

Heterogenetic knowledge classification Using Fuzzy inference for unified data clusters

Umer Farooq^{1,2,*}, Khalil Ahmad^{1,3}

¹National College of Business Administration & Economics, Lahore, Pakistan

²Department of Computer Science, Lahore Garrison University, Lahore, Pakistan

³Delta3T, Lahore, Pakistan

Abstract

Emerging technologies such as Cloud Computing, Internet of Things (IoT) and Big Data are developing a digital ecosystem. This ecosystem is catering diverse types and volumes of data that represents information segments. The essence of these segments become vital when transformed into knowledge units to provide a more meaningful and productive perspective. The transformed knowledge at this stage is heterogenetic in nature, consisting of functional and structural properties which needs to be arranged to formulate robust and efficient knowledge repositories. The heterogenetic knowledge can be transformed into classification clusters using structural properties by controlling the degree of heterogeneity. In this paper, Fuzzy Inference System (FIS) based classification approach is proposed for heterogenetic knowledge clustering.

Keywords: GPS, IoT, FIS, Knowledge Heterogeneity, Knowledge System

Received on 16 July 2019, accepted on 19 August 2019, published on 13 September 2019

Copyright © 2019 Umer Farooq *et al.*, licensed to EAI. This is an open access article distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/eai.13-7-2018.160072

*Corresponding author. Email: umerfarooq@lgu.edu.pk

1. Introduction

Emerging technologies doubling the data volume in every two years expected 800 TB by 2020. Data becomes more valuable than the money, fundamental property of an organization irrespective of the size of the organization. Social media, e-commerce, internet of things, sensors etc., are the major sources of the data generation (Gantz, Reinsel, & Arend, 2012). It is important to evaluate what data have been generated and how to use it. Thousands of devices equipped with IoT, “network of physical object or thing, embedded with software/hardware, sensors, and network connectivity which enable collection and exchange of data among the devices”. The IoT enabled transportation system automate the process of collection of data from the diverse location and aggregated very quickly. For example, Google collect real time traffic data from GPS and Google Maps

app enabled mobile phones make possible Google to evaluate live traffic condition and use it for the navigation.

The creation, distribution, manipulation and integration of information influencing human development, central activity as economic, political and cultural in information societies. The digital information explosion changing all aspect of our society.

The progressive development of information technology changed the traditional transformation of information advancing the contents and methods of information science. The enriched digital resources enhanced the boundaries of research to webpages, books, patents and standard literatures. These perspectives, knowledge management and services are goals of information science today. The standardization is an important source of knowledge carrier. Ontology and semantic technologies are useful for the standardization of information to knowledge. Semantic computing used to build concept network automatically in

existing knowledge base to provide standard knowledge service as knowledge mapping and knowledge retrieval.

Knowledge acquisition, exchange, sharing and protection are becoming vital areas of research in technology. Aligned with the same progress, digital progression in terms of cloud computing, internet of things and online business applications require a versatile approach to deal with the various types of knowledge structures adequately. Due to this expanding network of corporate structures and technology variants, it is very much desirable to consider such knowledge schemas which may serve the emerging needs. In corporate world, knowledge management is a comprehensive domain which is addressing this specific challenge, already there are various models and schemas in practice to generate, share and exchange knowledge among corporate entities (M., P., & R, 2002). It is important to note that knowledge in contrast to data is multifaceted and contains multiple semantic and associative dimensions based on objectivity, subjectivity or simply perspective of the knower. Knowledge heterogeneity is the key factors when we consider to make a knowledge system or a knowledge repository.

Another emerging segment of knowledge processing is known as “Big Data” as a resultant to the speedy growth of datasets, not only in corporate sector but also in general communication and online applications. As mentioned above, the emergence of cloud, the nature and conception of data/information processing has been changed. Cloud computing may reduce cost and processing deficit and already providing improved representation and management of knowledge. Various collaborated knowledge management systems are being developed, knowledge as a service (KaaS) for cloud computing is very much possible in upcoming applications and services by cloud service providers, which may lead to the provision of knowledge intensive customer care systems or enterprise knowledge resource planning (Rafiq, Bashar, & Shaikh, 2014). Internet of things (IoT) is another key element of massive amount of data producer in the future. The IoT have heterogeneous interconnected devices, which generates massive amount of data. The collected data will not be useful unless analyze, interpret, and understand. The data collected form the heterogeneous sensors provide heterogeneous sensing data. The heterogeneity of the IoT data processing brings challenges and several advantages and directions for system improvements (Wu, Xu, Feng, Du, & Wang, 2014).

Heterogeneity may generate redundancy in cross disciplinary scenario if not structured properly on the other side the same property may enhance knowledge sharing across various domains if structured in an appropriate fashion, therefore, a suitable degree of heterogeneity is a vital milestone to be achieved. There are various conditions and classifications to deal with this subject, collection of these condition is slow and tardy for making decision (Värzaru & Albu, 2010). This research is taking heterogeneity as the key challenge and valuable factor for the knowledge use beyond the boundaries. It is therefore, need to control the knowledge heterogeneity in order to take advantages in various scientific domains. Such controlling

mechanisms will enhance knowledge sharing, reusability and utilization on multi-disciplinary domains. It is important to identify or extract relationships among various heterogenetic entities to develop a rationale mechanism which may structure knowledge based on semantics and associative properties to attain a desirable degree of heterogenetic knowledge.

Heterogenetic knowledge can be categorized in two main variants i.e. structural and functional, here structural aspect is dealing with the inception, conception and association with other relevant knowledge entities while functional aspect is more specific to the processing of knowledge in respective domain. Primarily, controlling the structural heterogeneity may provide desirable results to generate a suitable knowledge schema for more sharable and exchangeable knowledge.

In this paper we presented FIS based knowledge classification approach for the heterogeneous knowledge. System will process heterogeneous knowledge to classify it by relation among entities.

2. Literature Review

Janowicza proposed a cross disciplinary knowledge platform that may act as a collector of knowledge from various domains into a unified, diverse and shareable scientific knowledge repository. It is required to develop more knowledge intensive and logical knowledge-based systems to deal with the emerging challenges and schema complexities which are referred as digital earth. Such cross disciplinary knowledge platform not only helps in providing interoperability among various scientific domains for creating, sharing and exchange of knowledge but it will also lead towards making homogenized conception and interpretation acceptable to parent domain as well as associative domains (Janowicza & Hitzlerb, 2012).

The creation, distribution and utilization of knowledge through an organization is important active to get sustainable competitive advantage. Explicit knowledge is focused in business intelligence while knowledge management deal with both implicit and explicit knowledge. Both concepts of knowledge stimulate learning, understanding and decision making. The analysis and comparison in (Herschel & Jones, 2013) explained BI is the subset of KM. The integration and of KM and BI will be help full to improves the mental models and understanding for better decision making.

Internet of things (IoT) is another key element of massive amount of data producer in the future (Wang et.al, 2018). The IoT have heterogeneous interconnected devices, which generates massive amount of data. The collected data will not be useful unless analyze, interpret, and understand (Rasool et.al 2019). The data collected form the heterogeneous sensors provide heterogeneous sensing data. The heterogeneity of the IoT data processing brings challenges and several advantages and directions for system improvements (Wu, Xu, Feng, Du, & Wang, 2014). In natural language understanding and human like reasoning

the commonsense knowledge is key in artificial intelligence. Hongyu Lin et al, prosed multi-knowledge reasoning model with heterogeneous knowledge. Uniform representation is generated by inference cost of each rule with the help of different relations in heterogeneous knowledge. Attention based mechanism developed for the selection of specific rule to be applied in specific context. The mining of commonsense knowledge from heterogeneous knowledge sources enabled by proposed multi-knowledge reasoning framework (Multi knowledge Framework). Yoo et al., has accepted the importance knowledge intensive ecosystem which specifically deals with heterogeneity and homogeneity of knowledge but also highlighted the complexity of schemas usable for cross-disciplinary approach due to the subjective and perspective nature of knowledge and its alignment of domain specific functions (Yoo, Boland, Jr., Lyytinen, & Majchrzak, 2012). Another important aspect highlighted by Pahl-Wostl et al., that is the need of readiness in the direct stakeholders like social scientists and other scientific community to understand the benefits and challenges of heterogenetic knowledge as in case of trans-disciplinary context, such ecosystems need coherence among trans-disciplinary knowledge along with highly complex and multifaceted schema for heterogenetic or homogeneity structures for sharing and exchange (Pahl-Wostl, et al., 2013). Entities of heterogeneous knowledge

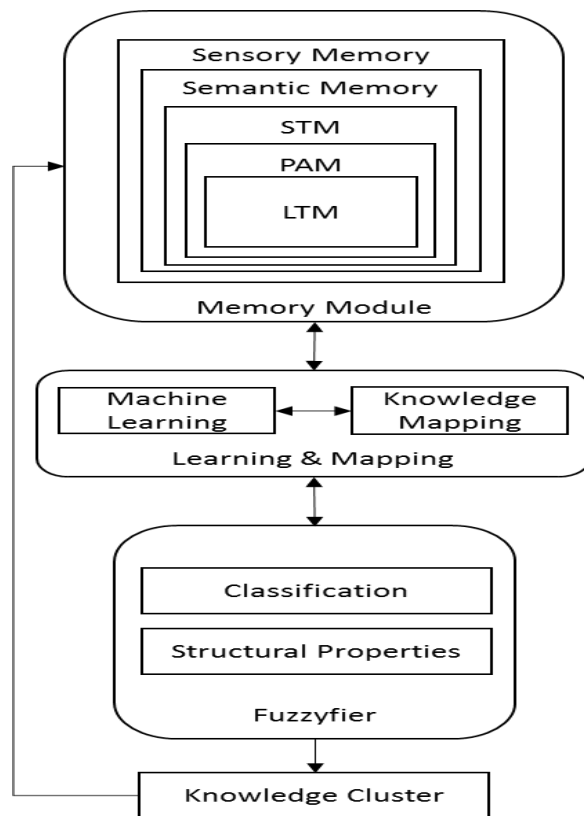


Figure 1. Proposed Model Heterogeneous Knowledge Classification System

bases identified by Haklae Kim using consistency, confidence, threshold filtering, One-to-One Mapping and belief-based approaches. These approaches used for the extraction and verification of identical relation of entities in heterogeneous knowledge bases (Kim, 2018).

Computational Intelligence approaches like Fuzzy system (Areej, et al, 2019) (Hussain et al, 2019), Neural Network (Ayesha, et al, 2019), Swarm Intelligence (Khan, et al, 2019) & Evolutionary Computing (Khan, et al, 2015) like Genetic Algorithm (Khan, et al, 2013) (Ali, et al, 2016) , DE, Island GA (Umair, et al, 2015), Island DE (Khan, et al, 2015) are strong candidate solution in the field of smart city (Wang et.al, 2019) (Kashif, et al, 2018), wireless communication (Asad, et al 2018) etc.

3. Methodology

As mentioned in previous sections that emergence of technologically advanced digital ecosystem has brought challenges for researchers to formulate adequate methods and approaches for data management and manipulation, more precisely, there is a need to look at digital systems beyond conventional data processing.

Especially in the domain of artificial general intelligence, where agency is adopting the concepts of cognitive science and developing agents inspired by human mind, therefore, instead of data, knowledge becomes the more appropriate

and interesting processing entity in agency. Following model has been proposed to deal with the degree of heterogeneity for the classified knowledge clustering.

Sensory memory module is an input module for the proposed model.it is not only receptor of the system also provide input for the semantic memory. Semantic memory understands the influx of semantic memory and existing knowledge residing in short term and long-term memory. The memory module explored for the patterns pattern with influx if not exist pattern already the supervised learning is required to provide the accurate semantic to incoming flux. In knowledge classification important phenomenon is to learning and mapping of heterogenetic knowledge i.e. to develop new classification. In this process the learning and

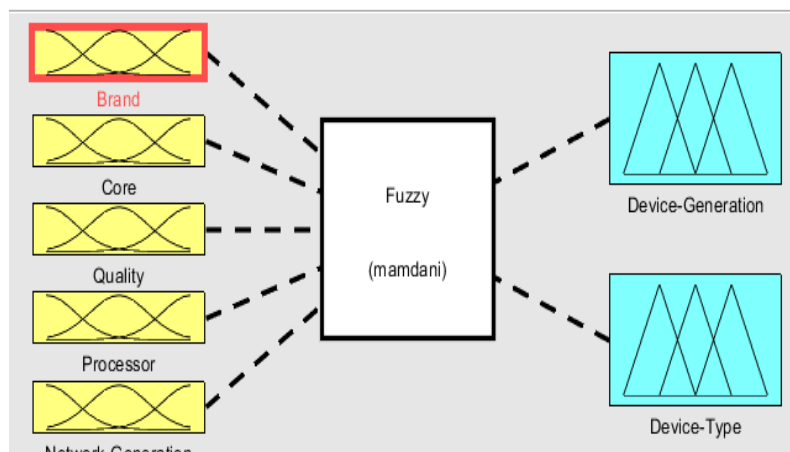


Figure 2. Proposed HKCFIS model

semantic memory are important to understand the classification. Fuzzier is used to generate classification which developed by structural properties of the knowledge. Proposed model is creating a heterogenetic knowledge cluster which is going to provide inter and intra linked structure of knowledge with autonomous classification generation and associative extraction. This knowledge is again stored in the long-term memory as a refined pattern for future reference.

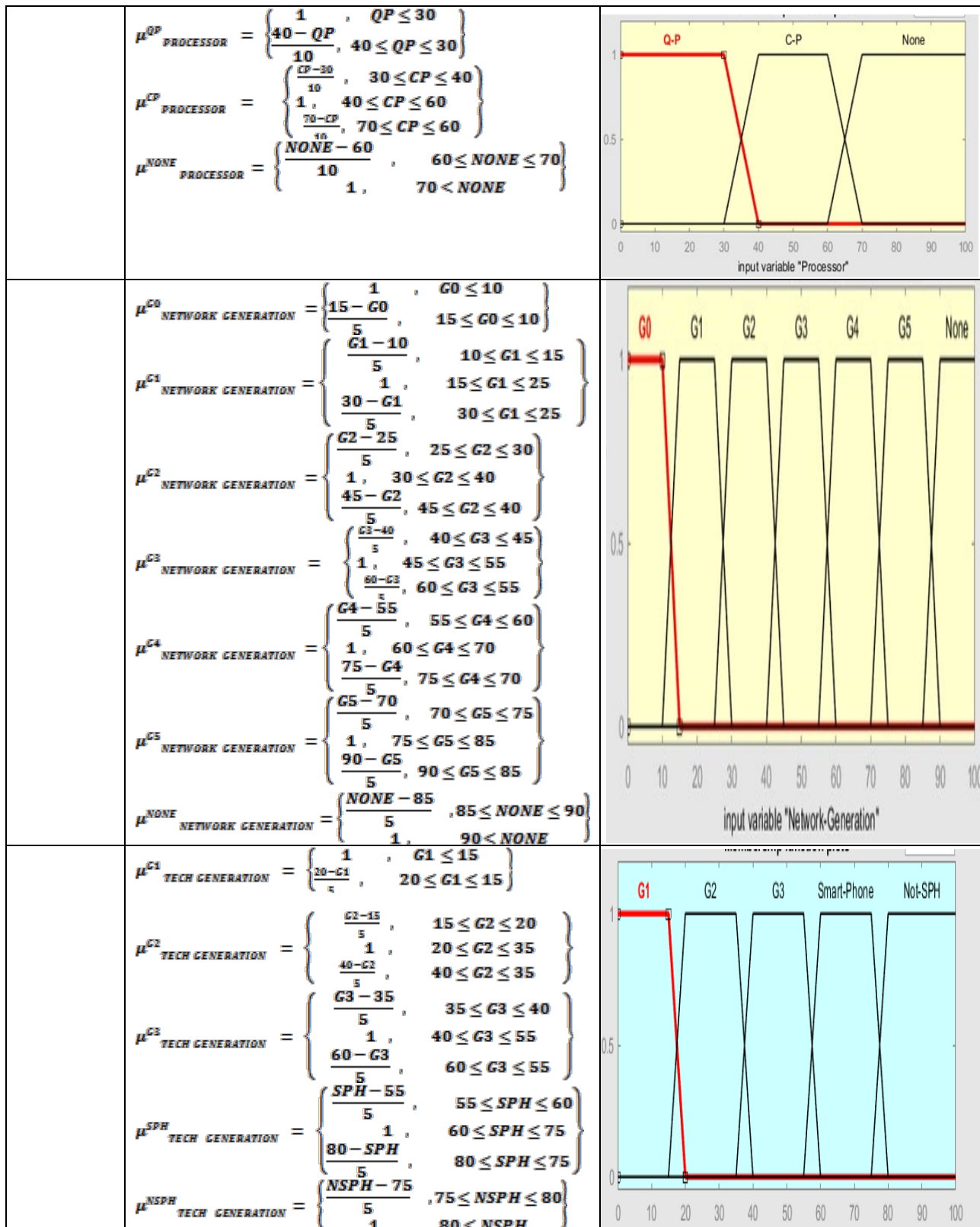
The input variables of HKCFIS are Brand, Core, RAM, Processor, and Network Generation, which are structure properties of digital devices. While the out puts are Tech Generation and Device Types as shown in Fig 2.

Graphical and mathematical representation of HKCFIS System member functions of I/O variables of both layers are shown in table 1.

4. Implementation and Results

Table 1. Input and Output Variables Membership Functions used in Proposed HKCFIS System

Variable	Membership Function (MF)	Graphical Representation of MF
Brand	$\mu^S_{Samsung} = \begin{cases} 1, & S \leq 30 \\ \frac{40-S}{10}, & 40 \leq S \leq 30 \end{cases}$ $\mu^{lg}_{LG} = \begin{cases} \frac{lg-30}{10}, & 30 \leq lg \leq 40 \\ 1, & 40 \leq lg \leq 60 \\ \frac{70-lg}{10}, & 60 \leq lg \leq 70 \\ 1, & 70 \leq lg \end{cases}$ $\mu^{lv}_{Lenovo} = \begin{cases} \frac{lv-60}{10}, & 60 \leq lv \leq 70 \\ 1, & 70 \leq lv \end{cases}$	
Core	$\mu^{c3}_{Core} = \begin{cases} 1, & c3 \leq 10 \\ \frac{20-c3}{10}, & 10 \leq c3 \leq 20 \end{cases}$ $\mu^{c5}_{Core} = \begin{cases} \frac{c5-10}{10}, & 10 \leq c5 \leq 20 \\ 1, & 20 \leq c5 \leq 30 \\ \frac{40-c5}{10}, & 40 \leq c5 \leq 30 \end{cases}$ $\mu^{c7}_{Core} = \begin{cases} \frac{c7-30}{10}, & 30 \leq c7 \leq 40 \\ 1, & 40 \leq c7 \leq 50 \\ \frac{60-c7}{10}, & 60 \leq c7 \leq 50 \end{cases}$ $\mu^{M1}_{Core} = \begin{cases} \frac{M1-50}{10}, & 50 \leq M1 \leq 60 \\ 1, & 60 \leq M1 \leq 70 \\ \frac{80-M1}{10}, & 80 \leq M1 \leq 70 \end{cases}$ $\mu^{M2}_{Core} = \begin{cases} \frac{M2-70}{10}, & 70 \leq M2 \leq 80 \\ 1, & 80 \leq M2 \leq 85 \\ \frac{95-M2}{10}, & 95 \leq M2 \leq 85 \end{cases}$ $\mu^{None}_{Core} = \begin{cases} \frac{NONE-85}{10}, & 85 \leq NONE \leq 95 \\ 1, & 95 \leq NONE \end{cases}$	
	$\mu^L_{Memory} = \begin{cases} 1, & L \leq 30 \\ \frac{40-L}{10}, & 40 \leq L \leq 30 \end{cases}$ $\mu^M_{Memory} = \begin{cases} \frac{M-30}{10}, & 30 \leq M \leq 40 \\ 1, & 40 \leq M \leq 60 \\ \frac{70-M}{10}, & 70 \leq M \leq 60 \end{cases}$ $\mu^H_{Memory} = \begin{cases} \frac{H-60}{10}, & 60 \leq H \leq 70 \\ 1, & 70 < H \end{cases}$	



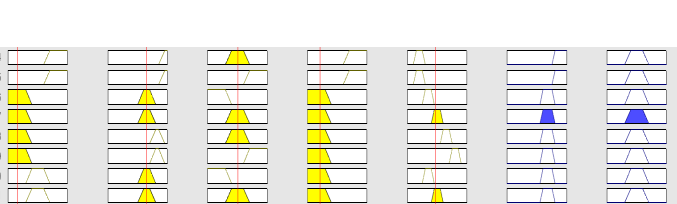
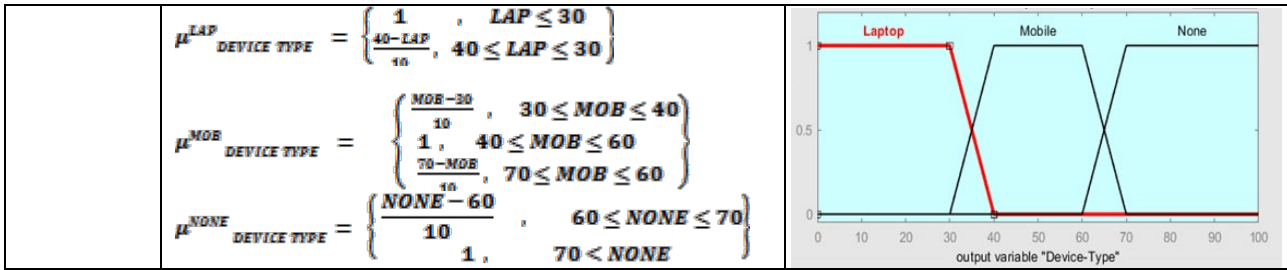


Figure 4. Proposed HKCFIS Model Lookup Diagram for 2nd generation of laptop



Figure 3. Proposed HKCFIS Model Lookup Diagram for Smart Mobile Phone

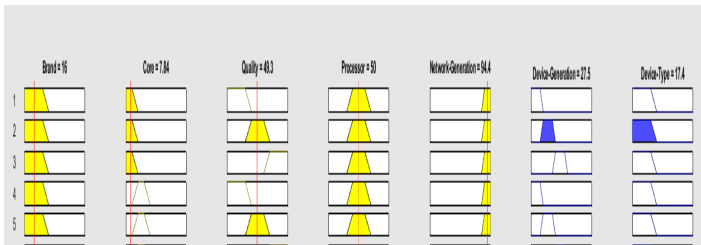


Figure 5. Proposed HKCFIS Model Lookup Diagram for 3rd generation of laptop

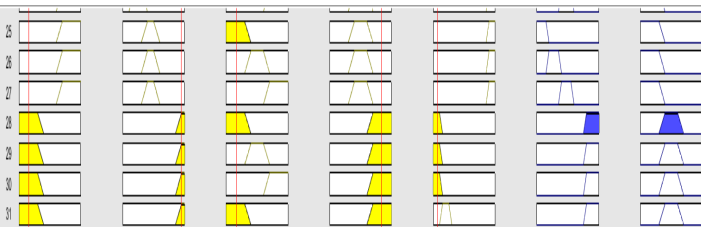


Figure 6. Proposed HKCFIS Model Lookup Diagram for Not Smart Mobile Phone

Table 2. Proposed HKCFIS Device-Generation Surface Views w.r.t various input variables

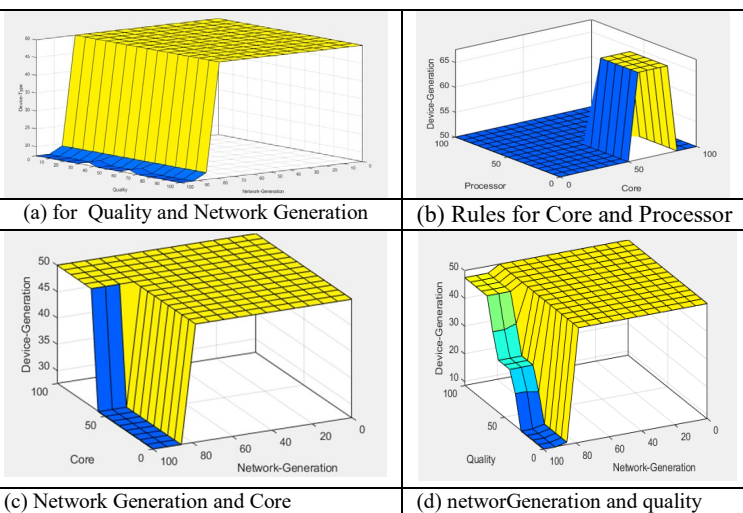


Table 2 (a, b, c & d) shows the Surface views of proposed HKCFIS Device generation with respect to different input variables. (a) Show the surface output of Device Type based on two inputs variables quality and network Generation. The bluish surface shows the 1st Generation and Yellowish color surface represent the Smart phone. (b) Show the surface output Device Generation based on two inputs core and processor. The bluish surface shows the 1st Generation and Yellowish color surface represent the Smart Phone. (c) Show the surface output of Device Generation based on two input variables core and network generation. The bluish color surface shows the 1st Generation and yellowish color surface represent the Smart Phone. Fig. 3(d) show the surface output of Device Generation based on two input variables quality and network generation. The bluish color surface shows the 1st Generation, sea-greenish color surface shows the 2nd Generation, greenish color surface shows the 3rd Generation and yellowish color surface represents the Smart phone.

Figure 3 shows the output device generation and device type based on five inputs brand, core, quality, processor and network generation. If the value of Brand is 16, the value of core is 7.84, the value of quality is 49.3, the value of processor is 50 and the value of network generation is 94.4 then device generation is 27.5 which represents the 2nd generation and device type is 17.4 which represents the laptop.

Figure 4 shows if the value of Brand is 16, the value of Core is 7.84, the value of Quality is 48.5, the value of Processor is 50 and the value of Network Generation is 94.4 then the device generation is 27.5 which represents 3rd generation and the value of device type is 17.4 which represents laptop.

Figure 5 shows if the value of Brand is 26.4, the value of Core is 90.4, the value of Quality is 15.2, the value of Processor is 78.6 and the value of Network Generation is

12 then the device generation is 88.4 which represents Not Smart Phone and the value of device type is 50 which represents mobile. Figure 6 shows if the value of Brand is 24.5, the value of Core is 61.2, the value of Quality is 52.6, the value of Processor is 50 and the value of Network Generation is 94.4 then the device generation is 67.5 which represents Smart Phone and the value of device type is 50 which represents Mobile.

5. Conclusion

The emerging technologies generates diverse types of data that transformed into information. The ultimate goal of information to attain knowledge. Knowledge units to provide more meaning full and productive perspective. The transformed knowledge at this stage is heterogenetic. On the bases of structural and functional properties of knowledge can be transformed into classified cluster to control degree of heterogeneity. The FIS based cluster classification model is proposed to attained classified cluster of knowledge.

References

- [1] Gantz, J., Reinsel, D., & Arend, C. (2012). The Digital Universe in 2020 : Big Data Bigger Digital Shadows, and Biggest Growth in the Far East.
- [2] Herschel, R. T., & Jones, N. E. (2013). Knowledge management and business intelligence: the importance of integration. *JOURNAL OF KNOWLEDGE MANAGEMENT*, 9(4), pp.45-55,.
- [3] Janowicza, K., & Hitzlerb, P. (2012). The Digital Earth as Knowledge Engine. *Semantic Web*, 3(3), 213–221.
- [4] Kim, H. (2018). A Combinational Method to Determining Identical Entities from Heterogeneous Knowledge Graphs. *Journal of Information Science Theory and Practice*, 6(3), pp, 06-15.
- [5] M., B., P, B., & R, C. (2002). Knowledge Nodes: The Building block of a Distributed Approach to Knowledge Management. *Journal of Universal Computer Science*, 8(6), 625-661.
- [6] Pahl-Wostl, C., Giupponi, C., Richards, K., Binder, C., Sherbinin, A. d., Sprinz, D., . . . Bers, C. v. (2013). Transition Towards a New Global Change Science: Requirements for Methodologies, Methods, Data and Knowledge. *Environmental Science & Policy*, 28, 36-47.
- [7] Rafiq, M., Bashar, A., & Shaikh, A. (2014). Innovative Trends in Knowledge Management: A Cloud Computing Perspective. Dubai: Proceedings of the First Middle East Conference on Global Business, Finance and Banking.
- [8] Vărzaru, M., & Albu, N. (2010). Knowledge Management, Cognitive Cartography and Value Creation in the Organization. *Annals of the University of Craiova, Economic Sciences Series*, 3.
- [9] Wu, Q., Xu, Y., Feng, S., Du, Z., & Wang, J. (2014). Cognitive Internet of Things: A New Paradigm Beyond Connection. *IEEE*.
- [10] Yoo, Y., Boland, R. J., Jr., Lyytinen, K., & Majchrzak, A. (2012). Organizing for Innovation in the Digitized World. *Organization Science*, 23(5), 1398-1408.
- [11] Kashif, I., Muhammad, A.K., Sagheer, A., Zahid, H., & Areej, F (2018). Intelligent Transportation System (ITS) for Smart-cities using Mamdani Fuzzy Inference System, *International Journal of Advanced Computer Science and Applications (IJACSA)*. ISSN: 2158-107X, Vol. 9, No. 2, (pp. 94-105), Digital Object Identifier (DOI): 10.14569/IJACSA.2018.090215.
- [12] Muhammad AsadUllah, Muhammad Adnan Khan, Sagheer Abbas, Atifa Athar, Syed Saqib Raza, and Gulzar Ahmad. Blind channel and data estimation using fuzzy logic-empowered opposite learning-based mutant particle swarm optimization. *Computational intelligence and neuroscience*, 2018, 2018.
- [13] Umair, M., Khan, M. A., & Choudry, M. A. S. (2013, January). GA backing to STBC based MC-CDMA systems. In *2013 4th International Conference on Intelligent Systems, Modelling and Simulation* (pp. 503-506). *IEEE*.
- [14] Khan, M. A., Umair, M., & Choudhry, M. A. S. (2015). GA based adaptive receiver for MC-CDMA system. *Turkish Journal of Electrical Engineering & Computer Sciences*, 23(Sup. 1), 2267-2277.
- [15] Khan, M. A., Umair, M., & Choudry, M. A. S. (2015, December). Island differential evolution based adaptive receiver for MC-CDMA system. In *2015 International Conference on Information and Communication Technologies (ICICT)* (pp. 1-6). *IEEE*.
- [16] Ayesha Atta, Sagheer Abbas, M Adnan Khan, Gulzar Ahmed, and Umer Farooq. An adaptive approach: Smart traffic congestion control system. *Journal of King Saud University-Computer and Information Sciences*, 2018.
- [17] Umair, M., Khan, M. A., & Choudry, M. A. S. (2015, December). Island genetic algorithm based MUD for MC-CDMA system. In *2015 International Conference on Information and Communication Technologies (ICICT)* (pp. 1-6). *IEEE*.
- [18] Ali, M. N., Khan, M. A., Adeel, M., & Amir, M. (2016). Genetic Algorithm based adaptive Receiver for MC-CDMA system with variation in Mutation Operator. *International Journal of Computer Science and Information Security*, 14(9), 296.
- [19] Hussain, S., Abbas, S., Sohail, T., Adnan Khan, M., & Athar, A. Estimating virtual trust of cognitive agents using multi layered socio-fuzzy inference system. *Journal of Intelligent & Fuzzy Systems*, (Preprint), 1-16.
- [20] Khan, M. A., Umair, M., Saleem, M. A., Ali, M. N., & Abbas, S. (2019). CDE using improved opposite based swarm optimization for MIMO systems. *Journal of Intelligent & Fuzzy Systems*, 37(1), 687-692.
- [21] Areej Fatima, Muhammad Adnan Khan, Sagheer Abbas, Muhammad Waqas, Leena Anum, and Muhammad Asif, Evaluation of Planet Factors of Smart City through Multi-layer Fuzzy Logic (MFL). *The ISC Int'l Journal of Information Security*. 51-58, 2019.
- [22] Wang, H., Zhang, Z., & Taleb, T. (2018). Special issue on security and privacy of IoT. *World Wide Web*, 21(1), 1-6.
- [23] Rasool, R., Najam, M., Ahmad, H. F., Wang, H., & Anwar, Z. (2019). A novel JSON based regular expression language for pattern matching in the internet of things. *Journal of Ambient Intelligence and Humanized Computing*, 10(4), 1463-1481.
- [24] Wang, Z. J., Zhan, Z. H., Lin, Y., Yu, W. J., Wang, H., Kwong, S., & Zhang, J. (2019). Automatic Niching Differential Evolution with Contour Prediction Approach

for Multimodal Optimization Problems. *IEEE Transactions on Evolutionary Computation*.