Adaptive Membership Selection Criteria using Genetic Algorithms for Fuzzy Centroid Localizations in Wireless Sensor Networks

Songyut Permpol, Kanokmon Rujirakul and Chakchai So-In
Department of Computer Science,
Faculty of Science, Khon Kaen University.
songyut_p@kkumail.com

***Abstract*—This paper investigates the effect of fuzzy inputs, i.e., signal strength, of various known nodes, to fuzzy logic systems in order to derive a proper weight for Centroid, properly used to approximate the location in wireless sensor networks with its key advantage on simplicity but with precision trade-off. Due to a fluctuation behavior of location estimation precisions with respect to a diversity of various inputs, here, we propose the use of heuristic approach applying genetic algorithms with mutation and cross-over steps to adaptively seek the optimal solution – a proper number of membership functions for fuzzy logic systems in weighted Centroid – to achieve higher location estimation accuracy. The performance of our methodology is effectively confirmed by the intensive evaluation on a large scale simulation in various topologies and node densities against fixed membership function scenarios including a traditional Centroid.**

***Index Terms*—Adaptive Membership Function Selection; Centroid; Fuzzy Logic; Genetic Algorithms; Wireless Sensor Networks.**

I. INTRODUCTION

For decades, wireless sensor networks (WSN) have been widely used in both research and industry while there exists several practical applications, e.g., health-care, surveillance, tracing, and tracking [1]. Their successes are from the commercial requirement satisfactions from the industry with beneficial gain, and most important, with several key features of WSNs, such as self-contained tiny sensor nodes including computing, storage, and transmission units, with dedicated power. The sensor can provide different functions, each of which can form a large-scale network towards un-wired communications [2].

Although a variety of WSN deployments has been put in practical usage, there remain several issues required for further improvements, in particular, with power constraint. Some of which include reliability, scalability, routing, and quality of services. One of the challenges is localization, especially without GPS functionality [3-4], or even if embedded, high complexity of the hardware logics and costs may be increased.

Without GPS, one of the promising location estimation schemes is based on range-free approximation with its key advantage on low cost, suitably for limited power sensor node. One of the pioneers is Centroid [5] but with key limitation on (high) estimation error. Nevertheless, for years, a research community has investigated on its improvement, and one of the promising techniques is to add extra weights to reflect additional factor besides the only node position information, i.e., the received signal strength as the indication (RSSI) corresponding to the location and distance of the known node (anchor node) [6].

One of the Centroid enhancements with regard to additional weights is based on fuzzy logic systems (FLS) since it can yield higher accuracy without scarifying computational complexity [7].The performance of the fuzzy can be relied on the design of input, output, and membership functions, and normally, these are fixed. For example, RSSI issued as the fuzzy input with fixed numbers of functions corresponding to the derived output; however, this fixed scheme may not reflect the effect of RSSI diversity.

Thus, this research investigates the possibility to construct an adaptive selection scheme to find out a proper number of membership functions. Here, we apply the concept of genetic algorithms (GA) to derive the optimal solution (number of functions). Then, the derived weight will be used further in weighted Centroid to determine the approximated location in WSNs.

This paper is organized as follows: a brief literature survey related to fuzzy centroid localization techniques is described in section 2. Then, in section 3, our methodology is presented with detailed description. Section 4 discusses the performance analysis and evaluation of our method including a comparative result against other existing methods, i.e., fixed numbers of membership functions including a traditional Centroid. Finally, section 5 provides the conclusion and possible future work.

II. RELATED WORK

The manuscript article should be written in English in the font of Times New Roman, which includes the following: abstract, introduction, literature review, objectives, research methodology, theory, testing and analysis, results and discussions, conclusion, acknowledgement and references. Manuscript should be prepared via the Microsoft Word processor. As briefly discussed, recently, several location estimation techniques have been proposed to improve the accuracy of WSNs [4], especially without GPS functionality.

One of the promising approaches is Centroid with its key advantage on simplicity but (low) precision estimation trade-off. Therefore, in this section, our review will focus on the use of FLSs to adjust the weight for Centroid approximation process.

Traditionally, in 2000, N. Bulusu *et al*. [5] proposed to use Centroid to estimate the outdoor location but without GPS functionality due to additional costs. This work did not apply any additional weights, i.e., there was no RSSIs involved in the calculation, and so low estimation accuracy.

To enhance the precision of the Centroid, in 2005, S. Yun *et al*. [6] considered the adjustable weight derived from FLSs using RSSIs as inputs. Here, Fuzzy Sugeno was implemented. The authors also applied GA to adjust the period of membership functions but with fixed five membership functions (also with 5rules).In addition, similarly, R. Monfared *et al*. [8] used Fuzzy Sugeno to derive the weight for Centroid but with fixed nine membership functions.

Another promising type of the FLSs, Fuzzy Mamdani, was also considered due to its simplicity. For example, D. F. Laripos *et al*. [9] not only considered the use of Mamdani but also took another computational step – including an additional constraint with another layer of RSSIs of received anchor nodes but with computational trade-off. In addition, V. Kumar *et al*. [10] investigated the performance of Fuzzy Sugeno and Fuzzy Mamdani and then reported the superior of the latter. The authors also showed the performance evaluation of the hybrid of these twos.

In 2013, similar to S. Yun *et al*. [6], A Patri *et al*. [11] manually searched for the best membership function and then reported that Sinc was superior. They also applied GA to adjust the membership function period of Sinc. Additionally note that in [12], GA was also used to minimize the error given the mathematic conversion of distance from RSSI; however, the accuracy may be not exact due to RSSI variation.

From all techniques discussed above, FLS has been widely used to derive weights. To improve the estimation precision, there are some promising techniques applying GA to adjust the membership function period or directly minimize the location error from RSSIs. However, these improvements will not reflect the diversity of inputs. Thus, this research focuses on the use of GA to select a proper number of membership functions.

III. ADAPTIVE MEMBERSHIP SELECTION CRITERIA FOR FUZZY CENTROID IN WIRELESS SENSOR NETWORKS

Our optimization scheme consists of three main stages as follows:

1. Beacon Announcement: the unknown node receives the beacon from anchor nodes given its coverage. This broadcasting is typically periodic containing anchor nodes' positions and RSSIs.
2. Fuzzy Centroid: RSSIs corresponding to the position of anchor nodes will be fed into the fuzzy logic system to generate the proper weights used to adjust Centroid estimation process. Note that here, we apply GA to seek the optimal number of membership functions.
3. Location Estimation: with generated weights, the location approximation process will be applied to finally predict the unknown location in (x, y) coordinates.

A. Centroid Localization

The key advantage of Centroid-based localization schemes is on low complexity in that the location approximation can only be derived based on the information of the actual location from known nodes (anchor) around the unknown position [5].Typically, the unknown location ($\hat{x}, \hat{y}$) node will receive the known positions i (x_i, y_i) from thei^{th} anchor node toward beacon announcements, and then apply the equation below.

$$\hat{x}, \hat{y} = \frac{\sum_{i=1}^{m} x_i}{m}, \frac{\sum_{i=1}^{m} y_i}{m} \quad (1)$$

Here, m denotes the total number of anchor nodes given the unknown node coverage. Note that only Centroid can mislead the estimation precision, especially without the concern of signal strengths. In general, the concept of weights (w) has been introduced to improve the precision as shown in equation below, so-called weighted Centroid [6].

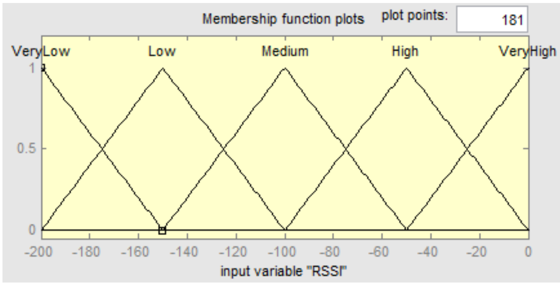
$$\hat{x}, \hat{y} = \frac{\sum_{i=1}^{m} x_i w_i}{\sum_{i=1}^{m} w}, \frac{\sum_{i=1}^{m} y_i w_i}{\sum_{i=1}^{m} w} \quad (2)$$

B. Fuzzy Weight Derivation

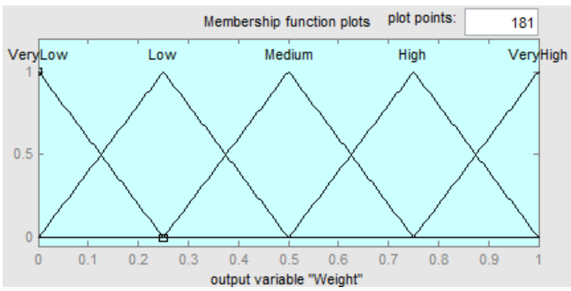
As previously stated, an additional weight is typically used to improve the location estimation performance. In this paper, we apply FLS [13] since their key advantage is on the simplicity, i.e., there is no high computational time required, and this is suitable for distributed sensor nodes (with limited power as constraint). In general, there are four main components of the fuzzy as follows.

1. Fuzzification: this component is used for input transformation into a proper range for membership function derivation (as examples shown in Figure 1).
2. Inference Engine: this component is used to validate the logic and rule to make justification.
3. Defuzzification: this component is used for output transformation into a proper range (as examples shown in Figure 2).
4. Knowledge: this component is used to collect the useful data consisting of two sub-components:
   a. Rule-based: this is used to generate the rule based on the expert as examples shown in Table 1.
   b. Database: this is used to prepare necessary information to control the rule construction and fuzzy logic process.

Note that with the recommendation provided by S. Yun *et al*.[6], there are five functions, i.e., Very low, Low, Medium, High and Very high, as inputs and five rules. However, in this research, the selection of membership function is based on the intensive evaluation using GA. The shape of membership function follows Triangular functions [14] (as shown in Figure 1 and 2). Fuzzy Mamdani [13] was our selection due to its key advantage of simplicity. As shown in Figure 2, the output generated from Fuzzy Mamdani will be the summation of each membership level based on rule aggregation using central of gravity (CoG).



(a)



(b)

Figure 1: Fuzzy Logic: Triangular function of input (RSSI) (a) and output (weight) (b)

Table 1
Fuzzy Rule: Example

Rule	IF: State of RSSI	THEN: State of Weight
1	Very low	Very low
2	Low	Low
3	Medium	Medium
4	High	High
5	Very high	Very high

Algorithm 1: Fuzzy Logic Systems
Input: $RSSI[N]$, l
Output: $output[N]$

1. Generate $l$membership functions for input $RSSI[N]$ and output weight$weight[N]$
2. Generate fuzzy rule
3. Apply fuzzy inference engine with aggregation method $(RSSI[N], weight[N])$
4. Calculate the output $(output[N])$ from CoG

Note that here, the fuzzy input is RSSI received from different anchor nodes given the coverage to derive the output weights in range from 0 to 1. Note that Algorithm 1 also shows the weight derivation (output) from fuzzy logic. Here, *l* denotes a number of membership functions corresponding to the inputs (RSSIs) from *N* anchor nodes.

C. Genetic Algorithms

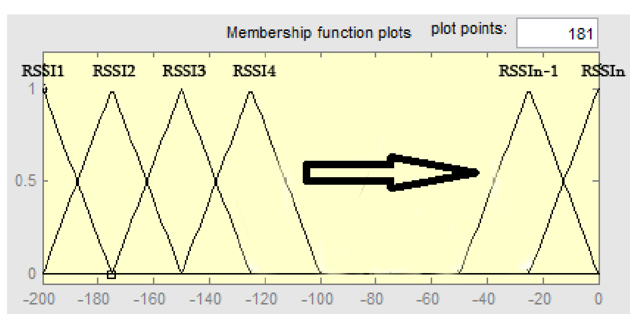
One of the pioneer heuristic approaches used to find out the optimal solution with regards to local minima is Genetic Algorithms (GAs) [15]. In general, GA imitates the nature selection process of chromosomes such that the evaluation of selection quality can be based on the fitness function. Their existing solutions will be selected and used to generate new solutions or genes. The selection process normally includes two main steps: cross-over and mutation [15].An overall procedure can be stated as follows.

1. At the beginning process, the population will be provided initially as a set. Here, the population represents the number of membership functions.
2. Each selected member in the population will be evaluated based on fitness function (*Fn*). Here, we use the maximization of outputs' standard deviations corresponding to the input RSSIs of FLSs.
3. A simple ranking process will be applied based on *Fn*.
4. A member from the population will be used as parent to create children (new generation), normally with cross-over and mutation operators
    a. Cross-over: our encoding scheme is to apply the equal-divider of bit scheme (binary representation) of number of membership functions of FLSs. For example, with *n* bits (*n*=6), two parents of $(28)_{10}$ = $(011100)_2$ and$(13)_{10}$ = $(001101)_2$, once applying cross-over (each half will be inter-changed), i.e., two children will be created as $(29)_{10}$=$(011101)_2$ and$(12)_{10}$=$(001100)_2$, respectively.
    b. Mutation: at this step, to create a diversity of the population, we randomly select the input from outside population.
5. Here, the population is increased, and again, the evaluation process is applied.
6. The selection of the best local fitness value will be performed based on *Fn*.
7. Return to Step 3 until the solution is satisfied or meeting the specific threshold.

D. Adaptive Membership Selection Criteria for Fuzzy Centroid

Our motivation towards the adaptive concept on the membership function selection is due to RSSI diversity received from different anchor nodes with respect to the distance to the unknown node. A fixed number of fuzzy membership functions (as examples shown, previously in Figure 1– 5 rules [6]) cannot represent the effect of this diversity.

Thus, in this research, our goal is to find out a proper number of membership functions based on the raw RSSIs in range of minimum and maximum RSSIs. Figure 3 shows examples when the number of membership functions is equal to the number of anchor nodes (*N*), i.e., around *N* triangular shapes.



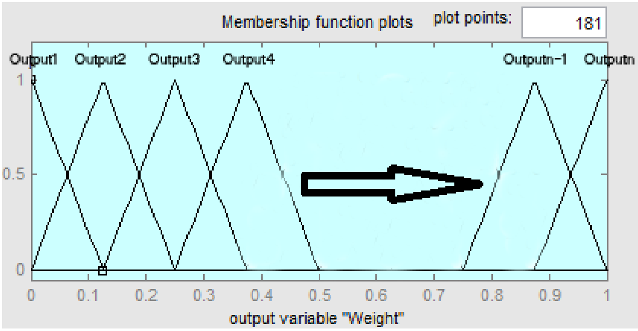


Figure 3: Fuzzy Logic: Triangular function of input (RSSI) and output (weight) with N members

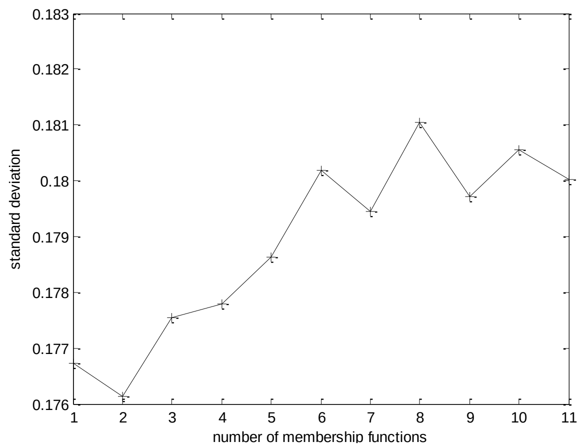


Figure 4: Standard deviation vs. number of membership functions

Algorithm 2 shows steps to determine the estimated location as follows: After the initial population (the number of membership functions) (line 1), *N* as examples, some random members will be selected from the population *l* (line 2), here, we use 3 as *l*. Lines 4 to 5 show the actual FLS used to finally create the output of defuzzifier process, and then generate the standard deviation of each fuzzy model. Line 6 shows a simple ranking algorithm (sorting)when applied in descending order.

Note that the best two models corresponding to the fitness function will be selected to perform cross-over operation to generate two more children (line 7). An encoding scheme uses 6 bits based on our intensive evaluation (at maximum of 2^6 = 64 functions). Then, another child will be generated based the mutation operation (random selection from outside current population) (line 8). This process will be iteratively apply given the stopping threshold. Finally, the derived weight based on the end of iteration will be used for fuzzy weighted Centroid to approximate the unknown location of $(\hat{x}, \hat{y})$.

Algorithm 2: Adaptive Membership Selection Criteria for Fuzzy Centroid

Input: $RSSI[N], l, threshold$

Output: $\hat{x}, \hat{y}$

1. Initial population size (number of membership functions)
2. Randomly select *l* members from the population
3. do
4. while *i* in *l* members
   Create Fuzzy Logic Model ($RSSI[N]$, *i*)
   Defuzzify *output[N]*
   Calculate Fitness function (standard deviation) of model *i*
5. end while
6. Sorting the models based on their *Fn* in descending order (model[*l*])
7. Perform Cross-over {model[*1*], model[*2*]} to create model[*l+1*] and model[*l+2*]
8. Perform Mutation to create model[*l+3*] from outside population
9. while (*threshold*)
10. Calculate the location estimation according to equation (2) using *output[N]*

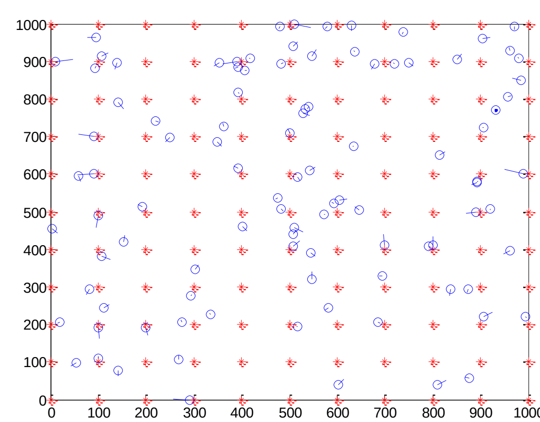
IV. Performance Evaluation

To confirm the validity of our methodology, this section provides an analysis of performance evaluation of our adaptive scheme, called Adaptive Fuzzy Centroid against the fixed schemes like five or nine rules [6, 8] including a traditional Centroid.

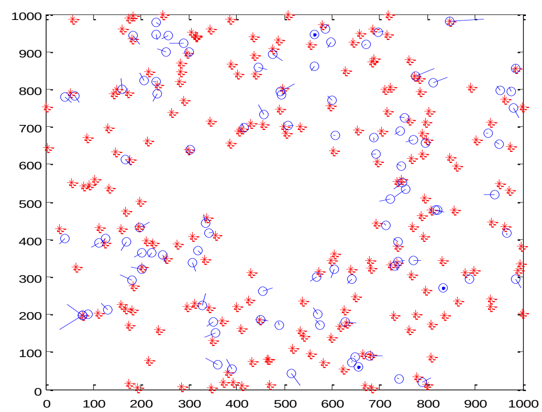
A. Simulation Configurations

Our testbed follows a standard simulation based on Matlab framework including standard libraries, and the recommendation provided by S. Gu *et al*. [16]. Our testbed is with a standard Windows 7 64-bit with Intel Core Q8400 2.66 GHz, 4 GB DDR-SDRAM, and 320GB 7200 rpm hard disk. Our simulation reflects a large-scale node distribution in area of 1000×1000 m^2dividing into two main topologies, i.e., grid and non-uniform (with five holes) [17],for the sake of paper length limitation (See Figure 5).

With two main topologies, a variation of anchor nodes was in range of 121, 196, and 441, corresponding to the grid deployment of 100×100, 75×75, and 50×50 m^2, respectively. This node density was also applied to the other topology. The number of unknown nodes is fixed to 100. The sensing coverage is at 100 meters. There is no mobility involved once deployed. The signal propagation model follows a log-distance path loss model [16]. There is no energy consumption factor in our focus - in computing and transmission logic - for localization model evaluation [18]. There is no limitation of routing techniques. The location estimation will be computed within each node in distributed manners.



(a) Grid Distribution



(b) Non-Uniform Distribution

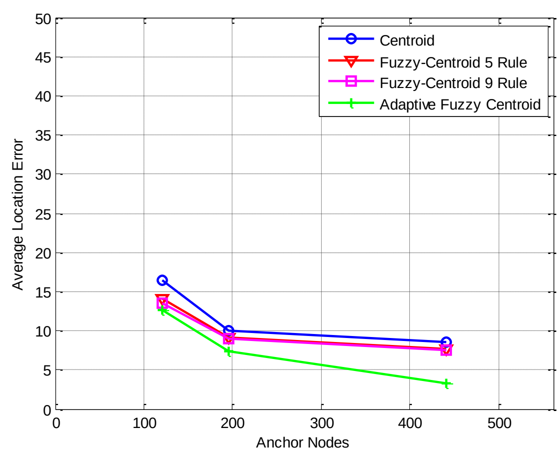
Figure 5: Node Distribution Deployment (cross = anchor; circle = unknown; line = absolute error)

In our performance analysis, Average Location Error (ALE) (in meters) is mainly used as our main metric measurement for location approximation error [14] stated in equation (3) below. Here, (x, y) and $(\hat{x}, \hat{y})$ represent the actual and estimated positions of unknown nodes. The simulation is over ten trials and the average of ALEs will be used. Again, our adaptive membership selection criteria were evaluated against fixed schemes as presented in [6, 8] including a traditional Centroid.

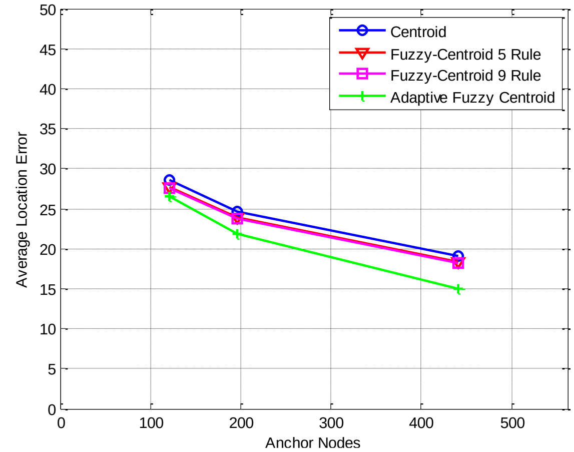
$$ALE = \frac{\sum\sqrt{(\hat{x}-x)^2+(\hat{y}-y)^2}}{\#sensor\ nodes} \quad (3)$$

B. Simulation Results

Figure 6 shows the performance evaluation results in terms of ALE in both grid and non-uniform distribution deployments with 100 meters in signal coverage. In general, with the increase of number of anchor nodes, the estimation precision tends to be high in both scenarios, such as from 15 to 30; and 4 to 18 meters, respectively.



(a) Grid Distribution



(b) Non-Uniform Distribution

Figure 6: Average Location Error with grid and non-uniform distributions

In addition, the error trends of grid deployments are higher than those of the other deployment scenario. This is justifiable since normally, with Centroid-based approaches, sensors in fixed deployment can represent the best error estimation scheme.

Specifically considering each comparative proposal, our adaptive scheme is superior than the others, and in order of Fuzzy Centroid with 9 and 5 rules, and finally a traditional Centroid, respectively. The estimation error of our scheme is in range between 26 and 15 meters; and 13 and just only 4 meters for both scenarios. However, with Centroid, the performance is the worst, i.e., from almost 20 to 30 meters and 9 to 17 meters, respectively. This is the fact that the fuzzy logic takes an effect with additional weights (RSSIs). In addition, the error of Fuzzy Centroid is obviously better than the only Centroid. The error of Fuzzy Centroid with 9 rules is slightly better than that of 5 rules but still worse than our adaptive scheme due to the effect of RSSI diversity.

V. CONCLUSION

A conclusion to review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions. One of the key limitations of a traditional Centroid is on high location estimation error even with simplicity gain. An adjustable weight has been widely used to mitigate that limitation. Here, we investigated the use of fuzzy logic systems to derive the adjustable weight. However, with fixed number of membership functions, the derived output may not reflect the effect of signal diversity from different anchor nodes in various locations and distances. Thus, we enhanced the membership selection criteria using genetic algorithms. With our intensive simulation in a large scale network and node densities, our proposal performance is superior than other fixed two schemes, and obviously the traditional one, i.e., in average of 12.89%, 14.08%, and 19.18%, respectively.

Although our performance validity can be confirmed towards the estimation improvement stated in our simulation results, i.e., low location estimation error, there still requires further investigation, such as with a variety of scenarios and constraints. Several issues regarding a practical deployment should be considered, such as scalability, network density and diversity, diverse topology, network dimension, mobility, and signal propagation model and coverage. In addition, additional transmission protocol overheads as well as computational time complexity trade-off should be further studied, and all of these are for possible future work.

REFERENCES

[1] Atzori, L., Iera, A., Morabito, G. "The Internet of things: a survey," *Comput. Netw*, vol. 54, pp. 2787-2805, 2010.

[2] Lindroos, V., Tilli, M., Lehto, A., Motooka, T. "Handbook of Silicon Based MEMS Mate-rials and Technologies," *William Andrew*, Burlington, MA, USA, 2010.

[3] Yick, J., Mukherjee, B., Ghosal, D. "Wireless sensor network survey," *Comput. Netw*, vol. 52, pp. 2292-2330, 2008.

[4] Han, G., Xu, H., Duong, T.Q., Jiang, J., Hara, T. "Localization algorithms of Wireless Sen-sor Networks: a survey," *Telecommun. Syst*. vol. 52, pp. 2419-2436, 2013.

[5] Bulusu, N., Heidemann, J., Estrin, D. "GPS-less low-cost outdoor localization for very small devices," *IEEE Pers. Commun*. vol. 7, no. 5, pp. 28-34, 2000.

[6] Yun, S., Lee, J., Chung, W., Kim, E. "Centroid localization method in wireless sensor net-works using TSK fuzzy modeling," *In: Proc. Int. Symp. on Advanced Intell. Syst*, pp. 971-974, 2005.

[7] Jang, H., Topal, E. "A review of soft computing technology applications in several mining problems," *Applied Soft Comput*, vol. 22, pp. 638-651 2014.

[8] Monfared, M.A., Abrishambaf, R., Uysal, S. "Range Free Localization of Wireless Sensor Networks Based on Sugeno Fuzzy Inference," *In: Proc. Int. Conf. on Sensor Technol. and Appl*, pp. 36-41, IEEE, Rome, Italy, 2012.

[9] Larios, D.F., Barbancho, J., Molina, F.J., León, C. "LIS: localization based on an intelligent distributed fuzzy system applied to a WSN," *Ad Hoc Netw*, vol. 10, no. 3, pp. 604-622, 2012.

[10] Kumar, V., Kumar, A., Soni, S. "A combined Mamdani-Sugeno fuzzy approach for locali-zation in wireless sensor networks," *In: Proc. Int. Conf. & Workshop on Emerging Trends in Technol*, pp. 798-803. ACM, NY, USA, 2011.

[11] Patri, A., Nayak, A. "A fuzzy-based localization in range-free wireless sensor network using genetic algorithm & Sinc membership function," *In: Int. Conf. on Green Comput., Commun. and Conservation of Energy*, pp. 140-145, IEEE, Chennai, India, 2013.

[12] Huanxiang, J., Yong, W., Xiaoling, T. "Localization algorithm for mobile anchor node based on genetic algorithm in wireless sensor network," *In: Int. Conf. on Intell. Comput. and Integrated Syst*, pp. 40-44. IEEE, Guilin, China, 2010.

[13] Zadeh, L.A. "Fuzzy algorithms. Information and Control. vol. 12, no. 2, pp. 94-102, 1968.

[14] So-In, C., Permpol, S., Rujirakul, K.: Soft computing-based localizations in wireless sensor networks," *Pervasive and Mobile Comput, (in press)* 2015.

[15] Mitchell, M. "An Introduction to Genetic Algorithms," *MIT Press*, 1999.

[16] Gu, S., Yue, Y., Maple, C., Wu, C. "Fuzzy logic based localization in Wireless Sensor Networks for disaster environments," *In: Proc. 18th Int. Conf. on Automation & Comput*, pp. 1-5. IEEE, CA, USA, 2012.

[17] Tran, D.A., Nguyen, T. " Localization in wireless sensor networks based on support vector machines," *IEEE Trans. Parallel and Distrib. Syst*, vol. 19, pp. 981-994, 2008.

[18] Pantazis, N.A., Nikolidakis, S.A., Vergados, D.D. "Energy-efficient routing protocols in wireless sensor networks," *IEEE Commun. Survey & Tutorials*, vol. 15, pp. 551--591, 2013.

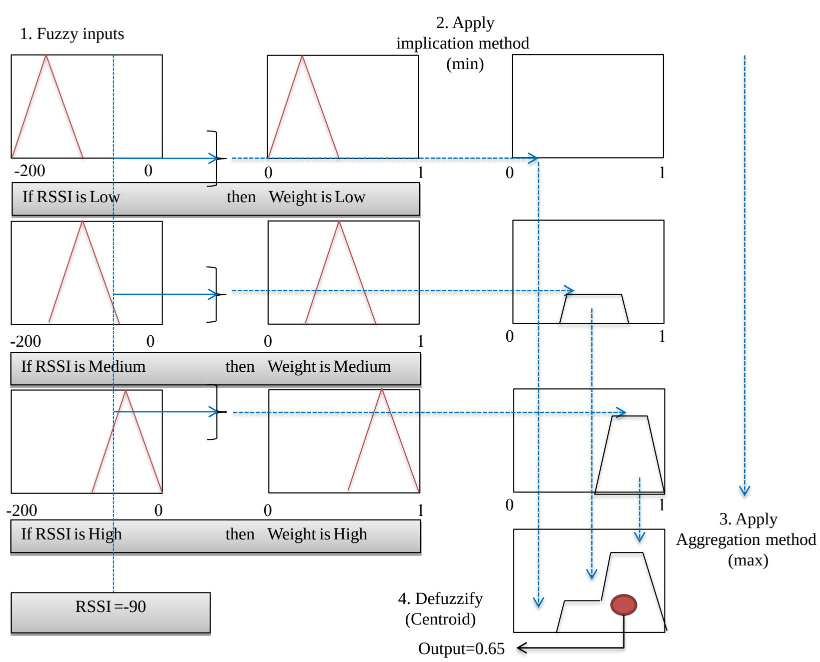


Figure 2: Fuzzy Logic: Fuzzy Inference Engine and Defuzzifier