

Original Research Paper

A Proposed Framework for Reducing Electricity Consumption in Smart Homes using Big Data Analytics

¹Mira Tamer Shaker, ²Ayman E. Khedr and ³Sherif Kholeif

¹Department of Management Information Systems, Faculty of Commerce and Business Administration, Future University in Egypt, Cairo, Egypt

²Department of Information Systems, Faculty of Computers and Information Technology,

Future University in Egypt, Cairo, Egypt

³Department of Information Systems, Faculty of Computers and Information, Helwan University, Cairo, Egypt

Article history

Received: 12-02-2019

Revised: 15-04-2019

Accepted: 26-04-2019

Corresponding Author:

Mira Tamer Shaker

Department of Management

Information Systems, Faculty

of Commerce and Business

Administration, Future

University in Egypt, Cairo,

Egypt

Email: mira.shaker@fue.edu.eg

Abstract: Smart homes with smart technologies can provide better insights into saving energy and improving the quality of our life. All connected appliances are Internet of Things (IoT) devices that support applications inside a smart home which produce an amount of data that shows what households are doing during their daily life. IoT and Big Data Analytics (BDA) became the most popular technologies in our smart life that are rapidly affecting all areas of technologies and businesses to increase the benefits for organizations and individuals. This research paper contains a review study for recent papers with different techniques that discusses BDA challenges and benefits of a smart home and its relationship with IoT. Moreover, the research paper also contains a proposed approach with its technique (clustering algorithm) for analysing data to solve the main research problem of the large power crisis in Egypt due to the high electricity consumption. Thus, using associate rule to recommend actions based on these data to reduce electricity consumption in different houses in Egypt based on each inhabitant's interest.

Keywords: Smart Home, Big Data, Big Data Analytics (BDA), Internet of Things (IoT)

Introduction

King (2003) defines a smart home as a “dwelling incorporating a communications network that connects the key electrical appliances and services and allows them to be remotely controlled, monitored or accessed”. While Harper (2003) acknowledged that “smart house technologies that most people are pleased with are connected with saving energy or money”. While mining data is one of the successful fields which has been currently one of the main vital solutions for many problems in the business field (Khedr *et al.*, 2015; 2018), the current research focuses on smart homes for which we need to extract information from raw data collected through wireless devices embedded in IoT applications in order to contribute and improve energy efficiency management, such as smart meter, which is one of the most cost-effective techniques that depends on an intelligent calculating device capable of reporting information about the amount of power consumed. It is connected to some sensor networks and collects information as measuring devices. Such objective is

beneficial for many organizations (Khedr and Kok, 2006; Hegazy *et al.*, 2015).

By monitoring the user's consumption habits, adjusting the consumption of final users is possible to smooth the demand curve so the consumption peaks can be reduced. This could be done by generating consumption predictions using different machine learning approaches. This objective has been applied in different fields such as in banking fields (Khedr *et al.*, 2016; Khedr *et al.*, 2014) and education (Khedr, Idrees, 2017A; 2017B), as well as smart homes.

In this paper, we present some recent techniques used in smart homes to collect data about consumers' consumption and their behavior. The paper is organized as follows: Section 2 is a background. Section 3 shows the literature review in the recent papers. Section 4 gives an overview of the approaches being used in recent papers. Section 5 presents research problem in Egypt. Section 6 presents the proposed approach and the data analysis. Section 7 presents the conclusion and the future work.

Background

Smart home system helps to reduce the energy consumed in the building as people consume a very large part of the electricity (Martirano, 2011). In fact, sensors are one of the IoT technology indicators of having smart homes and appliances including smart TVs, lighting control, home security system, temperature monitoring and fire detection. Householders can continuously monitor and control all home appliances, even outdoor and make any decision under any circumstances with the help of sensors of the appliances that send surveillance data to a central controller (Talari *et al.*, 2017) as shown in Fig. 1.

Devices and sensors enable information to be shared across in an appropriate manner and communicate with a smart environment, this is offered by the Internet of Things which is the next innovative technology that has

been adopted recently by different wireless technologies to benefit from the opportunities offered by the Internet technology (Marjani *et al.*, 2017). Home appliances such as air conditioners, TVs, water heaters and all the surrounding electronic equipment are connected to the Internet of Things that can be controlled remotely to facilitate daily life operations (Marjani *et al.*, 2017).

Sensors, applications and devices continuously generate huge amounts of structured, unstructured or semi-structured data that is strongly increasing known as the Big Data. Big Data is often described using five Vs: Volume, Velocity, Variety, Veracity and Value (Hadi *et al.*, 2015). Storing, processing and analyzing the growing amount of data or big data is inadequate using traditional database systems. Big data is somewhat new in IT and business but has been used in previous literature (Marjani *et al.*, 2017).

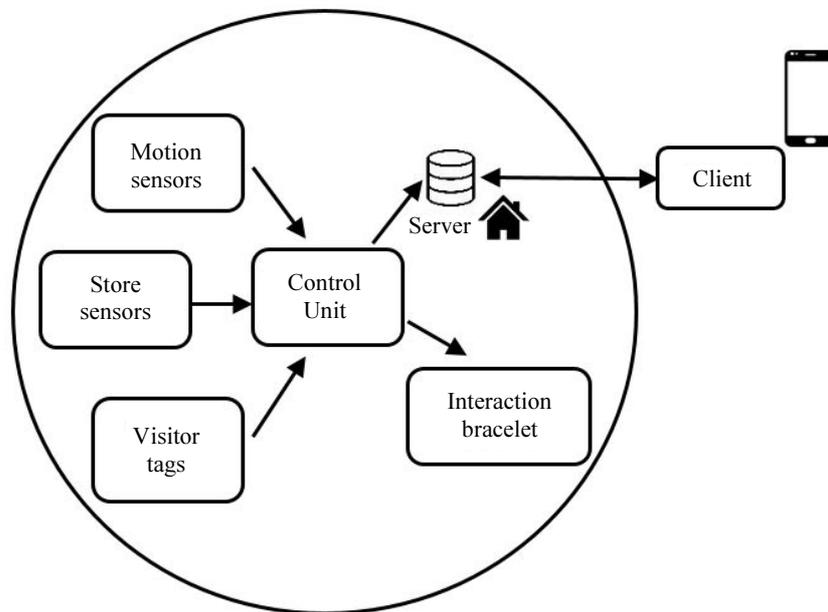


Fig. 1: The view of a smart home system (Chen *et al.*, 2016)

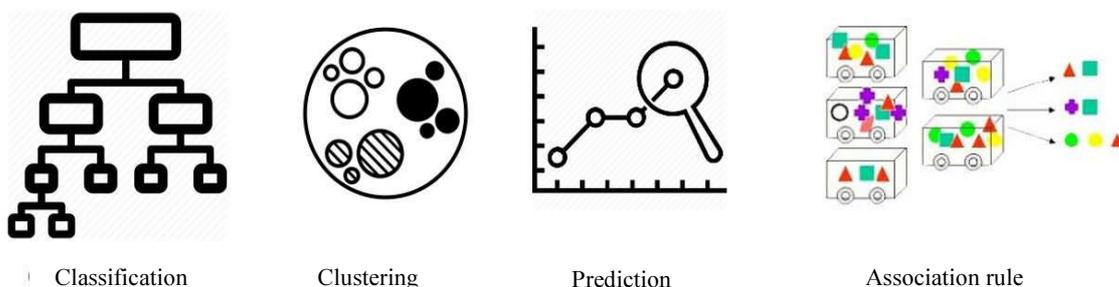


Fig. 2: Overview of big data analytics methods (Marjani *et al.*, 2017)

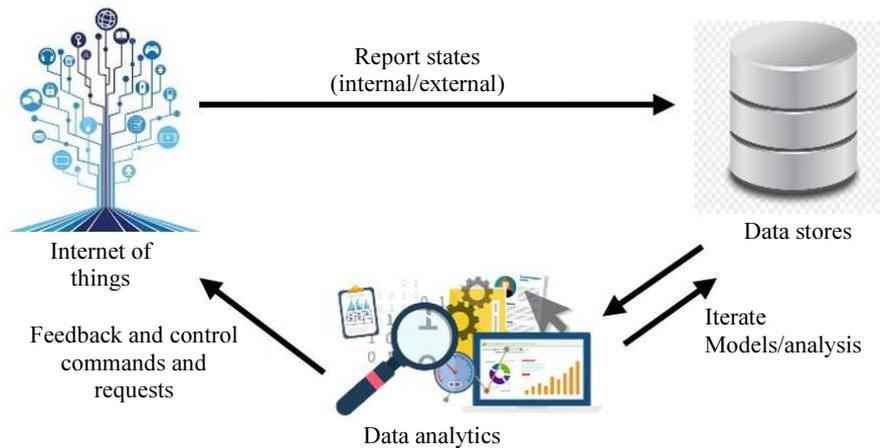


Fig. 3: Relationship between Internet of Things and Big Data Analytics (Marjani *et al.*, 2017)

BDA is the ideal first step towards a smarter city. It assures flexible and real-time data processing followed by intelligent decision procedures (Crammer *et al.*, 2013). A large volume of data can be collected from various sources as shown in Fig. 2.

BDA and IoT systems complement each other when combined, one could predict problems and fix them early and the other could react to problems. Supportive techniques are required for such a decision which many research has developed (El Seddawy *et al.*, 2013). Data analytics could draw the information that is delivered by internet of everything to build insights that are required (Shukla, 2017), (Bhati, 2017) as shown in Fig. 3.

Different smart devices like mobile phones, smart air-conditioners, etc. put smart homes into practice and apply the technology of Artificial Intelligence (AI). There are three generations of smart home technologies: First generation: Proxy server home automation approach and wireless technology (Bluetooth and Zigbee). Second generation: AI controls electrical devices. Third generation: Robot bubby “who” can interact with human beings and can stroll around Home. (Li, 2013; Moltagh *et al.*, 2015).

Research Problem

Recent researchers found that the power crisis is a prevailing problem world-wide. Thus, this is our main research problem specifically in Egypt that is currently observing a large power crisis due to the high electricity consumption that is increasing faster than the capacity expansion as shown in Fig. 4.

A research study was conducted at smart solution company in Maadi, Cairo, Egypt, where all engineers and employees have agreed that residents have low awareness in saving energy and electricity bills, where all residents consume high percentages of electricity and pay high bills. Moreover, they argued that households

demand smart devices or appliances not for the sake of saving energy but for some other reasons, such as home decoration, the need to control and monitor all devices easily and more advanced systems. In addition, households never ask for a full smart home, they ask only for some smart devices.

Egypt has many successful business fields (Nazier *et al.*, 2013; Khedr *et al.*, 2015), one of these fields is residential compounds such as Mountain View that doesn't sell fully smart homes, they only build the compound with an infrastructure that supports any smart automation systems, but it is useless as the awareness of having a smart home to save electricity and energy doesn't exist.

Air-Conditioners and Heaters

One of the most consuming electricity devices are air-conditioners and heaters. Companies has a program application that is installed in portable devices that enables clients to remotely control and monitor air-conditioners and heaters; it's a simple application that works by copying some numbers and data from the device's remote.

Lights

Companies installed motion sensors to automatically switch on the lights whenever the person enters a specific room or kitchen and switch it off whenever he/she comes out. The same problem exists in corridors. Companies need to install motion sensors instead of using multi-way switching. Other rooms such as bedrooms, living rooms and dining rooms don't need motion sensors, companies just install dim switches to control lights in each room which also helps in saving energy. Figure 5 shows the light switching system of an automated place captured from a smart solution company site in Egypt.

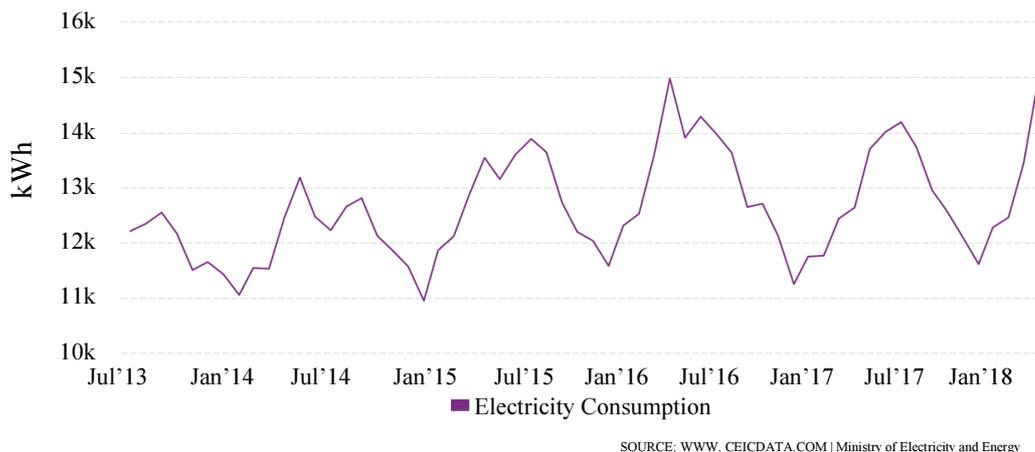


Fig. 4: Electricity consumption in Egypt KWh (2013-2018) (Energy, 2017)

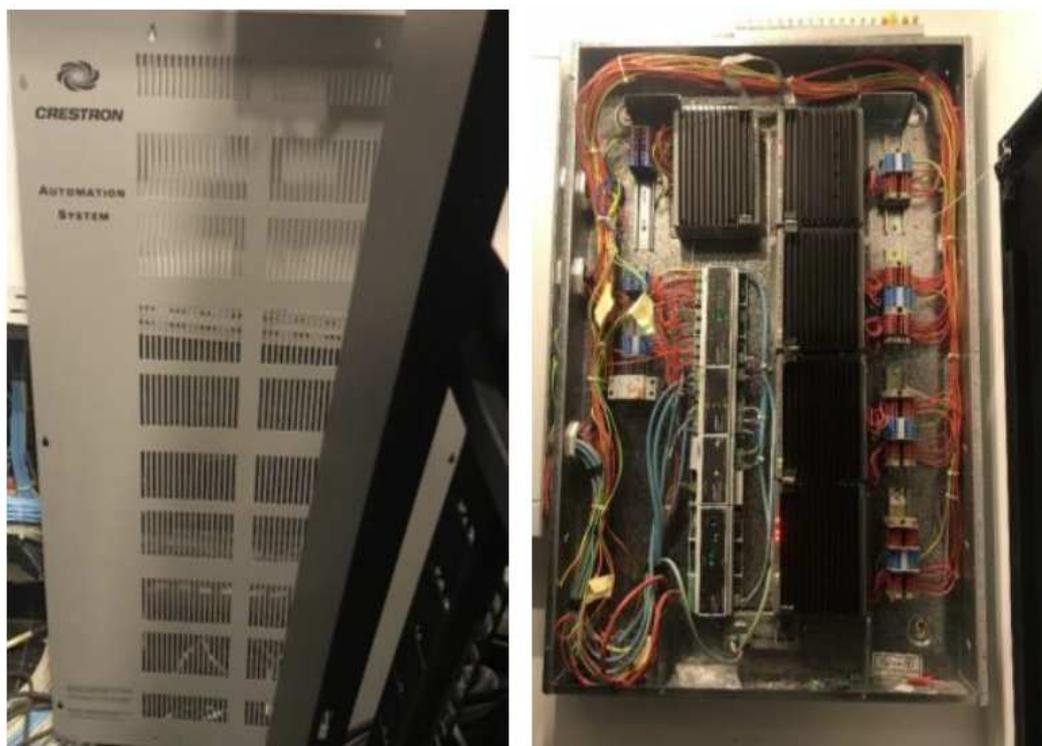


Fig. 5: The light switches in the right and its cover on the left

Smart Meter

All smart solution companies agreed that only few factories use smart meters that are never installed in homes regardless their huge benefits. Smart meters in homes would provide a report for every appliance in every room with the amount of electricity consumption across time and highlight which time is the highest consuming electricity. This will help in saving energy and reducing the electricity bills as people can control their devices for the rest of the month from the first week report.

Literature Review

This section explores literature, concepts, datasets, findings, proposed work and results from the field of smart home automation systems that were discussed and their existing approaches (mentioned in Table 1). The goal is to extract proven ideas for such framework to analyze it. In addition, techniques of BDA and IoT mentioned in Table 1 are discussed in details in Table 2.

Table 1: Comparative study of recent techniques

Papers	Proposed work	Dataset	Findings	Shortcomings
1- (Mayr, 2016).	Activity Recognition algorithm	WSU CASAS*	Proved that windowing technique is a solid basis for implementing a powerful tool to detect complex activities in Smart Homes.	They worked on 2 phases (offline and online) that increases the accuracy of the data.
2- (Cook <i>et al.</i> , 2017).			1. Ensuring clean data. 2. Annotating data. 3. Generating sufficiently varied data.	Checking home behavior datasets to find the challenges faced when generating datasets.
3- (Conrad, 2016).			Testing the algorithm on several days by using several sensors.	Create a healthy smart home to any enhance occupants' lives.
4- (Moltagh <i>et al.</i> , 2015).	Clustering Method	Smart Grid Smart City (SGSC)*	Indicates that many participants changed their time of use behaviors after subscribing to these products.	This methodology is to distinguish what exact changes occur in load profiles due to demand - response behaviors.
5- (Sami and Honarvar, 2016).	PrefixSpan: Projected Sequential Pattern Growth		Understanding the usage patterns of appliances for energy saving, sustainability, suspend the scheduled tasks to off-peak hours and automatic control the Home apparatus, etc.	Working on 2 clusters, with different cases to have several datasets (all day and every day).
6- (Talari <i>et al.</i> , 2017).	Integrating Social Network and Big Data Analytics.	Data derived from the literature reviews of smart homes and big data.	It is categorized into 5 perspectives: business, behavioral, interaction, technological and analytical.	Smart home technology makes strong use of big data techniques that help to master amount of data generated continuously.
7- (Chen <i>et al.</i> , 2016).	Weighted Hybrid Recommendation Algorithm.	Kuting Smart Home project.	Results show that the hybrid recommender system can optimize the distribution of weights of each component and get more reasonable Recall and precision rates.	More sensors deployed; more context information can be obtained. Data is collected every 20 sec over a period of 30 days.
8- (Wich and Karmer, 2016).	Load Disaggregation Algorithm (LDA). 25 Households	EU project called Smart HG. Installed in with data algorithms combined in the for city of Svebolle, Kalundborg, Denmark and 8 months of data	The framework allows a fast exploration and integration of a variety of machine learning know how to separate the recovery mechanisms improving the recognition's accuracy. It also generates natural language reports of the user's behavior from the recognized appliances.	Deducing user behavior from smart meter data to disaggregation analyzes from the natural language data processing. It also uses uniform interfaces to exchange data seamlessly between its items. It relies on a shared database which enables easy upgrades of the framework components.
9- (Ebeid <i>et al.</i> , 2016).	Machine Learning (ML) Approaches	Collected. Smart Home Energy (SHE) project, (SHE Consortium 2012).	Big Data allows large volumes of varied data to be managed and offers support for ML algorithm, Data mining visual tools, near to real-time monitoring and other information analysis and processing possibilities that fit perfectly with the requirements.	They proved that big data technology allows information to be analyzed in more detail than with traditional technology and the application of it to the energy sector is an innovative idea.

*Centre for Advanced Studies in Adaptive Systems (CASAS)

CASAS was established in the School of Electrical Engineering and Computer Science at Washington State University (WSU Pullman campus). CASAS tests technologies using smart homes to meet research needs. In WSU CASAS there are many tools can be used (Cook *et al.*, 2017).

*Smart Grid, Smart City (SGSC)

SGSC project is an Australia's first commercial-scale Smart grid. The project is funded by Ausgrid in partnership with the Australian Government, Energy Australia and their consortium partners (Moltagh *et al.*, 2015). The SGSC project tests smart grid in a real-world context by gathering robust information about costs and benefits. All electricity providers, governments, technology suppliers and consumers across Australia could use the outcomes of the project for future decisions (Sami and Honarvar, 2016), (Ayaz *et al.*, 2017).

Table 2: Techniques of big data analytics and internet of things

Technique	Description	Advantage	Disadvantage
1. Activity Recognition Algorithm	Activity recognition uses sensory observations that aims to identify and recognize the agents, actions and goals from a series of observations and environmental conditions (Mayr, 2016).	Activity recognition can detect concurrent activities and interleaved activities easily. It can observe and analyze human activities and interpret ongoing events successfully (Mayr, 2016).	Activity recognition requires a detailed analysis and understanding of the domain in which activities occur (Mayr, 2016).
2. Smart Windowing	A technique that can be used for online sensor streaming that consists of two phases: (1) an online phase that uses the same windowing technique to recognize activities and (2) an offline phase for analyzing and windowing the streaming data (Conrad, 2016).	It helps to avoid situations in which the sensor data of an activity is spread over several different chunks or the window size, i.e. the set of simple events, is not sufficient to predict a specific activity. Detect complex activities in smart houses (Conrad, 2016).	Expensive to manufacture and expensive to install.
3. Spatiotemporal feature Analysis	A set of different statistical spatiotemporal features is to recognize activities in real time (Mayr, 2016).	Allow to study the persistence of patterns over time and illuminate any unusual patterns.	Assessing both the temporal and spatial dimensions of data adds complexity to data analysis.
4. Clustering Method	Clustering where objects in the same group that are more like each other are grouped together (Moltagh <i>et al.</i> , 2015).	Automatic recovery from failure and easy to implement.	Complexity and failure to recover from database corruption.
5. Prefix Span	Prefix Projected Sequential Pattern Growth: It discovers the frequent single items, then wrappings this information into a frequent-pattern tree, or FP-tree. Pattern growth is a method of Pattern growth is a method of frequent-pattern mining that does not require candidate generation (Han and Lim, 2010).	PrefixSpan mines the complete set of patterns but greatly reduces the efforts of candidate subsequence generation. Moreover, prefix- projection substantially reduces the size of projected databases and leads to efficient processing.	The database will keep shrinking.

Table 2: Continue

6. Machine Learning Algorithm (NILM: Non- Invasive Load Monitoring)	Machine learning is the science of receiving computers to act without being openly programmed. Non- Invasive Load Monitoring(NILM) can distinguish devices from the combined data established from a single point of measurement as it is a smart method for energy disaggregation (Zoha <i>et al.</i> , 2012).	Cheap and easy to install since it uses one metering device for each energy entrance instead of metering device per room.	Requires a long period of observation to analyze the power characteristics of each device.
7. Collaborative Filtering	Collecting preferences or taste information from many users\using collaborative filtering method which can make automatic predictions. Therefore, by using what people liked or bought in the past collaborative filtering can predict the items that people will buy or prefer (Chen <i>et al.</i> , 2016). <ul style="list-style-type: none"> ●Item-based collaborative filtering. ● User-based collaborative filtering. 	<ul style="list-style-type: none"> ● Captures the change in user interests over time and not required to understand item content. ● It does not depend on analyzing the content provided by the user. 	<ul style="list-style-type: none"> ● Cold start problem. ● Cannot handle new users and new items.
8. Content-based Filtering	Could recommend items based on attributes that users bought or liked before (Chen <i>et al.</i> , 2016).	<ul style="list-style-type: none"> ● Comparison between items possible. ● Adaptive. ● More efficient, scalable. ● Easily updatable for recommendations. ● Constructed from simple components. 	<ul style="list-style-type: none"> ● May not recommend to new users. ● Overspecialization. ● Over-fitting. ● New user problem. ● Sparsity.
9. Weighted Hybrid Recommendation Algorithm	Combining content-based component prediction with collaboration filtering component prediction. The weight of each component can be allocated and expected using a Kalman Filter, that can estimate the system state with the measurements and this is the key of weighted hybridization (Chen <i>et al.</i> , 2016).		

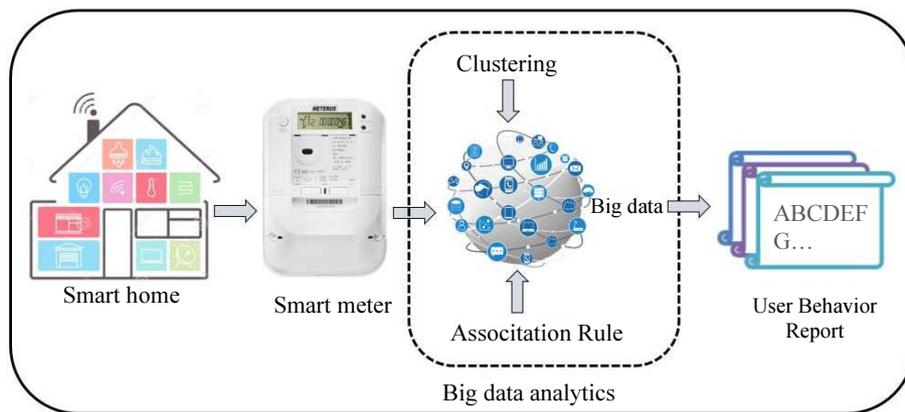


Fig. 6: Proposed framework of using big data analytics in smart home the case of Egypt

Big Data Analytics and Internet of Things Techniques

Proposed Approach and Data Analysis

In order to address the issue of electricity consumption due to high usage and consumption of the appliances mentioned in the research problem; we used clustering algorithm to analyze the data to find out their consumption pattern and then used association rule to help in recommending actions based on inhabitant’s interest as shown in Fig. 6. None of the recent papers mentioned in the literature review used both algorithms together.

Clustering algorithm is to group households based on their satisfaction level which is described by the

household’s characteristics and activities behavior by applying k-means on raw data. In addition, association rule (apriori algorithm) helps to find relationships between data items within large datasets in various types of databases.

Research Methodology

Dataset were collected from visiting five different Egyptian houses for several months and measuring the electricity consumption of their devices to know the highest electricity consuming device; we collected millions of data records every 6 sec. Every record was in UNIX time that is transferred into date and time using RStudio program and every 30 min of data are merged to find out the usage of each device.

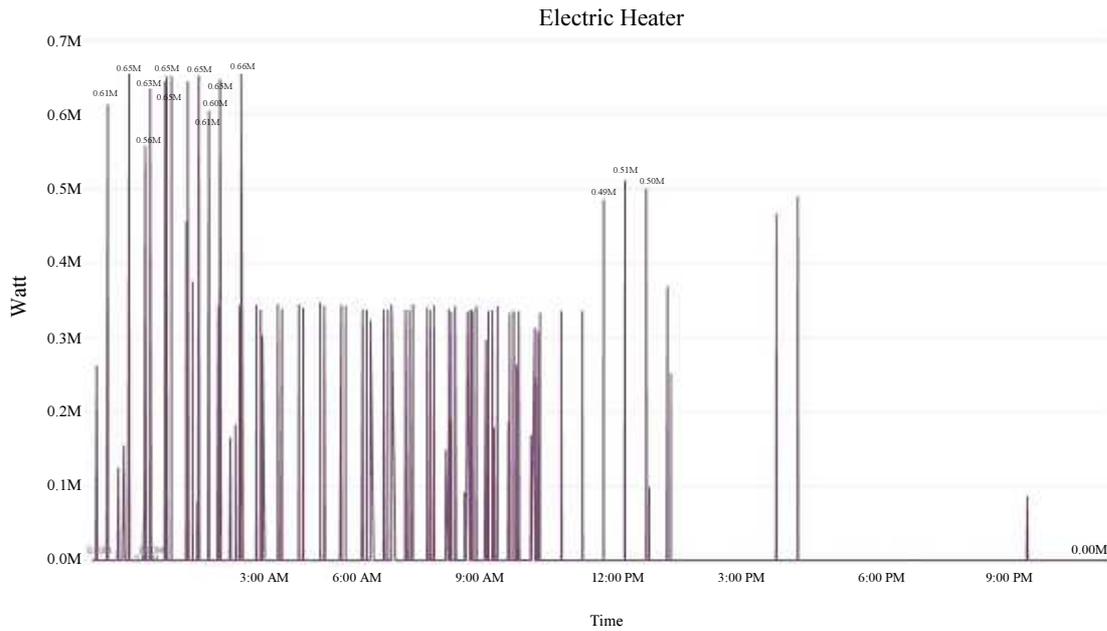


Fig. 7: Appliance power demand of electric heater in house 3

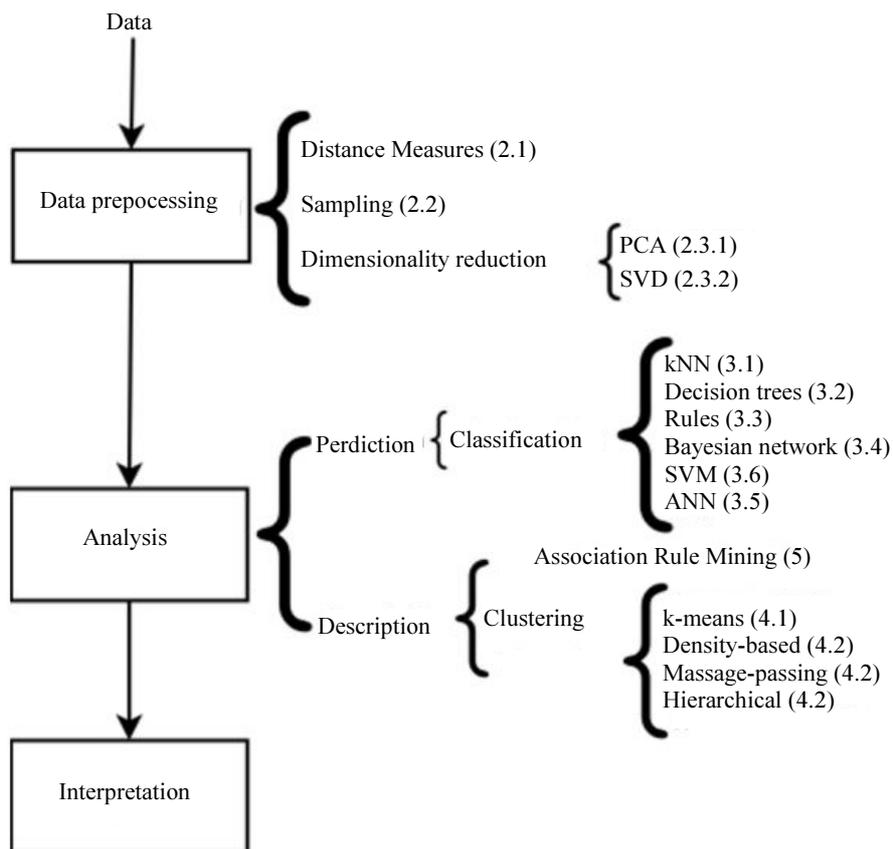


Fig. 8: Main steps and methods in a data recommender system (Amatriain *et al.*, 2011)

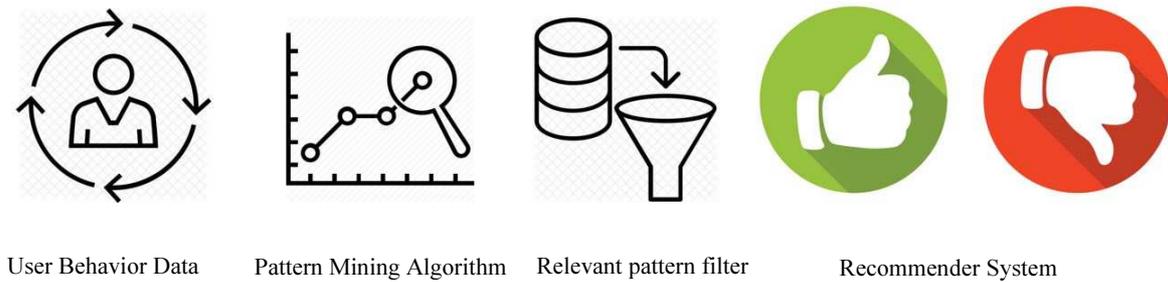


Fig. 9: Relevant behavior pattern filtering in the overall project (Zehnder *et al.*, 2015)

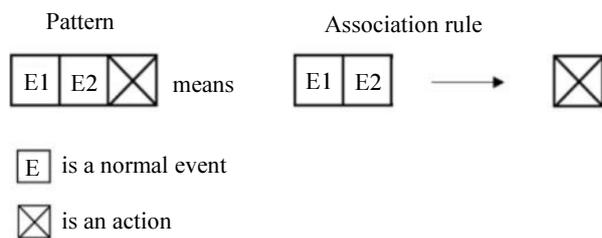


Fig. 10: Relevant patterns can be interpreted as an association rule (Zehnder *et al.*, 2015):

In each CSV file is a non-negative integer which records power demand of the downstream electrical load in Watt which is also merged every 30 min. All graphs for this analysis were produced using Power-BI program. We collected and analyzed the data to check the consumption time of each appliance in non-smart house as shown in Fig. 7 as an example of an electric heater in house 3.

Relevant Types of Behavior Patterns

To know the types of behavior patterns needed to suggest actions, we need to define the characteristics to identify the relevant patterns in the result set of the frequent (and/or periodic) pattern mining algorithm. The data analysis of a recommender system can be based on a variety of methods as shown in Fig. 8.

To be able to suggest actions by using relevant patterns, it must be composed of two main components as shown in Fig. 9. First, a relevant pattern must contain at least one action to lower energy usage in smart homes (action). Additionally, the pattern must consist of normal events served as a condition to suggest the action at the right time because a one-event condition is not enough to suggest an action. Pattern is created by at least three events, normally one action and two events, but in exceptional cases, three actions are possible as a relevant pattern consists of normal events (condition) and an action, it can be interpreted as the association rule Eq. (1) and Fig. 10.

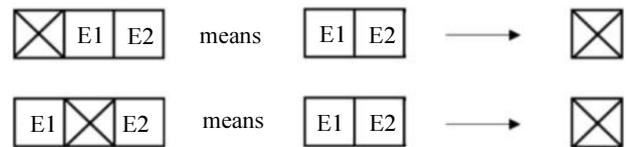


Fig. 11: Patterns with an action in the beginning or in the middle (Zehnder *et al.*, 2015)

An association rule is an implication of the form Equation 1:

$$X \rightarrow Y, \text{ where } X, Y \subset I, \text{ and } X \cap Y = \emptyset \quad (1)$$

The rule states that when X occurs, Y occurs with a certain probability (Deenadayalan *et al.*, 2014).

Association Rule doesn't consider the order of the items which defines time. Also, patterns can be explained as association rules. Patterns where the actions at the beginning or the center are defined as relevant patterns as shown in Fig. 11:

Patterns with Multiple Actions

As noted, patterns can also contain two or more actions. To apply the association rule approach for such patterns, one association rule is created for each action in a pattern. As actions are just a special form of events, the other actions must be treated as normal events and serve as a condition for the rule:

As shown in the example of Fig. 12, a four-event pattern containing 3 actions will result in 3 rules. For each action (1, 2 and 4) shown in the pattern, a rule is formed and the other actions areas normal events. To recommend an action, the condition must appear in the behavior data without an action.

Architecture

Recommender systems are used as information filtering systems to predict user preferences. In addition, it needs to suggest meaningful actions or items based on the interest of the user of the system.

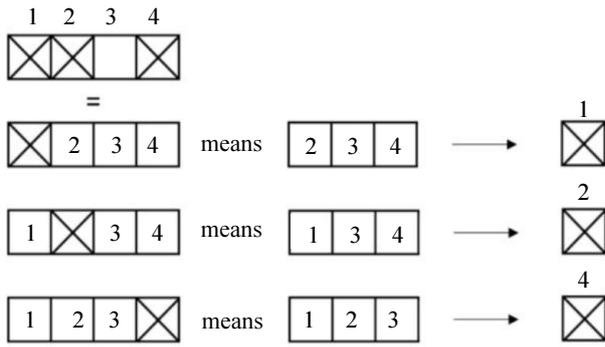


Fig. 12: A rule is created for each action in the pattern (Zehnder *et al.*, 2015)

The method chosen for this project is association rule. The design of a recommendation engine depends on the characteristics of the data. Data analysis methods and different techniques that can be classified into one of the three groups: Content-based systems, collaborative filtering systems and association rules-based systems (Ye and Huang, 2011). These methods can be combined in some exceptional cases.

The architecture of the recommender system developed in this project can be divided into three main parts:

- The storage of the association rules
- The event stream of the current behavior data inside the smart home
- The matching algorithm for both previous points

Prioritization

Designing of the recommender system allows matching more than one rule at the same time. For this reason, rules should be weighted to decide which rule is most suitable. Furthermore, excluding rules under a certain threshold could be done by prioritization criteria.

Some indicators can be considered when calculating the weight of a rule:

- 1) Support (count) of the pattern relative to all events
- 2) The number of events in a pattern
- 3) The position of the action in the pattern
- 4) Date when the rule was mined
- 5) The confidence of the pattern

As confidence must be calculated after mining by the recommendation system, the support (count) is calculated by the pattern-mining algorithm Window Sliding with De- Duplication (WSDD).

$$\text{Confidence } \left(\begin{array}{cccc} 1 & 2 & 3 & 4 \\ \square & \square & \square & \square \end{array} \right) \rightarrow \frac{\text{Support } \left(\begin{array}{cccc} 1 & 2 & 3 & 4 \\ \square & \square & \square & \square \end{array} \right)}{\text{Support } \left(\begin{array}{ccc} 1 & 2 & 3 \\ \square & \square & \square \end{array} \right)}$$

$$\text{Confidence } \left(\begin{array}{cccc} 1 & 2 & 3 & 4 \\ \square & \square & \square & \square \end{array} \right) \rightarrow \{0..1\}$$

$$\text{Because support } \left(\begin{array}{cccc} 1 & 2 & 3 & 4 \\ \square & \square & \square & \square \end{array} \right) \leq \text{Support } \left(\begin{array}{ccc} 1 & 2 & 3 \\ \square & \square & \square \end{array} \right)$$

Fig. 13: Formula to calculate the confidence of a rule (Zehnder *et al.*, 2015)

$$\text{Support } \left(\begin{array}{cccc} 1 & 2 & 3 & 4 \\ \square & \square & \square & \square \end{array} \right) \leq \text{Support } \left(\begin{array}{ccc} 1 & 3 & 4 \\ \square & \square & \square \end{array} \right)$$

Fig. 14: If the action is not the first or the last event, the support count is independent (Zehnder *et al.*, 2015)

$$\text{Confidence } \left(\begin{array}{cccc} 1 & 2 & 3 & 4 \\ \square & \square & \square & \square \end{array} \right) \rightarrow \frac{\text{Support } \left(\begin{array}{cccc} 1 & 2 & 3 & 4 \\ \square & \square & \square & \square \end{array} \right)}{\text{Support } \left(\begin{array}{ccc} 1 & 2 & 3 \\ \square & \square & \square \end{array} \right) + \text{Support } \left(\begin{array}{ccc} 1 & 3 & 4 \\ \square & \square & \square \end{array} \right)}$$

Fig. 15: Adapted formula for confidence (Zehnder *et al.*, 2015)

Confidence

Confidence in association rule mining shown in Eq. (2) is defined as follows: The rule holds in T with confidence conf if conf % of transactions that contain X also contain Y (Deenadayalan *et al.*, 2014):

$$\text{conf}(x \Rightarrow y) = \frac{\text{supp}(XY)}{\text{supp}(X)} \tag{2}$$

The confidence of a rule shows how often the action Y appears in pattern that contains event X (condition). If a rule has a confidence of 100% and without mining an action, no occurrence of a pattern. The confidence is calculated using pattern support (count) (Deenadayalan *et al.*, 2014) as shown in Fig. 13. The support count of a pattern including an action is divided through the support count of the pattern without action. This leads to the formula.

When mining overlapping patterns, the support of a longer pattern containing the same events as a shorter pattern must be equal or higher than the longer, the count of the shorter pattern will increase because of the occurrence of the longer pattern. Therefore, the result from the confidence calculated as shown in Fig. 13 will be between zero and one. However, the displacement of the action creates a new pattern with a support count that is completely independent from the original one in

case that the action is not at the beginning or at the end of the pattern.

The result of the confidence would not be between zero and one as shown in Fig. 14, therefore it won't be comparable to other patterns. Fig. 13 can be modified to be as shown in Fig. 15:

Functional prototype of the system must be developed to be used in the field test, it is not difficult to suggest or recommend an action. However, the action suggested must not decrease the comfort levels of inhabitants of the tested smart homes, we could allow the inhabitant to vote for each recommendation to measure if a suggestion decreases their comfort level.

Results

Recommended actions lowered energy usage in smart homes as the energy consumption of some devices was lower than before the execution of actions. This is shown in both figures: Fig. 16 during weekday and Fig. 17 during weekend. This was achieved either by reducing or turning off devices to lower their consumption. Turning off was the most obvious and effective approach to lower energy usage in a smart home provided that the actions were of the inhabitant's interest. Therefore, an action is categorized into 2 main attributes:

- Lower energy usage
- Does not decrease the comfort level

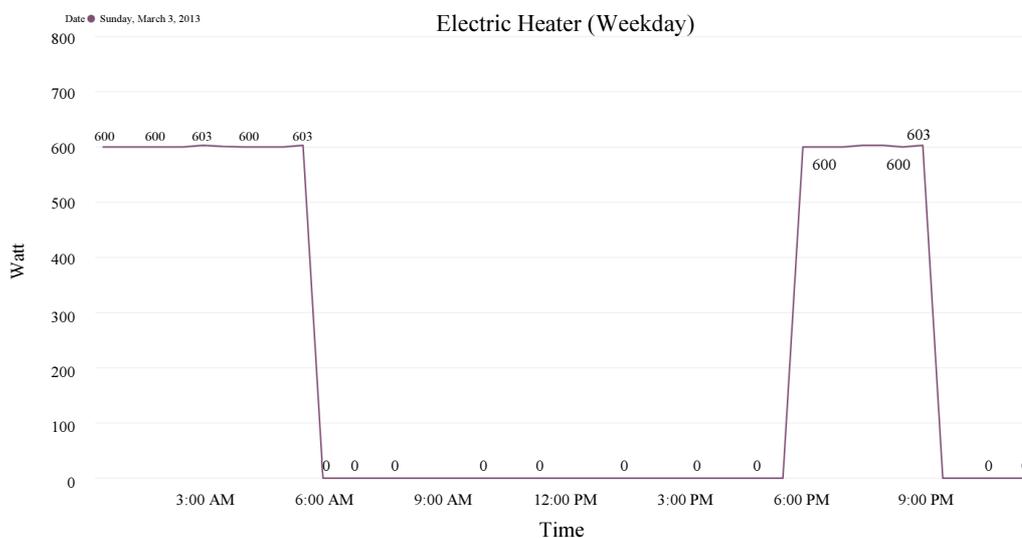


Fig. 16: Electric heater in house 3 (weekday) after applying recommended actions

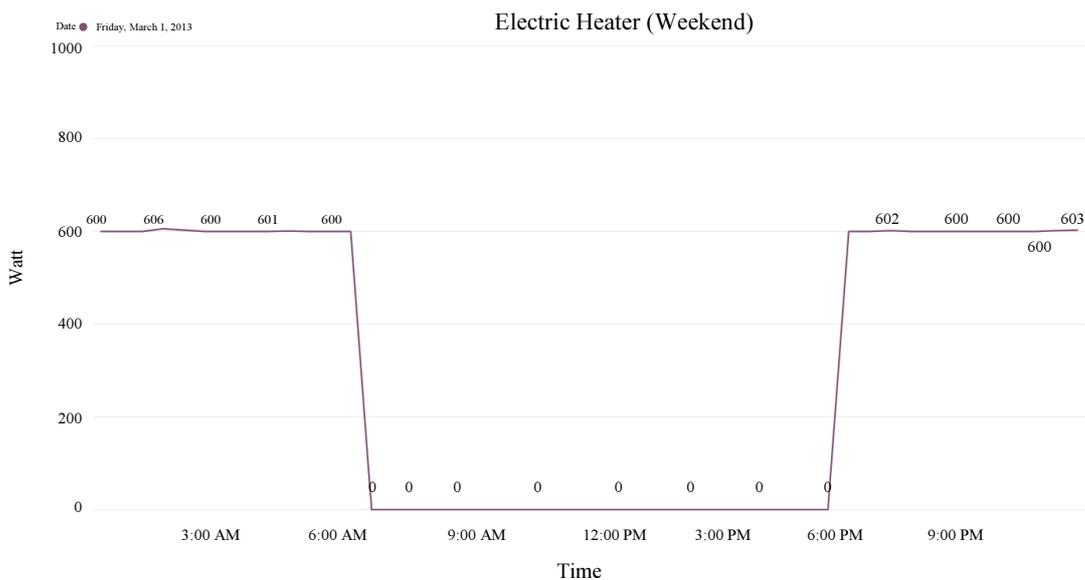


Fig. 17: Electric heater in house 3 (weekend) after applying recommended actions

Conclusion and Future Work

The proposed framework was applied on different houses and appliances that were measured and analyzed using k- means. Every house had its own usage which differed from other houses; some houses found that light was the highest consuming device, while others found the air-conditioner. Based on these findings, the recommender chose the most suitable action to lower the usage without decreasing the comfort level of inhabitants using apriori algorithm. After applying the recommended actions, most of the devices consumed less electricity.

Analysis to optimize decision making can be optimized by data analytics that uses data collection. To identify patterns, we need to review large amount of unorganized data that helps decision makers and enables a better understanding of behaviors. The idea of data analytics aims at handling huge volumes of data to identify trends, patterns and collect irreplaceable findings by applying BDA.

In future work, we need more information such as time between two actions, time of day, weekdays, or season when the pattern occurs; all of which will improve the accuracy of the suggestions of the recommender system by the mining algorithm. Better conclusions can be done with the help of such attributes to react on different behavior patterns in various times than just the date when the rule was mined (e.g., a pattern, which happened regularly 1 year ago on the same weekday as today could lead to better results as a pattern mined in recent times).

In addition, it would be better to recommend actions based on every inhabitant in the house and taking in consideration age, time, culture and gender as it may differ in recommending actions and results.

Author's Contributions

Ayman E. Khedr: Participation in highlighting the research problem, introducing and evaluating the research methodology, preparing the experiment, writing parts of the paper and reviewing the paper structure.

Sherif Kholeif: Participation in highlighting the research problem, introducing and evaluating the research methodology, preparing the experiment, writing parts of the paper and reviewing the paper structure.

Mira Tamer Shaker: Participation in highlighting the research problem, introducing, structuring and evaluating the research methodology, preparing and conducting the experiment, writing the paper and reviewing the paper structure.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

- Amatriain, X., A. Jaimes, N. Oliver and J.M. Pujol, 2011. Data mining methods for recommender systems. Research Gate.
DOI: 10.1007/978-0-387-85820-3_2
- Ayaz, K., M. Sulemani and N. Ahmed, 2017. Efficient energy performance within smart grid. *Scientific Res.*, 3: 75-86. DOI: 10.4236/sgre.2017.83005
- Bhati, A., 2017. Energy conservation through smart homes in a smart city: A lesson for Singapore households. *Energy Policy*, 104: 230-239.
DOI: 10.1016/j.enpol.2017.01.032
- Chen, H., X. Xie, W. Shu and N. Xiong, 2016. An efficient recommendation filter model on smart home big data analytics for enhanced living environments. *Sensors*, 16: 1706-1706.
DOI: 10.3390/s16101706
- Conrad, E., 2016. Presence detection from smart home motion sensor datasets: A model. *Proceedings of the 14th Mediterranean Conference on Medical and Biological Engineering and Computing, (BEC' 16)*, Springer, Cham, pp: 1249-1255.
DOI: 10.1007/978-3-319-32703-7_240
- Cook, D., M. Schmitter-Edgecombe, A. Crandall and B. Thomas, 2017. Collecting and disseminating smart home sensor data in the CASAS project. Researchgate.
- Crammer, K., A. Kulesza and M. Dredze, 2013. Adaptive regularization of weight vectors. *Adv. Neural Inform. Process. Syst.*, 91: 155-187.
DOI: 10.1007/s10994-013-5327-x
- Deenadayalan, E., D. Hanirex and K. Kaliyamurthie, 2014. Association rule mining in distributed database system. *Int. J. Comput. Sci. Mobile Comput.*, 3: 727-732.
- Ebeid, E., R. Heick and R. Jacobsen, 2016. Presenting user behavior from main meter data. *Proceedings of the IEEE International Conference on Smart Grid Communications*, Nov. 6-9, IEEE Xplore Press, Sydney, NSW, Australia, pp: 594-599.
DOI: 10.1109/SmartGridComm.2016.7778826
- El Seddawy, A.B., T. Sultan and A.Khedr, 2013. Enhanced K-mean algorithm to improve decision support system under uncertain situations. *Int. J. Computer Science and Network Security*, 13: 50-58.
- Energy, M.O., 2017. Egypt electricity consumption. CEIC.

- Hadi, J., H. Shnain and S. Hadishaheed, 2015. Big data and five v's characteristics. *Int. J. Adv. Electron. Comput. Sci.*, 2: 16-23.
- Han, D. and J. Lim, 2010. Smart home energy management system using IEEE 802.15.4 and zigbee. *IEEE Trans. Consum. Electron.*, 56: 1403-1410. DOI: 10.1109/TCE.2010.5606276
- Harper, R., 2003. *Inside the Smart Home*. 1st Edn.m Springer, ISBN-10: 1852336889, pp: 280.
- Hegazy, A.F., A.E. Khedr and Y. Al Geddawy, 2015, An adaptive framework for applying cloud computing in virtual learning environment at education a case study of "AASTMT". *Procedia Computer Science*, 65: 450-458.
- King, N., 2003. Smart home – a definition. Intertek Research and Testing Center, Milton Keynes.
- Khedr, A.E., M.A.Abdel-Fattah and Y. Nagm-Aldeen, 2015. A literature review of business process modeling techniques. *Int. J. Advanced Res. Computer Science Software*, 5: 43-47.
- Khedr, A.E., M. Kadry and G. Walid, 2015. Proposed framework for implementing data mining techniques to enhance decisions in agriculture sector Applied case on Food Security Information Center Ministry of Agriculture, *Procedia Computer Science*, 65: 633-642.
- Khedr, A.E. and J.N. Kok, 2006, Adopting knowledge Discovery in Databases for Customer Relationship Management in Egyptian Public Banks. In: IFIP International Federation for Information Processing, Professional Practice in Artificial Intelligence, J. Debenham (Eds.), Boston: Springer, pp: 201-208.
- Khedr, A.E., A.M. Idrees, A.E. Hegazy and S. El-Shewy, 2018. A proposed configurable approach for recommendation systems via data mining techniques. *Enterprise Information Syst.*, 12: 196-217. DOI: 10.1080/17517575.2017.1293301
- Khedr, A.E., Idrees, A.M. and A.I. El Seddawy, 2016. Enhancing Iterative Dichotomiser 3 algorithm for classification decision tree. *Wiley Interdisciplinary Reviews: Data Mining Knowledge Discovery*, 6: 70-79. DOI: 10.1002/widm.1177
- Khedr, A.E., A.M. Idrees and A.I. El Seddawy, 2014. Performance tuning of K-mean clustering algorithm a step towards efficient DSS. *Int. J. Innovative Research in Computer Science Technology*, 2: 2347-5552.
- Khedr, A.E. and A.M. Idrees, 2017A. Adapting load balancing techniques for improving the performance of e-learning educational process. *J. Computers*, 12: 250-257. DOI: 10.17706/jcp.12.3.250-257
- Khedr, A.E. and A.M. Idrees, 2017B. Enhanced e-learning system for e-courses based on cloud computing. *J. Computers*, 12: 10-19. DOI: 10.17706/jcp.12.1.10-19
- Li, R., 2013. The usage of automation system in smart home to provide a sustainable indoor environment: A content analysis in web 1.0. *Int. J. Smart Home*, 7: 47-60.
- Marjani, M., F. Nasaruddin, A. Gani, A. Karim and I. Abaker *et al.*, 2017. Big IoT data analytics: architecture, opportunities and open research challenges. *IEEE Access*, 5: 5247-5261. DOI: 10.1109/ACCESS.2017.2689040
- Martirano, L., 2011. A smart lighting control to save energy. *Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems*, Sept. 15-17, IEEE Xplore Press, Prague, Czech Republic, pp: 132-138. DOI: 10.1109/IDAACS.2011.6072726
- Mayr, F., 2016. Improving human activity recognition by smart windowing and spatio-temporal feature analysis. *Proceedings of the 9th ACM International Conference on Pervasive Technologies Related to Assistive Environments*, Jun. 29-Jul. 01, ACM, Corfu, Island, Greece, pp: 19-37. DOI: 10.1145/2910674.2910697
- Moltagh, O., G. Foliente and G. Grozev, 2015. Knowledge-Mining the Australian Smart Grid Smart City Data: A Statistical-Neural Approach to Demand-Response Analysis. In: Geertman, S., J. Ferreira, Jr., R. Goodspeed and J. Stillwell (Eds.), Springer International Publishing, Switzerland, pp: 189-207.
- Nazier, M.M., A. Khedr and M. Haggag, 2013, Business intelligence and its role to enhance corporate performance management. *Int. J. Management Information Technol.*, 3: 8-15.
- Sami, A. and A. Honarvar, 2016. Extracting usage patterns from power usage data of homes' appliances in smart home using big data platform, *Int. J. Inform. Technol. Web Eng.*, 11: 39-50. DOI: 10.4018/IJITWE.2016040103
- Shukla, D., 2017. Making the connection between big data analytics and the internet of things. *Customer Think*.
- Talari, S., M. Shafie-Khah, P. Siano and V. Loia, 2017. A review of smart cities based on the internet of things concept. *Energies*, 10: 421-421. DOI: 10.3390/en10040421
- Wich, M. and T. Karmer, 2016. Enrichment of smart home services by integrating social network services and big data analytics. *Proceedings of the 49th Hawaii International Conference on System Sciences*, Jan 5-8, IEEE Xplore Press, Koloa, HI, USA, pp: 425-434. DOI: 10.1109/HICSS.2016.59
- Ye, X. and J. Huang, 2011. A framework for cloud-based smart home. *Proceedings of the International Conference on Computer Science and Network Technology*, Dec. 24-26, IEEE Xplore Press Harbin, China, pp: 894-897. DOI: 10.1109/ICCSNT.2011.6182105

Zehnder, M., H. Wache, H. Witschel, D. Zanatta and M. Rodriguez, 2015. Energy saving in smart homes based on consumer behavior: A case study. Proceedings of the IEEE 1st International Smart Cities Conference, Oct. 25-28, IEEE Xplore Press, Guadalajara, Mexico, pp: 1-6.
DOI: 10.1109/ISC2.2015.7366231

Zoha, A., A. Gluhak, M. Imran and S. Rajasegarar, 2012. Non-intrusive load monitoring approaches for disaggregated energy sensing: A Survey. Sensors, 12: 16838-16868. DOI: 10.3390/s121216838