PRAJÑĀ - Journal of Pure and Applied Sciences Vol. 24 – 25 : 209 – 218 (2017)
ISSN 0975 - 2595

BENHANCING FOG COMPUTING WITH USERS PROFILE AND TAXONOMY OF EXPERIENCE

PRITI SRINIVAS SAJJA

GH Patel PG Department of Computer Science & Technology, Sardar Patel University, Vallabh Vidyanagar-388120
Email: priti@pritisajja.info

ABSTRACT

Modern information and communication facilities have opened up new dimensions in handling day to day transactions as well as spectacular activities. With the help of web platform, problem solving and decision making have become very effective. The Web has enabled use of remote resource such as hardware and software efficiently with the help of cloud computing paradigm; which has its own limitations. To overcome limitations of the cloud computing, fogging is introduced in this paper. This paper discusses fog computing, work done in the area, and proposes a generic framework for hybrid fuzzy fog computing. The architecture works in conjunction of users' profile and user experience defined with the help of fuzzy logic. To evaluate the proposed architecture, an experiment of the customized web content filtering is discussed with necessary technical details. The experiment also utilizes the artificial neural network. Technical details of the underlying neural network such as structure, training data set, etc. are also provided in this paper. At the end, the paper concludes with advantages, applications and possible future enhancements..

Keywords: Fog Computing, Users Profile, Fuzzy Logic, Neural Network, Web Filtering.

INTRODUCTION TO FOG COMPUTING AND STATE OF THE ART

Cloud computing has become buzz word in modern scenario as it provides facilities in various aspects such as platform, processes and infrastructure to many computing applications in order to store, use, and manage voluminous data/processes on network. However, in spite of benefits like readily available software, data and infrastructure, cost-effectiveness, portability and mobility, etc. there are issues too. The major issues with cloud computing is losing control on ones data, security, customization, reliability, and lack of standard problems. People would not like to transport their data on some unknown computing resources. Further, cloud computing needs networking infrastructure as bandwidth, which is poor sometimes. Above this, Internet and network latency also increase intensity of the issues.

When it comes to Internet of Things (IoT) and Internet of Everything (IoE), and many other personal applications based on domestic devices; they need data &other computing functionalities at the edge, not on the far cloud. Instead of centralized repository of various facilities, there is a need of decentralized computing as per the customized requirements of that edge. To facilitate this, technique of edge computing is suggested. As said, cloud may be logically far, but 'fog' is near earth, and hence the word fog computing/edge computing has evolved. Since such edge computing operates on end of the network, preferably at client's side, it eliminates need of establishing the full-fledged and dedicated channels for cloud. The term 'fog' is introduced by CISCO while suggesting new model of wireless computing. As per their view, "fog computing extends the cloud computing paradigm to the devices and applications, that do not conceptually fit into the paradigm" [1].

Figure 1 shows an architecture of the fog computing.

As stated, fog is comparatively nearer to the local devices and performs computing with less latency and quick response utilizing less bandwidth. Following are the major advantages of fogging.

Security and privacy: As the data are kept at the edge, not on the far cloud, security and privacy is better.

Efficiency: Since the fog node is at the edge, data can be efficiently communicated.

Decentralized communication: Instead of keeping all the data at the centralized server, the edge possesses the important data.

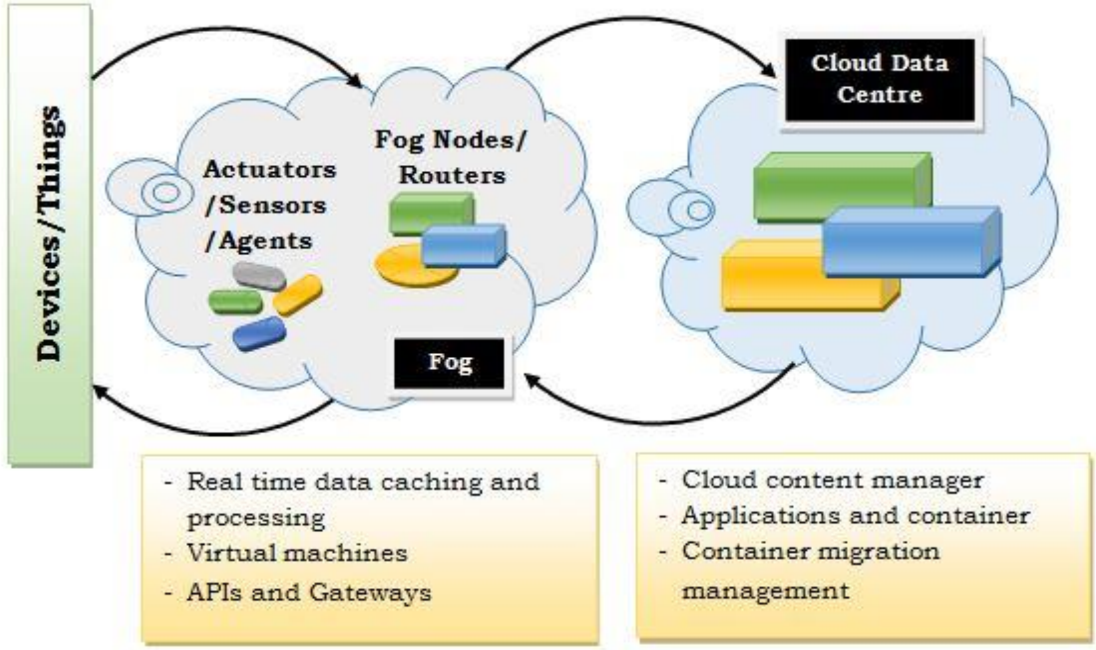

Figure 1: Fog Computing Architecture

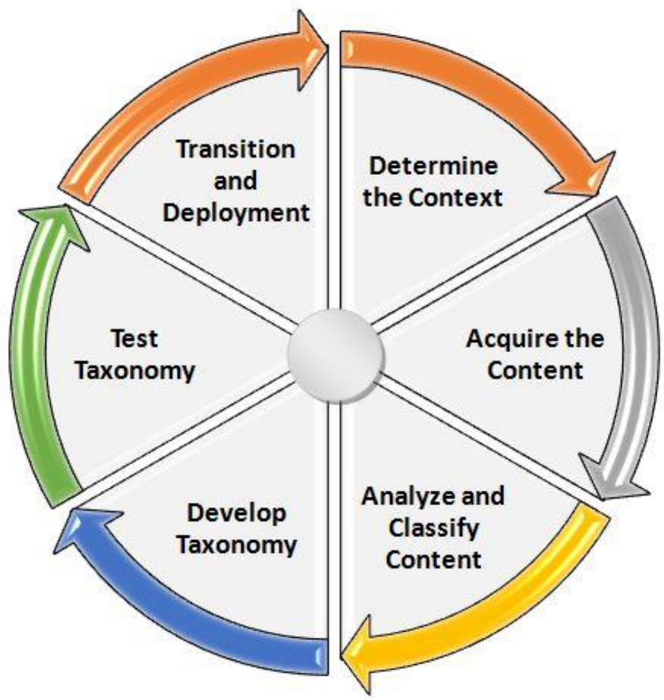

Figure 2: Lifecycle of Taxonomy

Modular approach: In between the client and server, there is an extra layer introduced, as an extension of the local network, which gives the advantages of structured approach.

Customized computing: Fog computing infrastructure allows development of customized systems for user.

Dependency: As per the OpenFog Consortium[1], fog computing is least dependent on protocol or radio systems, instead it is conceptual and resides at systems level.

Following are the some major areas, where fog computing is beneficial.

Opportunities in Pure Computer Science and Engineering

- Web filtering and content mining
- Mobile applications related to distributed intelligence
- Semantic web
- Monitoring and intrusion detention at end devices such as gate sensors, mobiles, etc.

Applied domains

- News summarization
- Health awareness and informatics
- Smart cities applications such as Smart traffic lights
- Self-maintaining transport and connected cars
- Tourism
- Crowd monitoring, etc.

Attracted from the advantages and versatility of the fog computing, many researchers have contributed; majority of them have chosen area of IoT/IoE and distributed intelligence as discussed by Flavio Bonomi and Rodolfo Milito [2]. Theoretical models and issues related to fogging are discussed by many researchers [3], [4], [5], [6], [7], and [8]. Fogging is also used for smart cities and vehicle monitoring applications by Xueshi Hou et al., [9]. Similar application for IoT is demonstrated by Antonio Brogi [10]. Further, the fog computing has become first choice for researcher working for 5G systems [11]. Modern scenario of the fog computing is surveyed by Ivan Stojmenovic & Sheng Wen [12] and Janakiram MSV [13].

Other approaches/techniques related with the fog computing are also been in practice. To name a few, "cloudlet: data centre in a box" and micro-data centre. Cloudlet was introduced at Carnegie Mellon University [14]. Similarly, Microsoft[2] also suggested use of micro data centre (modular data centre).

TAXONOMY, ITS APPLICATIONS AND STATE OF THE ART

Taxonomy is defined as systematic documentations of various classes of entities (things or concepts) under domain of consideration. For example, all biological species or organisms can be defined and documented in systematic way. Similarly for numbers also, there exist various taxonomies. In the domains such as economics, education, medical science, military and defense, taxonomies are used. Many of us have used the ACM Computing Classification System (CCS), which is classification system for various computer related subjects by Association for Computing Machinery (ACM).ACM has implemented the classification system in 1964, and now using the latest version[3] of it.

Taxonomy of experience is a field which is less traversed, when it comes to the generic classification. Application/domain specific taxonomy about experience of entities are generally developed and used. Long back in the year 1983, a canonical model of the emotional experience was proposed by Eleanor Daly et al., [15]. Taxonomy of prospection (envisioning future in closed domain] is also proposed by Karl Szpunar et al., [16]. Latest work includes modeling of taxonomy of news items [17].

Typical life cycle of taxonomy encompasses variety of phases, namely context and content determination, content acquisition, analysis and classification of the content, developing & testing taxonomy, and deployment. These phases are denoted in Figure 2.

[1]*https://www.openfogconsortium.org*

[2]*https://www.microsoft.com*
[3]http://www.acm.org/about/class/

To generate taxonomy of experiences of users related to webpages, following major aspects have been considered. This information can be collected from a well-organized survey on the target audience and formal documentation of experience of users can be done to generate sophisticated taxonomy of experience. Each entity is encoded and concern code is reoffered in user's profile.

- Topic and context
- Nature of content (reliable, efficient, timely, and useful)
- Type of content (research, teaching, general, advertisement, entertainment, personal, news, etc.)
- Sub-content
- Relevance of the content
- Length of content
- Format and interface (multi media) and copy editing, readability
- Level (as per user) and problem solving capacity
- Connection to social media and similar facilities, internal and external links, and comments
- Interests

The taxonomy may incorporate fuzzy words on need, as some of the features, may be vague and uncertain. Such fuzzy words are incorporated with help of fuzzy logic; which is a multi-valued logic invented by Lotfi Zadeh [18].

USE OF TAXONOMY OF EXPERIENCE FOR FOG COMPUTING

In this paper, a personalized taxonomy of experience of users for an entity or concept in a given domain is designed and utilized for further computing processes in order to make it more effective. The basic objective to encode experience of the users is to present useful information to the users in highly customized manner without giving unnecessary details. The taxonomy of experience of users encompasses various aspects such as domain and sub-domain information, description, usability, resources, cost, other users, issues, outcomes, risk involved, etc. These fields are generic in nature. Besides these fields, one also needs to encode domain specific experience. For example, if health awareness and monitoring is required, field like disease, symptoms, remedies, history of treatment, etc. are to be added. Some of the fields in general as well as domain specific categories are again vague and difficult to be measured in standard predefined units. To document such information fuzzy logic help is taken. With this, not only vagueness and uncertainty are managed, but also increases ease of operation. Here machine tries to identify meaning of the linguistic parameters, and hence enables machine learning. Later, based on the field experts' opinion and domain heuristics, classification of the collected experience information is done to generate taxonomy; which is partly domain dependent.

Once repository of the taxonomy of the experience about the target users in a domain is designed, it is ready to be encompassed with the fog computing framework. Along with repository of taxonomy, use of users profile repository is suggested. Basic idea behind such repository is to provide customized output as per users need. Along with the code of the taxonomy of the experience from the repository, a profile of users containing personal information, business information, web log/history, interests, priorities, habits, and preferences, health related information such as diagnosis, treatment history, etc. is also used. The users profile repository is generally stored at the client machine to save time and make the process more secure by hiding the users' personal information. The proposed framework of the fog computing in conjunction with the fuzzy user profile and taxonomy of experience is illustrated in Figure 3.

The fog computing framework takes the users query and validates it after necessary preprocessing such as correction of spellings and stop words deletions. Such preprocessing makes the query well formatted. The query then passes through the frequently visited pages (history - FVP) to check similarities. If similar result is readily available, instead of sending the request to cloud, the ready result is submitted to the user. Here the query may use fuzzy linguistic words, which are application specific and dependent on domain selected. The fuzzy user profile is also unique for every user and contains sufficient information by which the system knows the user. The fuzzy profile also consists of experience of users about entities of domain. Since it is challenging to store description of experience, which is machine readable, here the experiences about the events and concepts (called entities) are encoded as per repository of taxonomy. This is a kind of database normalization, achieved by generating dedicated taxonomy of experience, encode them and the code is used in user profile.

Second layer plays key role in understanding the fuzzy linguistic words used in the query with help of fuzzy membership functions after separating them by tokenizer. Further, the user profiles as well as the taxonomy are encoded in XML format, so as to make them machine readable. For that, we need metadata and Document Type Definitions (DTD).Activities related to decision making and problem solving are supported by fuzzy rules and inference engine[19]. The decision taken can be converted into crisp output and send to the application specific utility to the layer 3.

The fog computing framework demonstrated in Figure 3 is generic in nature and can be applied to any domain suitable for edge computing. One needs to define the domain specific applications, fuzzy membership function and rules, meta data and bag of words (synonyms). To demonstrate the working of the above mentioned generic framework, an application of web content filtering is selected and discussed in detail in the following section. The user profile and taxonomy of experience related to the domain are also defined and discussed.

EXPERIMENTING THE PROPOSED FRAMEWORK IN THE PERSONALIZED CONTENT FILTERING

Web has become a giant repository of everything. Earlier, the web was read only; great people used to contribute to the Web and many reads from it. Later the Web has become 'read-write' and people can write to and read from the Web. With the innovations of modern, secure, and platform independent languages such as java, it has become easy to execute some functions of Web; hence the Web has become 'read-write-execute'. So now, the Web has become repository of everything by everybody for everybody! Most of the users surf Web for content (content mining in a way); however, the web log mining, and web structure mining also in practice. For customized web content filtering, the system must know its users and presents the customized content, specific to users need [19]; [20].Further, for every user the output of the filtering must be unique and tailor made. In Figure 5 the proposed approach of the fog computing framework for effective and personalized content filtering is illustrated.

As stated, the system must 'know' its users. For that the user profile is created, which contains information on various aspects such as:

- Information specific to users such as age, gender, job type and location
- Users job and activities
- Users education information
- Family and social information (optional)
- Users history of surfing
- Users interests and hobbies
- Personal preferences
- Sequence of experience codes

It is to be noted that values of some of the fields are fuzzy in nature, and need to be represented with fuzzy linguistic variable.
The profile is encoded as shown in Figure 4.

The example illustrated in Figure 4 shows user's interest is in the field of "Research" with "High" value. This high value is fuzzy and can be encoded with help of fuzzy membership functions as shown in Figure 4.Once the user profile is encoded and necessary fuzzy membership functions are defined along with taxonomy of experience encoding, next step is to finalize the application logic. Here to meet the objective of customized web content filtering, an Artificial Neural Network (ANN) is used. Query given by users is first checked with the client local memories for frequently visited pages and history, if no match is found, the keyword of the query will be separated and given different weights by refereeing fuzzy rule base and bag of words. The normalized input vector is generated and sent to the underlying ANN. Since there are no generalized rules available for the domain, with large amount of the sample data only the web page filtering can be done. The ANN must be trained with good quality of valid vectors with the back propagation algorithm in supervised manner. The fuzzy variables are used directly in rules, which have structure as shown in Table 1[21].

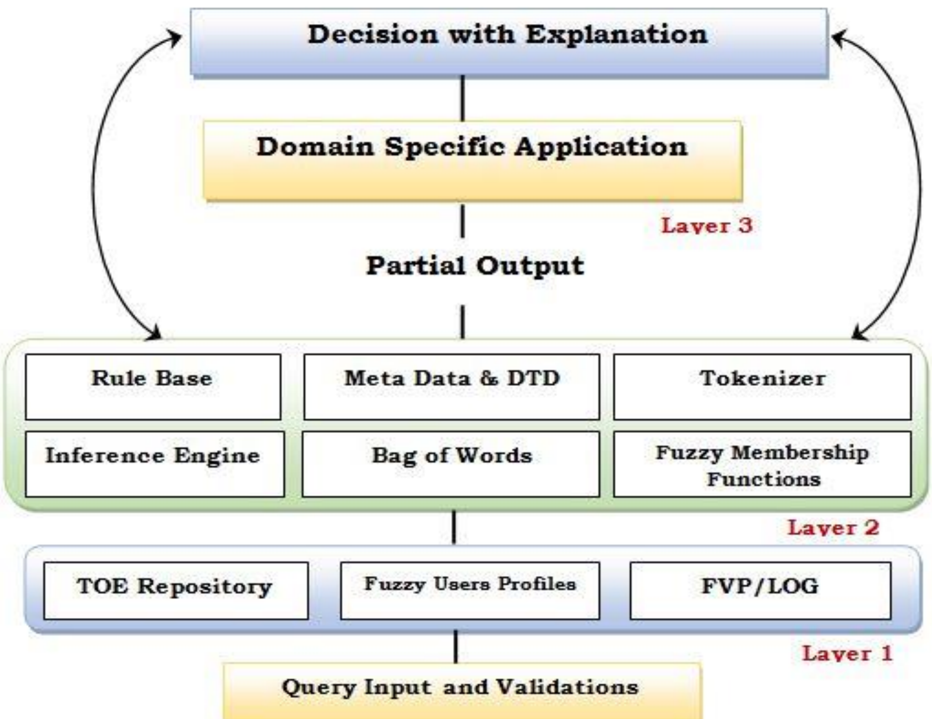

Figure 3: Proposed Fog Computing Framework with Taxonomy of Experience and Fuzzy Users Profile Repositories

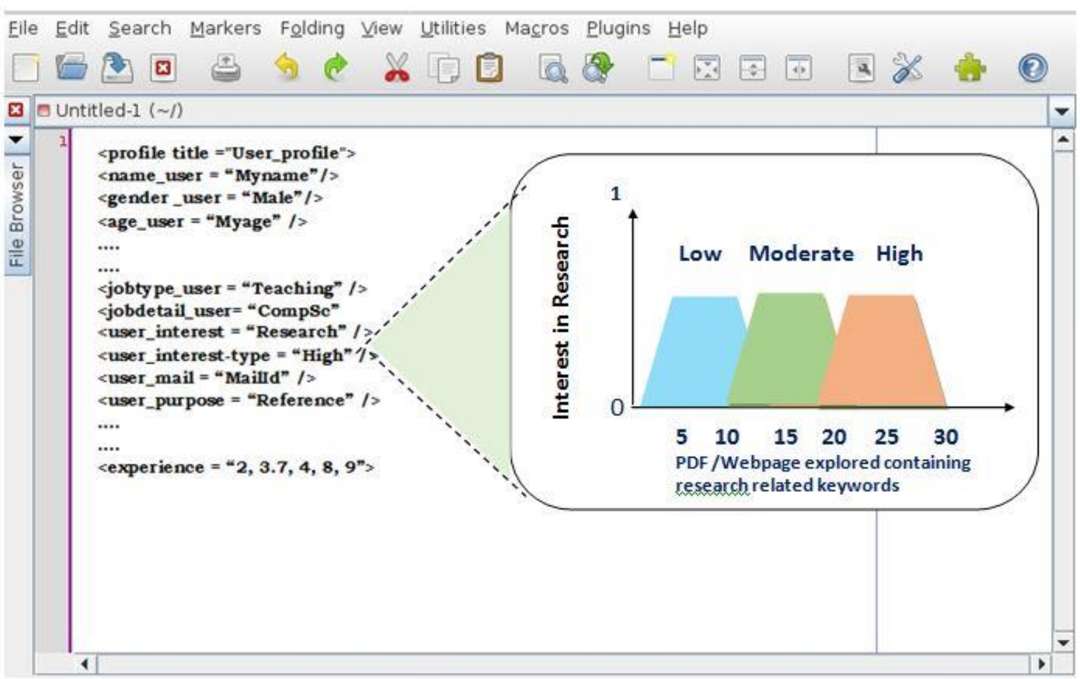

Figure 4: Fuzzy User Profile and Sample Membership Function

The fuzzy rule schema can be referred as IN Table 2.

Table 1: Fuzzy Rule Structure

```
<xs : schema id="Fuzzy_Rule" elementFormDefault= "qualified" xmlns :xs="http://www.w3.org /2001/XMLSchema">
<xs:complexType name = "Fuzzy_Rule">
<xs : sequence>
    <xs : element name ="rule "
          <xs:element name="antecedent" type="xs:string"
                <xs:element name="variable" type="xs:string"/>
                <xs:element name="value " type="xs:string"/>
                <xs:element name="acc_modifier" type="xs:string"/>
                ……../>
          <xs:element name="consequent " type="xs:string"
                <xs:element name="generally" type="xs:string"/>
                <xs:element name="specifically" type="xs:string"/>
                ……../>
    />
</ xs : sequence>
</ xs : complexType>
</ xs : element>
</ xs : schema>
```

Table 2: Fuzzy Rule Schema

```
"<xs:schema id="Sample" elementFormDefault="qualified"
xmlns
:xs="http://www.w3.org/2001/XMLSchema">
<xs:includeschemaLocation = "Fuzzy.xsd"/>"
```

The neural network classifies the web content by features and query provided into three categories in this application. The categories are recommendable, presentable, and rejected. Based on the users choice/level the web content/pages can be presented to the users. Optionally, with the help of fuzzy layers of the framework, the justification of the decision taken is also provided. The overall framework with application is illustrated as in Figure 5.

To prepare training data set, data are collected from various users about the choice of the suitable webpage and its placement into three pre-defined classes namely, recommendable, presentable, and rejected. Once users provide sample data, about query, webpage and its class (where the user thinks the webpage should belongs to); noise will be cleared from the data, synonyms of the keywords are found. The keywords are further given weights depending of their importance or frequency of occurrence. Locations of the keywords are also considered; for example, if the keyword is in title of the page, the weight is higher. Let the frequencies are f_1, f_2, $f_3,\ldots,f_n$ for the respective keywords $q_1,q_2,q_3,\ldots,q_n$. Following vector of normalized values will be considered as input to the ANN.

Training data set 1

{ (F), (Q), (D) }

The above training data contains the values of frequencies (F) $f_1,f_2,f_3,\ldots,f_n$ and (Q) $q_1,q_2,q_3,\ldots,q_n$. This is generalized equation; actual training data set contains real domain specific values and sufficient number of vectors to make base ANN learn from the data. (D) denotes category, here in the experiment, three categories are chosen. Category d1 stands for "Recommendable", d2 for "Presentable", and d3 for "Rejected".

Examples of some actual training data set used are as follows.

{ (1,2,0,1,2,1…), ("Xbook", "video", "age", "game", "birthday", "like", …), 0, 0, 1}

Here first two subsets represent input data and last three values represent sample output category. Further, "Xbook:" is name of site which is not permitted (eg. Visiting social networking sites during office time on office client/machine) and to be rejected.

As the web pages are to be classified in three categories, for preferable category, the output digit is 1, for other it is zero. In the above mentioned training data, the webpage is to be filtered out (rejected) and would not be presented to the user.

Here is another set of sample data.

{ (3,2,1,0,1,1…), ("SPU", "Semester", "Examinations", "Result", "Festival", "Course", …), 1, 0, 0}

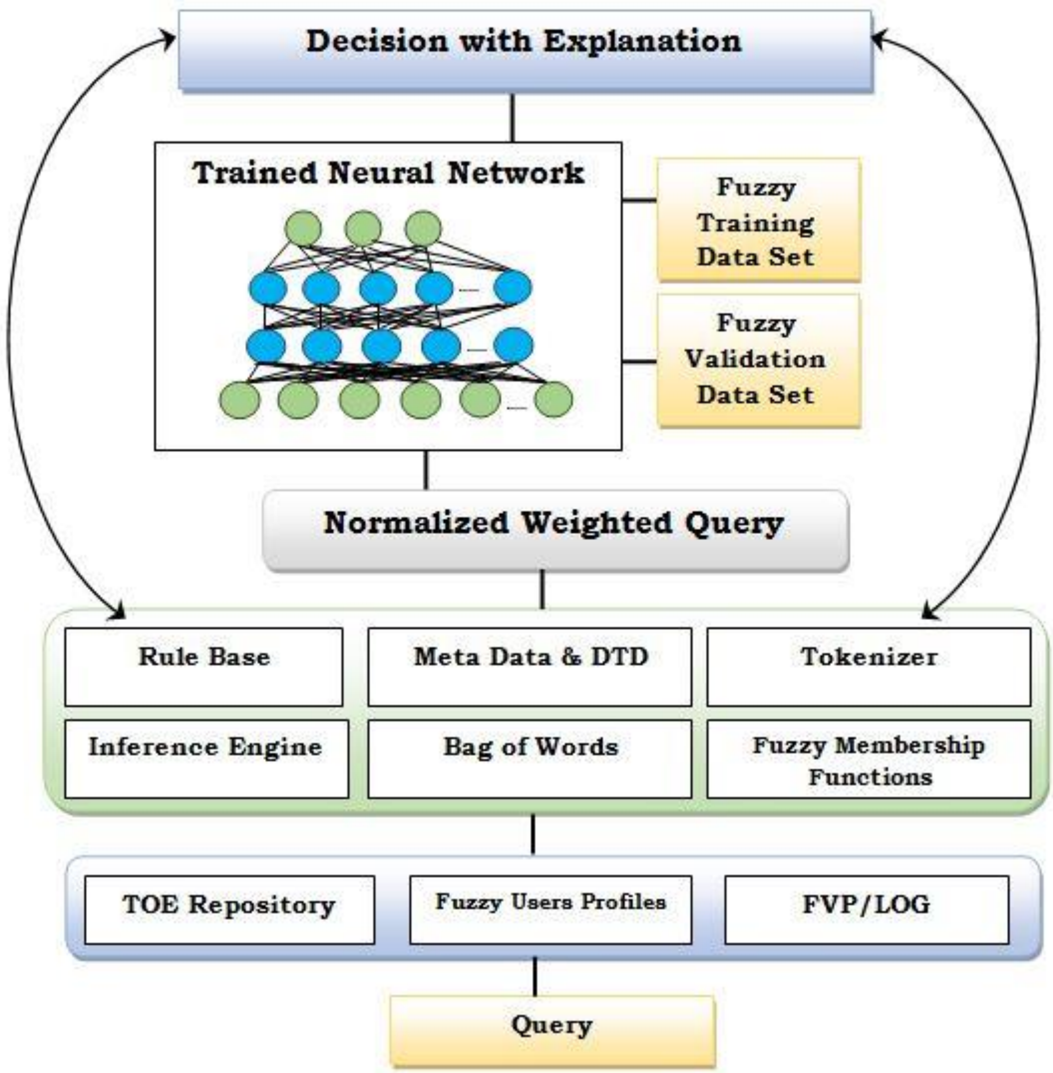

Figure 5: Personalized Web Content Filtering from Web through Fog Computing Framework

While training, the ANN considers each set of training data one by one and learns how to classify webpages into recommendable, presentable, and rejected categories. First the ANN takes only input data and calculates what it "thinks"! Later its output is compared with the actual output provided within sample data. It goes back, adjusts its weights of connections and reprocess in forward direction. After many such forward and backward passes, the ANN is able to categorize only one webpage. Such multiple and good quality training data sets are needed for completely and thoroughly trained ANN. Training data sets may vary as per the application. The base neural network is shown in Figure 6.

Later, when actual values are extracted by the fog node or at the client machine, the underlying ANN has only input. As per its training, the ANN provides correct output, classifies the webpage and presents only recommendable webpages. The experiment of webpage filtering will be helpful in many aspects. Such system can be used in institutes to identify web content which is malicious and prohibited. It has got domestic applications also, in order to block some unwanted sites, as per the parents/authorities need; automatically the web content will be put into rejected category and not shown to the users.

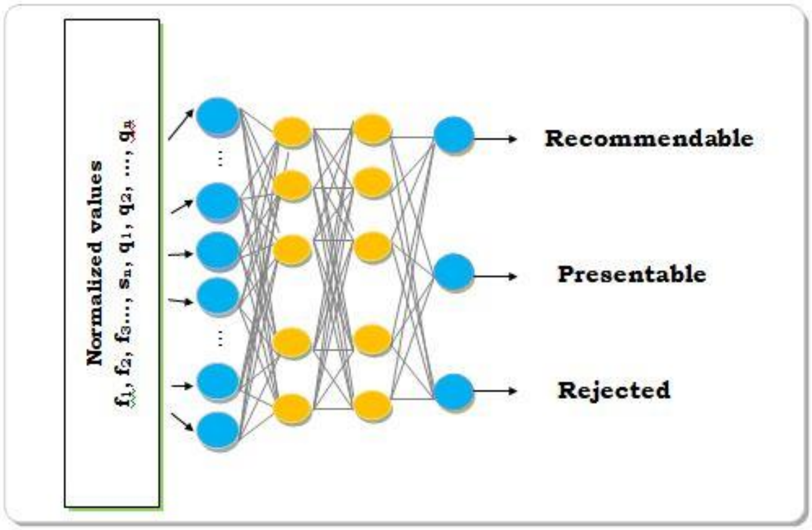

Figure 6: Underlying Artificial Neural Network

CONCLUSION

The architecture proposed is generic in nature and can be used for any applications. For demonstration, web content filtering is taken into consideration. Other example applications, which can be benefited with the proposed architecture, are as follows.

- Health care,
- Smart cities,
- IoT based systems,
- Customized data and Web mining,
- Mobile fogging,
- Vehicle and object tracking,
- etc.

At the edge of clients network/devices when cloud computing is to be used, one can go for such light weight cloud computing via fog. Open source software and forums such as fog project[4] and OpenFog are also available to guide developers.

In future, one can work on distributed intelligence systems for fog computing, in order to make fogging efficient and cost-effective. Fog computing research can also involve research in 5G systems, fog software agents, generic fog architecture, particularly layer dealing with abstraction, embedded systems within network devices, big data visualization in customized way, and applied IoT and IoE research in various domains. Further, the type 2 fuzzy logic can also be considered. Commercialization of the fog architecture and patenting can also be done.

[4] *https://fogproject.org/*

REFERENCES

[1] M. Abdelshkour, "IoT, from Cloud to Fog Computing," 18 September 2017. [Online]. Available: http://blogs.cisco.com/perspectives/iot-from-cloud-to-fog-computing. [Accessed 18 September 2017].

[2] B. Flavio and M. Rodolfo, "Fog Computing and its Role in the Internet of Things," in *Proceedings of the MCC workshop on Mobile Cloud Computing*, 2012.

[3] S. Sarkar and S. Misra, "Theoretical Modelling of Fog Computing: A Green Computing

Paradigm to Support IoT Applications," *IET Networks*, vol. 5, no. 2, p. 23–29, 2016.

[4] P. Lopez, A. Montresor, D. Epema, A. Datta, T. Higashino, A. Iamnitchi, M. Barcellos, P. Felber and E. Riviere, "Edge-Centric Computing: Vision and Challenges," *ACM SIGCOMM Computer Communication Review*, vol. 45, no. 5, pp. 37-42, 2015.

[5] M. Vaquero and R. Luis, "Finding Your Way in the Fog: Towards a Comprehensive Definition of Fog Computing," HP Laboratories Technical Report, 2014.

[6] J. Numhauser and J. Mesa, "XMPP Distributed Topology as a Potential Solution for Fog Computing," in *MESH 2013: The Sixth International Conference on Advances in Mesh Networks*, Spain, 2013.

[7] T. Perry, "What Comes After the Cloud? How About the Fog?," 08 February 2013. [Online]. Available: https://spectrum.ieee.org/tech-talk/computing/networks/what-comes-after-the-cloud-how-about-the-fog. [Accessed 13 October 2017].

[8] W. Shi, J. Cao, Q. Zhang, Y. Li and L. Xu, "Edge Computing: Vision and Challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, p. 637–646, October 2016.

[9] X. Hou, Y. Li, M. Chen, D. Wu, D. Jin and S. Chen, "Vehicular Fog Computing: A Viewpoint of Vehicles as the Infrastructures," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, p. 3860–3873, June 2016.

[10] B. Antonio and F. Stefano, "QoS-Aware Deployment of IoT Applications Through the Fog," *IEEE Internet of Things Journal*, vol. 1, no. 1, p. 99, 2017.

[11] S. Kitanov and T. Janevski, "State of the Art: Fog Computing for 5G Networks," in *Telecommunications Forum (TELFOR), 2016 24th*, 2016.

[12] I. Stojmenovic and S. Wen, "The fog Computing Paradigm: Scenarios and Security Issues," in *Federated Conference on Computer Science and Information Systems (FedCSIS)*, 2014.

[13] M. Janakiram, "Is Fog Computing the Next Big Thing in the Internet of Things," 18 April 2016. [Online]. Available: https://www.janakiram.com/posts/blog/is-fog-computing-the-next-big-thing-in-internet-of-things. [Accessed 5 October 2017].

[14] M. Satyanarayanan, P. Bahl, R. Caceres and N. Davies, "The Case for VM-Based Cloudlets in Mobile Computing," *IEEE Pervasive Computing*, vol. 8, no. 4, 2009.

[15] E. Daly, W. Lancee and J. Polivy, "A Canonical Model for the Taxonomy of Emotional Experience," *Journal of Personality and Social Psychology*, vol. 45, no. 2, pp. 443-457, 1983.

[16] K. Szpunar, R. Spreng and D. Schacter, "A Taxonomy of Prospection: Introducing an Organizational Framework for Future-oriented Cognition," in *Proceedings of the National Academy of Sciences of the United States of America*, 2014.

[17] W. Wang, "Understanding User Experience of News Applications by Taxonomy of Experience," in *Behaviour & Information Technology*, Taylor & Francis, 2017, pp. 1-11.

[18] L. A. Zadeh, "Fuzzy Sets," *Information and Control*, pp. 338-358, 1965.

[19] P. S. Sajja, "Intelligent Web Content Filtering Through Neuro-Fuzzy Approach," *International Journal of Data Mining and Emerging Technologies*, vol. 3, no. 1, pp. 33-39, 2013.

[20] P. S. Sajja and R. A. Akerkar, Intelligent Technologies for Web Applications, Boka Raton, FL, USA: CRC Press (Taylor & Francis Group), 2012.

[21] P. S. Sajja, "Knowledge Representation Using Fuzzy XML Rules in Web Based Expert System for Medical Diagnosis," in *Fuzzy Expert Systems for Disease Diagnosis*, Hershey, PA, USA, IGI Global Book Publishing,, 2014, pp. 138-167.