z M, Modeling Agent-based Collaborative Process , The 2nd International Conference on Computational Collective Intelligence Technology and Applications (ICCCI 2010), pp. 296-305, ISBN:3-642-16692-X 978-3-642-16692-1, 10-12 November, 2010 Taiwan.

[12] Jassim, O. A., Mahmoud, M. A., & Ahmad, M. S. (2016). AN EFFECTIVE RESEARCH SUPERVISION MANAGEMENT VIA A MULTI-

AGENT SYSTEM. Journal of Theoretical and Applied Information Technology, 89(1), 155.

[13] Itaiwi A. K., Ahmad M. S., Hamid N. H. A., Jaafar N. H., Mahmoud M. A., A Framework for Resolving Task Overload Problems Using Intelligent Software Agents, 2011 IEEE International Conference on Control System, Computing and Engineering, ICCSCE11,2011.

[14] Ahmed M., Ahmad M. S., and Yusoff M. Z. M., "A Collaborative Framework for Multiagent Systems." International Journal of Agent Technologies and Systems (IJATS), 3(4):1-18, 2011.

[15] Ahmed M., Ahmad M. S., and Yusoff M. Z. M., Mitigating Human-Human Collaboration Problems using Software Agents, The 4th International KES Symposium on Agents and Multi-Agent Systems – Technologies and Application (AMSTA 2010), pp. 203-212, ISBN:3-642-13479-3 978-3-642-13479-1, Gdynia, Poland, 23 – 25 June 2010.

[16] Itaiwi, A. M. K., Ahmad, M. S., Hamid, N. H. A., Jaafar, N. H., & Mahmoud, M. A. (2012, June). A multi-agent framework for dynamic task assignment and delegation in task distribution. In Computer & Information Science (ICCIS), 2012 International Conference on (Vol. 1, pp. 318-323). IEEE.

[17] Mahmoud, M. A., Ahmad, M. S., Mohd Yusoff, M. Z., & Mustapha, A. (2014). A Review of Norms and Normative Multiagent Systems. The Scientific World Journal, 2014.

[18] Mahmoud, M. A., Ahmad, M. S., Ahmad, A., Yusoff, M. Z. M., Mustapha, A., & Hamid, N. H. A. (2013, May). Obligation and Prohibition Norms Mining Algorithm for Normative Multi-agent Systems. In KES-AMSTA (pp. 115-124).

[19] Mahmoud, M. A., Ahmad, M. S., Ahmad, A., Yusoff, M. Z. M., & Mustapha, A. (2012). Norms Detection and Assimilation in Multi-agent Systems: A Conceptual Approach. In Knowledge Technology (pp. 226-233). Springer Berlin Heidelberg.

[20] Mahmoud, M. A., Ahmad, M. S., Ahmad, A., Mohd Yusoff, M. Z., & Mustapha, A. (2012, June). A norms mining approach to norms detection in multi-agent systems. In Computer & Information Science (ICCIS), 2012 International Conference on (Vol. 1, pp. 458-463). IEEE.

[21] Beer M., d'Inverno M., Jennings R.N., Luck M., Preist C., Schroeder M., Negotiation in multi-agent systems Knowledge Engineering Review, 14 (3) (1999), pp. 285–289.

[22] Z. Ren, C.J. Anumba, Multi-agent systems in construction—state of the art and prospects, Automation in Construction, 13 (2004), pp. 421–434

[23] M. Wang, H. Wang, D. Vogel, K. Kumar, D.K.W. Chiu Agent-based negotiation and decision making for dynamic supply chain formation Engineering Applications of Artificial Intelligence, 22 (7) (2009), pp. 1046–1055.

[24] Utomo C., Idrus A., A Concept toward Negotiation Support for Value Management on Sustainable Construction, Journal of Sustainable Development Vol 4, No 6 (2011).

[25] Victor Sanchez-Anguix , Vicente Julian , Vicente Botti , Ana García-Fornes, Tasks for agent-based negotiation teams: Analysis, review, and challenges, Engineering Applications of Artificial Intelligence, v.26 n.10, p.2480-2494, November, 2013.

[26] Coutinho, C., Cretant, A., Ferreira da Silva, C., Ghodous, P., & Jardim-Goncalves, R. (2014). Service-based negotiation for advanced collaboration in enterprise networks. Journal of Intelligent Manufacturing. doi:10.1007/s10845-013-0857-4.

[27] Dzeng, R. J., & Lin, Y. C. (2004). Intelligent agents for supporting construction procurement negotiation. Expert Systems with Applications, 27(1), 107-119.

[28] C.J. Anumba, Z. Ren, A. Thorpe, O.O. Ugwu, L. Newnham, Negotiation within a multi-agent system for the collaborative design of light industrial buildings Adv Eng Software, 34 (7) (2003), pp. 389–401.

[29] Mahmoud, M. A., Ahmad, M. S., & Yusoff, M. Z. M. (2016). A Conceptual Automated Negotiation Model for Decision Making in the Construction Domain. In Distributed Computing and Artificial Intelligence, 13th International Conference (pp. 13-21). Springer International Publishing.

Arazi Idrus

B Eng.(Hons) in Civil and Structural Engineering (Sheffield University, UK), M.Sc. (Cranfield University, UK), Ph.D (Imperial College, University of London, UK). Field of Specialization: Project Management, Construction It, Construction Productivity, Value Management; PPP/PFI, IBS Construction, Building Maintenance Management, Risk-Based Assessment, Concrete Repairs,Blast Design, and Load Resistance Factor Design of Offshore Structures.

Moamin A. Mahmoud

Moamin A. Mahmoud obtained his Bachelor in Mathematics from the College of Mathematics and Computer Science, University of Mosul, Iraq in 2007. He obtained his Master of Information Technology at the College of Graduate Studies, Universiti Tenaga Nasional (UNITEN), Malaysia in 2010, and PhD of Information and Communication Technology from Universiti Tenaga Nasional (UNITEN), Malaysia in 2013. His research interests are in the area of software agents, agent behavior in open societies, and social sciences simulation.

Mohd Sharifuddin Ahmad

Mohd Sharifuddin Ahmad received his B.Sc. in Electrical and Electronic Engineering from Brighton Polytechnic, UK in 1980. He started his career as a power plant engineer specializing in Process Instrumentation and Control in 1980. After completing his MSc in Artificial Intelligence from Cranfield University, UK in 1995, he joined UNITEN as a Principal Lecturer and Head of Dept. of Computer Science and Information Technology. He obtained his PhD from Imperial College, London, UK in 2005. He has been an associate professor at UNITEN since 2006. His research interests include applying constraints to develop collaborative frameworks in multi-agent systems, collaborative interactions in multi-agent systems and tacit knowledge management using AI techniques.

Azani Yahya

Field of Specialization Construction and Construction Materials - Construction Technology.

Hapsa Husen

Field of Specialization Statistics, earthquake engineering.

A Novel Hybrid Approach for Fast Block Based Motion Estimation

Shaifali Madan Arora[1], Kavita Khanna[2], Navin Rajpal[3]*

[1] Research Scholar, USICT, GGSIPU, Dwarka, & Assistant Professor, ECE Department, MSIT Janakpuri, New Delhi (India)
[2] Associate Professor, The NorthCap University Gurugram, Haryana (India)
[3] Professor, GGSIP University, Dwarka, Sector 16 C, New Delhi (India)

Received 26 December 2016 | Accepted 4 March 2017 | Published 30 March 2017

Abstract

The current work presents a novel hybrid approach for motion estimation of various video sequences with a purpose to speed up the entire process without affecting the accuracy. The method integrates the dynamic Zero motion pre-judgment (ZMP) technique with Initial search centers (ISC) along with half way search termination and Small diamond search pattern. Calculation of the initial search centers has been shifted after the process of zero motion pre-judgment unlike most the previous approaches so that the search centers for stationary blocks need not be identified. Proper identification of ISC dismisses the need to use any fast block matching algorithm (BMA) to find the motion vectors (MV), rather a fixed search pattern such as small diamond search pattern is sufficient to use. Half way search termination has also been incorporated into the algorithm which helps in deciding whether the predicted ISC is the actual MV or not which further reduced the number of computations. Simulation results of the complete hybrid approach have been compared to other standard methods in the field. The method presented in the manuscript ensures better video quality with fewer computations.

Keywords

Motion Estimation, Zero Motion Prejudgment, Initial Search Center, Decision Error, Peak Signal To Noise Ratio, Sum Of Absolute Differences.

DOI: 10.9781/ijimai.2017.03.011

I. Introduction

Rapid use of video based applications in broadcast and entertainment media has led to an overwhelming need to compress the video data. As a result many approaches came up towards video compression. Block based motion estimation is the most prevalent among the various techniques for motion estimation (ME). Due to the computational complexity of the ME process, an extensive research in the field has been conducted in the last two decades. Popularity of block based ME can be attributed to the simplicity and ease in hardware implementation of these algorithms. As a result, these algorithms have been used in many of the video coding standards including MPEG4 and H.264.

Block based motion estimation is based on the idea of reducing the temporal redundancy across the frames by matching the blocks in the current frame to the blocks in the reference frame in a specified search window. The brute force approach is to match all possible candidates in the search window and is known as full search [1]. This approach provides the optimum results but increases the computational overhead. The research then shifted over to finding the best match with the limited number of checking points in the search window. Three step search (TSS) [2], new three step search (NTSS) [3], four step search (4SS) [4] and diamond search (DS) [5] are the famous algorithms which resulted in high PSNR values and lesser computations. The main problem with all these approaches is of quality degradation due to trapping of search process in local minima as they use a fixed pattern for search. Adaptive rood pattern search [6] found a way out to this complication by using other search patterns in accordance with the estimated behavior of the current block. There after many algorithms have been given in this category which helped in decreasing the number of search points via dynamic search paths. In all the block matching algorithms motion estimation is established by locating the ideal match for the current block. The matching criteria that is used most widely and in current work is the sum of absolute differences (SAD) which needs to be minimized in order to maximize the performance which is measured with peak signal to noise ratio (PSNR).

Development of various fixed and adaptive search pattern based algorithms reduce the computational burden but at the same time they compromise with the video quality. The main aim of any motion estimation algorithm is to reduce the number of computations without deteriorating the video quality. Zero motion prejudgement (ZMP) and initial search centre prediction (ISC) have been proven to be beneficial in accelerating the process of motion estimation. ZMP helps to identify the stationary blocks before the calculation of the actual motion vector and thus saves the computations for calculating the motion vectors of the stationary blocks. On the other hand ISC aims to find an initial location in the search window so that a refined search is carried around this point instead of centre of search window.

All the above proposed algorithms use the center of the search window for starting the process of finding the best matching block. It has been observed that there exists spatial as well as temporal coherence between the adjacent neighboring blocks and hence the motion of the current block can be predicted by utilizing the motion information of

* Corresponding author.
E-mail addresses: navin_rajpal@yahoo.com

the neighboring blocks. Using this information motion vector for the current block can be predicted. This predicted location is expected to be in the region of global minima, this reduces the number of search steps to attain global minima and thus the number of computations for motion estimation. Further, accuracy in determination of ZMP and ISC enhances the accuracy in determination of motion vectors which enhances the accuracy and quality of regenerated frame at the receiver end. Hence bit coding error, which refers to the difference in actual frame and regenerated frame at the receiver, should be reduced.

In the current work we have used a hybrid approach for fast block matching motion estimation. The idea is to firstly identify the stationary blocks and stop the search process for these blocks. For this purpose the dynamic threshold prediction technique as given in [7] has been used. The technique is not only simple but is also efficient in identifying the number of stationary blocks and thus helps to reduce the decision error.

After the identification of stationary blocks, initial search centers have been predicted with an approach as given in [8]. The advantage of this approach lies in its precise and accurate prediction of initial search centers which aids in speeding up the entire process of motion estimation.

Novelty of the proposed hybrid scheme is established with the help of increased PSNR, SSIM and search efficiency in comparison to various state-of-art algorithms in the field of fast block matching motion estimation. Also the number of computations is reduced when compared to the other standard methods

This paper is organized as follows. Section 2 describes the concepts of zero motion pre- judgment (ZMP) and dynamic threshold estimation technique of ZMP. General concept of initial search center (ISC) prediction along with the method of predicting ISC is given in section 3. A brief introduction to half way search termination and small diamond search algorithm are given in section 4 and 5 respectively. The detailed hybrid algorithm designed for the solution of the problem is presented in section 6. Simulation results along with the analysis and comparisons are shown in section 7. Section 8 concludes the presented work.

II. Zero Motion Pre-judgment

Zero motion pre-judgment has been extensively used in the literature to identify stationary blocks early in the video sequences so as to save unnecessary computations. It has been established in [6] that block distortion for stationary blocks is very less in comparison to moving blocks which plays a key role in identifying stationary blocks. The SAD value of the current block to the stationary block represents the block distortion and this SAD value is compared to a predetermined threshold for detecting stationary blocks. Different approaches in literature have used different thresholds for ZMP.

The concept of fixed threshold based Zero motion pre-judgment was firstly taken by Nie and Ma [6] while proposing adaptive rood pattern search algorithm. This approach is based on using a fixed threshold of 512 but using this threshold a large number of moving blocks could be detected as stationary blocks especially for slow motion sequences. The concept of fixed threshold has also been used by Luo et al [9] along with the search priority assigned to each point. The dis-advantages of fixed threshold have led to the use of dynamic thresholds. Ahmed et al [10] have used an adaptive threshold which is determined by finding the highest or lowest of SAD values of the adjacent MBs based on specified conditions. Ismail et al [11] have used three level thresholds on the basis of three categories of SAD values. Dynamic early stop termination technique is also proposed in [11] to dynamically update the threshold by using the following equation [11]:

$$T_s = \min[\max(SAD_{0,0avg}, 400), SAD].\lambda + \varepsilon \; ; \quad (1)$$

where $SAD_{0,0avg}$ is the average of all the previous stationary blocks, λ is used to slow down or accelerate the ME process and ε is empirically taken as zero.

Two static thresholds based on motion contents have been given by Lin et al [12] which have been determined as per the static experimental results.

In yet another advancement of predicting threshold adaptively Ismail et al [13, 14] have given a formula based on average SAD scores of all the stationary blocks. The threshold value T_s is given as [13, 14]:

$$T_s = \min[\max(SAD_{ISCavg}, 512), SAD_{ISCcurrent}.\alpha + \beta \; ; \quad (2)$$

where parameters $\alpha = 0.75$ and $\beta = 128$.

But these thresholds do not guarantee the accurate results. A further refined dynamic threshold estimation technique given in [7] is based on the following observations:

1. A block which is having SAD below a particular threshold is not necessarily a stationary block.
2. SAD value of a stationary block w.r.t. its collocated block is least when compared to the SAD value of the stationary block with respect to its vertical and horizontal neighbors taken in the reference frame.

These drawbacks have been alleviated by using a two level threshold estimation technique given in [7]. We have used only a single level of the technique and incorporated in the proposed hybrid ME technique. The reason is that in [7] only the issue of ZMP has been taken up but here we are using other techniques along with ZMP to fasten the process of motion estimation. Use of both the levels incurs lot of complexity in terms of number of computations and thus only one level has been found sufficient when used with other techniques.

This single level of threshold determination for ZMP is explained with the help of following algorithm:

$T_1 = \max(SADa, 256) * \alpha + \beta$

if $SAD_c < T_1$

check if SAD_c is equal to min(SAD_c, SAD_l, SAD_r, SAD_t, SAD_b)

then Declare the block as stationary

if abs (T_1-SAD_c) < α

then update SADa

else SADa remains same,

else Move to motion estimation algorithm

Here T_1 is determined by modifying equation (2) defined above. Here 256 is taken, instead of 512 in max operator so that moving blocks in slow motion sequences with small distortions can be appropriately determined. SAD_c is the SAD between the current block and its collocated block in the reference frame; SAD_l, SAD_r, SAD_t, SAD_b represent the SADs between current block w. r. t its left, right, top and bottom neighboring blocks in the reference frame. SADa represents the average distortion which is initially given a value of 512 from the results in [6] for fixed threshold so as to find the first stationary block. This value is updated and assigned the SAD value of first stationary block encountered. Max operator assists to pursue changes in SADa. SADa is updated based on the difference between SAD_c and T_1. If this difference is greater than α then SAD_c is not considered for updating the average distortion SADa. As a consequence the effect of very large or very small distortion values of the current stationary block would not affect the average variation of threshold. Parameters α is taken as 0.75 and β is taken as 128.

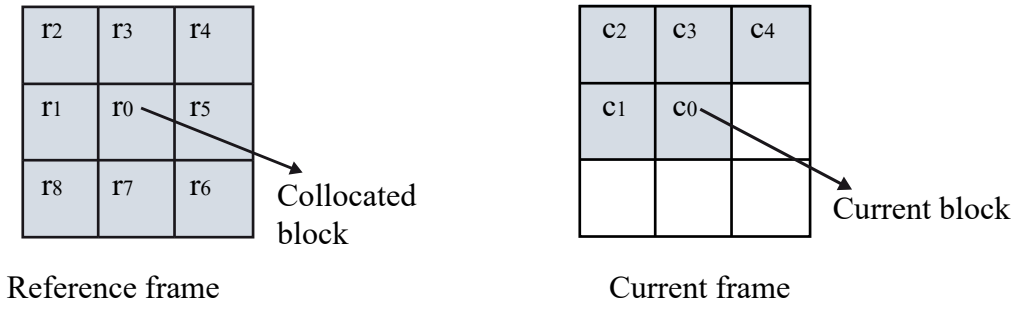


Fig. 1. Spatial and temporal neighbors of current block.

III. Initial Search Center (ISC) Prediction

In most of the recent approaches using ZMP and ISC for fast motion estimation, initial search centers are identified before the stationary blocks. But using ZMP as a post processing step to ISC, leads to the identification of search centers even for the stationary blocks. This is the reason we have taken this step of predicting ISC after the ZMP.

Initial search center prediction helps in faster attainment of actual MV. ISC is predicted on the notion that there exist a lot of similarities in the neighboring video frames. These similarities may be spatial or temporal. Figure 1 shows the temporal and spatial neighboring relations.

Thus the current block will exhibit similar motion as compared to its surrounding blocks. So the motion prediction of the current block can be done from the motion of neighboring blocks in the current and temporal frame. Various methods have been proposed in literature for finding the ISC. A tabulated summary of these methods is given in [8].

The method used for ISC prediction in the current approach is the one given in [8]. This method has the following advantages over the previous methods:

1. The method makes use of the future points from reference frame to account for the fact that motion of an object is possible in any of the neighboring directions. No method in literature has used this concept.
2. The method works in two stages. First stage works by finding the suitable MVs whereas second stage finds the best among the previously found MVs.

A procedure used in [8] for finding the ISC is as follows:

(1) Find an initial estimation of the motion vectors denoted by MPISC as:

$$MPISC = \text{median}\ (c_1, c_2, c_3, c_4, r_0, r_5, r_6, r_7, r_8) \quad (3)$$

Find the variation of MPISC with all the neighboring MVs (9 blocks):

$$V_i = \text{abs}\ (MPISC_x - c_{ix}/r_{ix}) + \text{abs}(MPISC_y - c_{iy}/r_{iy}) \quad (4)$$

The blocks for which $V_i > T_2$ are the suitable blocks for further processing; where $T_2 = 2$.

These candidate blocks are denoted by $CISC_i$ (candidates for ISC).

(2) Find SAD of the current block with MPISC and with all the CISC.

$$SAD_i = SAD(c_0, MPISC/CISC_i) \quad (5)$$

Minimum SAD implies that the probability of movement is in this particular direction. Therefore ISC is assigned MV in accordance to macro block with least SAD.

IV. Half-way Early Search Termination

Predicted ISC can be the position of actual MV. If this can be detected early then search can be terminated early. To do so, the SAD value of the predicted ISC is checked, if it is below a predefined threshold Td, then the current block may be assumed to have high correlation with that particular neighboring block. Same MV can be declared for this block as that of the best matched neighboring block and search is terminated thereafter saving huge computations. Threshold Td in the proposed manuscript is taken same as that T_1 calculated above.

V. Small Diamond Search (SDS) Algorithm

Once the initial search center is predicted with the proposed technique, there is high probability that it lies near the global minima. So the actual MV could be obtained by using a fixed and small search pattern to perform a refined search rather than using some fast BMA. Two types of fixed small search patterns have been defined in literature – four point pattern as in small diamond search (SDS) [15] and eight point square search pattern as in block based gradient descent search (BBGDS) [16]. We have used SDS rather than BBGDS to perform the refined search for MV. It is based on the comparative analysis of SDS and BBGDS given by Nee and Ma [6] indicating clearly that performance, in terms of PSNR, of both the algorithms is almost same whereas BBGDs incurs 40-80% more complexity in terms of number of calculations. Figure 2 shows the two fixed search patterns

So with the ISC and SDS, minimum distortion point (MDP) is obtained which is then considered as the new search center. This recursive procedure continues till the MDP is the center of the fixed SDS pattern or search window boundary is met.

Fig. 2. Fixed small search patterns.

VI. Proposed Hybrid ME Algorithm

The proposed hybrid algorithm based on ZMP, ISC, Half way search termination and SDS works in following steps:

1. Find SAD_c of current block and its collocated block in the reference frame. If this $SAD_c < T1$ AND SADc is equal to min(SAD_c, SAD_l, SAD_r, SAD_t, SAD_b), block is declared as stationary block. Search is terminated thereafter and go to step 7 otherwise go to step 2.

2. Find the MPISC = median (c_1, c_2, c_3, c_4, r_0, r_5, r_6, r_7, r_8) and identify the points using V_i= abs ($MPISC_x$- c_{ix}/r_{ix}) + abs($MPISC_y$- c_{iy}/r_{iy}) , which are distant apart from MPISC. The points for which V_i is above a threshold will be the candidate points for ISC (CISC) prediction.
3. Compute the SAD of C_0 with MPISC and CISCs and then find minimum SAD. Declare ISC as the point corresponding to minimum SAD.
4. Check whether ISC could be the location of actual MV by comparing its SAD with a predicted dynamic threshold. If true, declare the position of ISC as MV for current block and go to step 7 otherwise to next step.
5. Create a search pattern as shown in fig using four symmetrical horizontal and vertical points at distance unit distance around ISC. Compute minimum distortion point (MDP).
6. Set MDP from step 5 as the center for new unit size small diamond search pattern including four symmetrical adjacent points in horizontal and vertical directions. The search is applied repeatedly using this new search pattern taking MDP at the center until the MDP comes out to be the center point of the search pattern or the search window boundary is met. The search is terminated at any point if SAD<Td. Declare the displacement of this point from the center of search window as the MV for this block.
7. Exit

The steps followed in the proposed algorithm are depicted graphically in Figure 3.

VII. Performance Analysis and Comparison Results

The main goal of any algorithm is to lower the computational complexity while maintaining the video quality as that of FS algorithm. The performance of the proposed algorithm has been evaluate by doing

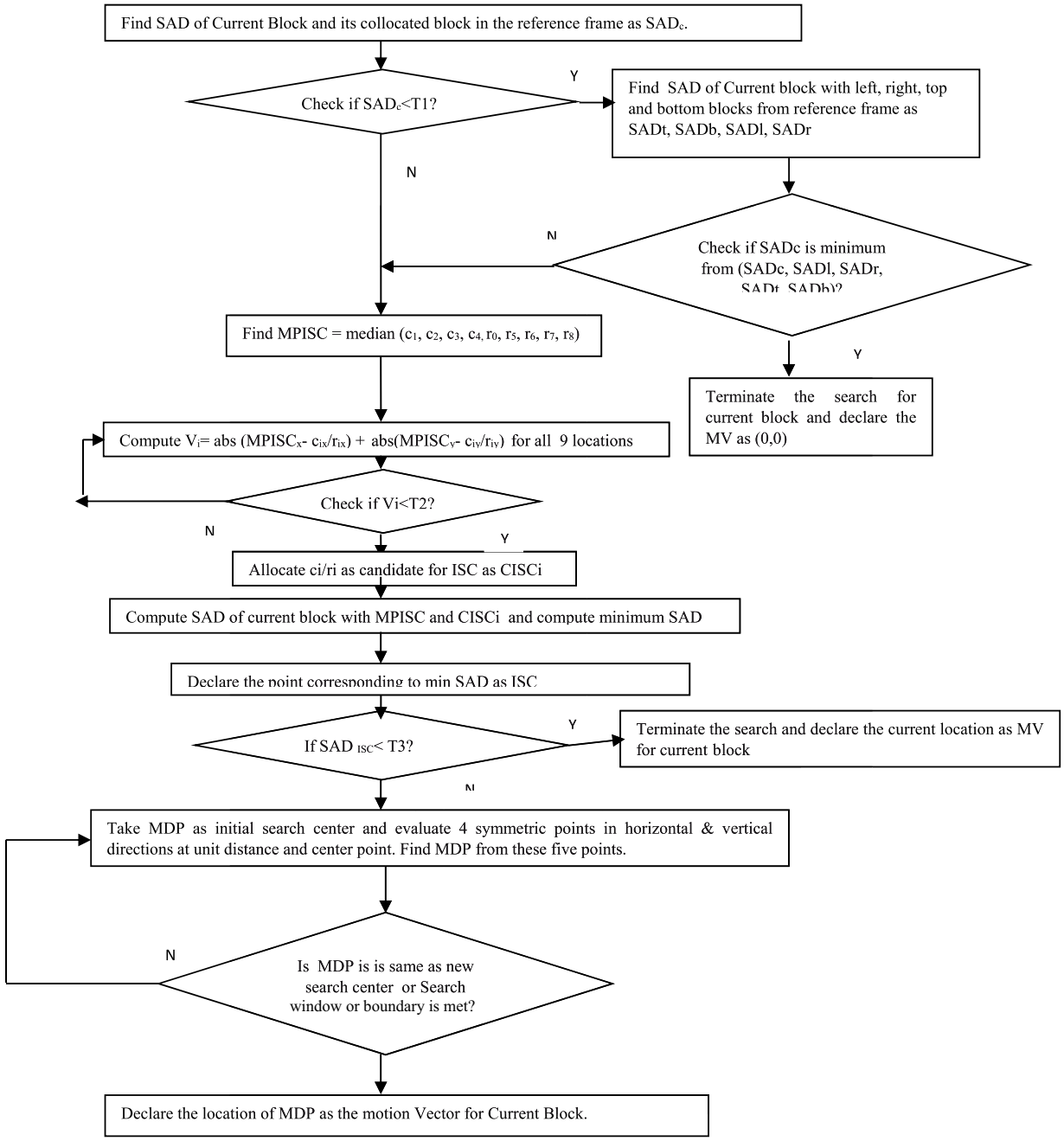


Fig. 3. Block diagram of the proposed algorithm.

simulations on various standard YUV test video sequences containing different motion characteristics, listed in Table 1. Experimental set up for simulations include a 15fps, size of ±7 for search window and16×16 for blocks. Proposed algorithm is compared with fixed size algorithms like FS, TSS, NTSS, 4SS, DS, and predictive motion based vector adaptive search pattern algorithms like ARPS, DPS and recently proposed APSP and FPS algorithms.

To measure the performance of proposed ME algorithm following parameters are evaluated – computational complexity and search efficiency, video quality in terms of average PSNR per frame, structural similarity index measurement (SSIM) per frame, average number of bits required per pixel to represent the residual frame (Difference between the actual and the motion compensated frame) and distance between the actual and predicted Motion Vector.

Computational complexity of a ME algorithm can be evaluated in terms of average number of search points required per block to estimate the MVs.

Search efficiency can be evaluated by finding the distance between the actual MV obtained from FS algorithm and the estimated MV using a fast BMA.

$$Distance\_actualMV\_predictedMV(DAP) = \frac{(|MV_x - MV_{fx}| + |MV_y - MV_{fy}|)}{NB} \quad (6)$$

Where (MV_x, MV_y), (MV_{fx}, MV_{fy}) represents MVs of FS algorithm and fast BMA resp; NB represents the total number of blocks in a frame.

TABLE I. Test Video Sequences Used in Experiment Having Different Motion Characteristics

Sequence	Motion Characteristics
QCIF Video Sequences	
Akiyo, Clair, Miss America	Static Background with very small moving objects, maximum quasi stationary blocks
Salesman, Silent, Foreman	Static Background with objects having Large Motion
Football, Tennis	Background and Objects with Large motion in horizontal and vertical Directions
CIF Video Sequences	
News, Mother-Daughter	Static Background with small moving objects, maximum quasi stationary blocks
Hall	Static Background with large moving objects
Stefan, Mobile Calendar	Large motion in horizontal direction with Rich Motion Content
Highway	Large motion in vertical Directions

TABLE II. Comparison of Various Parameters for QCIF Video Sequences

Sequences	Parameters	FS	DS	ARPS	DPS	ASPS	FPS	Proposed
Akiyo	Comp	184.56	11.44	1.44	4.90	1.47	0.96	0.97
	PSNR	42.16	42.16	42.16	42.16	42.10	42.16	42.16
	SSIM	0.992	0.992	0.992	0.992	0.992	0.992	0.992
	DAP		0.20	0.33	0.20	0.40	0.37	0.30
Clair	Comp	184.56	11.48	1.39	5.05	1.45	0.96	0.98
	PSNR	42.75	42.69	42.68	42.68	42.57	42.61	42.68
	SSIM	0.991	0.991	0.991	0.991	0.991	0.991	0.991
	DAP		5.31	5.32	5.34	5.95	5.81	5.33
Miss America	Comp	184.56	12.21	2.01	6.24	1.48	1.49	1.76
	PSNR	40.62	40.62	40.53	40.39	39.49	39.95	40.57
	SSIM	0.981	0.981	0.981	0.981	0.981	0.981	0.981
	DAP		31.82	60.73	64.42	65.99	66.59	51.94
Silent	Comp	184.56	11.87	2.32	5.53	1.84	1.30	1.94
	PSNR	36.20	35.80	35.72	35.70	35.52	35.62	35.82
	SSIM	0.968	0.968	0.967	0.967	0.966	0.966	0.967
	DAP		6.32	7.86	7.62	11.60	11.34	6.20
Salesman	Comp	184.56	11.51	1.80	5.05	1.48	1.49	1.47
	PSNR	38.61	38.59	38.57	38.57	38.57	38.55	38.57
	SSIM	0.979	0.979	0.979	0.979	0.979	0.979	0.979
	DAP		0.37	0.80	0.70	0.77	0.93	0.63
Foreman	Comp	184.56	12.21	4.08	6.03	6.07	5.83	4.00
	PSNR	28.71	28.50	28.45	28.42	28.42	28.32	28.48
	SSIM	0.908	0.907	0.907	0.907	0.907	0.905	0.908
	DAP		26.93	39.53	39.27	40.93	42.93	30.57
Football	Comp	184.56	17.52	10.59	10.05	12.57	14.78	10.46
	PSNR	21.65	21.43	21.35	21.10	21.07	20.12	21.55
	SSIM	0.669	0.656	0.655	0.638	0.634	0.603	0.661
	DAP		94.06	98.66	113.34	129.94	171.32	83.36

Further Peak Signal to noise ratio (PSNR) is evaluated as [17, 18]:

$$PSNR = 10\log_{10}\frac{Max^2}{MSE} \quad (7)$$

Where the value of Max is taken as 255 representing the maximum possible pixel value in a video frame. MSE is the mean square error between the original frame and the motion compensated frame. FS algorithm gives the best MVs, hence best video quality and maximum PSNR. Therefore provides standard PSNR with which the PSNR values calculated from other BMAs is compared.

Structural similarity index measurement (SSIM) is also a means to measure the similarity between two images. SSIM between the two blocks C (block in original frame) and R(block in motion compensated frame) is given as:

$$SSIM(C,R) = \frac{2(u_C u_R + p1)(2\sigma_{CR} + p2)}{(u_C^2 + u_R^2 + p1)(\sigma_C^2 + \sigma_R^2 + p2)} \quad (8)$$

It has been observed from fixed search pattern algorithms like TSS that these use fixed search points to attain actual MV. Early search termination and search near center feature is added in NTSS and 4SS to reduce the search points especially in slow motion sequences. DS algorithm incorporated above features along with special search pattern further lowered the search points and became the most favorable algorithm in various standards. ARPS with zero motion prejudgment and DPS algorithms modified the large diamond search pattern in DS and able to reduce the search points. Recently proposed ASPS and FPS algorithms incorporated ISC & ZMP techniques to reduce the search points. In these algorithms ISC point is found first and then ZMP is implemented. However in case of slow motion sequences most of the blocks don't possess motion or are stationary. Therefore finding ISC before determining a block to be stationary or not would involve unnecessary ISC computations for ZMP. Therefore in the current approach ZMP is implemented first and ISC is calculated only for the blocks which are not stationary.

The simulation results are shown in Tables 2 and 3 for qcif and cif sequences respectively. Comparisons have been done on the basis of four parameters that are computations, PSNR, SSIM and distance between actual and predicted MV. It can be observed from the results that there is 9-11 times reduction in search locations with the proposed approach for very slow motion video sequences having stationary background like "Akiyo", "Clair", "Miss America" compared to DS algorithm in qcif video sequences. The reduction in CIF video sequences like "News" & "Mother-Daughter" is 6-7 times. Such high reduction in computations is possible because of accuracy in prediction of ISC which leads this predicted point to be in the region of global minima and search followed by small number of search points to attain the position of actual motion vector. The computations are slightly larger than recently proposed FPS algorithm because in the proposed ZMP technique, four additional neighboring points need to be checked. This small overhead increases accuracy in determination of slow moving blocks and stationary blocks otherwise very slow moving blocks whose distortion is less than dynamic threshold have high probability to be identified as stationary blocks. This influence

Table III. Comparison of Various Parameters for CIF Video Sequences

Sequences	Parameters	FS	DS	ARPS	DPS	ASPS	FPS	Proposed
Mother Daughter	Comp	204.28	13.60	2.56	6.77	2.19	2.55	2.82
	PSNR	39.85	39.70	39.65	39.55	39.00	38.78	39.68
	SSIM	0.973	0.973	0.972	0.972	0.968	0.969	0.972
	DAP		444.31	504.70	515.00	521.33	538.87	484.49
News	Comp	204.28	12.62	2.02	5.63	2.03	2.23	1.97
	PSNR	37.26	37.06	37.01	36.88	36.61	36.74	37.04
	SSIM	0.982	0.982	0.982	0.982	0.981	0.981	0.982
	DAP		27.77	42.49	34.29	63.03	56.60	36.54
Hall	Comp	204.28	13.02	7.34	6.27	6.083	6.6431	5.89
	PSNR	34.37	34.15	34.11	34.07	34.06	34.01	34.14
	SSIM	0.941	0.941	0.941	0.941	0.941	0.940	0.941
	DAP		161.13	213.53	244.33	226.67	270.80	193.50
Mobile Calender	Comp	204.28	15.22	6.41	8.30	7.29	7.01	7.39
	PSNR	33.88	33.60	33.50	33.20	32.67	32.24	33.53
	SSIM	0.933	0.931	0.930	0.927	0.920	0.915	0.931
	DAP		360.11	446.68	445.30	552.32	604.46	412.93
Highway	Comp	204.28	15.47	6.94	8.64	7.89	5.65	7.99
	PSNR	35.62	34.52	34.28	34.04	34.01	33.63	34.69
	SSIM	0.906	0.901	0.899	0.898	0.897	0.889	0.901
	DAP		586.798	686.010	662.727	727.697	874.970	650.465
Stefan	Comp	204.28	19.81	8.39	8.77	9.41	15.59	7.59
	PSNR	23.82	20.94	23.10	23.03	19.52	19.01	23.63
	SSIM	0.850	0.750	0.825	0.823	0.645	0.654	0.844
	DAP		994.50	437.90	392.30	1199.30	1481.20	271.10

can be observed as improvement in video quality in terms of PSNR, SSIM. For sequence like "Miss America", "Mother-Daughter" there is appreciable improvement in PSNR to recently proposed FPS or ASPS algorithms. This improvement is possible because of accuracy in determination of actual motion vectors and reduction in trapping in local minima. Further the proposed algorithm shows appreciable improvement in search speed and video quality especially for fast motion sequences.

VIII. Conclusion

In this paper a hybrid technique for fast motion estimation is proposed. The technique is based on using improved dynamic techniques for determination of zero motion blocks, improved accuracy in prediction of initial search center prediction, early search termination and small diamond search pattern. Proposed technique enhances the video quality in terms of PSNR and SSIM. Further it increases the search efficiency and reduces the number of computations required to estimate the motion vectors. Simulation results show the superiority of the proposed technique to the existing techniques in literature.

References

[1] M. Ahmadi and M. Azadfar, "Implementation of fast motion estimation algorithms & comparison with full search method in H.264", IJCSNS International Journal of Computer Science & Network Security, vol. 8(3), pp. 139-143, 2008.

[2] T. Koga, K. Iinuma, A. Hirano, Y. Iijima and T. Ishiguru, "Motion compensated inter frame coding for video conferencing", in Proc of NTC81. New Orleans, LA, pp. C9.6.1–C9.6.5, 1981.

[3] R. Li, B. Zeng and M. L. Liou, "A new three step search algorithm for block motion estimation", IEEE Transactions on Circuits and Systems for Video Technology, vol. 4, pp. 438-442, 1994.

[4] L. Po and W. Ma, "A novel four step search algorithm for fast block Motion estimation", IEEE Transactions on Circuits and Systems for Video Technology, vol. 6(3), pp. 313-317, 1996.

[5] S. Zhu and K. Ma, "A new diamond search algorithm for fast block Matching motion estimation", IEEE Transactions on Image Processing, vol. 9(2), pp. 287–317, 2000.

[6] Yao Nie, and Kai-Kuang Ma, "Adaptive Rood Pattern Search for Fast Block-Matching Motion Estimation", in IEEE Trans. on Image Processing, vol 11(12), pp. 1442 -1448, 2002.

[7] Shaifali Madan Arora, Navin Rajpal and Kavita Khanna "A new approach with enhanced accuracy in zero motion prejudgment for motion estimation in real time applications", Real Time Image Processing, Springer, 2016, pp 1-17, DOI: 10.1007/s11554-016- 0593-z.

[8] Shaifali Madan Arora, Navin Rajpal, Kavita Khanna and Ravinder Purwar, "Improved Accuracy in Initial Search Center Prediction to Fasten Motion Estimation in h.264", IETE journal of Research, Published online DOI: http://dx.doi.org/10.1080/03772063.2016.1205962.

[9] Luo, Jun, Xiaohua Yang, and Liheng Liu, "A fast motion estimation algorithm based on adaptive pattern and search priority", Multimedia Tools and Applications, vol. 74(24), pp. 11821–11836, 2015.

[10] Ishfaq Ahmad, , Weiguo Zheng, Jiancong Luo, and Ming Liou, "A fast adaptive motion estimation algorithm", Circuits and Systems for Video Technology, IEEE Transactions on, vol. 16(3), pp. 420-438, 2006.

[11] Ismail, Yasser, Mohamed A. Elgamel, and Magdy A. Bayoumi, "Fast variable padding motion estimation using smart zero motion prejudgment technique for pixel and frequency domains", Circuits and Systems for Video Technology, IEEE Transactions on, vol. 19(5), pp. 609-626, 2009.

[12] Lin, Lili, I-Chyn Wey, and Jing-Hua Ding. "Fast predictive motion estimation algorithm with adaptive search mode based on motion type classification", Signal, Image and Video Processing, vol. 10(1), pp. 171-180, 2016.

[13] Yasser Ismail, Mohamed Elgamel and Magdy Bayoumi, "Adaptive techniques for a fast frequency domain motion estimation", In Signal Processing Systems, 2007 IEEE Workshop on, pp. 331-336. IEEE, 2007.

[14] Yasser Ismail, Jason B. McNeely, Mohsen Shaaban, Hanan Mahmoud and Magdy A. Bayoumi, "Fast motion estimation system using dynamic models for H. 264/AVC video coding", Circuits and Systems for Video Technology, IEEE Transactions on, vol. 22(1), pp. 28-42, 2012.

[15] S. Zhu and K.-K. Ma, "A new diamond search algorithm for fast block-matching motion estimation," in Proc. 1997 Int. Conf. Information, Communications and Signal Processing (ICICS), vol. 1, Sept. 9–12, 1997, pp. 292–296.

[16] Wang Y, Wang Y, Kuroda H, "A globally adaptive pixel decimation algorithm for block motion estimation", IEEE Transactions on Circuits and Systems for Video Technology, vol. 7(3), pp. 477-488, 1997.

[17] Kamble, Shailesh D., Nileshsingh V. Thakur, and Preeti R. Bajaj. "A Review on Block Matching Motion Estimation and Automata Theory based Approaches for Fractal Coding", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4(2), pp.91-104, 2016.

[18] Kamble, Shailesh D., Nileshsingh V. Thakur, and Preeti R. Bajaj. "Modified Three-Step Search Block Matching Motion Estimation and Weighted Finite Automata based Fractal Video Compression", International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4(4), pp. 27-39, 2017.

Shaifali Madan Arora

Dr Shaifali Madan Arora is an Assistant Professor at MSIT, New Delhi. She is a Ph. D from GGSIPU, Dwarka, New Delhi. She has done her B.Tech from GNDU, Amritsar, India and M.Tech from GNDEC, Ludhiana, India. She has teaching experience of more than 14 years. She is a life member of ISTE. Her areas of interest include Digital image and signal processing, Artificial intelligence, Microprocessors and controllers. She has various research publications in quality national and international conferences and journals.

Kavita Khanna

Dr. Kavita Khanna is, at present, an associate professor with North Cap University, Gurgaon, Haryana. She has done her doctorate from GGSIPU, Dwarka, New Delhi. She has 17 years of teaching experience during which she has published more than 30 research papers and guided 20 M. Tech students in their research work. Her research areas include artificial neural networks, digital image processing, computer graphics and design and analysis of algorithms. Apart from that she is working as a Radio Jockey with All India Radio FM Gold.

Navin Rajpal

Prof. Navin Rajpal is Professor at USICT, GGSIPU, New Delhi. He did his B. Sc. (Engineering) in Electronics & Communication from R.E.C. Kurukshetra, now known as NIT, Kurukshetra. He did his M. Tech and Ph.D. from Computer Science & Engineering Department, IIT, Delhi. He served in various capacities and has more than 24 years of experience in teaching and research. He has worked as Senior Scientific Officer for more than eight years at Centre for Applied Research in Electronics IIT Delhi on various sponsored and Consultancy projects. Before joining this university in July 2000 as Reader, he worked for more than 4 years as Assistant Professor at C.R. State College of Engineering, Murthal, where he was In-charge Computer Science and Engineering Department for about two years. In G.G.S. Indrapratha University he has worked as Reader, USICT and In-charge Computer Center from July 2000 to August 2004. He has also worked as Head CS&E at IGIT from Jan. 2005 to Dec. 2007. He has supervised several M. Tech. and nine Ph.D. Students. He has published / presented more than 100 research papers in National and International Journals / Conferences. He is a life member of CSI and ISTE. His areas of interest are Computer Vision, Image Processing, Pattern Recognition, Artificial Neural Networks, Computer Graphics, Algorithms Design and Digital Hardware Design.

Multi-agent Systems for Arabic Handwriting Recognition

Youssef Boulid*, Abdelghani Souhar and Mohamed Elyoussfi Elkettani

University Ibn Tofail, Faculty of science, 14000, Kenitra (Morocco)

Received 1 March 2017 | Accepted 17 March 2017 | Published 30 March 2017

Abstract

This paper aims to give a presentation of the PhD defended by Boulid Youssef on December 26th, 2016 at University Ibn Tofail, entitled "Arabic handwritten recognition in an offline mode". The adopted approach is realized under the multi agent paradigm. The dissertation was held in Faculty of Science Kénitra in a publicly open presentation. After the presentation, Boulid was awarded with the highest grade (Très honorable avec félicitations de jury).

Keywords

Artificial Intelligence, Arabic Handwritten Documents, Handwriting Recognition, Image Processing, Pre-treatment, Segmentation, Feature Extraction, Recognition, Multi-Agent Systems.

DOI: 10.9781/ijimai.2017.03.012

I. Introduction

On December 26th, 2016, Boulid Youssef defended his PhD thesis related with Arabic handwritten recognition [1]. The thesis was supervised by Prof. Mohamed Elyoussfi and co- supervised by Prof. Abdelghani Souhar. The assessing committee of the PhD dissertation was composed of Prof TOUAHNI Raja, Prof SADIQ Abdelalim, Prof AIT KERROUM Mounir, Prof BENATTOU Mohammed from Faculty of Science, Kénitra and Prof TABII Youness from National School of Applied Sciences, Tétuan. The thesis has been read and approved by this committee. All of them were present at the presentation. The main publications associated with the PhD thesis are [2-5].

II. Thesis Summary

Handwritten recognition is a very broad subject of research and depending on the quality of the document to recognize, there is a multitude of problems that can be encountered. The pattern recognition process is often adopted in the design of handwriting recognition systems, which consists mainly of four stages: the pre-processing stage, which concerns the preparation of the document in terms of normalization and suppression of noise, the segmentation stage, which concerns the detection of lines, words and then the segmentation of those words into characters. The third stage concerns the feature extraction, in which the designer must choose or design the adequate characteristics to be extracted from the character that allow to minimize the intra-class variance while maximizing the interclass variance. The fourth stage involves learning and testing, in which a learning algorithm is used to recognize new letters or new words based on those already learned. To these stages is added a post-processing stage which concerns the verification of the recognized words using a lexical and semantic analysis.

Several researchers propose techniques that respond under certain conditions to a precise problem of a given stage of the process. While the major problem lies in the collaboration between these different techniques since the followed process is often executed in a sequential manner, which is a handicap since the errors in the first stages will propagate in the following stages and thus influence the result of the recognition.

As for a human reader, he has several strategies when facing a document, such as: complete reading (word-by-word reading), the inspection (searching in specific regions of the document), and the overview of the document which give him the capability to read documents he has not seen before.

From this point of view we are interested in analyzing the problems of the recognition of handwritten document by taking inspiration from the mechanisms of what we think the human reader uses during the reading process. This problem is modeled under the multi-agent systems paradigm while taking into consideration the specific characteristics of the Arabic language.

In this context, the contribution of the thesis concerns the recognition of handwritten Arabic documents and precisely the pre-processing, the line segmentation and the character recognition stages [1].

Generally, there are two ways for document noise removal; either detection and suppression of noise, this is possible when the patterns of the noise have independent characteristics that could differentiate them from the textual content, or extraction of textual content while ignoring the noise, in this case the contextual information and the prior

* Corresponding author.
E-mail address: y.boulid@gmail.com

knowledge about the text are required.

For the problem of text line segmentation, generally there are three approaches: the first one focuses on the regions separating the text lines, the second one searches the connected component that constitutes the lines, while the third one searches for the baseline of each word and regroup those that participate in the same line.

Feature extraction methods could be classified in two categories: the structural features, that extract geometrical and topological properties such as the number and position of diacritical points, number of connected components, presence of loops, orientation of curves, location of intersections; and the statistical features such as histograms of projection profile of transition, moments, histogram of gray level distribution, Fourier descriptor, freeman chain code…etc.

Based on the mechanisms of scanning, which humans can use when reading a document, we have divided the problem of noise removal into two collaborative agents. The first one is responsible for the estimation of global parameters of the document and the creation and affectation of noise removal agents into different regions of the document. Based on the nature of the Arabic script, we have found that the notion of intersection (pixel position between characters in a cursive word) could be used in a manner that allows us to distinct between textual and non-textual content. The percentage of intersections according to the total area of the component is higher than 50% in the case of noise. The treatments that the agents execute are: the suppression of salt and pepper noise based on the average stroke width of the text, the classification of content into textual and non-textual based on the percentage of intersections and finally, the distinction of noise similar to text using the contextual information [2].

Based on the mechanisms of reading word after word, we have modeled the problem of line segmentation as an agent based on utility that integrates the Markov Decision Process. The proposed approach detects the connected component in the same line by using knowledge about characteristics and disposition of the components in the document [3].

Problems may occur when adjacent line are touching due to narrow gaps between them, where we may find words that belongs to different lines, but are linked together. To overcome this and inspired by the mechanisms of perception involved in the process of reading, we have modeled the problem as three collaborative agents. The first one is responsible for the estimation of global parameters of the document and for the line extraction. The second one is responsible for the detection of the first component in the line and the detection of components that belong to the same line. As to the third one is responsible for splitting and segmentation of touching characters and words [4].

Based on the fact that Arabic is written from right to left, we have found that extracting features from the right portion rather than from the whole character's image allows enhancing the recognition rate. The Arabic script uses the information about the baseline to differentiate between some character having similar shapes. Once the word is correctly segmented, we recognize each one of its characters, but when the character is isolated we no longer have the information about the baseline. Also, the existing datasets for Arabic handwritten letters do not integrate this information. Extracting textural and structural features after a proper decomposition of the character allows us to increase the recognition rate. This solution allows compensating the lack of the baseline information [5].

To overcome the problem that resides in the traditional approaches, which is the use of the phases in the recognition process in a sequential manner, we have proposed an agent-based modeling offering the possibility to implement different strategies of human reading. According to the homogeneity of the document, this latter is divided into regions, where each region contains a set of agents from different level of the recognition process that could collaborate locally and also between different regions. Each agent has a memory allowing him to track different actions that he performs and the possibility to undo an action in order to correct it if necessary.

The obtained results are encouraging, although we are still in the preliminary stages of the design of a handwritten Arabic recognition system achieving at least the human performance.

Finally, the need of a platform, allowing collaboration between the different stages of the recognition process, is necessary. We believe that such a platform should be based on multi-agent systems offering the possibility of implementing and integrating the different recognition stages in parallel.

References

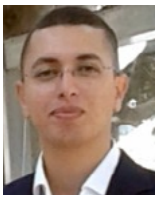
[1] Boulid, Y. Arabic handwritten recognition in an offline mode. PhD thesis, Faculty of Science Kénitra, University Ibn Tofail, December 2016.

[2] Boulid, Y., & Elkettani, M. Y. (2014). Approach for Arabic Handwritten Image Processing: Case of Text Detection in Degraded Documents. International Journal of Computer Applications, Vol. 101, N° 14, Pages 35-42.

[3] Boulid, Y., Souhar, A., & Elkettani, M. Y. (2016). Detection of Text Lines of Handwritten Arabic Manuscripts using Markov Decision Processes. International Journal of Interactive Multimedia and Artificial Intelligence, (Special Issue on Artificial Intelligence Underpinning), Vol. 4, Nº1 Pages 31-36.

[4] Boulid, Y., Souhar, A., & Elkettani, M. Y. (2016). Segmentation approach of Arabic manuscripts text lines based on multi agent systems. International Journal of Computer Information Systems and Industrial Management Applications, Vol. 8, Pages 173 - 183.

[5] Boulid, Y., Souhar, A., & Elkettani, M. Y. (2017). Handwritten Character Recognition Based on the Specificity and the Singularity of the Arabic Language. International Journal of Interactive Multimedia and Artificial Intelligence, (Regular Issue), Vol. 4, Nº4 Pages 45-53.



Youssef Boulid

Youssef Boulid received his M.S. degree in Decision Support Systems and Project Management in 2012 and PhD degree in Computer Science in 2016 from University Ibn Tofail, Faculty of science, Kenitra- Morocco. Now he works as a researcher at Maxware Technology Kénitra-Morocco. His research interests include image processing, handwritten document analysis, Arabic handwritten recognition and artificial intelligence.

Abdelghani Souhar

Abdelghani Souhar received M.S. degree in applied Mathematics in 1992, PhD degree in computer science in 1997 from University Mohammed 5 in Rabat - Morocco. Now he is a Professor at university Ibn Tofail in Kenitra - Morocco. His research interests include CAE, CAD and Artificial Intelligence.

Mohamed Elyoussfi Elkettani

Mohamed Elyoussfi Elkettani received M.S. degree in applied mathematics in 1980 and PhD degree in Statistics in 1984 from Orsay Faculty of Science, University of Paris XI. Now he is a Professor at university Ibn Tofail in Kenitra - Morocco. His research interests include Multivariate statistics and Image recognition algorithms.

N-grams Based Supervised Machine Learning Model for Mobile Agent Platform Protection against Unknown Malicious Mobile Agents

Pallavi Bagga[1], Rahul Hans[2], Vipul Sharma[3] *

[1] Assistant Professor, Department of Computer Science and Engineering, Lovely Professional University, Phagwara, Punjab (India)
[2] Assistant Professor, Department of Computer Science and Engineering, DAV University, Jalandhar, Punjab (India)
[3] Assistant Professor, Department of Computer Science and Engineering, Jaypee University of Information Technology, Solan (India)

Received 8 February 2017 | Accepted 20 March 2017 | Published 30 March 2017

unir
LA UNIVERSIDAD EN INTERNET

Abstract

From many past years, the detection of unknown malicious mobile agents before they invade the Mobile Agent Platform has been the subject of much challenging activity. The ever-growing threat of malicious agents calls for techniques for automated malicious agent detection. In this context, the machine learning (ML) methods are acknowledged more effective than the Signature-based and Behavior-based detection methods. Therefore, in this paper, the prime contribution has been made to detect the unknown malicious mobile agents based on n-gram features and supervised ML approach, which has not been done so far in the sphere of the Mobile Agents System (MAS) security. To carry out the study, the n-grams ranging from 3 to 9 are extracted from a dataset containing 40 malicious and 40 non-malicious mobile agents. Subsequently, the classification is performed using different classifiers. A nested 5-fold cross validation scheme is employed in order to avoid the biasing in the selection of optimal parameters of classifier. The observations of extensive experiments demonstrate that the work done in this paper is suitable for the task of unknown malicious mobile agent detection in a Mobile Agent Environment, and also adds the ML in the interest list of researchers dealing with MAS security.

Keywords

Classification, Malicious Mobile Agents, Nested Cross Validation, N-gram Feature Extraction.

DOI: 10.9781/ijimai.2017.03.013

I. Introduction

A collection of executable programs known as a Mobile agent (MA) migrates from one execution platform to another in a heterogeneous network to perform various tasks on the behalf of its user [1]. The employment of mobile agents introduce many benefits to the distributed computing including network load reduction, overcoming network latency, executing dynamically, asynchronously and autonomously [2]. In many respects, a mobile agent is analogous to a computer virus, since it travels from one computer to another and it utilizes computer resources or it creates clones of itself to achieve its goals. The major difference between both is the usefulness of mobile agent and its friendly behavior. However, the mobile agents while moving in the network, brings with them the fear of Trojan horses, viruses and other invasive means or entities [3]. This is because the attacks can be occurred when the mobile agent traverses in the communication channel and there may be some muggers earwigging the network either to gain some of the information carried by the agent or information stored in the agent platform (i.e. passive attack) or mutating that information for their own advantage (i.e. active attack) [4, 37]. In recent years, numerous researchers have done considerable studies in order to prevent malicious mobile agents causing any harm to Mobile Agent Platform (MAP).

Wahbe et al. [11] proposed a Sandboxing technique, which offers an isolated environment (or a restricted area) for the execution of suspected mobile agents. This isolation prevents the mobile agent from accomplishing specific code exercises, for example local file system interaction, and accessing system properties. Noordende et al. [12] proposed a Mansion API where the agents execute in a protected environment like Sandboxing technique. Additionally, the agents are authenticated based on the trust level between agent owners as well as platform owners. Marikkannu et al. [13] suggested a Dual checkpoint mechanism involving two gates, inner and outer for the mobile agent verification consisting of Digital signatures as well as checksum ensuring the validity of a mobile agent. Alfalayleh et al. [14] recommended a Code Signing mechanism in which the sign of originator on code is checked by agent platform for verifying that it has not been modified. Lee et al. [15] proposed a technique in which the agent byte code compiles the proof carried by mobile agent with the platform's security policy. Upon receipt, the agent platform employs a proof checker for the purpose of checking and verifying the security proof of incoming agent byte code. Ordille [16] proposed the use of Path history that enables the platform either to run

* Corresponding author.
E-mail addresses: pallavibagga315@gmail.com (Pallavi Bagga), rahulhans@gmail.com (Rahul Hans), vipuls85@gmail.com (Vipul Sharma).

the agent or discard it; and to decide the trust level, privileges, resources and services that should be acknowledged to the agent if it is permitted to. Path History contains the identities of the current platform as well as the next platform in the itinerary. Cao et al. [17] proposed the use of agent path history information on the role activation and permission activation. The roles activated for an agent will be filtered by path patterns whereas the permissions for roles will be finely tuned by a set of host patches. Furthermore, Idrissi et al. [34, 36] proposed an authentication process based on Diffie–Hellman Key Exchange integrated with digital signature DSA to prevent the vulnerabilities arisen due to the unavailability of authentication, which makes it well resistant to the Man in the Middle attack; as well as another mobile agent platform security technique based on Elliptic Curve Cryptography (ECC) and dynamic role assignments using Role Based Access Control (RBAC) policy.

Venkatesan et al. [18] proposed Malicious Identification Police (MIP) that uses Attack Identification Scanner (AIS) to scan the incoming agent byte code in order to diagnose the maliciousness in it. In Policy Based MIP proposed by Venkatesan et al. [19], the privileges of an agent are also checked in addition to AIS [18], to know if it wants to do more than the privileges granted to it. Otherwise, Intelligent AIS (IAIS) decides to start the lexical analyzer by its own decision, where agent byte code is turned into tokens and diagnose the non-match tokens by comparing with tokens present in the Knowledge Base (KB). Afterwards, unknown tokens are executed and tested in an isolated environment to check for their malicious intentions and updates the KB containing malicious codes, with the newly diagnosed (if any) vicious code. In order to fend off this waiting time of the agent, Venkatesan et al. [20] further included the agent clones to handle multiple incoming agents simultaneously. Additionally, the pipelining concept was introduced by separating the operations i.e. the tasks of scanning, pattern extracting and detecting unknown codes are performed by different agents, which ultimately reduces the time complexity. Clearly, many researchers have been buckled down in the field of MAP security. However, the unknown malicious mobile agent detection before invading the MAP is still a challenge and a concern owing to the growth of malicious agents in recent years.

Nowadays, malicious code detection techniques employ one of these two approaches: Signature-based or Behavior-based. Signature-based methods involve the identification of distinctive tokens in the binary code [5]; whereas Behavior-based methods rely on the rules created by the experts that define the malicious behavior or non-malicious behavior of code [6]. While being very precise, signature-based methods are unable to diagnose previously unknown malicious codes whereas behavior-based methods can only detect the presence of malicious content after the code has been executed [7]. Realizing the necessity of a detection method for the unknown malicious code, in recent years, the machine learning algorithms or Classification Algorithms were magnificently employed which was highly inspired by the Text categorization problem [8]-[9], [23]-[25].

In this paper, an attempt has been made for detecting unknown malicious mobile agents using Machine Learning algorithms, which represents a novel contribution in the field of MAP security as per the survey done by the authors. This attempt addresses several facets of the detection challenge: mobile agent representation, classification and performance evaluation. The present work is also influenced by the objective to achieve very high classification accuracy rate while maintaining the low false negatives (i.e. misclassifying malicious agent as non-malicious). Though there are various representations of executable files: "Portable Executable (PE)", "Byte Sequence n-grams", and "plain-text string features" [9]; in this paper, n-gram representation of the agent executable is considered to be used as features for the classification process, since an extensive n-grams analysis is also one of the major focuses of this paper. N-grams are overlapping substrings obtained in a sliding window fashion [10].

The extracted n-gram features are then fed into four commonly used Classification algorithms: Naive Bayesian, SMO, IBK, J48 Decision Tree, for discriminating between two categories of agent classification (malicious mobile agent and non-malicious mobile agent), which is supported by WEKA tool [21]. The extensive experiments are performed on a collection of 80 files, in which half of the total files are malicious. The experimental results are evaluated based on standard performance evaluation measures such as "Sensitivity Rate", "Specificity Rate", "Positive Predictive Value", "Negative Predictive Value", "F-score", "Receiver Operating Characteristics – Area Under Curve", "Miss Rate", "Fall out" and "Accuracy Rate", while employing the 5-fold nested cross validation scheme.

II. Material and Methods

A. Dataset Used

To the best of author's knowledge, there is no standard data set available for the detection of malicious mobile agents. Therefore, the benchmark dataset of malicious files known as CSDMC2010[1] API sequence corpus, containing Windows API/System-Call trace files, is selected for the purpose of classification. The dataset contains 388 files involving 320 malware traces as well as 68 benign traces (considered as non-malicious in this paper). For the training dataset, only 40 malicious files and 40 non-malicious files are collected after random sampling (equal number for malicious and non-malicious files is considered in order to avoid the Class-imbalance problem). This standard dataset is preferable for the proposed approach since agent byte code can be viewed as a sequence of agent API function calls. This assumption is made on account of the previous studies of extracting API call sequences from byte codes [31],[32].

B. Performance Evaluation Measures

To evaluate the classification performance of detecting malicious mobile agents successfully, it is necessary to identify appropriate performance metrics. The measures derived from the Confusion Matrix (Figure 1) to calculate and be applied to classifier evaluation are described in Table I [26]. The confusion matrix indicates the correct and incorrect classification outcomes predicted by the classifier when compared with the actual classification outcome. The measures other than Accuracy Rate and Misclassification Rate are considered to figure out whether the present framework holds good for the classification of either malicious mobile agents or non-malicious mobile agents or both.

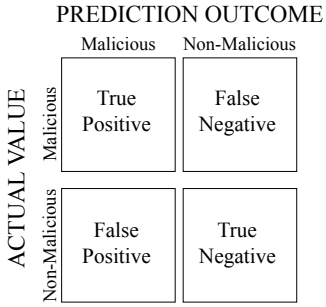


Fig. 1. Confusion Matrix to evaluate the performance of classifier.

- *True Positives* (TP): Number of malicious agents classified as malicious.
- *True Negatives* (TN): Number of non-malicious agents classified as non-malicious.
- *False Negatives* (FN): Number of malicious agents classified as non-malicious.

1 http://www.csmining.org/index.php/malicious-software-datasets-.html

TABLE I. PERFORMANCE EVALUATION MEASURES FOR CLASSIFICATION OF MALICIOUS MOBILE AGENTS

Metric	Definition	Formula	Expected Value
Sensitivity	Also known as True Positive Rate (TPR) or Recall. It evaluates the ability of a classifier to correctly identify an agent as malicious.	$TP/(TP+FN)$	Maximum
Specificity	Also known as True Negative Rate (TNR). It evaluates the ability of a classifier to correctly identify an agent as non-malicious.	$TN/(TN+FP)$	Maximum
Positive Predictive Value (PPV)	Also known as Precision. It is the percentage of agents classified as malicious which are truly malicious.	$TP/(TP+FP)$	Maximum
Negative Predictive Value (NPV)	It is the percentage of agents classified as non-malicious which are truly non-malicious.	$TN/(TN+FN)$	Maximum
Miss Rate	Also known as False Negative Rate (FNR). It evaluates the proportion of malicious agents that are classified as non-malicious.	$FN/(TP+FN)$	Minimum
Fall out	Also known as False Positive Rate (FPR). It evaluates the proportion of non-malicious agents that are classified as malicious.	$FP/(FP+TN)$	Minimum
ROC-AUC	The curve is drawn by plotting the TPR against the FPR at different threshold settings.	*NA*	Between 0.9 and 1
Accuracy	It evaluates the ability of a classifier in classifying the whole dataset.	$(TP+TN)/(TP+FP+TN+FN)$	Maximum
F-measure	Also known as F-score. It is an evaluation of classifier's accuracy, which combines both the precision as well as the recall as a harmonic mean.	$2.Precision.Recall/(Precision+Recall)$	Maximum

Note: NA means Not Applicable

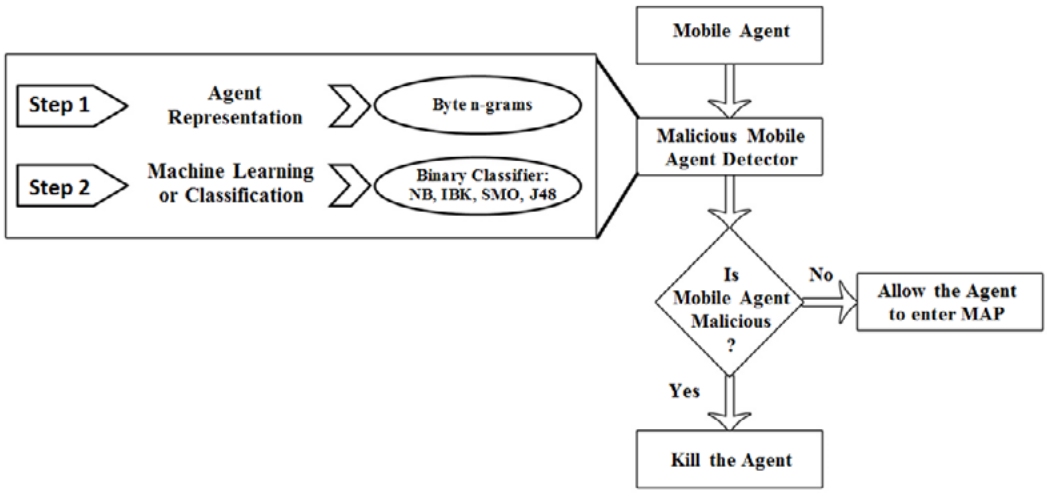


Fig. 2. Present Methodology for Malicious Mobile Agent Detection.

- *False Positives* (FP): Number of non-malicious agents classified as malicious.

C. Methodology

The major objectives of present methodology are as follows:

- To automate the detection of malicious mobile agents before they conquer the Mobile Agent Platform.
- To evaluate the performance of n-gram representation of mobile agent.
- To use Machine Learning algorithms for the task of unknown malicious mobile agent detection.
- To scrutinize the performance of various classifiers for classifying the mobile agents.

The methodology used in this paper is shown in Figure 2. It mainly consists of two consecutive steps: n-gram feature extraction of mobile agent and classification. These steps are described in detail in subsequent sub-sections.

1. Mobile Agent Representation using Byte n-grams – Data Preparation

A standard n-gram analysis is used to extract features from the malicious and non-malicious files. This method is purely machine-learning based method and exploits Natural Language Processing (NLP) also [30]. The n-grams are extracted in a sliding-window fashion, where a window of fixed length (n) slides one byte at a time. In general, n-grams are all substrings of a larger string with length "n" [24]. In present context, byte n-grams are viewed as API call based features. Many researches in recent years have released the importance of n-gram based methods in malware detection, since this technique of extracting features is simple and easy to implement. Each n-gram is analogous to a word or a term of a text document in the Text Categorization problem. For instance, there are eight 3-grams in the text "abc_dabc_e": "abc", "bc_", "c_d", "_da", "dab", "abc", "bc_" and "c_e". For the preparation of data, the unique n-grams are identified in all the mobile agent files and are merged together. In above example, there are only six distinct n-grams i.e. "abc",

"bc_, "c_d", "_da", "dab", "c_e". The procedure of n-gram extraction repeats for different values of n. To limit the experiments for present study, the varying n-grams are employed with the value of n ranging from 3 to 9 only. This is because if the value of n increases, the number of unique n-gram features also increases. The number of distinct n-grams extracted from dataset files is 1403, 2236, 3074, 4055, 5137, 6445 and 7727 for 3-gram, 4-gram, 5-gram, 6-gram, 7-gram, 8-gram and 9-gram respectively.

2. Classification

Since the unknown mobile agent can be classified either malicious or non-malicious, the Binary Classification is taken into account. The standard commonly used classification algorithms such as Naïve Bayesian [22], Instance based Learner [22], Sequential Minimal Optimization [27]-[29], and J48 Decision Tree [22], are implemented. These classification algorithms differ in performance within different domains. In this paper, the best fitted algorithm for the dataset has been identified by the experimentation as shown in the subsequent section.

i) *IBK*

IBK is a WEKA implementation of k-Nearest Neighbor (k-NN). In general, the nearest neighbor classifiers compare a given test tuple with the identical training tuples. The training tuples are characterized using n features. Each tuple represents a point in an n-dimensional space. Hence, all the training tuples are exemplified in an n-dimensional feature space. When an unknown tuple is given as an input, a k-NN classifier explores the feature space for the closest k training tuples to the unknown tuple [22].

The closeness is defined in terms of Distance Metrics such as Chebyshev distance, Manhattan distance, and Euclidean distance. The unknown tuple is labeled with the most common class among its k-nearest neighbors. The value of k is usually an odd number to avoid tied votes; however, choosing the value of k is very analytical. The smaller value of k indicates the higher influence of noise on the result whereas the larger value of k makes the classification computationally very expensive. The pseudo code of IBK is shown in Algorithm 1.

Algorithm 1. IBK classification algorithm

Algorithm IBK(k, X, Y, x)
//Input: k- an integer odd value (number of nearest neighbors), X- Training data consisting of n tuples, Y- Class Labels of X, x- Unknown tuple
//Output: Class label of x
1. **for** i ← 1 to n do
2. compute distance (X_i, x)
3. **end for**
4. sort the distances in ascending order
5. select the first k points from the sorted list (these are the k nearest training tuples to unknown tuple)
6. **return** class label that belongs to the majority of k selected tuples

ii) *Naïve Bayesian*

This classifier is so called because it relies on the Bayesian Theorem. Moreover, it is called "Naive", because it assumes the independence between every pair of attributes (or features), which is known as "class conditional independence" [22]. The classifier takes an unknown tuple as an input, for which the class label is not known, and returns a class label as an output for which the maximum probability is obtained as per the probabilistic calculations. This classifier is particular suited for the higher dimensionality of features. The pseudo code of Naïve Bayesian is shown in Algorithm 2.

Algorithm 2. Naïve Bayesian classification algorithm

Algorithm naiveBayesian(D,X,m,n,A)
//Input: D - Training set of tuples and their associated class labels, n- number of attributes, A - set of attributes (A_1,A_2,….A_n), X - n-dimensional attribute vector (x_1,x_2,….x_n), m – number of classes
//Output: Class of tuple X
1. **for** i ← 1 to m do
2. calculate the Posterior probability conditioned on X i.e. $P(C_i|X)$ using
Bayes' theorem

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \quad (1)$$

where in Equation (1),
P(X) is constant for all classes,
$P(C_i) = \frac{|C_{i,D}|}{|D|}$, where $|C_{i,D}|$ is the number of training tuples of Class C_i in D

$$P(X|C_i) = \prod_{k=1}^{n} P(x_k|C_i) = P(x_1|C_i) \times P(x_2|C_i) \times \ldots \times P(x_n|C_i) \quad (2)$$

and,in Equation (2), x_k refers to the value of attribute A_k for tuple X
3. **end for**
4. **return** class label with maximum $P(C_i|X)$

iii) *J48 Decision Tree*

A Decision tree classifies the tuples as per a set of tree-structured if-then-rules. Each internal node represents a test on an attribute (or a feature), each branch represents the outcome of test, whereas each leaf node holds a class label. J48 is a WEKA implementation of the C4.5 decision tree. It consists of two steps: the decision tree induction and the tree pruning [22]. The pseudo code of decision tree algorithm is shown in Algorithm 3.

Algorithm 3. J48 Decision Tree classification algorithm

//Input: D- set of training tuples and their associated class labels,
A- attribute list
//Output: Decision Tree
1. create a node N
2. **if** all tuples of D belong to same class C, **then**
3. **return** N as a leaf node labeled with class C and terminate.
4. **if** A is empty, **then**
5. **return** N as a leaf node labeled with the most common class in D
(Majority Voting)
6. apply Gain Ratio feature selection method to find the best criterion 'a' ∈ A
7. label N with 'a'
8. **for** each value j of 'a' **do**
9. grow a branch from N with condition a=j
10. let Dj be the set of tuples in D with a=j
11. **if** Dj is empty **then**
12. add a leaf node labeled with most common class in D to N
13. **else** add the node returned by decisionTree(Dj, A-a) to N
14. **end for**
15. **return** N

iv) *SMO*

SMO is a WEKA implementation of Support Vector Machine (SVM). SVM was originated from statistical learning theory with the objective to find the solution of interested problem without solving a more difficult problem as an intermediate step [27]. The statistical learning theory offers a framework that helps to choose the hyper plane space such that it diligently symbolizes the underlying function in the target space [28].

SVM avoids over-fitting to the training dataset.

Given a training set of instance-label pairs (x_i, y_i), i=1,…,N where $x_i \in R^n$ and $y \in \{1,-1\}^l$, SVM finds the solution of an optimization problem as shown in Algorithm 4 [29]. The training tuples x_i are represented into a higher (possibly infinite) dimensional space using function ∅. In other words, the kernel ∅ is used to transform data from input to the feature space. There are four basic kernels: Linear, Polynomial, Radial Basis Function (RBF) and Sigmoidal.

Algorithm 4. SMO Optimization Problem

Optimization Problem:
Given Constraints in Equation (3):

$$y_i(w^T \emptyset(x_i) + b) \geq 1 - \xi_i \text{ and } \xi_i \geq 0 \quad (3)$$

Minimize the Error Function as mentioned in Equation (4):

$$\frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i \quad (4)$$

where,
C>0 is the penalty parameter of the error term, and is called Capacity Constant. The choice of C is made carefully in order to avoid over-fitting. Larger the C, more the error is penalized.
w is the vector of coefficients.
b is a constant.
ξ_i represents parameters for handling non-separable data (inputs)
i labels the N training cases
x_i represents the independent variables.

III. Results and Discussions

The classification of mobile agent into two categories based on their n-gram features has been performed on 80 agent files of dataset of API calls sequence. An extensive setting of parameters is done to optimize the performance of each classification algorithm (NB, SMO, IBK and J48), such as "value of k", "distance measure", or "nearest neighbor search algorithm" in IBK, "pruning", or "confidence factor" in J48 decision tree, "complexity parameter", or "kernel" in SMO. The nested five-fold cross validation scheme is performed to obtain unbiased evaluation results [33]. In nested five-fold cross validation method, the data is randomly divided into five disjoint folds. The four folds are used for tuning of classifier parameters (using cross validation scheme) and then the tuned classifier is validated on left out fold.

This procedure repeats for five times, each time with different left-out folds. This nesting of cross validation loops avoids so-called resubstitution-bias [33]. Additionally, the standard parameters such as Sensitivity, Specificity, ROC, PPV, NPV, FNR, FPR, F-measure and Accuracy, evaluate the performance results and the results of all iterations are averaged to get the final outcome. It has been evidenced that the performance of present work highly depends on the choice of classifier.

The results of various classifiers (NB, SMO, IBK and J48) for different values of n for n-grams i.e. n=3 to 9 have been investigated and are presented in Table II. To help with nested cross validation, WEKA tool has been used to adjust the classifier settings repeatedly in order to get the results on suitable parameter values. The results demonstrate that IBK classifier, J48 classifier, SMO classifier and NB classifier gives maximum accuracy rate of 96.25%, 97.50%, 95.00% and 93.75% respectively (as shown in Figure 3), while maintaining the miss rate of 2.50%, 2.50%, 5.00% and 7.50% for trigrams, 9-grams, 5-grams to 9-grams and 7-grams respectively (as depicted in Figure 4) in distinguishing malicious files from non-malicious. The value for Area under ROC curve is more than 0.92 for each classifier. IBK gives maximum sensitivity of 97.50 % and specificity of 95.00% for 3-gram features. J48 gives maximum sensitivity of 97.50 % as well as specificity of 97.50% for 9-gram features as shown in Figure 5 and Figure 6. Moreover, the highest values of PPV and NPV (97.50% each) belong to J48 classifier for 9-grams.

TABLE II. Results of Different Classifiers for Different Values of N in N-Gram

n	Classifier	Accuracy Rate (%)	Miss Rate (%)	Fall out (%)	PPV (%)	NPV (%)	F-measure (%)	ROC area
3	Bayesian	85.00	15.00	15.00	85.00	85.00	85.00	0.94
	SMO	92.50	7.50	7.50	92.50	92.50	92.50	0.93
	J48	95.00	2.50	7.50	92.86	97.37	95.12	0.94
	IBK	96.25	2.50	5.00	95.12	97.44	96.30	0.97
4	Bayesian	87.50	12.50	12.50	87.50	87.50	87.50	0.93
	SMO	93.75	5.00	7.50	92.68	94.87	93.83	0.94
	J48	93.75	5.00	7.50	92.68	94.87	93.83	0.94
	IBK	93.75	5.00	7.50	92.68	94.87	93.83	0.94
5	Bayesian	90.00	12.50	7.50	92.11	88.10	89.74	0.94
	SMO	95.00	5.00	5.00	95.00	95.00	95.00	0.95
	J48	95.00	2.50	7.50	92.86	97.37	95.12	0.94
	IBK	93.75	5.00	7.50	92.68	94.87	93.83	0.93
6	Bayesian	91.25	10.00	7.50	92.31	90.24	91.14	0.95
	SMO	95.00	5.00	5.00	95.00	95.00	95.00	0.95
	J48	96.25	2.50	5.00	95.12	97.44	96.30	0.95
	IBK	92.50	2.50	12.50	88.64	97.22	92.86	0.93
7	Bayesian	93.75	7.50	5.00	94.87	92.68	93.67	0.96
	SMO	95.00	5.00	5.00	95.00	95.00	95.00	0.95
	J48	95.00	5.00	5.00	95.00	95.00	95.00	0.94
	IBK	92.50	2.50	12.50	88.64	97.22	92.86	0.93
8	Bayesian	92.50	7.50	7.50	92.50	92.50	92.50	0.96
	SMO	95.00	5.00	5.00	95.00	95.00	95.00	0.95
	J48	96.25	2.50	5.00	95.12	97.44	96.30	0.96
	IBK	92.50	2.50	12.50	88.64	97.22	92.86	0.94
9	Bayesian	92.50	10.00	5.00	94.74	90.48	92.31	0.97
	SMO	95.00	5.00	5.00	95.00	95.00	95.00	0.95
	J48	97.50	2.50	2.50	97.50	97.50	97.50	0.97
	IBK	92.50	5.00	10.00	90.48	94.74	92.68	0.93

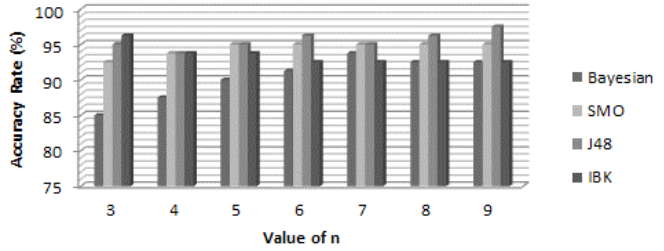


Fig. 3 Graph showing Accuracy Rate of Different Classification Algorithms.

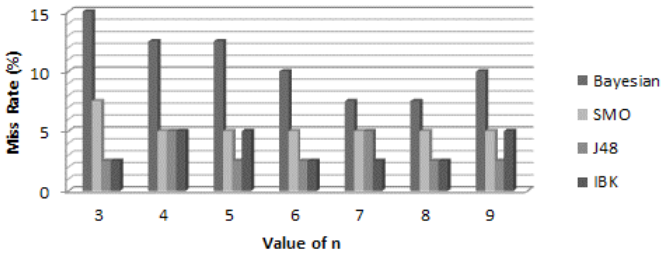


Fig. 4 Graph showing Miss Rate of Different Classification Algorithms.

This means IBK and J48 decision tree are the best in distinguishing actual malicious agents (positives) and actual non-malicious agents (negatives) using 3-gram and 9-gram features respectively. The performance of other classification algorithms reduces with the increase in number of features.

It is shown in Figure 5 that the sensitivity rate increases up to n=7 using NB classifier and then decreases. Using SMO, the sensitivity rate

increases up to n=4 and then remains the same with further increase in value of n. Using J48 and IBK classifiers, sensitivity rate is minimum for n=3, which decreases and then again increases with increase in value of n. The Figure 6 indicates that the minimum specificity rate (85.00%) is obtained using NB classifier for 3-grams, which further increases with the increase in value of n. SMO and J48 classifiers provide the minimum specificity rate of 92.50%. From 5-gram to 9-grams, specificity is constant with SMO classifier whereas from 6 to 8-grams, IBK increases with increase in value of n.

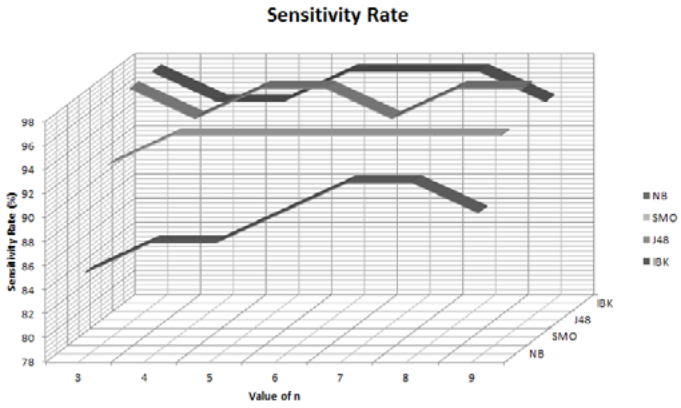


Fig. 5 Graph showing Sensitivity Rate of Different Classification Algorithms.

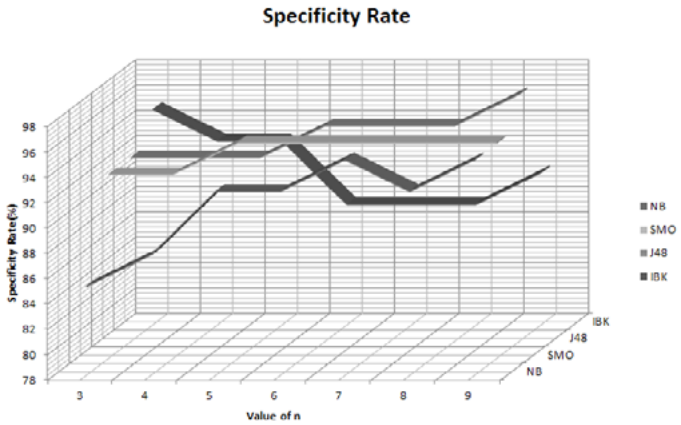


Fig. 6 Graph showing Specificity Rate of Different Classification Algorithms.

IV. Conclusions and Future Scope

This paper aims for probing the suitability of machine learning algorithms for the task of classification of mobile agent either malicious or non-malicious in a Mobile Agent Environment using a specific dataset. In particular, n-grams are used as features during the classification process. Different classification algorithms used are Naïve Bayesian, Sequential Minimization Optimization, Instance Based Learner, and J48 Decision Tree. J48 decision tree algorithm gives higher accuracy rate of 97.50% and miss rate of 2.50% for 9-grams whereas IBK gives higher accuracy rate of 96.25% and miss rate of 2.50% for 3-grams, as compared to other classification algorithms. In other words, out of all the classifiers, J48 gives better result with 7727 features whereas IBK with 1403 features, but J48 increases the computational cost due to large feature space. Therefore, IBK is declared to be the best classifier. Clearly, the optimistic results boost the use of present research for MAP protection.

In near future, the work can be extended with the use of more different classifiers with higher values of n for n-gram features in order to evaluate the performance of classification task. Moreover, large number of n-gram features burdens the classification process; therefore, feature selection methods can be applied in future. The work can even be done on different datasets, since the classifiers may give different results on different datasets.

References

[1] A. Aneiba, and S.J. Rees, "Mobile Agents Technology and Mobility", in *Proceedings of the 5th Annual Postgraduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting,* 2004. pp. 14-20.

[2] D.B. Lange, and M. Oshima, "Seven good reasons for Mobile Agents", *Communications of the ACM*, vol. 42, no. 3, pp. 88-89, 1999.

[3] L.L. Thomsen, and B. Thomsen, "Mobile Agents – The new paradigm in computing", *ICL- The Systems Journal*, vol. 12, pp. 14-40, 1997.

[4] R. Oppliger, "Security issues related to mobile code and agent-based systems", *Computer Communications*, vol. 22, no. 12, pp. 1165-1170, 1999.

[5] D. Venugopal, and G. Hu, "Efficient signature based malware detection on mobile devices", *Mobile Information Systems*, vol. 4, no. 1, pp. 33-49, 2008.

[6] W. Ma, P. Duan, S. Liu, G. Gu, J. Liu, "Shadow attacks: automatically evading system-call-behavior based malware detection", *Journal in Computer Virology*, vol. 8, no. 1, pp. 1-13, 2012.

[7] R. Moskovitch, C. Feher, N. Tzachar, E. Berger, M. Gitelman, D. Shlomi, and Y. Elovici, "Unknown Malcode Detection Using OPCODE Representation", *Intelligence and Security Informatics*, vol. 5376, pp. 204-15, 2008.

[8] A. Shabtai, R. Moskovitch, Y. Elovici, and C. Glezer, "Detection of malicious code by applying machine learning classifiers on static features: A state-of-the-art survey", *Information Security Technical Report*, vol. 14, no. 1, pp. 16-29, 2009.

[9] J.Z. Kotler, and M.A. Maloof, "Learning to detect malicious executables in the wild", in *Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* Aug 22, New York, ACM Press, pp. 470-478, 2004.

[10] S. Jain, and Y.K. Meena, "Byte Level n-Gram Analysis for Malware Detection", in *Proceedings of the 5th International Conference on Information Processing*, ICIP 2011, Aug 5-7, Banglore, India, Springer Berlin Heidelberg, pp. 51-59, 2011.

[11] R. Wahbe, S. Lucco, T.E. Anderson, and S.L. Graham, "Efficient Software-Based Fault Isolation", in *Proceedings of the Fourteenth ACM Symposium on Operating Systems Principles*, SOSP '93, Dec 5-8, Asheville, North Carolina, USA, ACM Press, pp. 203-216, 1993.

[12] G. Noordende, F. M. Brazier, and A.S. Tannenbaum, "A security framework for a mobile agent system", in *Proceedings of the 2nd International Workshop on Security in Mobile Multiagent Systems*, SEMAS 2002, Bologna, Italy, pp. 43-50, 2002.

[13] P. Marikkannu, and A. Jovin, "A Secure Mobile Agent System against Tailgating Attacks", *Journal of Computer Science*, vol. 7, no. 4, pp. 488-492, 2011.

[14] M. Alfalayeh, and L. Brankovic L, "An overview of security issues and techniques in mobile agents", in *Proceedings of the 8th IFIP TC-6 TC-11 Conference on Communications and Multimedia Security,* pp. 59-78, 2005.

[15] P. Lee, and G. Necula, "Research on proof-carrying code for mobile-code security", in *DARPA workshop on foundations for secure mobile code,* pp. 26-28, 1997.

[16] J.J. Ordille, "When agents roam, who can you trust?", in *Proceedings of the IEEE First Annual conference on Emerging Technologies and Applications in Communications,* pp. 188-191, 1996.

[17] C. Cao, and J. Lu, "Path-history-based access control for mobile agents," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 21, no. 3, pp. 215-225, 2006.

[18] S. Venkatesan, and S. Chellappan, "Protection of mobile agent platform through Attack Identification Scanner (AIS) by Malicious Identification Police (MIP)", in *Proceedings of the IEEE First International Conference on Emerging Trends in Engineering and Technology (ICETET '08)*, IEEE, pp. 1228-1231, 2008.

[19] S. Venkatesan, C. Chellappan, T. Vengattaraman, P. Dhavachelvan, and Vaish, "Advanced mobile agent security models for code integrity

and malicious availability check", *Journal of Network and Computer Applications*, vol. 33, no. 6, pp. 661-671, 2010.

[20] S. Venkatesan, R. Baskaran, C. Chellappan, A. Vaish, and P. Dhavachelvan, "Artificial immune system based mobile agent platform protection," *Computer Standards & Interfaces*, vol. 35, no. 4, pp. 365-373, 2013.

[21] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reuteman, and I.H. Witten, "The WEKA data mining software: an update", *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10-18, 2009.

[22] J. Han, "Data mining: concepts and Techniques", 2000.

[23] D. Gavriluţ, M. Cimpeşu, D. Anton, and L. Ciortuz, "Malware Detection Using Machine Learning", in *Proceedings of the International Multiconference on Computer Science and Information Technology,* Oct 12, IEEE, pp. 735-741, 2009.

[24] T. Abou-Assaleh, N. Cercone, V. Kešelj, and R. Sweidan, "N-gram-based detection of new malicious code", in *Proceedings of the IEEE 28th Annual International Conference on Computer Software and Applications*, pp. 41-42, 2004.

[25] I. Santos, Y.K. Penya, J. Devesa, and P.G. Bringas. "N-grams-based File Signatures for Malware Detection", in *Proceedings of the 2009 International Conference on Enterprise Information Systems*, pp. 317-320, 2009.

[26] M. Sokolova, and G. Lapalme, "A systematic analysis of performance measures for classification tasks", *Information Processing & Management*, vol. 45, no. 4, pp. 427-437, 2009.

[27] R. Burbidge, B. Buxton, "An Introduction to Support Vector Machines for Data Mining", Keynote Speakers, young OR12, pp. 3-15, 2001.

[28] V. Jakkula, "Tutorial on Support Vector Machine (SVM)", School of EECS, Washington State University, 2006.

[29] C.W. Hsu, C.C. Chang, C.J. Lin, "A Practical Guide to Support Vector Classification", *Technical Report*, Department of computer Science, National Taiwan University, 2003.

[30] C.Y. Suen, "n-Gram Statistics for Natural Language Understanding and Text Processing", *IEEE Transactions on Pattern Analysis*, vol. PAMI-1, no. 2, pp.164-172, 1979.

[31] T.T. Nguyen, H.V. Pham, P.M. Vu, and T. T. Nguyen, "Learning API Usages from Bytecode: A Statistical Approach", in *Proceedings of the 38th International Conference on Software Engineering (ICSE '16),* May 14-22, Austin, Texas, ACM Press, pp. 416-427, 2016.

[32] Y. Qiao, Y. Yang, J. He, C. Tang, and Z. Liu, "CBM: Free, Automatic Malware Analysis Framework Using API Call Sequences", in *Proceedings of the Seventh International Conference on Intelligent Systems and Knowledge Engineering,* Dec 2012, Beijing, China, Springer Berlin Heidelberg, pp. 225-236.

[33] S. Varma, and R. Simon, "Bias in error estimation when using cross validation for model selection", *BMC Informatics*, vol. 7, no. 1, 2006.

[34] H. Idrissi, E. M. Souidi, and A. Revel, "Security of mobile agent platforms using access control and cryptography", in *Proceedings of the 9th KES international conference (KES-AMSTA)*, pp. 27–39, 2015.

[35] H. Idrissi, A. Revel, and E. M. Souidi, "Security of Mobile Agent Platforms using RBAC based on Dynamic Role Assignment", *International Journal of Security and Its Applications*, vol. 10, no. 4, pp.117-134, 2016.

[36] M. Malathy, S.J. Smilee, J. N. Samuel, "Secure Mobile Agent in M-Commerce over Internet", in *Proceedings of International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS), pp. 1-5, 2016.*

Pallavi Bagga

Pallavi Bagga is presently working as an Assistant Professor in the Department of Computer Science and Engineering at Lovely Professional University, Phagwara, Punjab, India. She has accomplished her Master of Technology in Computer Science & Engineering from DAV University, Jalandhar, Punjab, India in 2016. Earlier, she has completed her Bachelor of Technology in Computer Science & Engineering from Punjab Technical University, Punjab, India in 2014. Her predominant areas of research interests include Distributed Computing and Machine Learning. Nowadays, she is actively pursuing her research career in the sphere of Mobile Agents System Security using Machine Learning Algorithms.

Rahul Hans

Rahul Hans is currently working as an Assistant Professor in the Department of Computer Science and Engineering at DAV University, Jalandhar, Punjab, India. He has received his B.Tech degree in Computer Science and Engineering from Punjab Technical University, Jalandhar, Punjab, India and subsequently his M.Tech degree in Computer Science and Engineering from Guru Nanak Dev University, Amritsar, Punjab, India. He has a teaching experience of more than four years. He is also a member of Computer Society of India (CSI). He has published 16 research papers in various International Conferences and Journals. His research areas include Security and Fault tolerance in Mobile Agent Systems in Distributed Computing as well as the Machine Learning.

Vipul Sharma

Dr. Vipul Sharma is currently working as an Assistant Professor (Senior Grade) in the Department of Computer Science and Engineering at Jaypee University of Information Technology, Solan, India. He has received his B.Tech. Degree in Computer Science and Engineering from Punjab Technical University, Jalandhar, Punjab, India and subsequently his M.E. and Ph.D. Degrees in Computer Science & Engineering from UIET, Panjab University, Chandigarh, India. His Ph.D. thesis focused on use of machine learning techniques to associate low level image features with the query concepts, relevance feedback techniques to learn users' intention and feature selection to find an optimal set of features that can not only capture the texture of images but also enhances the semantic interpretability for reducing the semantic gap in medical image retrieval. He is actively pursuing research in Medical Image Processing. He is the awardee of TEQIP-II fellowship during Ph.D. He also received the UIET, Panjab University Research Award-2015 and a cash prize of Rs 10000 for his outstanding publication in the Medical & Biological Engineering & Computing Journal (SCI Journal). He is a member of many professional bodies including Society of Photo-Optical Instrumentation Engineers, USA; Engineering in Medicine & Biology Society; Association for Computing Machinery; IEEE; IEEE Computer Society Technical Committee on Pattern Analysis and Machine Intelligence; Multimedia Computing; IEEE Young Professionals. He has published 16 research articles in reputed journals and conferences. His current research includes development of effective classification and retrieval techniques for the retrieval of medical images from large image databases.

A Topic Modeling Guided Approach for Semantic Knowledge Discovery in e-Commerce

V. S. Anoop[1]* and S. Asharaf[2]

[1] Data Engineering Lab, Indian Institute of Information Technology and Management – Kerala (IIITMK), Thiruvananthapuram (India)
[2] Indian Institute of Information Technology and Management – Kerala (IIITM-K), Thiruvananthapuram (India)

Received 27 February 2017 | Accepted 6 April 2017 | Published 28 April 2017

Abstract

The task of mining large unstructured text archives, extracting useful patterns and then organizing them into a knowledgebase has attained a great attention due to its vast array of immediate applications in business. Businesses thus demand new and efficient algorithms for leveraging potentially useful patterns from heterogeneous data sources that produce huge volumes of unstructured data. Due to the ability to bring out hidden themes from large text repositories, topic modeling algorithms attained significant attention in the recent past. This paper proposes an efficient and scalable method which is guided by topic modeling for extracting concepts and relationships from e-commerce product descriptions and organizing them into knowledgebase. Semantic graphs can be generated from such a knowledgebase on which meaning aware product discovery experience can be built for potential buyers. Extensive experiments using proposed unsupervised algorithms with e-commerce product descriptions collected from open web shows that our proposed method outperforms some of the existing methods of leveraging concepts and relationships so that efficient knowledgebase construction is possible.

Keywords

Text Mining, Latent Dirichlet Allocation, Web Mining, Semantic Graphs, Semantic Web, e-Commerce.

DOI: 10.9781/ijimai.2017.03.014

I. Introduction

A knowledgebase or a relational database storing useful patterns which are extracted from unstructured data such as plain text has great potential in business domains on which a large number of customer centric services can be embedded. Consider a knowledgebase storing factual information about e-commerce products which are extracted from unstructured descriptions available from web. Services such as meaning aware product search and discovery [1] and feature based filtering [2] can be attached to this knowledgebase for enhancing customer shopping experience. Appraising such prospective applications, a large number of researches have been reported in the recent past which focuses on mining large unstructured text repositories for finding useful patterns and leveraging them to a structured form for immediate use by the applications. It is well established that among all other forms of unstructured data, text is considered to be very rich in information and diverse applications are producing and consuming text data. There are still a lot of avenues where we can exploit the text data in its full potential for extracting content rich patterns. As the amount of text that is being generated grows exponentially, we need more efficient and scalable algorithms to process such data.

This is the era of data explosion thus organizations are already flooded with data and the growth of such data is exponentially increasing. Leveraging useful knowledge from such data using traditional algorithms are inefficient and time consuming thus we need more efficient algorithms to cop up with these scenario. In text mining, concepts can be defined as a sequence of words that is used to represent real or imaginary entities found in plain text. Extraction of such concepts is an important step in bringing out useful patterns from text because knowledge engineering applications make use of these concepts for enriching the associated knowledge. Another use of concept mining is that a variety of information retrieval, opinion mining and classification systems make use of such concepts. Even though many such extraction systems are available, when dealing with large amount of data, these algorithms may perform poor when it comes to the extraction of relevant and semantically rich concepts. Thus we need more efficient and intelligent algorithms to work with these large text archives.

Topic modeling algorithms has attained a special interest among text mining researchers and practitioners because of its text understanding nature and the ability to deal with large text archives. Topic models [3] are suite of algorithms which can bring out hidden thematic structures from large text repositories and they are mostly unsupervised when it comes to the learning paradigm. Since its inception, researchers have extended topic modeling to many dimensions and as a result various implementations of the same are reported in the past literatures. Probabilistic topic models [4] and Latent Dirichlet Allocation (LDA) [5] are such implementations of the basic topic modeling. Among these, LDA algorithm is most widely used by text mining enthusiasts because of the assumption it has in modeling topics and also the easiness in integrating it with widely used programming languages such as Python and Java.

* Corresponding author.
E-mail address: anoop.res15@iiitmk.ac.in

Main Contributions of this paper: In this work, the authors propose a new approach for extracting semantically rich and close to real world concepts from e-commerce product descriptions which are publically available. A bootstrapping relation extraction algorithm is also proposed which looks for a specific set of seed relations that connects these concepts. The advantage of this proposed framework is that it is completely unsupervised and the need for a tagged corpus can be eliminated. The proposed method is also scalable that make this method efficient when dealing with large datasets. When compared with state-of-the-art methods in concept extraction, the proposed method outperforms them in terms of quality of concepts leveraged.

Organization of this paper: The rest of this paper is organized as follows. Section 2 defines the problem and briefly reviews the works that have been reported in the recent past which are closely related to our work and also a quick view of the LDA algorithm. Our proposed framework for extracting concepts and relations which is guided by topic modeling is explained in Section 3. Detailed explanation of our experimental setup and implementation of the same on e-commerce product descriptions are shown in Section 4. Section 5 showcases a detailed evaluation of the results and our conclusion and potential future works are given in Section 6.

II. Problem Definition

In this paper, the authors propose an approach for leveraging useful concepts and relation patterns from product descriptions available in e-commerce websites which are unstructured in nature. The proposed hybrid approach combines topic modeling, tf-itf weighting - a newly introduced topic word scoring scheme - and some basic linguistic processing such as POS tagging for mining potential concepts from reasonably high volumes of text. The method then extracts relationships which links these concepts and explores the possibility of building a product knowledgebase which has many potential applications such as meaning aware product search [1] and semantic question answering in e-commerce. In a nutshell this paper aims to answer the following questions:

1. Is it possible to leverage close to real-world and better human interpretable concepts from large collection of webpages using topic modeling?
2. Is it possible to extract semantic relations connecting such concepts from same webpages using a bootstrapping method and create a knowledgebase storing these patterns?
3. Is it possible to generate a semantic graph connecting such leveraged concepts and relations so that meaning aware applications can be built?

Towards the identification of semantically rich concepts and relation patterns, many approaches have been proposed. In the following subsection we describe the major ones that are close to our proposed algorithms.

III. State of the Art in Concept Extraction

Automatic concept extraction techniques has attained a high interest among knowledge management and engineering enthusiasts. Due to potential applications, a large number of research literatures have been reported in the field of concept extraction or concept mining which proposed many algorithms with varying degrees of success. In this section, we discuss past literatures on topic modeling guided concept extraction, automated concept mining and also a systematic review on recent past literatures reported in relation extraction.

Since topic modeling algorithms such as Latent Dirichlet Allocation (LDA) [5] and Probabilistic Latent Semantic Indexing (PLSI) have come up with well defined statistical and mathematical foundation, a good number of works that make use of such algorithms have been reported. The very first approach that thinks beyond traditional 'bag-of-word' method was the bigram topic model [3]. In this model, a new topic word is generated from a context given by a hidden topic associated with it and the previous word. Later a new statistical model called topical n-gram [6] was introduced that make use of a variable called switching variable for the identification of a new n-gram. The major drawback of this method was with the post-processing assumption it had and failed to attend the situation in practical scenario that the words within an n-gram usually won't share same topic.

Another approach called phrase discovering topic model [7] that uses a pitman-yor process for generating a topic-word matrix has been introduced. Performance of this method when dealing with large unstructured text data was not satisfactory and consumed significant amount of time to construct the matrix. A new model that performs phrase segmentation along with topic modeling was introduced [8] but failed to work with datasets having large number of files. TopMine [9], a system capable of mining topical phrase mining, is introduced recently, which implements a two step process for discovering phrases from text and for training a tradition topic model such as LDA. The assumption associated with this system, which says words in the same phrase must be assigned with the same topic, is not happening most of the time in practical scenario.

A graph based commonsense concept extraction and detection of semantic similarity [11] was introduced which uses a manually labeled dataset containing 200 multi-word concept pairs for evaluating their proposed parser. The method could leverage semantically and syntactically related concepts. The major shortfall of this algorithm is the supervised nature and requires human effort for tagging the dataset. A key phrase extraction technique called Automatic Concept Extractor (ACE) [12] was introduced by Ramirez and Mattmann which could extract concepts from HTML pages and their method used text body and visual clues such as bold, italic texts, etc., for identifying potential concepts. Their method could outperform some of the reported algorithms that were prevalent at that time and another system named GenEx [13] which employed a genetic algorithm supported rule learning mechanism for concept extraction. Since this relies on human crafted rules, the method could not perform well when dealing with diverse concepts.

A Naive Bayes learning model based key phrase extraction system called Automatic Keyphrase Extraction (KEA) [14] was developed which uses the model created with known key phrases extracted from training documents for inferring phrases from new set of documents. Another widely used method was introduced by Frantzi et al., which extracts multi-word terms from medical documents and is named as C/NC method [15]. The algorithm uses a POS tagger POS pattern filter for collecting noun phrases and then uses some statistical measures for determining the term-hood of candidate multi-words. This method could extract medical concepts from small collection of text documents but performance was degraded when dealing with large archives.

Parameswaran et al. introduced a system capable of extracting concepts from user tags and query log dataset which make use of technique similar to association rule mining [16]. The authors use features such as frequency of occurrences and the popularity among users for extracting concepts and they build a web of concepts. Another method, which uses a bag of word approach, was introduced by Gelfand et al., which can extract concepts from unstructured text by forming a closely tied semantic relations graph [17]. On applying this method specifically for some classification tasks, the authors found that their method produces better concepts than the Naive Bayes text classifier.

IV. State of the Art in Relation Extraction

A heavy number of real world applications in information retrieval and natural language processing require the proper identification of semantic relations connecting entities or concepts. There is an invaluable potential vested interest in the conversion of data from unstructured to structured form so that intelligent applications can use this for business purposes. Earlier methods of relation extraction were heavily dependent on supervised methods so that the creation of labeled data was time consuming and also too expensive to create in large quantities. Later bootstrapping based methods were introduced which start with a set of seed relation patterns and then learn more relations from unstructured text data. Here we discuss some of the semi-supervised relation mining algorithms which have similar method of working as our relation extraction algorithm.

Brin proposed a relation extraction system, DIPRE (Dual Iterative Pattern Relation Expansion) [18] for extracting author - book pairs from the web. This system starts with a small set of author - book pairs and then crawl the web for finding the occurrence of such pairs. If a new relation has been found, DIPRE adds it to the seed and continue crawling until there are no new seed relations found or a specified threshold has been met. Another system named Snowball [19] was introduced which works in the same direction of DIPRE but for extracting organization - location relation on plain text. The difference is that Snowball represents relation tuples as a vector and then uses a vector similarity function to group related tuples. For labeling a new data, Snowball first executes named entity recognition over the data to identify location and organization entities. The advantage of Snowball compared to DIPRE is that instead of searching for exact matches, Snowball searches for pairs having slight variations in token or punctuations. A large scale web based information extraction system called KnowItAll [20] was introduced that could label training examples using a small set of domain independent relation extraction patterns. Relation specific extraction rules built from generic patterns are used for learning domain specific extraction rules. These rules are then applied to web pages filtered through search engine queries and a probability value calculated using point-wise mutual information was assigned to the extracted patterns.

Algorithms discussed above such as DIPRE, Snowball and KnowItAll are relation specific systems where the user has to specify the relations of interest such as author - book or organization - location. To overcome this issue, another system named TextRunner [21] was introduced which learns the relations, classes and entities from its corpus in a self - supervised manner. It first labels training data as positive or negative and a classifier is trained using this data and the model is used by a pattern extractor. This extractor then generates candidate relations from sentences and chooses the positive relations tagged in the first step. A two stage bootstrapping algorithm for relation extraction [22] was introduced by Ang Sun. The first step of the algorithm is a commonly used bootstrapping method starting with a small set of seed relations and a large corpus to learn relation patterns, and a second stage bootstrapping which takes as input the relation patterns learned in the first stage and aims to learn relation nominals and their contexts. The author showed that this method could achieve a 2% gain in the f-measure.

This proposed approach incorporates both statistical methods such as topic modeling, tf-itf weighting and linguistic processes such as POS tagging for leveraging product concepts from product descriptions that are collected from e-commerce websites. We expect the learnt concepts are close to the real world understanding of products and quantify them using standard measures such as precision, recall and f-measure. For relation extraction, we propose a bootstrap based algorithm similar to one proposed by Ang Sun [22]. Starting with 11 manually chosen relation words which are commonly found in product descriptions as seed relations, we extract all concepts which are specified by these seeds. Then relation word and the associated concepts are extracted and pipelined for a knowledgebase construction process. While existing methods use a two stage process for tagging and creating a "bag-of-concepts" and then trains a topic model, our method uses a single stage lightweight process for inferring concepts from unstructured data that is guided by topic modeling.

V. Background: Latent Dirichlet Allocation (LDA)

A good number of topic modeling algorithms are introduced in the recent past which varies in their method of working mainly with the assumptions they adopt for the statistical processing. An automated document indexing method based on a latent class model for factor analysis of count data in the latent semantic space has been introduced by Thomas Hofman [23]. This generative data model called Probabilistic Latent Semantic Indexing (PLSI), considered as an alternative to the basic Latent Semantic Indexing has a strong statistical foundation. The basic assumption of PLSI is that each word in a document corresponds to only one topic. Later, a new topic modeling algorithm known as Latent Dirichlet Allocation (LDA) [5] was introduced which is more efficient and attractive than PLSI. This model assumes that a document contains multiple topics and such topics are leveraged using a Dirichlet Prior process. In the following section, we will briefly describe the underlying principle of LDA.

Even though a LDA works well on broad ranges of discrete datasets, the text is considered to be a typical example to which the model can be best applied. The process of generating a document with n words by LDA can be described as follows [5]:

1. Choose the number of words, *n*, according to Poisson Distribution;
2. Choose the distribution over topics, θ, for this document by Dirichlet Distribution;
   a) Choose a topic $T^{(i)} \sim$ Multinomial(θ)
   b) Choose a word $W^{(i)}$ from $P(W^{(i)} \mid T^{(i)}, \beta)$

Thus the marginal distribution of the document can be obtained from the above process as shown in (1):

$$\int_{\theta} \prod_{i=1}^{n} \sum_{T(i)} P(W(i) \mid T(i), \beta).P(T(i) \mid \theta)P(\theta \mid \alpha)d\theta \quad (1)$$

where $P(\theta \mid \alpha)$ is derived by Dirichlet Distribution parameterized by α, and $P(W(i) \mid T(i), \beta)$ is the probability of $W^{(i)}$ under topic $T^{(i)}$ parameterized by β. The parameter α can be viewed as a prior observation counting on the number of times each topic is sampled in a document, before we have actually seen any word from that document. The parameter β is a hyper-parameter determining the number of times words are sampled from a topic [5], before any word of the corpus is observed. At the end, the probability of the whole corpus *D* can be derived by taking the product of all documents' marginal probability as given in (2):

$$P(D) = \prod_{i=1}^{M} P(di) \quad (2)$$

where $P(di)$ is the probability of i^{th} document in the corpus.

VI. Topic Modeling Guided Semantic Knowledge Discovery

Even though the power of Latent Dirichlet Allocation algorithm has been used extensively for harnessing the topics from large text datasets, there are very few studies that have been reported for extending LDA for leveraging semantically rich concepts. Our proposed framework moves into this direction and we try to map statistically generated "topics" to semantically rich "concepts" and then extracts relationships and relation mentions that connect these concepts. Further we extend this for the creation of an e-commerce product knowledgebase which employs relational tables and contains pairs of feature and values as facts. Proposed framework can be classified into two sub-modules (i) concept extraction from publically available e-commerce product descriptions and (ii) relation extraction and knowledgebase construction. The overall work flow of the proposed method is depicted in Figure.1.

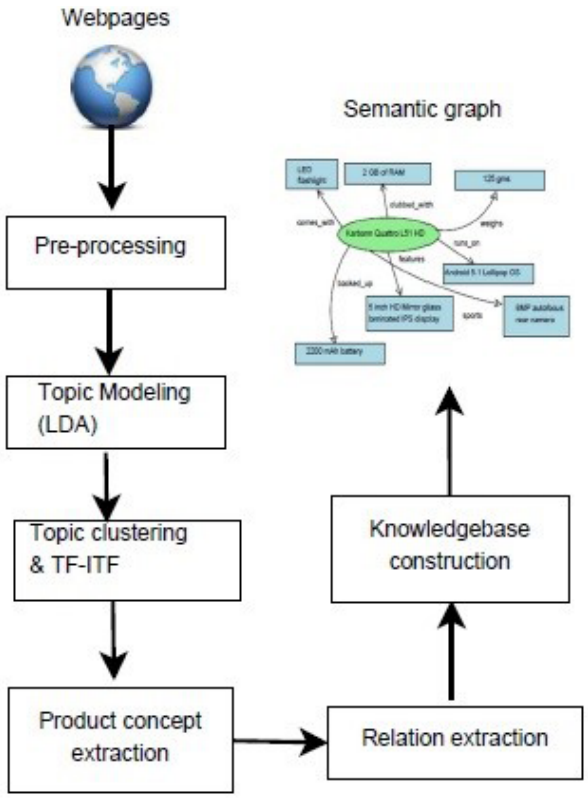


Fig. 1. Overall workflow of the proposed method.

A. Product Concepts Extraction

In this module, we introduce a topic to concept mapping procedure for leveraging potential concepts from statistically computed topics which are generated by the LDA algorithm. The first step of the proposed framework deals with the preprocessing of data which is meant for removing unwanted and irrelevant data and noises. Latent Dirichlet Allocation algorithm is executed on top of this preprocessed data which in turn generates topics through the statistical process. A total of 50 topics have been extracted by tuning the parameters of LDA algorithm. Once we have got the sufficient topics for the experiment, for each topic, we have created a topic - document cluster by grouping the documents which generated such a topic and the same process has been executed for all topics under consideration. Now, we introduce a new weighting scheme called *tf-itf* (term frequency - inverse topic frequency), which is used for finding out highly contributing topic word in each topic. We bring this weighting scheme to filter out the relevant candidate topic words. Term frequency *tf* is the total number of times that particular topic word comes in the topic - document clusters. Normalized term frequency, N_{tf} of a topic word T_w can be calculated as in (3):

$$Ntf = \frac{N\_Tw}{N} \tag{3}$$

where N_Tw is the number of times Tw occurs in the cluster and N is equal to the total number of terms in the cluster.

Inverse topic frequency I_{tf} is calculated as:

$$Itf = \frac{N\_Td}{N\_dTw} \tag{4}$$

where N_Td is the total number of documents in the cluster and N_dTw is the number of documents with term Tw in the cluster.

This step is followed by a sentence extraction process in which all the sentences which contain the topic words which have high *tf-itf* weight are extracted. Next, we apply parts of speech (POS) tagging on these sentences and extract only noun and adjective tags as we are only concentrating on the extraction of concepts. In linguistic pre-processing step, we take Noun + Noun, Noun + Adjective and (Adjective | Noun) + Noun combinations of words from the tagged collection. Concept identification is the last step in the process flow in which we find out the term count of all the combinations of Noun + Noun, Noun + Adjective and (Adjective | Noun) + Noun. A positive term count implies that the current multi word can be a potential "concept" and if we get a zero term count, then that multi word can be ignored. The newly proposed algorithm for extracting the concepts is shown in Algorithm 1.

Algorithm 1: Product Concept Extraction

1. For each topic t, create topic – document clusters
2. Compute *tf – itf* for topic words in all clusters
3. Choose top weighted topic words
4. Extract sentence from corpus containing top weighted topic words
5. Parts_of_Speech_Tag(sentence) and extract NN, NNP, NNS and JJ tags from the result
6. Take all combinations of Noun + Noun and Adjective + Noun and create a collection of terms
7. Calculate termcount for each of these terms
8. If termcount > 0, then add the term to concept repository
   Else remove the term from repository

B. Relation Extraction

Here, we propose a bootstrapping algorithm for extracting potential relationships and relation mentions from the product description dataset. The method starts with a specific set of relation words called seed relations that are manually collected and then searches the entire dataset and e-commerce websites for finding the occurrence of these seed relations. We have defined such seed relationships that are specific to e-commerce, especially for mobile phone descriptions. Once it finds a match while scanning the dataset and the web page, the algorithm extracts concepts that are mentioned using this seed relation. The proposed bootstrapping algorithm for relation extraction is shown in Algorithm 2. The seed relations along with example patterns we have used for this experiment is shown in Table 1.

Algorithm 2: Relation Extraction

1. Generate array of selected seed relations
2. Generate array of concepts extracted using Algorithm 1
3. index = 0
4. While True do
   Crawl through website URLs specified
   If seed_relations[index] found then
    Split instance with seed_relations[index]
    Add to collection of relation patterns
    Confidence_Score = count(seed_relations[index])
   End If
    index = index + 1
   End While

TABLE I. Shows Top Occurring Seed Relations And Sentences Containing Those Relations Which Are Used for Describing Mobile Phone Features in E-Commerce Websites.

Relation	Example patten
Features	The Karbonn Quattro L51 HD features 5-inch HD Mirror glass laminated IPS display
Clubbed_with	clubbed with 2GB of RAM
Runs_on	The smartphone runs on Android 5.1 Lollipop operating system
Backed_up	backed up by a 2200 mAh battery
Sports_an	The Quattro L51 HD sports an 8-megapixel auto-focus rear camera
Offers	the smartphone offers Dual SIM, 4G LTE (B3,B5,B40)
Measures	The handset measures 144.5 x 71.5 x 7.15 mm
Weighs	weighs around 125 grams
Comes in	The Quattro L51 HD comes in Black colour.
Comes with	whereas the front camera comes with LED flashlight
Powered with	It is powered with 1.3GHz dual-core processor
Powered by	It is powered by a 1.3 GHz quad-core processor

C. Knowledgebase Construction

Knowledge base construction is the process of creating an organized collection of facts extracted from unstructured data. User-centric applications can access these facts and assertions for providing users a knowledge driven experience. For example, as we show in this work, a knowledgebase storing facts and assertions about mobile phones may help in providing a meaning aware product discovery experience to a potential online buyer. Rather than using the currently available syntactically tagged features for a particular product, applications can use the knowledgebase automatically constructed for providing such services. Consider the natural language query a potential buyer can pass on to such an application, say, "Which mobile has a 13 MP camera and 2 GB of RAM ?". By conceptualizing this query and searching in a knowledgebase storing mobile phone features, we can show the customers all the mobile phones having the user specified features and thus provide a semantic search experience. Identifying the usability of such type of applications, we proceed further with organizing the concepts and relations identified by our proposed algorithms into a knowledgebase thus making the system complete and potentially useful.

This module of the proposed method organizes the extracted concepts, entities and associated relations in a relational database table format where each tuple represents the relation word (feature) and the concept (value). As shown in Figure 2, for each product, we create relational tables having 4 attributes - an identifier, the relation term extracted, the associated concept and a confidence score. The confidence score is the number of times a particular relation occurs in a given dataset along with a particular product concept, a high confidence score denotes that the particular relation can be considered as a valid relation. Our knowledgebase construction process is depicted in Figure 2.

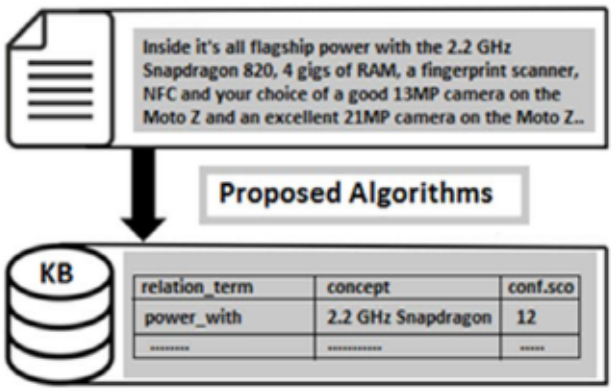


Fig. 2. Knowledgebase Construction Process.

VII. Experimental Setup

This section details the experimental setup used for implementing our proposed method. All methods described in this paper were implemented in Python 2.7. The experiments were run on a server configured with AMD Opteron 6376 @ 2.3GHz / 16 core processor and 16 GB of main memory.

A. Dataset Collection and Pre-processing

We are using mobile phone descriptions collected from publicly available e-commerce websites on the internet such as gsmarena.com, fonearena.com. The crawler we have created specifically for this task crawled 21855 webpages in total. Since such websites contain lot of hyperlinks and text styling information, a thorough pre-processing has been done for cleansing and extracting relevant and useful text descriptions. We have removed stop words, URLs and all other special characters and created an experiment ready copy of the original product descriptions.

B. Topic Modeling, Clustering and tf-itf Weighting

In this phase, Latent Dirichlet Allocation (LDA) algorithm has been executed on top the pre-processed product description dataset for generating topics for the experiment. We used a total number of 50 topics for this experiment and the numbers of iterations were set to 300 as Gibbs sampling [24] method normally approaches the target distribution after 300 iterations. Here, document clusters have been created in such a way that the documents which contributed to the creation of a particular topic are clustered together. This clustering has been done for all the 50 topics generated for the experiment and at the end, we have a total of such 50 document clusters. Now, a new weighting scheme called tf-itf (term frequency - inverse topic frequency) is introduced which ranks the topic words generated against each topic and is used for finding out the highly relevant topic words.

C. Sentence Extraction, POS Tagging and Concept Extraction

We consider topic words having highest tf-itf weights and extracts sentences which contain these topic words. This is done for filtering out unwanted or irrelevant sentences and a parts-of-speech tagging process has been applied on these extracted sentences. Since our aim is to find out potential product concepts, only words tagged as noun or adjectives are selected for further experiment. Natural Language Toolkit (NLTK) [25] is used in this experiment for POS tagging, which contains a good number of libraries for natural language processing with Python programming language.

Once we collect noun and adjective tags from the above step, all possible combinations of noun + noun, adjective + noun and noun/adjective + noun are taken. The term count of all these multi-words are calculated against the original product description dataset and getting a positive count implies that this specific word can be considered as a potential concept else otherwise. The same process has been repeated for all identified multi-words. Top 10 concepts leveraged are shown in Table 2.

TABLE II. TOP 10 CONCEPTS EXTRACTED USING ALGORITHM 1

Sl.No.	Concept	Sl. No.	Concept
1	5-inch HD Mirror glass laminated IPS display	6	Dual SIM, 4G LTE
2	2 GB RAM	7	125 grams
3	Android 5.1 Lollipop operating system	8	Black colour
4	2200 mAh battery	9	LED flashlight
5	8-megapixel auto-focus rear camera	10	1.3GHz dual-core processor

D. Relation Extraction, Knowledgebase Construction and Semantic Graph Creation

Once product concepts are extracted, our next aim is to identify the relation patterns that best associate these concepts so that we can proceed to the knowledgebase construction. A bootstrapping based relation pattern extraction algorithm shown in Algorithm 2, is used which starts with a specific set of relation patterns which are commonly found in e-commerce product descriptions, specifically in mobile phone descriptions. Such relation words we have identified for this experiment are shown in Table 1. The second column of the table shows the pattern and third column depicts the example sentences extracted from product descriptions which contain these relation patterns. Our proposed relation extraction algorithm leveraged 13743 such sentences where these patterns occur by crawling 21855 web pages we have collected where mobile phone descriptions are available.

Further we explore the possibility of creating a knowledgebase which contains the concepts and relations extracted from unstructured product descriptions collected from e-commerce and related websites. Relational database tables can be created for each product where each tuple can contain product concept and associated relation. For example, consider a sentence extracted from the product description of Karbonn Quattro L51 HD mobile phone - "The Karbonn Quattro L51 HD features 5-inch HD Mirror glass laminated IPS display" for which we have extracted "5-inch HD Mirror glass laminated IPS display" as the concept and "features" as relation keyword. We map this to a relational tuples as shown in Figure 3 and a confidence score has also been recorded against each tuple, which is the number of occurrences of specific concept and relation together that shows the relevance of the same.

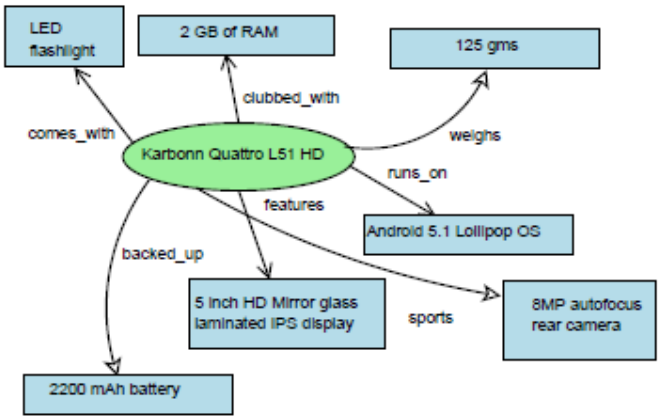


Fig. 3. Semantic graph representation of product concepts and relations.

VIII. EVALUATION OF RESULTS

For a rigorous evaluation, the proposed algorithm was compared against two topic modeling guided concept extraction algorithms and two other algorithms which are not using topic modeling. The proposed method was compared against the following two topic modeling guided baseline methods:

1. **Topical n-gram** [6] - Topical n-gram model automatically determines whether to form an n-gram or not by considering its surrounding text. Experimental results show that this algorithm generates topics which are more interpretable than traditional LDA model.
2. **TopMine** [9] - This algorithm first extracts phrases using a method similar to frequent pattern mining and then train a modified LDA model on the "bag-of phrases" input. This algorithm is found to be performing better than topical n-gram model (TNG) and a number of phrase discovering topic models such as phrase discovering topic model (PDLDA)

For further evaluation, we compared our method with the following two algorithms which are not guided by topic modeling:

1. **ACE** [12] - This algorithm was designed to improve search engines by automatically identifying concepts associated with search engine results. They emphasize a concept through the amount of times that appears on a webpage and using HTML tags. We use this method as baseline since our implementation uses data crawled from e-commerce webpages.
2. **KEA** [14] - KEA is an algorithm for automatically extracting key-phrases from plain text. This algorithm identifies candidate key-phrases using lexical methods, calculates feature values for each candidate, and uses a machine learning algorithm to predict which candidates are good key-phrases.[14].

A. Qualitative Evaluation

We have implemented the aforementioned algorithms using our dataset of mobile phone descriptions and the results show that this proposed method outperforms these two topic modeling guided algorithms as well as two general concept extraction algorithms in terms of extraction of relevant and semantically rich concepts. A human labeled concept collection was created from the dataset and then each of these algorithms was executed, and our proposed method shows better accuracy in identifying potential concepts. We make use of precision (P) and recall (R) measures for verifying the performance of our algorithms but F-measure (F) is calculated when it is analyzed that it is practically difficult to achieve high precision and recall at the same time. Here, true positive is defined as the number of overlapped concepts between human authored concepts and concepts generated by the algorithm, false positive is the number of extracted concepts that are not truly human authored concepts, and false negative is the human authored concepts that are missed by the concept extraction method. The result of precision, recall and f-measure comparison with general concept extraction algorithms such as ACE and KEA are shown in Table 3 and Figure 4. The result of precision, recall and f-measure comparison with topic modeling guided algorithms such as topical n-gram and TopMine are shown in Table 4 and Figure 5 respectively.

TABLE III. COMPARISON OF ACE, KEA AND OUR PROPOSED METHOD

Algorithm	Precision	Recall	F-measure
ACE	0.6913	0.7257	0.7080
KEA	0.7019	0.8139	0.7537
Proposed	**0.8533**	**0.8876**	**0.8701**

TABLE IV. Comparison of Topical N-gram, Topmine and Our Proposed Method

Algorithm	Precision	Recall	F-measure
Topical n-gram	0.7115	0.6251	0.6655
TopMine	0.8011	0.7562	0.7780
Proposed	**0.8955**	**0.8798**	**0.8875**

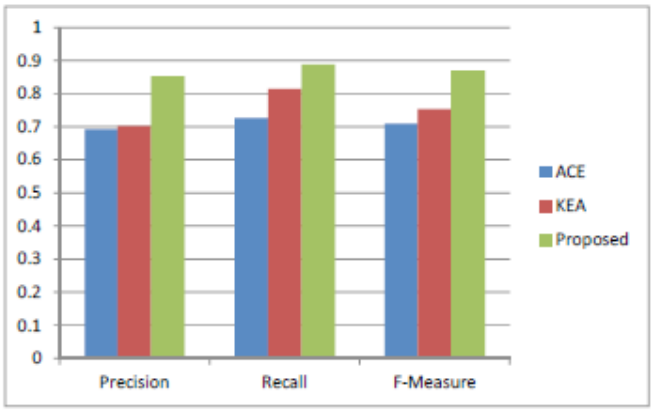


Fig. 4. Comparison of Precision, Recall and F-measure for ACE, KEA and our proposed method.

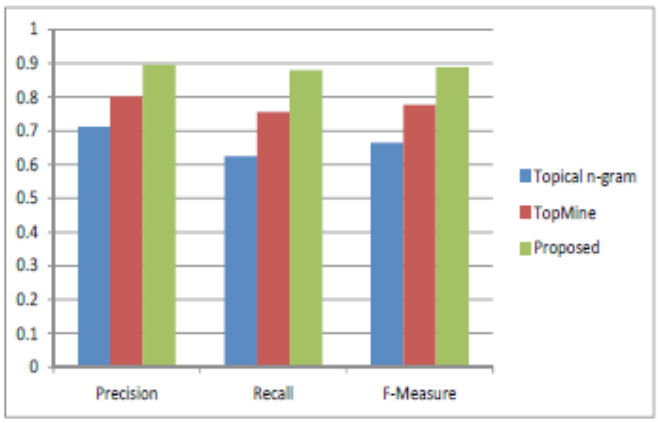


Fig. 5. Comparison of Precision, Recall and F-measure for Topical n-gram, TopMine and our proposed method.

B. Evaluation Based on Topical Coherence

Topical coherence measures the real-world understandability of topics or concepts generated by topic modeling algorithms. It is a well studied area in text mining and thus good number of literatures are available that discusses topical coherence measures. Very recently Roder et.al. proposed a new coherence measure [26] based on a combination of some already known approaches. They have shown that their proposed algorithm outperforms some state-of-the-art methods for measuring coherence including point-wise mutual information (PMI) [27]. They calculate co-occurrence of given words using Wikipedia and employ a normalized PMI value for calculating coherence. This paper uses this newly introduced coherence measure to compare the interpretability of our method with the topic modeling guided baseline methods. We have also employed some advanced statistical analysis [28] for validating the results with existing systems.

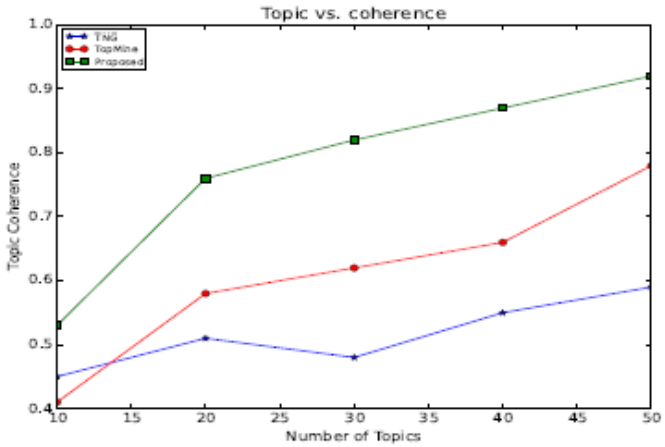


Fig. 6. Topical coherence comparison for Topical n-gram, TopMine and proposed method.

The topical coherence comparison graph for topical n-gram, TopMine and our proposed algorithm for 50 topical concepts is shown in Fig. 6. It is evident from the graph that there is an increase in coherence values when the number of topics is increased. From Figure 6, it is clear that, among the two baseline methods, topical n-gram shows least coherence and TopMine algorithm outperforms topical n-gram algorithm. Interestingly in all the cases, our proposed algorithm showcases superior topical coherence for the dataset under consideration and achieved better results.

IX. Conclusions

This paper introduced a novel algorithm for concept extraction from unstructured product descriptions available on e-commerce websites and other gadgets websites. A bootstrapping based algorithm is also proposed which is capable of finding out potential relationships starting with specific seed relations. We also explored the possibility of creating a product knowledgebase using these leveraged concepts and relationships so that more sophisticated semantic search and product discovery experience can be given to the user. Extensive and systematic experiments with large dataset comprised of product descriptions show that this topic modeling guided algorithm can better extract concepts from text dataset compared to other state of the art methods.

References

[1] Asharaf, S., Anoop, V. S. and Afzal, A. L. “A Framework for Meaning Aware Product Discovery in E-Commerce”. In I. Lee (Ed.), Encyclopedia of E-Commerce Development, Implementation, and Management (pp. 1386-1398). Hershey, PA: Business Science, 2016 Reference. doi:10.4018/978-1-4666-9787-4.ch098.

[2] Nasery, M., Braunhofer, M., and Ricci, F. “Recommendations with Optimal Combination of Feature-Based and Item-Based Preferences” InProceedings of the 2016 Conference on User Modeling Adaptation and Personalization (pp. 269-273). ACM, 2016.

[3] Wallach, H. M. “Topic modeling: beyond bag-of-words”. InProceedings of the 23rd international conference on Machine learning (pp. 977-984). ACM., 2006.

[4] Blei, D. M. “Probabilistic topic models”. Communications of the ACM, 55(4), 77-84, 2012.

[5] Blei, D. M., Ng, A. Y., and Jordan, M. I. “Latent dirichlet allocation”. Journal of machine Learning research, 3(Jan), 993-1022, 2003.

[6] Wang, X., McCallum, A., and Wei, X. “Topical n-grams: Phrase and topic discovery, with an application to information retrieval”. In Seventh IEEE International Conference on Data Mining (pp. 697-702). IEEE, 2007.

[7] Lindsey, R. V., Headden III, W. P., and Stipicevic, M. J. “A phrase-discovering topic model using hierarchical pitman-yor processes”. In

Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (pp. 214-222). Association for Computational Linguistics, 2012.

[8] Jameel, S., and Lam, W. "An unsupervised topic segmentation model incorporating word order". In Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval (pp. 203-212). ACM, 2013.

[9] El-Kishky, A., Song, Y., Wang, C., Voss, C. R., and Han, J. "Scalable topical phrase mining from text corpora". Proceedings of the VLDB Endowment, 8(3), 305-316, 2014.

[10] Griffiths, T. L., and Steyvers, M. "Finding scientific topics". Proceedings of the National academy of Sciences, 101(suppl 1), 5228-5235, 2004.

[11] Rajagopal, D., Cambria, E., Olsher, D., and Kwok, K. "A graph-based approach to commonsense concept extraction and semantic similarity detection". In Proceedings of the 22nd International Conference on World Wide Web (pp. 565-570). ACM, 2013.

[12] Ramirez, P. M., and Mattmann, C. A. "ACE: improving search engines via Automatic Concept Extraction". In Information Reuse and Integration, 2004. IRI 2004. Proceedings of the 2004 IEEE International Conference on (pp. 229-234). IEEE, 2004.

[13] Turney, P. D. "Learning algorithms for keyphrase extraction". Information retrieval, 2(4), 303-336, 2000.

[14] Witten, I. H., Paynter, G. W., Frank, E., Gutwin, C., and Nevill-Manning, C. G. "KEA: Practical automatic keyphrase extraction". In Proceedings of the fourth ACM conference on Digital libraries (pp. 254-255). ACM, 1999.

[15] Frantzi, K., Ananiadou, S., and Mima, H. "Automatic recognition of multi-word terms: the c-value/nc-value method". International Journal on Digital Libraries, 3(2), 115-130, 2000.

[16] Parameswaran, A., Garcia-Molina, H.,and Rajaraman, A. "Towards the web of concepts: Extracting concepts from large datasets". Proceedings of the VLDB Endowment, 3(1-2), 566-577, 2010.

[17] Gelfand, B., Wulfekuler, M., and Punch, W. F. Automated concept extraction from plain text. In AAAI 1998 Workshop on Text Categorization (pp. 13-17), 1998.

[18] Brin, S. "Extracting patterns and relations from the world wide web". In International Workshop on The World Wide Web and Databases (pp. 172-183). Springer Berlin Heidelberg, 1998

[19] Agichtein, E., and Gravano, L. "Snowball: Extracting relations from large plain-text collections". In Proceedings of the fifth ACM conference on Digital libraries (pp. 85-94). ACM, 2000.

[20] Etzioni, O., Banko, M., Soderland, S., and Weld, D. S. "Open information extraction from the web". Communications of the ACM, 51(12), 68-74, 2008.

[21] Yates, A., Cafarella, M., Banko, M., Etzioni, O., Broadhead, M., and Soderland, S. "Textrunner: open information extraction on the web". In Proceedings of Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations (pp. 25-26). Association for Computational Linguistics, 2007.

[22] Sun, A. A Two-stage Bootstrapping Algorithm for Relation Extraction. In RANLP (pp. 76-82), 2009.

[23] Hofmann, T. "Probabilistic latent semantic indexing". In Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval (pp. 50-57). ACM, 1999.

[24] George, E. I., and McCulloch, R. E. "Variable selection via Gibbs sampling". Journal of the American Statistical Association, 88(423), 881-889, 1993.

[25] Bird, S. "NLTK: the natural language toolkit". In Proceedings of the COLING/ACL on Interactive presentation sessions (pp. 69-72). Association for Computational Linguistics, 2006.

[26] Roder, M., Both, A., and Hinneburg, A "Exploring the space of topic coherence measures". In Proceedings of the eighth ACM international conference on Web search and data mining (pp. 399-408). ACM, 2005.

[27] Bouma, G. "Normalized (pointwise) mutual information in collocation extraction". Proceedings of GSCL, 31-40, 2009.

[28] Semwal, Vijay Bhaskar, et al. "An optimized feature selection technique based on incremental feature analysis for bio-metric gait data classification." Multimedia Tools and Applications: 1-19, 2016.

Anoop V.S.

Anoop V.S. is a Ph.D Research Scholar at Data Engineering Lab, Indian Institute of Information Technology and Management – Kerala (IIITM-K), Technopark, Thiruvananthapuram, India. He received his Masters in Computer Applications (MCA) from IGNOU, New Delhi and Master of Philosophy (M.Phil) in Computer Science from Cochin University of Science and Technology (CUSAT), Kerala in 2014. His research interests include Information Retrieval, Text Mining and Natural Language Processing. He has several research publications in international journals, conference proceedings and book chapters.

Asharaf S.

Asharaf S is an Associate Professor at Indian Institute of Information Technology and Management – Kerala (IIITM-K), Technopark, Thiruvananthapuram, India. He received his Ph.D and Master of Engineering degrees in Computer Science and Engineering from Indian Institute of Science, Bangalore. He is a recipient of IBM Outstanding-PhD student award. His areas of interest include algorithms, business models and software systems related to data mining, data analytics, information retrieval, computational advertising, soft computing and machine learning. He has published over 40 research papers across major venues in machine learning, and knowledge management.

Heuristics Considering UX and Quality Criteria for Heuristics

Frederik Bader[1], Eva-Maria Schön[2], Jörg Thomaschewski[3]*

[1] Verivox GmbH (Germany)
[2] University of Seville, Seville (Spain)
[3] University of Applied Sciences Emden/Leer (Germany)

Received 28 March 2017 | Accepted 22 April 2017 | Published 12 May 2017

Abstract

Heuristic evaluation is a cheap tool with which one can take qualitative measures of a product's usability. However, since the methodology was first presented, the User Experience (UX) has become more popular but the heuristics have remained the same. In this paper, we analyse the current state of heuristic evaluation in terms of heuristics for measuring the UX. To do so, we carried out a literature review. In addition, we had a look at different heuristics and mapped them with the UX dimensions of the User Experience Questionnaire (UEQ). Moreover, we proposed a quality model for heuristic evaluation and a list of quality criteria for heuristics.

Keywords

User Interfaces, Human Computer Interaction, Ergonomics, Heuristic Evaluation, User Experience, User Experience Questionnaire, Usability, Quality Criteria, Literature Review.

DOI: 10.9781/ijimai.2017.05.001

I. Introduction

The term User Experience (UX) wasn't officially defined until 2010 in DIN EN ISO 9241-210. Before that usability, as defined in DIN EN ISO 9241-10 in 1995, was regarded as the contemporary paradigm. Through the definition of the UX, usability has become expanded (e.g., [1, 2]).

There are various methods to ensure good usability, such as laboratory or field studies, which are often time-consuming and expensive. With heuristic evaluation [3], a type of expert-based evaluation has been established which requires no participants and is inexpensive as well as quick to carry out. Nevertheless, with this method, systematic results can be targeted. A set of rules, heuristics, form the basis of this method. Heuristics ensure that certain, desired characteristics of an object being investigated that lead to positive usability are examined.

In literature, there are some articles that deal with heuristic evaluation and recommend the use of heuristics [1-6]. In the analysis of the related work, it is striking that current heuristics are almost completely focussed on the evaluation of usability. Here, a research gap is shown, because heuristics have not yet been considered in depth for the evaluation of UX [2] and there are no studies on it that are worth naming.

This article is aimed at analysing the current state of research on heuristics in order to evaluate UX. In addition, we would like to determine what quality criteria for the evaluation of UX heuristics look like. For this reason, we will answer the following research questions:

* Corresponding author.
E-mail address: joerg.thomaschewski@hs-emden-leer.de

- RQ1: What heuristics are used to measure the User Experience?
- RQ2: Which quality criteria exist for heuristics?

A literature research was conducted to answer the research questions. In a further step, the heuristics that were found are analysed and mapped with dimensions of the UX. In addition, quality criteria were extracted from the literature research, consolidated and made concrete in a recommendation, through additional criteria.

The main contribution of this research is a mapping of heuristics with dimension of the UX (RQ1) and, as far as we know, the first quality model for heuristic evaluation with a suggestion of quality criteria for heuristics (RQ2).

The paper is structured as follows: section II provides an overview of related work. Section III describes the methodological process and the selection criteria. In section IV, the results of both research questions are presented and analysed before they are summarised in section V.

II. Related Work

In 1990, Nielsen and Molich [3] already described heuristic evaluation as a convenient way to evaluate the usability of a system. Since then, many things have changed and it is often not just usability that is evaluated, but rather the entire UX [2]. Since then, the early heuristics from Nielson or Shneiderman have hardly developed further, yet are still frequently cited and used (e.g., [4]-[6]).

New heuristics are often further developments of established heuristics and emerge, for example, by making an adjustment to another context being examined, such as mobile end devices [7], persuasive technologies [8] or patient safety for medical devices [9].

In doing so, the established heuristics are used for the validation of one's own results [7], [8]. These heuristics expand upon the area being evaluated. However, they do not take UX heuristics into consideration, because they do not cover the pragmatic and hedonic quality [1], which are components of the UX.

There are initial approaches to evaluate the UX with the help of heuristics [4], [5], [7], [10]-[13]. Two publications [10], [12] deal with special heuristics for the evaluation of the UX. Väänänen-Vainio-Mattila and Wäljas [10] predominantly relates it to a model in accordance with Hassenzahl [1], whereas Arhippainen [12] orientates herself on the definition of ISO 9241-210 [2].

There are various approaches for the creation of new heuristics:

- Other heuristics serve as the basis for the new heuristics [8], [9], [12], [13]
- Practical experiences are used in establishing them [8], [10], [12], [13]
- The establishment of heuristics on the basis of literature research [8], [10], [12], [14]
- The results from questionnaires by experts serve as the basis for new heuristics [8], [13]
- Heuristics are obtained from the validation of applications [8], [14]

Characteristics for the good application of heuristics or quality criteria are only named individually in literature on the topic [12], [14]-[16].

More frequently, the problem of generalisation is seen. General heuristics are easy to understand and implement [12], [16]; however, many specific topics could be overlooked [7].

III. Research Methodology

A literature review was conducted to find heuristics and quality criteria for heuristics used by the HCI community. The search string was defined as follows:

((Heuristic OR Heuristik) AND ("User Experience" OR UX)) OR "heuristic evaluation" OR ((Heuristic OR Heuristik) AND (quality OR "quality factor" OR "quality criteria"))

In October 2016, searches were launched in the following search spaces:

- scholar.google.com
- Elsevier ScienceDirect
- IEEE Xplore Digital Library (IEL)
- SpringerLink

The filtering process consisted of: (i) reading the title, (ii) reading the abstract, and (iii) reading the complete study. Studies were included if they met the following criteria:

1. The study reported how heuristic evaluation was used;
2. The study evaluated heuristics or heuristic evaluation;
3. The study proposed a method similar to heuristic evaluation;
4. The study proposed new heuristics;
5. The study was written in English.

The studies that did not meet the inclusion criteria or the full text of which was not available were excluded. After the selection of studies, backward snowballing [17] was used. For data analysis, we classified the paper according to objective, methodology (e.g., literature review, survey, heuristic evaluation) and relevant artefacts or results.

Our selection criteria for papers proposing new heuristics (RQ2) were: (i) heuristics needed to be presented in the paper, (ii) heuristics are formulated as rules, as mentioned in [3], (iii) heuristics differ distinctly (in terms of content) from existing heuristics and are general or easily adaptable to other contexts which are to be selected.

The articles found on RQ1 (What heuristics are used to measure the User Experience?) were examined to see which dimensions of the UX are evaluated with heuristics. In doing so, each heuristic was examined for its connections to the established UX dimensions of the UEQ [18]. These UX dimensions are:

- Attractiveness
- Perspicuity
- Efficiency
- Dependability
- Stimulation
- Novelty

The UX dimensions dependability, perspicuity and efficiency are associated with the pragmatic quality and described in accordance with the model from Hassenzahl: the usability [1]. In comparison with this, stimulation and novelty are amongst the hedonistic quality characteristics and, in doing so, describe the UX. In contrast, attractiveness is a dimension of valency.

Furthermore, it was determined whether heuristics take a further dimension of the UX into consideration in a targeted way, which is not a part of the UX dimensions of the UEQ [18]. In this case, it was noted in an additional column in the table, under the header "Other (not related to UX dimensions").

IV. Results

The major studies that remain after application of the selection criteria for the RQ1 (What heuristics are used to measure the user experience?) are listed in Table 1.

TABLE I.
Works Relevant for RQ1 (What Heuristics are Used to Measure the User Experience?)

Title	Author	Year
Designing the user interface: Strategies for effective human-computer interaction [19]	Shneiderman et al.	1987
Heuristic evaluation of user interfaces [3]	Nielsen and Molich	1990
Developing an Expert Evaluation Method for User eXperience of Cross-Platform Web Services [10]	Väänänen-Vainio-Mattila and Wäljas	2009
Heuristic Evaluation of Persuasive Health Technologies [8]	Kientz et al.	2011
Usability Heuristics Validation through Empirical Evidences: A Touchscreen-Based Mobile Devices Proposal [16]	Inostroza et al.	2012
Ten User Experience Heuristics [12]	Arhippainen	2013

The works relevant to RQ2 (What qualification criteria exist for heuristics?) are depicted in Table 2.

TABLE II. WORKS RELEVANT FOR RQ2 (WHICH QUALIFICATION CRITERIA EXIST FOR HEURISTICS?)

Title	Author	Year
Using Heuristics to Evaluate the Playability of Games [15]	Desurvire et al.	2004
Usability Heuristics Validation through Empirical Evidences: A Touchscreen-Based Mobile Devices Proposal [14]	Inostroza et al.	2012
Ten User Experience Heuristics [12]	Arhippainen	2013
Developing SMASH: A set of SMArtphone's uSability Heuristics [16]	Inostroza et al.	2016

A. RQ1: Heuristics and UX Dimensions

After the use of the selection criteria (see section III, on Research Methodology), six articles remain; they can be classified as follows:

- Established usability heuristics from Nielsen [3] and Shneiderman [19]
- Heuristics specially created for mobile end devices [7]
- Heuristics for the examination of persuasive technologies [8]
- Specific heuristics for the examination of the UX [10], [12]

Amongst the articles selected, only two publications can be found [10], [12] that developed heuristics of the measurement of the UX.

In the following sub-sections, the six heuristics will now be presented and their connection to the UX dimensions of the UEQ examined.

1. Shneiderman

In 1987, in his book "Designing the userface" [19], Shneiderman created eight golden rules of interface design.

TABLE III. SHNEIDERMAN HEURISTICS AND THEIR CONSIDERATION IN THE UX DIMENSIONS OF THE UEQ

Heuristic / UX dimension	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty
Strive for consistency				+		
Seek universal usability			+			
Offer informative feedback		+		+		
Design dialogs to yield closure		+	+	+		
Prevent errors		+	+	+		
Permit easy reversal of actions		+				
Keep users in control			+	+		
Reduce short-term memory load		+				

According to Shneiderman, this guide was created in order to gain interaction design and therefore, also receive improved usability. Therefore, the intention was not to create heuristics for the heuristic evaluation in accordance with Nielsen and Molich, but these rules could still be used for it. Just like Nielsen's heuristics, these eight rules are considered established. The results in Table 3 show that all eight rules lie in the area of classic usability.

2. Nielsen und Molich

In 1990, Nielsen and Molich developed the method of heuristic evaluation [3] and empirically examined it in four experiments. The heuristics were expanded by Nielsen in 1994 [20] and are considered established. The expanded heuristics are reflected upon in this work. In this work, the expanded heuristics [20] are considered. The heuristics are also used as a benchmark of new heuristics [7], [8]. In Table 4, the contrast between the heuristics (lines) to the UX dimensions (columns) is shown. It should be noted that the heuristic "Aesthetic and minimal design" goes beyond classic usability [2] and can be attributed to the "attractiveness" dimension.

TABLE IV. NIELSEN HEURISTICS AND THEIR CONSIDERATION OF THE UX DIMENSIONS OF THE UEQ

Heuristic / UX dimension	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty
Visibility of system status		+		+		
Match between system and the real world		+		+		
User control and freedom		+	+	+		
Consistency and standards		+				
Error prevention		+	+	+		
Recognition rather than recall				+		
Flexibility and efficiency of use			+	+		
Aesthetic and minimalist design	+					
Help users recognize, diagnose, and recover from errors		+		+		
Help and documentation		+		+		

3. Väänänen-Vainio-Mattila and Wäljas

With her heuristics, Väänänen-Vainio-Mattila and Wäljas aims [10] at the evaluation of the UX of web services. They are based on text research and informal examinations of existing web services. Different from Nielsen [3], each heuristic was sub-divided into hedonistic and pragmatic aspects, which took place in accordance with Hassenzahl's UX model [1]. After a case study, the heuristics were expanded.

Of the selected heuristics, this is one of two that is explicitly created with the goal of measuring the UX. Table 5 shows that four heuristics go beyond classic usability.

TABLE V. VÄÄNÄNEN-VAINIO-MATTILA AND WÄLJAS HEURISTICS AND THEIR CONSIDERATION OF THE UX DIMENSION OF THE UEQ

Heuristic / UX dimension	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty	Other (not related to UX dimensions)
Composite and linked services			+	+			
Cross-platform service access			+				
Social interaction and navigation					+		
Dynamic aspects of the service					+		
Contextual aspects of the service							+
Service usability		+	+	+			
Trust and safety				+			
Technical issues affecting UX							+
Service and content suitability			+				

4. Kientz et al.

Kientz et al. [8] present heuristics for the evaluation of persuasive health technology. In a literature review, a list of all heuristics is selected, with regard to persuasive health technologies. As a result, 13 experts were asked to list their respective 10 most important heuristics on this list. It could also be expanded through their own. Kientz et al. consolidated these 130 heuristics to a list of 10. These 10 heuristics

were used in two case studies and compared with Nielsen's heuristics [20]. The control group could discover more issues while carrying out the evaluation, yet they were less severe. Table 6 shows that three heuristics go beyond classic usability.

TABLE VI. KIENTZ ET AL. HEURISTICS AND THEIR CONSIDERATION OF THE UX DIMENSIONS OF THE UEQ

Heuristic / UX dimension	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty	Other (not related to UX dimensions)
Appropriate functionality		+		+			
Not irritating or embarrassing				+			
Protect users' privacy							+
Use of positive motivation strategies					+		
Usable and aesthetically appealing design	+						
Accuracy of information				+			
Appropriate time and place		+		+			
Visibility of user's status		+		+			
Customisability				+			
Educate users		+		+			

5. Inostroza et al.

In 2016, with "SMASH", Inostroza et al. [16] developed a set of heuristics that have to do with usability within the context of smartphones. The 12 heuristics are based on prioritised and formally-represented characteristics of smartphones from literature and case studies and, in its approaches, have a connection to UX.

The heuristics were validated in a case study and every heuristic finally refined using the criteria of utility, clarity, ease of use and the need of additional elements (checklists). In Table 7, the 12 heuristics and their connection to the UX dimensions is depicted. It shows that two of the twelve heuristics go above and beyond the dimensions of classic usability.

TABLE VII. INOSTROZA ET AL. HEURISTICS AND THEIR CONSIDERATION OF THE UX DIMENSIONS OF THE UEQ

Heuristic / UX dimension	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty
Visibility of system status		+		+		
Match between system and the real world		+		+		
User control and freedom		+	+	+		
Consistency and standards		+				
Error prevention		+	+	+		
Minimize the user's memory load		+		+		
Customization and shortcuts			+	+		
Efficiency of use and performance			+			
Aesthetic and minimalist design	+					
Help users recognize, diagnose, and recover from errors		+	+	+		
Help and documentation		+		+		
Physical interaction and ergonomics		+		+	+	

6. Arhippainen

In a tutorial, Arhippainen [12] presented 10 user experience heuristics. The aim of the heuristics is to support designers shaping aspects of the UX. They are based on empirical studies from mobile services. In addition to those from Väänänen-Vainio-Mattila and Wäljas [10], they are the only ones that focus on the analysis of UX. The results from Table 8 also show that only five of the ten heuristics for classic usability are included.

TABLE VIII.
ARHIPPAINEN HEURISTICS AND THEIR CONSIDERATION OF THE UX DIMENSIONS OF THE UEQ

Heuristic / UX dimension	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty	Other (not related to UX dimensions)
Ensure usability		+	+	+			
Provide utility matching with the user's values		+		+			
Surpass the user's expectations					+		
Respect the user		+		+			
Design the product or service to fit the intended contexts			+	+			
Provide several ways to interact, leave choice for the user		+					
Respect the user's privacy and security							+
Support the user's activities - do not force					+		
Go for a perfect visual design	+						
Give a surprise gift	+						

7. Discussion

In the overview (Table 9), it appears that the established heuristics, with one exception, exclusively cover the usability dimensions of the UEQ. The UX dimensions mostly remain ignored.

TABLE IX.
OVERVIEW OF HEURISTICS AND THEIR CONSIDERATION OF THE UX DIMENSIONS OF THE UEQ

Heuristic / UX dimension	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty	Other (not related to UX dimensions)
Shneiderman	0	5	4	5	0	0	0
Nielsen and Molich	1	7	3	8	0	0	0
Väänänen-Vainio-Mattila and Wäljas	0	1	4	3	2	0	2
Kientz et al.	1	4	0	7	1	0	1
Inostroza et al.	1	9	5	9	1	0	0
Arhippainen	2	4	2	4	0	0	1

Beyond all of the heuristics, it is shown that the dimension of dependability, with 35 considerations, is the heuristic that is observed

the most. The UX dimension of the UEQ, which shows the fewest connections to the heuristics, is novelty (see Table 10).

TABLE X.
NUMBER OF THE HEURISTICS THAT SHOW A CONNECTION TO THE UX DIMENSIONS OF THE UEQ

UX dimension	Number of heuristics
Attractiveness	5
Perspicuity	30
Efficiency	18
Dependability	36
Stimulation	6
Novelty	0
Other (not related to UX dimensions)	4

The heuristics from Inostroza et al. and Arhippainen consider 5 of 6 UX dimensions, Shneiderman considers a total of 3 UX dimensions and Nielsen and Molich, Väänänen-Vainio-Mattila and Wäljas as well as Kientz et al. 4 UX dimensions.

The number of individual heuristics varies between 8 [19] and 12 [16]. For better comparability, in Table 11, the considerations of the UX dimensions of the UEQ are weighted against the total number of connections. In the overall view of the heuristics, the number of dimensions is given that a heuristic takes into account, on average.

TABLE XI
OVERVIEW OF THE HEURISTICS AND THEIR WEIGHTED CONSIDERATION OF THE UX DIMENSIONS OF THE UEQ

Heuristic / UX dimension	Attractiveness (weighted)	Perspicuity (weighted)	Efficiency (weighted	Dependability (weighted)	Stimulation (weighted)	Novelty (weighted)	total (weighted)
Shneiderman	0	0.63	0.5	0.63	0	0	1.75
Nielsen und Molich	0.1	0.7	0.3	0.8	0	0	1.90
Väänänen-Vainio-Mattila and Wäljas	0	0.11	0.44	0.33	0.22	0	1.11
Kientz et al.	0.1	0.4	0	0.7	0.1	0	1.30
Inostroza et al.	0.08	0.75	0.42	0.75	0.08	0	2.08
Arhippainen	0.2	0.4	0.2	0.4	0.2	0	1.40

Accordingly, the heuristics from Inostroza et al. [16] show the highest coverage of the UX dimensions of the UEQ. The heuristics from Nielsen and Molich [20] as well as Shneiderman [19], which were a basis for Inostroza et al. [16], follow.

The heuristics examined predominantly move in the directions of the UX dimensions of dependability and perspicuity, which can also be organised with the dimension of efficiency of usability. The dimensions of stimulation and novelty are to be attributed to the UX and are hardly taken into consideration.

Consequently, the heuristics from Inostroza et al., Nielsen and Molich as well as Shneiderman view usability to be the strongest, and Arhippainen and Väänänen-Vainio-Mattila and Wäljas, the UX. The connection to the pragmatic dimensions is more strongly pronounced than in the hedonistic. Nielsen and Molich as well as Shneiderman, do not view any UX dimensions of the UEQ that are attributed to the hedonistic quality.

B. RQ2: Quality Criteria for Heuristics

Based on the selected studies, as far as we know, the first quality model for heuristic evaluation (Fig. 1) was created. This depicts the central elements of heuristics, the test object, evaluators, users and the relationships between them.

Because in the selected papers, only individual characteristics but no list of quality criteria could be found for heuristics, quality criteria were deviated from the literature selected and the quality model for heuristic evaluation (Fig. 1). They should ensure the efficient use of heuristics:

- Complying with the heuristics must help the user
- Validated through comparison with established methods / heuristics [15]
- Suitable for the context / object being examined [12], [14], [16], without negatively influencing the comprehensibility
- Comprehensible and easy / simple to implement (also for non-specialists) [16]

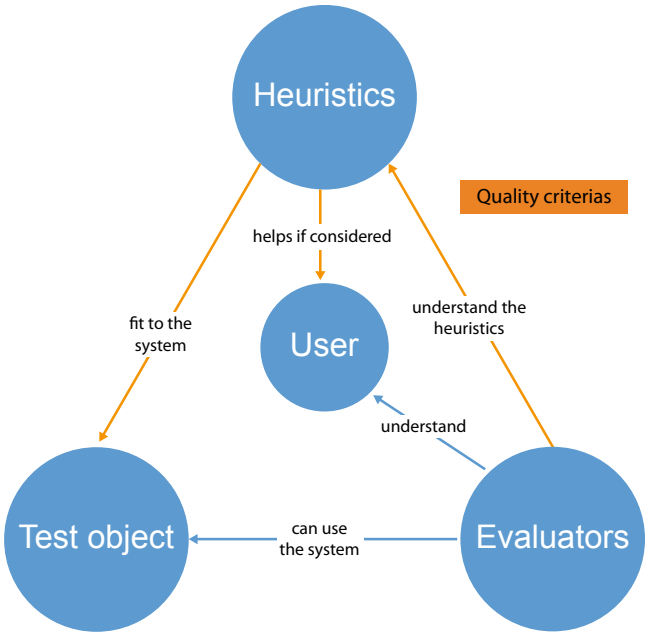


Fig. 1 Quality model of heuristic evaluation.

Furthermore, the additional following quality control criteria are recommended, which expand these quality criteria:

- Selective
- Not more than 10-15 heuristics

Additionally, quality criteria are recommended in order to ensure identical construction. In addition to this, heuristics should have the following elements:

- Designation (a semantic summary, not coded)
- Description
- Example

This recommendation for quality criteria should be ensured when considering the quality and the good applicability of heuristics in a heuristic evaluation.

V. CONCLUSION AND FUTURE WORK

Within the scope of this paper, we have examined the status of research regarding heuristic evaluation. In doing so, we have focused

on heuristics to measure UX. The heuristics identified were mapped with the dimensions of the UEQ for further analysis.

With heuristics, we examined as to whether not just the usability, but rather also the UX can be evaluated. To do so, every heuristic related to the UX dimensions was examined in accordance with the UEQ.

In this article we also introduced a quality model for heuristic evaluation and proposed a list of quality criteria for heuristics, which are based on research in literature.

With the heuristics identified, the pragmatic quality, which is attributed to usability, was examined most frequently. While research in the literature showed that many heuristics are particularly often based on the heuristics of Nielsen and Molich as well as those from Shneiderman which were created with a focus on usability.

The heuristics from Nielsen and Molich are still widely used for heuristic evaluation and as a foundation for other heuristics. They can be applied to most test objects. Other heuristics seem to be more focused on a particular domain. We recommend using the heuristics from Nielsen and Molich [3], as well as those from Inostroza et al. [16] when it is intended to use heuristic evaluation for UX as they have the highest weighted consideration of UX dimensions of the UEQ (Table 11).

The results of this study show that initial approaches exist to apply heuristics in a targeted way for the evaluation of the UX. However, this research field still requires further empirical research.

In the future, it would be important to examine how the results of a heuristic evaluation using a heuristic with a strong relationship to the dimensions of the UX appear in comparison to the established heuristics.

References

[1] M. Hassenzahl, "The Thing and I: Understanding the Relationship Between User and Product," vol. 3, pp. 31–42, 2003.

[2] Ergonomics of human-system interaction, ISO 9241-210:2010, 2010.

[3] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces" in *CHI '90 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 249-256, 1990.

[4] E. T. Hvannberg, E. L.-C. Law, and M. K. Lárusdóttir, "Heuristic evaluation: Comparing ways of finding and reporting usability problems" in *Interacting with Computers*, vol. 19, no. 2, pp. 225–240, 2007.

[5] M. González, L. Masip, A. Granollers, and M. Oliva, "Quantitative analysis in a heuristic evaluation experiment" in *Advances in Engineering Software*, vol. 40, no. 12, pp. 1271–1278, 2009.

[6] H. Idyawati, Murni Mahmud, and Abu Osman Md Tap, "User Experience: Assessing the Effectiveness of Internet Booking Service" in *International Conference on User Science and Engineering (i-USEr)*, pp. 1-6, 2010.

[7] R. Inostroza, C. Rusu, S. Roncagliolo, C. Jimenez, and V. Rusu, "Usability Heuristics for Touchscreen-based Mobile Devices," in *Ninth International Conference on Information Technology: New Generations (ITNG)*, pp. 662–667, 2012.

[8] J. A. Kientz, E. K. Choe, B. Birch, R. Maharaj, A. Fonville, C. Glasson and J. Mundt, "Heuristic Evaluation of Persuasive Health Technologies" in *Proceedings of the 1st ACM International Health Informatics Symposium*, pp. 555-564, 2010.

[9] J. Zhang, T. R. Johnson, V. L. Patel, D. L. Paige, and T. Kubose, "Using usability heuristics to evaluate patient safety of medical devices" in *Patient Safety*, vol. 36, no. 1–2, pp. 23–30, 2003.

[10] K. Väänänen-Vainio-Mattila and M. Wäljas, "Developing an Expert Evaluation Method for User eXperience of Cross-Platform Web Services" in *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era*, 2009.

[11] H. Rantavuo and A. Harder, "Heuristic Evaluation of User Experience – Case Nokia" in *SIGCHI Conference Paper*, pp. 1101–1106, 2014.

[12] L. Arhippainen, "Ten User Experience Heuristics," in *Introduction to Tribology*, B. Bhushan, Ed., Chichester, UK: John Wiley & Sons Ltd, pp. 1–8, 2013.

[13] M. P. González, T. Granollers, A. Pascual, and J. Lorés, "Testing Website Usability in Spanish-Speaking Academia through Heuristic Evaluation and Cognitive Walkthroughs" in *Journal of Universal Computer Science*, vol. 14, no. 9, pp. 1513-1528, 2008.

[14] R. Inostroza, C. Rusu, S. Roncagliolo, C. Jimenez, and V. Rusu, "Usability Heuristics Validation through Empirical Evidences: A Touchscreen-Based Mobile Devices Proposal," in *31st International Conference of the Chilean Computer Science Society (SCCC)*, pp. 60–68, 2012.

[15] H. Desurvire, M. Caplan, and J. Toth, "Using Heuristics to Evaluate the Playability of Games" in *CHI 2004 Conference on Human Factors in Computing Systems*, pp. 1509-1512, 2004.

[16] R. Inostroza, C. Rusu, S. Roncagliolo, V. Rusu, and C. A. Collazos, "Developing SMASH: A set of SMArtphone's uSability Heuristics" in *Computer Standards & Interfaces*, vol. 43, pp. 40–52, 2016.

[17] S. Jalali and C. Wohlin, "Systematic Literature Studies: Database Searches vs. Backward Snowballing" in *Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement*, pp. 29–38, 2012.

[18] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Construction of a benchmark for the User Experience Questionnaire (UEQ)" in *International Journal of Interactive Multimedia and Artifical Intelligence*, vol. 4, no. 4, pp. 40-44, 2016.

[19] B. Shneiderman, C. Plaisant, M. Cohen, S. M. Jacobs, and N. Elmqvist, "*Designing the user interface: Strategies for effective human-computer interaction*". Boston: Pearson, 1987.

[20] J. Nielsen, "Usability inspection methods" in *CHI '94 Conference Companion on Human Factors in Computing Systems*, pp. 413–414, 1994.

Frederik Bader

Frederik Bader is a MSc student in media technology with a primary focus in human-computer interaction at the University of Applied Science Emden/Leer. He works as a usability engineer at Verivox (Germany). His research interests are human-computer interaction and methods of usability and user experience evaluation.

Eva-Maria Schön

Eva-Maria Schön received her MSc in Computer Science and Media Applications at the University of Applied Science Emden/Leer in 2014. She is a PhD student at the University of Seville (Spain) and also works as a lead consultant at CGI (Germany). She focuses on agile coaching and human-centred design. Her research interests are agile software development, requirements engineering and human computer interaction.

Jörg Thomaschewski

Jörg Thomaschewski received his PhD in physics from the University of Bremen (Germany) in 1996. He became Full Professor at the University of Applied Sciences Emden/Leer (Germany) in September 2000. His research interests are Internet applications for human-computer interaction, e-learning, and software engineering. Dr Thomaschewski is the author of various online modules, such as "Human-Computer Communication," which are used by the Virtual University (online) at six university sites. He has broad experience in usability training, analysis and consulting.

Anomaly based Intrusion Detection using Modified Fuzzy Clustering

B. S. Harish and S. V. Aruna Kumar *

Department of Information Science and Engineering, Sri Jayachamarajendra College of Engineering, Mysuru, Karnataka (India)

Received 22 March 2015 | Accepted 3 May 2017 | Published 12 May 2017

Abstract

This paper presents a network anomaly detection method based on fuzzy clustering. Computer security has become an increasingly vital field in computer science in response to the proliferation of private sensitive information. As a result, Intrusion Detection System has become an indispensable component of computer security. The proposed method consists of three steps: Pre-Processing, Feature Selection and Clustering. In pre-processing step, the duplicate samples are eliminated from the sample set. Next, principal component analysis is adopted to select the most discriminative features. In clustering step, the network samples are clustered using Robust Spatial Kernel Fuzzy C-Means (RSKFCM) algorithm. RSKFCM is a variant of traditional Fuzzy C-Means which considers the neighbourhood membership information and uses kernel distance metric. To evaluate the proposed method, we conducted experiments on standard dataset and compared the results with state-of-the-art methods. We used cluster validity indices, accuracy and false positive rate as performance metrics. Experimental results inferred that, the proposed method achieves better results compared to other methods.

Keywords

Anomaly Detection, Intrusion Detection, Fuzzy Clustering, Principal Component Analysis, Robust Spatial Kernel Fuzzy C-Means.

DOI: 10.9781/ijimai.2017.05.002

I. Introduction

Computer security has become an increasingly vital field in computer science in response to the proliferation of private sensitive information. The term "Intrusion" refers to any unauthorized access which attempts to compromise confidentiality, integrity and availability of information resources [1] [14] [32]. Traditional intrusion prevention techniques such as firewalls, access control and encryption have failed to fully protect systems from sophisticated attacks. As a result, Intrusion Detection System has become an indispensable component of computer security which is used to detect the aforementioned threats. In 1987, Denning [7] first proposed an intrusion detection model. Since then many researchers have been focusing on developing efficient and accurate Intrusion Detection System (IDS) models. The intrusion detection techniques fall under two types, Misuse or Signature Based and Anomaly Based methods. Signature based methods detect only known intrusion attacks whose signatures are stored in the database. These methods fail to detect unknown intrusions. On the other hand, anomaly based methods detect the attacks based on the signature deviation.

In early days, intrusion detection is done using rule based approaches, where experts define a set of rules for normal and abnormal conditions. These systems work better for known attacks but fail to detect unknown attacks. In later 1990's researchers concentrated to develop automatic intrusion detection methods. Many researchers used data mining and machine learning algorithms to detect unknown attacks. Among various intrusion detection techniques, Fuzzy Logic based methods play a very important role. From literature review it is found that clustering methods are widely used approaches in intrusion detection system. Jianliang et al, [13] developed an intrusion detection system using K-means clustering algorithm. The experimentation was carried out on standard KDD-99 dataset. Cluster to class mapping, No class and Class Dominance are the key problems in K-means clustering. To overcome these drawbacks, Bharti et al., [4] developed two variants of traditional K-means algorithm. Ren et al., [23] developed a Fuzzy C-Means (FCM) algorithm to detect intrusions. The intrusion detection model was built through carrying out fuzzy partition and clustering of data. The experimental result shows that the algorithm can effectively separates normal and abnormal data. To overcome cluster centre initialization and convergence problem of FCM, Wang et al., [29] proposed a hybrid algorithm for intrusion detection system. This hybrid method combines FCM with Quantam behaved Particle Swarm Optimization. The Particle Swarm Optimization algorithm is used to overcome the drawback of FCM and to achieve global optimization and fast convergence. Guorui et al., [10] developed a semi supervised Fuzzy C-Means clustering algorithm for intrusion detection. This method overcomes the drawbacks of FCM i.e Sensitivity to the initial values and converging to the local minima by using few labelled data to improve the learning ability of the Fuzzy C-Means. Sampat and Sonawani [25] developed an intrusion detection system using Improved Dynamic Fuzzy C-Means (IDFCM) clustering. The IDFCM is a variant of the traditional FCM which adaptively updates the cluster centres. Experimental result shows the IDFCM gives better detection accuracy rate than traditional FCM. Hameed et al., [11] developed an hybrid clustering algorithm for intrusion detection. This hybrid algorithm combines Modified Fuzzy Possiblistic C-Means (MFPCM) and symbolic fuzzy clustering. This method uses 30 features with optimal sensitivity and highest discriminatory power.

Ganapathy et al., [9] proposed a intrusion detection system based on Weighted FCM and Immune Genetic Algorithm (GA). The Weighted FCM is a modification of FCM which builds a system for more accurate

* Corresponding author.
E-mail addresses: bsharish@sjce.ac.in (B. S. Harish), arunkumarsv55@gmail.com (S. V. Aruna Kumar).

attack detection. Immune GA improves the performance, probability of gaining the global optimal values and solves the high dimensionality problem. Khazaee and Rad [17] developed a novel method based on Fuzzy C-Means for improving the intrusion detection performance. The experiments were conducted on KDD Cup-99 dataset. Kumar et al., [19] developed a novel method of intrusion detection which involves fuzzy feature clustering. In this method fuzzy features clustering method is used to reduce the dimensionality of system calls. Pandeeswari and Kumar [21] developed hybrid detection system for cloud environment. This hybrid model works in three steps. In the first step, to improve the learning capability of ANN, fuzzy clustering is used. In the second step, various ANN modules are trained according to their cluster values. In the last step, fuzzy aggregation module is used to combine the results of various ANN.

Karthik and Nagappan [16] developed an intrusion detection system using Kernel Fuzzy C-Means and Bayesian Neural Network. This hybrid model consists of two step. In the first step, Fuzzy Bisector Kernel Fuzzy C-means is used to obtain the cluster centers. In the second step the centroids obtained from previous step are used for the learning of Bayesian network. The experiments were conducted on standard KDD Cup 99 Dataset. Rustam and Talita [24] proposed an intrusion detection algorithm based on Fuzzy Kernel C-Means (KFCM). Khazaee and Faez [18] developed an hybrid classification method for network intrusion detection. This hybrid model combines fuzzy clustering with multilayer perceptron neural network. Here training samples are initially clustered using fuzzy clustering and the inappropriate data will be detected and moved to another dataset. Further, Multilayer Perceptron will be trained using new labels. To classify outlier fuzzy ARTMAP neural network is employed. Surana., [27] developed an hybrid algorithm which combines Fuzzy C-Means and Neural Network for intrusion detection. This approach divides the training data into smaller groups using Fuzzy C-Means. Later, Neural Networks are trained using these subsets. Finally individual neural network results are aggregated. Kumar and Harish [2] proposed Robust Spatial Fuzzy C-Means algorithm which considers spatial information and uses kernel distance metric. Xie et al., [31] developed an intrusion detection using hybrid clustering method. This hybrid method is the combination of Fuzzy C-Means Clustering, Average Information Entropy, Support Vector Machine and Fuzzy Genetic algorithm.

It is evident from the above discussion that, the researchers have been attempting to come out with more efficient and robust intrusion detection techniques. Further, most of the existing methods are based on supervised and unsupervised learning approaches. Supervised learning requires a large volume of training samples and they fail to detect unknown attacks. On the other hand, unsupervised learning has got the advantage of detecting unknown attacks. On the contrary, the main limitations of the unsupervised learning are as follows: higher False Alarm Rate (FAR), fails to identify the specific type of attacks and worstly affected by curse of dimensionality. Further, most of the existing unsupervised methods uses euclidean as a distance metric. Unfortunately, euclidean metric is very sensitive to noise which results in degrading the system accuracy.

With this backdrop, to eliminate the above said problems in this paper, we present network anomaly detection method based on fuzzy clustering. The proposed method consists of three steps: Pre-Processing, Feature Selection and Clustering. To evaluate the proposed method, we conducted experiments on standard intrusion detection dataset and compared the results with other variants of FCM methods.

In summary, the main contributions of this paper are as follows:

- A Modified Fuzzy Clustering method (RSKFCM) for anomaly detection is presented. The method can also identify a specific type of attacks.
- To handle curse of dimensionality, we employed Principal Component Analysis (PCA) as feature selection method.
- To overcome the existing drawbacks of FCM methods, we considered neighborhood information which in turn reduces the false alarm rate and increases the system accuracy.
- We used Gaussian kernel as distance measure to compute the distance between cluster center and samples. The advantage of using Gaussian kernel is that it reduces the effect of noise.
- We validated the proposed method on standard intrusion dataset using four cluster validity functions, accuracy, and false alarm rate. Further, we also compared our results with the contemporary methods.

The rest of the paper is organized as follows: Section 2 presents details about the KDD Cup 99 Dataset. Traditional Fuzzy C-Means is presented in Section 3. Section 4 presents the proposed method. Experimental setup, dataset used for experimentation and results are presented in section 5. Conclusions are drawn in section 6.

II. KDD Cup 99 Dataset

Since 1999, Many researchers evaluated the intrusion detection models on the KDD Cup 99 dataset [26]. This dataset was originally created by 1998 DARPA intrusion detection evolution program. The dataset contains five and two millions of training and testing samples respectively. Each sample as a set of 41 features derived from each connection and a label which specifies connection as normal or attack. These 41 features can be categorized into 4 groups.

- Basic features: These features can be derived from packets header without inspecting the payload.
- Content features: These features contains the information present in the payload of the original TCP packets and extracted using domain knowledge.
- Time based Traffic features: These features contains the service information which inspect the connection in the past two seconds that have the same service as the current connections.
- Host based Traffic features: These features examine the connections which established in the past two seconds and which have the same destination host as the current connections.

The dataset contains 21 different types of attacks which can be categorized into 4 types.

- Denial of Service (DoS): Attacker tries to forbid the legitimate users from utilizing the requested services/resources
- Probe (PRB): Attackers attempt to gain information about the target host.
- Remote to Local (R2L): Attackers do not have the account in victim computer, so they try to get access to the computer.
- User to Root (U2R): Unauthorized user tries to gain access to local super user(root) through the network.

III. Traditional Fuzzy C-Means

Fuzzy C-Means clustering algorithm is based on the traditional fuzzy set which was proposed by Bezdek [3]. FCM is the improvement of K-means algorithm which groups the data points based on the membership value. In FCM, the membership of a data point depends on the similarity of the data point to a particular class relative to all other class. Let $X = \{x_1, ...x_j, ...x_n\}$ be the N data points and $V = \{v_1, ...v_i, ...v_j\}$ be the set of c centroids. FCM partitions X into c clusters by minimizing the objective function in (1).

$$J = \sum_{j=1}^{N}\sum_{i=1}^{c} u_{ij}^{m} \left\| x_j - v_i \right\|^2 \quad (1)$$

where u_{ij} is the degree of membership of x_j in the cluster i, v_i is the i^{th} cluster center, $\|.\|$ is a distance metric and m is a constant which controls the fuzziness of the resulting partition. Using the Lagrangian method Bezdek derived two necessary condition for minimizing the objective function J as follows:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\left\|\frac{x_j - v_i}{x_j - v_k}\right\|\right)^{\frac{1}{m-1}}} \quad (2)$$

$$v_i = \frac{\sum_{j=1}^{N} u_{ij}^{m} x_j}{\sum_{j=1}^{N} u_{ij}^{m}} \quad (3)$$

The Clustering process begins by randomly choosing the c cluster centers. Further, the membership are calculated based on the relative distance of data point x_j to the centroid v_i using equation (2). The data points which are close to centroids are assigned highest membership value, where as data points far from the centroids are assigned low membership value. After computing the membership of all the data points, the cluster centers are updated using equation (3). The clustering process stops when the successive difference of objective function is less than the pre defined threshold value (ε).

IV. Proposed Method

The technique of fuzzy clustering has become very important in the application of intrusion detection. This is due to the large role of uncertainty nature of an attack. Motivated by this, in this paper we proposed fuzzy clustering based anomaly intrusion detection method. The proposed method consists of three steps: Pre-processing, Feature Selection and Clustering. Fig. 1 shows the block diagram of the proposed system.

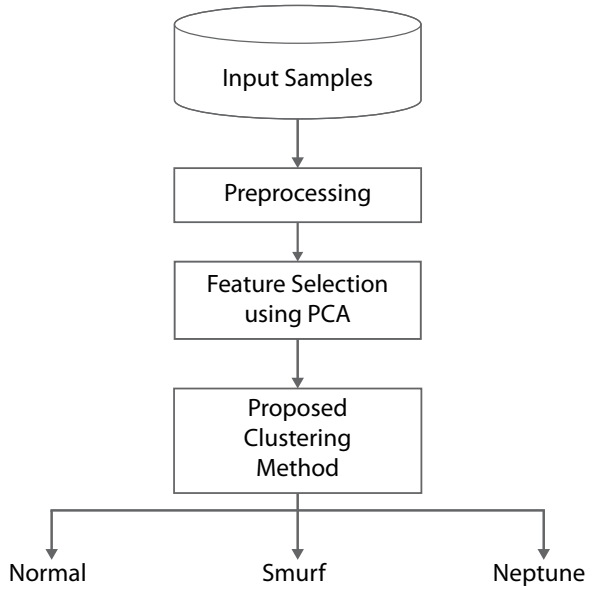


Fig. 1. Block diagram of the Proposed System.

A. Pre-processing

The pre-processing step is performed to make the dataset convenient to be use in the clustering step. The pre-processing will have great impact on the intrusion detection efficiency. In this step we performed two pre-processing steps that are: removing duplicate samples and filling missing values. Dataset contains a large number of duplicate samples. If the dataset contains duplicate samples, then clustering takes more time and also gives inefficient results. To achieve accurate results we removed duplicate samples from the dataset. To fill the missing values, first we divide the whole dataset according to their class and compute the mean value for each feature. Further the missing value is replaced with corresponding feature's mean value.

B. Feature Selection

In this paper, we employed Principal Component Analysis (PCA) [12][15] to select the most discriminative features. PCA is a widely used feature selection method and it is based on linear transformation, which maps data from a high dimensional feature space to lower dimension. The first PCA as the highest contribution to the variance in the original dataset and each succeeding components have remaining variance as possible. The primary advantages of PCA are it is simple, non paramteric and it can preserve large percentage of the total variance with only a few components. Thus, these characteristics motivated us to adopt PCA for feature selection.

Let us consider a training network sample $X = \{x_1, x_2, ...x_m\}$ and it has n user behaviour features $F = \{f_1, .., ,, , f_n\}$. So, totally we have $m \times n$ feature matrix. First we compute the mean of the observations X as follows:

$$\mu = \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \vdots \\ \vdots \\ \overline{x}_m \end{pmatrix} \quad (4)$$

Next we compute the covariance matrix i.e

$$S = \frac{1}{n}\sum_{i=1}^{n} \Phi_i \Phi_i^{T} \quad (5)$$

where Φ_i is the standard deviation which is computed as indicated in (6):

$$\Phi_i = X - \mu \quad (6)$$

Φ_i^{T} is the transpose of the standard deviation.

Further, eigen value and corresponding eigen vector of the covariance matrix S is computed. Let $\{(\lambda_1, \xi_1), \ldots (\lambda_k, \xi_k) \ldots (\lambda_m, \xi_m)\}$ are m eigen values and corresponding eigen vectors pairs of the covariance matrix. Next, we choose k eigen vectors corresponding to the largest eigen values. Afterwards, we form $m \times k$ matrix D, whose column consist the k eigen vectors. The representation of data by principal component consist of projecting the original data on to the k dimensional subspace h_k such that:

$$Y_i = D^T \Phi_i = D^T (X - \mu) \quad (7)$$

In the next step we employed the Robust Spatial Kernel FCM (RSKFCM) to cluster these reduced feature matrix.

C. Robust Spatial Fuzzy C-Means

Traditional Fuzzy C-Means (FCM) leads to its non robust result mainly due to: not utilizing the neighbourhood information and use of Euclidean distance. To overcome these problems, Kumar and Harish

[2] proposed a Robust Spatial Kernel FCM (RSKFCM) technique. RSKFCM incorporates spatial information to the conventional FCM membership function and uses kernel distance metric. Experimental results reveal that, RSKFCM gives better clustering results than other FCM variants. Thus, inspired by the good performance presented in [2], we applied RSKFCM method to detect network anomalies.

The main aim of the RSKFCM is to minimize the following objective function J :

$$J = \sum_{i=1}^{c}\sum_{j=1}^{N} w_{ij}^{m} \left\| \Phi(x_j) - \Phi(v_i) \right\|^2 \tag{8}$$

Where Φ is an implicit nonlinear map, and

$$\left\| \Phi(x_j) - \Phi(v_i) \right\|^2 = K(x_j, x_j) + K(v_i, v_i) - 2K(x_j, v_i) \tag{9}$$

Where $K(x,y) = \Phi(x)^T \Phi(y)$ is an inner product kernel function. If we adopt the Gaussian RBF kernel i.e $K(x,y) = \exp\left(-\| x - y \|^2 / \sigma^2 \right)$, then $K(x,x) = 1$. The simplified objective function becomes :

$$J = 2\sum_{i=1}^{c}\sum_{j=1}^{N} w_{ij}^{m} \left(1 - K\left(x_j, v_i\right)\right) \tag{10}$$

Where w_{ij} is the new membership function which combines traditional FCM membership function and neighbourhood membership function. The new membership value is computed as:

$$w_{ij} = \frac{u_{ij}^{p} s_{ij}^{q}}{\sum_{k=1}^{c} u_{kj}^{p} s_{kj}^{q}} \tag{11}$$

Where u_{ij} is the Kernel FCM (KFCM) membership function and s_{ij} is the neighbourhood membership function. Kernel FCM is variant of the FCM, unlike in FCM it uses kernel function as distance metric. The membership function of KFCM is calculated as:

$$u_{ij} = \frac{\left(1 - K\left(x_j, v_i\right)\right)^{-1/(m-1)}}{\sum_{k=1}^{c}\left(1 - K\left(x_j, v_k\right)\right)^{-1/(m-1)}} \tag{12}$$

To compute the neighbourhood membership function, we calculated distance from each samples to other samples and considered k nearest samples. The neighbourhood memebrship function is defined as follows:

$$s_{ij} = \sum_{k \in NK(x_j)} u_{ik} \tag{13}$$

where $NK(x_j)$ represents a array of k nearest samples from x_j. This spatial function represents the probability of sample x_j belongs to i^{th} cluster. In new membership function p and q parameters controls the relative importance of both functions.

The RSKFCM algorithm is carried out through an iterative optimization of the objective function shown in equation (10) with the update of membership value and cluster centers. The cluster centers are updated using equation (14):

$$v_i = \frac{\sum_{j=1}^{N} w_{ij}^{m} K\left(x_j, v_i\right) x_j}{\sum_{j=1}^{N} w_{ij}^{m} K\left(x_j, v_i\right)} \tag{14}$$

The clustering process stops when the successive difference of the objective function is less than the pre defined threshold value (ε). The individual stages of Robust Spatial Kernel Fuzzy C-Means (RSKFCM) are described in Algorithm 1.

Algorithm 1: RSKFCM Clustering Method

Input: Intrusion Samples

Output: Label

Initialize cluster centers, ε , m

Repeat

{

Step 1 Compute all membership values u_{ij} of each sample against centers as:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\left\| \frac{x_j - v_i}{x_j - v_k} \right\|\right)^{\frac{1}{m-1}}} \tag{15}$$

Step 2 Compute the new membership value w_{ij} using equation (11)

Step 3 Calculate the objective function J as follows:

$$J = 2\sum_{i=1}^{c}\sum_{j=1}^{N} w_{ij}^{m} \left(1 - K\left(x_j, v_i\right)\right) \tag{16}$$

Step 4 Calculate new cluster center values v_i according to expression in (17).

$$v_i = \frac{\sum_{j=1}^{N} w_{ij}^{m} K\left(x_j, v_i\right) x_j}{\sum_{j=1}^{N} w_{ij}^{m} K\left(x_j, v_i\right)} \tag{17}$$

}

Until $\left\{ J(i) - J(i-1) \right\} < \varepsilon$

V. Experimental Results

To evaluate the proposed method, we conducted experiments on EDA dataset [5]. This dataset is the variant of original KDD dataset. Since, in original KDD dataset smurf, neptune and normal traffic represents 99.3% of the total samples, EDA dataset contains only these three classes.

To handle categorical features, this dataset adds one dummy variable per category into the original KDD dataset, which in turn increases the feature size into 122.

The performance of the proposed method is evaluated using four cluster validity indices, accuracy and false positive rate. For all algorithms in comparison, we set the fuzzy co-efficient m to widely used value 2. All the cluster centers are initialized randomly. We set stopping criteria ε = 0.0001. We implemented and simulated all the algorithms with matlab2013.

A. Evaluation using Cluster Validity Indices

In this section we evalated performance of the propsoed method using four cluster validity indices. These cluster validity indices help

to validate whether clustering method accurately presents the structure of the data set or not. Wide varieties of cluster validity indices are proposed in the literature. In this paper we have used four widely used cluster validity functions, namely Partition Coefficient (V_{pc}), Partition Entropy (V_{pe}), Fukuyama-Sugeno function (V_{fs}), and Xie-Beni function (V_{xb}).

Bezdek [20][22] proposed Partition Coefficient (V_{pc}) and Partition Entropy (V_{pe}) which uses only the membership values to evaluate the -cluster validity as indicated in (18) and (19):

$$V_{pc}(U) = \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{c} u_{ij}^{m} \tag{18}$$

$$V_{pe}(U) = \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{c} u_{ij}^{m} \log u_{ij} \tag{19}$$

The value of V_{pc} varies between $[\frac{1}{c},1]$ where c indicates the number of clusters. The value of V_{pe} ranges between $[0,\log_a c]$ where c is the number of cluster and a is the base of the logarithm. When V_{pc} is maximal or V_{pe} is minimal, the optimal clusters are achieved.

The Fukuyama-Sugeno function (V_{fs}) [8] is given by:

$$V_{fs}(U,V;X) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m}\left(\left\|x_j - v_i\right\|^2 - \left\|v_i - \bar{v}\right\|^2\right) \tag{20}$$

where $\bar{v} = \frac{1}{c}\sum_{i=1}^{c} v_i$. V_{fs} uses both the membership information and input data. When V_{fs} value is minimum, the better clustering results are achieved.

The Xie-Beni function (V_{xb}) function, which was initially proposed by Xie-Beni (XB) in [30] and modified by Pal and Bezdek in [18], is defined as indicated in equation (21):

$$V_{xb}(U) = \frac{\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m}\left\|x_j - v_i\right\|^2}{n\left(\min_{i\neq k}\left\{\left\|v_i - v_k\right\|^2\right\}\right)} \tag{21}$$

In V_{xb} the numerator indicates the compactness of the fuzzy partition and denominator indicates the strength of the separation between clusters. When V_{xb} is minimal, the best clustering result is achieved.

To evaluate the performance, we compared our proposed algorithm with traditional Fuzzy C-Means (FCM), Kernel FCM (KFCM) and Spatial FCM (SFCM) methods. Table I presents the performance comparison of the proposed method.

B. Performance Comparison with State-of-the-art Methods

We compared our proposed method with six unsupervised anomaly detection methods. The methods used in comparison are as follows: K-Means [28], Improved K-Means [28], K-Medoids [28], Expectation Maximization [28], Fuzzy C-Means [6], Fuzzy Rough Clustering [6]. We used accuracy and False positive rate as evaluation metrics. Fig. 2 shows the comparison of the proposed method with other methods using accuracy. Fig. 3 shows the comparison of the proposed method with other methods using False Positive Rate.

TABLE I
PERFORMANCE COMPARISON

Method	V_{pc}	V_{pe}	V_{xb} $[1\cdot10^{-3}]$	V_{fs} $[-1\cdot10^{6}]$
FCM	0.826	0.201	59.349	2.163
KFCM	0.847	0.196	53.821	2.263
SFCM	0.891	0.121	51.153	2.281
RSKFCM	**0.915**	**0.108**	**32.142**	**2.346**

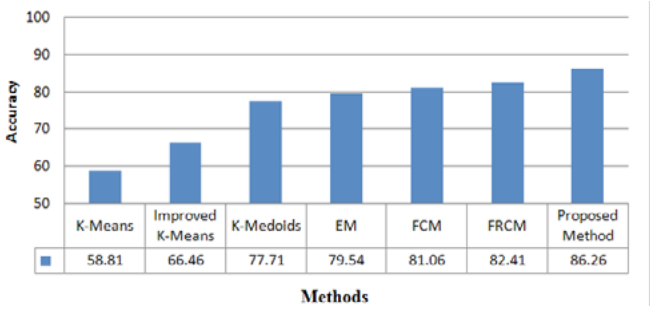


Fig. 2. Comparison of Accuracy.

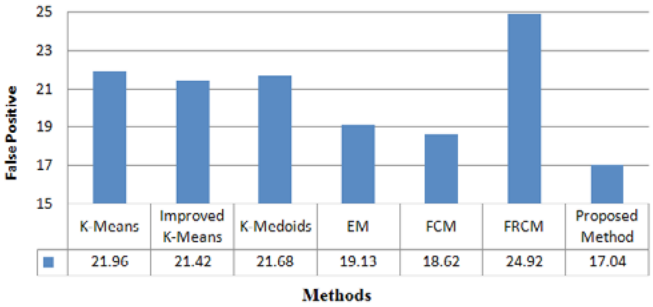


Fig. 3. Comparison of False Positive Rate.

C. Discussion

In Table 1, Fig. 2 and Fig. 3 we can observe that our proposed method outperforms all other methods. As mentioned earlier, our proposed method uses neighborhood information and kernel distance metric. The other methods in comparison, cluster a given sample based on membership value or distance value. Whereas our proposed method considers neighborhood membership value along with the membership value of that sample. This neighborhood information increases the accuracy of the proposed method. On the other hand, other methods use Euclidean distance which is very sensitive to noise. Whereas our proposed method uses Gaussian kernel distance metric. This reduces the noise effect and in turn increases the accuracy.

VI. CONCLUSION

The existing computer security technologies fail to prevent the threats completely. As a result, Intrusion Detection System (IDS) becomes important component in network security. IDS reduces the manpower needed for monitoring and increases the detection efficiency. In this paper, we presented Fuzzy C-Means based intrusion detection system. Principal Component Analysis is employed to select the most discriminate features. Afterwards a Robust Spatial Kernel FCM is used to cluster the network samples. To evaluate the efficiency of the proposed method, we conducted experiments on standard dataset and compared the results with variants of traditional Fuzzy C-Means methods and other clustering methods. The results inferred that the

proposed method outperforms the other methods. The advantage of the proposed clustering method is it considers the neighbourhood membership value and uses kernel distance metric which increases the clustering accuracy and reduces the noise effect. However, the performance of RSKFCM algorithm depends on the cluster center initialization. In future work, Evolutionary algorithm can be used to initialize the cluster centers.

References

[1] Ahmed, Mohiuddin, Abdun Naser Mahmood, and Jiankun Hu. "A survey of network anomaly detection techniques," *Journal of Network and Computer Applications, vol.* 60, pp. 19-31, 2016.

[2] S. V. Aruna Kumar and B. S. Harish, "Segmenting MRI Brain Images using novel Robust Spatial Kernel FCM (RSKFCM)," *Eighth International Conference on Image and Signal Processing*, pp. 38–44, 2014.

[3] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The Fuzzy C-Means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191–203, 1984.

[4] K. Bharti, S. Shukla, and S. Jain, "Intrusion detection using unsupervised learning," *International Journal on Computer Science and Engineering*, vol. 1, no. 2, pp. 1865–1870, 2010.

[5] J. Camacho, "Visualizing big data with compressed score plots: approach and research challenges," *Chemometrics and Intelligent Laboratory Systems*, vol. 135, pp. 110–125, 2014.

[6] W. Chimphlee, A.H. Abdullah, M.N.M Sap, S. Srinoy and S. Chimphlee, Anomaly-based intrusion detection using fuzzy rough clustering. IEEE *International Conference on Hybrid Information Technology, ICHIT'06.* vol. 1, pp. 329-334, 2006.

[7] D. E. Denning, "An intrusion-detection model," *IEEE Transactions on software engineering*, no. 2, pp. 222–232, 1987.

[8] Y. Fukuyama and M. Sugeno, "A new method of choosing the number of clusters for fuzzy c-means method," *In Proceedings of Fifth Fuzzy Systems Symp*, pp. 247–250, 1989.

[9] S. Ganapathy, K. Kulothungan, P. Yogesh, and A. Kannan, "A no accuracy weighted fuzzy c–means clustering based on immune genetic algorithm or intrusion detection," *Procedia Engineering*, vol. 38, pp. 1750–1757, 2012.

[10] F. Guorui, Z. Xinguo, and W. Jian, "Intrusion detection based on the semi-supervised fuzzy c-means clustering algorithm," in *Consumer Electronics, Communications and Networks (CECNet), 2012 2nd International Conference on* IEEE, pp. 2667–2670, 2012.

[11] S. M. Hameed, S. Saad, and M. F. AlAni, "An extended modified fuzzy possibilistic c-means clustering algorithm for intrusion detection," *Lecture Notes on Software Engineering*, vol. 1, no. 3, p. 273, 2013.

[12] H. Hotelling, "Analysis of a complex of statistical variables into principal components." *Journal of educational psychology*, vol. 24, no. 6, pp. 417, 1933.

[13] M. Jianliang, S. Haikun, and B. Ling, "The application on intrusion detection based on k-means cluster algorithm," in *Information Technology and Applications, 2009. IFITA'09. International Forum on*, vol. 1. pp. 150–152, 2009.

[14] Y. Jun-lan, "Intrusion detection technology and its future trend," *Information Technology*, vol. 4, pp. 172–176, 2006.

[15] P. Kushwaha and R.. Welekar. "Feature Selection for Image Retrieval based on Genetic Algorithm." *International Journal of Interactive Multimedia and Artificial Intelligence*, Vol.4, pp. 16- 21.2016.

[16] G. Karthik and A. Nagappan, "Intrusion detection system using kernel fcm clustering and bayesian neural network," *International Journal of Computer Science and* Information *Technology and Security*, vol. 3, no. 6, pp. 391–399, 2013.

[17] S. Khazaee and M. S. Rad, "Using fuzzy c-means algorithm for improving intrusion detection performance," in *2013 13th Iranian Conference on Fuzzy Systems*, pp. 27–29, 2013.

[18] S. Khazaee and K. Faez, "A novel classification method using hybridization of fuzzy clustering and neural networks for intrusion detection," *International Journal of Modern Education and Computer Science*, vol. 6, no. 11, p. 11, 2014.

[19] G. R. Kumar, N. Mangathayaru, and G. Narsimha, "An approach for intrusion detection using fuzzy feature clustering," IEEE *International Conference on Engineering & MIS (ICEMIS)*, pp. 1–8, 2016.

[20] N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy c-means model," *IEEE Transactions on Fuzzy systems*, vol. 3, no. 3, pp. 370–379, 1995.

[21] N. Pandeeswari and G. Kumar, "Anomaly detection system in cloud environment using fuzzy clustering based ANN," *Mobile Networks and Applications*, vol. 21, no. 3, pp. 494–505, 2016.

[22] J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media, 2013.

[23] W. Ren, J. Cao, and X. Wu, "Application of network intrusion detection based on fuzzy c-means clustering algorithm," *Third International Symposium on Intelligent Information Technology Application, 2009. IITA*, vol. 3. IEEE, 2009, pp. 19–22. 2009.

[24] Z. Rustam and A. S. Talita, "Fuzzy kernel c-means algorithm for intrusion detection systems," *Journal of Theoretical and Applied Information Technology*, vol. 81, no. 1, p. 161, 2015.

[25] R. Sampat and S. Sonawani, "Network intrusion detection using dynamic fuzzy c means clustering." *Network*, vol. 2, no. 4, 2015.

[26] H Seth, and S. D. Bay. "The UCI KDD Archive [http://kdd. ics. uci.edu]. Irvine, CA: University of California." Department of Information and Computer Science 152, 1999.

[27] S. Surana, "Intrusion detection using fuzzy clustering and artificial neural network," Advances in Neural Networks, Fuzzy Systems and *Artificial Intelligence, ISBN-978-960-474-379-7*, 2013.

[28] I. Syarif, A. Prugel-Bennett and G. Wills. Unsupervised clustering approach for network anomaly detection. *Networked Digital Technologies*, pp.135-145. 2012.

[29] H. Wang, Y. Zhang, and D. Li, "Network intrusion detection based on hybrid fuzzy c-mean clustering," *Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSKD),* vol.1., pp. 483–486. 2010.

[30] X. L. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 13, no. 8, pp. 841–847, 1991.

[31] L. Xie, Y. Wang, L. Chen, and G. Yue, "An anomaly detection method based on fuzzy c-means clustering algorithm," in *Second International Symposium on Networking and Network Security (ISNNS, 10) Jinggangshan, PR China*, pp. 89–92. 2010.

[32] W. L. Yaseen, A.O Zulaiha and A N M Zakree. "Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system." *Expert Systems with Applications*, vol.67, pp. 296-303, 2017.

B S Harish

He obtained his B.E in Electronics and Communication (2002), M.Tech in Networking and Internet Engineering (2004) from Visvesvaraya Technological University. He completed his Ph.D. in Computer Science (2011); thesis entitled "Classification of Large Text Data" from University of Mysore. He is presently working as an Associate Professor in the Department of Information Science & Engineering, JSS Science & Technology University, Mysuru. He was invited as a Visiting Researcher to DIBRIS - Department of Informatics, Bio Engineering, Robotics and System Engineering, University of Genova, Italy from May-July 2016. He delivered various technical talks in National and International Conferences. He has invited as a resource person to deliver various technical talks on Data Mining, Image Processing, Pattern Recognition, Soft Computing. He is also serving and served as a reviewer for National, International Conferences and Journals. He has published more than 40 International reputed peer reviewed journals and conferences proceedings. He successfully executed AICTE-RPS project which was sanctioned by AICTE, Government of India. He also served as a secretary, CSI Mysore chapter. He is a Member of IEEE (93068688), Life Member of CSI (09872), Life Member of Institute of Engineers and Life Member of ISTE. His area of interest includes Machine Learning, Text Mining and Computational Intelligence.

S V Aruna Kumar

He received his B.E degree in Computer Science and Engineering from Kuvempu University and M.Tech degree in Software Engineering from Visvesvaraya Technological University, India. He is currently pursuing Ph.D degree in Computer Science from Visvesvaraya Technological University. His area of research includes Machine Learning, Pattern Recognition, Image Processing and Soft Computing.

Combining Fuzzy AHP with GIS and Decision Rules for Industrial Site Selection

Aissa Taibi* and Baghdad Atmani

Laboratoire d'Informatique d'Oran (LIO), University of Oran1 Ahmed Ben Bella, BP 1524 EL Mnaouer Oran (Algeria)

Received 13 April | Accepted 9 June 2017 | Published 30 June 2017

Abstract

This study combines Fuzzy Analytic Hierarchy Process (FAHP), Geographic Information System (GIS) and Decision rules to provide decision makers with a ranking model for industrial sites in Algeria. A ranking of the suitable industrial areas is a crucial multi-criteria decision problem based on socio-economical and technical criteria as on environmental considerations. Fuzzy AHP is used for assessment of the candidate industrial sites by combining fuzzy set theory and analytic hierarchy process (AHP). The decision rule base serves as a filter that performs criteria pre-treatment involving a reduction of their numbers. GIS is used to overlay, generate criteria maps and for visualizing ranked zones on the map. The rank of a zone so obtained is an index that guides decision-makers to the best utilization of the zone in future.

Keywords

Geographic Information Systems (GIS), Multi-Criteria Decision Analysis (MCDA), Industrial Site Selection, MCDA-GIS Integration, Fuzzy Analytic Hierarchy Process (FAHP), Decision Rules.

DOI: 10.9781/ijimai.2017.06.001

I. Introduction

Site selection is the process of finding the suitable sites for a project establishment depending on socioeconomic and environmental criteria [1]. Author in [2] identifies some determining factors in project management literature: 1) The complexity of projects does not depend only on the industry; 2) Project development is not just the realm of the project managers and their team, but there is an important role of stakeholders to play; 3) Uncertainty and continuous change in projects creates a difficulty to plan, and accomplish schedules, resources and budgets. This study takes these factors into account with others for selecting industrial sites and shows that a well measured location reduces the project complexity.

Issues related to the field are problems of zonal aptitudes in the context of decision support. Complexity in decision making for selecting industrial sites derives from the inherent trade-offs between socioeconomic, technical and environmental criteria. Chaotic location causes epidemics and attacks on the health of the citizens. The linear model of Simon (intelligence, design and choice) and its extensions are insufficient to deal with the problem complexity [3].

Each zone is a spatial action only if it is defined by its geographical position, shape and spatial relations [4]. Majority of criteria and all actions of the problem under study have a geographical character. We adopt the approach of coupling between GIS and MCDA methods as the latter is favored by this problem.

The conceptual idea beyond MCDA-GIS integration approach is twofold: on the one hand, it is to use the functions of GIS to prepare geographic data necessary in inputs for the MCDA methods, and on the other; it is to employ GIS visualization potentialities to map the analysis results [5].

Geographic information systems (GIS) are needed to model, store, manage, view, analyze, and represent objects or collection of spatial objects. Action assessments according to geographic criteria are based on the most important feature of GIS: Mapping. A map is a model of the reality privileging the geometrical representation of objects with a graphic and semiotic logic [6]. For example the seismic value of an industrial zone derives from its geographical position on the seismic map.

Analytic hierarchy process (AHP) is a popular method used as a tool for multi-criteria decision making (MCDM) or as a weight estimation technique. AHP is used in many areas such as site selection, territory planning, new energy, and so on [7]. AHP allows the solution of problems with geographic nature or GIS-based [8]. This method has mathematical properties and allows total ranking, it requires a hierarchy of the decision problem and a pairwise comparisons of entities in every node of the hierarchy [9].It is remarkable that integrated AHPs are better than the stand-alone AHP [10].

This paper aims to provide decision makers with a ranking model for industrial site selection based on GIS-FAHP integration. For this model to be more efficient in supporting decision making we propose:

1. A mixed integration mode between GIS and AHP to give more flexibility to take into account the geographical character of the data.
2. A set of decision rules is modeled and used as a filter for criteria in input for the AHP method.
3. To deal with fuzziness and uncertainty, we adopt the Chang fuzzy extension [11] of AHP (FAHP) which is based on Fuzzy numbers

* Corresponding author.
E-mail address: taibiissa@yahoo.fr

with triangular membership functions.

The remainder of this article is presented as follows: Section 2 and 3 are devoted to related work and background, the proposed approach and proposed Spatial Decision Support System (SDSS) architecture are in sections4 and 5, a case study is illustrated in section 6, we end with a conclusion and perspectives.

II. Related Work

Spatial decision problems give rise to the GIS-based multi-criteria decision approach. Author in [12] underlines that in the context of the synergistic capabilities of both GIS and MCDA that theoretical and applied research on GIS-MCDA advances. In this section, we present some works on the theoretical aspects of the approach as well as some applications and we finish by our proposal.

A. Theoretical Research

The earlier works on this approach, according to [12] are that of Diamond and Wright (1988), Janssen and Reitveld (1990), Carver (1991), Langevin et al (1991).

Among the 363 articles inventoried since 1990 until 2004 in [12], there are 34 (9.4%) that uses AHP and 47% uses vector mode for geographical representation. Author in [13] develop a Decision Making Support System (DMSS) for land reform in South Africa, he gives new article statistics since 1990 up to 2015 showing the increase in using the GIS-MCDA integration approach. Author in [14] investigates on the motivation behind the approach and interrogates: What this integration is exactly needed for, and if it is really interesting to invest labor in the development of such solutions. In [15] F.Joerin links between land use planning and GIS-MCDA integration and affirms that GIS and MCDA constitute a very interesting package for land management. Author in [5] develops a strategy of GIS-MCDA integration and formulates three integration modes.

B. Applied Research

Among works using GIS-MCDA approach for industrial site selection we cite:[16], in this work , after declaration that site selection influences the life style of surrounding communities, for this, analysts must strive to determine the optimum location. The proposed approach entails two phases: screening and evaluation, the author integrates three decision tools: Expert system (ES), GIS and MCDA and uses COM (Component Object Model) concept for connecting them, Visual Rule Studio to develop the expert system, ArcGIS 8.2 to provide a GIS platform and Microsoft Excel to provide tools to implement AHP. In [17], following the work cited above, authors value the good selection of an Industrial Site and estimate that 80% of data used in this field by managers are geographical in nature, they feel that the synergistic effect generated by coupling GIS and MCDA contributes to the efficiency and quality of spatial analysis for industrial site selection, in this work AHP and WLC (Weighted Linear Combination) methods are used.

In the optic of environmental protection, we cite: [18] where the objective is to evaluate the ecological impacts of projects in the context of sustainability assessments. GIS and Analytic Network Process (ANP) are used to identify potential ecological corridors in the Piedmont Region (Northern Italy). In [19] author highlights that environmental research benefits from the integration between GIS and MCDA, he indicates that both MCDA tools and GIS potentialities are needed for spatial multi-criteria evaluation, he presents a new GIS-MCDA integration tool for wastewater land application in agriculture. In [20], the author integrates AHP with GIS to locate and rank suitable sites for soil aquifer treatment by considering technical, socio-environmental and cost criteria. In [21] authors attempt to demonstrate how GIS-MCDA integration approach helps to set apart the preferences of regional stakeholders to simplify the decision making process in planning park in Canada.

Another work relating to the energy diversification is to design a model based on GIS-MCDA integration for conducting wind energy project [22]. In [23] risk index and cost are the main criteria used to determine the preferable route of power transmission line (PTL) using conjointly AHP and PROMETHEE, GIS and MCDA are used synergistically to generate the best solution. In [24] authors couple AHP and OWA(Ordered Weighting Averaging) with GIS for mapping accessibility patterns of housing development sites in Cammore Alberta, the proposed system supports housing developers to trade off between benefits and costs to access facilities by residents.

Most works cited above use AHP, which is based on the use of crisp numbers while the modeling of data in decision field requires fuzzy numbers. Since fuzziness is a common characteristic of decision making problems, the FAHP method was developed to address this problem. In [25] author makes a state-of-the-art survey of FAHP, he underlines that 190 research application papers are published between 2004 and 2016 to argue that FAHP has been applied to a wide range of applications. In [26] the author assesses vulnerability to earthquake hazards through spatial multi-criteria analysis of urban areas; he uses FAHP and GIS to treat criteria for social risks, induced risks and systemic vulnerability. In [27] the problem is to select the best location for wastewater lift station by designing a model based on FAHP and GIS. Author in [28] highlights the impact of landfill selection on the environment, economy and ecology, he proposes a two-step solution: (i) use of GIS to determine candidate areas; (ii) use of fuzzy multi criteria decision-making (FAHP) to rank them. In [29] the authors propose a system which integrates fuzzy set theory (FST), factor rating system (FRS) and simple additive weighting (SAW) to evaluate facility locations alternatives. The purpose in [30] and [32] is to protect Turkish lakes by integrating GIS and AHP for landfill site selection. In [31] author focuses on the final disposal as a stage of municipal solid waste (MSW) and uses a fuzzy multi-criteria method to determine the best location. In [33] the authors propose an application of a new hybrid fuzzy AHP Model to locate international distribution trade centers. In the field of site selection for tourist hotels, authors in [34] treat 21 criteria using FAHP and give a clear direction for investors. In [35] there is application of FAHP approach for selecting the best underground mining method in IRAN, taking into consideration decision maker subjectivity. The purpose of the authors in [36] is to select an appropriate location to implant industrial corporation by combining FAHP and On Line Analytical Processing (OLAP) analysis.

In this paper a mixed integration is proposed: indirect integration in screening phase and full integration in the evaluation phase.

After developing the hierarchy for ranking industrial zones in Algeria using AHP, we propose a decrease of this hierarchy by reducing technical criteria with decision rules. In this way we can:

- Reduce the complexity and have more coherence during the construction of pairwise comparison matrixes.
- Use qualitative data with subjective importance required by AHP, technical criteria with objective importance are not necessary.

Linguistic variables and triangular fuzzy numbers [11] are adopted to model the decision maker's (DM) hesitance and subjectivity attitudes. We will describe in detail, our approach and its implementation in section IV and V.

III. Background

We begin this section by arguing the necessity to use the GIS-MCDA integration approach to get the best solution for spatial decision problems and then basic concepts of FAHP and decision rules are presented.

A. GIS-MCDA Integration

Solution of spatial decision problems means combining and transforming geographical as well as decision data (input) to have a resulting decision (output).

GIS is an excellent tool for acquisition, storage, manipulation and spatial analysis of geographic data, but has a lack to deal with spatial decision problems especially with conflicting objectives in the decision-making process. Below some critics addressed to GIS technology [5].

- Decision data like decision maker's preferences are not taken into account by current GIS.
- Assessment and comparison of different scenarios are not permitted by GIS. The solutions given by GIS satisfy all criteria simultaneously.
- Analytic functionalities found in most GIS are oriented towards the management of data, but not towards an effective analysis of them.

MCDA is a technique to assess and structure actions according to a set of conflicted criteria and a proposed decision maker's preference. MCDA comes to overcome the lack of GIS in tackling spatial decision problems. Salem in [5] underlines the necessity of GIS-MCDA integration and proposes three modes: (1) An indirect GIS-MCDA integration mode; (2) A built-in GIS-MCDA integration mode; (3) A full GIS-MCDA integration mode.

GIS and MCDA are two different areas of research, but they complement each other for solving spatial decision problems.

The conceptual idea on which GIS-MCDA integration is based is to use the capabilities of GIS to prepare inputs necessary for multi-criteria method and exploit the potential of GIS for results visualization.

B. AHP and its Fuzzy Extension

AHP has been studied extensively and used in almost all the applications related with MCDA in the last 20 years [37]. AHP has the flexibility to combine quantitative and qualitative data, to handle different groups of actors, to combine the opinions expressed by many experts, and can help in stakeholder analysis [38]. AHP is based on the additive weighting model and can be used in GIS environment in the same way that the WAC (Weighted Additive Combination) method.

Defect in original AHP is the use of crisp numbers, while uncertainty and vagueness of the experts' opinion are the prominent characteristic of decision making problems. FAHP is developed to address this problem and overcomes this defect. In applications it is often convenient to work with Triangular Fuzzy Numbers (TFNs) because of their computational simplicity, and they are useful in promoting representation and information processing in a fuzzy environment. In addition, TFNs are the most utilized in FAHP studies [39]. Linguistic variables and triangular fuzzy numbers are used to model the decision maker's (DM) hesitance and subjectivity attitudes when constructing pairwise comparison matrixes [11]. This paper adopts TFNs in the FAHP and describes their algebraic.

A TFN can be defined by a triplet (l, m, u) and the membership function $\mu_M(x)$ is illustrated in Fig. 1 and defined in equation (1) [39]:

$$\mu_M(x) = \begin{cases} 0 \;\; si\; x < l \\ \frac{x-l}{m-l} \;\; si\; l \le x \le m \\ \frac{h-x}{h-m} \;\; si\; m \le x \le h \\ 0 \; si\; x \ge h \end{cases} \quad (1)$$

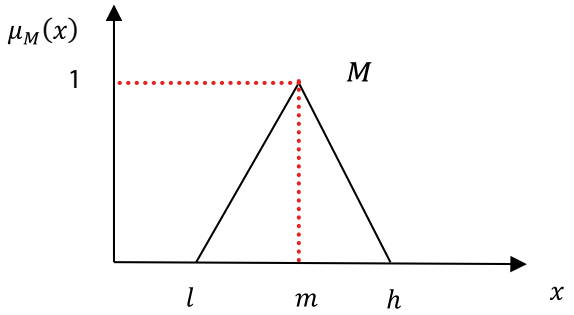


Fig. 1. Fuzzy triangular function.

There are various operations on TFNs; four operations used in this paper are illustrated.

Let two TFNs A and B be defined by the triplets:

A = (a, b, c) and B = (d, e, f).

Then:

$$A \oplus B = (a,b,c) \oplus (d,e,f) = (a+d, b+e, c+f)$$
$$A \ominus B = (a,b,c) \ominus (d,e,f) = (a-d, b-e, c-f)$$
$$A \otimes B = (a,b,c) \otimes (d,e,f) = (a.d, b.e, c.f)$$
$$A/B = (a,b,c)/(d,e,f) = \left(\frac{a}{d}, \frac{b}{e}, \frac{c}{f}\right)$$

The two sets:

$X = \{x_1, x_2, x_3, \ldots, x_n\}$ as an object set, and

$G = \{u_1, u_2, u_3 \ldots, u_m\}$ as a goal set. According to the principles of Chang's extent analysis [44], each object is considered correspondingly and extensive analysis for each of the goals, g_i, is executed. It means that *m* extent analysis values for each object can be obtained using the following signs:

$$M_{g_i}^1, M_{g_i}^2, \ldots M_{g_i}^m \qquad i=1,2,\ldots,3 \quad (2)$$

Where $M_{g_i}^j \;\; j = 1 \ldots m$ are triangular fuzzy numbers and:

The value of fuzzy synthetic extent with respect to the i^{th} object is represented as:

$$S_i = \sum_{j=1}^m M_{g_i}^j \otimes \left[\sum_{i=1}^n \sum_{j=1}^m M_{g_i}^j\right]^{-1} \quad (3)$$

The fuzzy addition operation of m extent analysis values must be performed for particular matrix to obtain $\sum_{j=1}^m M_{g_i}^j$ such that:

$$\sum_{j=1}^m M_{g_i}^j = \left(\sum_{j=1}^m l_j, \sum_{j=1}^m m_j, \sum_{j=1}^m h_j\right) \quad (4)$$

Then, we perform the fuzzy edition operation of m extent analysis values for a particular matrix to obtain

$$\sum_{i=1}^n \sum_{j=1}^m M_{g_i}^j = \left(\sum_{i=1}^n l_i, \sum_{i=1}^n m_i, \sum_{i=1}^n h_i\right) \quad (5)$$

And the inverse of the vector in Eq. (3) is computed such that

$$\left[\sum_{i=1}^n \sum_{j=1}^m M_{g_i}^j\right]^{-1} = \left(\frac{1}{\sum_{i=1}^n h_i}, \frac{1}{\sum_{i=1}^n m_i}, \frac{1}{\sum_{i=1}^n l_i}\right) \quad (6)$$

C. Decision Rules

The kind of preferential information modeled by decision rules seems to be close to the natural reasoning. This type of rules can be used alone to tackle decision problem. In [40] author presents an approach called Dominance based Rough Set Approach (DRSA) based on decision rules and dominance principle for decision support. Author in [5] proposes a model based on rules for choosing the appropriate method

according to the proposed problem. The rules cover three categories of facts: i) the characteristics of spatial decision-making problem, ii) the characteristics of decision makers and iii) characteristics of the method itself. The simplicity in modeling by rules, facilitate for experts to formulate knowledge. Author in [41] develops and implements a rule base for selection of landfill sites in Switzerland; he raised practically the problem of explosion in the number of rules needed for this specific case (for 10 criteria there are 310 combinations). To find weights of deducted criteria (criteria in conclusion) he adds the weights of basic criteria (criteria in premises) without returning to the decision maker.

It seems advantageous to combine the MCDA techniques with rule base. The rule base serves as a filter that performs a pre-treatment of information and consequently reduces the amount of data to be processed by MCDA.

IV. Proposed Approach

In this section, we present the main aspect of our contribution. As mentioned in the previous section, Salem in [5] proposes three integration modes, indirect integration, built-in integration and full integration. In this paper a mixed integration is proposed: Preparing geographic criteria to support decision making in weighting is made by GIS independently (indirect integration) in screening phase while visualization function is integrated directly (full integration) in the MCDA module in the evaluation phase and is considered as a finality of the decision analysis (Fig. 2).

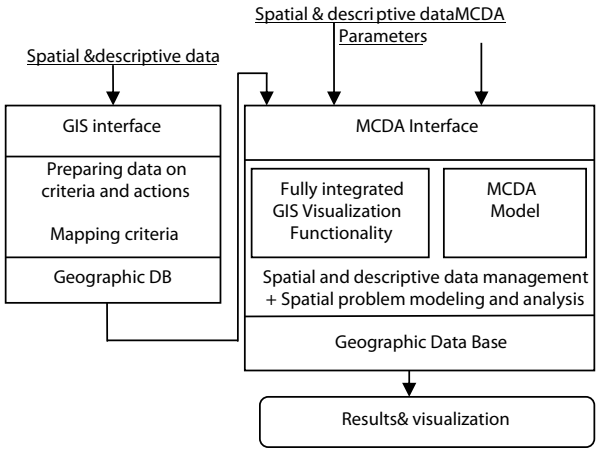


Fig. 2. GIS-MCDA proposed integration mode.

The proposed approach consists of two phases: Screening and evaluation as explained below and depicted in (Fig. 3)

- Screening phase: After the choice of zones at the beginning of the decision-making process, zones studies begins by collecting data about criteria and actions from geographic, socioeconomic and climate databases as archives of regions, the maps for the criteria are built using GIS. A field of expertise is available. A rule base is constructed in order to reduce the number of criteria before decision analysis.
- Site evaluation phase: consists of the total ranking by combining FAHP and a specific rule base. The AHP procedure involves six essential steps [9, 42]:

1. Definition of the problem
2. Represent the problem by an hierarchy structure
3. Construct pairwise comparison matrixes
4. Estimate the relative weights
5. Check the consistency

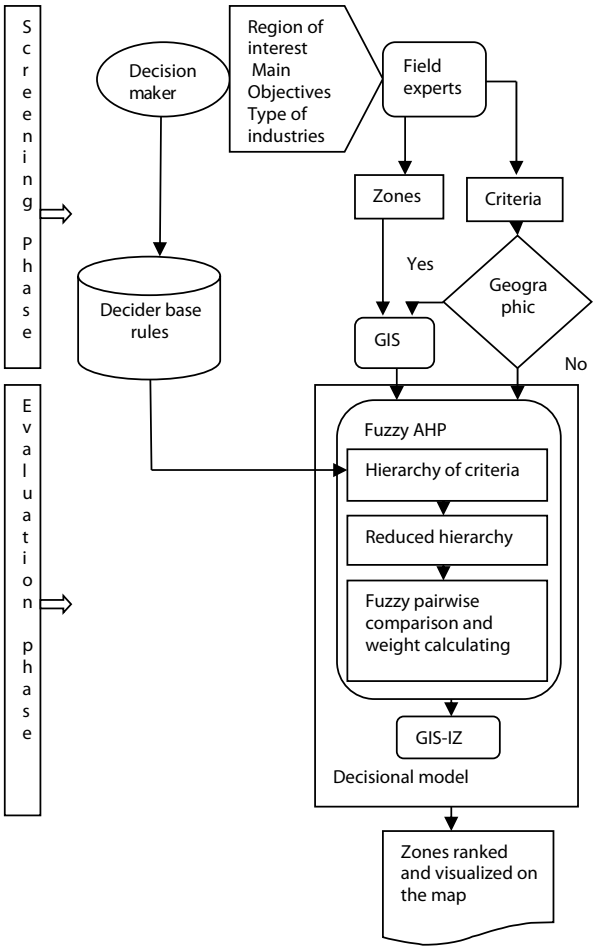


Fig. 3. Framework of Proposed Approach.

6. Obtain the overall rating

After developing hierarchy for a specific problem, we propose a reduction of technical criteria by using decision rules. The objective is to give more flexibility for studying geographical criteria and mapping alternatives. These rules are modeled and used as a filter for criteria in input, which allows a considerable reduction of the problem hierarchy. The purpose is to reduce the technical criteria into reduced ones with subjective importance, it is to be noted that AHP is a subjective methodology (Fig. 4). Thus a step named “Reducing the problem hierarchy” is added after step2.

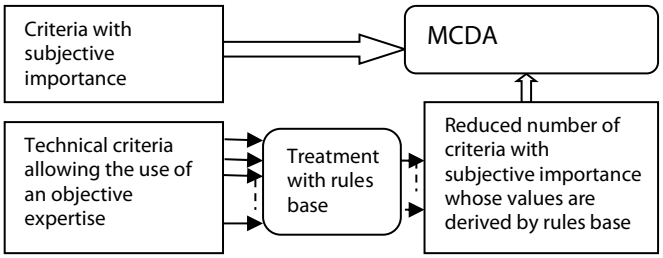


Fig. 4. Combining multi-criteria analysis and rule base inspired from [43].

To deal with uncertainty and decision maker hesitation in pairwise comparison (Step 3) there are many fuzzy-AHP methods proposed by various authors (Buckley, 1985; Chang, 1996; Cheng, 1997; Deng, 1999; Leung and Cao, 2000; Mikhailov, 2004; VanLaarhoven and Pedrycz, 1983). Chang Fuzzy AHP extension [11] as explained in subsection (III, A) is conforming to this study.

Decision maker gives his pairwise comparison matrixes according to Saaty9-unit scale as shown in Table I [9].

TABLE I. Scale of Relative Preference Based on Saaty [9]

Intensity of preferences Numerical value	Definition (Verbal Scale)
1	Equally preferred
3	Weak preference of one over other
5	Strongly preferred
7	Very strongly preferred
9	Extremely preferred
2,4,6,8	Intermediate values between the two adjacent judgments

The result matrixes of the pairwise comparisons made by Decision Maker are transformed into Fuzzy ones by means of linguistic variables, represented by triangular fuzzy numbers listed in Table II.

Extent analysis method (in subsection III, A) is used to evaluate the importance weight and consequently the rank of each action (Industrial zone).

TABLE II. Triangular Fuzzy Numbers for Pairwise Comparisons

Linguistic variables	TFNs	Reciprocal TFNs
Extreme strongly preferred	(7,9,9)	(1/7,1/9,1/9)
Intermediate	(6,8,9)	(1/6,1/8,1/9)
Very strongly preferred	(5,7,9)	(1/5,1/7,1/9)
Intermediate	(4,6,8)	(1/4,1.6,1/8)
Strongly preferred	(3,5,7)	(1/3,1/5,1/7)
Intermediate	(2,4,6)	(1/2,1/4,1/6)
Moderately preferred	(1,3,5)	(1,1/3,1/5)
Intermediate	(1,2,4)	(1,1/2,1/4)
Equally preferred	(1,1,1)	(1,1,1)

V. Proposed Spatial Decision Support System (SDSS)

Inspired from the general Decision Making Support System Architecture in [43], taking into account the geographical aspect of data, the proposed DMSS has the architecture shown in Fig. 5.

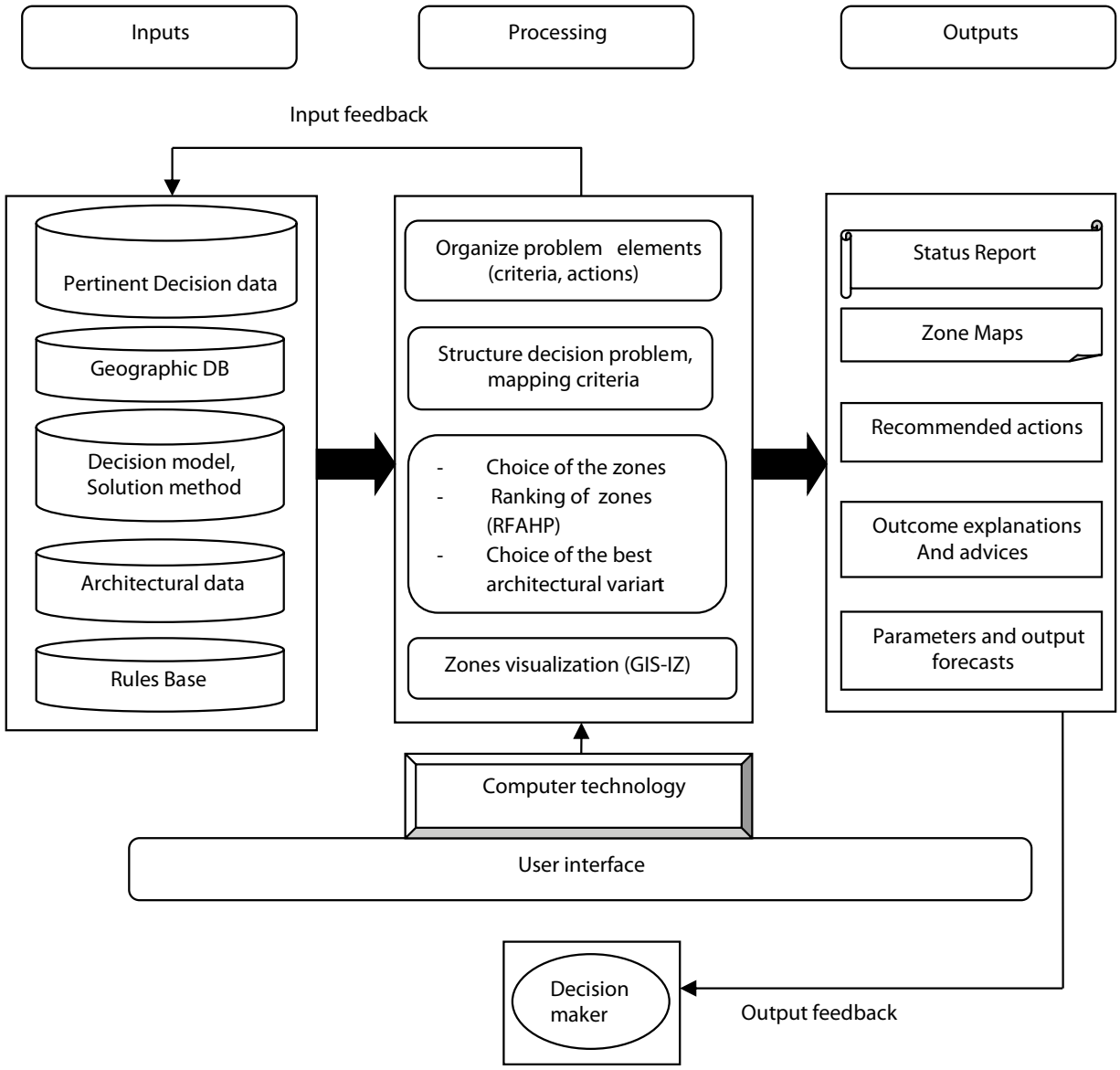


Fig. 5. General Spatial Decision Support System Architecture.

DMSS Inputs: include a database of pertinent decision data, a geographic database for actions and criteria, a rule base and an appropriate method (AHP in this case).The decision maker utilizes computer technology via a user interface to access at various bases or executes the processing.

DMSS Processing: Process in three phases:

First phase: It is the phase of determining the aptitude of candidate zones in different regions. Zonal aptitude is defined by the properties of an area to satisfy pre-established conditions or to be favorable to the development of a given phenomenon [6].

The Boolean zoning is a multi-criteria decision-making approach using a single logic operator (∩) to combine the criteria for selecting zones, for example the aptitude binary index $C_{slope,j}$ for a zone j according to the slope criterion is given such that:

If $C_{slope,j} \leq 20°$ Then 1, else 0.

The aptitude of a zone is calculated using the intersection of all aptitude criteria binary values such that:

$$I_{apt,j} = C_{1,j} \cap C_{2,J} \cap \ldots \cap C_{k,j}$$

$C_{k,j}$ is the aptitude binary value of the zone j according to the criterion k.

Second phase (Aim of this paper): It is the total ranking of zones using conjointly decision rules and RFAHP (Ranking FAHP) unit to develop a ranking for zones.

Decision rules set specified for this case is used to reduce the number of criteria, this reduction improves results, reduces the complexity and decreases inconsistencies. Technical criteria allowing the use of an objective expertise are reduced using a set of adequate rules; we have in result, criteria with subjective importance whose values are derived by decision rules . After this step FAHP deals with reduced subjective criteria only.

Third phase: Consists to choose one of three available architectural variants, the selection criteria are: architecture, management cost, number of fragmented islands and the types of planned investments.

GIS-IZ module ensures the display of zones before and after each decision-making phase. To accomplish this task the vector mode is adopted, each industrial zone is considered as a geographical entity of the abstract spatial type "POINT", it is implemented with their geographical position (latitude and longitude) using Geo-Tools in three steps:

1. Introduce the file with shape file (SHP) extension, which represents the Algerian administrative division map.
2. Insert another file with the same extension representing industrial zones.
3. Make a projection between the two thematic maps.

DMSS Outputs: GIS-IZ unit displays the zones with their corresponding ranks on the map of Algeria as in Fig.6.

VI. Case Study

Industrial zones under study are programmed by the Algerian state in 2013 to satisfy the expectations of investors. The objective is to reach the economic regional balance and population stabilization with respect to environment [44].

A. Actions

Industrial zones (Actions) under study are: Action1 (A1): SBA, Action2 (A2): Horchaia, Action3 (A3): Ras Elma, Action4 (A4): Maghnia, Action5 (A5): Kolea.

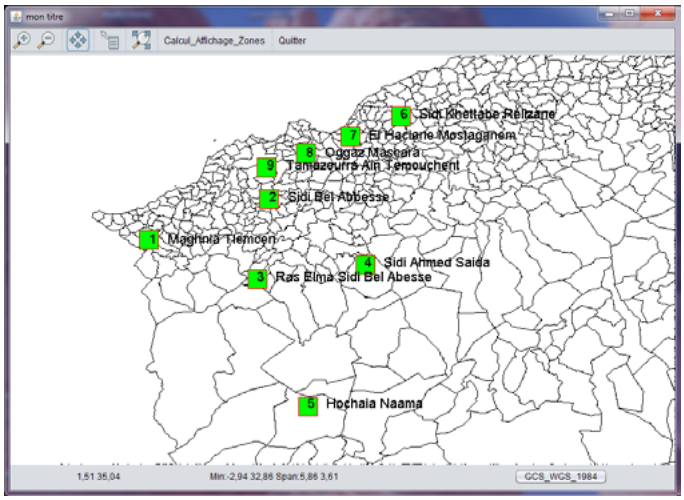


Fig. 6. Main visualizing interface.

B. Criteria

Criteria used in this study are classified into four categories:

Natural risk (Cr1), Socio-economic (Cr2), Environmental impact (Cr3) and Climate characteristics Cr4). According to these categories, 11 different evaluation criteria are defined. Fig. 7 shows the hierarchical structure of this problem according to the first step of AHP.

(Cr11): Risk groundwater pollution.

(Cr12): fauna and flora problem.

(Cr13): Citizens noises.

(Cr21): Seismicity.

(Cr22): Flood.

(Cr31): Temperature.

(Cr32): Rainfall.

(Cr33): bioclimatic Floor.

(C41): Cost management.

(Cr42): Equipment and development potentiality.

(Cr43): Transport Infrastructure.

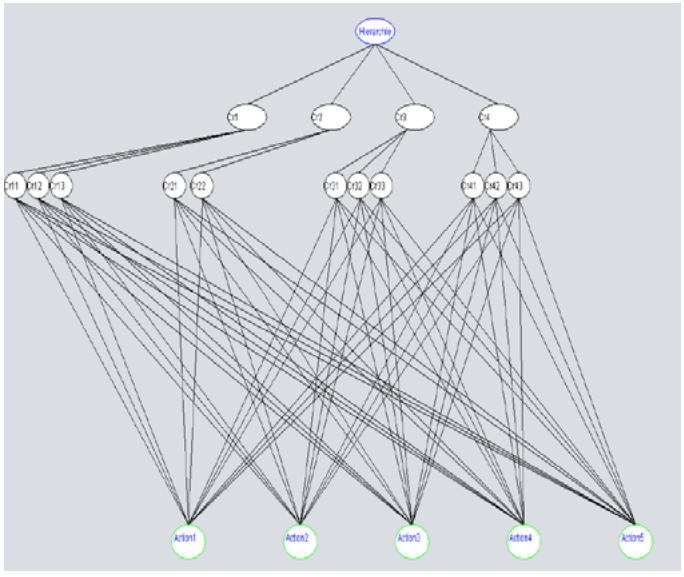


Fig. 7. Representation of the hierarchical structure, source (our SDMSS).

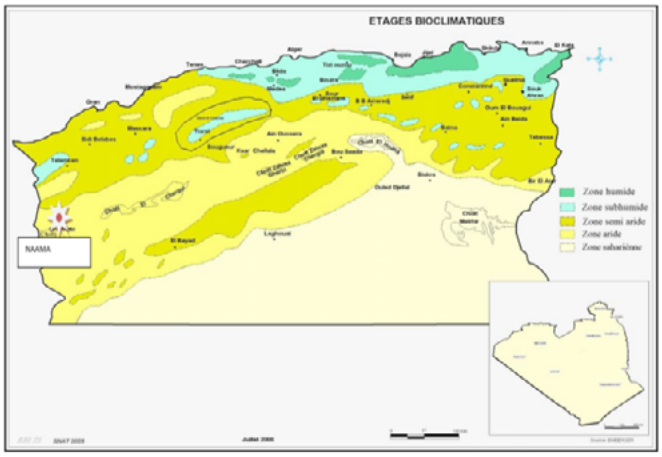


Fig. 8. Representation of bioclimatic floors of Algeria [44].

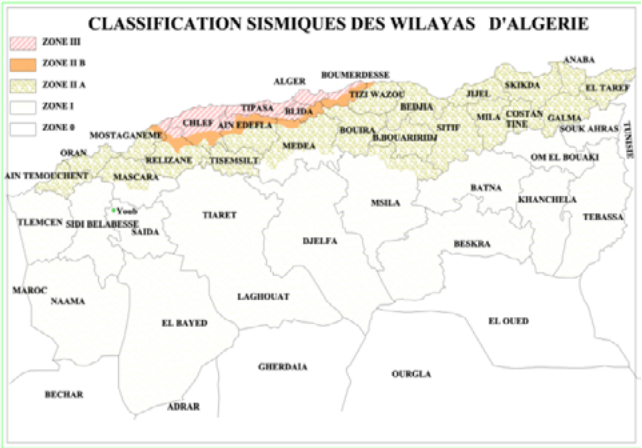


Fig. 9. Representation of seismic classification of Algeria.

C. Rule base

Rule base must be developed by experts since the deductions must be based on the field expertise. Problems encountered during development of this rule base are:

- High number of rules is necessary when tackling this problem, especially when the criteria number is elevated.
- There is a lot of possible criteria value in premises rules.
- Difficulty of ensuring, maintenance, coherence and completeness of the rule base.

To overtake these problems, more general rules are considered and when modeling rules premises, interval values are used instead of simple values if it is possible. Below the list of rules used in this study:

1. Rules set according to environmental impacts

   Rules structure:

   In conclusion rules, environmental impact (Cr1) can be strong, low or moderate.

   Premises rules contain a combination of the following criteria values:

   (Cr11): Risk groundwater pollution (present, absent) (Cr12): Fauna and flora problem (existent, inexistent) (Cr13): Noise Citizen (strong, low)

   Proposed rules:

   R1: If (Cr11= present) and (Cr12= exist) and (Cr13= strong) Then Cr1= strong.

   R2: If (Cr11 = present) and (Cr12 = inexistent) and (Cr13= strong) Then Cr1 = strong.

   R3: If (Cr1=absent) and (Cr2= inexistent) and(Cr3= low) Then Cr1 = low.

   In other cases, Cr1 = moderate.

2. Rules set according to natural risks.

   Rules structure:

   In conclusion rules, natural risks (Cr2) can be strong, low or moderate.

   Premises rule contains a combination of the following criteria values:

   (Cr21): Seismicity (low to moderate, low, strong)

   (Cr22): Flood (low, moderate, strong)

   Proposed rules:

   R1: If (Cr21 = low to moderate) and (Cr22 = moderate) Then Cr2 = moderate.

   R2: If (Cr21 = strong) and (Cr22 = strong) Then Cr2 = strong.

   R3: If (Cr21 = strong) and (Cr22 = moderate) Then Cr2 = strong.

   R4: IF (Cr21 = strong) and (Cr22 = low to moderate) Then Cr2= strong.

   R5: If (Cr21 = low and Cr22= low to moderate) Then Cr2 = moderate.

   In other case, Cr2=moderate.

3. Rules set according to climate characteristics

   Rules structure:

   In conclusion rules, climate characteristics (Cr3) can be favorable, unfavorable or little favorable.

   Premises rules contain a combination of the following criteria values:

   (Cr31): Temperature (numeric values)

   (Cr32): Rainfall (numeric values interval)

   (Cr33): Bioclimatic floor (Arid, Semi Arid).

   Proposed rules:

   R1: If (15 < Cr31< 20) and (50<Cr32<200) and (Cr33 =semi arid) Then Cr3= favorable.

   R2: If (15<Cr31<40) and (Cr32<400) and (Cr33=semi arid) Then Cr3 = favorable.

   R3: If (Cr31<10) Then Cr3= unfavorable.

   R4: If (Cr31>50) Then Cr3 unfavorable.

   In other cases, Cr3 = little favorable

4. Rules set according to socio-economic criterion

   Rules structure:

   In conclusion rules, socio economic criterion (Cr4) can be very favorable, favorable, little favorable or unfavorable.

   Premises rules contain a combination of the following criteria values:

   (Cr41): Managing fee (numeric values)

   (Cr42): Development potentiality (encouraging, not encouraging, little encouraging).

   (Cr43):Transport Infrastructure (available, few available, unavailable)

   Proposed rules:

   R1: If (Cr41>3E+9) and (Cr42 =not encouraging) (Cr43= Few available) Then Cr4 = unfavorable.

R2: If (Cr41>3E+9) and (Cr42 =encouraging) and (Cr43= few available) Then Cr4 = few favorable.

R3: If (Cr41>3E+9) and (Cr42 = encouraging) and (Cr43= available) Then Cr4 = favorable.

R4: If (15E+8<Cr41<3E+9) and (Cr42 = encouraging) and (Cr43 = few available) Then Cr4 = favorable

R5: If (Cr41<15E+8) and (Cr42 = encouraging) and (Cr43 = available) Then Cr4 = very favorable.

R6: If (Cr41<15E+8) and (Cr42 =not encouraging) and (Cr43 = unavailable) Then Cr4= not favorable.

R7: If (Cr41<1E+9) and (Cr42 =not encouraging) and (Cr43= available) Then Cr4= very favorable.

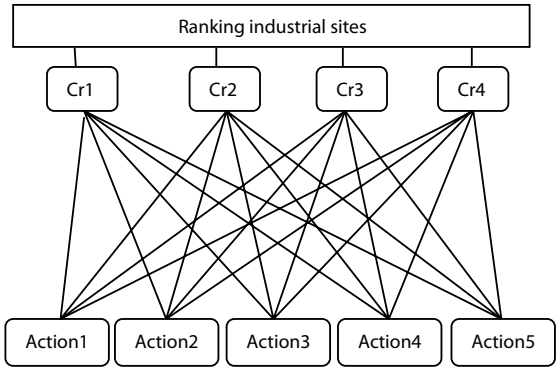


Fig. 10. The hierarchy structure after reduction.

D. Pair Wise Comparisons Matrices

All associated elements in the low hierarchy of each element in the tree structure are pairwise compared according to decision maker (DM's) opinion and Saaty 9-unit scale in Table I. Obtained matrices are transformed into fuzzy ones using the scale in Table II. Computing the normalized value of row sums (i.e. fuzzy synthetic extent) is done by fuzzy arithmetic operations. The pairwise comparison matrix for main criteria is given in Table III.

TABLE III. Pairwise Comparisons Between Main Criteria Based on the DM's Opinion.

Criterion	Env-impact (Cr1)	Nat-risk (Cr2)	Climate (Cr3)	Soc-eco (Cr4)
Env-impact(**Cr1**)	1	5	3	1/3
Nat-risk(**Cr2**)	1/5	1	1/3	1/7
Climate(**Cr3**)	1/3	3	1	5
Soc-eco(**Cr4**)	3	7	1/5	1
Total	4,53	16	4,53	6.47

The fuzzy pairwise comparisons for main criteria are illustrated in Table IV.

TABLE IV. Fuzzy Pairwise Comparisons for Main Criteria.

Criterion	Env-impact (Cr1)	Nat-risk (Cr2)	Climate (Cr3)	Soc-eco (Cr4)
Env-impact(**Cr1**)	(1,1,1)	(2,5,7)	(1,3,5)	(1/5,1/3,1/2)
Nat-risk(**Cr2**)	(1/7,1/5,1/3)	(1,1,1)	(1/5,1/3,1/2)	(1/9 ,1/7,1/5)
Climate(**Cr3**)	(1/5,1/3,1/2)	(2,3,5)	(1,1,1)	(3,5,7)
Soc-eco(**Cr4**)	(2,3,5)	(5,7,9)	(1/7,1/5,1/3)	(1,1,1)
Total	(1,1,1)	(2,5,7)	(1,3,5)	(1/5,1/3,1/2)

The values of fuzzy synthetic extent are evaluated using Eq. 1-6 as follows:

$$\sum_{j=1}^{m} M_{g_i}^{j} = \begin{pmatrix} (4.2; 9.33; 13.5) \\ (1.45; 1.67; 2.03) \\ (6.2; 9.33; 13.5) \\ (8.14; 11.2; 15.33) \end{pmatrix}$$

$$S_i = \begin{bmatrix} \frac{4.2}{44.26} & \frac{9.33}{31.53} & \frac{13.5}{20} \\ \frac{1.45}{44.26} & \frac{1.67}{31.53} & \frac{2.03}{20} \\ \frac{6.2}{44.26} & \frac{9.33}{31.53} & \frac{13.5}{20} \\ \frac{8.14}{44.26} & \frac{11.2}{31.53} & \frac{15.33}{20} \end{bmatrix} = \begin{bmatrix} 0.095 & 0.3 & 0.67 \\ 0.03 & 0.053 & 0,10 \\ 0.14 & 0.3 & 0.67 \\ 0.18 & 0.35 & 0.76 \end{bmatrix}$$

The last step is devoted to converting fuzzy values into crisp ones. It is done by applying the average function which is adequate with the attitude of the decision maker. Fuzzy weights and their corresponding crisp weights are illustrated in Table V.

TABLE V. Weight Calculating for Main Criteria

Criterion	Fuzzy weight			Crisp weight
Env-impact(Cr1)	0.095	0.3	0.67	0.355
Nat-risk(Cr2)	0.03	0.053	0.10	0.061
Climate(Cr3)	0.14	0.3	0.67	0.37
Soc-eco(Cr4)	0.18	0.35	0.76	0.43

The fuzzy pairwise comparisons for alternatives over all criteria are illustrated in Table VI.

TABLE VI.
Fuzzy Pairwise Comparison Matrix for Alternatives Over Criteria

Cr1	**A1**	**A2**	**A3**	**A4**	**A5**
A1	(1,1,1)	(1,3,5)	(1/6,1/4,1/2)	(1/5,1/3,1)	(1/9,1/7,1/5)
A2	(1/5,1/3,1)	(1,1,1)	(1,3,5)	(1/7,1/5,1/3)	(3,5,7)
A3	(2,4,6)	(1/5,1/3,1)	(1,1,1)	(1/9,1/7,1/5)	(3,5,7)
A4	(1,3,5)	(3,5,7)	(5,7,9)	(1,1,1)	(1,2,4)
A5	(5,7,9)	(1/7,1/5,1/3)	(1/7,1/5,1/3)	(1/4,1/2,1)	(1,1,1)
Cr2	**A1**	**A1**	**A1**	**A1**	**A1**
A1	(1,1,1)	(1/4,1/2,1)	(1/5,1/3,1)	(1/5,1/3,1)	(2,4,6)
A2	(1,2,4)	(1,1,1)	(1/5,1/3,1)	(1/5,1/3,1)	(1/4,1/2,1)
A3	(1,3,5)	(1,3,5)	(1,1,1)	(1/6,1/4,1/2)	(1,3,5)
A4	(1,3,5)	(1,3,5)	(2,4,6)	(1,1,1)	(1,3,5)
A5	(1/6,1/4,1/2)	(1,2,4)	(1/5,1/3,1)	1/5,1/3,1)	(1,1,1)
Cr3	**A1**	**A1**	**A1**	**A1**	**A1**
A1	(1,1,1)	(1,3,5)	(3,5,7)	(1,3,5)	(1/4,1/2,1)
A2	(1/5,1/3,1)	(1,1,1)	(1/4,1/2,1)	(2,4,6)	(1/5,1/3,1)
A3	(1/7,1/5,1/3)	(1,2,4)	(1,1,1)	(1/9,1/7,1/5)	(1/5,1/3,1)
A4	(1/5,1/3,1)	(1/6,1/4,1/2)	(5,7,9)	(1,1,1)	(1/5,1/3,1)
A5	(1,2,4)	(1,3,5)	(1,3,5)	(1,3,5)	(1,1,1)
Cr4	**A1**	**A1**	**A1**	**A1**	**A1**
A1	(1,1,1)	(3,5,7)	(4,6,8)	(1/7,1/5,1/3)	(1/5,1/3,1)
A2	(1/7,1/5,1/3)	(1,1,1)	(1/7,1/5,1/3)	(1,2,4)	(5,7,9)
A3	(1/8,1/6,1/4)	(3,5,7)	(1,1,1)	(1,3,5)	(1/6,1/4,1/2)
A4	(3,5,7)	(1/4,1/2,1)	(1/5,1/3,1/2)	(1,1,1)	(1/7,1/5,1/3)
A5	(1,3,5)	(1/9,1/7,1/5)	(2,4,6)	(3,5,7)	(1,1,1)

The priority calculating of alternatives over the different criteria are illustrated in Table VII.

To calculate the final rank of each action A_i (i = 1...5) we consider all paths that link each action with the goal in the graph of the hierarchy (Fig. 10.) according to the following formula:

$$\boldsymbol{Poid_{Flou}(Ai)} = ((\boldsymbol{Poid_{C_1}} \otimes \boldsymbol{Poid_{C1A_i}}) \oplus (\boldsymbol{Poid_{C_2}} \otimes \boldsymbol{Poid_{C2A_i}}) \oplus (\boldsymbol{Poid_{C_3}} \otimes \boldsymbol{Poid_{C3A_i}}) \oplus (\boldsymbol{Poid_{C_4}} \otimes \boldsymbol{Poid_{C4A_i}}))$$

TABLE VII. WEIGHTS AND RANKS FOR THE ALTERNATIVES OVER CRITERIA

Cr1	Fuzzy weight			Crisp weight	Rank
A1	0.032	0.091	0.243	0.122	5
A2	0.0713	0.184	0.45	0.235	4
A3	0.084	0.2	0.48	0.25	3
A4	0.147	0.34	0.82	0.43	1
A5	0.0872	0.34	0.36	0.26	2
Cr2	**Fuzzy weight**			**Crisp weight**	**Rank**
A1	0.11	0.16	0.44	0.23	3
A2	0.041	0.10	0.35	0.1636	4
A3	0.065	0.26	0.73	0.3516	2
A4	0.093	0.36	0.97	0.47	1
A5	0.04	0.10	0.33	0.156	5
Cr3	**Fuzzy weight**			**Crisp weight**	**Rank**
A1	0.092	0.29	0.795	0.3923	1
A2	0.053	0.14	0.42	0.20	4
A3	0.036	0.08	0.27	0.13	5
A4	0.096	0.20	0.52	0.272	3
A5	0.0735	0.27	0.83	0.3911	2
Cr4	**Fuzzy weight**			**Crisp weight**	**Rank**
A1	0.11	0.20	0.53	0.28	2
A2	0.097	0.17	0.45	0.24	3
A3	0.07	0.153	0.42	0.214	4
A4	0.06	0.11	0.30	0.1566	5
A5	0.095	0.21	0.588	0.297	1

For example, the result fuzzy weight, crisp weight and rank of A_1 (SBA) is calculated such that:

$$
\begin{aligned}
&\boldsymbol{Poid_{Flou}(SBA)} = \\
&((0.09 \quad 0.3 \quad 0.67)\otimes(0.03 \quad 0.09 \quad 0.24))\oplus \\
&((0.03 \quad 0.05 \quad 0.10)\otimes(0.11 \quad 0.16 \quad 0.44))\oplus \\
&((0.14 \quad 0.3 \quad 0.67)\otimes(0.09 \quad 0.29 \quad 0.79))\oplus \\
&((0.18 \quad 0.35 \quad 0.76)\otimes(0.11 \quad 0.20 \quad 0.53)) = \\
&(0.0027 \quad 0.027 \quad 0.1603)\oplus \\
&(0.0033 \quad 0.008 \quad 0.044)\oplus \\
&(0.0126 \quad 0.087 \quad 0.53)\oplus \\
&(0.02 \quad 0.07 \quad 0.4028) \\
&\boxed{\boldsymbol{Poid_{Flou}(SBA)} = (0.0386 \quad 0.192 \quad 1.1371)} \\
&\xrightarrow{Normalisation} \boxed{\boldsymbol{Poid_{Floue}(SBA)} = (0.028 \quad 0.14 \quad 0.83)}
\end{aligned}
$$

In the same way we calculate the weights and consequently the ranks of the actions A_2, A_3, A_4 and A_5.

Obtained ranks of zones according to our approach are illustrated in Table VIII.

TABLE VIII
THE FINAL RANKS FOR THE ACTIONS (OBTAINED BY AHP AND FAHP).

Industrial zones	AHP Weights (%)	AHP Ranks	FAHP Weights (%)			FAHP Ranks
SBA (A1)	21.4	3	2	14	83	3
Hourchia (A2)	17.02	4	3	16	95	1
Koléa (A3)	11.38	5	2	14	89	2
Ras Elma (A4)	25.7	1	2	14	82	5
Maghnia (A5)	21.50	2	2	16	81	4

VII. CONCLUSION

This paper proposes an approach based on the FAHP, GIS and Decision rules for ranking industrial sites in Algeria. The process of decision-making consists of a two-stage analysis:

An initial site screening followed by a detailed assessment of the suitability candidate sites. Because the human decision-making process usually contains fuzziness and vagueness, the FAHP is used to overcome this problem. GIS is used to prepare geographic data in screening phase and to visualize ranked zones on a map in the evaluation phase. Because it is advantageous to combine the MCDA techniques with rule base, a set of rules is utilized and serves as a filter that performs a pre-treatment of criteria and consequently reduces the problem complexity. In terms of complexity we save to build $11 \times (5 \times 5)$ matrixes with their fuzzy extent and priority calculating in this case. This research allowed us to determine the usefulness of the approach study for a site selection sector where the decision is important and dangerous, and intersects with the geography and even history. The rank of an industrial zone so obtained is an index with which we can:

- Criticize the choice of zones.
- Alert the planners and builders of zones.
- Assign the zone to adequate investment projects.

As a future direction we will engage Data Mining for the construction of the rule base. The conjunctive rules obtained after learning will be used for reduction criteria by exploiting the transition functions of the cellular machine CASI (Cellular Automation for Symbolic Induction). In response to the limitations of various approaches to rules simplifications, we use CASI, which makes it possible to eliminate redundant and incoherent information in order to produce an optimal set of rules [46, 47].

REFERENCES

[1] Khalid, E., "Exploring multi-criteria decision strategies in GIS with linguistic quantifiers: An extension of the Analytical Network Process using Ordered Weighted Averaging operators", International Journal of Geographical Information Science, 815356, 2013.

[2] D. Magaña Martínez and J.C. Fernández Rodríguez, "Artificial Intelligence applied to project success: a literature review", International Journal of Artificial Intelligence and Interactive Multimedia, Vol. 3, No.5, PP. 77-82, 2015.

[3] R. M. Dawes, "The Robust Beauty of Improper Linear Models in Decision Making", American psychologist, Vol. 34, No. 7, PP. 571-582, 1979.

[4] R. Balzarini, P.-A. Davoine, and M. Ney, "Evolution et développement des méthodes d'Analyse spatiale multicritère pour des modèles d'aptitude : L'exemple des applications en Géosciences", presented at Conférence Francophone ESRI, Oct. 2011, Versailles, France.

[5] Chakhar, S. , "Cartographie Décisionnelle Multicritère : Formalisation Et Implémentation Informatique ", Thèse de doctorat, D.F.R. Sciences des Organisations, Université Paris Dauphine, France,2006.

[6] R. Caloz and C. Collet, Analyse spatiale de l'information géographique, Presse polytechniques et universitaires romandes, 2011.

[7] Y.-M. Wang, Y. Luo and Z. Hua, "On the extent analysis method for fuzzy AHP and its applications", European Journal of Operational Research, PP.735–747, 2008.

[8] M.H. Vahidniaa, A. Alesheikhb, A. Alimohammadic and A. Bassirid, "A Fuzzy Analytical Hierarchy Process in GIS Application", the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol.37.part B2. Beijing, 2008.

[9] T. L. Saaty, "Decision making with the analytic hierarchy process", Int. J. Services Sciences, Vol. 1, No. 1, PP.83-98, 2008.

[10] Y. Ge, Q. Xu and H. Li, "The Design and Application of a Generic AHP Evaluation System", in Proc. 4th International Conference on Wireless Communications, Networking and Mobile Computing, Dalian. China, 2008, PP 248-252.

[11] D.Y. Chang, "Applications of the extent analysis method on fuzzy AHP", European Journal of Operational Research, 95, PP. 649-655, 1996.

[12] Malczewski, J., "GIS-based multi-criteria decision analysis: a survey of the literature", International Journal of Geographical Information Science, Vol. 20, No. 7, PP. 703-726, 2006.

[13] Musakwa, W., Makoni. E. N., Kangethe, M. and Segooa L., "Developing a decision support system to identify strategically located land for land reform in south Africa", the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL, No.1-3, PP.

196-203, 2014.
[14] Karim, L., "On the motivation behind MCDA and GIS integration", Int. J. Multicriteria Decision Making, Vol. 3, No. 2/3, PP. 101-113, 2013.
[15] Joerin, F., Thériault, M. and Musy, A., "Using GIS and outranking multicriteria analysis for land-use suitability assessment", International Journal of Geographical Information Science, Vol. 15, No. 2, PP. 153-174, 2001.
[16] Khalid, E., "A COM-based Spatial Decision Support for Industrial Site Selection", Journal of Geographic Information and Decision Analysis, Vol. 7, No. 2, PP. 72 – 92, 2003.
[17] A. Rikalovic, I. Cosic, and D. Lazarevic, "GIS Based Multi-Criteria Analysis for Industrial Site Selection", in Proc.24th DAAAM International Symposium on Intelligent Manufacturing and Automation, 2013, PP. 1054-1063.
[18] V. Ferreti, "Integrating Multicriteria Analysis and Geographic Information Systems for studying ecological corridors in the Piedmont Region", presented at the 74th Meeting of the European Working Group, Multiple Criteria Decision Aiding.
[19] G. Massei, L. Rocchi, L. Paolotti, S. Greco and A. Boggia, "Decision Support Systems for environmental management: A case study on wastewater from agriculture", Journal of Environmental Management Vol. 14, No.6, PP. 491-504, 2014.
[20] M. Anane, H. Kallali and S. Jellali, "Ranking suitable sites for Soil Aquifer Treatment in Jerba Island (Tunisia) using remote sensing, GIS and AHP-multicriteria decision analysis", Int. J. Water, Vol. 4, Nos. 1/2, 2008.
[21] C. Prévil, M. Thériault and J. Rouffignat, "Analyse multicritère et SIG pour faciliter la concertation en aménagement du territoire : vers une amélioration du processus décisionnel", Les Cahiers de géographie du Québec, Vol. 47, No. 130, PP. 35-61, 2003.
[22] M.L. Vazquez, J.-P. Waaub, A. Ilinca and J.-L. Chaumel, "Analyse spatiale et approche d'aide multicritères et multi acteurs a la négociation pour évaluer des scenarios d'implantation des parcs éoliens", présenté a la 1ere Conférence Intercontinentale d'Intelligence Territoriale «Interdisciplinarité dans l'aménagement et développement des territoires», PP.10, Oct. 2011, Gatineau, Canada.
[23] Dedemen, Y., "A multi-criteria decision analysis approach to GIS-based route selection for overhead power transmission lines", Thesis submitted to the Graduate School of Natural and Applied Sciences of Middle East Technical University. Department of Geodetic and Geographic Information Technologies, 2013.
[24] Y. Meng, J. Malczewski and S. Boroushaki, "A GIS-Based Multicriteria Decision Analysis Approach for Mapping Accessibility Patterns of Housing Development: A Case Study in Canmore, Alberta", Journal of Geographic Information System, Vol. 3, PP. 50-61, 2011.
[25] S. Kubler, J. Robert, W. Derigent, A. Voisin and Yves Le. Traon, "A state-of the art & testbed of fuzzy AHP (FAHP) applications", Experts Systems with Application, Vol. 65, PP. 398-422, 2016.
[26] T. Rashed and J. Weeks, "Assessing vulnerability to earthquake hazards through spatial multicriteria analysis of urban areas", International Journal of Geographical Information Sciences, Francis and Taylor edition, 2003.
[27] E. H. Ibrahim, S. E. Mohamed, and A. A. Atwan, "Combining fuzzy analytic hierarchy process and GIS to select the best location for a wastewater lift station in Elmahalla El-kubra, North Egypt", International Journal of Engineering & Technology, Vol.11, No. 05, PP. 37-43, 2011.
[28] V. Akbari, M.A. Rajabi, S.H. Chavoshi and R. Shams, "Landfill Site Selection by Combining GIS and Fuzzy Multi Criteria Decision Analysis, Case Study: Bandar Abbas, Iran", World Applied Sciences Journal, Vol. 3, PP. 39-47, 2008.
[29] S.-Y. Chou, Y.-H. Chang, C.-Y. Shen, "A fuzzy simple additive weighting system under group decision-making for facility location selection with objective/subjective attributes", European Journal of Operational Research, Vol. 189, PP.132-145, 2008.
[30] S. Sener, E. Sener, B. Nas and R. Karaguzel "Combining AHP with GIS for landfill site selection: A case study in the Lake Beys_Ehircatchment area (Konya, Turkey)", Waste Management journal, Vol. 30, 2010.
[31] S. Sener, E. Sener and R. Karagüzel, "Solid waste disposal site selection with GIS and AHP methodology: a case study in Senirkent–Uluborlu (Isparta) Basin, Turkey", Environ Monit Assess journal, 173, PP. 533-554, 2011.
[32] M. Ekmekciog˘lu, T. Kaya and C. Kahraman, "Fuzzy multicriteria disposal method and site selection for municipal solid waste", Waste Management journal Vol. 30, PP. 1729-1736, 2010.
[33] C.-C. Chou and K.-W. Yu , "Application of a New Hybrid Fuzzy AHP Model to the Location Choice", Hindawi Publishing Corporation Mathematical Problems in Engineering journal, Vol. 2013, Article ID 592138, 12 pages, 2013.
[34] T.-Y. Choua, C.-L. Hsu, and M.-C. Chen, "A fuzzy multi-criteria decision model for international tourist hotels location selection", International Journal of Hospitality Management, Vol. 27, PP. 293-301, 2008.
[35] M. Z. Naghadehi, R. Mikaeil and M. Ataei, "The application of fuzzy analytic hierarchy process (FAHP) approach to selection of optimum underground mining method for Jajarm Bauxite Mine, Iran", Expert Systems with Applications, Vol. 36, PP. 8218-8226, 2015.
[36] O. Boutkhoum, M. Hanine, T. Agouti and A. Tikniouine, "An improved hybrid multi-criteria multidimensional model for strategic industrial location selection: Casablanca industrial zones as a case study". Springer Plus, 2015.
[37] W. Ho, "Integrated analytic hierarchy process and its applications – A literature review", European Journal of Operational Research, Vol. 186, PP. 211-228, 2008.
[38] R. Ramanathan, "A note on the use of the analytic hierarchy process for environmental impact", Assessment Journal of Environmental Management, Vol.63, PP. 27–35, 2001.
[39] Y.-C. Tang and M. J. Beynon, "Application and Development of a Fuzzy Analytic Hierarchy Process within a Capital Investment Study", Journal of Economics and Management, Vol. 1, No. 2, PP. 207-230, 2005.
[40] S. Greco, B. Matarazzo and R. Słowiński, "Decision rule approach", New York: Springer-Verlag, 2005, ch. 13 pp. 507-562.
[41] K. Frakling, "Modélisation et apprentissage des préférences par réseaux de neurones pour l'aide à la décision multicritère". Thèse doctorat, Ecole nationale des mines de Saint-Etienne, 1996.
[42] A.H.I. Lee, W.-C. Chen and C.-J. Chang, "A fuzzy AHP and BSC approach for evaluating performance of IT department in the manufacturing industry in Taiwan", Expert Systems with Applications, Vol.34, PP.96-107, 2008 .
[43] Guisseppi, A. and Forgionne, R. R., "Effective DMSS Guidance for Financial Investing", International Journal of Decision Support System Technology,(IJDSST), Vol.1, No. 1, PP. 1-14, 2009.
[44] ANIREF (Agence Nationale d'Intermédiation et de REgulation Foncière), "Etude preliminaire d'amenagement du parc industriel", technical report, 2013, Alger, Algerie.
[45] A. Taibi, B. Atmani, "Multicriteria Decision Aided System for ranking industrial zones (RPRO4SIGZI)", in pro, The Fourth International Conference on Database and Data Mining (DBDM 2016). Dubai, UAE, April 23~24, 2016.
[46] Atmani, B., and Beldjilali, B., "Knowledge Discovery in Database: Induction Graph and Cellular Automaton", Computing and Informatics Journal, Vol.26, No. 2, PP. 171-197, 2007.
[47] Benfriha, H., Barigou, F. and Atmani, B., "A text categorization framework based on concept lattice and cellular automata", Int. J. of Data Science, Vol. 1, No.3, pp. 227 –246, 2016.

Taibi Aissa

He is a computer science teacher at the Faculty of Social Science of Oran2 University (Algeria). He earned his Master of Science degree in 2002 from Oran1 Ahmed Benbella University. He is currently a Ph.D. candidate in the Computer Science Department at the same university. His research interests focus on Multi-criteria decision analysis and geographic information systems, Expert Systems, social data analysis and fuzzy reasoning.

Baghdad Atmani

He is a Professor of Computer Science at the University of Oran 1 Ahmed Benbella. His field of interests is Data Mining and Machine Learning Tools. His research is based on Knowledge Representation, Knowledge-based Systems and CBR, Data and Information Integration and Modelling, Data Mining Algorithms, Expert Systems and Decision Support Systems. His research is guided and evaluated through various applications in the field of control systems, scheduling, production, maintenance information retrieval, simulation, data integration and spatial data mining.

Supporting Multi-agent Coordination and Computational Collective Intelligence in Enterprise 2.0 Platform

Seddik Reguieg* and Noria Taghezout

LIO Laboratory, Dept. of Computer Science, University of Oran1, Ahmed Ben Bella, BP 1524, El Mnaouer, Oran (Algeria)

Received 24 January 2017 | Accepted 24 June 2017 | Published 18 July 2017

Abstract

In this paper, we propose a novel approach utilizing a professional Social network (Pro Social Network) and a new coordination protocol (CordiNet). Our motivation behind this article is to convince Small and Medium Enterprises managers that current organizations have chosen to use Enterprise 2.0 tools because these latter have demonstrated remarkable innovation as well as successful collaboration and collective intelligence. The particularity of our work is that is allows employer to share diagnosis and fault repair procedures on the basis of some modeling agents. In fact, each enterprise is represented by a container of agents to ensure a secured and confidential information exchange between intra employers, and a central main container to connect all enterprises' containers for a social information exchange. Enterprise's container consists of a Checker Enterprise Agent (ChEA), a Coordinator Enterprise Agent (CoEA) and a Search Enterprise Agent (SeEA). Whereas the central main container comprises its proper agents such as Selection Agent (SA), and a Supervisor Agent (SuA). JADE platform is used to allow agents to communicate and collaborate. The FIPA-ACL performatives have been extended for this purpose. We conduct some experiments to demonstrate the feasibility of our approach.

Keywords

Collaboration, Collective Intelligence, Coordination Protocol (CordiNet), Enterprise 2.0, Pro Social Network, Multi-Agent System (MAS).

DOI: 10.9781/ijimai.2017.07.001

I. Introduction

Recently, an ecosystem of employees, partners, suppliers, and customers has begun to proliferate by using web 2.0 technologies, to develop capabilities by collectively generate, share and refine business knowledge. This ecosystem is named as an enterprise 2.0 collaboration platform [1].

These kinds of platforms provide enterprises with new models and tools for an emergent collaboration and co-creation to assure and to harness collective intelligence [2], [3], [4].

In addition, the mobilization of collective intelligence in enterprises 2.0 platforms encourages teamwork and knowledge sharing. The adoption of web 2.0 technologies as new collaborative management practices allows turning the emphasis on the development of human capital such as the knowledge, experience and skills of members and on the development of social capital – like relationships, communities and norms that facilitate collective action and collaborative behavior.

To be effective, our work covers a wide range of goals and tasks in a more sophisticated way to lead the world and the current industry to a new win-win spirit as an agreement whereby each enterprise is also concerned with the interests of the other in an equally favorable to its own interests. In other words, it is an agreement that increases the earnings of each.

Enterprise 2.0 supports the informal organization as a set of technologies i.e. social networks [5]. These social networks are combined with other resources as areas of collaboration, connection and communication to improve the organization of knowledge, skills of users and encouraging collaboration and innovation [6], [7]. Furthermore, the ability of a social group to develop its human capital has become a brand of collective intelligence within the meaning of [8], [9]: distributed intelligence everywhere, constantly valued, coordinated in a real time, resulted in a "full" mobilization of skills.

Thus, according to the study that is provided in [10], a social network is defined as a set of relations of a specific type (e.g., collaboration, support, consulting, control, or influence) between a set of actors. Each actor in the network is required to create links with other actors; analysis of these links can predict the characteristics of the actors or the appearance of links between them.

Furthermore, social network can play an important role in supporting different types of decision making, as they provide their users various forms of support, ranging from instrumental to the emotional and informational. The synergy of these themes provides an innovative and unique perspective on the actual process of decision-making within social network.

An effective decision-making is often based on the control of information over time. This action aims to capture the current state of the most important pieces of information that are updated in different ways. In Enterprise 2.0 platform, employers are looking for the right information at the right time under confidentiality and security. However, employers don't have all necessary information such as diagnosis or failure solution procedures. This information can be found intra enterprise (coordination), or inter enterprise (collaboration). To ensure exchange knowledge intra enterprise we must first ensure coordination.

* Corresponding author.
E-mail address: rseddiko@gmail.com

The coordination can be expressed by different techniques, which are summarized in: coordination languages, coordination algorithms, and coordination protocols. Several works have appeared in the field of coordination [11], [12], [13], which propose solutions that can provide consistency in production systems. Among those works, we can cite [14]; this approach is a centralized one. It is composed of a single agent which is responsible of the coordination system. This is on the one hand but on the other, several architectures that are [15] based on the distributed aspect were developed in order to provide a large number of interactions.

This work contributes within several principles and technicalities to build an enterprise 2.0 platform and to achieve collective intelligence via information sharing among trusted contacts in several enterprises.

Our platform consists on two levels: a collaborative environment, that allows users to exchange knowledge and information inter employers in different enterprises using their relations in social network, and coordination environment for exchanging knowledge and information intra enterprise which is based on a multi-agent system and coordination protocol. The motivation of our platform's components is to make an efficient decision by offering effective diagnosis and failure solution procedures with a secured and confidential support.

In this article, we focus on the study of two basic components of the framework: the professional social network (Pro Social Network) and the coordination protocol (CordiNet). Pro Social Network is dedicated to companies that want to share their knowledge and expertise on the industrial diagnosis. CordiNet represents a coordination protocol that manages interactions between agents.

After a brief introduction that describes the context. Section 2 of this paper discusses some approaches of enterprise 2.0, collective intelligence, and multi-agents systems, emphasizing their strengths and limitations. The proposed approach is being presented in section 3 of the article. Herein, Pro Social Network as well as CordiNet are described. In Section 4, the most important performatives of the protocol are provided with some experiments that have been conducted with the proposed protocol. Finally, a conclusion regarding the research works as well as future works are mentioned in Section 5.

II. Background

This section presents an overview of related work (social networks, multi-agents systems and enterprise 2.0) and presents our major contribution as a continuation of this background.

A. Social Networks and Multi-agents Systems

Many studies have integrated agents in social networks in different ways. In this section, we concentrate on the most relevant works that influence our research.

Authors in their literature review on [16] internet based collaboration tools have mentioned that online collaboration tools are divided into 2 classes; the first one regrouped some tools that are based on internet-based technologies while the second class comprised the tools that are based on Web 2.0 technologies. These latter provides a support for a beneficial collaboration. Some well-known examples of these communication tools are wikis, blogs, forums, RSS feeds, community chats and social networking. Collaboration 2.0 has been initiated by the use of Web 2.0 tools. This new form of collaboration allows a robust real time collaboration between dispersed participants as it was well explained in the work [17].

More precisely, in the work that was presented in [17], the domain experts were given a collaborative WEB interface in order to save and capitalize their business knowledge as business rules through an agent-based platform. Furthermore, domain ontology has been developed in order to generate the business model corresponding to the enterprise and check the consistency of business rules. The modeling was based on some agents such as Expert agent, Evaluator agent, Translator and Security agents. The main objective was to increase the execution speed of processes and effective response [17].

Among the early works on software agents supporting social networking system, the papers [18], [19] and [20] are particularly relevant.

In paper [18], author presented a multi-agent system that implemented a distributed social network system. By using such system, users are considered as the sole owners of the information they provide when they address privacy. The main particularity of this system is that users are represented by agents that both mediate access to private data and proactively negotiate with other agents in order to extend their user's social network. A distributed connection discovery algorithm is also presented. This algorithm is developed to suggest connections to the users for constructing a social network through the information that is stored in their profile. This constitutes the main advantage of this approach. The major limitation is that the privacy is not guaranteed by only the fact that the users can specify which data shall be used in order to rich their social network especially when we deal with friends of friends.

The study that was provided in [20] is quite different from other works which supported the social networking. This is due to the key points that were addressed in the corresponding paper. Our deeper review of the content gives us a precise point of view of the advantages of the use of enterprise 2.0 in project management such as the sharing of information and expertise between operators and senior managers and time flexibility.

Furthermore, this research offers a good list of examples of well-known companies that opened their road to success by promoting the web 2.0 model in their projects and activities related to their organization.

In the work presented in [19], a simple framework has been proposed to help understand the collaboration that was offered by Web 2.0 technologies.

The study was based on the development of a Web-based Decision Support System (DSS) which included five agents: Analyzer agent, Proposal agent, Resource agent, production agent and Coordinator agent. These latter were used to examine how human participants can create, share and exchange experiences or diagnosis on resources failures which each other to have new ideas or useful information for the decision-making process. The limitation of this work consists in improving the communication and collaboration between the human operators and integrating a case base reasoning to take benefit from past experiences.

B. Research on Enterprise 2.0

As defined in [20] [21], Enterprise 2.0 or E2.0 is represented as a set of organizational and technological approaches to implement new organizational models that are based on mass participation, technology-optimized collaboration, knowledge sharing, and exploitation of professional social networks.

From our review on enterprise 2.0 notion and its application in industrial domain, we find that this technology has encountered various difficulties like any other technological innovation (technological, organizational, cultural, social...) [22].

It seems that the more imminent challenge is the compatibility of Web 2.0 with the enterprise culture and existing technological infrastructure more precisely with the manager's attitude. As well explained in [20], the acquisition of Web 2.0 tools is not enough to say that a company is an Enterprise 2.0. Corresponding managers and employers need to be motivated. The adoption of Enterprise 2.0 concept is conditioned by culture and certain attitudes.

The review of the current proposals in the literature shows that the work presented in Schauer et al. [23] offers very interesting features; Authors introduced open-source enterprise 2.0 tools to support knowledge workers in the execution of cognitive tasks and information sharing. They also presented a study of the growing market for enterprise 2.0 systems using an open source license. In their study, they used a set of 97 individual features and criteria that are grouped around the central functionalities of communication, coordination and collaboration, in order to analyze the suitability of a representative sample for the average knowledge worker.

Even more, among the basic functionalities of enterprise 2.0 platforms, we cite electronic collaboration. In this context, a new study was described in [24], which can be considered as a real contribution for a better understanding by providing an overview of the market for E-Collaboration software packages. Based on the classification that was done in [24], five system classes have been identified to characterize the market for Collaboration systems.

The first identified class regrouped Everyday systems which are used continuously by the group as primary systems to support everyday activities. Such systems provide basic features to support all three processes of communication, coordination and collaboration.

Integrated systems were considered as the second class of systems that share many features (e.g. E-Mail, calendars, address books and task lists) with everyday systems, but their constituent characteristic is the support of synchronous communication.

Another class of coordination systems has been identified. As mentioned in [24], this class comprises a total of 24 systems which focus on group coordination. Particularly, task coordination systems and process coordination systems are constituents of this class.

A set of 13 systems are grouped in a class named meeting systems that support online and video conferences. The systems are based on extensive synchronous communication featuring such as text chat, audio and video communication.

Regarding these 4 classes, we can say that our approach belongs to the subclass that regroups cases converging from one class to another, such as the systems being at the convergence of the coordination systems to the integrated systems.

As clearly explained in [24], the systems integrating project features are quite similar to coordination ones. The collaboration and coordination processes are supported by the coordination systems, whereas the integrated systems add the communication aspect that we find in the WEB 2.0 technologies integrated in our project. In addition to their project management capabilities, integrated systems as we see them are systems that provide a set of features ranging from a simple to knowledge management functionality regarding to the organization of an enterprise. This category of systems easily adheres to the use of WEB 2.0 tools augmented by a set of additive project coordination functions.

Whereas Xu et al. [25] presented a survey on asynchronous collaboration tools that allow users to collaborate at different times. They structured their features in four major functional categories such as: communication, information sharing, electronic calendar and project management.

Furthermore, authors in [26] advocated the idea that enterprise collaboration must evolve towards a new paradigm in which knowledge workers are considered as co-producers of information, software services and applications without involving IT departments. Their prototype named EzWeb consisted on an implementation of an open Enterprise 2.0 collaboration platform that empowered its users to co-produce and share instant application.

In the study that was provided in [27], authors proposed a model of Enterprise 2.0 as a mash up developer which is supported in searching for assistance from developers owning specific knowledge, according to typical collaboration patterns. Their study focused on collaborative development in the Enterprise 2.0 contexts which include Web APIs for sharing Web sites, information about developers based on a specialization of ontology and relationships among developers and Web APIs.

In other works such as [35], authors were interested by sharing and managing knowledge in enterprise 2.0 platform. In this paper, author presented a model integrating the interpersonal network in enterprise 2.0. They took the knowledge governance as a regulator variable between enterprise knowledge system and its environment, a relationship between the performance of knowledge collaboration and structure entropy has been also built. In their paper, they constructed the self-organization concept model and logistic process analysis model of knowledge system in enterprise 2.0 environment, which are to explore the evolution pattern of knowledge collaboration system in favor of the organization mechanism design for enterprise knowledge governance. However, in enterprise 2.0 platform, sharing knowledge must be done under confidentiality and security which is not describe in their work. Authors mentioned also that their work lacks presumption appropriate control parameters and formulation of governance structure and strategies.

C. Our Contribution and Comparison with other Works

In current small and medium enterprises that exist in our country, collaboration became a very important way to proliferate or succeed. These Organizations used to ensure a lot of activities around blogs and wikis with a focus on collaboration and sharing of knowledge. Managers are usually encouraging employees to be connected with Facebook or LinkedIn social network but they don't feel ready to consider a project based on Enterprise 2.0.

Our motivation behind this work is to convince senior managers that Enterprise 2.0 is a real solution which allows exploiting the WEB 2.0 technologies.

In collaborative platforms or social networks, Web servers are constantly overloaded with client requests. An Enterprise 2.0 platform is essentially based on a web server that provides several services namely: internal tools and applications such as online ERP, monitoring, collaboration and coordination applications. As a result, services are multiplying in response to the number of companies and users who are members benefiting from all the functionalities offered by this platform.

To simplify the work of the web server, we propose a multi-agent system in order to perform complex tasks while ensuring the security and integrity of each company's data. By sharing tasks with the web server, which will only be dedicated to displaying and processing replies to users' requests, the multi-agent system takes over and searches for information in each database of each company, in an invisible way to the customer.

The originality of our approach lies in the suggested coordination protocol (CordiNet) which regroups a set of primitives that are deployed by realizing extensions of the performatives of the FIPA-ACL.

Moreover, the Supervisor agent (SuA) of the central container plays a very important role; it makes it possible to efficiently manage the work of every coordinator agent of the enterprise. The search for information in our platform is supported by the selection agent (SA) which executes a search orientation algorithm according to the semantics of the request and its priority in the system.

In Table 1, we present a comparison between some related works and ours. We just sort the items by using '+' to indicate that the option exists and '-' to indicate that the option doesn't exist. We put our approach in the bottom of Table 1. Some indications are given with the Table1 in order to clarify some criteria that are used in the comparison.

A: Enterprise 2.0

B: Social network

C: Coordination

D: Multi-Agents System

E: Interaction Protocol

F: Ontology

G: Learning Aspect

H: Information retrieval mechanism

I: Recommendation System

J: Knowledge sharing

As we can see in Table 1, no work has introduced the concept of multi-agents and protocols in enterprise 2.0 platform.

TABLE I. COMPARISON BETWEEN SOME RELATED WORKS AND OUR APPROACH

Works[1]	A	B	C	D	E	F	G	H	I	J
1	+	+	+	-	-	-	-	-	-	+
2	-	+	+	-	-	-	-	-	-	+
3	-	+	+	-	-	-	-	+	-	+
4	+	+	-	-	-	+	-	-	+	+
5	+	+	-	-	-	+	-	+	-	+
6	+	-	+	-	-	-	-		-	+
7	+	+	+	+	+	-	-	+	+	+

[1]Works: 1: Schauer et al. [23], 2: K. Riemer [24], 3: Xu et al. [25], 4: J. Soriano et al. [26], 5: D. Bianchini et al. [27], 6: X. Yun et al. [35]; 7: Our work.

Our contribution is to integrate multi-agents system into enterprise 2.0 platforms and more precisely into its social network.

An Enterprise 2.0 platform consists of several subsystems and components that are all connected to a Web server. The increase in the number of companies adhering to our platform implies an increase in the number of users; this could either cause a system crash or generate no response because the Web server cannot handle all requests (multiple users, multiple components). To avoid this, we have introduced a multi-agent system and an interaction protocol to perform a distributed processing which does not necessarily includes the web server, such as the search for information (in our case, the search for Industrial diagnostics).

In the next section, we describe and motivate the problem on which we will focus in the development of a collaborative platform that also supports the coordination of its agents.

III. PROPOSED APPROACH

The business model of enterprises has changed from an old one which is based on commands and information management to a new one which integrates a new paradigm of collaboration that is based on the WEB 2.0 technologies. In Our approach, we find two different types of enterprises information flows namely: external flow which will describe collaboration environment and internal flow for coordination environment. Figure 1 presents our platform that consists on several spaces which allow enterprises to collaborate, coordinate and control their internal or external flows. In this paper, we will focus and develop the collaboration and coordination space.

A. Pro Social Network Design

In the first environment, we implement a social network as a support to provide companies of new models and tools for collaboration, participation and co-creation.

This collaboration is established between the various participants such as company employees, customers, suppliers, experts with the ability to create groups that can contain these profiles.

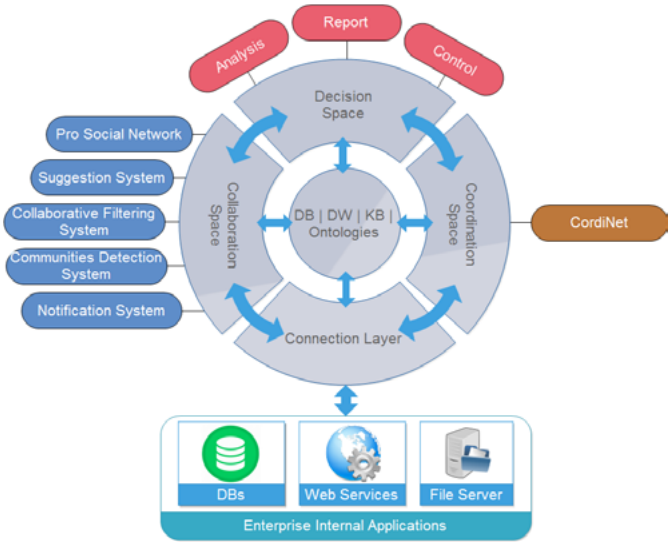


Fig. 1. Architecture of our Enterprise 2.0 platform.

The profile is given throughout an interface that represents a user, group or company on our network. This latter corresponds to an identity that is established on the network. It may be a reflection of a real or virtual identity. The virtual identity means that any user can create an account, in this case user does not belong to a company which is adhered to our platform, and we cannot identify his real identity, it is said Freelancer while the real identity means that the account is given to users who belong to a company that adheres to our platform, an identity verification must be done in the administrative space of this company, the HRD account validates the creation of this account.

Our collaborative environment ensures the allocation of a digital professional identity to different profiles. It allows the establishment of the collaboration while enriching the relationship between the different profiles; the main objective behind this is to maintain the continuity of the professional links.

To fully enhance the collaborative environment, we have introduced a feature that allows the user to adopt good practices during knowledge sharing and participation in collaborative work.

Figure 2 presents an example of exchanging information between actors. For example, let's us consider a simple user named as User X, this latter is in relationship with some other users of the same or different enterprise. Thus, he can share information on his wall, enterprise wall or his groups' wall.

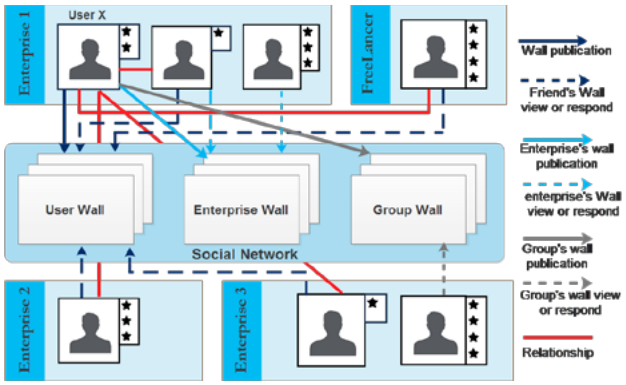


Fig. 2. Information's exchange in our social network.

Our social network disposes some features for enterprise management and users like: profiles, expert base, relationship

management, flow of activities and foremost, asynchronous and synchronous communication.

In what follows, we present the general structure and features that characterize a user's account (Figure 3).

Fig. 3. User's account features.

More precisely, to access our platform, we provide two ways ; (i) by creating an account and enriching it, (ii) by using APIs that are offered by social networks (Google+, Facebook, etc) in order to retrieve the existing profiles from these networks.

Our collaborative environment allows participants to share and exchange diagnosis under confidentiality and security. Each enterprise can manage its employers in our platform, and each user can share diagnosis with its collaborators (different enterprise) or colleagues (same enterprise). By using a catalog, this latter can set, update and delete diagnosis and failure solution procedures in an easy way.

To give more confidentiality for user intervention on the use of catalog, many options are given as below:

1. Me only: It means that user is the only one who can see this publication
2. Colleague: It means that a colleague who is in relationship with this user can get this diagnosis.
3. Collaborator: it means that collaborators who are in relationship with this user can get this diagnosis.
4. Freelancers: it means that the freelancers who are in relationship with this user can get this diagnosis
5. Open: all users that are connected to the platform can get this diagnosis.

User can use a customized type to combine between confidentialities in case of need (except for Me only and Open) eg. Confidentiality = colleague and collaborators, Confidentiality = colleagues and freelancers, Confidentiality = collaborators and freelancers.

When user searches an industrial diagnosis of any machine, the result must respect diagnosis' confidentiality that is established by their editors.

As it can be seen a lot information are exchanged during the treatments, it's why we opt to use agents as modeling entities in order to control the whole system and we develop a coordination protocol for this task.

B. CordiNet Design

In the second environment, we want to set up a coordination environment (CordiNet) to exchange knowledge and information between employers of the same enterprise. The goal is to make the system more secure and confidence, reactive and produce a quick execution. In this section we present our coordination environment as a multi-agent system and the corresponding coordination protocol.

Our approach uses reactive and cognitive agents which need an elaborated language to be able to exchange messages. However, a structural format of a message is not sufficient in itself to be able to formalize all the conversations between agents. The protocol concept was introduced to support such conversations, allowing speech acts. We can define coordination protocols as shared conventions on messages that are exchanged by agents working together in a coordinated manner. More protocol is effective; less information needs to be transmitted and less time is spent in the communication as mentioned in [28].

1. Agent-based Modeling

The use of a multi-agents system that is opened in an application for Internet offers three main advantages: scalability, stability and load balancing.

The advantages of Multi-agents system (MAS) such as autonomy, adaptability, knowledge level, learning aspects and availability of interaction protocols have encouraged us to adopt them in our platform to balance overload the web server [29].

Our multi-agent architecture is developed as a blackboard behind the Web server. The proposed model mainly consists of two types of containers; enterprise container which controls internal enterprise flow and central main container which controls external flow of our collaborative environment (see Figure 4).

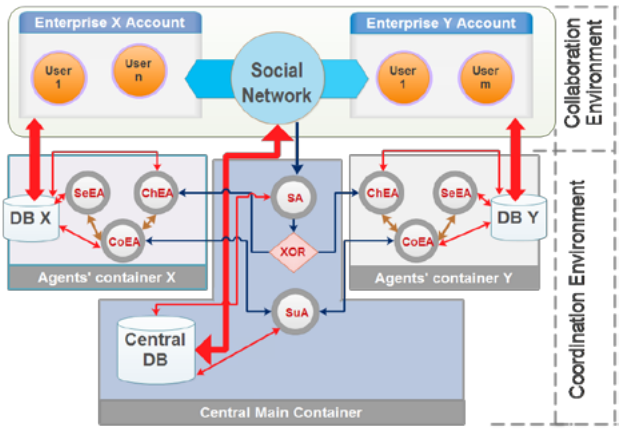


Fig. 4. Architecture of our Multi-agent system to support our social network.

Each enterprise that is connected to our platform is represented by a container. In fact, every enterprise Container consists of three components: Checker Enterprise Agent (ChEA), Search Enterprise Agent (SeEA), and Coordinator Agent Enterprise (CoEA). The central main container consists of two components: Selection Agent (SA), and Supervisor Agent (SuA).

We have also developed a coordination protocol ensuring the exchange of messages between internal flow and external flow. The SuA has a learning aspect. This mechanism facilitates decision-making in the system and requires less execution time. Coordination in multi-agent systems can be performed in a centralized manner in which circulation of information between enterprises is assured by the central main container's agents.

As we have previously described, we had used two types of container: enterprise's container and central main container.

1. Enterprise Container: : It is a container that regroups agents that control internal flows in a given enterprise, when an enterprise adheres to our platform then enterprise's container's agents are created by Supervisor Agent (SuA) of central container. These

agents are described as follows:

- Checker Enterprise Agent (ChEA): It is a reactive agent which detects if a new diagnosis is required by enterprise user.
- Search Enterprise Agent (SeEA): It is a cognitive agent which searches if diagnosis exists in enterprise's database and returns it.
- Coordinator Enterprise Agent (CoEA): It is a cognitive agent which works with the Central Coordinator Agent in order to receive diagnosis' request and processing. It verifies confidentialities and undesired messages. The internal architecture of CoEA is illustrated in Figure 5.

2. Central Main Container: includes 2 agents that control external flows coming from different enterprises such as exchanging of diagnosis namely:
   - Selection Agent (SA): It is a reactive agent which selects and directs the diagnosis' request to the corresponding agent enterprise according to any user of our social network.
   - Supervisor Agent (SuA): It is a cognitive agent which controls enterprises' containers, by creating and supervising certificated agents. In addition of its specific role of supervision, this agent executes some tasks such as verification of confidentiality, detection and filtering of undesired messages and control of all messages. The internal architecture of SuA is illustrated in figure 6.

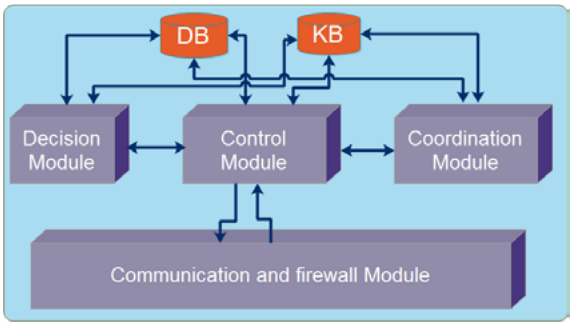


Fig. 5. Internal architecture of COordinator Enterprise Agent (CoEA).

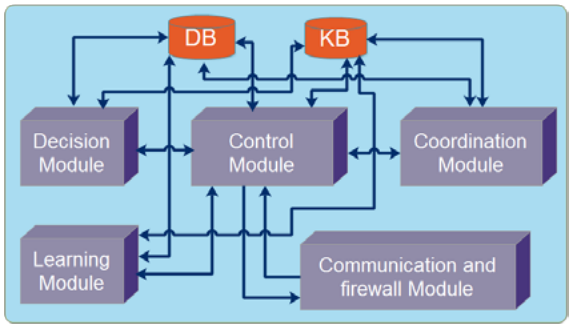


Fig. 6. Internal architecture of Supervisor Agent (SuA).

2. The Proposed Protocol

An agent sends a message in two cases: (i) When receiving a message that needs an answer; and (ii) When executing an action and needs information. Each message sent of our protocol is designed using the language communication FIPA-ACL [30].

In our coordination protocol, each type of agent has some behavior classes that are ready to use. Classes contain a set of methods in order to monitor each step of the protocol. They are called when specific messages are received or are to be sent and must be treated in order to adapt the protocol to its context of use. In what follows, we give a general description of our coordination protocol.

1. The sending of messages: An agent sends a message in both cases: Following receipt of a message that requires a response or when performing an action.
2. Structure of a message that is sent: Each message which is sent in our protocol is designed using the communication language FIPA-ACL [30]. In our MAS, each message has the following syntax:
   - Content Object: Content of message.
   - Protocol: Protocol used to exchange ACL messages.
   - Conversation-id: Conversation identifier (expression) that is used for the following acts of communication.
   - Reply-with: The term used by the agent as to identify this message.
   - Receiver: the name of the agent who receives the message.
   - Sender: the name of the agent who sends the message.
3. Secure communication: In a replay attack gained knowledge of data content is used to modify data that has been transported before to acquire unauthorized information. In case of a spoofing and masquerading attack, we use certificated messages where each agent has a communication and firewall module. Firewall module helps to eliminate undesirable messages by identifying a digital signature.
4. The rules: The rules define the possible message exchange between agents under certain conditions. This representation is taken from [31] and appears to be most suitable for the description of our protocol due to its ease of writing and reading, and for its rigorous notation that facilitates coding. The upper part of the rule 1 defines the sending of a message from X to Y. The lower part defines actions (a1 ... an) that can be executed by Y, after that, Y sends a response to Z if the condition (C) which is expressed at the right of the rule is verified (Z is optionally X).

$$NameofRule \frac{X \rightarrow Message \rightarrow Y}{A1, A2, \ldots, An - \mathrm{Re}sponse \rightarrow Z} C \quad (1)$$

5. Classes of our protocol: We distinguish 3 classes that composed our protocol in order to manage different situations for each agent. Some of the classes of our protocol represent finite state machines, in fact with state and transitions.
   - panneUserInitiator Class : It is used by CoEA when there is a search of a diagnosis launched by user. Figure 7 shows the finite state machine of this class and Table 2 gives its description.

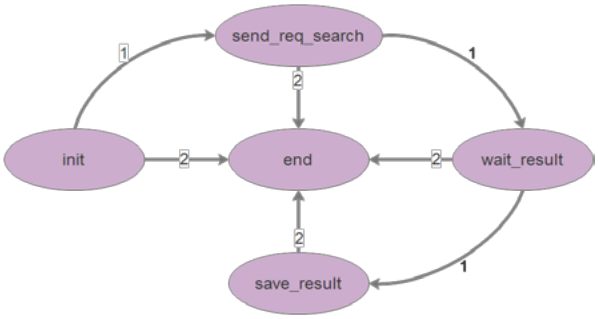


Fig. 7. Finite state machine of panneUserInitiator.

TABLE II. PANNEUSERINITTIATOR'S STATES AND TRANSITIONS DESCRIPTIONS

State[1]	Description	Transition
1	Initialization of protocol	1: Preparing information search 2: Error of initialization
2	Sending information search	1: Information search request send 2:Error of sending
3	Waiting result of the searched information	1: Result received with success 2: Error of waiting
4	Saving result and forwarding feedback	1 : Operation succeed 2 : Error of saving or forwarding
5	End of protocol	None

[1]States: 1: init, 2: send_req_search, 3: wait_result, 4: save_result, 5:end.

- panneUserResponder Class: It is used by CoEA to send the diagnosis after receiving request from panneUserInitiator. This class replaces the user behavior while respecting its confidentialities put in place.
- panneCoordinatorCentral Class: It is used by Supervisor Agent (SuA) for receiving a diagnosis's request, collecting available CoEA, sorting CoEA list by learning mechanism, sending diagnosis requests to list of CoEA, receiving diagnosis and solutions. Figure 8 shows the finite state machine of this class while Table 3 gives its description.

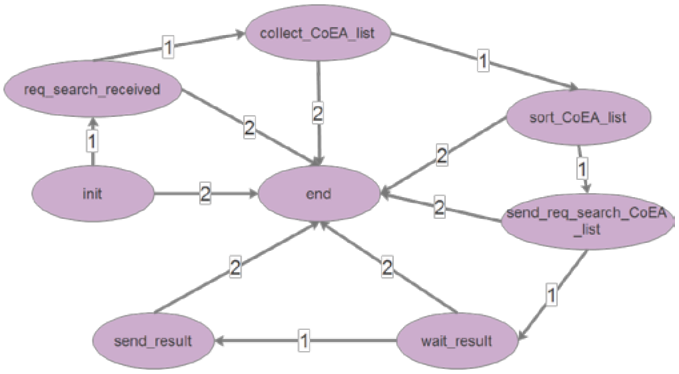


Fig. 8. Finite state machine of panneCoordinatorCentral.

TABLE III.
PANNE COORDINATOR CENTRAL'S STATES AND TRANSITIONS DESCRIPTIONS

State[1]	Description	Transition
1	Initialization of protocol	1: Initialization request waiting 2: Error of initialization
2	Waiting information search request	1: Waiting of request information search 2:Error of waiting
3	Collecting list of active CoEA	1: Success of collect 2: Error of collect
4	Sorting CoEA list with search strategy max uniform cost	1 : Establishing search strategy to sort CoEA list to get optimal solution 2 : Error strategy or sorting list
5	Sending information search to CoEA list	1: Send information search request to CoEA succeed 2: Error sending
6	Waiting results	1: Waiting until results received 2: Error of waiting
7	Sending results to CoEA	1: Sending results succeed 2: Error of sending results
8	End	None

[1]State: 1: init, 2: req_search_received, 3: collect_CoEA_list, 4: sort_CoEA_list, 5: send_req_rearch_CoEA_list, 6: wait_result, 7: send_result, 8: end.

3. Communication between Agents

Finding diagnosis and failures solutions bring us to information seeking context. This notion is an important part of our framework. To describe how diagnoses are founded, AUML sequence diagram (Figure 9) and UML sequence diagram (Figure 10) are given.

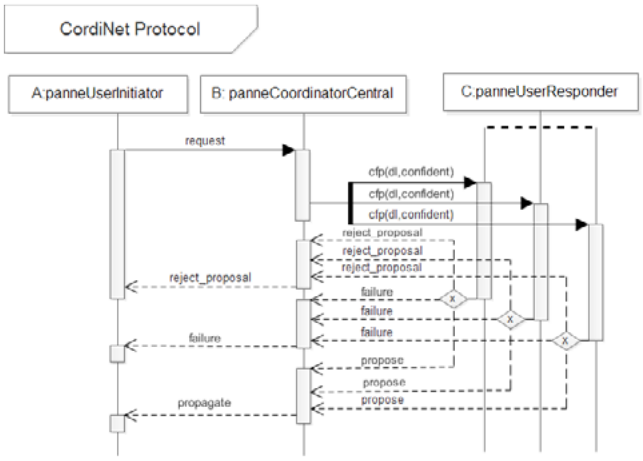


Fig. 9. AUML sequence diagram of CordiNet Protocol.

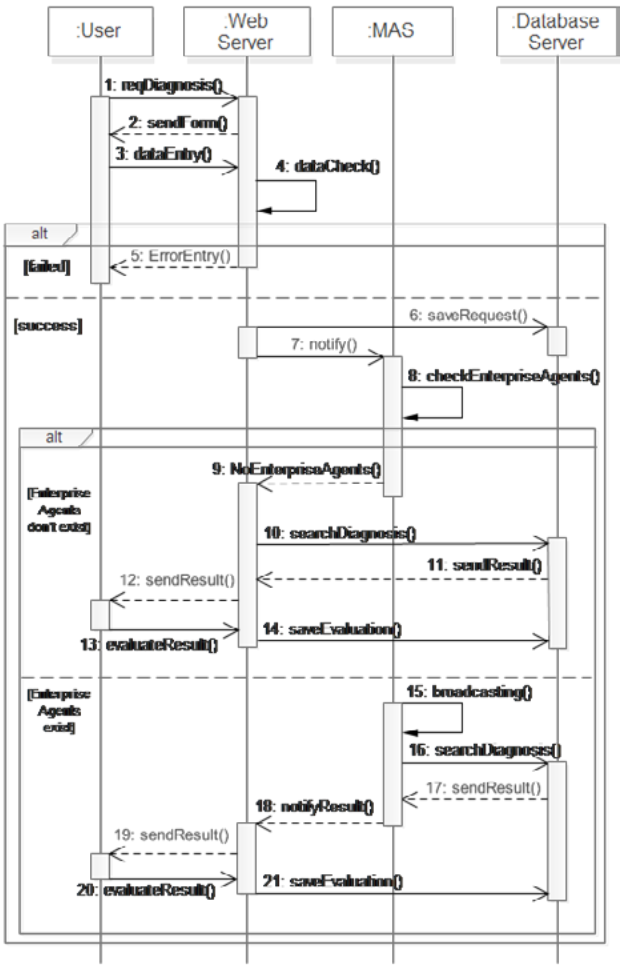


Fig. 10. Form for publishing and adding a diagnosis to a catalog.

In the sequence diagram that is presented in Figure 9, we distinguish three actors: a panneUserInitiator which corresponds to the initiator agent (CoEA of the applicant user enterprise), a panneCoordinatorCentral which corresponds to the Supervisor agent

of main container (SuA), and panneUserResponder which corresponds to all participants agents (all CoEA of all enterprises).

After the reception of a request from the panneUserInitiator to panneCoordinatorCentral, a cfp message is sent to all panneUserResponder with a deadline and confidentialities.

We distinguish three types of response:

- reject_proposal if time of response is out.
- propose if a solution is found then this set of solution is propagated to panneUserInitiator.
- Failure in others cases (agent error, database error, etc...).

Figure 10 shows interaction between user, webserver, multi-agents system and database server. The internal process of MAS actor is given by the previous AUML (Figure

4. General Process of Information Seeking

To describe how information is sought we adopted the description from [32] and defined information seeking task as: the act of seeking information as a consequence of a need to satisfy some goal. Furthermore, search process has been studied extensively [33], [34]. Figure 11 summarizes our process of information retrieval.

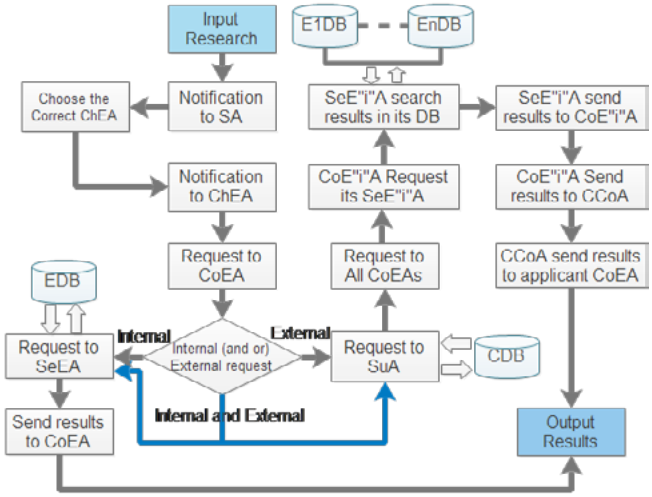


Fig. 11. General process of information retrieval.

Each company named "i" has its proper database (EiDB). When the user initiates the search (Input Research) from his own account, a notification is then sent to the agent CChA. The CoEA agent directs the request both to the external stream (to SuA) and the internal stream (to SeEA). SuA is then responsible for sending a message to the other CoEAs to initiate the search for all the necessary confidentiality information in their respective databases. The result of this search makes it possible to return solutions to the agent SuA which will propagate them to the CoEA that is identified as the requestor.

IV. Experiments

This paper describes the major experiments that are done to demonstrate the feasibility of our suggested approach. For that a simulation study was launched on the basis of Java and JADE platform for multi-agents; Html5, Css3, AJAX and Php for social network.

During the simulation, we generate some databases that are implemented using MySQL. The agents communicate using the JADE platform, which includes several predefined performatives communication. In this work, Sniffer Agent provided by JADE was employed to monitor the communication among agents on the agent platform.

We illustrated three enterprises and 30 connected employers. JADE platform shows containers created for each enterprise (sbo, ametal, inotis) and main container for central agents as it appears in figure 12.

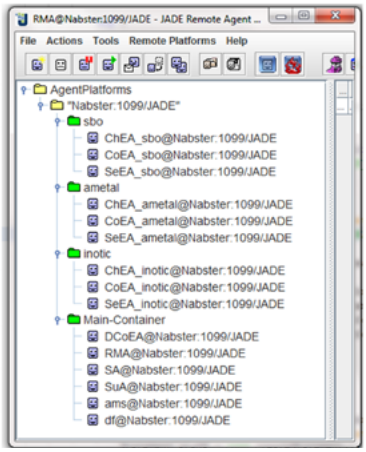


Fig. 12. JADE Platform showing main and enterprise containers.

Figure 13 shows a form which helps a simple user for adding diagnosis or failure solution to his personal catalog.

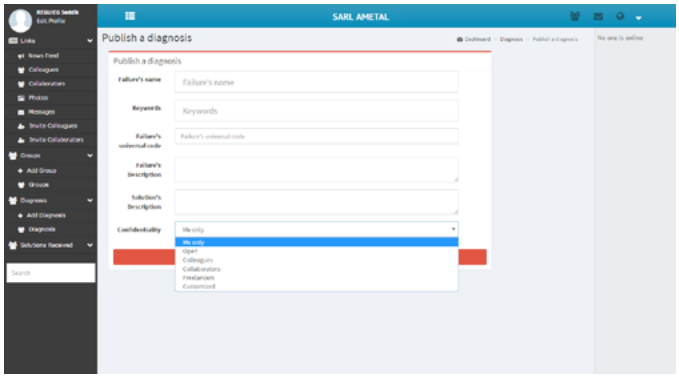


Fig. 13. Form for publishing and adding a diagnosis to a catalog.

Figure 14 gives an overview on the diagnosis' catalog for a user of an enterprise with his confidentialities.

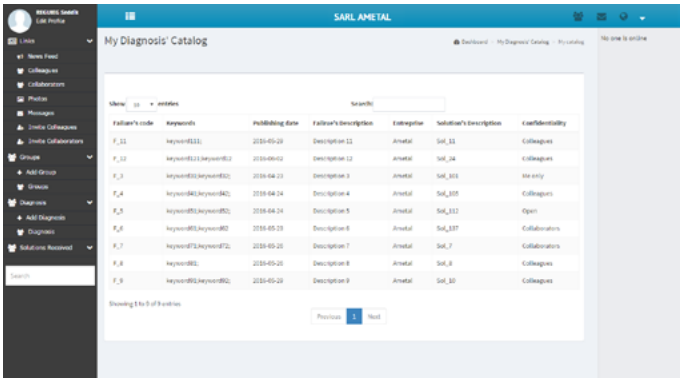


Fig. 14. User's diagnosis' catalog.

Figure 15 shows JADE's sniffer agent which describes the exchange of messages between agents from user's request until finding solutions.

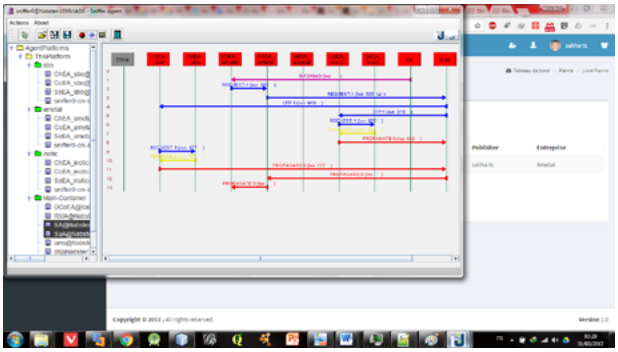


Fig. 15. Sniffer Agent for monitoring messages' exchange between agents.

Figure 16 visualizes some solutions that are received after seeking information from several users coming from different enterprises.

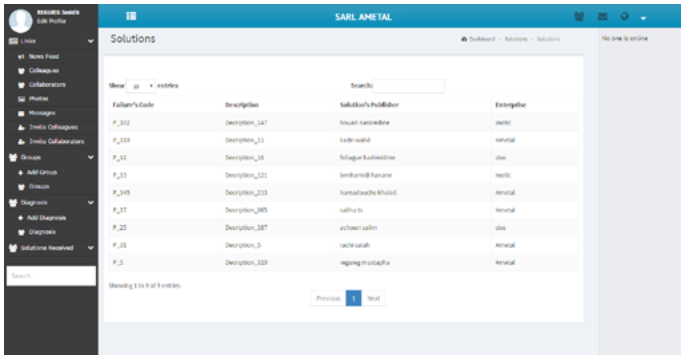


Fig. 16. Solutions received from several users.

V. Results

We plot several aspects of the results in series of figures. First, we consider an evaluation of our multi-agents system (MAS). Then, we present a simple comparison between using a webserver and its combination with MAS. Secondly we present an evaluation of our interaction protocol, and finally an evaluation of our social network is given by using a questionnaire.

A. Multi Agents and Webserver Evaluation

As we had described in previous sections, the adoption of multi agents system decreases the webserver overhead.

Table 4 illustrates the ratio between number of enterprises which adhere in our platform and number of agents created. As we see number of agents increase proportionally with number of enterprises. The messages are generated considering one user's request.

TABLE IV. Ratio between Number of Enterprises, Number of Generated Agents and Generated Messages

Number of Enterprises	Number of generated agents	Number of generated messages
2	8	9
10	32	41
50	152	201
100	302	401
200	602	801
300	902	1201
600	1802	2401

Table 5 and Figure 17 provide an indication to measure the performance of the web server by considering the contribution of MAS. For that, we had developed our approach with and without using MAS. As we notice, for a single user request, the overhead of the web server which is expressed by the number of queries that are performed using the agents keeps its stability, whatever the number of companies that adhere to our platform. Whereas if we ignore this contribution, the webserver is overloaded by all executed queries. We can explain that all other treatments of research information and respect for confidentialities are made by agents.

TABLE V. Comparison between Number of Executed Queries with/ without using MAS

	Approach 1 : simple webserver application	Approach 2: Merging webserver application and SMA	
Number of Enterprises	Number of webserver's queries without MAS	Number of MAS queries	Number of webserver's queries using MAS
2	118	13	105
10	158	53	105
50	358	253	105
100	608	503	105
200	1108	1003	105
300	1608	1503	105
600	3108	3003	105

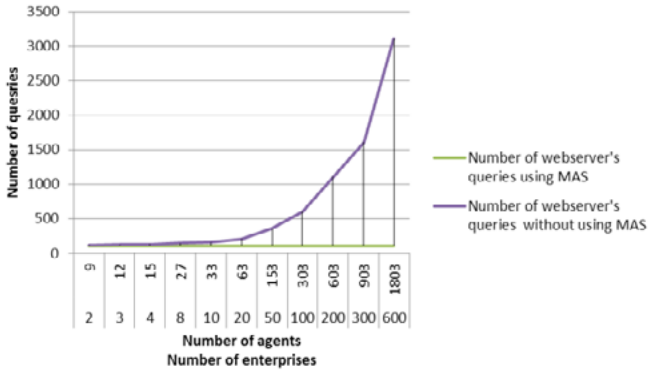


Fig. 17. Number of executed queries with/without using MAS.

B. Social Network Evaluation

An evaluation of the quality of the developed social network was carried out using a questionnaire. We developed our questions according to several criteria for the ergonomic evaluation of the multi-user interface. The questionnaire was distributed to 30 users of three enterprises.

Figure 18 shows the usability results obtained from the questionnaire, in which a question is generally followed by an inverse question to reveal opposite facts. For each question we assigned a weight. At the end, we count the sum of weights to find the level of satisfaction of the user. Some of the questions in the questionnaire are given in the following tables (Table 6, Table 7, and Table 8).

TABLE VI. Some Relevant Questions related to Accessibility

Number	Question
Q1	Is content structurally separated from navigational elements?
Q2	Is the website cross-browser compatible?
Q3	Is the website adapted for every computer resolution?
Q4	Is the URL short and simple?
Q5	Is the time of loading the home page correct?

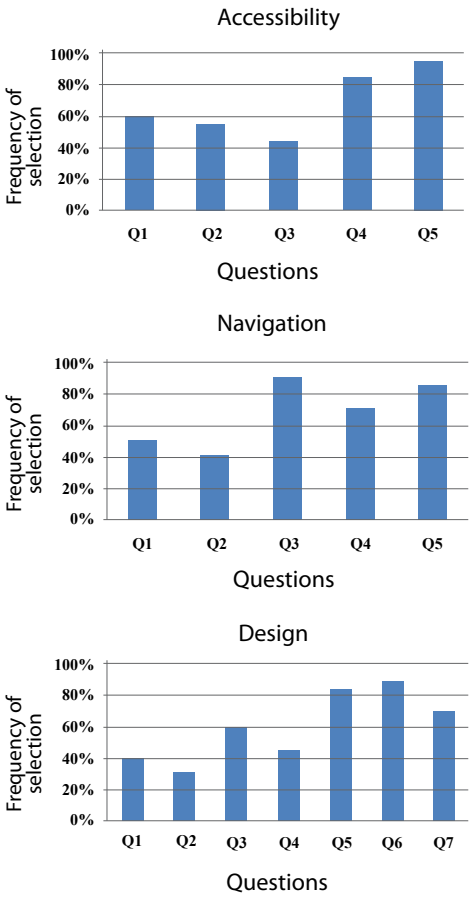


Fig. 18. Statistics of the responses of our platform's users.

TABLE VII. SOME RELEVANT QUESTIONS RELATED TO DESIGN

Number	Question
Q1	Is the site's design aesthetically appealing?
Q2	Are the colors used harmonious and logically related?
Q3	Are the color choices visually accessible? (For example high enough in contrast to assist the colorblind and visually impaired in reading the site appropriately)
Q4	Is the design audience appropriate?- The standard text size should be readable for visitors who don't know how to adjust their browsers.
Q5	Are the fonts easy to read on various screen resolutions?
Q6	The site is homogeneous from one page to another (no visual break)
Q7	The number of colors is limited (correct = 4)?

TABLE VIII. SOME RELEVANT QUESTIONS RELATED TO NAVIGATION

Number	Question
Q1	Are links labeled with anchor text that provides a clear indication of where they lead without over using exact match anchor text?
Q2	What is the maximum number of clicks it takes to reach a page within the depths of the site?
Q3	Is a response given immediately (0.1 seconds) after a click is made on a hyperlink?
Q4	Do clickable items stylistically indicate that they are clickable?
Q5	How intuitive is it to navigate? Are signs obvious or obscured? Buttons/Links Like Text that are not clickable and vice versa, links/buttons that cannot be identified as such?

We proposed a questionnaire consisting of six (6) categories for thirty (30) adhered users. We collected the answers to the questions and we obtained the results that are presented in Figure 19.

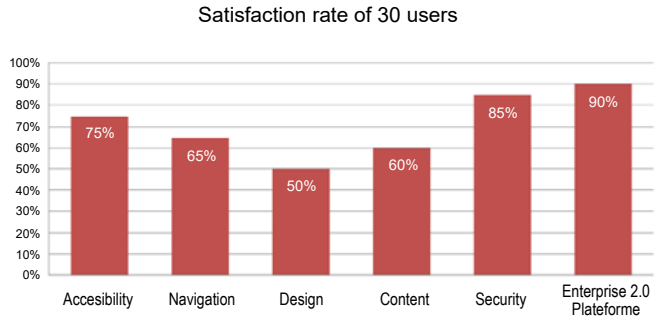


Fig. 19. Evaluation of the usability of our platform.

According to the results, we concluded that the system could be improved to increase design and content. This can be achieved by adding new collaborative environmental assessment questions.

VI. CONCLUSION AND FUTURE WORKS

The emergence of Web 2.0 and its recent success in public-sector organizations (Enterprise 2.0) has served to fulfill objectives that are planned for an organization. It is shown in this research that the application of an enterprise 2.0 project (that is based on Web 2.0 technologies) could assist organizations improve business processes and increase employee's productivity, communications, and information sharing.

More precisely, the emergence of Enterprise 2.0 platforms has given SMEs a big step forward in providing solutions to overcome hard-to-solve problems. In this paper, some of them have been tackled especially collective intelligence, social interaction, and knowledge management. Our suggested solution implies three domains of research: enterprise 2.0, multi-agents systems (MAS), and social network.

Our motivation behind the choice of the MAS is to lighten the Web server tasks in enterprise2.0 platform, and to provide more security to information exchange with some respect to confidentialities intra and inter enterprises. The objective of our work is to propose a helpful approach utilizing new coordination protocol (CordiNet) and computational collective intelligence for enterprise 2.0 design. Firstly, we implement a collaborative environment (Pro Social Network) that allows employers to share diagnosis and fault repair procedures. Secondly, we propose a coordination environment that is based on multi-agents system and interaction protocol.

This study is based on capitalization of user's intelligence that is expressed in catalogs in order to be used by other users with a specialized authorization. The MAS is responsible of managing information traffic inter or intra enterprises. To optimize and reorder the information research, a max cost uniform algorithm has been used with a slight modification.

Some companies are candidates and interested by our prototype such as: Unilever (Oran-Algeria), Inotis (Oran-Algeria) for a real implementation.

For further research, this framework should be improved by using learning coordinator agent which capitalizes knowledge from past experiences. Furthermore, even if our framework consists on a learning feature to find solution, a semantic feature is very essential. We intend to integrate some tools that treat the Semantic Web.

References

[1] A. P. McAfee. "Enterprise2.0: The Dawn of Emergent Collaboration". MIT Sloan Management Review, 2006.

[2] T. O'Reilly, "What is web 2.0: Design patterns and business models for the next generation of software". Communication & Strategies 1, 2007.

[3] A. Sheun, "Web 20: A strategy guide : Business thinking and strategies behind successful web 2.0 implementations". O'Reilly Media, 2008.

[4] D. Beer, R. Burrows, "Sociology and, of and in web 2.0: some initial consideration". Sociological Research Online 12, 2007.

[5] R. L. Cross, and A. Parker, "The Hidden Power of social networks: Understanding How Work Really Gets Done in Organizations". Harvard Business Press. Boston, MA:USA, 2004.

[6] J. Payne, "Using wikis and blogs to improve collaboration and knowledge sharing". Strategic HR Review 7, 2008, pp. 5-12.

[7] A. M. Zhang, Y. Zhu, H. Hildebrandt, "Enterprise networking web sites and organizational communication in Australia". Business Communication Quarterly 72, 2009, pp. 114-119.

[8] P. Lévy, "L'intelligence collective : pour une anthropologie du cyberspace", la Découverte, 1997.

[9] D. Billouard-Fuentes, "Emergence d'Intelligence Collective dans un Environnement Numérique de Travail", 7ème colloque du chapitre français de l'ISKO International Society for Knowledge Organization, Lyon, France, pp.10, June 2009.

[10] E. Lazega, "Analyse de réseaux et sociologie des organisations". Revue française de sociologie, 1994, pp. 35-23.

[11] R. Sikora, and M.J. Shaw, "A multi-agent framework for the coordination and integration of information systems". Management science, vol 44, 1998, pp.65-78.

[12] T. Blecker, and G. Graf, "Multi-agent systems in internet based production environments – an enabling infrastructure for mass customization". Paper presented at the second Interdisciplinary World congress on Mass Customization and Personalization Munich, Germany. In proceeding, 2003, pp, 1-27.

[13] S. Liu, and R.I.M. Young, "Utilizing information and knowledge models to support global manufacturing coordination decisions". Int. J. of computer Integrated Manufacturing, Vol. 17, 2004, pp.479-492.

[14] P.M. Francisco, and H.N. Douglas, "Multi-agent mediator architecture for distributed manufacturing", J. of Intelligent Manufacturing, Vol. 7, 1996, pp. 257-270.

[15] D. Wang, S.V. Nagalingam, and G.C.I. Lin, "Development of an agent-based Virtual CIM architecture for small to medium manufacturers", Robotics and Computer-Integrated Manufacturing, Forthcoming, doi.10.1016/j.rcim.2005.09.001.

[16] L. Wahl and A. Kitchel, "Internet based collaboration tools", In International journal of e-Collaboration, Vol 12, Issue1, January-march 2016.

[17] H.N. Sad and N. Taghezout. "Integrating Agents into a Collaborative Knowledge-based System for Business Rules Consistency Management" In International Journal of Interactive Multimedia and Artificial Intelligence , Vol 4 N°2, pp 61-72, December 2016.

[18] E. Franchi "A multi-Agent Implementation of Social Networks", CEUR Workshop Proceedings, 2010, vol 621.

[19] I. Bessedik and N. Taghezout : "Web 2.0 Integration in a WEB-Based Decision Support System: Effect Study of Social Networking on Decision-Making" In IJWA, Vol 4, N°2, pp 69-77, (2012).

[20] N. Flor, D. Piraquive, R. G. Crespo and Víctor H. M. García. "Managing Projects and Organizations within Enterprise 2.0", Pensee Journal Vol 76, No. 5; May 2014.

[21] M. Corso, A. Martini and Andrea Pesoli, "Enterprise 2.0 : What models are emerging?" The results from a 70 case-based research, In International journal of Knowledge and Learning 4(6), pp: 595-612, January 2008.

[22] N. Ramadan and N. Al-Qirim, "New Hybrid Web 2.0 Adoption Governance Framework for Public Sector", In International journal of e-Collaboration 11, 1, pp 8-21, January-march 2015.

[23] B. Schauer, M. Zeiller, and R. Matzinger, "Open source tools for enterprise 2.0: A Feature-Based Adaptable Analysis". A. Fred et al. (Eds): IC3K 2011, Springer-Verlag Berlin Heidelberg. CCIS 348, pp. 407-422, 2013.

[24] K. Riemer, "The Market for E-Collaboration Systems – Identification of System classes using cluster analysis". In Österle, H. Schelp, J., Winter, R. (eds.) Proceeding Fifteenth European Conference on Information Systems, 2007, pp. 346-357.

[25] J. Xu, J. Zhang, T.Harvey, J. Young, "A survey of Asynchronous Collaboration Tools". Information Technology Journal 7(8), 2008, pp. 1182-1187.

[26] J. Soriano, D. Lizcano, M. A. Canas, M. Reyes, and J. J. Hierro. "Fostering Innovation in a MAshup-oriented Enterprise 2.0 Collaboration Environment". System and Information sciences Notes, no 1, 2007, pp. 62-68.

[27] D. Bianchini, V. De Antonellis, M. Melchiori, "Collaborative mashup development in Enterprise 2.0". Proceedings of the Fifth Interop-Vlab. It Workshop on Complexity of Systems, Complexity of Short Papers for Seminar IM. Volume 2013.

[28] J.L. KONING and S. PESTY, "Modèles de communication, in Principes et architectures des systèmes multi-agents", édité par Jean-Pierre Briot et Yves Demazeau, Collection IC2, Hermes Science Publications, Paris, 2001.

[29] Gray RS., et al., "Mobile-agent versus client/server performance: scalability in an information-retrieval task". Mobile Agents: 5th International Conference, MA, Atlanta, Georgia, number 2240 in Lecture Notes in computer Science, 2001, pp. 229-243.

[30] The FIPA ACL Message Structure Specifications. http://www.fipa.org/specs/fipa00061/, 2002.

[31] P.M. Francisco, and H.N. Douglas, "Multi-agent mediator architecture for distributed manufacturing", J. of Intelligent Manufacturing, Vol. 7, 1996, pp.257–270.

[32] T. D. Wilson, "Human information behavior". Information Science Research 3, 2000, pp. 49-55.

[33] D. O. Case, "Looking for Information, Second Edition: A Survey of Research on Information Seeking, Needs and Behavior", Library and Information Science, academic Press, December 2006.

[34] Fisher, K. E., S. Erdelez, and L.E. Mckechine. "Theories of Information Behaviour". Asist Monograph. Information Today, August, 2005.

[35] X. Yun and C. Jianbin. "A Self-organization Analysis of Knowledge Collaboration with Enterprise 2.0 Application". In International Journal of u- and e- Service, Science and Technology Vol.8, No.1, 2015, pp.211-220.

Seddik Reguieg

Seddik Reguieg was born on Dec 5, 1985 in Oran (Algeria); he earned his Computer Science Engineering Degree with specialization Industrial Computer Science in 2008 from Oran 1 Ahmed Ben Bella University (Algeria), he also received Computer Science Magister Degree with specialization in Diagnostic and Decision-Making Assistance and Human Interaction Machine in 2011 from the same university. He is currently an assistant teacher at University of Hassiba Benbouali of Chlef and a Ph.D. student in Computer Science Department at University of Oran 1 Ahmed BenBella. His fields of special interest include Artificial Intelligence, Collaborative filtering, Recommendation Systems, WEB 2.0 Technologies, and Graph theory.

Noria Taghezout

Noria Taghezout is an assistant professor at university of Oran 1 Ahmed Ben Bella, Algeria. She holds her doctorate thesis in MITT at PAUL SABATIER UNIVERSITY in France in 2011. She also received another doctorate thesis in Distributed Artificial Intelligence from university of Oran 1 Ahmed Ben Bella in 2008. She holds a Master degree in Simulation and Computer aided-design. She conducts her research at the LIO laboratory as a chief of the research group in Modeling of enterprise process by using agents and WEB technologies. Since she studied in UPS Toulouse, she became a member of the EWG-DSS (Euro Working Group on Decision Support Systems). She is currently lecturing Collaborative decision making, Enterprise management and Interface human machine design. Her seminars, publications and regular involvement in Conferences, journals and industry projects highlight her main research interests in Artificial Intelligence.

Workforce Optimization for Bank Operation Centers: A Machine Learning Approach

Sefik Ilkin Serengil[1]*, Alper Ozpinar[2]

[1] Research and Development Center, SoftTech A.S., Istanbul (Turkey)
[2] Mechatronics Engineering Department, Istanbul Ticaret Universitesi, Istanbul (Turkey)

Received 28 October 2015 | Accepted 23 December 2016 | Published 29 December 2017

unir
LA UNIVERSIDAD EN INTERNET

Abstract

Online Banking Systems evolved and improved in recent years with the use of mobile and online technologies, performing money transfer transactions on these channels can be done without delay and human interaction, however commercial customers still tend to transfer money on bank branches due to several concerns. Bank Operation Centers serve to reduce the operational workload of branches. Centralized management also offers personalized service by appointed expert employees in these centers. Inherently, workload volume of money transfer transactions changes dramatically in hours. Therefore, workforce should be planned instantly or early to save labor force and increase operational efficiency. This paper introduces a hybrid multi stage approach for workforce planning in bank operation centers by the application of supervised and unsupervised learning algorithms. Expected workload would be predicted as supervised learning whereas employees are clustered into different skill groups as unsupervised learning to match transactions and proper employees. Finally, workforce optimization is analyzed for proposed approach on production data.

Keywords

Artificial Neural Networks, Forecasting, Machine Learning, Predictive Models, Time Series Analysis.

DOI: 10.9781/ijimai.2017.07.002

I. Introduction

Technology revolution within the last decades makes the use of smart phones, phablets, tablets, computers in every instance of our daily and business lives. Finance sector and banks are also highly affected from this revolution and they adapt their systems to these new trends. More retail customers use digital banking channels with different layers of digital security and approvals result in nearly human free operation. On the other hand, commercial customers' behavior still tends to use bank branches for money transfer transactions due to the nature of their organizational structures.

A customer could fax an instruction including money transfer transactions and authorized signature to the bank branch. This service is accepted as the customer to be situated in branch physically and promised to be commited in 90 minutes (SLA – Service Level Agreement). These instructions mostly consist of large count of transaction orders and also with money amounts much greater than the electronic transfer limits. Branch employees validate the signature on instruction, scan and then deliver it to the operation center. That's the way how branch employees spend much more time for customer relationships.

Daily operation deals with this NP Type problem of planning the turnover and shifts according to banks standard operation procedures by forecasting the expected workload. Thus, proper workforce planning decreases the labor force costs whereas service quality and customer satisfaction increases. Call Centers and Operation Centers are the most common fields of this kind of businesses [1].

Banks aim to reduce the operational workload of branches through operation centers. Expert employees are appointed in operation centers to serve faster, more accurate and high quality service. Operation center employees digitalize the hard copy money transfer instruction and commit the transaction. Delivering operational transactions centrally provides service quality enhancement, improves customer satisfaction and saves serious amount of labor force.

Currently, operation center managers usually predict the workload and assign the workforce manually depending on their previous experience of the team and operation leaders. Mostly, they reschedule and change their plans at the moment the workload density is observed. This type of delayed planning results in non-optimized working environment which should be avoided. Considering the hundreds of average employee numbers in operational and call centers this optimization cannot be ignored.

Workforce optimization problem in bank operation centers is pretty similar to Call Center studies. Transaction volumes of the both business fields have dramatically changed [2]. Related work by authors already mentions predictability of transaction counts [3]. Moreover, formulas for inbound transaction volume are generalized by extracting time based attributes of historical data in recent works [4] [5]. Similarly, some other studies including smoothing methods would be adapted to the problem [6]. Also, these business fields have high turnover and shift rates of labour force [7]. Herein, skill based approaches exist to allocate optimized labour force [8] [9].

This paper introduces an approach to plan workforce in bank operation centers based on forecasting workload with a machine learning algorithm and analyzes the workforce optimization on production data. Also, a generic workload forecasting system is

* Corresponding author.
E-mail address: serengil@gmail.com

developed to involve different transaction types or business fields. The worked on data obtained from Isbank's operation center date between 2012, Jan 01 – Present. The historical data includes transaction date time and transaction volume information.

II. Background

Human based experience and intelligence can be replaced by artificial intelligence and expert systems in many areas with the improvements and findings in these areas within the last decades [10] [11]. A special area of artificial intelligence which is mostly based on statistics is the machine learning. This discipline is strong about modelling NP type problems.

In machine learning problems, a mathematical function is modeled by given historical data examples and the obtained learned function forecasts the outputs of future examples without known outputs. The function is derived from the affecting factors of the problem that would be predicted. Determining affecting factors specifies the complexity of function and this is the key factor of model success. If the function is not complex enough to cover the state space of the problem, there would be underfitting problem. This means that the function would not even successfully forecast for the given historical examples. Function should be re-modeled again as to become more complex for this case. In contrast, there would be overfitting problem if the function is too complex. This means that function would forecast successfully for given historical examples but could not predict successfully for unknown examples with large computational and time complexity.

For the case of machine learning adaptation and usage for the expert system of the bank operational centers, there are several factors affecting on transaction count that have already been discovered. Firstly, by the nature of business operations, day of week is one of the most important factors affecting transaction count as demonstrated (see Fig. 1). Transaction count peaks up on Friday and Monday due to the weekends. Moreover, if the public holidays shift the first or last work day of week, transaction counts of the following days are affected dramatically. That's why; Boolean first or last work day parameter should be additionally included in input layer. Furthermore, half work days shift every year because of the hijri calendar. That's why; Boolean half day parameter is added in input.

Secondly, Morning hours have low transaction volume whereas dinner hours bottom out. There is an obviously seen trend on evening hours where transaction count peaking up as illustrated (see Fig. 1). Furthermore, transaction counts show a change depending on month of year as shown in Fig. 2. Also, day of month causes to change in transaction count. Generally, customers tend to transfer money on beginning, ending and the middle of the month as shown in Fig. 2.

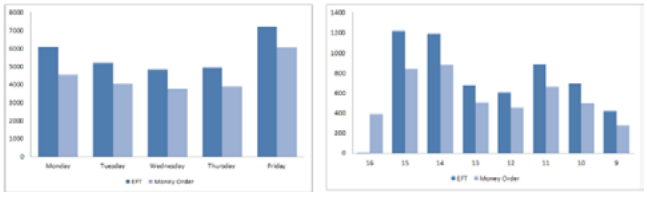


Fig. 1 Weekday and Work Hour Effect.

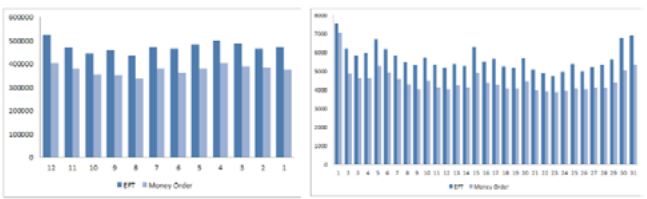


Fig. 2 Month of Year and Day of Month Effect.

Thirdly, yearly deviation is thought to be useful as input to catch the trend. Deviation is calculated by the difference between current and previous year's average transaction count for 10 days period. Finally, transaction count of the previous hours should be included in the input to retrieve future values in time series problems. Thus, transaction counts of previous three hours (h-1, h-2, and h-3) are included into the network. Finally, the model is based on aiming to retrieve transaction count of hour h. Thus, the output of the network should be transaction count (h).

TABLE I.
Correlative Variables for Money Transfer Transactions

Var	Description	Min	Max	Corr.
X1	Hour	8	17	0.0945
X2	Day	1	31	-0.0590
X3	Month	1	12	0.0029
X4	Is Monday	0	1	0.0564
X5	Is Tuesday	0	1	-0.0581
X6	Is Wednesday	0	1	-0.1127
X7	Is Thursday	0	1	-0.0928
X8	Is Friday	0	1	0.2118
X9	Year	2012	2016	-0.2681
X10	Yearly Deviation	-1213	694	-0.0799
X11	Week of year	1	52	-0.0193
X12	Week of month	1	6	-0.0696
X13	Is Spring	0	1	0.0215
X14	Is Summer	0	1	0.0109
X15	Is Fall	0	1	-0.0521
X16	Is Winter	0	1	0.0218
X17	Is Last Workday of Month	0	1	0.0201
X18	Is First Workday of Month	0	1	0.0989
X19	Middle Workday of Month	0	1	0.0343
X20	Trx count (h-1)	0	3718	0.2832
X21	Trx count (h-2)	0	3518	0.0278
X22	Trx count (h-3)	0	2656	0.3234
X23	Is After Religious Day	0	1	-0.0417
X24	Is Before Religious Day	0	1	0.0883
X25	Is Between Long Holiday	0	1	-0.0083
X26	Is First of Last Workday	0	1	0.2217
X27	Is Halfday	0	1	-0.0082

The variables in TABLE I are thought to affect money transfer transaction counts. Minimum, maximum and correlation coefficient values are computed for Electronic Fund Transfer (EFT) transactions. Correlation indicates the strength of relationship between two variables. It ranges from -1 to +1. Coefficient closes -1 or +1 for strongly related datasets. Sign of the coefficient states the direction of relation. Neutral means no relationship between datasets. None of the correlation coefficient of variable closes to ±1. This means there is no directly relation between these variables and transaction count.

III. Motivation

A basic neural network cell has ability to learn, remember and predict. A neuron consists of multiple inputs and an output as illustrated in Fig. 3. Each input (x) would be involved in network through own weight (w) which specifies the strength of input on output. Learning is provided by adjusting the weight values positively or negatively. Assembly function calculates the net input (o) which is derived from the sum of the multiplying the inputs and their own weights. Activation function (commonly sigmoid function) computes the net output (y). Finally, the output of the neuron is calculated by the formulas (1) and (2) for vectors $w = [w_0, w_1, w_2, \ldots, w_n]$ and $x = [x_0, x_1, x_2, \ldots, x_n]$

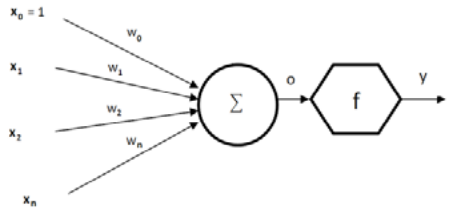


Fig. 3 Basic Neural Network Cell.

$$o = w^T x \quad (1)$$

$$y = 1 / (1 + e^{-w^T x}) \quad (2)$$

A complex neural network system consists of multiple neuron cells. It provides a satisfactory way to forecast and predict. Designing input - output parameters and modeling neural network are dramatically important to have successful results.

The challenge in this study is predicting transaction volumes. In other words, it is a regression problem. Neural networks can be applied on both regression and classification studies. On the other hand, linear regression can be applied for regression studies but this algorithm is not convenient for non-linear time series. What's more, some statistical methods can be applied on non-linear time series such as exponential smoothing methods but these models are based on discovering trends and seasonal effects on data set. What if some activities might repeat irregularly (e.g. random day in a week) but these activities may have high effect on result such as sport activities or religious days. In this case, statistical approach might fail whereas neural networks can be successful if event days are defined in input parameters. That's why, neural networks and exponential smoothing methods are applied on the problem in this study.

Network design plays a pivotal role to have successful results. Three layered network is modeled with node numbers 27, 18, 1. Firstly, nodes in input layer correspond to the variables illustrated in Table I. Secondly, the number of hidden neurons should be 2/3 the size of the input layer, plus the size of the output layer as Heaton declared [12]. That's why, there are 18 nodes in hidden layer. Thirdly hidden nodes are connected with an output layer. Finally, output node calculates expected transaction count.

Additionally, sigmoid function is selected as activation function and back propagation algorithm and stockastic gradient descent is applied to implement learning. Furthermore, some configuration parameters of the network model is mentioned in the Table II. So, neural network models is built in 260 seconds for following parameters and historical data consisting of 7K instances. Network is built on a machine with following capabilities: Core i7 CPU, 16 GB RAM and 64 bit OS.

TABLE II.
CONFIGURATION PARAMETERS OF THE NEURAL NETWORK MODEL

Configuration Parameter	Value
Learning rate	0.01
Momentum	0.1
Epoch	5000

IV. FORECASTING RESULTS

Neural network model is trained with 7151 instances date between [2013-01-14, 2016-08-22] whereas model is evaluated for 801 production instances on runtime date between [2016-08-23, 2016-12-31]. In other words, the following results show how the built model works on production on a bank. Performance of the network is demonstrated in Table III. Moreover, prediction and actual values are demonstrated in Fig. 4 for a day and Fig. 5 for overall. Also, the same historical data set is modeled with triple, double and single exponential smoothing methods and the model success is compared. It seems the neural network model forecasts much more successful than the others. Mean Absolute Error and Correlation metrics are calculated to evaluate the performance of the system.

Suppose that p is the prediction set and a is the actual set. Performance of the system would be calculated by formulas (3) and (4).

$$MAE = (1/n).\sum_{i=1}^{n} | p_i - a_i | \quad (3)$$

$$r = \sum_{i=1}^{n}(p_i - \bar{p})(a_i - \bar{a}) / \sqrt{\sum_{i=1}^{n}(p_i - \bar{p})^2 \sum_{i=1}^{n}(a_i - \bar{a})^2} \quad (4)$$

Fig. 4 NN Prediction Result on a day (2015-12-04).

TABLE III
PERFORMANCE OF THE PREDICTION MODELS

Trx	Metric	NN	TES	DES	SES
EFT	MAE	71.05	283.69	290.09	294.14
	MAE/Mean (%)	14.27	56.97	58.26	59.07
	Correlation (%)	97.02	30.08	-1.16	12.43
	Mean	497.91			
	Instances(hour)	801			
Money Order	MAE	48.80	161.34	161.42	156.05
	MAE/Mean (%)	13.71	48.01	48.04	46.44
	Correlation (%)	95.70	18.46	4.48	26.80
	Mean	336.00			
	Instances(hour)	801			

To sum up, forecast and actual sets have 97.02% correlation for EFT and 95.70% correlation for Money Order transactions. Moreover, the neural network model would forecast with distance ±71.05 and ±48.80, and also MAE / Mean ratios are 14.27% for EFT and 13.71% for Money Order respectively on a dataset consisting of 801 instances. In contrast, triple, double and single exponential smoothing methods cannot get close to the neural networks results as seen in Table III.

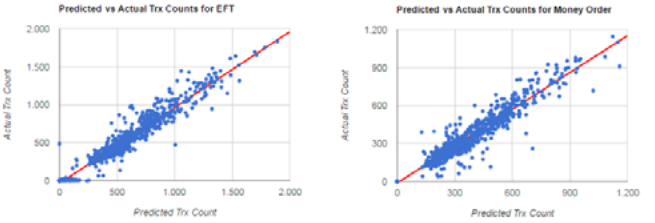


Fig. 5 NN Results between [2016-08-23, 2016-12-31].

V. UNSUPERVISED LEARNING

There are class labels for dataset in supervised learning. Machine

learning algorithms create a generalized function from historical data and this function predicts the class labels for unknown examples. In this case, predicted classes could be compared with actual classes and the success performance of the function could be calculated. In contrast, there are not class labels for dataset in unsupervised learning. In this case, dataset could be grouped into different clusters. Newly examples are assumed to be members of these created clusters. However, function performance could not be calculated for unsupervised learning problems. There are commonly known algorithms such as k-means or c-means to implement unsupervised learning.

Satisfactory results are retrieved with neural networks as shown is previous section. The requirement is matching expected work with the correct employees to optimize the problem.

Employees should be evaluated by their skills and the individual performance should be considered while distributing work to have more sensitive results. In this section, employee skills are examined by unit performing time of a work and the average count of completed works by itself on an hour period. Two different approaches are proposed for workforce planning. Firstly, aggressive mode is proposed to handle high transaction volumes to reduce the queue immediately. Secondly, moderate mode is proposed to encourage personal development. Applying aggressive or moderate modes should be decided by expected transaction volumes.

Fig. 6 is retrieved when the employee skills are analyzed between 06/27/14 - 08/27/14 for 265 employees. Every node represents an employee and every shape (star, square and triangle) stands for different skill groups. Unsupervised k-means algorithm is applied to cluster employees. Simply, employees are clustered into three different sets depending on their skills. Star shaped employees (Cluster 1) seem to be in the highest performance work group. They complete large number of work in a short time period.

In aggressive mode, distributing work should begin with employees in cluster 1; continue with square shaped employees (Cluster 2), triangle shaped employees (Cluster 3) respectively. Suppose that employee array is ordered by cluster priority, sorted with respect to the unit count column from greatest to smallest and also sorted with respect to the unit perform time column in seconds from smallest to greatest respectively. Proposed employee assignment process is illustrated in Fig. 7 as pseudo code. PN is expected transaction count in next hour; PQ is waiting transaction count on queue from previous hours. The algorithm basically proposes to start with reserving the most powerful employee and repeat it until expected work handled. If the expected work could not be handled with all employees, then exit iteration.

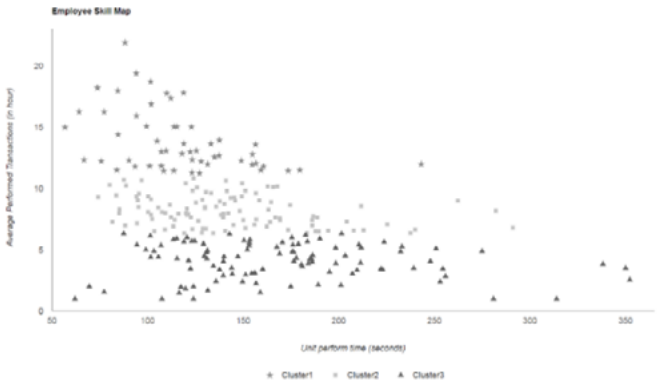


Fig. 6 Employee Skill Map.

```
int workforce = 0,  i = 0, break = 0

while(workforce < (PN + PQ) AND break = 0)
    if(i < total_number_of_employees)
        if(isAvailable(Employee[i]))
            workforce += Employee[i].UnitCompletedWork
            Reserve(Employee[i])
        end if
        i = i+1
    end if
    else
        break = 1
    end else
end while
```

Fig. 7 Pseudo code of Employee Assignment for Aggressive Mode.

TABLE IV. Workforce Clusters

Cluster	Employees	Avg Commited Trx
Cluster 1 (star)	53	15.00
Cluster 2 (square)	106	9.00
Cluster 3 (triangle)	106	5.00

Needed workforce should be computed by average completed work count of each employee instead of unit perform time. In this way, delays between works would not be ignored. This approach guarantees to reduce the queue immediately. However, talented employees would handle more transactions in this approach. Alternatively, fairer moderate mode is proposed for workforce planning.

In moderate mode, employees are group into different teams as demonstrated in Table IV. Cluster 1 consists of 53 employees, and Cluster 2 and Cluster 3 both consist of 106 employees. Each team would consist of a Cluster 1 member, two Cluster 2 members and Cluster 3 members. In this way, there would be 53 different and equal talented teams and each team consists of 5 members. Moreover, Cluster 1 employees complete almost 15 transactions per hour whereas Cluster 2 and Cluster 3 members complete 9 and 5 transactions respectively. In other words, each team could handle averagely 43 transactions per hour. Furthermore, SLA time is 90 minutes for money transfer transactions. In other words, 2/3 the size of expected work should be completed for each hour. Besides, assigned teams should be membered 2 / (3x43) the size of expected transaction count and each transaction should be assigned to a team randomly. A generalized algorithm for moderate type distribution is illustrated in Fig. 8. This approach is seemed to be fairer to aggressive mode and it also encourages the personal development. However, it does not guarantee to reduce the queue immediately. This approach should be applied on low transaction volume.

VI. Workforce Optimization

If the aggressive mode workforce planning method is applied on production data for dates between Sep 01, 2015 and Oct 20, 2015 on operation center, workforce cost would reduce almost 6.5%. Moreover, SLA (promised time to commit the transaction) would decrease from 90 minutes to 60 minutes in aggressive mode; this means work time is optimized 33.3%. This describes the moments proposed operator count larger than reserved operators.

```
sub create_teams() //clustering employees into 3 classes
        cluster1[int(num_of_employees/5)]
        cluster2[2*int(num_of_employees/5)]
        cluster3[2*int(num_of_employees/5)]

        //employee array is sorted by their skills from greatest to smallest
        for i=1 to num_of_employees
            if i <= cluster1.length
                cluster1[i] = employee[i]
            else if i <= (cluster1.length + cluster2.length)
                cluster2[i-cluster1.length] = employee[i]
            else
                cluster3[i - cluster1.length - cluster2.length] = employee[i]
        end for

        team[cluster1.length] //creating teams
        for i=1 to cluster1.length
            index1stMember = random (1, cluster1.lenght +1 - i)
            team[i].add(cluster1[index1stMember]);
            cluster1.remove(index1stMember);
            for j=1 to 2
                index2ndMember = random (1, cluster2.lenght +j - i)
                team[i].add(cluster2[index2ndMember]);
                cluster2.remove(index2ndMember);
            end for
            for j=1 to 2
                index3rdMember = random (1, cluster3.lenght +j - i)
                team[i].add(cluster3[index3rdMember]);
                cluster3.remove(index3rdMember);
            end for
        end for
end sub

sub reserve_teams
        threshold = 43 //a team could handle in a hour
        reserved = (60/SLA) * expected_work / threshold
        reserved_team[reserved] //new reserved team array

        for i = 1 to reserved
            temp_team = random(1, team.length
            reserved_team[i] = team[random (1, temp_team)]
            team.remove(temp_team)
        end for
end sub

sub assign_work
        for i=1 to queuedWork.length
            team[random (1, team.length)].assign(queuedWork[i])
        end for
end sub
```

Fig. 8 Pseudo Code of Employee Reservation for Moderate Mode.

Fig. 9 describes the overall optimization on system for aggressive mode and for moderate mode. The blue lines state the moments proposed operator count less than the reserved operator count, and the red lines express the moments proposed operator count larger than reserved operator count. X-axis corresponds to the work hour instances, and Y-axis corresponds to the workforce optimization for work hour. Also, workforce optimization is calculated by formula (5).

$$optimization = (proposed\ employee - reserved\ employee)/reserved\ employee \quad (5)$$

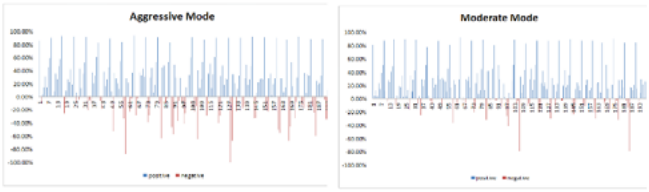


Fig. 9 Workforce Optimization (x-axis: instance, y-axis: optimization percentage).

To sum up, planning workforce with less number of employees is possible to handle the more workload as shown in Table V.

TABLE V.
WORKFORCE OPTIMIZATION RESULTS

Metric	Aggressive Mode	Moderate Mode
Reduced Employees	3422	4318
Instance (hour)	197	197
Total employees	265	265
Optimization (Employee/hour)	17.37	21.91
Optimization (%)	6.55%	8.27%
SLA (minutes)	60	90

VII. SYSTEM ARCHITECTURE

The proposed hybrid multilevel expert system is designed as illustrated in Fig. 10. Firstly, historical data should be regularly retrieved and stored. This operation is handled by ETL module. Secondly, machine learning algorithm is run by machine learning module. Training network and predicting is applied on this module. ETL module also provides to be run machine learning module with current data. Thirdly, AI module is responsible for reserving and assigning required workforce based on employee skills and expected workload calculated by ML module. Also, applying moderate or aggressive mode is decided on AI module, too. Then, feedback module validates how correct machine learning module predicts. Thus, network can be retrained if validation fails. After then, these operations are implemented on service oriented architecture and service layer communicates with data layer to collect and store data.

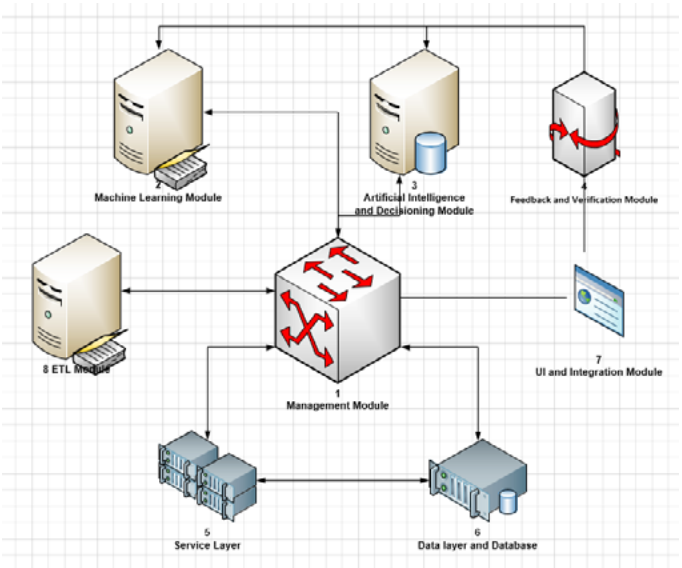


Fig. 10 System Architecture.

Deeply, Generic Machine Learning module provides to create dynamic neural network model for different transaction types. Thus,

new business processes or transaction types would be modeled easily in the expert system. Common time based attributes (e.g. weekday, transaction hour) are defined strictly whereas custom exceptional days (e.g. holiday for USD region, religious day) could be defined and involved in neural network model as seen in Fig. 11. At this point, some parameters help to handle time series problems such as has previous hours as input whereas correlative depth parameter indicates how many previous hour would be included in network as input.

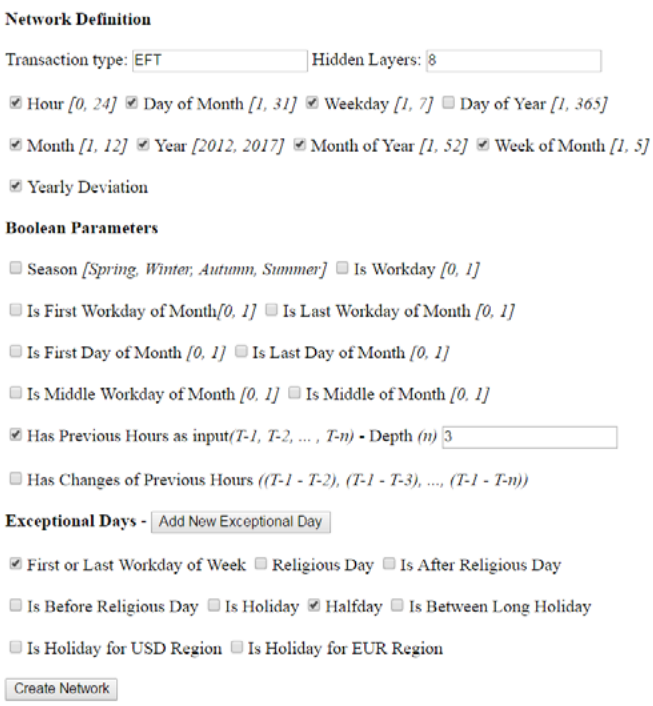


Fig. 11 Dynamic Network Definition.

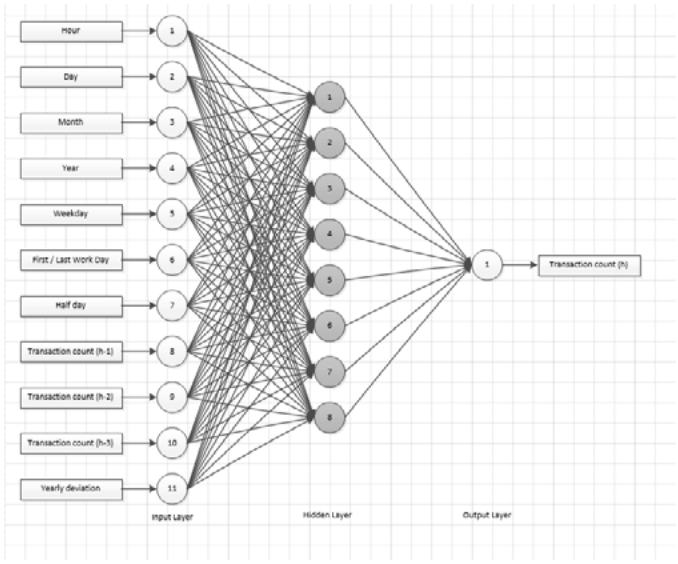


Fig. 12 Created Neural Network Model.

Dynamically defined network parameters create a custom neural network model. Thus a generic solution would be provided as seen in Fig. 12. Finally, neural network learning is applied by the pseudo code as illustrated in Fig. 13.

```
For i=0 to trxs.length //e.g. trxs[0] = EFT, trxs[1] = Money Order
  For k=0 to trxs[i].retrieveHistoricalData.length
        For j=0 to trxs[i].retrieveParams.length
//Retrieve selected parameters on dynamic network definition
//Suppose that Hour, Day and Month Params are picked up.
          trxs[i].attributes[k][j]
          =Extract(trxs[i].retrieveHistoricalData[k].getTrxDate()
          , trxs[i].retrieveParams[j])
        End for
        trxs[i].attributes[k][j+1]
          = retrieveHistoricalData[k].getTrxCount()
          //append transaction count
  End for
    /*Then EFT's attributes (trxs[0].attributes) are illustrated below
    11, 30, 12, 1315
    10, 30, 12, 968
     9, 30,12, 593 */
    applyNeuralNetworkLearning(trxs[i].attributes)
End for

Sub Extract()
  ...
   /* This method extracts custom parameter from transaction date. E.g. Extract(2016-12-30 11:00, Hour) returns 11,  Extract(2016-12-30 11:00, Day) returns 30 and Extract(2016-12-30 11:00, Month) returns 12  */
End sub

Sub retrieveHistoricalData()
  ...
   /* This method returns all historical data including transaction date and transaction count for a transaction type. Suppose that the following instances are historical data of EFT
  2016-12-30 11:00, 1315
  2016-12-30 10:00, 968
  2016-12-30 09:00, 593 */
End sub
```

Fig. 13 Pseudo Code of Neural Network Training.

VIII. Conclusion

In this paper, a hybrid multi-level machine learning based expert system approach is introduced to plan workforce management for bank operation centers based on applying supervised and unsupervised machine learning algorithms for forecasting workload and clustering workforce. Chosen supervised machine learning algorithm, neural network, is compared with alternative exponential smoothing algorithms to evaluate how successful results are. Furthermore, workforce optimization is analyzed on production data. Satisfactory results are obtained for both workload forecasting and workforce optimization.

Although, this paper mainly focuses on forecasting transaction volumes of money transfer transactions, a generic architecture is developed to workload forecasting. Thus, newly transaction types or business fields should be easily involved in workload forecasting lifecycle. A similar approach is thought to be adopted in turnover and shift requiring work areas such as Call Centers.

Acknowledgments

This work has been conducted by SoftTech A.S. under the project number 5059, and supported by Turkish government organization TUBITAK TEYDEB (Technology and Innovation Funding Programs

Directorate of The Scientific and Technological Research Council of Turkey) in scope of Industrial Research and Development Projects Grant Program (1501) under the project number 3150070.

The authors would like to thank Isbank Enterprise Architecture Division and Banking Operations Division for their help and support.

References

[1] Z. Aksin, M. Armony and V. Mehrotra, «The modern call center: A multi-disciplinary perspective on operations management research,» Production and Operations Management, vol. 16, no. 6, pp. 665-688, 2007.

[2] L. Brown, N. Gans, A. Mandelbaum, A. Sakov and H. Shen, «Statistical analysis of a telephone call center: A queueing-science perspective,» Journal of the American statistical association, vol. 100, no. 469, pp. 36-50, 2005.

[3] S. I. Serengil and A. Ozpinar, «Planning Workforce Management for Bank Operation Centers with Neural Networks,» in Proceedings of 15th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Databases, Venice, Italy, 2016.

[4] D. Barrow and N. Kourentzes, «The impact of special days in call arrivals forecasting: a neural network approach to modelling special days,» European Journal of Operational Research, 2016.

[5] F. M. Bianchi, S. Scardapane, A. Uncini and A. Rizzi, «Prediction of telephone calls load using Echo State Network with exogenous variables,» Neural Networks, vol. 71, pp. 204-213, 2015.

[6] C.-C. Lin, C.-L. Lin and J. Z. Shyu, «Hybrid multi-model forecasting system: A case study on display market,» Knowledge-Based Systems, vol. 71, pp. 279-289, 2014.

[7] T. L. Tuten and P. E. Neidermeyer, «Performance, satisfaction and turnover in call centers: The effects of stress and optimism,» Journal of Business Research, vol. 57, no. 1, pp. 26-34, 2004.

[8] B. Legros, O. Jouini and Y. Dallery, «A flexible architecture for call centers with skill-based routing,» International Journal of Production Economics, vol. 159, pp. 192-207, 2015.

[9] M. A. Valle, S. Varas and G. A. Ruz, «Job performance prediction in a call center using a naive Bayes classifier,» Expert Systems with Applications, vol. 39, no. 11, pp. 9939-9945, 2012.

[10] C. Pruenza, M. Teulón, L. Lechuga, J. Díaz and A. González, «Development of a Predictive Model for Induction,» International Journal of Interactive Multimedia and Artificial Intelligence, vol. In Press, 2017.

[11] S. Bhattacharya and S. Nath, «Intelligent e-Learning Systems: An Educational,» International Journal of Interactive Multimedia and Artificial Intelligence, vol. In Press, 2017.

[12] J. Heaton, Introduction to Neural Networks for Java, Heaton Research, 2000.

Sefik Ilkin Serengil

Sefik Ilkin Serengil received the MSc degree in computer engineering from Galatasaray University, Istanbul, Turkey, in 2011. He has been working as a Software Developer for SoftTech A.S. since 2010. His current research interests are Machine Learning and Cryptography.

Alper Ozpinar

Alper Ozpinar received the PhD degree in me-chanical engineering from Yildiz Technical University, Istanbul, Turkey, in 2007. He is currently Asst. Prof. Dr. in Mechatronics Engineering Department at Istanbul Ticaret University. His current research interests are Renewable Energy Sources, Artificial Intelligence, Machine Learning and System Optimization.

Smart Algorithms to Control a Variable Speed Wind Turbine

Nabil Farhane[1], Ismail Boumhidi[1], Jaouad Boumhidi[2] *

[1] LESSI Laboratory, Department of Physics, Faculty of Sciences, Sidi Mohammed Ben Abdellah, Fez, (Morocco)

[2] LIIAN Laboratory, Department of Computer Science, Faculty of Sciences Sidi Mohammed Ben Abdellah, University of Fez (Morocco)

Received 2 June 2017 | Accepted 15 July 2017 | Published 21 August 2017

Abstract

In this paper, a robust adaptive fuzzy neural network sliding mode (AFNNSM) control design is proposed to maximize the captured energy for a variable speed wind turbine and to minimize the efforts of the drive shaft. Fuzzy neural network (FNN) is used to improve the mathematical system model, by the prediction of model unknown function, which is used by the Sliding mode control approach (SMC) and enables a lower switching gain to be used despite the presence of large uncertainties. As a result, the used robust control action did not exhibit any chattering behavior. This FNN is trained on-line using the backpropagation algorithm (BP). The particle swarm optimization (PSO) algorithm is used in this study to optimize the learning rate of BP algorithm in order to improve the network performance in term of the speed of convergence. The stability is shown by the Lyapunov theory and the trajectory tracking errors converge to zero without any oscillatory behavior. Simulations illustrate the effectiveness of the designed method.

Keywords

Adaptive Fuzzy Neural Network Sliding Mode (AFNNSM), Sliding Mode Control (SMC), Particle Swarm Optimization (PSO), Variable Speed Wind Turbine.

DOI: 10.9781/ijimai.2017.08.001

I. Introduction

Wind energy is an abundant renewable source of new electrical generation capacity in the world, and it is exploited by converting the kinetic energy of moving air mass into electricity; therefore, it is necessary to introduce tools to make these installations more profitable [1]. Practically, there are two main types of horizontal axis wind turbines: fixed speed and variable speed [2]. In this study, we consider the case of variable speed, due to its great ability in the extraction of energy. In addition to that, variable speed system is more complex and requires an efficient control strategy [3]. Several studies have been devoted to the control of the aeroturbine mechanical as well as the electrical components. This work is devoted to the mechanical part (aeroturbine), with the main objective of designing a controller in order to maximize the energy captured from the wind and minimize the stress on the drive train shafts that takes into consideration the nonlinear nature of the system behavior and the flexibility of the drive-train shaft. The proposed control structure also overcomes the drawbacks of some existing control methods.

The sliding mode control [4-7], is widely used in the control of a variable speed wind turbine. This is due to its property of robustness with respect to uncertainties and disturbances [8],[9]. The chattering behavior is especially the main problem in the design of SMC. One possible method to solve this problem is the boundary layer approach [6]. This technique has given good results when the system uncertainties are small. However, when these uncertainties are large, a high gain is needed and higher amplitude of chattering is produced. In order to solve this problem, we proceed to use FNN for the estimation of the unknown model function so that the system uncertainties can be kept small and hence enable a lower switching gain to be used. The designed method is a combination of traditional SMC and FNN with online adaptation of the parameters.

Intelligent systems such as fuzzy systems and neural networks are being used successfully in an increasing number of application areas [10]-[14]-[23]. Fuzzy neural networks (FNNs) have the low-level learning and computation power of neural networks, and the high-level human-like thinking and reasoning of fuzzy theory. The proposed control consists of the so called equivalent control added with robust control term, the FNN predicted unknown terms are incorporated in the equivalent control component, thus enable the robust component to be used with a small gain which is responsible to compensate the network errors prediction. The FNN is trained on-line using the backpropagation algorithm (BP). The learning rate is one of the parameters of BP algorithm which have a significant influence on results; learning rate which is too small or too large may not be favourable for convergence. Particle swarm optimization (PSO) algorithm is used in this work to optimize this parameter in order to improve the training speed.

This study is organized as follows. The next section presents the wind turbine modelling. Section 3 shows the design of the proposed adaptive fuzzy neural network sliding mode control. In Section 4, simulation results are provided to demonstrate the robust control performance of the proposed (AFNNSM) control. Finally, in section 5 a concluding remark is given.

* Corresponding author.

E-mail addresses: nabilfarhane84@gmail.com (Nabil Farhane), iboumhidi@hotmail.com (Ismail Boumhidi), jaouad.boumhidi@usmba.ac.ma (Jaouad Boumhidi).

II. Wind Turbine Modelization

Wind energy across a surface A_v depends on the cube of the wind speed v, and the density of the air ρ. This energy is given by:

$$P_v = \frac{1}{2} A_v \rho v^3 \tag{1}$$

where:

$$A_v = \pi R^2 \tag{2}$$

R is the radius of the rotor. A variable speed wind turbine is composed of aeroturbine, a gearbox, and a generator. The aerodynamic power captured by the rotor is given as follows:

$$P_a = \frac{1}{2} \rho \pi R^2 C_p(\lambda, \beta) v^3 \tag{3}$$

Where C_p is the power coefficient, β is the pitch angle and the tip-speed ratio, λ, is given as follows:

$$\lambda = \frac{\omega_r R}{v} \tag{4}$$

With ω_r is the rotor speed. By using the relationship:

$$P_a = \omega_r T_a \tag{5}$$

The aerodynamic torque expression is:

$$T_a = \frac{1}{2} \rho \pi R^3 C_q(\lambda, \beta) v^2 \tag{6}$$

Where:

$$C_q(\lambda, \beta) = \frac{C_p(\lambda, \beta)}{\lambda} \tag{7}$$

$C_q(\lambda, \beta)$ is the torque coefficient.

In the literature, the wind turbine is presented by the modelling of the mechanical part [15]-[17], or by that of the electrical part [18], [19]. In this paper, we are interested in the modelling of the mechanical part of the variable speed wind turbine, which is presented by a two-mass model that is shown in Fig. 1.

The dynamics of the rotor is characterized by a differential equation of the first order:

$$J_r \dot{\omega}_r = T_a - T_{ls} - K_r \omega_r \tag{8}$$

J_r and K_r are respectively the rotor inertia and the rotor friction coefficient.

The low-speed shaft torque T_{ls} resulting effects of friction and torque generated by the differences between the rotor angular velocity ω_r and that of the output shaft ω_{ls} (see Fig. 1):

$$T_{ls} = B_{ls}(\theta_t - \theta_{ls}) + K_{ls}(\omega_r - \omega_{ls}) \tag{9}$$

B_{ls} and K_{ls} are respectively the shaft stiffness coefficient and the shaft damping coefficient.

θ_r and θ_{ls} are respectively the rotor side angular deviation and the gearbox side angular deviation.

The relationship between the high-speed shaft torque T_{hs} and the generator electromagnetic torque T_{em} is given by:

$$J_g \dot{\omega}_g = T_{hs} - K_g \omega_g - T_{em} \tag{10}$$

J_g, ω_g and K_g are respectively the generator inertia, the generator speed and the generator friction coefficient. Let's assume an ideal gearbox with transmission ratio n_g, we get:

$$n_g = \frac{T_{ls}}{T_{hs}} = \frac{\omega_g}{\omega_{ls}} = \frac{\theta_g}{\theta_{ls}} \tag{11}$$

θ_g is the generator side angular deviation.

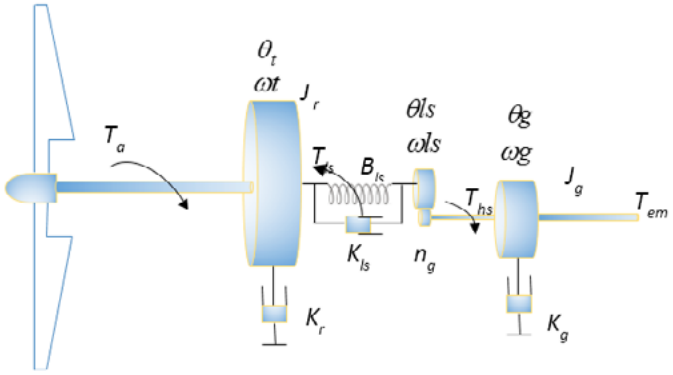


Fig. 1. Two-mass model of wind turbine.

Replacing the time derivative of Tls from (9) and using (10) and (11) we get the following dynamic system model (12) :

$$\begin{bmatrix} \dot{\omega}_r \\ \dot{\omega}_g \\ \dot{T}_{ls} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \omega_t \\ \omega_g \\ T_{ls} \end{bmatrix} + \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} T_a + \begin{bmatrix} b_{12} \\ b_{22} \\ b_{32} \end{bmatrix} T_{em}$$

Where:

$$a_{11} = -\frac{K_r}{J_r}, \quad a_{12} = 0, \quad a_{13} = -\frac{1}{J_r}, \quad a_{21} = 0,$$

$$a_{22} = -\frac{K_g}{J_g}, \quad a_{23} = \frac{1}{n_g J_g}, \quad a_{31} = \left(B_{ls} - \frac{K_{ls} K_r}{J_r} \right),$$

$$a_{32} = \frac{1}{n_g} \left(\frac{K_{ls} K_g}{J_g} - B_{ls} \right), \quad a_{33} = -K_{ls} \left(\frac{J_r + n_g^2 J_g}{n_g^2 J_g J_r} \right),$$

$$b_{11} = \frac{1}{J_r}, \quad b_{12} = 0, \quad b_{21} = 0, \quad b_{22} = -\frac{1}{J_g}, \quad b_{31} = \frac{K_{ls}}{J_r},$$

$$b_{32} = \frac{K_{ls}}{n_g J_g}$$

The nonlinear character of this system is due to the aerodynamic torque as follows:

$$T_a = \frac{1}{2}\rho\pi R^3 C_q(\lambda,\beta)v^2 \tag{12}$$

This depends on a strongly nonlinear way, the rotor speed ω_r, the blade pitch angle β and the wind speed v which is a not controllable input, random and strongly fluctuating.

III. Design of The Fuzzy Neural Network Sliding Mode Control

The main control objective is to maximize the captured wind energy and minimize the efforts of drive train shafts. The power coefficient curve $C_p(\lambda,\beta)$ has a unique maximum which corresponds to the optimal wind energy [20]:

$$C_p(\lambda_{opt},\beta_{opt}) = C_{p_{opt}} \tag{13}$$

The rotor, thus, provides maximum aerodynamic power, only to the tip-speed λ_{opt}:

$$\lambda_{opt} = \frac{\omega_{r_{opt}} R}{v} \tag{14}$$

To maximize the captured energy of the wind, the variables λ and β must be maintained at their optimal values in order to ensure maximum value of C_p. So, the blade pitch angle is fixed at its optimal value β_{opt}. The tip-speed λ depends on both of the wind speed v and the rotor speed ω_r. As the wind speed is not a controllable input, the rotor speed ω_r must be adjusted by T_{em} (As seen in Fig. 2), to track the optimal reference given by:

$$\omega_{r_{opt}} = \frac{\lambda_{opt}}{R} v \tag{15}$$

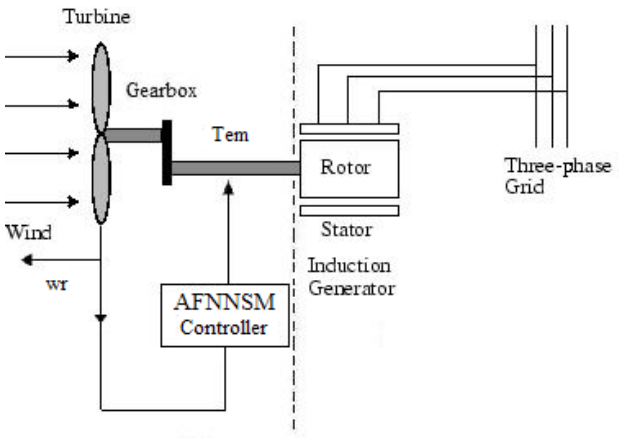


Fig. 2. Basic configuration of wind turbine system.

A. Controller Design

Let define $u = T_{em}$, $x_1 = \omega_r$ and $x_3 = \omega_g$, the dynamic system model described by (12) can be rewritten in the state space as follows:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f(x_1,x_2,x_3) + bu + d(t) \\ \dot{x}_3 = a_{22}x_3 + \dfrac{a_{23}}{a_{13}}(x_2 - a_{11}x_1 - b_{11}T_a) + b_{32}u \\ y = x_1 \end{cases} \tag{16}$$

$u \in R$ and $y \in R$ are respectively the input and the output of the system, $\underline{x} = [x_1,\ x_2]^T \in R^2$ is the state vector of the system which is assumed to be available for measurement, and $f(x_1,x_2,x_3)$ is the nominal representation of the system given as:

$$f(x_1,x_2,x_3) = \left(a_{13}a_{31} - \frac{a_{13}a_{33}a_{11}}{a_{13}}\right)x_1 + \left(a_{11} + \frac{a_{13}a_{33}}{a_{13}}\right)x_2 + a_{13}a_{32}x_3 + \left(a_{13}b_{31} - \frac{a_{13}a_{33}b_{11}}{a_{13}}\right)T_a + b_{11}\dot{T}_a \tag{17}$$

$d(t)$ represents the unknown model part of the system (uncertainties and external disturbances) and $b = a_{13}b_{32}$ is known constant.

In the following, the tracking error is defined as:

$$e = \omega_r - \omega_{r_{opt}} \tag{18}$$

The relative degree of the system (16) is $r = 2$.

The sliding variable can be defined as:

$$S = \dot{e} + \gamma e = (x_2 - \dot{\omega}_{topt}) + \gamma(x_1 - \omega_{topt}) \tag{19}$$

Where γ is a positive constant.

Differentiating S with respect to time, we have:

$$\dot{S} = f(x_1,x_2,x_3) + bu + d(t) + \gamma x_2 - (\ddot{\omega}_{topt} + \gamma\dot{\omega}_{topt}) = F(x,t) + bu \tag{20}$$

where the function:

$$F(x,t) = f(x_1,x_2,x_3) + \gamma x_2 - (\ddot{\omega}_{topt} + \gamma\dot{\omega}_{topt}) + d(t) \tag{21}$$

The following condition (22) must be satisfied [6] to guarantee the existence of sliding mode in finite time:

$$S\dot{S} < -\eta|S| \tag{22}$$

Where η is a small positive constant. The control law that satisfies Eq. (20) is given by:

$$u = u_{eq} + U_s \tag{23}$$

Where $\hat{u}_{eq} = -\dfrac{\hat{F}(x,t)}{b}$, with $\hat{F}(x,t)$ is the Fuzzy neural network prediction of smooth unknown nonlinear function F, given in the next section. The additive control is:

$$U_s = -\frac{\alpha}{b} sat(S) \tag{24}$$

Where $sat(.)$ is the saturation function given by:

$$sat(S(x,t)) = \begin{cases} S/L & if\ |S| < L \\ \text{sgn}(S) & \text{otherwise} \end{cases} \tag{25}$$

L is the boundary layer thickness and α is chosen according to the following theorem.

The fuzzy neural network prediction error is denoted:

$$F - \hat{F} = \varepsilon_F \tag{26}$$

With $|\varepsilon_F| < \varepsilon_F^*$ and ε_F^* is the upper bound of the network error prediction assumed known.

The block diagram of the AFNNSM system is depicted in Fig. 3.

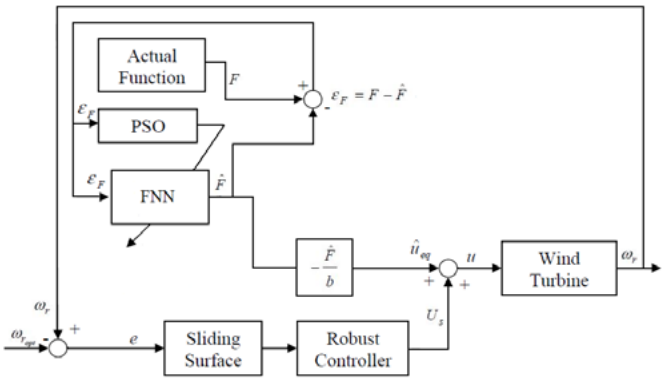


Fig. 3. Block diagram of the proposed AFNNSM control wind turbine.

Theorem: Consider the system described by (16) in the presence of large uncertainties. If the system control is designed as:

$$u = \hat{u}_{eq} - \frac{\alpha}{b} sat(S) \tag{27}$$

Where $\hat{u}_{eq} = -\frac{\hat{F}(x,t)}{b}$, with $\hat{F}(x,t)$ is the output of the proposed fuzzy neural network, $\varepsilon_F^* + \eta < \alpha$; the trajectory tracking errors will converge to zero in finite time.

Proof: Consider the candidate Lyapunov function:

$$V = \frac{1}{2} S^2$$

$$\dot{V} = S\dot{S}$$

$$\dot{V} = S(F(x,t) + bu)$$

By replacing the expression of u given in the theorem we have:

$$\dot{V} = S(F(x,t) - \hat{F}(x,t) - \alpha sat(S))$$

$$= S\varepsilon_F - \alpha S sat(S) \le |S||\varepsilon_F| - \alpha S sat(S)$$

$$< |S|\varepsilon_F^* - \alpha S sat(S)$$

By choosing $\varepsilon_F^* + \eta < \alpha$ we have: For any $L > 0$, if $|S| > L$, $sat(S) = sign(S)$ the function $\dot{V} = (\varepsilon_F^* - \alpha)|S| < -\eta|S|$. However, in a small L-vicinity of the origin [6], $sat(S) = \frac{S}{L}$ is continuous, the system trajectories are confined to a boundary layer of sliding mode manifold $S = 0$.

B. Fuzzy Neural Network Representation

In this paper, we consider a FNN with five-layered of adjustable weights (Fig. 4). The $\{x_i\}_{i=1,2}$ are input variables, $\hat{F} = \hat{\xi}(\underline{x}, t)$ is the output variable. Fig. 4 shows the structure of the FNN, which is comprised of the input, the membership, the rule, the normalized and the output layers. To give a clear understanding of the signal propagation and mathematical function in each layer, the following section describes FNN functions layer by layer.

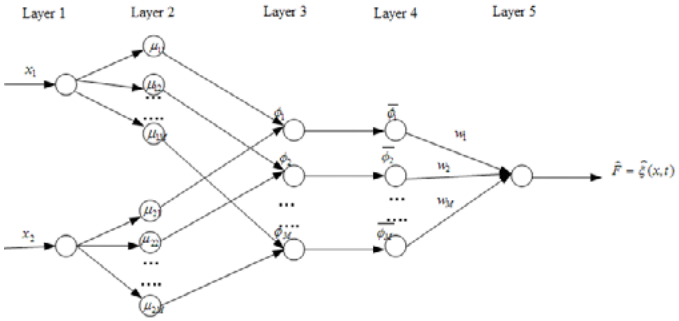


Fig. 4. The architecture of the proposed fuzzy neural network for the prediction of uncertain part.

Layer 1: This layer, which corresponds to one input variable $x_k (k = 1,2)$, only transmits input values to the next layer directly.

Layer 2: In this layer, each node performs a membership function and acts as a unit of memory. The Gaussian function is adopted as the membership function. For the kth input, the corresponding net input and output of the lth node can be expressed as:

$$\mu_k^l(x_k) = \exp[-\frac{(x_k - c_k^l)^2}{(\sigma_k^l)^2}] \tag{28}$$

where c_k^l and σ_k^l ; $l = 1,2,\cdots,M$ are the mean and standard deviation of the Gaussian function of the lth partition for the kth input variable x_k , respectively, and M is the total number of fuzzy rules.

Layer 3: This layer represents one fuzzy logic rule and performs precondition matching of a rule.The output of a rule node in this layer is calculated by the product operation as follows.

$$\phi_l(\underline{x}) = \prod_{k=1}^{2} \mu_k^l(x_k) \tag{29}$$

Layer 4: In this layer, each of these firing strengths of the rules is compared with the sum of all the firing strengths. Therefore, the normalized firing strengths are computed in this layer as,

$$\bar{\phi}_l(\underline{x}) = \frac{\phi_l(\underline{x})}{\sum_{l=1}^{M} \phi_l(\underline{x})} \tag{30}$$

Layer 5: This layer (or output layer) calculates the summation of its input values from the previous.

$$\hat{\xi}(\underline{x}) = \sum_{l=1}^{M} w_l \bar{\phi}_l(\underline{x})$$

The aim of the learning algorithm is to adjust the weights of w_{ξ}^{l}, c_{ξ}^{l} and σ_{ξ}^{l}.The on-line learning algorithm is a gradient descent (GD) search algorithm in the space of network parameters [21]. The essence of (GD) consists of iteratively adjusting the weights the direction opposite to the gradient of error, so as to reduce the discrepancy according to:

$$\dot{w}_l = \eta_{w_l} \varepsilon_F \phi_l \tag{31}$$

Where $\eta_{w_y^l}$ is the leaning rate for w_y^l

The updated laws of c_{ς}^{l} and σ_{ς}^{l} also can be obtained by the gradient decent search algorithm:

$$\dot{c}_k^l = 2\eta_{c_k^l} \varepsilon_F \phi_l w_l \frac{(x_k - c_k^l)}{(\sigma_k^l)^2}$$

$$\dot{\sigma}_k^l = 2\eta_{\sigma_k^l} \varepsilon_F \phi_l w_l \frac{(x_k - c_k^l)^2}{(\sigma_k^l)^3}$$

Where $\eta_{c_k^l}$ and $\eta_{\sigma_k^l}$ are the learning-rate parameters of the mean and the standard deviation of the Gaussian functions.

C. Particle Swarm Optimization Algorithm

The objective is to optimize the learning rate η_{w_l} by the PSO algorithm.

In PSO, *m* particles fly through an *n*-dimensional search space. For each particle *i*, there are two vectors: the velocity vector $V_i = (v_{i1}, v_{i2}, ..., v_{in})$ and the position vector $X_i = (x_{i1}, x_{i2}, ..., x_{in})$. Similar to bird socking and fishes schooling, the particles are updated according to their previous best position $P_i = (p_{i1}, p_{i2}, ..., p_{in})$ and the whole swarm's previous best position $P_g = (p_{g1}, p_{g2}, ..., p_{gn})$. This means that particle *i* adjusts its velocity V_i and position X_i in each generation according to the equations bellow [22]:

$$v_{id}(t+1) = v_{id}(t) + c_1 \times rand(\ )_1 \times (p_{id} - x_{id}) + c_2 \times rand(\ )_2 \times (p_{gd} - x_{id}) \tag{32}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \tag{33}$$

Where $d = 1,2,...,n$; c_1, c_2 are the acceleration coefficients with positive values, which we take as: $c_1 = 1$ and $c_2 = 2$ in this paper; $rand(\ )_1$, $rand(\ )_2$ are random numbers between 0 and 1. The new velocity and position for each particle are calculated using the equations (32) and (33) based on its velocity $v_{id}(t)$, best position P_{id} and the swarm's best position P_{gd}.

In order to calculate the optimized parameter of learning rate η_{w_l} given in equation (29), the PSO is used *on-line* to minimize the neural network prediction error.

We define the quadratic errors e_{rq} as: $e_{rq}(t_i) = \varepsilon_F^2(t_i)$

The objective function f to be minimized is chosen as the norm of the quadratic error: $f = norm(E_{rq})$ with E_{rq} the vector that contains all errors $e_{rq}(t_i)$.

The PSO algorithm tests the search space using *m* particles according to (30) and (31). Each particle i moves in search space and stores its best position P_{id}, then, it compares all positions to finally take out the chosen $\eta_{k-optimum}$ that gives the minimum value of the objective function f.

IV. Simulation Results

Different simulation cases are considered for the case of a two-mass model of the wind turbine with:

- Presence of constant additive control disturbance d of 10 $kN.m$,
- Presence of an additive measurement noise on ω_g, with a SNR of approximatively 7 *dB*.
- A wind speed profile of 7 m/s means value.

The two-mass model parameters are presented below:

$$R = 21.65\ m,\ \rho = 1.29\ kg/m^3,$$

$$J_r = 3.25 \times 10^5\ kg.m^2,\ J_g = 34.4\ kg.m^2,$$

$$K_{ls} = 9500\ N.m/rad/s,\ B_{ls} = 2.691 \times 10^5\ N.m/rad,$$

$$K_r = 27.36\ N.m/rad/s,\ K_g = 0.2\ N.m/rad/s,$$

$$n_g = 43.165.$$

Three membership functions have been used for each of the two inputs of the FNN. The associated fuzzy sets with Gaussian membership function for each input signal, have initial input for the mean and the standard deviation:

$$c_1 = [-6.4470 \quad 1.2640 \quad 21.9693]$$

$$\sigma_1 = 1.0e+005 * [0.0043 \quad 0.0241 \quad 2.1675]$$

$$c_2 = [-130.1245 \quad -275.7842 \quad -30.8045]$$

$$\sigma_2 = 1.0e+005 * [0.0140 \quad 0.0261 \quad 4.1794]$$

The initial output weights :

$w = [-10.9070 \quad 119.7829 \quad -118.3859]$, being the initial weights, which are obtained hors-line identification system.

Fig. 5 presents the wind speed profile used in this study. Fig. 6 shows the estimation error for the predicting function F.

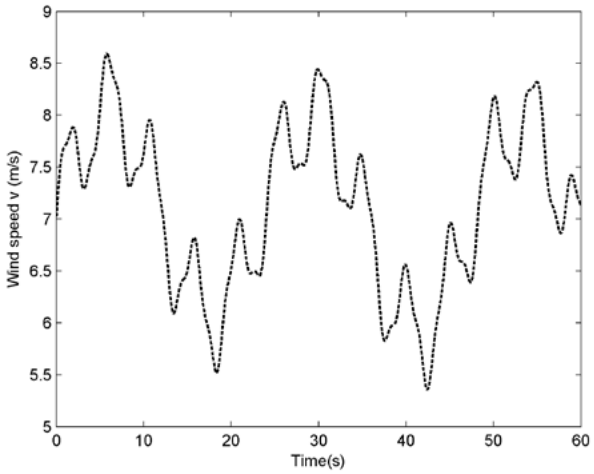


Fig. 5. Wind speed profile of 7 m/s mean value.

Fig. 5 is proposed for an average speed, it was considered the case where the wind speed is highly variable and the case where the wind speed does not change rapidly.

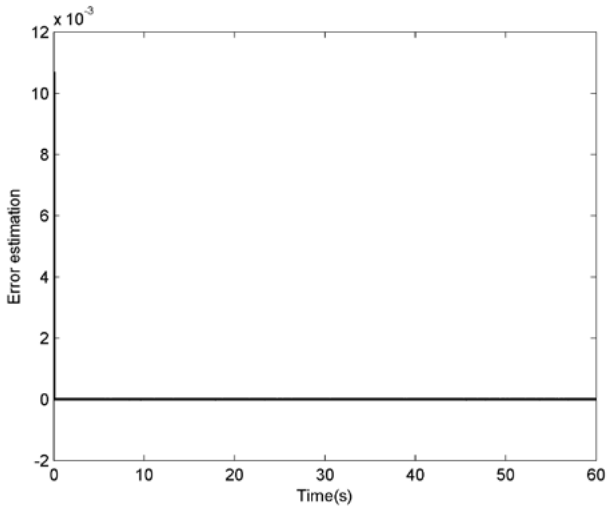


Fig. 6. Error estimation of the predicting function F.

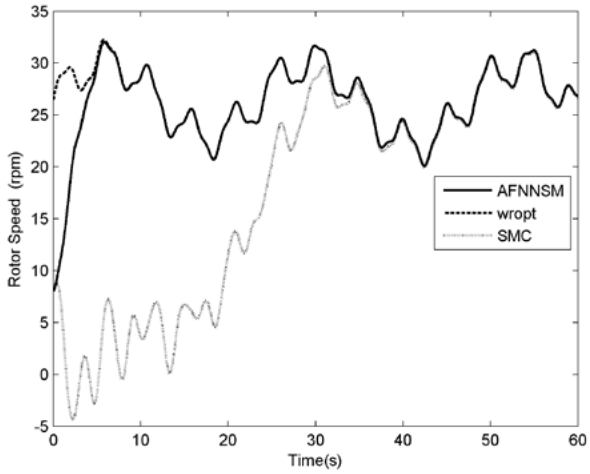


Fig. 7. Rotor speed with $\alpha = 1.2$.

To gain a large enough alpha = 1.2 we see the output controlled by the standard sliding mode converges slowly. However that commissioned by the proposed method NNSMC converges rapidly to the desired output

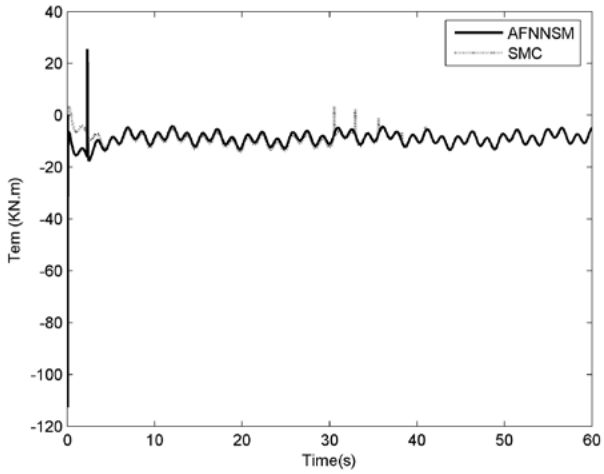


Fig. 8. Electromagnetic torque with $\alpha = 1.2$.

Fig. 8 shows the actual and optimal trajectory of the rotor speed, we find that the best tracking performance is obtained when applying the RFNNSM.

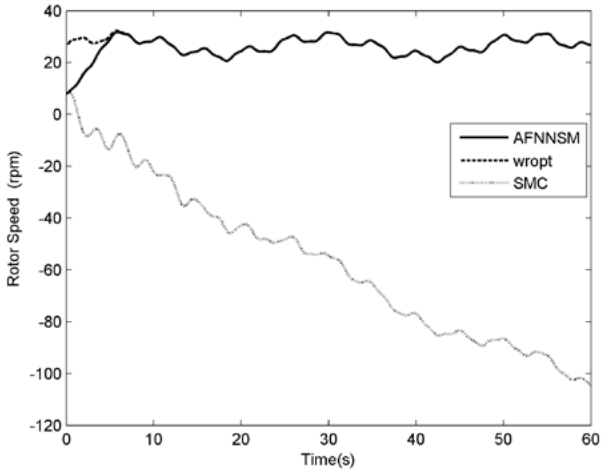


Fig. 9. Rotor speed with $\alpha = 0.9$.

For comparison we have considered in the control law, for both AFNNSM and traditional SMC controllers, the same gain $\alpha = 1.2$(see the obtained outputs in Fig. 7) and $\alpha = 0.9$ (see the obtained outputs in Fig. 9). From these figures, it can be seen that, the best tracking performance is obtained when the proposed AFNNSM controller is applied. The corresponding control signals are given in Fig. 8 for $\alpha = 1.2$, and in Fig. 10 for $\alpha = 0.8$. Fig. 9 shows that the standard SMC controller totally failed to control the system.

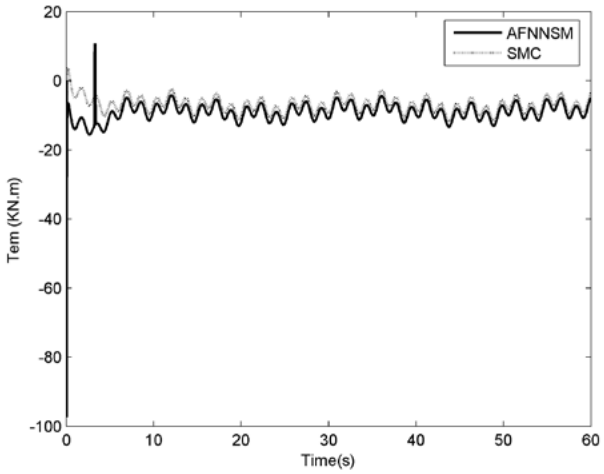


Fig. 10. Electromagnetic torque with $\alpha = 0.9$.

By using the a gain aloha = 0.9, Fig. 10 shows the optimal trajectory of the rotor speed, otherwise we find that the best tracking performance is obtained when the RFNNSM is applying.

Table 1 contains both the gain used in this study as well as the number of the iteration in order to find the global optimum particle of learning rate, so we can remark that the global optimum is found from the 21st iteration.

TABLE I. The Global Optimum Particle of Learning Rate

Iteration number	η_{w_l} for $\alpha = 0.9$	η_{w_l} for $\alpha = 1.2$
1	0.7395	0.4492
2	0.7100	0.4665
3	0.6821	0.4804
4	0.6745	0.4940
5	0.6745	0.5156
6	0.6794	0.5264
7	0.6897	0.5230
8	0.6997	0.5142
9	0.7035	0.5092
10	0.7011	0.5102
11	0.6967	0.5113
12	0.6953	0.5119
13	0.6949	0.5121
14	0.6954	0.5139
15	0.6956	0.5145
16	0.6968	0.5147
17	0.6975	0.5143
18	0.6976	0.5138
19	0.6975	0.5136
20	0.6972	0.5137
21	0.6971	0.5138
22	0.6971	0.5138
23	0.6972	0.5138
24	0.6972	0.5138
25	0.6972	0.5138
26	0.6972	0.5138
27	0.6972	0.5138
28	0.6972	0.5138
29	0.6972	0.5138
30	0.6972	0.5138

V. Conclusion

This paper addressed the robust optimal reference tracking problem for a variable speed wind turbine. The designed method is a combination of traditional sliding mode control approach and fuzzy neural network. The later is employed to approximate the unknown nonlinear model function with online adaptation of parameters via BP learning algorithm. This provides a better description of the plant, and hence enables a lower switching gain to be used despite the presence of large uncertainties. The particle swarm optimization (PSO) algorithm is used to optimize the learning rate of BP algorithm in order to get faster convergence. The comparison with the traditional sliding mode control has been realized and simulation results have shown a good performance of the proposed method to track the optimal reference without any chattering problem.

References

[1] J. F. Manwell, J. G. McGowan, A. L. Rogers, Wind Energy Explained: Theory, Design and Applications (John Wiley & Sons, 2002).

[2] G. Ofualagba, E. U. Ubeku, Wind energy conversion system- wind turbine modelling, Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, IEEE, 2008, 1-8.

[3] T. Burton, D. Sharpe, N. Jenkins, E. Bossanyi, Wind Energy Handbook (John Wiley & Sons, 2001).

[4] S. Abdeddaim, A. Betka. Optimal Tracking and Robust Power Control of the Doubly Fed Induction Generator wind Turbine, Electric Power Systems Research, vol. 49, 2013, pp. 234–242.

[5] V. I. Utkin, Sliding Modes in Control Optimization (Springer-Verlag, 1992).

[6] J.J. Slotine, Sliding Controller Design for Non-linear Systems, International Journal of Control, vol. 40, n. 2, 1984, pp. 421–434.

[7] V. Utkin, J. Guldner, J. Shi, Sliding Mode Control in Electromechanical System (Taylor & Francis, London, 1999).

[8] P.F. Puleston, F. Valenciaga, Chattering Reduction in a Geometric Sliding Mode method. A Robust Low-chattering Controller for an Autonomous Wind System, Control and Intelligent Systems, vol. 37, n. 1, 2009, pp. 39-45.

[9] H. Amimeur , D. Aouzellag , R. Abdessemed, K. Ghedamsi, Sliding Mode Control of a Dual-stator Induction Generator for Wind Energy Conversion Systems, International Journal of Electrical Power and Energy Systems, vol. 42, n. 1, 2012, pp. 60–70.

[10] Y. C. Chen, C. C. Teng, A Model Reference Control Structure Using a Fuzzy Neural Network. Fuzzy Sets and Systems, vol. 73, n. 3, 1995, pp. 291–312.

[11] V. Gorrini, H. Bersini, Recurrent fuzzy systems. In Proceedings of the IEEE international conference on fuzzy systems, vol. 1, 1994, 193–198.

[12] F.B. Duh, C.T. Lin, Tracking a Maneuvering Target Using Neural Fuzzy Network, IEEE Transactions on Systems, Man, and Cybernetics, vol. 34, n. 1, 2004, pp. 16–33.

[13] C.F. Juang, C.H. Hsu, Temperature Control by Chip-implemented Adaptive Recurrent Fuzzy Controller Designed by Evolutionary Algorithm, IEEE Transactions on Circuits System, vol. 52, n. 11, 2005, pp. 2376–2384.

[14] S.J. Lee, C.L. Hou, A Neural-fuzzy System for Congestion Control in ATM Networks, IEEE Transactions on Systems, Man, and Cybernetics, vol. 30, n. 1, 2000, pp. 2–9.

[15] J.M. Rubio, & LT. Aguilar, Maximizing the Performance of Variable Speed Wind Turbine with Nonlinear Output Feedback Control, Procedia Engineering, vol. 35, 2012, pp. 31-40.

[16] L. Munteanu, A. L. Bratcu, N-A. Cutululis, E. Ceanga, Optimal Control of Wind Energy Systems, Advances in Industrial Control (Springer, 2008).

[17] J.F. Adrià, G.B. Oriol, S. Andreas, S. Marc, M. Montserrat, Modeling and Control of The Doubly Fed Induction Generator Wind Turbine, Simulation Modelling Practice and Theory, vol. 18, n. 9, 2010, pp. 1365–1381.

[18] Ph. Delarue, A. Bouscayrol, A. Tounzi, X. Guillaud, Modelling, Control and Simulation of an Overall Wind Energy Conversion System, Renewable Energy, vol. 28, n. 8, 2003, pp. 1169-1185.

[19] A. Gaillard, P. Poure , S. Saadate , M. Machmoum, Variable Speed DFIG Wind Energy System for Power Generation and Harmonic Current Mitigation, Renewable Energy, vol. 34, n.6, 2009, pp. 1545–1553.

[20] K. A. Stol, Geometry and Structural Properties for the Controls Advanced Research Turbine (CART) from Model Tuning (Subcontractor Report SR-500-32087, National Renewable Energy Laboratory, Golden, CO, September. 2004).

[21] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland (Ed), Parallel Distributed Processing, (Cambridge, MA, (1): MIT Press, 1986, 318–362).

[22] M. A. Cavuslua, C. Karakuzub, Neural Identification of Dynamic Systems on FPGA with Improved PSO Learning. Applied Soft Computing, vol. 12, 2012, pp. 2707–2718.

[23] Esmaeilpour, M., and A. R. A. Mohammadi, Analyzing the EEG Signals in Order to Estimate the Depth of Anesthesia using Wavelet and Fuzzy Neural Networks, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, n.2, 2016, pp. 12-15. Doi: 10.9781/ijimai.2016.422.

Nabil Farhane

Nabil Farhane received his M.Sc. degree from Sidi Mohammed ben Abdellah University of Fez in 2011. He is currently pursuing the Ph.D. degree at Sidi Mohammed ben Abdellah University. His research interests include fuzzy logic, neural networks, non-linear control, etc.

Ismail Boumhidi

Ismail Boumhidi is a professor of electronics at the Faculty of Sciences, Fez, Morocco. He received his Ph.D. degree from Sidi Mohammed ben Abdellah University, Faculty of Sciences, in 1999. His research areas include adaptive robust control, multivariable non-linear systems and fuzzy logic control with applications.

Jaouad Boumhidi

Jaouad Boumhidi received his Ph.D. degree in Automatic Control and Computer Sciences from Sidi Mohammed ben Abdellah University in 2005. He is now a professor at Sidi Mohammed ben Abdellah University. His research interests include neural networks, fuzzy logic, robust control, etc.

Segmentation of Arabic Handwritten Documents into Text Lines using Watershed Transform

A. Souhar[1]*, Y. Boulid[2], ElB. Ameur[1] and Mly. M. Ouagague[2]

[1]University Ibn Tofail, Faculty of science, 14000, Kenitra (Morocco)
[2]Research and Development Unit Maxware Technology, 14000, Kenitra (Morocco)

Received 1 June 2017 | Accepted 13 July 2017 | Published 21 August 2017

Abstract

A crucial task in character recognition systems is the segmentation of the document into text lines and especially if it is handwritten. When dealing with non-Latin document such as Arabic, the challenge becomes greater since in addition to the variability of writing, the presence of diacritical points and the high number of ascender and descender characters complicates more the process of the segmentation. To remedy with this complexity and even to make this difficulty an advantage since the focus is on the Arabic language which is semi-cursive in nature, a method based on the Watershed Transform technique is proposed. Tested on «Handwritten Arabic Proximity Datasets» [21] a segmentation rate of 93% for a 95% of matching score is achieved.

Keywords

Text Line Segmentation, Arabic Script, Handwritten Character Recognition, Connected Component Analysis, Projection Profile, Watershed Transform.

DOI: 10.9781/ijimai.2017.08.002

I. Introduction

In order to recognize texts in a document, often character recognition systems go through four stages: the pre-processing stage, which concerns the preparation of the document in terms of normalization and suppression of noise. The segmentation stage that allows the detection of lines, words and also the segmentation of those words into characters. The third stage concerns the feature extraction from the character, this feature allows to minimize the intra-class variance while maximizing the inter-class variance. The fourth stage involves learning and testing to recognize new letters or new words based on machine learning algorithms. To these stages is added a post-processing stage to verify recognized words using a lexical, syntax and semantic analysis.

All these phases are crucial since if an error is made in one it will strongly influence the subsequent ones [23]. In this paper we focus on the text line segmentation which could be defined according to [1] in the process of assigning the same label to units that are partially aligned. We are interested in Arabic handwritten documents which are more challenging than Latin documents and that is mainly due to the semi-cursive nature of the Arabic script which is characterized by its calligraphy and the presence of ascending and descending character, the overlapping between piece of Arabic words and also the diacritical points located either above or below characters [24].

In addition to these characteristics, the different writing styles and the inclination within the same line make the process of text line extraction from Arabic handwriting challenging (Fig. 1).

* Corresponding author.
E-mail address: houssouhar@gmail.com

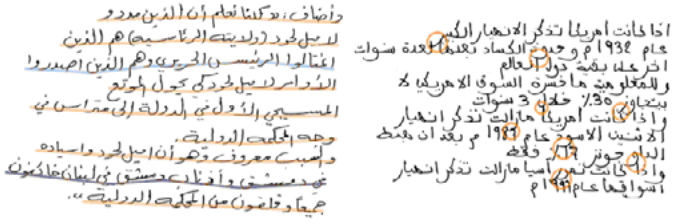


Fig. 1. Some problems of Arabic handwritten text lines extraction.

To cope with this complexity several techniques have been proposed in the literature. Generally we can categorize them into three approaches (Fig. 2).

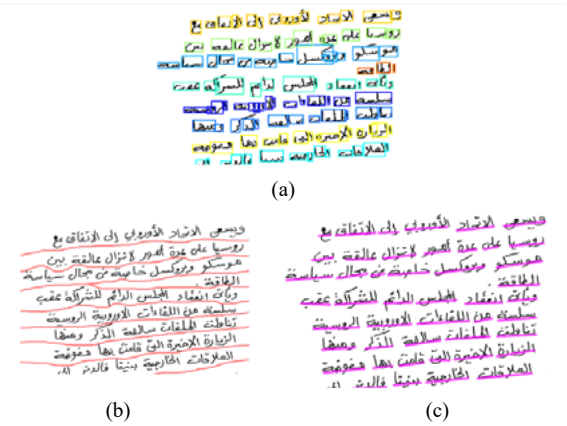
The first approach is based on the analysis of the arrangement of the connected components in order to construct successively the lines. A multi-agent system to detect and group connected components that belong to the same line is proposed in [2]. The first agent estimates global parameters and extracts the lines, the second one searches and detects adjacent components and the third one segments the touching lines. A similar technique is called Smearing method where the goal is to apply some transformation on the text bloc in order to group together homogeneous blocks constituting the lines [3, 4, 5, 6].

The second approach tries to search between interlines in order to separate the adjacent lines. The algorithm in [7] computes the distance transform directly from the gray scale images and generates two types of seams: the medial seams determine the text lines and the separating seams define the upper and lower boundaries of the text lines. Similarly, the work in [8] splits the document into vertical slices and applies a matching method on the result of projection profile in order to estimate the medial seam, then a modified version of seam carving procedure is used to compute the separating seam. The final seams will go through the regions between text lines.

A method was proposed in [9] that uses ALCM algorithm which gives a mask corresponding to text line locations. Then, the text lines are extracted by superimposing the components with text line pattern mask. Another method is based on convolution neural network with watershed transform, which is proposed in [10] to estimate the text area between the baseline and the corpus line.

The third approach uses the baseline of words and tries to connect those that participate to the same lines. Other methods represent the problem as a graph and search the orientations of text components and uses Breadth First Search algorithm and affinity propagation clustering method to assign the components to text lines [11, 12]. The work in [13] adapts the problem of text line extraction from binary Arabic handwritten documents as a Markov Decision Processes using knowledge about the features and arrangement of the components belonging to the same line.

The papers in [1, 14] give a complete study of several text lines segmentation methods of handwritten document.



(a) (b) (c)

Fig. 2. The main approaches for handwritten text lines extraction.

To respond to the specificity of the Arabic script, we investigate the use of the Watershed Transform which is a well known image segmentation approach. The idea is that when considering the text components as markers, the catchment basin (regions) of watershed occupy a space that adapts with the shape and size of the component with respect to its neighbors. This effect could be exploited to estimate the locations of text lines.

To respond to the specificity of the Arabic script, we found interesting to investigate the approach based on the analysis of text component using watershed transform as preprocessing step to get the regions of the text component in order to localize the neighboring components within the same line. We found that the work in [15] first estimates the baseline using adaptive head-tail connection in order to link the component with each other and then uses watershed on the result of the baseline detection in order to extract the lines.

We would like also to investigate the use watershed transform as a post processing step in order to localize the lines estimated by a matching horizontal projection profile technique based on the work in [8].

The goal here is to compare if it is interesting to use watershed transform as preprocessing or post processing steps. This paper is an extension of our work in [22], and it is organized as follows: section 3 describes the proposed approaches. Section 4 presents and discusses the results and finally section 5 concludes the paper.

II. Proposed Approach

According to a local vision, a text line can be perceived as a set of aligned words. Thus we focus on the detection of Piece of Arabic words that constitute this line. From this point of view, the adequate approach will be the one that is based on the analysis of the arrangement of the connected components (Fig. 2.a)

According to a global vision, a text line can only be defined by means of the following line. Here we are not interested to the constituent of the line, but to its neighbors. From this point of view the adequate approach will be the one that approximates the baseline (Fig. 2.b and 2.c).

Therefore we investigate the use of watershed transform technique from two visions, local and global.

A. Watershed Transform

In topography watershed means the ridge that divides areas drained by different river systems. Here we refer to image segmentation technique in the field of mathematical morphology. The idea is to treat the image as a topographic surface in which the dark pixels are considered low elevations and lighter ones are considered high elevations and when the flooding starts from the low elevations and the merging of the waters from different locations is prevented, the resulted image will be portioned into catchment basins and watershed lines as illustrated in Fig. 3.

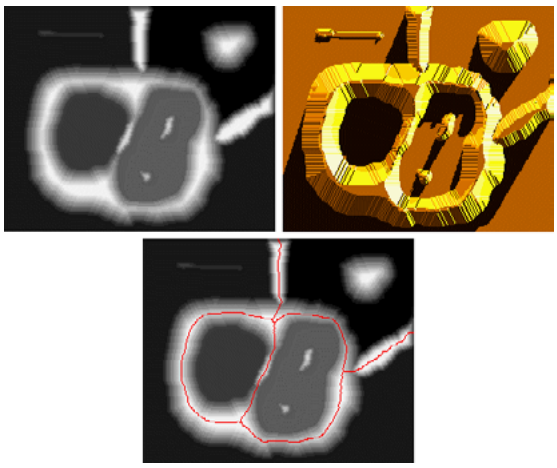


Fig. 3. Illustration of watershed segmentation: (top-left) gray scale image, (top-right) topographic surface, (bottom) watersheds in red.

Fernand Meyer [16] introduced an algorithm based on a set of defined markers in order to overcome the problem of over segmentation which can be summarized as follows:

1. Each of the markers is given a different label.
2. The pixels around each marked area are inserted into a priority queue with priority level that corresponds to the gradient magnitude of the pixel.
3. The pixel having the lowest priority is extracted from the queue and if all its neighbors that have already label have the same label, then it will be labeled with their label. The non-marked neighbors that are not in the priority queue are put in it.
4. Redo step 3 until the queue is empty.

The resulted image will contain non-labeled pixels which correspond to watershed lines.

B. Local Vision

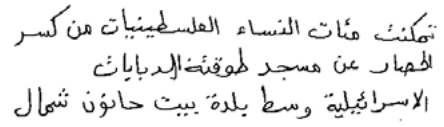
As stated in [22], the idea of using watershed transform comes from the observation of its effect on binary images. If the text components are set as marker, the resulted regions of watershed allow them to occupy a region which adapts according to their size and shape. These regions can help in localizing the neighboring components that may belong to the same line.

Since the approach is based on the analysis of the connected component in order to extract the lines, diacritical points may influence the results of the analysis and therefore should be removed before applying watershed transform. Local linear regions are detected by analyzing the spatial relationship of each resulted region with its neighbors and if one of the neighboring text components is located in its field of view then it will be linked to it by a line thus resulting in a new image where all neighboring text components are linked together. Finally watershed transform is applied again on this new image which results in detecting regions of the text lines.

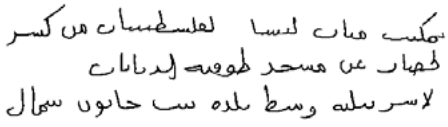
1. Diacritical Points Removal

The proposed approach is sensitive to the presence of diacritical points since it is based on the analysis of the arrangement of the text component. So as a preprocessing step to remove these points, we use the technique in [2] which is based on the estimation of the stroke width of text component. Fig. 4 illustrates the document without diacritics.

After locating the lines, these points will be assigned to their corresponding words.



(Original image)



(Resulted image)

Fig. 4. Result of diacritical points removal.

2. Finding Local Linear Regions

After removing the diacritical points, we set the markers as text components and apply Watershed Transform and in order to prevent the flooding of the edge regions in the corners of the image (Fig. 5.a), we set a new flooding area represented by the inverse of the convex hull of all the text components. The result of this process is shown in Fig. 5.b.

(a)

(b)

Fig. 5. Result of watershed transform on binary document.

One of the characteristic of the watershed transform is that it keeps the alignment of the text lines. For each region, the analysis of the spatial arrangement of text components that may participate to the same line is restricted only to the neighboring regions which correspond to those intersecting the contour of the region in question after dilating its contour with a disk of two pixels. Fig. 6 illustrates an example of neighboring regions detection.

Fig. 6. Example of localizing the neighboring regions of the text component in the middle.

For any linear writing, words and characters in a local region are more or less aligned, and this remains valid even for documents with curved lines. Taking into consideration this fact, the idea is to search, for each text component, among its neighbors those that intersect with its field of view. As for the example in Fig. 7 the blue rectangle corresponds to the current components and the green rectangles correspond to text components that intersect the field of view and therefore are probably located in the same line as the current component.

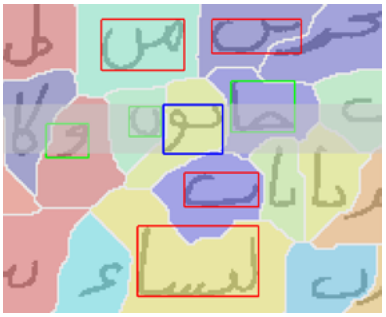


Fig. 7. Example of detection of the neighboring text components included in the field of view of the text component in the middle.

Applying this criterion as it is on the whole components, gives errors in the case of narrow gaps between adjacent lines where the bounding box of the neighboring component in the adjacent lines may intersect with the field of view of the component in question. This is illustrated in Fig. 8.

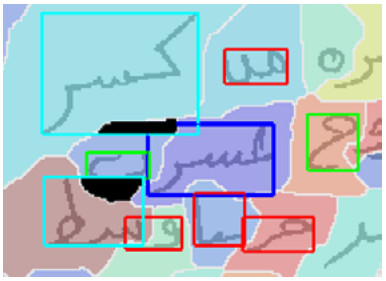


Fig. 8. Example of removing text component in the adjacent lines (black regions correspond to intersection of the components in cyan with others in the neighborhoods).

In order to resolve this issue, the idea is to remove text component in the neighboring regions with bounding box that contains portion of other text component region that intersect the field of view. In the case where tow text components participating in the field of view and intersecting each other we keep them both.

As shown in Fig. 9, other problems occur where text parts may be located in a place which makes it participate at the same time in more than on field of view thus allowing to link component in adjacent lines which result in merging the two lines together. Thus to resolve this, the component will be linked to the component in which it participate the most based on the percentage of participation described in [13].

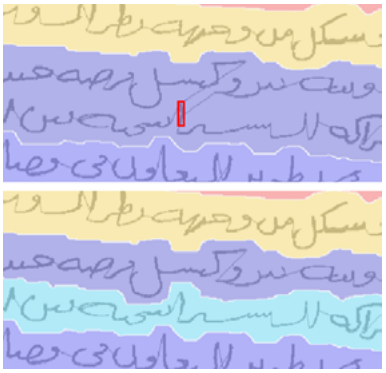


Fig. 9. Example of detection of text part (red rectangle) that participate in two adjacent lines.

3. Extraction of the Lines

As a result of finding local linear regions, we get a vector that contains for each text component the local ones that participate to the same line. A recursive function is used on this vector in order to find all components that are related to each other that will be linked together from their centroids by a line. Fig. 10 illustrates this process. Finally the watershed transform is again applied to the new image in order to estimate the location of the lines.

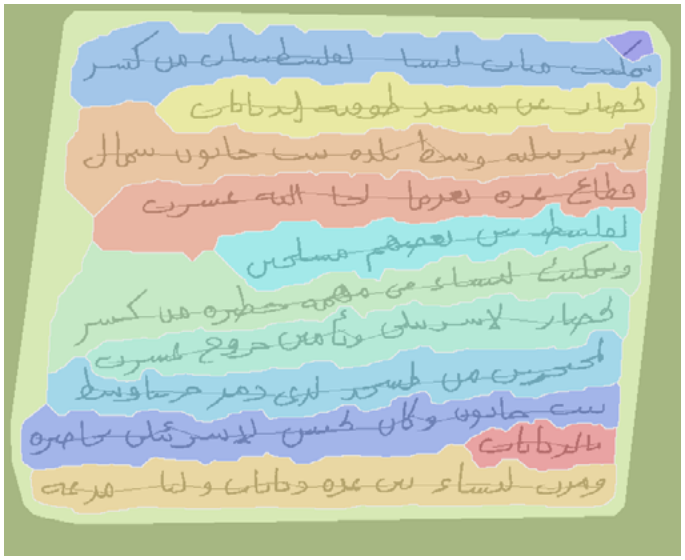


Fig. 10. Example of text lines detection on binary documents.

In some cases parts of characters or big diacritics may be located in interline and therefore remain isolated (not linked to any other component). Based on their width and height which is often smaller than the average words width and the average line height, these isolated components are detected and reassigned to the nearest line (Fig. 11).

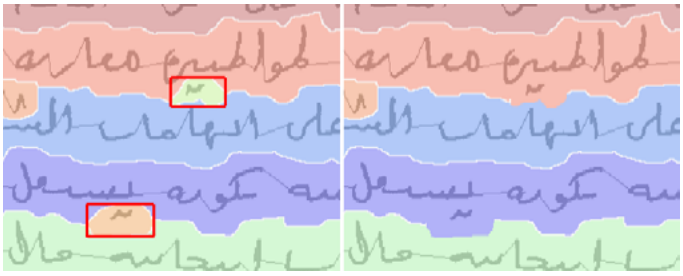


Fig. 11. Example of detection of diacritical points and broken characters (red rectangles).

C. Global Vision

We found interesting to investigate the use of horizontal projection profile as a global technique for the approximation of the base lines and then applying watershed transform on the extracted paths in order to extract the lines.

1. Projection Profile

One of the most used methods in text segmentation is called Projection Profile. Horizontal projection profile produces an histogram that represents for each line in the image: the number of black pixels [17, 18], the number of transitions black to white [19]. Locations of the maxima and minima values are detected, then the space between two consecutive minimums correspond to the location of the text line. This technique is well adapted for printed documents that contain straight lines. But for handwritten documents that present curved and short lines, the peaks do not reflect the location of the lines, also the presence of diacritical points in the Arabic document makes this technique too sensitive to this kind of documents (Fig. 12).

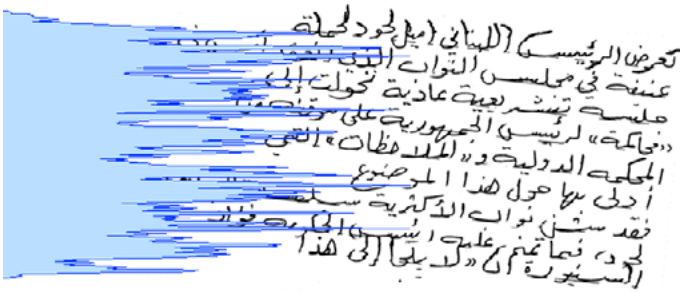


Fig. 12. Example of the histogram (in blue) of the horizontal projection profile.

In order to cope with the problem of curved lines, the idea is to divide the document into vertical slices and then analyzing each slice separately [20, 8].

2. Estimation of the Baseline

The first stage in [8] detects the medial seams (approximation of the orientation of each text line) using a projection profile matching approach. After splitting the page into slices a smoothed horizontal projection profiles is computed for each slice independently. Local maxima in two consecutive slices are matched in both directions, if a matching is found between two peaks a line is drawn between them, thus creating for each line a curve that goes through its peaks.

Due to the presence of diacritical points and ascenders and descenders, the histogram of the horizontal projection profile may contain lot of peaks, so in order to have a histogram that reflects the number and position of the lines, as in [8], we apply a cubic spline that smooths the histogram. Example of this process in shown in Fig. 13.

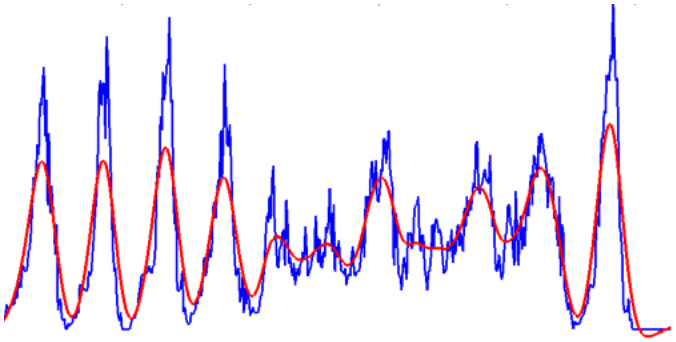


Fig. 13. Example of the histogram of the horizontal projection profile (in blue) and its smoothed version (in red).

We adapt the method in [8] by changing the direction of the matching between peaks of the slices in order to take into consideration the writing direction of the Arabic script which is from right to left, also we ignore the step of removing lines that start from intermediate column since in Arabic documents we may face this kind of situation. The step of extending the small lines is modified by simply taking the last coordinate of each line and creating a straight line between it and the corresponding coordinate in the last column of the image.

After extracting the diacritical points, the horizontal projection profile is applied as shown in Fig. 14.

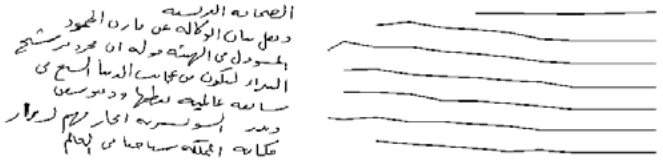


Fig. 14. Example of lines detection using matching horizontal projection profile.

3. Extraction of the Lines

The curves of the projection profile are dilated vertically and marked as locations for the watershed transform where the flood must start. Finally, the text lines correspond to the regions of the watershed transform. The result of this process is illustrated in Fig. 15.

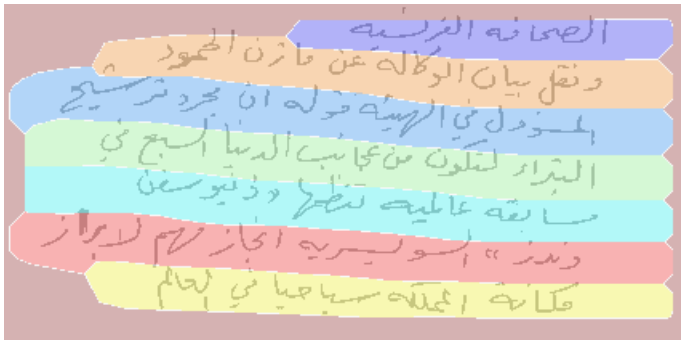
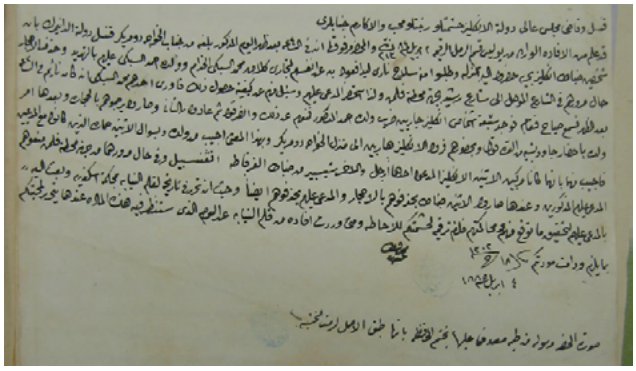
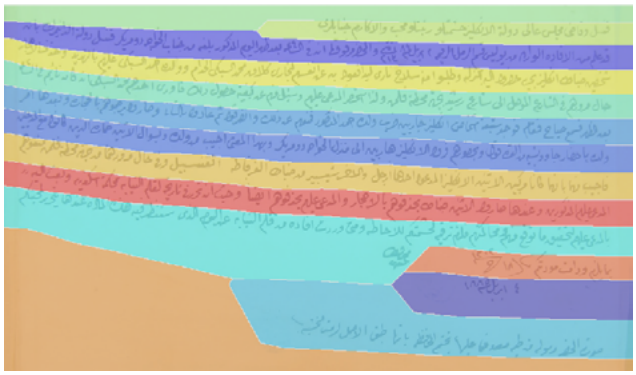


Fig. 15. Result of watershed transform on the projection profile lines superposed on the original image.

One of the advantages of this approach is its ability to work on grayscale documents. As illustrated in Fig. 16, the method detects efficiently the line event in the presence of noise.



(a)



(b)

Fig. 16. Example of text lines detection on gray level documents.

III. Comparative Analysis

In order to assess the accuracy of the proposed approach we use a subset of the publicly available data set "Handwritten Arabic Proximity Datasets" [21]. The F1-mesure score explained in [13] is used here too. Table I compares the score of the proposed method with previously proposed methods tested on the same samples.

TABLE I. Results of the Segmentation Rate using F1-score

Methods	F1-mesure
The method in [13]	90.5%
The method in [2]	94.3%
Watershed as pre-processing	89.4%
Watershed as post-processing	93.3%

As can be observed from the result it is clear that using watershed transform as post-processing step to the horizontal projection profile gives better result than using it as pre-processing step to the analysis of the linearity of connected component. This could be justified by the fact that the projection profile treats the document from a global vision by analyzing the peaks of text lines as if it searches the useful part of words which result in an approximation of the baseline. While by analyzing the connected components, the document is treated from a local vision and thus when constructing the lines we may deviate and this is mainly caused by the presence of touching lines where a given component may be linked to other, which is located in adjacent lines.

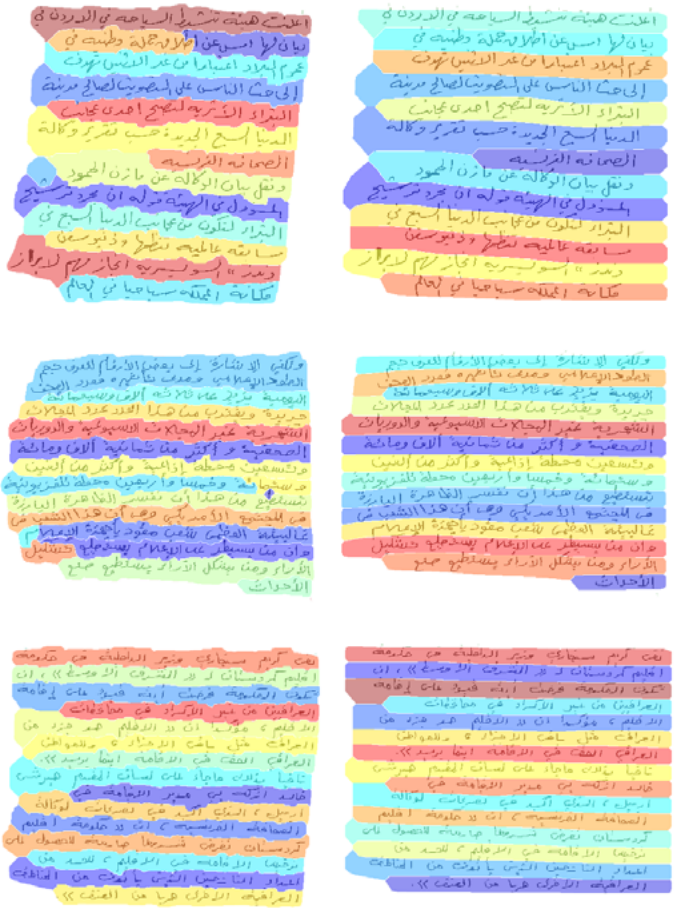
Another advantage of this approach is to treat an also gray level document which is benefic in the case of the presence of complicated noise that could not be removed in the binarized version of the document. Meanwhile, the drawback is that the approximation of the line is global; therefore, some parts of words may be broken in the extracted lines.

Table II recaps the main differences of the two approaches.

TABLE II. DIFFERENCES BETWEEN USING WATERSHED TRANSFORM

	As pre-processing	As post-processing
Gary scale or Binary	Binary	Both
Sensitive to noise	Yes	No
Sensitive to touching lines	Yes	No
Broken characters	No	Yes
Multi-languages	Yes	Yes
Simple and effective	No	Yes

Fig. 17 illustrates some results of the two approaches for text lines detection from Arabic handwritten documents.



(a) (b)

Fig. 17. Examples of handwritten text lines detection using the two proposed approaches using watershed transform as preprocessing (a) and as pos-processing (b).

IV. CONCLUSION

This paper investigates the use of watershed transform in the context of text lines segmentation from handwritten Arabic documents. The first approach treats the document from a local perspective by analyzing the arrangement of the connected component locally; while the second approach treats it from a global perspective by estimating the baseline. Tested on the same samples, the second approach even if it is coarse achieved better results and showed its less sensitivity to the presence of noise and touching lines.

As future work, we are willing to enhance the approach in order to prevent the broking of the characters and also to be able to treat historical and complex documents.

REFERENCES

[1] Likforman-Sulem, L., Zahour, A., & Taconet, B. (2007). Text line segmentation of historical documents: a survey. International Journal on Document Analysis and Recognition, 9(2), 123-138.

[2] Boulid, Y., Souhar, A., & Elkettani, M. Y. (2016) Segmentation approach of Arabic manuscripts text lines based on multi agent systems. International Journal of. Computer Information Systems and Industrial Management, 8, 173–183.

[3] Khayyat, M., Lam, L., Suen, C. Y., Yin, F., & Liu, C. L. (2012, March). Arabic handwritten text line extraction by applying an adaptive mask to morphological dilation. In Document Analysis Systems (DAS), 2012 10th IAPR International Workshop on (pp. 100-104). IEEE.

[4] Nikolaou, N., Makridis, M., Gatos, B., Stamatopoulos, N., & Papamarkos, N. (2010). Segmentation of historical machine-printed documents using Adaptive Run Length Smoothing and skeleton segmentation paths. Image and Vision Computing, 28(4), 590-604.

[5] Shi, Z., & Govindaraju, V. (2004). Line separation for complex document images using fuzzy runlength. In Document Image Analysis for Libraries, 2004. Proceedings. First International Workshop on (pp. 306-312). IEEE.

[6] Stamatopoulos, N., Gatos, B., Louloudis, G., Pal, U., & Alaei, A. (2013, August). ICDAR 2013 handwriting segmentation contest. In Document Analysis and Recognition (ICDAR), 2013 12th International Conference on (pp. 1402-1406). IEEE.

[7] Saabni, R., Asi, A., & El-Sana, J. (2014). Text line extraction for historical document images. Pattern Recognition Letters, 35, 23-33.

[8] Arvanitopoulos, N., & Süsstrunk, S. (2014, September). Seam carving for text line extraction on color and grayscale historical manuscripts. In Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on (pp. 726-731). IEEE.

[9] Shi, Z., Setlur, S., & Govindaraju, V. (2009, July). A steerable directional local profile technique for extraction of handwritten arabic text lines. In Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on (pp. 176-180). IEEE.

[10] Pastor-Pellicer, J., Afzal, M. Z., Liwicki, M., & Castro-Bleda, M. J. (2016, April). Complete System for Text Line Extraction Using Convolutional Neural Networks and Watershed Transform. In Document Analysis Systems (DAS), 2016 12th IAPR Workshop on (pp. 30-35). IEEE.

[11] Kumar, J., Abd-Almageed, W., Kang, L., & Doermann, D. (2010, June). Handwritten Arabic text line segmentation using affinity propagation. In Proceedings of the 9th IAPR International Workshop on Document Analysis Systems (pp. 135-142). ACM.

[12] Kumar, J., Kang, L., Doermann, D., & Abd-Almageed, W. (2011, September). Segmentation of handwritten textlines in presence of touching components. In Document Analysis and Recognition (ICDAR), 2011 International Conference on (pp. 109-113). IEEE.

[13] Boulid, Y., Souhar, A., & El Kettani, M. E. Y. (2016). Detection of Text Lines of Handwritten Arabic Manuscripts using Markov Decision Processes. International Journal of Interactive Multimedia and Artificial Intelligence, 4(1), 31-36.

[14] Razak, Z., Zulkiflee, K., Idris, M. Y. I., Tamil, E. M., Noor, M. N. M., Salleh, R., ... & Yaacob, M. (2008). Off-line handwriting text line segmentation: A review. International Journal of Computer Science and Network Security, 8(7), 12-20.

[15] Oh, K., Kim, S., Na, I., & Kim, G. (2014). Text Line Segmentation using AHTC and Watershed Algorithm for Handwritten Document Images. International Journal of Contents, 10(3), 35-40.

[16] Meyer, F. (1994). Topographic distance and watershed lines. Signal processing, 38(1), 113-125.

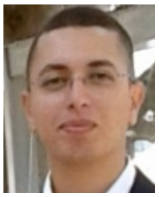
[17] Bennasri, A., Zahour, A., & Taconet, B. (1999). Extraction des lignes d'un texte manuscrit arabe. In Vision Interface (Vol. 99, pp. 42-48).

[18] Nicolaou, A., & Gatos, B. (2009, July). Handwritten text line segmentation by shredding text into its lines. In Document Analysis and Recognition,

2009. ICDAR'09. 10th International Conference on (pp. 626-630). IEEE.
[19] Marti, U. V., & Bunke, H. (2001). On the influence of vocabulary size and language models in unconstrained handwritten text recognition. In Document Analysis and Recognition, 2001. Proceedings. Sixth International Conference on (pp. 260-265). IEEE.
[20] Zahour, A., Likforman-Sulem, L., Boussellaa, W., & Taconet, B. (2007, September). Text line segmentation of historical Arabic documents. In Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on (Vol. 1, pp. 138-142). IEEE.
[21] Handwritten Arabic Proximity Datasets. Language and Media Processing Laboratory. http://lampsrv02.umiacs.umd.edu/projdb/project.php?id=65.
[22] Boulid, Y., Souhar, A., Ameur, Elb. & Ouagague, Mly. M. (2017) Watershed transform for text lines extraction on binary Arabic handwritten documents. In Proceedings of BDCA'17, Tetouan, Morocco, March 29-30, 2017, 6 pages. https://doi.org/10.1145/3090354.3090444
[23] Boulid, Y., Souhar, A., & Elkettani, M. E. Multi-agent Systems for Arabic Handwriting Recognition. International Journal of Interactive Multimedia and Artificial Intelligence, (2017, In Press), http://dx.doi.org/10.9781/ijimai.2017.03.012.
[24] Boulid, Y., Souhar, A., & Elkettani, M. E. (2017). Handwritten Character Recognition Based on the Specificity and the Singularity of the Arabic Language. International Journal of Interactive Multimedia and Artificial Intelligence,4(4), 45-53.

Abdelghani Souhar

Abdelghani Souhar is a full Professor of computer science at the University of Ibn Tofail, Faculty of science Kenitra Morocco. He received the M.S. degree in applied Mathematics in 1992, PhD degree in computer science in 1997 from the University of Mohammed 5 in Rabat-Morocco. His habilitation thesis concerned Modeling complex systems As Arabic handwritten recognition and an intelligent system for generating mesh.
His research interests include automatic processing of Arabic language, Modeling complex systems, CAE/CAD, and Artificial Intelligence.



Youssef Boulid

Youssef Boulid received his M.S. degree in Decision Support Systems and Project Management in 2012 and PhD degree in Computer Science in 2016 from University Ibn Tofail, Faculty of science, Kenitra-Morocco. Now he works as a researcher at Maxware Technology Kénitra – Morocco.
His research interests include image processing, handwritten document analysis, Arabic handwritten recognition and artificial intelligence.

El Bachir Ameur

El Bachir Ameur is a full Professor of computer science at the University of Ibn Tofail, Faculty of science Kenitra (Morocco). In 2002 he received the Ph. D. degree in numerical analysis and computer science from the University of Mohamed I Oujda (Morocco).
His Ph. D. concerned approximation of curves and surfaces by spline and wavelet functions.
His research interest concerns approximation and reconstruction of 2D/3D surfaces by spline and wavelet, signal and image processing, watermarking and steganography.

Mly Moustafa Ouagague

Mly Moustafa Ouagague received his M.S. degree in Decision Support Systems in 2001 from Blaise Pascal University. Now he works as a researcher at Maxware Technology Kénitra - Morocco.
His research interests include artificial intelligence, complex systems, image processing and serious games.

Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S)

Martin Schrepp[1], Andreas Hinderks[2], Jörg Thomaschewski[3]*

[1] SAP AG (Germany)
[2] University of Seville (Spain)
[3] University of Applied Sciences Emden/Leer (Germany)

Received 19 June 2017 | Accepted 12 July 2017 | Published 11 September 2017

unir
LA UNIVERSIDAD EN INTERNET

Abstract

The user experience questionnaire (UEQ) is a widely used questionnaire to measure the subjective impression of users towards the user experience of products. The UEQ is a semantic differential with 26 items. Filling out the UEQ takes approximately 3-5 minutes, i.e. the UEQ is already reasonably efficient concerning the time required to answer all items. However, there exist several valid application scenarios, where filling out the entire UEQ appears impractical. This paper deals with the creation of an 8 item short version of the UEQ, which is optimized for these specific application scenarios. First validations of this short version are also described.

Keywords

User Experience, UEQ, Questionnaire, Short Version.

DOI: 10.9781/ijimai.2017.09.001

I. Introduction

The appearance of modern devices that offer quite natural and easy-to-learn interactions, as for example smart phones or tablets, have taken the general expectation of users concerning user experience of user interfaces to a new level. Today's users simply expect a high level of satisfaction during their interaction with a user interface, even if it is a complex business application or a programming environment.

In order to be successful in a highly competitive market environment, it is thus no longer sufficient to offer products of new and powerful functionality. Users also expect that they can learn how to use the application without much effort, solve their tasks fast and efficiently, and are able to control the interaction at each point. In addition to these goal-oriented interaction qualities, it is also important that the product catches the user's attention and interest and that using the product is interesting and stimulating. Consequently, hedonic, not directly goal-oriented interaction qualities have to be considered as well in order to be successful [1].

For example, in a study concerning business software [2] it was shown that pragmatic quality aspects and hedonic quality aspects equally influence the attractiveness and preference for a product. Thus, user experience with all its facets is an important aspect that must be considered during product design and as a part of quality control.

This raises the question of how to measure user experience. All aspects of user experience are highly subjective evaluations. A product that is seen as easy to learn and understand by one person can be judged as quite complicated and difficult to learn by another person. This can, for example, be due to different levels of expertise or knowledge. Another reason can be a different level of expertise with similar products.

* Corresponding author.
E-mail address: joerg.thomaschewski@hs-emden-leer.de

The same is true for the perceived performance of a product. A product that a user perceives as slow and annoying can be seen as performing adequately by another user. In this respect, users vary widely in their expectations and personal preferences. Thus, any measurement of user experience must consider the feedback of a representative and large enough group of users.

Therefore, questionnaires are a simple method to collect such user feedback [3]. They can be distributed rather efficiently to larger groups of users, especially if they are designed as online questionnaires. In addition, analyzing the numerical data from such questionnaires is highly standardized and thus efficient as well.

In this paper we describe the design and evaluation of a short version of the User Experience Questionnaire, which is a widely used tool to measure user experience.

II. The User Experience Questionnaire (UEQ)

The objective of the UEQ is to allow a quick assessment done by end users covering a preferably comprehensive impression of user experience. It should allow the users to express feelings, impressions, and attitudes that arise when experiencing the product under investigation in a very simple and immediate way.

The UEQ can be used as a paper-pencil version, but is also short enough to be used as an online questionnaire. It consists of 26 items (Fig. 1) that are grouped into 6 scales.

Each item of the UEQ consists of a pair of terms with opposite meanings, for example:

Not understandable o o o o o o o *Understandable*

Efficient o o o o o o o *Inefficient*

Participants rate each item on a 7-point Likert scale. The answers are scaled from -3 (fully agree with negative term) to +3 (fully agree with positive term). Half of the items start with the positive term, the others with the negative term (in randomized order).

The original version of the UEQ was designed in German [4], [5], but has so far been translated to several languages like Spanish [7] and Portuguese [11]. The English version of the UEQ is shown in Fig. 1.

annoying	o o o o o o o	enjoyable	1
not understandable	o o o o o o o	understandable	2
creative	o o o o o o o	dull	3
easy to learn	o o o o o o o	difficult to learn	4
valuable	o o o o o o o	inferior	5
boring	o o o o o o o	exciting	6
not interesting	o o o o o o o	interesting	7
unpredictable	o o o o o o o	predictable	8
fast	o o o o o o o	slow	9
inventive	o o o o o o o	conventional	10
obstructive	o o o o o o o	supportive	11
good	o o o o o o o	bad	12
complicated	o o o o o o o	easy	13
unlikable	o o o o o o o	pleasing	14
usual	o o o o o o o	leading edge	15
unpleasant	o o o o o o o	pleasant	16
secure	o o o o o o o	not secure	17
motivating	o o o o o o o	demotivating	18
meets expectations	o o o o o o o	does not meet expectations	19
inefficient	o o o o o o o	efficient	20
clear	o o o o o o o	confusing	21
impractical	o o o o o o o	practical	22
organized	o o o o o o o	cluttered	23
attractive	o o o o o o o	unattractive	24
friendly	o o o o o o o	unfriendly	25
conservative	o o o o o o o	innovative	26

Fig. 1. English version of the UEQ.

The original German version of the UEQ was designed using a data analytics approach to ensure the practical relevance of the constructed scales. Each scale represents a distinct UX quality aspect.

An initial set of more than 200 potential items related to UX was created with usability experts in two brainstorming sessions. A number of these experts then reduced the selection to a raw version of 80 items.

The raw version was used in several studies (with a total of 153 participants) on the quality of interactive products, including a statistics software package, cell phone address books, online collaboration software, or business software. Finally, the scales and the items representing each scale were extracted from this data set by principal component analysis [4], [5].

This analysis yielded the final questionnaire with 26 items arranged into six scales:

- *Attractiveness*: Overall impression of the product. Do users like or dislike it? Is it attractive, enjoyable or pleasing?
- *Perspicuity*: Is it easy to get familiar with the product? Is it easy to learn? Is the product easy to understand and unambiguous?
- *Efficiency*: Can users solve their tasks without unnecessary effort? Is the interaction efficient and fast? Does the product react to user input quickly?
- *Dependability*: Does the user feel in control of the interaction? Can he or she predict the system's behavior? Does the user feel confident when working with the product?
- *Stimulation*: Is it exciting and motivating to use the product? Is it enjoyable to use?
- *Novelty*: Is the product innovative and creative? Does it capture the user's attention?

Scales are not assumed to be independent. In fact, a user's general impression is recorded by the *Attractiveness* scale, which should be influenced by the values on the other 5 scales (see Fig. 2).

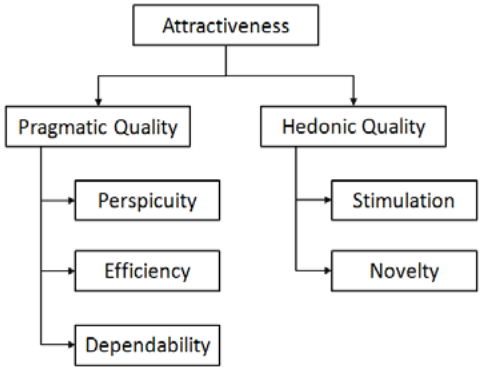


Fig. 2. Assumed scale structure of the User Experience Questionnaire (UEQ).

Attractiveness is a pure valence dimension (emotional reaction on a pure acceptance/rejection dimension). *Perspicuity*, *Efficiency,* and *Dependability* are pragmatic quality aspects, i.e. they describe interaction qualities that relate to the tasks or goals the user aims to reach when using the product. *Stimulation* and *Novelty* are hedonic quality aspects, i.e. they do not relate to tasks and goals, but describe aspects related to pleasure or fun while using the product [2], [6].

For details concerning the design and validation of the UEQ see [4], [5]. Helpful hints on using the UEQ are also available in [7], [8]. There is a benchmark available as well, which is described in [12].

For a semantic differential like the UEQ, it is very important that participants can fill it out in their native language. Thus, several contributors created a number of translations.

The UEQ is currently available in 19 languages (German, English, French, Italian, Russian, Spanish [7], Portuguese [11], Turkish, Chinese, Japanese, Indonesian [10], Dutch, Estonian, Slovene, Swedish, Greek, Polish, Hindi, and Bulgarian).

The UEQ in all available languages, an Excel sheet for data analysis, and the UEQ Handbook are available free of charge at www.ueq-online.org.

III. Scenarios Requiring a Short Version?

Usually, 3-5 minutes are sufficient to fill out the UEQ including some demographic data [3]. Thus, the UEQ is already a quite efficient method to capture the opinion of a user towards the user experience of a product, leading to the obvious question why a shorter version is needed at all?

In the last couple of years we received a number of requests for a shorter version, and some users even created their own short version by removing a few items (which is not a recommended practice for a standardized questionnaire like the UEQ [9]). Accordingly, there seem to be some cases in which a full UEQ is considered to be too time consuming.

All these requests came from three different generic application scenarios in which only a very small number of items could be used to measure user experience.

1. The first scenario is collecting data when the user leaves a web shop or web service. For example, the user has just ordered something in a web shop and logs out. After pressing the log out button, the user is asked to fill out a short questionnaire concerning the user experience of the shop. In such scenarios it is crucial that the user has the impression that filling out the questionnaire can be

done extremely fast. Otherwise, users will refuse to give feedback (they are finished with their initial task and are in the process of leaving the shop, so motivating them to spend some more time on feedback is difficult). Presenting an entire UEQ with all 26 questions in such a scenario will severely reduce the number of users willing to give feedback.

2. In the second and quite frequent scenario, a questionnaire concerning user experience should be included in an already existing product experience questionnaire. Typically, such a questionnaire is sent out after a customer has purchased a product and has already used it for some time.

   Such questionnaires try to collect data about the entire product experience, asking, for example, why the customer chose the product, if the functionality of the product fulfills the expectations, if the purchasing process was pleasant, if the customer wants to be informed about similar or other products of the company in the future, etc. As a result, such questionnaires tend to be quite lengthy. Thus, it is difficult to add a full 26 item user experience questionnaire in such cases.

   On the other hand, it is often not possible to collect data concerning user experience in a separate questionnaire, since the number of customer interactions cannot exceed a certain limit (customers can get easily annoyed if they receive such marketing e-mails too often).

   Thus, including a very short user experience section in such a customer experience questionnaire is often the only way for UX practitioners to collect feedback on their customers' user experience.

3. A third scenario mentioned sometimes are experimental settings where a participant is asked to judge the user experience of several products or variants of a product in one session. In such scenarios the products or product variants are presented to the participant in a random order one after the other, and they have to fill out a questionnaire concerning user experience for each of them. In such a setting, the number of items must be kept to a minimum. Otherwise the participant will be stressed and the quality of answers will decrease quickly.

All of these scenarios share the requirement that the number of items must be small. In addition, any instruction must be simple and quick to read.

IV. Construction of the Short Version

The short version should contain only a limited number of items, but it should still cover the spectrum of product qualities measured by the UEQ.

To shorten the UEQ it was decided to skip the measurement of the single dimensions and to concentrate on the measurement of the two meta-dimensions *pragmatic* and *hedonic quality*. For each of these dimensions four items are chosen. Thus, the short version of the UEQ (henceforth: UEQ-S) will only contain eight items, grouped into two scales. In addition, the mean value of the eight items will be given as an overall UX value.

A data set with 1867 data records was collected by the German UEQ in previous studies. Each data record reflects the evaluation of a product by a participant. In total, 21 different products were assessed (business software, web shops, household appliances, etc.).

A main component analysis was performed on all twelve UEQ items from the *Efficiency*, *Perspicuity* and *Dependability* scales. For the analysis the number of factors was set to 1, and the four items that showed the highest loading on this factor were chosen. These were the items 11, 13, 20, and 21 (see Fig. 1) of the UEQ. They therefore represent the *Pragmatic Quality* scale of the short version UEQ-S.

The same procedure was repeated for the eight UEQ items from the *Stimulation* and *Originality* scales. The items 6, 7, 10, and 15 (see Fig. 1) showed the highest loadings and thus represent the *Hedonic Quality* scale of the UEQ-S.

The UEQ-S thus consists of items shown in Fig. 3:

obstructive	o o o o o o o	supportive	1
complicated	o o o o o o o	easy	2
inefficient	o o o o o o o	efficient	3
clear	o o o o o o o	confusing	4
boring	o o o o o o o	exiting	5
not interesting	o o o o o o o	interesting	6
conventional	o o o o o o o	inventive	7
usual	o o o o o o o	leading edge	8

Fig. 3. The short version UEQ-S.

The first four items represent the pragmatic quality scale and the last four items the hedonic quality scale.

To check the cross-loadings the data set was reduced to these eight items and a main component analysis (varimax rotation) was performed with two factors. Table 1 shows the items' loadings on these factors.

TABLE I. Loadings of the Items of the Short Version UEQ-S on the Factors Pragmatic and Hedonic Quality

Item	Pragmatic	Hedonic
clear / confusing	0.71	0.21
inefficient / efficient	0.63	0.39
complicated / easy	0.79	0.10
obstructive / supportive	0.69	0.41
boring / exiting	0.29	0.74
not interesting / interesting	0.36	0.75
conventional / inventive	0.19	0.82
usual / leading edge	0.19	0.86

Thus, the items show the intended scale structure. Only the item *obstructive/supportive* yields a relevant, but still relatively small cross-loading to the other factor. The other items load strongly on the factor they belong to and only weakly to the other factor.

In the original UEQ half of the items start with the positive term and the other half of items start with the negative term. In addition, the order of the items is randomized in the questionnaire. This was done to be able to detect participants that do not answer seriously [9] and to force users to carefully read the alternatives. However, this also has some disadvantages. The change of polarity must be explained in the instruction and in addition, it is cognitively more demanding for the participants.

In order to simplify the instruction and make it easier to fill in the questionnaire, it was decided that all items have the same polarity. The left side reflects the negative term and the right side the positive term (see Fig. 3). In addition, the order is not randomized: the first 4 items reflect the pragmatic quality and the items 5 to 8 the hedonic quality.

V. Prediction Quality

As a first evaluation we calculated how well the scales of the short version (UEQ-S) approximate the corresponding scales of the full version (UEQ). Therefore, in our data set that was used to design the short version, we calculated the difference between the mean value of all 8 items in the short version and the mean value of all 26 items (here 12 items belong to pragmatic quality, 8 belong to hedonic quality and

6 belong to the scale *Attractiveness*, which is neither pragmatic nor hedonic) in the full UEQ for each participant. The same was done for the scales pragmatic and hedonic quality of the short version.

Regarding the pragmatic quality we calculated the difference between the mean of the four items of the pragmatic quality scale of the UEQ-S and all twelve items of the *Efficiency*, *Perspicuity* and *Dependability* scales in the full UEQ for each participant.

Following that pattern, we compared the difference between the four items of the hedonic quality scale of the UEQ-S with the mean of the eight items of the *Stimulation* and *Originality* scales in the full UEQ.

The distribution of these differences (kernel density plots) is shown in Fig. 4, Fig. 5 and Fig. 6.

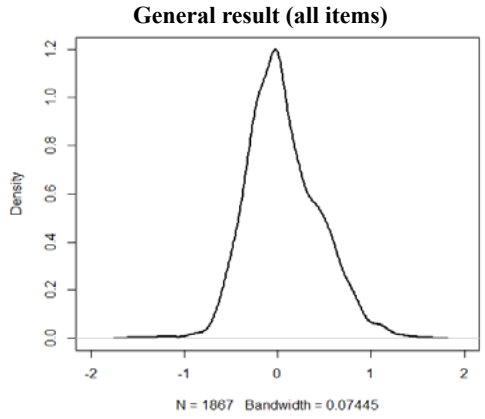


Fig. 4. Distribution of the difference per participant between the full UEQ and the short version UEQ-S for the overall value.

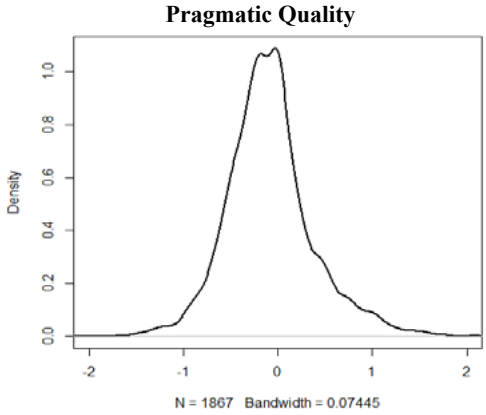


Fig. 5. Distribution of the difference per participant between the full UEQ and the short version UEQ-S for pragmatic quality.

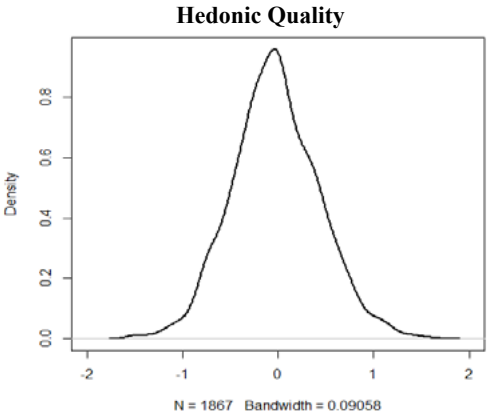


Fig. 6. Distribution of the difference per participant between the full UEQ and the short version UEQ-S for hedonic quality.

The mean and standard deviation of the observed differences are 0.06 (0.39) for all items (Fig. 4), -0.09 (0.46) (Fig. 5) for the items concerning pragmatic quality and -0.03 (0.45) (Fig. 6) for the items concerning hedonic quality. Please note that the UEQ scale ranges from -3 to +3, so these differences concerning the scale means are quite small.

In all three cases the distribution of the observed differences is nearly symmetrical around zero, thus there is no systematic over- or underestimation based on the reduced number of items in the short version UEQ-S. It is evident that the short version is able to predict the values of the full version quite accurately.

VI. An Evaluation Study

In a first study with the short version 47 students judged the user experience of different well-known products. Each student could choose to judge either Amazon, Skype or Wikipedia with an online version of the UEQ-S. We only report results for Amazon, since for the other two products there were simply not enough data to draw meaningful conclusions.

The consistency of the pragmatic quality and hedonic quality scales was reasonably high. The corresponding Cronbach Alpha values were 0.85 (pragmatic quality) and 0.81 (hedonic quality).

The scale means for Amazon (N=31 students decided to judge Amazon) were 1.09 for pragmatic and 0.51 for hedonic quality. These values are quite similar to the values obtained in an older study with a similar target group (German students) and the full UEQ. In this study the mean value for the *Efficiency*, *Perspicuity* and *Dependability* UEQ scales was 1.17 and the mean value for *Stimulation* and *Originality* was 0.66. Thus, the short version UEQ-S seems to approximate the long version expectedly well.

A main component analysis of this data set shows the expected factor structure once more. The loading of the items of the two extracted factors (factors were extracted according to the Kaiser-Guttman criterium, loadings after varimax rotation) are shown in Table 2.

TABLE II. Factor Loadings on the Two Extracted Factors of the Short Version UEQ-S for the Full Data Set of the Study (47 Participants)

Item	Factor 1 pragmatic quality	Factor 2 hedonic quality
clear / confusing	**0.825**	-0.209
inefficient / efficient	**0.833**	0.255
complicated / easy	**0.849**	-0.033
obstructive / supportive	**0.789**	0.215
boring / exiting	0.274	**0.701**
not interesting / interesting	0.220	**0.806**
conventional / inventive	-0.047	**0.841**
usual / leading edge	-0.221	**0.809**

VII. Language Versions

The items of the UEQ-S are a subset of the UEQ items. Accordingly, all the available translations of the UEQ can be used, i.e. the UEQ-S is directly available in all languages for which a full version exists (German, English, French, Italian, Russian, Spanish, Portuguese, Turkish, Chinese, Japanese, Indonesian, Dutch, Estonian, Slovene, Swedish, Greek, Polish, Hindi, and Bulgarian).

The translated version of the UEQ-S can simply be created by choosing the corresponding items from the full UEQ of the desired language.

However, the question remains if the selected items predict the

behavior of the full UEQ as well as the German version. We cannot verify this for all translations yet, since we do not have access to sufficiently large data sets for all of them. So far, this is only possible for some languages.

Table 3 shows the measured deviation per participant between the short version and the corresponding values for the long version for five languages (same computation method as described in the validation of the German short version above).

TABLE III. Deviations between Short Version UEQ-S and Full Version

Language	All	Pragmatic	Hedonic	n
Portuguese	0.19 (0.33)	0.07 (0.41)	0.13 (0.41)	206
Turkish	0.13 (0.39)	-0.06 (0.44)	0.22 (0.47)	943
Spanish	0.16 (0.38)	0.02 (0.49)	0.1 (0.38)	377
Indonesian	0.14 (0.33)	-0.15 (0.38)	0.25 (0.38)	212
English	0.15 (0.34)	0.03 (0.43)	0.17 (0.36)	224

The data shows that for these languages the fit between the short version and the full version of the UEQ is good enough to allow a practical application of the UEQ-S. For example, for the English version we can expect that the mean of the eight items of the UEQ-S deviates 0.15 (on average) from the mean of all 26 UEQ items. For the four UEQ-S items of the pragmatic quality scale the deviation from the mean of all twelve UEQ items of the *Efficiency*, *Dependability* and *Perspicuity* scales is 0.03 on average. For the four UEQ-S items of the hedonic scale the deviation from the mean of all eight UEQ items of the scales *Stimulation* and *Originality* averages at 0.17. Thus, as in the case of the German version, the approximation is quite good.

VIII. Limitations of the Short Version UEQ-S

We described the design of the UEQ's short version UEQ-S. For a UX professional who wants to plan an evaluation the question arises which of the two versions should be used. Obviously, the short version has some advantages concerning the number of questions and accordingly the time the participants need to fill out the questionnaire.

However, this comes at a price. The full UEQ gives a detailed feedback concerning 6 different aspects of UX, i.e. measures on the *Attractiveness*, *Efficiency*, *Perspicuity*, *Dependability*, *Stimulation* and *Originality* scales. This is lost in the short version that only distinguishes between pragmatic and hedonic quality.

Given the fact that even a full UEQ requires only 3-5 minutes, the usage of the UEQ-S should be limited to the scenarios described in the beginning of this paper. The short version should only be used in situations where a full UEQ can not be applied at all. Otherwise, the loss of detailed information is not compensated by saving time in filling out the questionnaire.

IX. Conclusion

We described the design and validation of a short version of the UEQ. It consists of only eight of the 26 items of the UEQ. The short version, which is named UEQ-S, contains two subscales (pragmatic and hedonic quality; 4 items each) and a total value reflecting the overall user experience.

It was shown that the short version is able to predict the behavior of the full version concerning pragmatic and hedonic quality. The mean value obtained by the four items of the short version approximates the values obtained by averaging all 12 pragmatic items (from the *Efficiency*, *Perspicuity* and *Dependability* scales) and all 8 hedonic items (from the *Stimulation* and *Originality* scales) of the full version.

In a first application study concerning Amazon done with German students, the scales showed a high level of consistency. In addition, the measured mean for the pragmatic and hedonic quality approximates the values obtained by the full UEQ collected in a previous study.

The short version UEQ-S is only intended for specific scenarios which do not allow employing a full UEQ. The UEQ-S does not allow measuring the detailed UX qualities *Attractiveness*, *Efficiency*, *Perspicuity*, *Dependability*, *Stimulation* and *Novelty*, which are part of the UEQ report. It is, in general, quite useful to gather these detailed values when it comes to interpreting the results and define areas of improvement [3].

Thus, the short version UEQ-S only allows a rough measurement on higher level meta-dimensions. Our recommendation therefore is to only use the short version UEQ-S in the scenarios described in this paper. The short version should not replace the usage of the full version in standard scenarios, for example after usability tests. In such scenarios, the small gain in efficiency does not compensate for the loss of detailed information on the single scales and therefore more detailed quality aspects.

References

[1] Preece, J.; Rogers, Y. & Sharpe, H. (2002): Interaction design: Beyond human-computer interaction. New York: Wiley.

[2] Schrepp, M.; Held, T. & Laugwitz, B.: The influence of hedonic quality on the attractiveness of user interfaces of business management software. Interacting with Computers 18, pp. 1055-1069 (2006).

[3] Schrepp, M.; Hinderks, A. & Thomaschewski, J. (2014). Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios. In: Marcus, A. (Ed.): Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. Lecture Notes in Computer Science, Volume 8517, pp. 383-392, Springer International Publishing.

[4] Laugwitz, B.; Schrepp, M. & Held, T. (2006). Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. [Construction of a questionnaire for the measurement of user experience of software products] In: A.M. Heinecke & H. Paul (Eds.): Mensch & Computer 2006 – Mensch und Computer im Strukturwandel. Oldenbourg Verlag, pp. 125-134.

[5] Laugwitz, B.; Schrepp; M. & Held, T. (2008). Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (Ed.): USAB 2008, LNCS 5298, pp. 63-76.

[6] Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. International Journal of Human-Computer Interaction, 13(4), pp. 481-499.

[7] Rauschenberger, M.; Schrepp, M.; Perez-Cota, M.; Olschner, S. & Thomaschewski, J. (2013): Efficient Measurement of the User Experience of Interactive Products. How to use the User Experience Questionnaire (UEQ). Example: Spanish Language Version. In: International Journal of Interactive Multimedia and Artificial Intelligence, 2(1), pp. 39-45.

[8] Schrepp, M. (2015). User Experience Questionnaire Handbook. DOI: 10.13140/RG.2.1.2815.0245. Available under: https://www.researchgate.net/publication/281973617_User_Experience_Questionnaire_Handbook.

[9] Schrepp, M.; Hinderks, A. & Thomaschewski, J. (2016). User Experience mit Fragebögen evaluieren - Tipps und Tricks für Datenerhebung, Auswertung und Präsentation der Ergebnisse. S. Hess & H. Fischer (Hrsg.): Mensch und Computer 2016 – Usability Professionals. http://dx.doi.org/10.18420/muc2016-up-0016.

[10] Santoso, H.B.; Schrepp, M.; Isal, R.Y.K.; Utomo, A.Y. & Priyogi, B. (2016). User Experience Questionnaire: Development of an Indonesian Version and Its Usage for Product Evaluation. The Journal of Educators Online-JEO January 2016 ISSN 1547-500X Vol 13 Number 1, pp. 58-79.

[11] Pérez Cota, M.; Thomaschewski, J.; Schrepp, M. & Goncalves, R. (2014). Efficient Measurement of the User Experience. A Portuguese Version. Procedia Computer Science. 5th International Conference on Software Development and Technologies for Enhancing (27), pp. 491-498. DOI: 10.1016/j.procs.2014.02.053

[12] Schrepp, M.; Hinderks, A. & Thomaschewski, J. (2017). Construction of a benchmark for the User Experience Questionnaire (UEQ). International

Journal of Interactive Multimedia and Artificial Intelligence, 4(4), pp. 40-44. DOI: 10.9781/ijimai.2017.445

Martin Schrepp

Martin Schrepp has been working as a user interface designer for SAP AG since 1994. He finished his diploma in mathematics in 1990 at the University of Heidelberg (Germany). In 1993 he received a PhD in Psychology (also from the University of Heidelberg). His research interests are the application of psychological theories to improve the design of software interfaces, the application of Design for All principles to increase accessibility of business software, measurement of usability and user experience, and the development of general data analysis methods. He has published several papers concerning these research fields.

Andreas Hinderks

Andreas Hinderks graduated with a diploma in Computer Science and is Master of Science in Media Informatics by the University of Applied Science Emden/Leer. He has worked as a business analyst and a programmer from 2001 to 2016. Back then, he focused on developing user-friendly business software. Currently, he is a freelancing business analyst and senior UX architect. Furthermore he is a Ph.D. student at the University of Seville, Spain. He is involved in research activities dealing with UX questionnaires, process optimization, information architecture, and user experience since 2011.

Jörg Thomaschewski

Jörg Thomaschewski received a PhD in physics from the University of Bremen (Germany) in 1996. He became Full Professor at the University of Applied Sciences Emden/Leer (Germany) in September 2000. His research interests are Internet applications for human-computer interaction, e-learning, and software engineering. Dr. Thomaschewski is the author of various online modules, e.g., “Human-Computer Communication,” which are used by the Virtual University (online) at six university sites. He has wide experience in usability training, analysis, and consulting.

A System to Generate SignWriting for Video Tracks Enhancing Accessibility of Deaf People

Elena Verdú Pérez[1]*, B. Cristina Pelayo G-Bustelo[2], M. Ángeles Martínez Sánchez[3], and Rubén González Crespo[1]

[1] Universidad Internacional de la Rioja, Logroño (Spain)
[2] University of Oviedo, Oviedo (Spain)
[3] University of Granada, Granada (Spain)

Received 1 June 2017 | Accepted 18 August 2017 | Published 11 September 2017

Abstract

Video content has increased much on the Internet during last years. In spite of the efforts of different organizations and governments to increase the accessibility of websites, most multimedia content on the Internet is not accessible. This paper describes a system that contributes to make multimedia content more accessible on the Web, by automatically translating subtitles in oral language to SignWriting, a way of writing Sign Language. This system extends the functionality of a general web platform that can provide accessible web content for different needs. This platform has a core component that automatically converts any web page to a web page compliant with level AA of WAI guidelines. Around this core component, different adapters complete the conversion according to the needs of specific users. One adapter is the Deaf People Accessibility Adapter, which provides accessible web content for the Deaf, based on SignWritting. Functionality of this adapter has been extended with the video subtitle translator system. A first prototype of this system has been tested through different methods including usability and accessibility tests and results show that this tool can enhance the accessibility of video content available on the Web for Deaf people.

Keywords

Accessibility, Deaf, Machine Translation, Multimedia Content, Sign Language, SignWriting.

DOI: 10.9781/ijimai.2017.09.002

I. Introduction

WEB content, specifically multimedia content, has increased much on the Internet during recent years [1]. New tools facilitating video editing and uploading to Internet have led naturally to the massive generation of videos, many of them of didactic, informative or illustrative nature [2]. Unfortunately, the poor accessibility level of much multimedia content increases the gap existing in access to information between not impaired and impaired people.

Therefore, there are organizations promoting recommendations in order to reach the adequate accessibility level, being the Web Content Accessibility Guidelines (WCAG) 2.0 [3], proposed by the W3C, one of the most important recommendations.

Moreover, in different countries, legislations exist that state that public administration websites, as well as those of other institutions that benefit from public financing, must comply with a certain level of accessibility. For example, in Spain, the websites of public administrations or other institutions receiving public financing or giving general services, must comply with the AA level of the WCAG. In United States, some states have codified laws to ensure web accessibility while others inform about standards and guidelines, applying to state entities [4].

However, in spite of these efforts, most multimedia content on the Internet is not accessible. Developing an accessible website is not only expensive but also difficult because there are some aspects that depend on interpretation. For example, while a textual description of an image could be adequate for a visual impaired person, it could not be appropriate for a person with cognitive problems [5].

In order to face these drawbacks, Crespo et al. [5] propose a new approach allowing the collaborative creation of different accessible adaptations of websites. A step-by-step assistant or wizard guides the semiautomatic adaptation that can be done by any user and that is saved on a repository, in which all adapted profiles of websites are shared.

A web platform that also aims at facilitating the delivery of accessible content via web is presented in [6] and [7]. This platform is a global solution targeting every disability, as different adapters can be integrated to provide accessible content according to the different limitations of users. Specifically the Deaf People Accessibility Adapter has already been developed and described in [7]. This component adapts the application content for people with severe auditory disability by translating standard web applications to web applications based on SingWriting, a way of writing Sign Language. The present paper describes a solution that intends to complement this adapter, eliminating existing barriers in video subtitling. If a web page provides a standard subtitling file, the proposed solution translates the available plain text to vector graphics, representing SingWriting, which accompany the video sequence, allowing the perception of audio information of the video by Deaf people. Specifically, the adapter supports translation of subtitles written in English language into SignWriting for American Sign Language

* Corresponding author.
E-mail address: elena.verdu@unir.net

(ASL), one of the visual-gestural languages most commonly used by the Deaf communities in the United States of America.

Next section of the paper explains reasons why Deaf People experience difficulties when consuming multimedia web content. Besides, some state-of-art solutions are described. Section III gives an overview of the accessibility global platform and the auditory disability accessibility adapter, both mentioned in previous paragraph. Besides, section III presents the solution that makes subtitled videos accessible for the Deaf, whose evaluation is presented in Section IV. Section V presents main conclusions and future works.

II. Context: Problems of Deaf People Accessing Video Content and State-of-Art Solutions

A. SignWriting: a Bridge to Mitigate the Difficulties Experienced by Deaf People

Most multimedia content on the Internet is not accessible for Deaf people, who remain behind the digital barrier. Complex texts and multimedia content without alternative text are the most critical barriers for hearing impaired users according to the study of Pascual et al. [8]. Including captions or transcriptions to a video can make it more accessible [1] but reading oral language is still difficult for some deaf people, especially for those with prelocutive auditory disability, who find many difficulties using web sites. Prelocutive indicates that deafness appears before learning of language. Technical advances such as cochlear implants allow profoundly deaf to hear, as it simulates the function of the inner ear, providing sound signals to the brain. At present, many young deaf people benefitting from cochlear implants are achieving good results when learning oral languages. However not all people can be implanted and not all implanted people obtain an adequate proficiency level [9].

Oral languages are based on letters that represent sounds, which for prelocutive deaf, are completely unknown. Therefore, for these people, learning reading oral languages is merely memorizing sequences of letters [7]. This has consequences such as the fact that deaf children are also affected in their linguistic and, therefore, cognitive development because language is the tool that organizes our thoughts. This is used to describe the concepts we learn and mediates between the subjects and the objects of knowledge [10].

Therefore, for those with auditory disability who cannot benefit from the cochlear implant, a visual language will be their preferred language as it is received through vision, which is the natural communication channel for them. A sing language is a visual-gestural language expressed with hands, body and facial gestures, which is used, among others, by the Deaf. Sign languages allow the Deaf to communicate unhindered, focusing on the message and not on the medium. Moreover, sign languages are organized in the brain as spoken languages. Although many signs mimic the meaning of the concept expressed, the position and movement of the hands for many of them do not reveal the meaning. Therefore, a Deaf person using Spanish Sign Language will not understand someone using American Sign Language, as it happens with spoken language [11].

As well as there is not a unified way of signing in every country, Deaf people have not yet stablished a unified way of writing sign languages. SignWriting is a proposal candidate for that [12]. As above mentioned, reading spoken languages is a very difficult task for the Deaf because written words are made of letters representing sounds, thus learning to read is a task consisting in memorizing. SignWriting represents signs on the signer's perspective, that is, how the signers see their own hands when they sign. Also, it represents head, shoulder and trunk movements, as well as facial expressions, which can change totally the meaning of a sign [13]. Fig. 1 contains the word 'SignWriting' expressed in SingWriting. In this specific case, the facial expression does not take a role when expressing the sing so the graphic does not include the face expression. The upper part of the graphic shows the positions and configuration of hands and the arrows indicate the movement of the hands. Fig. 2 represents the word 'Hello', which includes the smiling facial expression.

As it can be seen, the elements composing SignWriting can be generated as vector graphics, easily treated by current web applications. Therefore, the proposed system described in this paper will use vector graphics to show SignWriting as subtitles of video content, enhancing video visualization experience for Deaf people.

Fig 1. 'SignWriting' sign in SignWriting.

Fig 2. 'Hello' sign in SignWriting.

B. Related Works

There are several systems that translate texts and video subtitles to Sign Language automatically. San Segundo et al. [14] have developed a machine translation system from Spanish to Spanish Sign Language. The system uses speech recognition, which is automatically translated and interpreted by an avatar animation module. Araujo et al. [15] propose a solution that translates video subtitles and generates a sequence in Brazilian Sign Language in real time, respecting linguistic constructions and grammatical rules. Signs are interpreted by an avatar and the authors are working on strategies that allow human collaboration to maintain, improve and extend automatic translations rules of the system.

López-Ludeña et al. [16] propose a bidirectional system that translates speech into sign language and generates speech from sing language. The system recognizes speech and uses a memory-based translation strategy, using a set of sentences in oral language and their translation into sign language as a learning stage that will enable the system to translate similar sentences.

Most of translators to Sign Language correspond to automatic translators focused on specific domains of activity, in which vocabulary and expected sentence constructions are quite limited, such as a hotel reception [16], a bus company [17], renewing the identity card or the driver's license [14], for example.

Another approach which can be applied to open domains and intends to be cost effective, is the one described in [18]. To solve the needs of Deaf people accessing to multimedia content, an architecture is proposed that includes intelligence in the streaming server so that it distinguishes clients' characteristics and properties of communication to provide the most suitable content. Specifically, for clients with hearing disabilities, a middleware is proposed to translate the subtitles tracks to SignWriting. The work described in this paper completes this work, taking elements from this architecture to implement a translator to SignWriting for subtitled videos, complementing the adapter that translates standard web applications to applications based on SignWritting [7], that is, the graphical representation of Sign Language.

III. Solution to Global Accessibility

There is not a universal accessibility, given that the needs of each disability are very different and therefore accessibility requirements are also very different. The work of Crespo et al. [6] [7] describes a solution for those hearing impaired but it intends to be a global solution to accessibility based on two main principles:

- Separating content and presentation: Content has to be firstly presented in a language that can be processed by a machine and not by the user. The purpose is that the user is able to select the representation of the content according to his needs (this can be SignWriting or HamNoSys for example).
- Automatic generation of content in a specific format: According to the targeted user, a specific formatted representation of content is generated. For a hearing impaired user, the representation can be based on sign language symbols coded as SVG format, as in the specific solution described in this paper.

Therefore, the problem of constructing an accessible website is separated into different parts (see Fig. 3). First, there is a core component, which facilitates the collaborative adaptation of websites. This is called the *WAI Accessibility Adapter,* which automatically converts any web application, regardless its accessibility level, to the expected accessibility level. As Fig. 3 shows, this adapter is an intermediate component between the application server and the client. It is totally independent of the specific requirements of the users according to their disability. This component transforms code served by the application server so that it conforms with level AA of WAI guidelines. One of the main problems of implementing accessible websites is covered by this component, that is, facilitating the delivery of accessible web content through a gateway, saving the costs that would involve to make accessible each website separately by each company. However, this component does not allow covering all the problems specific to the different disabilities by itself.

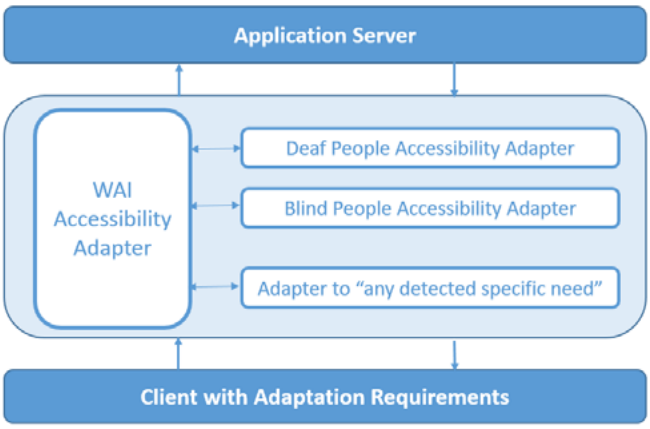


Fig. 3. Overall architecture of the global accessibility solution (Own elaboration based on [7].

The special features required by a specific disability are achieved by different targeted-disability adapters that would act as kind of user agents. Then, there can be a specific component for those with visual disability and another for those auditory impaired. Any specific accessibility requirement should be solved by the corresponding specific adapter.

As the present paper describes a solution to accessibility problems of those with auditory disability, from this point the paper will focus on the Deaf People Accessibility Adapter.

A. WAI Accessibility Adapter

The WAI Accessibility Adapter is an intermediate component between a web application and its clients. When a client makes a request to a web application, instead of directly sending the request to the web application, the request is sent to the WAI Adapter. Then, the adapter sends the request to the Web application and apply some transformations to the content returned so that this content is accessible. The adapter sends the accessible content to the client as response. This adapter has two main components [7]:

- Proxy adaptor: it receives the requests of the clients, forwards them to a specific browser instantiated in the pool component, obtaining the HTML generated from that request by the web application. Then, it applies the transformations to deliver accessible contents to the client.
- Pool of browsers: it receives the proxy requests and instantiates browsers, managing their assignments to specific clients. Instantiated browsers send requests to the web application and receive the corresponding HTML code, which is delivered to the proxy adaptor.

B. Deaf People Accessibility Adapter

The Deaf People Impairment Adapter is a middleware that allows translating typical web applications into SignWriting web applications. This adapter is based on SWMLSVG (SignWriting Markup Language for Scalable Vector Graphics), a language defined to visualize signwriting graphics in any web browser supporting the visualization of SVG.

SWML (SignWriting Markup Language) is based on XML [19]. It is an effort to allow the interoperability of web applications using SignWriting. SWMLSVG language extends SWML to render SignWriting symbol sequences as vector graphics SVG. It allows easy visualization of SignWriting elements in any web browser supporting visualization of SVG graphics.

The Adapter uses SVG format because it is more adaptable than other formats. This development is intended to be a solution for the Deaf but these people can also suffer from partial lack of vision, for example. Therefore, enlarging the size of images without losing quality could be needed for some users. Other benefits of using vector drawings instead of rasters are smaller storage space for simple images or easy edition [7].

The overall architecture of this component in charge of the adaptation of web content to SignWriting is shown in Fig. 4. The HTML document received from the WAI Adapter is processed by the SWMLSVG Adapter. This transforms the input HTML document into a HTML document with SWMLSVG inside. Two databases are used while undertaken the transformation:

- The standard database SSS-XXXX-SVG, which contains, in written form, specifically in SVG format, the different possible configurations a sign can have. SignWriting dictionaries are sorted by Sign-Symbol-Sequence (SSS). Therefore, SSS has the same function for sign languages as the alphabet has for oral languages. A file SSS-XXXX-SVG serves as database of SVG images, which compose the SignWriting alphabet.

- A database (SBML), which is a kind of cache of symbols previously configured and then stored for quick use [7].

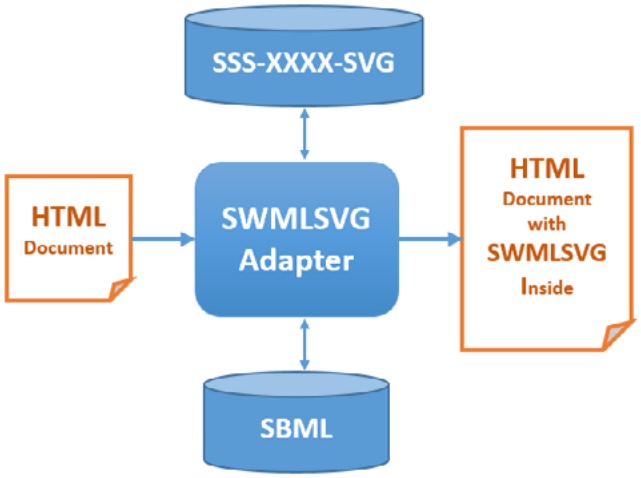


Fig. 4. Overall view of Deaf People Accessibility Adapter (Own elaboration based on [7].

C. SignWriting Subtitling System for Video Streaming

As described previously, many webpages contain video content currently and, therefore, the Deaf People Accessibility Adapter has been extended to support subtitle translations. This section describes this new component of this adapter, a SignWriting subtitling system based on HTML5 <video> tag and Synchronized Multimedia Integration Language (SMIL) Timesheets. When the adapter detects that a webpage contains a video with subtitles, it takes the subtitles expressed in an oral language and generates subtitles in SignWriting format, timely synchronized with the video stream.

This synchronization is achieved thanks to SMIL TimeSheets, which can orchestrate documents compound of HTML and SVG drawings, separating timing and synchronization of elements in the document from content and presentation [20].

SMIL is a W3C standard XML-based language defined to describe the temporal behaviour of a multimedia presentation allowing synchronizing multimedia content [21]. SMIL Timesheets, also proposed by W3C, reuse some SMIL timing primitives but separate content, styling and timing, with the objective of facilitating authoring and handling of multimedia contents. While CSS (Cascading Style Sheets) defines the spatial layout and format of the elements of a web page, SMIL Timesheets play a similar role in the temporal aspect. SMIL Timesheets define which elements are shown at a certain time and can be reused in different documents, as CSS [22].

HTML5 specification introduced the <video> tag, which facilitates video integration into webpages. The video element includes the <track> tag in which subtitles and caption files can be specified so that they are shown while media is playing. The adapter will take the subtitle file referenced in the track element and generate a web document with SVG graphics showing a SignWriting sequence synchronized with video. HTML5 does not support including SVG code in the HTML track element, therefore SMIL Timesheets are used to synchronize vector drawing on web page with video stream. As SMIL is not natively supported by web browsers [20], the use of JavaScript is required to govern the document timing from SMIL data. Specifically the open-source JavaScript library timesheet.js is used to integrate SMIL and render SVG drawings synchronized with video stream, following the approach proposed by Cazenave et al. [20]. For translation of oral language subtitles to SignWriting, the databases used by SWMLSVG adapter shown in Fig. 4 are used (see Fig. 5).

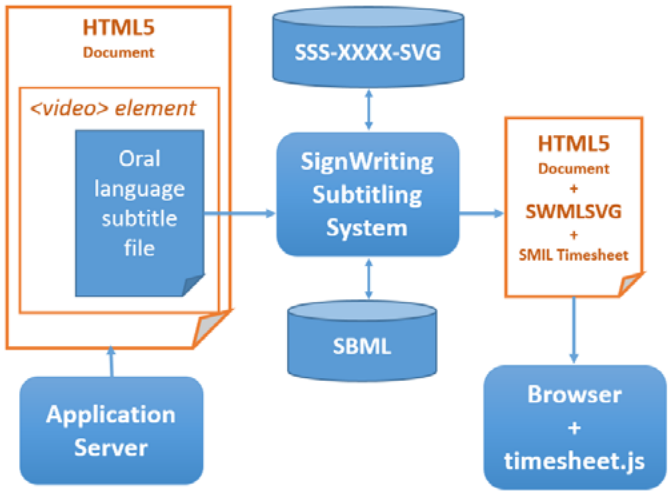


Fig. 5. Overall view of SignWriting Subtitling System.

Currently the first prototype of the SignWriting Subtitling System has been developed and tested. It provides an interface to upload a video and a subtitle file, supporting subtitles expressed in SubRip Text format and written in English language, translating them into SignWriting for American Sign Language (ASL). Fig. 6 shows how SVG graphics are visualized synchronized with video. This first prototype has been tested providing a first and prompt feedback for developers. The experiment undertaken for testing is described in next section. The support of WebVTT is under development and complete integration into Deaf People Accessibility Adapter is in progress.

Fig. 6. SignWriting-subtitled video streaming.

IV. Evaluation

A. Method

The application performance of first prototype of the SignWriting Video Subtitling System has been validated by using different types of tests:

- Unit tests: some unit tests on the most important components of the system are done.
- Usability tests: to evaluate the system by testing on users.
- Accessibility tests: both using an automatic evaluation accessibility tool and manually evaluating with a usability and accessibility checklist.
- Integration tests: these are carried out each time a specific functionality is implemented.

This section specifically describes the methods used for accessibility and usability testing, as these are crucial taken into account the target user and objective of the application, which is improving accessibility to video content of Deaf People.

Two questionnaires were used during usability testing with users. The first one included general questions to know the computer literacy and computer habits of participants in the experiment, that is, if they use computer regularly and the type of activities they do with computers (working, leisure, social networking…).

Therefore, this first questionnaire served to classify the 13 participants according to their computer literacy profile and habits:

- User 1: person in the age range of 20 to 35, regular user of computers and social networks, but without experience in digital certificates. 6 participants belong to this profile.
- User 2: person in the age range of 35 to 55, regular user of computers and Internet for leisure activities, but without experience in social networks and digital certificates. This group is made up of 3 participants.
- User 3: person in the age range of 35 to 55 with little experience using computers, Internet and social networks. This group is made up of 4 participants.

During the next stage, the participants visualized three videos of short duration by using the developed interface. After that, they answered the second questionnaire, which contained 17 1-to-5 rating scale questions to measure the following different usability aspects:

- Ease of use: is browsing simple? Is it easy to access to a specific content? …
- Functionality: are there some missed options? Are tasks done fluently? …
- User interface: is size of signs adequate? Are colors of interface adequate? Is interface nice and clean? Is the application well-structured? Is transition between pages well structured? Is the needed information delivered at each moment? Does the interface help to focus on the task that is being done?

For accessibility evaluation, WAVE accessibility evaluation tool is used [23], which is an online web service that helps to determine the accessibility of web content. This includes many checks for compliance issues contained in the Section 508 [24] and WCAG 2.0 guidelines. As automatic tools cannot check all issues of these guides, moreover only persons can determine true accessibility, the guide of evaluation of usability authored by Hassan and Martín [25], which includes a section to evaluate accessibility, is also used for validation. This is an extensive form that considers different aspects such as identity and information, language and writing, informative labels or signs, or page layout. Tests are done in several browsers such as Internet Explorer, Firefox or Google Chrome.

B. Results

Table 1 shows the average score obtained for each evaluated usability aspect through the questionnaire of 17 questions previously mentioned. The questions were rated with a 1-to-5 rating scale, being 1 the more negative response and 5, the more positive. Therefore satisfaction of participants with every aspect is in general high, being every aspect rated above 4. Fig. 7 shows the average score obtained for each aspect by the different groups of users. Even with a different computer literacy background, responses of the participants belonging to groups User 1 and User 2 were very similar, with a high score for every aspect. As expected due to their lack of experience with Internet and computers, participants of profile User 3 rated lower the different aspects, as shown in Fig 7. Anyway, average scores for the three aspects were higher than 3.5 and no question received a score lower than 3 for this group.

TABLE I. AVERAGE SCORE OBTAINED IN USABILITY QUESTIONNAIRE (1-TO-5 RATING SCALE, 1-STRONGLY DISAGREE AND 5 - STRONGLY AGREE)

FACTOR	Average Score
Ease of use	4.42
Functionality	4.28
User Interface	4.27

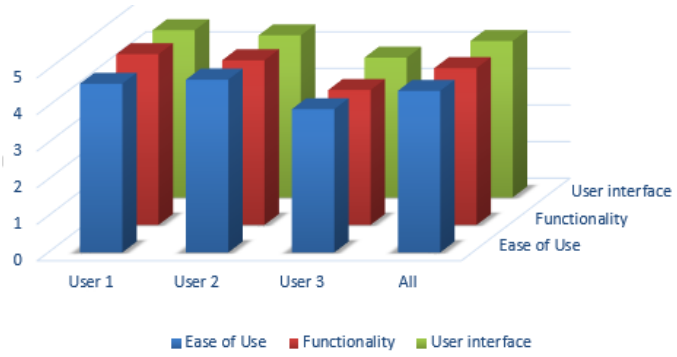


Fig. 7. Average score by factor and type of user obtained in usability questionnaire (1-to-5 rating scale, 1- strongly disagree and 5 - strongly agree).

WAVE tool is also used to check the accessibility of the application and no errors are detected. However, as said before, the absence of errors reported by an automatic tool does not mean that the page is accessible. Moreover some warnings are informed by this tool.

Therefore, accessibility is also verified with the guide of evaluation of usability previously mentioned [25]. All criteria of the checklist of this guide are complied except the following ones:

- There is no a high contrast between the font color (SignWriting graphic) and background.
- Users have not total control over the interface.

The results of usability testing show that participants are satisfied with the SignWriting subtitling tool. Not every word had its translated symbol, and this may be the reason why some users did not give a high score to question "Is the needed information delivered at each moment?". A dictionary generator tool allows to modify and extend dictionaries fluently once symbols are available, so this result can be improved in the future. Besides, as accessibility tests show, the application does not suffer from accessibility problems.

V. CONCLUSION

This paper describes an extension of a Deaf People Accesibility Adaptor that translates oral language subtitles of videos embedded in web pages to SignWriting. This extension is under development, and a first prototype has already been tested to give a prompt feedback to developers. This first test has probed that users are satisfied with the tool and that this potentially will enhance multimedia content browsing experience of Deaf Users.

More tests, with a higher number of participants, will be done when the tool is completely integrated into the Deaf People Accessibility Adapter. The translator failed to translate some words during tests already done, as they were not available in the dictionary. In order to measure these types of errors, next tests will include new metrics such as Sign Error Rate (SER). SER is the percentage of wrong signs in the translation output compared to a reference, which is provided by a human expert [14]. Besides, new recent instruments [26] [27] to evaluate usability and accessibility will be considered to assure an efficient evaluation of user experience in next experiments.

As future work, some complementary existing solutions will be explored. For example, the proposed system may be complemented with the system proposed by Bouzid and Jemni [28], who describe an avatar-based system for automatically generating 3D signing animations from SignWriting annotations expressed with SWML (SignWriting Markup Language). Besides, some measures to optimize the performance of the system will be considered and evaluated, such as the fusion of the repositories SMBL and SSS-XXXX-SVG as proposed in [29].

Specifically the solution described in this paper is applied to improve the web accessibility for people with hearing impairment. However, thanks to the mentioned separation of content and representation, by using the WAI Accessibility Adaptor or core component, which can be extended by any specific adaptor, the approach can be extended to target other types of impairments, as intellectual disability. Therefore, the proposed translator could be adapted to support other types of graphics accompanying video, which would be adapted to each specific need.

References

[1] Shiver, B. N., & Wolfe, R. J. (2015). Evaluating Alternatives for Better Deaf Accessibility to Selected Web-Based Multimedia. Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility (pp. 231–238). New York, NY, USA: ACM Press. doi: 10.1145/2700648.2809857

[2] Cordero, A., Lluch, C.J., Sanabria-Codesal, E., & Torregrosa, J.R. (2015). Towards a better learning models through OCWs and MOOCs. International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 3, No. 4. pp. 26-30. doi: 10.9781/ijimai.2015.345

[3] Caldwell, B., Cooper, M., Guarino Reid, L., & Vanderheiden, G. (Eds.). (2008). Web Content Accessibility Guidelines (WCAG) 2.0. W3C Recommendation. Retrieved from http://www.w3.org/TR/WCAG20/

[4] Fulton, C. (2011). Web Accessibility, Libraries, and the Law. Information Technology and Libraries, Vol. 30, No. 1, pp. 34-43. doi: 10.6017/ital.v30i1.3043

[5] Crespo, R.G., Espada, J.P., Burgos, D., & Verdú, E. (2016). Social4All: Collaborative Platform for Improving Web Accessibility. In Proceedings of the XVII International Conference on Human Computer Interaction (pp. 33:1-33:2). New York, NY, USA: ACM. doi: 10.1145/2998626.2998629

[6] Crespo, R.G., Aguilar, L.J., & Martínez, O. S. (2012). Improving access to IT services for people with disability through software aids. Journal of Ambient Intelligence and Smart Environments, Vol. 4, No. 6, pp. 563-564. doi: 10.3233/AIS-2012-0181

[7] Crespo, R. G., Martíne, O. S., Lovelle, J. M., García-Bustelo, B. C., Díaz, V. G., & Ordoñez de Pablos, P. (2012). Improving Cognitive Load on Students with Disabilities through Software Aids. In Cakir, A., & Ordóñez de Pablos, P. (Eds.), Social Development and High Technology Industries: Strategies and Applications (pp. 163-175). Hershey, PA, USA: IGI Global. doi: 10.4018/978-1-61350-192-4.ch011

[8] Pascual, A., Ribera, M., & Granollers, T. (2015). Impact of web accessibility barriers on users with a hearing impairment. DYNA, Vol. 82, No. 193, pp. 233–240. doi: 10.15446/dyna.v82n193.53499

[9] Báez-Montero, I., & Fernández Soneira, A.M. (2010). Spanish deaf people as recipients of closed captioning. In A. Matamala & P. Orero (Eds.), Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing. (pp. 25-44). Bern (Switzerland): Peter Lang.

[10] Oliveira, F. C. de M.B., Gomes, G.N.C., de Freitas, A.T., de Oliveira, A.C., Silva, L.C., & Queiroz, B. (2015). A Comparative Study of the Acceptability of Signs for the Brazilian Sign Language Created in Person and Remotely. In Proceedings of the 46th ACM Technical Symposium on Computer Science Education (pp. 207-211). New York, NY, USA: ACM. doi: 10.1145/2676723.2677293

[11] Fromkin, V., Rodman, R., & Hyams, N. M. (2014). An introduction to language (Tenth edition). Boston, MA, USA: Wadsworth/Cengage Learning.

[12] Bianchini, C.S., Borgia, F., & de Marsico, M. (2012). SWift - A SignWriting Editor to Bridge between Deaf World and E-learning. In 2012 IEEE 12th International Conference on Advanced Learning Technologies (ICALT) (pp. 526-530). IEEE. doi: 10.1109/ICALT.2012.235

[13] Parkhurst, S., & Parkhurst, D. (2010) A Cross-Linguistic Guide to SignWriting: A Phonetic Approach. Retrieved from http://www.signwriting.org/archive/docs7/sw0617_Cross_Linguistic_Guide_SignWriting_Parkhurst.pdf

[14] San-Segundo, R., Montero, J. M., Córdoba, R., Sama, V., Fernández, F., D'Haro, L. F., López-Ludeña, V., Sánchez, D., & García, A. (2012). Design, development and field evaluation of a Spanish into sign language translation system. Pattern Analysis and Applications, Vol. 15, No. 2, pp. 203-224.

[15] de Araújo, T.M.U., Ferreira, F.L.S., Silva, D.A.N.S., Oliveira, L.D., Falcão, E.L., Domingues, L.A., Martins, V.F., Portela, I.A.C., Nóbrega, Y.S., Lima, H.R.G., Filho, G.L.S., Tavares, T.A., & Duarte, A. N. (2014). An approach to generate and embed sign language video tracks into multimedia contents. Information Sciences, Vol. 281, pp. 762-780. doi: 10.1016/j.ins.2014.04.008

[16] López-Ludeña, V., González-Morcillo, C., López, J. C., Ferreiro, E., Ferreiros, J., & San-Segundo, R. (2014). Methodology for developing an advanced communications system for the Deaf in a new domain. Knowledge-Based Systems, Vol. 56, pp. 240–252. doi: 10.1016/j.knosys.2013.11.017

[17] López-Ludeña, V., González-Morcillo, C., López, J. C., Barra-Chicote, R., Cordoba, R., & San-Segundo, R. (2014). Translating bus information into sign language for deaf people. Engineering Applications of Artificial Intelligence, Vol. 32, pp. 258–269. doi: 10.1016/j.engappai.2014.02.006

[18] Crespo, R.G., Fernández, G.G., Martínez, O.S., Franco, E.T., & Aguilar, L.J. (2009). Intelligent Streaming Server for Non Accessible Contents Stored on Web Servers to Disabled People: Signwritting Case. In J.M. Cochado et al. (Eds.), International Symposium on Distributed Computing and Artificial Intelligence 2008 (DCAI 2008) (pp. 512-520). Berlin Heidelberg: Springer-Verlag.

[19] da Rocha Costa, A. C., & Pereira Dimuro, G. (n.d.). Supporting Deaf Sign Languages in Written Form on the Web. Retrieved from: http://www.signwriting.org/forums/software/swml/swml01.html

[20] Cazenave, F., Quint, V., & Roisin, C. (2011). Timesheets.js: when SMIL meets HTML5 and CSS3. In Proceedings of the 11th ACM symposium on Document engineering (pp. 43-52). New York, NY, USA: ACM. doi: 10.1145/2034691.2034700

[21] Bulterman, D., Jansen, J., Cesar, P., Mullender, S., Hyche, E., DeMeglio, M., Quint, J., Kawamura, H., Weck, D., García-Pañeda, X., Melendi, D., Cruz-Lara, S., Hanclik, M., Zucker, D.F., & Michel, T. (2008). Synchronized Multimedia Integration Language (SMIL 3.0), W3C Recommendation. W3C. Retrieved from: http://www.w3.org/TR/REC-smil/

[22] Jalava, T., Honkala, M., Pohja, M., & Vuorimaa, P. (2005). Timesheets: XML Timing Language. W3C. Retrieved from: http://www.w3.org/Submission/xml-timing/

[23] WAVE. (2017). WAVE Web accessibility evaluation tool. Retrieved from: http://wave.webaim.org

[24] Section 508. (2017). Section 508 Standards. Retrieved from: https://www.section508.gov/summary-section508-standards

[25] Hassan, Y., & Martín, F.J. (2003). Guía de Evaluación Heurística de Sitios Web. Retrieved from: http://www.nosolousabilidad.com/articulos/heuristica.htm.

[26] Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Construction of a Benchmark for the User Experience Questionnaire (UEQ). International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 4, No. 4, pp. 40-44. doi: 10.9781/ijimai.2017.445

[27] Bader, F., Schön, E.M., & Thomaschewski, J. (2017). Heuristics Considering UX and Quality Criteria for Heuristics. International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 4, No. 6, pp. 48-53. doi: 10.9781/ijimai.2017.05.001

[28] Bouzid, Y., & Jemni, M. (2013). An animated avatar to interpret signwriting transcription. 2013 International Conference on Electrical Engineering and Software Applications (ICEESA) (pp. 1-5). doi: 10.1109/ICEESA.2013.6578394

[29] Crespo, R.G. & Martínez, O.S. (2010). The Web 3.0 in the service of the people with severe deafness through the accessibility guidelines 2.0. Sociedad y Utopía, No. 36, pp. 153-172.

Elena Verdú Pérez

Dr. Elena Verdú Pérez received the Master's and Ph.D. degrees in Telecommunications Engineering from the University of Valladolid, Spain, in 1999 and 2010, respectively. She is currently an Associate Professor at Universidad Internacional de la Rioja (UNIR). For more than 10 years, she has worked on research projects at both national and European level. She has contributed to the design and development of several telematic systems such as QUESTournament competitive learning system, a distributed remote evaluation system, an intelligent tutoring system, or an intelligent system for map web services. Her research interests include intelligent tutoring systems, competitive and collaborative e-learning systems, web accessibility, machine learning, and intelligent systems applied to robotics.

Cristina Pelayo G-Bustelo

Dr. B. Cristina Pelayo G-Bustelo is an Associate Professor in the Computer Science Department of the University of Oviedo, Ph.D. in Computer Science and Computer Engineer. Her research interests include Web Engineering, Modelling Software with Domain-Specific Languages, Model-Driven Engineering and Business Process Management, Internet of Things and Data Science.

María Ángeles Martínez Sánchez

She received the M.S. degree in Social Work in 1999 and the Ph.D. degree in Social Sciences in 2011, both from the University of Granada (Granada, Spain). Currently, she is professor in the Department of Social Work and Services at the Faculty of Social Work in University of Granada. Additionally, she is a sign interpreter and worked for 11 years as a social worker in the deaf association of Granada. Her main research topics are digital libraries, bibliometrics, ICT applied in deaf population, and problems of inclusion in deaf population.

Rubén González Crespo

Dr. Rubén González Crespo has a PhD in Computer Science Engineering. Currently he is Dean of Higher School of Engineering from UNIR and Director of the AENOR (Spanish Association for Standardization and Certification) Chair in Certification, Quality and Technology Standards. He is advisory board member for the Ministry of Education at Colombia and evaluator from the National Agency for Quality Evaluation and Accreditation of Spain (ANECA). He is member from different committees at ISO Organization.