International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

An Approach for Situation Aware IOT Service based on SOA

Trupti P. Mokal[1], Poonam Gupta[2]

[1]Department of Computer Engineering, G.H. Raisoni College of Engineering and Management, Wagholi, Pune

[2]Professor, Department of Computer Engineering, G.H. Raisoni College of Engineering and Management, Wagholi, Pune

Abstract: ***Since late eighties of the last century Situation Awareness (SA) has started gaining the interest of several researchers from different scientific fields. Significance of Situation awareness (SA) is well established in different sectors such as aviation, armed forces and air traffic. In this paper, a weather department data is considered where weather department person keep entering the daily routine data into situation aware system to a web application on the basis of some attributes like temperature, humidity, wind speed, etc. An automatic Situation aware system validates the weather data in set time for any kind of disaster prediction using following four steps: Preprocessing, k-means clustering, HMM (Hidden Marcov Model) and Fuzzy logic. An existing system is based on Hardware based device like sensor and communicators. There is no backup plan for failure of these devices whereas in this system, event generation is been triggered by well applied data mining techniques. This paper introduces situation awareness due to data entered into situation aware system to a web application. The different patterns based on the effects of event is been identified based on pre-defined protocols using frequent dataset mining at distributed network server. The Publish/Subscribe broker module broadcast event risks to all subscribers using Gaussian distribution model via SMS or email. This paper proposes an automatic situation aware system, alert generator software, due to which an alert arise. This is what an Event-driven is.***

Keywords: K-means clustering, HMM (Hidden Marcov Model), Fuzzy logic, Pattern Identification, Gaussian distribution

1. Introduction

In this modern century, Internet of Things (IoT) can be thought of as billions of devices connected together where the devices have to be intelligent and they need to interact with environment and interact with people. And when we take these three things together and have an intelligent system you can react and do things with benefit, we call this nothing but IoT. For instance, consider one scenario of a weather department data where there are three entities: weather admin, publisher and subscriber. Weather admin keep entering the daily routine data into situation aware system to a web application on the basis of some attributes like temperature, wind speed, humidity, etc. Publisher accesses permission from weather department and subscribers will register to their respective publisher. This weather data is validated for any kind of dis-aster prediction using following steps such as preprocessing, k-means clustering, HMM Baum-Welch and fuzzy logic for decision technique. Then pattern identification is done using frequent dataset mining. Finally, using Gaussian distribution model event risks are broadcasted to all the subscribers via SMS or email.

Hidden Marcov Model (HMM) is a machine learning algorithm which extracts the hidden states by observing the observed states. For instance, consider the observed states such as heavy cloud, low wind and high humidity. Then hidden marcov model observes these observed states and extracts the hidden state. The hidden state will be rain. Similarly, if the observed states are heavy cloud, high wind and high humidity then HMM will observe these observed states and extract the hidden state. The hidden state will be cyclone. HMM is also used in speech recognition system. There are four main phases when constructing a speech recognition system: data prepa-ration, training, testing and analysis. Hidden Marcov Model Toolkit is a toolkit that provides sophisticated facilities for speech analysis, HMM training, testing and results analysis. User can control the modules either defining an option in the command line or defining the desired parameters in a text file. This is how HMM works in speech recognition system. In this way, HMM has achieved remarkable performances in some pattern recognition fields.

K-means clustering algorithm is the classical algorithm which depends on selection of initial center point and de-termination of optimal clustering number. It is required to find initial center selection method of high stability based on partitioning in clustering analysis. If we apply k-means clustering algorithm for above given example, then there will be two clusters formation. Cluster 1 will include heavy cloud, low wind and high humidity whereas cluster 2 will include heavy cloud, high wind and high humidity. Its working is defined in following five steps. First step defines data mining object and purpose. In second step, we need to carefully collect data information at ordinary times. In third step, process transforms the collected data set into a data model; the data model is prepared according to algorithm. Fourth step includes clustering of data which is put to data model into multiple groups composed of similar objects. Finally, fifth step include clustering results analysis.

Fuzzy logic approach is a computational intelligence tech-nique which is a natural way of representing and processing uncertain information. Fuzzy logic has been widely used in control engineering and many other fields of study. In fuzzy logic approach, we are setting numerical values to every observed state. Considering above example its scrutinization is done in five parts: 0-0.2 very low, 0.21-0.4 low, 0.41-0.6 medium, 0.61-0.8 high, 0.81-1 very high. If the time is set to three hours then we check it after every three hours since

from morning 6 a.m. if we started till next three hours i.e. 9 a.m., 12 a.m., 3 a.m., and so on. Once we get value which represents high then there is situation and alert arise. This is the way how fuzzy logic works.

Gaussian distribution model is a probability distribution function suitable for wavelet coefficient probability statistics. It is of bell curve in shape. Gaussian distribution model is used to broadcast event risks to all the subscribers via SMS or email.

This paper is classified as below where section II discusses past works as literature survey. Section III reveals the detailing of our proposed methodology. The evolution of our method-ology is performed in section IV. Finally section V concludes this paper with some scope for future extensions.

2. Literature Survey

This section analyzes many of the past works in below manner.

[1] Introduces an optimization scheme which solved the proposed objective function with global optimality and convergence. This paper proposes a novel multi-view clustering method called Discriminately Embedded K-Means (DEKM). This paper designs a weighted multi-view Linear Discriminant Analysis (LDA). This paper develops an unsupervised cluster-ing scheme to learn the common clustering indicator. This paper proposes an unsupervised clustering framework which embeds multiple subspaces learning into multi-view K-Means clustering to construct a unified framework.

[2] Introduces K-Means algorithm in ideological and polit-ical education management. This paper states that K-Means algorithm depends on selection of initial center point and determination of optimal clustering number. This paper pro-poses a kind of improved K-Means algorithm which is used in clustering analysis of ideological and political education management. In this paper, an experiment was performed to find the initial center selection method based on improved K-Means algorithm. This paper states one problem that number of clustering is uncertain. This has an impact on effect of clustering.

[3] Proposes configurable many-core hardware/software ar-chitecture to execute commonly used K-Means clustering al-gorithm. This paper states that the system is fully scalable and capable of achieving much higher speed-ups by increasing its parallelism. In this paper, proposed solution chooses a center split as it provides more scalable solution. This paper tested the solution for its ability to accelerate the clustering algorithm. This paper performs some experiments whose results showed that the proposed solution can accelerate the algorithm, even with a few number of accelerators.

[4] Proposes an idea of using HMM like palmprint iden-tification with Hidden Marcov Model. This paper employed Down Sliding Window (DSW) sampling technique to ob-tain observation sequences. In this paper an experiment was performed in which 50 individuals, 1000 palmprint images were randomly selected. This paper discusses the influences of HMM and DSW parameters on system performance. This paper states that Hidden Marcov Model is feasible for palm-print identification.

[5] Describes the use of HMM for speech recognition system for the Mongolian language. In this paper, the acoustic and language models based on Hidden Marcov Models are trained. This paper evaluates the performance of isolated word recognition with context independent and context dependent models. This paper describes speech data pre-processing and feature extraction. This paper states that system is not robust due to deficiency of training data.

[6] Proposes an algorithm for adaptive training of a HMM in an Automatic Speech Recognition (ASR) system. This paper introduces the features of HMM. This paper states that HMM is a statistical model which consists of a process that moves on a Marcov chain. This paper states one main disadvantage of HMM method that it requires very large databases for system training. This paper describes an adaptive algorithm which allows the determination of an adaptation policy to a given problem.

[7] Introduces infrared techniques on failure detection in power grid. This paper proposes a method based on wavelet generic Gaussian and maximum posterior probability estima-tion for noise removing of insulator infrared images. In this paper, some experiments were performed whose result indicated that the proposed method can effectively remove the infrared image noise. In this paper, an image denoising algorithm based on generalized Gaussian distribution and MAP estimation is proposed. Also, results showed that proposed method is more efficient than noise filtering.

[8] Narrates that sensor node reporting plays a vital role in the purpose of wireless sensor networks. In this paper, two dimensional Gaussian distribution based on Dynamic Node Deployment model is developed. This paper identifies the directional position based on the angle measurement of the sensor node location. The identified directional position further places the deployed sensor nodes with Gaussian distribution model. This paper carries out simulation on computing the experimental value. This paper states that the usage of 2D-GDDNE model reduces power consumption by 15 % in comparison with existing methods.

[9] Introduces the features of current transport through a barrier in Au/native fluoride/HgCdTe Schottky diode. This paper describes that an ideality factor decreases and barrier height increases with the increase in temperature. This paper states that Gaussian distribution is used to describe inhomo-geneity of barrier height. This paper describes inhomogeneity described by Gaussian function. This inhomogeneity may be related with semiconductor surface potential fluctuations.

[10] Highlights the principles of fuzzy logic applications in the area of channel estimation. This paper is survey on fuzzy logic applications in mobile and wireless communication area. This paper gives special emphasis on growing Long

Term Evolution LTE communication networks. This paper states an overview of fuzzy logic control and its classification. In this paper, fuzzy logic is used to perceive future areas to be investigated further by researcher.

[11] Proposes a two stage fuzzy logic approach. In this paper, the device tries to learn and fit customer habits in order to discover outlier warning signals. The two stages approach builds a reference of its condition. This paper accomplishes a real-time monitoring and analysis of gathered data from body sensors. This paper presents a wearable smart device in the IOT domain when some outliers are discovered.

[12] Discusses current trends in handover research. This paper proposes a fuzzy logic based approach for handover for wireless cellular networks. This paper presents an overview of published work on handover performance and control. This paper describes an application of fuzzy logic for control. This paper evaluates the performance of S-FBHS by simulations.

[13] Explains the interaction between local positioning and IOT (Internet of Things). This paper shows the principle, performance and use of recent local positioning systems in the context of the IOT in detail. This paper develops ultra wideband (UWB) locating system based on a pulsed frequency modulated continuous wave (FMCW) radar principle. This paper states that wireless localization and identification are major source of information for the IOT. This paper produces the best possible localization in any situation.

[14] Proposes a self-organized IOT aware system for online shopping. In this paper, awareness is achieved by monitoring/analyzing the data of users behavior. This paper focuses on analysis of the users eye gaze. The objective is to derive insights about the products the user is interested in or not. This paper aims to make the online shopping system adaptive to its circumstances efficiently and dependably by even non-IT professional shop managers.

[15] Presents a novel trust-based cognitive mechanism making the objects of IOT infrastructure capable to build their situational awareness. This paper uses this knowledge for appropriate reaction to detected threats. This paper demonstrates the efficiency of proposed solution. This paper describes the situation aware system which performs three main functions: Collecting information about environment, trust evaluation and classification and suggesting reactions to identified threats. This is done to ensure communication security.

3. Proposed methodology

Proposed methodology of SOA for situation awareness for IOT can be explained as follows.

Basically there are two servers in this project: one is web server and the other one is situation awareness server. Web server is connected to several machines through Wi-Fi or wireless and these machines can easily access their web pages. This project consists of three entities: weather admin, publisher and subscriber. Weather admin defines services such as cyclone, rain etc. These services are registered by publisher and subscriber registers their respective publisher. Situation awareness server is purely a standalone system and it takes all the weather data and performs following steps: preprocessing, k-means clustering, HMM, fuzzy logic, pattern identification, Gaussian distribution. HMM (Hidden Marcov Model) extracts the hidden states by observing the observed states. For example, if observed states are heavy cloud, low wind and high humidity then hidden state will be rain. And if observed states are heavy cloud, high wind and high humidity then hidden state will be cyclone. Now, if we will make cluster of these then there will be two clusters based on K-Means clustering algorithm. Cluster 1 includes heavy cloud, low wind and high humidity whereas cluster 2 includes heavy cloud, high wind and high humidity. Fuzzy logic is used for setting numerical values to every observed state. After pattern identification is done, Publish/subscribe broker module broadcast event risks to all the subscribers using Gaussian distribution model via SMS or email.

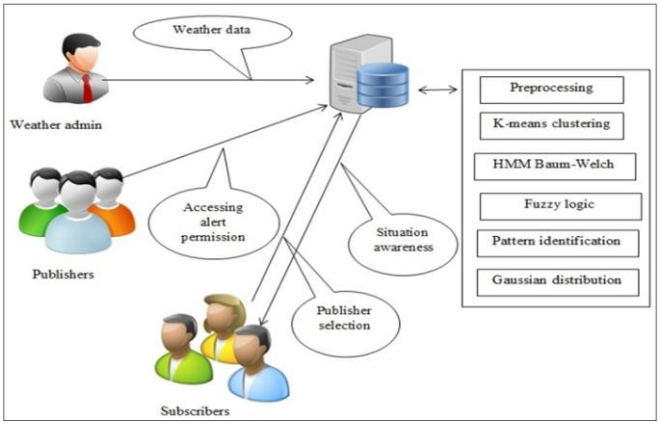

Figure 1: System Overview of Proposed methodology

System architecture represents a three-tier architecture which consists of three layers: Presentation layer, Business layer and Data access layer. Basically there are three entities: Weather department admin, publisher/broker and subscriber. Weather admin defines services such as cyclone, rain etc. These services are registered by publisher. Subscriber under these publishers will register to their respective publisher. Publisher and broker are same entity, who accesses permission from weather department. These three entities are getting data from data access layer. From coordination of event generation module and database, event is generated. These events are met to business layer. Finally, we are getting data through GUI in presentation layer (i.e. registration, login, etc.)

Here, first of all proposed system designs a situation of weather report distribution for different entities. Here in the system mainly three roles are considered like Weather Admin – offers type of services such as fisherman, farmers, cricket association. In case of cricket association, accordingly test matches are started. In summer days it start early and in winter it starts late. Based on these there are different people who need weather data. These sectors are registered with weather admin and they are called as broker.

Secondly publisher is the entity who takes the service of the weather admin based on the requirement like formers agency for rain or cyclone, Fisheries agency for cyclone and many more. Finally subscribers are the last entity in the hierarchy who takes the services of the particular publisher or the broker. There can be multiple brokers

Once the system is triggered weather data from database is done on the basis of the requirement by selecting and labeling required attributes.

After that using K- means clustering this preprocessed data is allowed to form cluster to get the abstract clusters. These clusters are been then feed to the machine learning model to identify the hidden states of the situation.

Baum –welch algorithm is used to extract the hidden states from the clustered data. Initially forward probability of the data is been identified using equation 1

$$\alpha = O\ (t+1) \sum_{j=1}^{N} \mathrm{Osi(O1)} \qquad (1)$$

Where α = Forward Probability
O(t+1) = For every next element of Observed state
N= Size of the Observed state List
$\mathrm{Osi(O1)}$ = For Every observed state

Then, backward probability is also identified using equation 2

$$\beta = \lambda\, O\ (t+1) \sum_{j=1}^{N} \mathrm{Osi(O1+1)} \qquad (2)$$

Where β = Backward Probability
O(t+1) = For every next element of Observed state
N= Size of the Observed state List
$\mathrm{Osi(O1)}$ = For Every observed state
λ = Frequency of the occurred state

Then by using forward and backward probabilities hidden states are been evaluated using Baum-welch algorithm as shown in the below algorithm 1.

Algorithm 1 : BAUM- WELCH ALGORITHM

// Input : Data Set **D**,
Observed States $\mathbf{O_s}$ = { O_{s1} , O_{s2},O_{s3}}
Step 0: **Start**
Step 1: Identify the Observed state Attribute $\mathbf{O_{si}}$
Step 2: **FOR** i=0 to size of **D**
Step 3: Identify Attribute $\mathbf{O_{si}}$ and put in separateList $\mathbf{O_{SL}}$
Step 4: **END FOR**
Step 5: Transaction count $\mathbf{T_c}$=0
Step 6: **FOR** i=0 to size of $\mathbf{O_{SL}}$
Step 7: identify $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$
Step 8: Compute $\boldsymbol{\gamma}$ using Equation3
Step 9: IF $\boldsymbol{\gamma}$ belongs to O_s
Step 10: **THEN** add H**s (Hidden State)** to list
Step 11: END FOR
Step 12: **Stop**

Once the hidden states are been extracted then they are scrutinized for the patterns using pattern mining definition. Using fuzzy classification theory, classification of these identified patterns is done for proper situation awareness identification

Once the situation is identified then its distribution is done among the publisher and subscribers on the basis of the past regressions using Gaussian distribution theory. The Publish/Subscribe broker module broadcast event risks to all subscribers using Gaussian distribution model via SMS or email.

4. Implementation

Basically there are two servers, one is web server in which Apache Tomcat is installed and the other one is message server plus mail server. Both are connected to the switch. Also three entities like weather admin, publishers and subscribers are linked to this switch. Weather admin is working in weather department. He will login into the system and a panel is provide to him so that he will add all the details of the day like date, time temperature, humidity, wind speed etc. These entities will be stored in database i.e in webserver machine. There are two roles given web server and situation awareness plugin. This situation awareness tool is triggered for set time. First of all data is preprocessed, that means we are fetching required attributes. Then clustering is done as shown in fig 3. In this process first of all labeling is done. For that purpose four attributes are selected. Those attributes are temperature, visibility, humidity and wind speed. A specific condition is set for each of these attributes. For example, for temperature if >20 and <34 then labeling is -2. For visibility if medium then 0 else 1. For humidity and wind speed if high then labeling is 1 if medium then 2 and if low then 3.Then we need to compute Euclidean distance then find smallest and biggest Euclidean distance. After that find range between these smallest and biggest. In this way we get five clusters by using k-means algorithm.

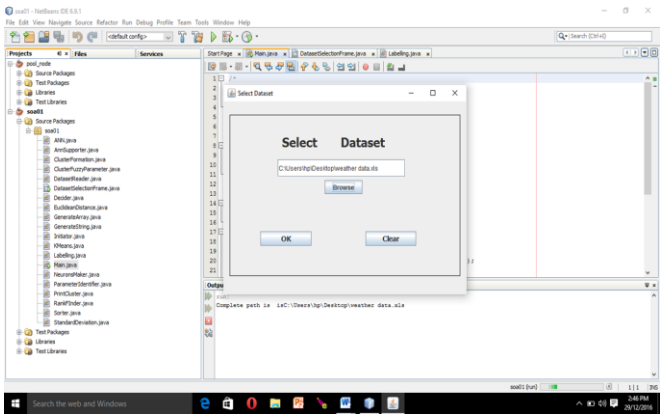

Figure 2: Accessing data from weather data.xls.

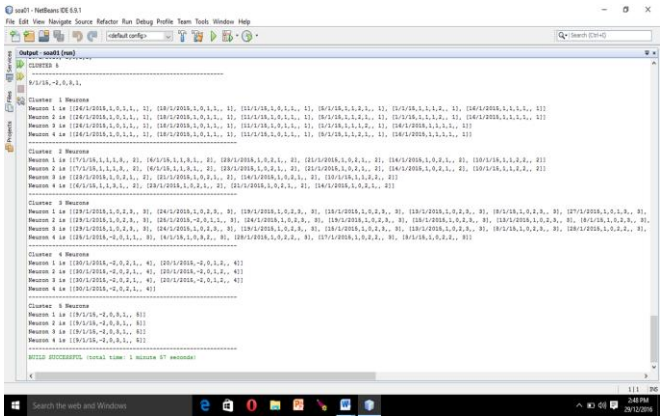

Figure 3: Cluster formation by k-means clustering.

5. Conclusion

In today's world situation awareness remains as the most appreciated and highly demanding area. Due to tremendous growth of population and thereby need of technology always arises the risk factors in many sectors. So predicting the situation awareness without using any hardware accessories becoming major research area.

So the proposed approach of evaluation of situation awareness in IOT using SOA by incorporating strict protocols of machine learning algorithm by utilizing timely data of a system adds fuel to the situation awareness system. This system can be enhancing in the future by incorporating multiple scenarios in distributed paradigm to create vast architecture of situation awareness in IOT using SOA.

References

[1] Z. Liang, W. Taiqing , "DSW feature based Hidden Marcov Model: An application on object Identification", 2011 International Conference of Soft Computing and Pattern Recognition (SoCPaR), Dalian, pp. 502-506

[2] A. Ayush, B. Damdinsuren, "Design and Implementation of HMM based Mongolian Speech Recognition System", Ifost, Ulaanbaatar, 2013, pp. 341-344, doi:10.1109/IFOST.2013.6616910.

[3] D. Shaked, A.Cohen, "Speaker adaptation for Hidden Marcov Mod-els",Sixteenth Conference of Electrical and Electronics Engineers in Israel, vol. 9, Issue 2, pp. 1-4, 2013.

[4] J. Xu, J. Han, "Discriminatively Embedded K-Means for Multi-view Clustering",2016 IEEE Conference on Computer Vision and Pattern Recognition, pp. 5356-5364, doi:10.1109/CVPR.2016.578.

[5] L. Jianzhu, L. Jing, "Ideological and political education manage-ment research based on Improved K-Means Clustering", 2016 Inter-national Conference on Robots and Intelligent System, pp. 213-216, doi:10.1109/ICRIS.2016.114.

[6] J. Canilho, M. Vestias and H. Neto,"Multi-core for K-Means Clustering on FPGA", 2016 26th International Conference on Field Programmable logic and Applications (FPL), Lausanne, 2016, pp. 1-4.

[7] A. Mudassir, S. Akhtar, "A survey on fuzzy logic applications in wireless and mobile communication for LTE networks", 2016 10th International Conference on Complex, Intelligent and Software Intensive Systems, doi 10.1109/CISIS.2016.106.

[8] P. Raimondo, F. Rango, "A two stages fuzzy logic approach for IoT wearable devices", 2016 IEEE 27th Annual International Symposium on PIMRC, pp. 1-6.

[9] T. Inaba, D. Elmazi, "Application of fuzzy logic for secure Handover in wireless Cellular networks", 2016 10th International Conference on Complex Intelligent and Software Intensive Systems, pp. 509-515.

[10] H. He, W. Lee and D. Luo, "Insulator Infrared Image Denoising Method Based on wavelet Generic Gaussian distribution and Map Estimation", 2016 IEEE Industry Applications Society Annual Meeting, Portland, OR, pp. 1-6, 2016.

[11] Rajaram , Prakasam, "Self-Motivated Node Exploitation using Two-Dimensional Gaussian Distribution in WSN", International Conference on Electrical, Electronics and Optimization Techniques (ICEEOT), pp. 1030-1034, 2016.

[12] E. Zakirov, V. Kesler, "The Gaussian Distribution of Barrier Heights in Au/Native Fluoride/nHgCdTe Schottky Diodes", 17th International Conference on Micro/nano technologies and Electron devices EDM, pp. 11-15, 2016.

[13] F. Kirsch, R. Miesen and M. Vossiek, "Precise Local-Positioning for Autonomous Situation Awareness in the Internet of Things", 2014 IEEE MIT-S International Microwave Symposium (IMS2014), Tampa, FL, 2014, pp. 1-4.

[14] Y. Yamamoto, A. Yoshitaka, "IoT-aware Online Shopping System En-hanced with Gaze Analysis", 2016 World Automation Congress (WAC), Rio Grande, 2016, pp. 1-6.

[15] J. Glowacka, J. Krygier, M. Amanowicz, "A Trust-based situation awareness system for Military applications of IOT", 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT), Milan, 2015, pp. 490-495.