IoT Big Data Value Map: How to Generate Value from IoT Data

Nastaran Hajiheydari
York Business School
York St John University
York, United Kingdom
N.hajiheydari@yorksj.ac.uk

Mojtaba Talafidaryani
Faculty of Management
University of Tehran
Tehran, Iran
Mojtabatalafi@ut.ac.ir

SeyedHossein Khabiri
Faculty of Management
University of Tehran
Tehran, Iran
Hosseinkhabiri@ut.ac.ir

ABSTRACT

Huge sources of business value are emerging due to big data generated by the Internet of Things (IoT) technologies paired with Machine Learning (ML) and Data Mining (DM) techniques' ability to harness and extract hidden knowledge from data and consequently learning and improving spontaneously. This paper reviews different examples of analyzing big data generated through IoT in previous studies and in various domains; then claims their business Value Proposition Map deploying Value Proposition Canvas as a novel conceptual tool. As a result, the proposed unprecedented framework of this paper entitled "IoT Big Data Value Map" shows a roadmap from raw data to real-world business value creation, blossomed out of a kind of three-pillar structure: IoT, Data Mining, and Value Proposition Map. The result of this study paves the way for prototyping business models in this field based on value invention from huge data analysis generated by IoT devices in different industries. Furthermore, researchers may complete this map by associating proposed framework with potential customers' profile and their expectations.

CCS Concepts

• Information systems→ Data analytics

Keywords

Internet of Things (IoT); Data Mining; Value Proposition Canvas; Value Proposition Map; IoT Big Data Value Map

1. INTRODUCTION

The Internet of Things (IoT) creates huge streams of data increasingly and it deserves the magnitude of "big data" [15]. Reference [25] calls it the new dominant IT paradigm. Firms like IBM, HP, and Oracle are working on smart things, from smart machines to smart cities [15]. Reference [10] forecasts 26 billion units in IoT by 2020 and this means a significant impact on the data available for the world; thus we will face a huge dizzy amount of data that should be organized and harnessed to yield knowledge and applications.

To fulfill the requirement of using the continuous streams of big data generated by IoT, Data Mining, and Machine Learning techniques are beneficial options, as they comprise techniques for discovering knowledge and then automatically learn and improve with experience and more data. Reference [29] addresses on-line Machine Learning as an answer to IoT data streams. Reference [23] also addresses novel Machine Learning techniques as an answer to harness big data. Similarly, reference [12] conducts a research about the role of big data in smart cities mentioning Machine Learning and Data Mining for smart city big data analytics.

Obtaining knowledge with the help of Data Mining techniques through the big data generated by the IoT creates and proposes value to customers and businesses. From the business point of view, a business model is a conceptual tool, which clears the business logic [20] and one of the most vital elements of business model is the value proposition. In this paper, we have reviewed researches on the applications of Data Mining in IoT and then we have extracted their hidden value proposition and introduced it in an organized framework called Value Proposition Map in the way that [22] suggest.

Therefore, what we have done in this paper is to propose a strong three-pillar structure (IoT, Data Mining and Value Proposition Map), from generating raw data to delivering real-world value to customers, which shapes business models for plenty of near-future businesses in domains like knowledge management, agriculture, food industries, energy management, urban management, health and medical sciences.
The present study is organized as follows: Section 2, 3, and 4 provide an overview on IoT, Data Mining, and Value Proposition Map concepts respectively. Section 5 discusses research method and presents study's findings. Section 6 draws conclusion of the paper.

2. INTERNET OF THINGS (IoT)

IoT is a global network of interconnected sensing and actuating devices (cheap sensors which not only they can sense but also act), any product or physical object (uniquely addressable and visible to the globe) and machines based on standard communication protocols that gather and send the data for analysis with the purpose of smarter, more comfortable, more efficient and more effective world [11, 13, 15, 16].

IoT notable technologies are Radio Frequency Identification (RFID), cheap sensors (wearable, implanted, and environmental), Wireless Sensor and Actuator Networks (WSAN), cloud computing and 3D virtual reality technology [2, 3, 17, 18].

IoT architecture is shown in Figure 1 and it has four layers:

- Devices: This layer consists of devices that gather data from the environment.

ICSLT 2019, January 10–12, 2019, Vienna, Austria

ACM ISBN 978-1-4503-6235-1/19/01…$15.00
https://doi.org/10.1145/3312714.3312728

- Networking and communications: This layer consists of infrastructures for delivering data to higher layers.
- Platform and data storage: This layer provides a facility for data access and storage.
- Data management and processing: This layer provides access to services for end users [17].

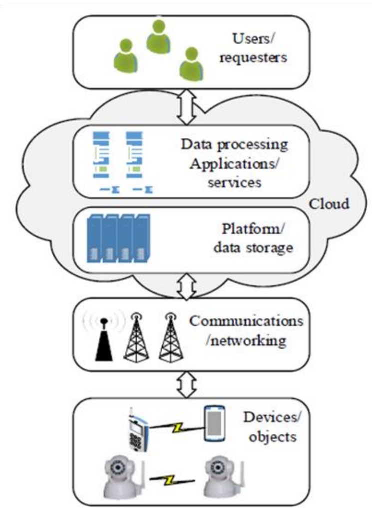

Figure 1. An IoT architecture [17]

As it is shown in Figure 1, one of the most important layers of IoT architecture is related to data processing, which makes it meaningful and provides services or appoints an application for users. Big sources of data might be generated by this layer, which purposeful analytical tools and techniques can make value generation possible.

3. DATA MINING

Data Mining is a broad concept of discovering knowledge from raw data. Using Data Mining techniques and algorithms can end up from raw data to five types of information: associations (happenings triggered by a single event), sequences (outcomes that happen over time triggered by an event), classifications (describes distinct groups of items), clusters (recognizes patterns in which data can be separated), and forecasts (forecasting future values using existing values) [15].

Two types of methods can be used to deal with Machine Learning and Data Mining problems: supervised and unsupervised learning techniques. In supervised learning, a user defines a target value and classifies each instance with explicitly predefined labels and then these techniques put the future instances into proper classes. On the other hand, in unsupervised learning techniques, instead of users, it is the machine that automatically classifies instances based on similarities [5]. K-Nearest Neighbors (kNN), decision trees, Bayesian classifiers, logistic regression, Support Vector Machines (SVM), and Artificial Neural Networks (ANN) are some algorithms among the supervised learning techniques. In unsupervised learning techniques, we can refer to association rule mining, k-means, Bayesian non-parametric, Latent Dirichlet Allocation (LDA) and Locality-Sensitive Hashing (LSH) [1].

Four primary steps of a Data Mining process are data gathering, data preprocessing, model learning, and testing and validation [1]. Based on this logical process, it is mandatory to manipulate data in order to make models, which can learn and deliver much more value rather than raw materials.

4. VALUE PROPOSITION MAP

Reference [21] proposed Business Model Canvas to represent the logic of value generation and making money by businesses. Companies like P&G and Nestle used the canvas because it is a way to move from product-centric thinking to business model approach [19].

Components of the business model canvas are customer segments, value propositions, channels, customer relationships, revenue streams, key resources, key activities, key partnerships and cost structure [22]. To concentrate on strategic and core components of the business model and to be more precise on value generation logic, reference [22] defines in summary “value propositions” and “customer segments” as:

- Value propositions are based on a bundle of products and services that create value for a customer segment.
- Customer segments are the groups of people and/or organizations a company or organization aims to reach and create value for with a dedicated value proposition.

Zoom into “value propositions” and “customer segments” definitions, there is a subtle linkage regarding transferring value from products/services to customers. Reference [22] proposed Value Proposition Canvas to scrutinize and introduce Value Proposition Map.

Reference [22] proposes The Value (Proposition) Map, which breaks value proposition down into products and services, pain relievers, and gain creators and defines them:

- Products and services are simply a list of what a business offers.
- Pain relievers describe how exactly a business’s products and services alleviate specific customer pains. They explicitly outline how a business intends to eliminate or reduce some of the things that annoy its customers before, during, or after they try to complete a job or that prevent them from doing so.
- Gain creators describe how a business products and services create customer gains. They explicitly outline how a business intends to produce outcomes and benefits that its customer expects, desires, or would be surprised by, including functional utility, social gains, positive emotions, and cost savings.

Figure 2 shows a schematic of The Value Proposition Map.

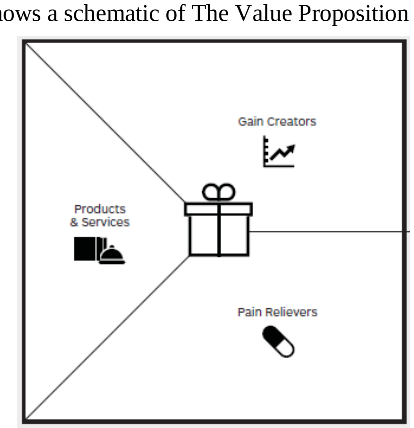

Figure 2. Value Proposition Map [22]

Table 1. IoT Data Mining applications survey from the value map perspective

Year	Authors	Data Mining techniques	IoT devices	Products or Services	Gain creators	Pain relievers
2014	[8] Farruggia et al.	Bayesian Naive classifier	medical imaging devices	text-based indexing system	indexes and classifies medical images and documents	responses to need for content-based medical images retrieval and classification
2016	[7] Edwards-Murphy et al.	decision tree	Wireless Sensor Networks (WSN)	honeybee health monitoring system	protects the population of honey bees and maximizes their productivity	identifies colony health problems
2016	[26] Timoteo et al.	support vector regression	digital scanning receiver	wireless positioning system	estimates the position of a mobile terminal in cellular networks	decreases distance prediction errors
2016	[28] Wang & Yue	association rule mining	RFID tags	food safety pre-warning system	maintains the quality and safety of food products	prevents food safety risks and incidents
2017	[4] Baecke & Bocca	logistic regression, random forests, artificial neural networks	In-Vehicle Data Recorders (IVDRs)	telematics-based insurance product	improves an insurer's risk selection process	reduces a customer's insurance risk
2017	[24] Terroso-Saenz et al.	k-means, random forest, artificial neural networks, support vector regression	sensors and smart meters	IoT Energy Platform (IoTEP)	processes manages and analyses energy data in buildings	reduces energy consumption
2017	[9] Fernández-Ares et al.	Self-Organizing Map (SOM)	smartphones and hands-free devices	mobility monitoring system	optimizes traffic flows, improves security or energetic issues	addresses problems that would be solved in a Smart City
2018	[6] De Vito et al.	neural networks, support vector regression, Gaussian processes regression, multiple linear regression and reservoir computing	multi-sensory devices	chemical multi-sensor devices calibration model	calibrates chemical multi-sensor devices to estimate gas concentrations	solves chemical multi-sensor devices complex calibration problem
2018	[14] Kumar & Gandhi	logistic regression	wearable sensor devices, RFID, 5G mobile networks	health monitoring system	identifies the most significant clinical parameters to get heart disease	prevents impending heart diseases
2018	[27] Uzelac et al.	random forest	smartphones, physical environment sensors, Bluetooth headset, accelerometer	system for recognizing lecture quality	improves lecturer's achievement or performance	decreases students' dissatisfaction with the lecture quality

5. IoT BIG DATA ANALYSIS VALUE MAPPING

To propose a framework representing business value of IoT data analysis, we have reviewed previous researches on the applications of Data Mining in IoT and their contribution have been evaluated systematically from business point of view and based on the Table 1 structure. To this end, we have conducted a systematic search on Science Direct database with the search string: ("Internet of Things" OR "IoT") AND ("Big Data" OR "Data Mining") in Title, Abstract, and Keywords fields and just kept research articles published during 2014-2018 (recent 5 years).

Then we identified and examined products or services that researchers have introduced in their study. In other words, each product or service is investigated from the perspective of gains that customers expect and the pains that can be relieved. Consequently, the hidden value proposition of these emerging products has been convinced. To be sure about the proposed concepts, review and coding have been done by two researchers separately and the results have been checked to choose the best meaning conveying real value of the surveyed application. In order to make the hidden value in these studies perceptible and transfer a scientific perception, a value map is proposed. All the information in Table 1 is demonstrated and summarized in Figure 3 to be able to grasp at a glance considering the logical process of Data Mining. Related labels in Figure 3 are distinguished with different colors.

The framework proposed in Figure 3 can be used as a helpful roadmap for future studies. In other words, this framework states that researchers should first collect data from IoT devices, then these raw data might be processed in order to be prepared for feeding the various Data Mining algorithms, and then, Data Mining discovers hidden patterns and rules in big data. In the next step, results should be evaluated by domain experts to be applied in decision-making. At the end, gained knowledge should be implemented in the form of a product or service to resolve real-world problems. To ensure that the results have significant value to be introduced to a market and it meets customer needs, Value Proposition Map might be deployed to represent the product and service specifications including gain creators and pain relievers.

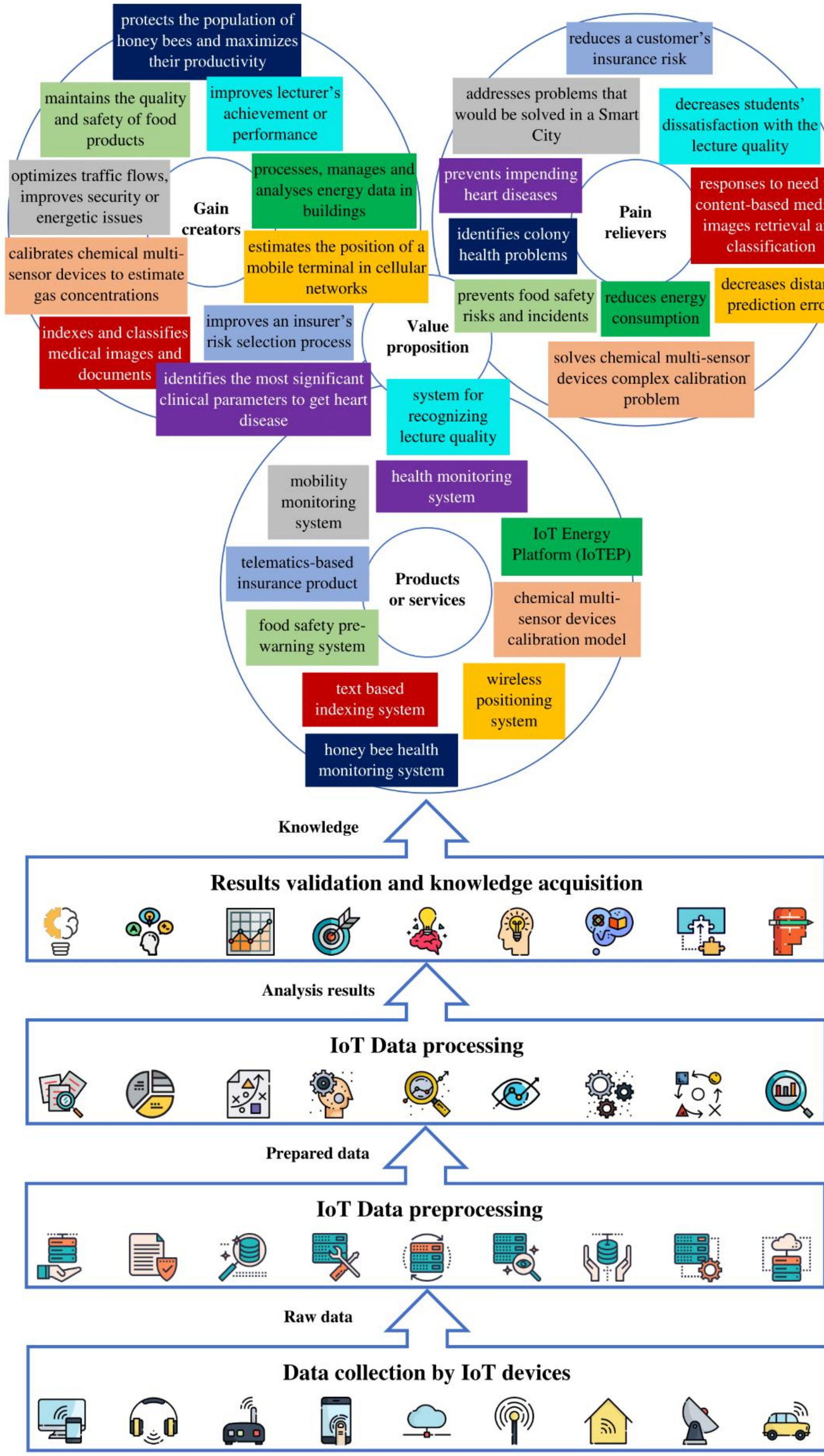

Figure 3. IoT big data analysis value mapping

6. CONCLUSION

IoT technologies and examples have penetrated in different aspects of our today's life and a huge amount of data is generated by these devices that can be taken into account for creating value. The purpose of this paper is to clarify and grab attentions to this emerging source of business value, which is being created and is growing exponentially and uncontrollably in the near-future business world. Considering this objective, the business value of emerging applications of IoT big data analysis in different domains is expressed in a scientific form using Value Proposition Canvas tool. As an example, using huge RFID data analysis in food industry implement a food safety pre-warning system and enables maintaining the quality and safety of food products as a value for customers [28]. Other implications in different industries such as agriculture, energy management, urban management, health, and medical sciences have been analyzed and related value examined in this survey. Moreover, the gain creator and the pain reliever related to each application are investigated applying the value map framework. Businesses may consider this study result as a preface for starting a new line of value invention from huge data generated by IoT devices. Further research can work on customer profile side of value map and present customer jobs, customer gains and customer pains to pave the way for prototyping business models in this field. In addition, by concentrating on one single industry different aspects of generating value from IoT data can be formulated and proposed for practitioners' attention.

7. REFERENCES

[1] Amatriain, X., Jaimes, A., Oliver, N., & Pujol, J. M. (2011). Data Mining methods for recommender systems. In *Recommender Systems Handbook* (pp. 39-71). Springer. Boston, MA, USA. DOI: https://doi.org/10.1007/978-0-387-85820-3_2

[2] Amendola, S., Lodato, R., Manzari, S., Occhiuzzi, C., & Marrocco, G. (2014). RFID technology for IoT-based personal healthcare in smart spaces. *IEEE Internet of Things Journal*, *1*(2), 144-152. DOI: https://doi.org/10.1109/JIOT.2014.2313981

[3] Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A survey. *Computer Networks*, *54*(15), 2787-2805. DOI: https://doi.org/10.1016/j.comnet.2010.05.010

[4] Baecke, P., & Bocca, L. (2017). The value of vehicle telematics data in insurance risk selection processes. *Decision Support Systems*, *98*, 69-79. DOI: https://doi.org/10.1016/j.dss.2017.04.009

[5] de Souza, P. S. S., dos Santos Marques, W., Rossi, F. D., da Cunha Rodrigues, G., & Calheiros, R. N. (2017, January). Performance and accuracy trade-off analysis of techniques for anomaly detection in IoT sensors. In *Proceedings of the International Conference on Information Networking* (pp. 486-491). Da Nang, Vietnam. IEEE. DOI: https://doi.org/10.1109/ICOIN.2017.7899541

[6] De Vito, S., Esposito, E., Salvato, M., Popoola, O., Formisano, F., Jones, R., & Di Francia, G. (2018). Calibrating chemical multisensory devices for real world applications: An in-depth comparison of quantitative Machine Learning approaches. *Sensors and Actuators B: Chemical*, *255*, 1191-1210. DOI: https://doi.org/10.1016/j.snb.2017.07.155

[7] Edwards-Murphy, F., Magno, M., Whelan, P. M., O'Halloran, J., & Popovici, E. M. (2016). b+ WSN: Smart beehive with preliminary decision tree analysis for agriculture and honey bee health monitoring. *Computers and Electronics in Agriculture*, *124*, 211-219. DOI: https://doi.org/10.1016/j.compag.2016.04.008

[8] Farruggia, A., Magro, R., & Vitabile, S. (2014). A text based indexing system for mammographic image retrieval and classification. *Future Generation Computer Systems*, *37*, 243-251. DOI: https://doi.org/10.1016/j.future.2014.02.008

[9] Fernández-Ares, A., Mora, A. M., Arenas, M. G., García-Sánchez, P., Romero, G., Rivas, V., ... & Merelo, J. J. (2017). Studying real traffic and mobility scenarios for a Smart City using a new monitoring and tracking system. *Future Generation Computer Systems*, *76*, 163-179. DOI: https://doi.org/10.1016/j.future.2016.11.021

[10] Gartner. (2013, December). Gartner says the Internet of Things installed base will grow to 26 billion units by 2020. Retrieved from https://gartner.com/newsroom/id/2636073 on 24 September 2018.

[11] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, *29*(7), 1645-1660. DOI: https://doi.org/10.1016/j.future.2013.01.010

[12] Hashem, I. A. T., Chang, V., Anuar, N. B., Adewole, K., Yaqoob, I., Gani, A., ... & Chiroma, H. (2016). The role of big data in smart city. *International Journal of Information Management*, *36*(5), 748-758. DOI: https://doi.org/10.1016/j.ijinfomgt.2016.05.002

[13] INFSO D.4 Networked Enterprise & RFID INFSO G.2 Micro & Nanosystems in co-operation with the Working Group RFID of the ETP EPOSS. (2008, May). Internet of Things in 2020: Roadmap for the Future. Retrieved from https://docbox.etsi.org/erm/Open/CERP%2020080609-10/Internet-of-Things_in_2020_EC-EPoSS_Workshop_Report_2008_v1-1.pdf on 24 September 2018.

[14] Kumar, P. M., & Gandhi, U. D. (2018). A novel three-tier Internet of Things architecture with Machine Learning algorithm for early detection of heart diseases. *Computers & Electrical Engineering*, *65*, 222-235. DOI: https://doi.org/10.1016/j.compeleceng.2017.09.001

[15] Laudon, K. C., & Laudon, J. P. (2016). Management information systems: Managing the digital firm. *Pearson*.

[16] Lee, I., & Lee, K. (2015). The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Business Horizons*, *58*(4), 431-440. DOI: https://doi.org/10.1016/j.bushor.2015.03.008

[17] Luong, N. C., Hoang, D. T., Wang, P., Niyato, D., Kim, D. I., & Han, Z. (2016). Data collection and wireless communication in Internet of Things (IoT) using economic analysis and pricing models: A survey. *IEEE Communications Surveys & Tutorials*, *18*(4), 2546-2590. DOI: https://doi.org/10.1109/COMST.2016.2582841

[18] Miorandi, D., Sicari, S., De Pellegrini, F., & Chlamtac, I. (2012). Internet of Things: Vision, applications and research challenges. *Ad Hoc Networks*, *10*(7), 1497-1516. DOI: https://doi.org/10.1016/j.adhoc.2012.02.016

[19] Muhtaroglu, F. C. P., Demir, S., Obali, M., & Girgin, C. (2013, October). Business model canvas perspective on big data applications. In *Proceedings of the IEEE International Conference on Big Data* (pp. 32-37). Silicon Valley, CA, USA. IEEE. DOI: https://doi.org/10.1109/BigData.2013.6691684

[20] Osterwalder, A., & Pigneur, Y. (2003, September). Modeling value propositions in e-business. In *Proceedings of the International Conference on Electronic Commerce* (pp. 429-436). Pittsburgh, Pennsylvania, USA. ACM. DOI: https://doi.org/10.1145/948005.948061

[21] Osterwalder, A., & Pigneur, Y. (2010). Business model generation: A handbook for visionaries, game changers, and challengers. *John Wiley & Sons*. Hoboken, NJ, USA.

[22] Osterwalder, A., Pigneur, Y., Bernarda, G., & Smith, A. (2014). Value proposition design: How to create products and services customers want. *John Wiley & Sons*. Hoboken, NJ, USA.

[23] Qiu, J., Wu, Q., Ding, G., Xu, Y., & Feng, S. (2016). A survey of Machine Learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, *2016*(1), 67. DOI: https://doi.org/10.1186/s13634-016-0355-x

[24] Terroso-Saenz, F., González-Vidal, A., Ramallo-González, A. P., & Skarmeta, A. F. (2017). An open IoT platform for the management and analysis of energy data. *Future Generation Computer Systems*. In press. DOI: https://doi.org/10.1016/j.future.2017.08.046

[25] The Economist Intelligence Unit. (2013, October). The Internet of Things business index: A quiet revolution gathers pace. Retrieved from https://eiuperspectives.economist.com/technology-innovation/internet-things-business-index/white-paper/internet-things-business-index on 24 September 2018.

[26] Timoteo, R. D., Silva, L. N., Cunha, D. C., & Cavalcanti, G. D. (2016). An approach using support vector regression for mobile location in cellular networks. *Computer Networks*, *95*, 51-61. DOI: https://doi.org/10.1016/j.comnet.2015.12.003

[27] Uzelac, A., Gligorić, N., & Krčo, S. (2018). System for recognizing lecture quality based on analysis of physical parameters. *Telematics and Informatics*, *35*(3), 579-594. DOI: https://doi.org/10.1016/j.tele.2017.06.014

[28] Wang, J., & Yue, H. (2017). Food safety pre-warning system based on Data Mining for a sustainable food supply chain. *Food Control*, *73*, 223-229. DOI: https://doi.org/10.1016/j.foodcont.2016.09.048

[29] Yasumoto, K., Yamaguchi, H., & Shigeno, H. (2016). Survey of real-time processing technologies of IoT data streams. *Journal of Information Processing*, *24*(2), 195-202. DOI: https://doi.org/10.2197/ipsjjip.24.195