

# Localization of Indoor Areas Using Wi-Fi Signals, Type-2 Fuzzy Inference Systems and Data Mining

Manuel Castañón-Puga, *Member, IAENG*, Abby Stephanie Salazar, Leocundo Aguilar,  
Carelia Gaxiola-Pacheco and Guillermo Licea

**Abstract**—Recent advances in communication and Wi-Fi mobile technology has led the Wireless Local Area Network to become more ubiquitous. These networks provide a potential infrastructure that enables the indoor location estimate for users who wear wireless devices such as smartphones. This infrastructure goes beyond GPS (Global Position System) where signal is weak or is not available in indoors. This paper describes an alternative and unconventional method to people indoor localization designed for persons using Wi-Fi RSSI fingerprinting by a mobile application as well as an estimation based on Type-2 fuzzy inference systems. Wi-Fi Fingerprinting creates a radio map of a given area based on the RSSI data from several access points (APs) and generates a set of RSSI data for a given zone location. Consequently the Type-2 fuzzy inference system is generated using the obtained set of data; thus new RSSI values are introduced to the Type-2 fuzzy inference system to obtain an estimation of the user zone location.

**Index Terms**—Wireless Networks, Location-based Services and Positioning, Wireless/mobile networked applications, Type-2 Fuzzy Inference System, Data Mining.

## I. INTRODUCTION

With recent advances in technology, there is a possibility of connecting different devices using wireless networks to transparently interact with the environment and seamlessly with people. The challenge in these ideas is to find ways to meet the particular context in time and space, to make a decision based on it. Devices require algorithms with the ability to understand the context information so that a given system can adapt to it and behave in a certain way.

In order to understand the context it is necessary to take in mind different factors, which vary depending on the activity to be performed, such as current weather or temperature, atmospheric pressure, lighting from one place, position of an object, the place where a person is, among others. Considering a future implementation that helps to estimate the context, an indoor location method that works within a mobile device using Wi-Fi signals is proposed.

### A. Situational Awareness

Since a longtime the interest of creating things that could function and perform actions without the need for direct intervention of man exists. In the past, objects that could interpret their environment, communicate and respond appropriately based on that, looked distant due to the complexity involved. However, at present this is more possible mainly

Manuscript received March 14, 2014; revised March 20, 2014. This work was supported in part by the Internal Fund for Research Projects (Grant No. 300.6.C.135.17) of the Autonomous University of Baja California, México. M. Castañón-Puga, A. S. Salazar, L. Aguilar, C. Gaxiola-Pacheco and G. Licea are with the Chemistry and Engineering Faculty, Autonomous University of Baja California, Tijuana, BC, 22390, México, e-mail: {puga, abby.salazar, laguilar, cgaxiola, glicea}@uabc.edu.mx

by two factors: because in the last years the technological advancement is increasing and also because of the great increase of devices connected to different networks, such as Internet [1].

The first one gives us devices less invasive. Additionally, the second factor is an advantage for communication between people and devices, in order that relevant information can be taken and shared for different purposes as to understand the context of a situation.

The context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves [2].

Devices with the capability of understanding the context are an important key for the non-human intervention in technology tasks. The context parameter of this work is based on is the location of an entity.

### B. Technologies

Actual location techniques commonly use GPS, Bluetooth, or Radio Frequency Identification (RFID) technologies, among others [3]. However, a disadvantage of GPS is that the satellite signals are blocked by obstacles such as walls, in addition to those variations in weather or the presence of buildings results in approximations with errors of meters, so it is not possible to use this system as a method for indoor location [4] [5].

On the other hand, Bluetooth technology has limited coverage, this communication is focused on very short distances to achieve the location.

Finally RF is an expensive solution since it involves the install of different sensors in the area where you will perform the estimate of location, so it is not an economically viable method [4].

Consequently acceptable alternative techniques are required to fit with established infrastructures and satisfy the functionality of indoor location with ease of use and an affordable cost, as the use of technologies based on Wi-Fi [3].

### C. Location Methods

Triangulation and Trilateration: These methods map RSSI as a function of distance that requires a steep linear characterization curve in order to be properly implemented. Functions describing these curves are then used with live RSSI values as input to generate an (x,y) location prediction [3].

One disadvantage of these methods is to do the synchronization of these values and needs a model to determine the distance according to RSSI values.

This work is an alternative solution for localization using a radio map of a given area based on Wi-Fi RSSI data from three (or more) Access Points. Some advantages are that this method works on indoor environments with acceptable coverage, a minimum or null modification of the area infrastructure is required, and is an option less expensive than others [4].

The method proposed is based on a FIS (Fuzzy Inference System) using a clustering method [6], that is a fast, one-pass algorithm for estimating the number of clusters and the cluster centers in a set of data.

#### D. Type-2 Fuzzy Inference Systems

The concept of a Type-2 fuzzy set was introduced by Zadeh [7] as an extension of the concept of usually Type-1 fuzzy sets. A Type-2 fuzzy set is characterized by a membership function whose membership value for each element of the universe is a membership function within the range [0, 1], unlike the Type-1 fuzzy sets where the value of membership is a numeric value in the range [0, 1].

The creation of a fuzzy set depends on two aspects: the identification of a universe of appropriate values and specifying a membership function properly. The choice of membership function is a subjective process, meaning that different people can reach to different conclusions on the same concept. This subjectivity derives from individual differences in the perception and expression of abstract concepts and it has little to do with randomness. Therefore, subjectivity and randomness of a fuzzy set are the main difference between the study of fuzzy sets and probability theory [8].

In Type-1 fuzzy sets, once the membership function is defined for a concept, this is based on the subjective opinion of one or more individuals and it shows no more than one value for each element of the universe. In doing so, it loses some of the ambiguity of the discussed concepts, especially where people may have a slightly different opinion, but they are considered valid. The Type-2 fuzzy sets allow handling linguistic and numerical uncertainties. Figure 1 depicts two graphics of fuzzy sets: a) with Type-1 fuzzy logic, and b) with Type-2 fuzzy logic.

In a) the values set shown are  $A = \{(x, \mu_A(x)) | x \in X\}$  where  $A \subseteq X$ ,  $X \subseteq X$ ,  $X$  is the universe of valid values and  $\mu_A(x)$  is the membership function (MF) that contains a map of each value of  $X$  with its membership value corresponding to a value between [0, 1]. For b) the values of the set are  $\tilde{A} = \{(x, u, \mu_{\tilde{A}}(x, u))\}$ , where MF  $\mu_{\tilde{A}}(x, u)$  has a membership value for each element of the universe as a function of membership in the range [0, 1], so the footprint can be seen around the curve of a).

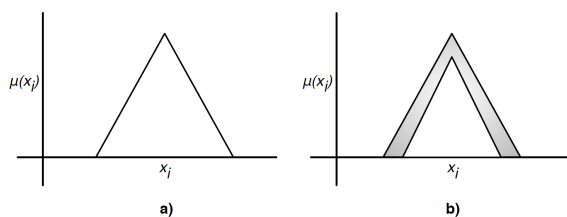


Fig. 1: Type-1 fuzzy set and Type-2 fuzzy set with uncertainty.

#### E. Data Mining and Fuzzy C-Means Clustering

In recent years, the use of new information technologies comes to help in the handling of large amount of data, one evolution of these technologies is the data mining extraction that permits representing knowledge of data store implicitly in big databases. Data mining is a multidisciplinary field that combines techniques from machine learning, pattern recognition, statistics, database and visualization, to direct it to the extraction and interpretation of a huge database. The data mining focuses on filling the need to discover, predict and forecast the possible actions with some confidence factor for each prediction [9].

Moreover, it contributes to the making of tactical and strategic decisions provided by power users, which allows to measure actions and results in the best way, generating descriptive models to explore and understand the data and identify patterns, relationships and dependencies that affect the final results. It creates predictive models that allow undiscovered relationships through the data mining process that are expressed as possible business rules[10].

1) *Fuzzy C-Means Clustering Algorithm for Data Mining:* The fuzzy proposals represent an important place in data mining providing intelligible results. One of these proposals is the Fuzzy C-Means clustering algorithm (FCM) [11] [12] that makes use of membership function and centroid computation procedure iteratively to find the best centroid. The FCM is one of the popular clustering algorithms, the effectiveness of the clustering method relies on the distance measure. FCM is the result of combination of the c-means approach with the handling of fuzzy data. The result of this combination is adequate because it considers the uncertainty presented in the data, avoiding incorrect results and creating crisp partitions in a correct way [11]. Additionally the FCM is used to acquire the adequate levels of the set clustering parameters [13].

#### F. JT2FIS Framework

1) *JT2FIS Class Library:* JT2FIS is a class library developed for Java. The main purpose is to deploy a library to building interval Type-2 fuzzy inference systems with an object-oriented programming language.

A fuzzy inference system (FIS) is based on logical rules that can work with numeric values or fuzzy inputs, these rules and individual results are evaluated together to form a fuzzy output, then, a numerical value must be passed through a process of defuzzification if necessary.

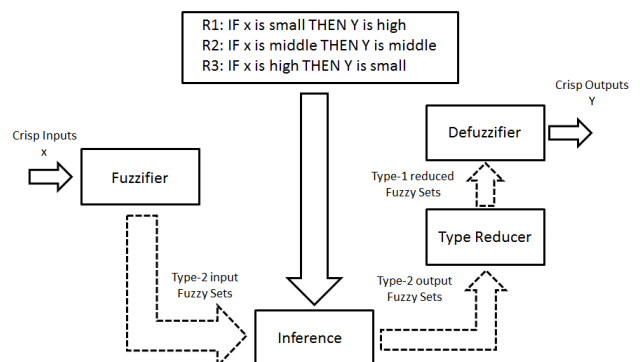


Fig. 2: Type-2 fuzzy inference system block diagram.

JT2FIS architecture, design and implementation is described in detail in [14].

2) *JT2FISClustering Class Library*: JT2FISClustering is a class library developed for Java. The main purpose is to deploy a library to build interval Type-2 fuzzy inference systems with an object-oriented programming language from data mining process. A clustering method is a data mining classic technique used to discover fuzzy sets and rules to configure the FIS from real data. The library implements a fuzzy c-Means clustering algorithm for data mining.

## II. CASE OF STUDY

There is an interest about knowing the location of people in indoor spaces such as shopping malls and museums in order to generate statistics about the interests of visitors in order to guide the new ideas and improvements of the place to cover those estimated interests. Another advantage of knowing the location of a person is to show them relevant or interesting information related to that place. One reason of why a person may spend much time in a zone is probably because that person likes something there.

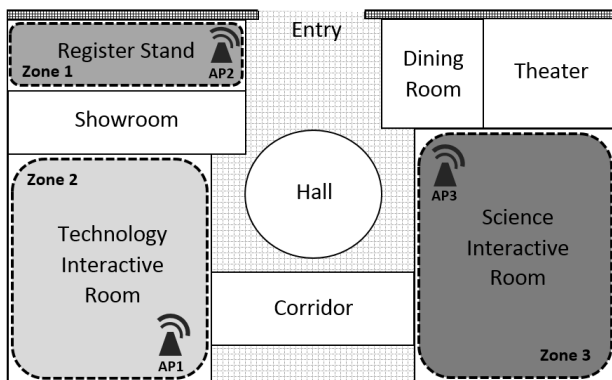


Fig. 3: Interactive museum zones

Therefore, this case of study proposes the location of children inside an interactive museum with an android mobile device using wireless signals from 3 different Access Points (APs) in the area. There are three different zones in Figure 3, zone 1 is the register stand room, zone 2 and 3 are representing two different themed rooms. Moreover, 3 APs are strategically positioned for having enough coverage of the zones to do the location of the kids. The objective is to have enough RSSIs values from the selected APs that describes each desired zone, so a collection of data should be done in order to generate the Mamdani Type-2 FIS to determine the location.

An important detail is that the location outside the selected zones is possible, because when the membership of two outputs is approximately equal and greater than the other output zones, means the child or device is between them two. For example, the case the output of of approximately 50% in zone 1, the output of zone 2 about 50%, and a low membership of zone 3, results of a position in the showroom, meanwhile a 50% in zone 2 and zone 3, could be a zone between them that is corridor or hall. Then a collection and training of data in every area from the place of study is not needed.

## III. MODELLING WiFi DATA WITH TYPE-2 FUZZY SYSTEM

### A. Collecting the Data from Mobile Application

The proposed mobile application uses the JT2FIS class library to implement an inference fuzzy system to do the location process. The system should be trained before it can be used. Wireless signals vary because of different factors, then the environment conditions when values for training were taken should be as similar as possible when doing the location, or should do a new training of the application for best results. The application detects the wireless signals from the three configured APs in the area and shows the percentage of location to the user with the mobile device as shown in Figure 4. Detecting the location of the user, related information of that zone is shown on the screen of the application. Additionally, this information can be recollected and used for the administrator to generate statistics about the most visited zones.

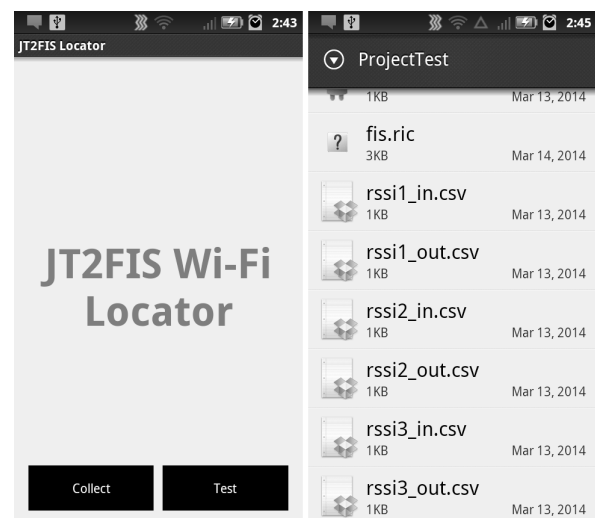


Fig. 4: Location application in Android device

The collected data was taken in three different zones using a function to do a Wi-Fi scan using the Android device searching RSSIs of three selected APs. The collection of every set of data from each zone was based on the time, some factors as the total area of the zone, determine how many data you will need to have a good coverage of valid points for that zone. More quantity and diversity of data can help to have less uncertainty to generate and have a better evaluation of the FIS. The three different inputs obtained for the collected data correspond to each of the three AP signals. At the same time, three different outputs were generated with a value of one in the corresponding zone the input data was collected. These were saved as .csv files. Table I shows a sample of the collected and generated data.

### B. Data Mining the Collected Data

Once the inputs and outputs are obtained the next step is use the JT2FIS class library specific methods inside the Android application to do the data mining in-line. The first to do is the clustering, which consist in detect the different sets or groups of data in all the collected data.

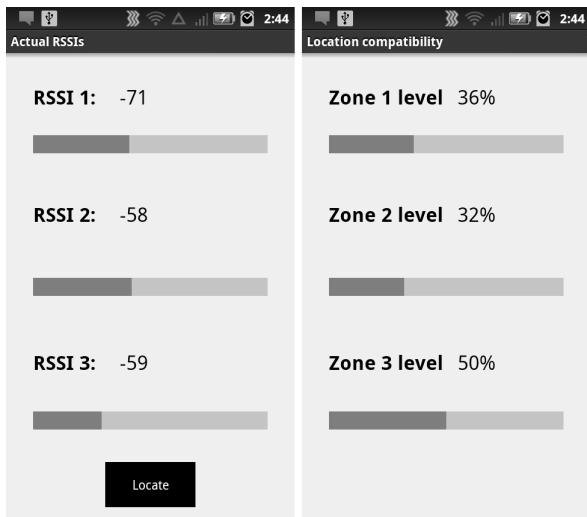


Fig. 5: Location application in Android device

TABLE I: Sample of collected data for each zone and the generated outputs

Zone	Input1	Input2	Input3	Output1	Output2	Output3
1	-78	-74	-54	1	0	0
1	-80	-72	-45	1	0	0
1	-72	-66	-58	1	0	0
2	-76	-60	-58	0	1	0
2	-74	-59	-59	0	1	0
2	-65	-56	-56	0	1	0
3	-65	-66	-64	0	0	1
3	-66	-67	-60	0	0	1
3	-68	-65	-67	0	0	1
...	...	...	...	...	...	...

Listing 1: Fuzzy C-Means and FIS configuration

```
FuzzyCMeans configureFuzzyCMeans =
    new FuzzyCMeans(numberClusters);
GenerateMamdaniFis configureMamdaniFis =
    new GenerateMamdaniFis(configureFuzzyCMeans);
```

Listing 1 shows the configureFuzzyCMeans object that represents the clustering method used based on Fuzzy C-Means to determine the number of clusters of the FIS. If there is already a desired number of clusters, then we can send it as parameter to the object constructor. The second object is required for set the architecture of the FIS and is constructed with the clustering set up from FuzzyCMeans object.

Listing 2: Generate Mamdani

```
Mamdani fis = configureMamdaniFis.
    generateMamdaniFisGaussUncertaintyMeanMember
    Function(inputList, outputList);
```

In Listing 2 the new object invokes the method to generate and return a Mamdani FIS with inputs and outputs obtained from the collected RSSI data in each zone.

Listing 3: Locate new values with FIS

```
ArrayList<Defuzzy> result=
    fis.evaluateFisSingletonCentroid(inputStack,
    points,Fis.AndMethod.MIN,Fis.OrMethod.MAX,
    Fis.ImpMethod.MIN,Fis.AggMethod.MAX);
```

Finally, in Listing 3 the FIS is evaluated to get a location of the user getting first the corresponding input of the values scanned at real time with the device with the application and then the returned array has the grade of membership of the set of inputs with each zone.

In this case there are three clusters of the input data to generate the Mamdani type 2 FIS. In addition, there is the Test option in this example application to obtain three new fresh values of the same APs selected to evaluate the FIS with this RSSIs using the option Locate. The results are three different output values which are the compatibility grade for the input set to each zone, there is an example in Figure 5.

The statistics generated can be used to show patterns about the spent time in each zone for every user. In Figure 6 each concentric circle represents one kid. The 360 degrees represents the 24 hours of a day, then we can see range of times in each zone. The museum availability is from 9:00 a.m. to 7:00 p.m. so the activity of the kids location is always between this time. The colors representing each zone is shown at the right of the circle diagram. There could be richer statistics in the future to identify the modules inside each zone, then determine the exact modules where the kid is spending time.

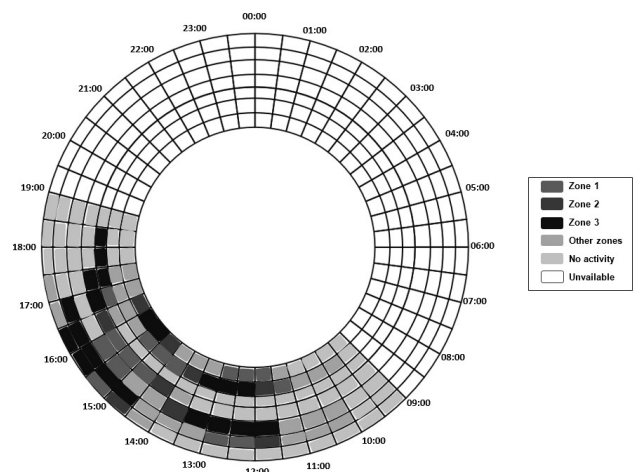


Fig. 6: Circular graphic of visited zones

#### IV. THE TYPE-2 FUZZY INFERENCE SYSTEM

On this section, we are going to show a Type-2 Fuzzy Inference System for the interactive museum case of study. The Figure 7 shows the FIS obtained from c-Means clustering method. We used JT2FIS Type-2 Fuzzy Inference System Visual Component [14] to illustrate the example.

Figure 8 shows an input configuration view of the system. We can see that function members capture input and output uncertainty from collected data

The obtained fuzzy system is used later to evaluate future signals on mobile devices in order to decide the current localization area. Finally, we build a Java fuzzy inference system on Android application in order to validate the approach.

Figure 9 shows a view of how the system evaluates crisp data inputs to get a crisp output that represents the probability of being in the determinate interactive museum area.

We build a Java fuzzy inference system on Android application in order to validate the approach using a Java

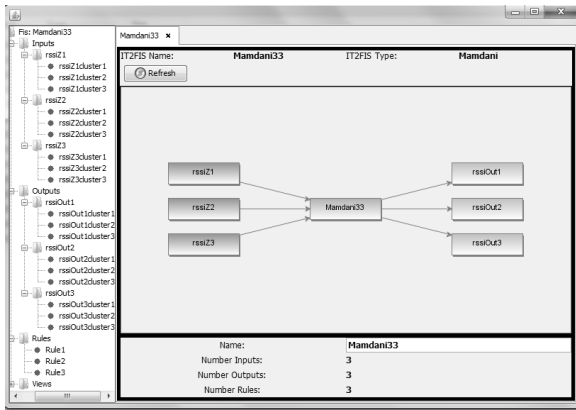


Fig. 7: Fuzzy Inference System for interactive museum example.

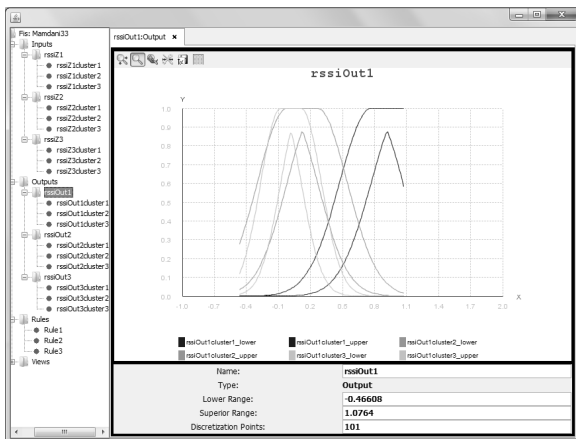


Fig. 8: Input setup for interactive museum example.

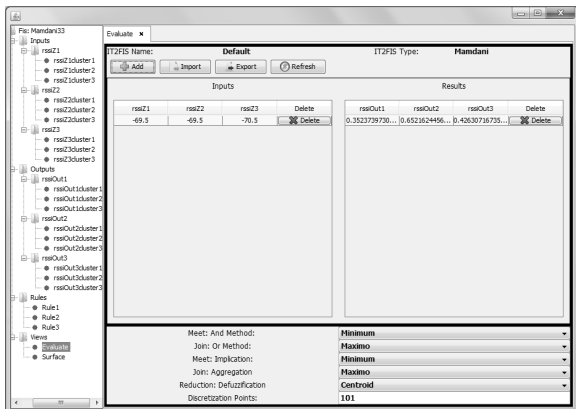


Fig. 9: Evaluation of crisp input data using the obtained FIS to get the location area.

Fuzzy Inference System class library (JT2FIS) [14] (See Figure 4).

## V. CONCLUSION AND FUTURE WORK

This paper introduced a localization in indoor areas using WI-Fi signal networks, and Type-2 fuzzy inference system approach.

On this approach, data was collected by mobile devices that sensed WI-Fi signals from real indoor areas. Then, we built a Type-2 fuzzy inference system using c-Means clustering algorithm. The obtained fuzzy system is later used

to evaluate future signals on mobile devices in order to decide the current localization area. Finally, we build a Java fuzzy inference system on Android application in order to validate the approach.

As future work, we are planning to use a neuro-fuzzy system in order to improve precision on how to produce the fuzzy inference system from data. Increasing the complexity of tests is necessary, so we are planning to add more Wi-Fi signals in order to increment inputs on the FIS, test on different day times, and configure experiments with different Wi-Fi access points locations in order to evaluate optimal conditions.

## ACKNOWLEDGMENT

This work was supported in part by the Internal Fund for Research Projects (Grant No. 300.6.C.135.17) of the Autonomous University of Baja California, México.

## REFERENCES

- [1] "Inspiring the Internet of Things," 2011, URL: <http://www.alexandra.dk/uk/> [accessed: 2014-01-02].
- [2] A. K. Dey, "Understanding and Using Context," *Future Computing Enviroments Group*, 2001.
- [3] W. Chai, J. Zhou, C. Chen, H. Nies, and O. Loeld, "Continuous Indoor Localization and Navigation Based on Low-cost INS/Wi-Fi Integration," *International Conference of Positioning and Indoor Navigation*, 2011.
- [4] E. Navarro, B. Peuker, and M. Quan, "Wi-Fi Localization Using RSSI Fingerprinting," 2011.
- [5] S. Hwang and Y. Donghui, "GPS Localization Improvement of Smartphones Using Built-in Sensors," *International Journal of Smart Home*, vol. 6, 2012.
- [6] S. Chiu, "Fuzzy Model Identification Based on Cluster Estimation," *Journal of Intelligent and Fuzzy Systems*, vol. 2, 1994.
- [7] L. Zadeh, *FuzzySets, Information and Control*, 1965.
- [8] J. Jang, C. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. MATLAB Curriculum Series, Prentice Hall, 1997.
- [9] J. Han and M. Kamber, *Data Mining: Concepts and Teiques*, 2nd ed. Morgan Kaufmann, 2006.
- [10] Two-Crows, "Introduction to Data Mining and Knowledge Discovery," *Two Crows Corporation*, 1999.
- [11] S. Bozkir and E. Sezer, "FUAT - A fuzzy clustering analysis tool," *Expert Systems with Applications, FUZZYSS11: 2nd International Fuzzy Systems Symposium*, vol. 40, p. 403: 842849, 2013.
- [12] J. Bezdek, "FCM: The fuzzy c-means clustering algorithm," *Computers and Geosciences*, p. 1023 and 191203, 1984.
- [13] X. Yin, L. Khoo, and Y. Chong, "A fuzzy c-means based hybrid evolutionary approach to the clustering of supply chain," *Computers and Industrial Engineering*, p. 664 and 768780, 2013.
- [14] M. Castañón-Puga and J. Castro and M. Flores-Parra and C. Gaxiola-Pacheco and L. Martínez-Méndez and L. Palafox-Maestre, "JT2FIS A Java Type-2 Fuzzy Inference Systems Class Library for Building Object-Oriented Intelligent Applications," in *Advances in Soft Computing and Its Applications*, ser. Lecture Notes in Computer Science, vol. 8266. Berlin Heidelberg: Springer, 2013, pp. 204–215.