A Study into Smart Grid Consumer-User Profiling for Security Applications

Mutinta Mwansa, William Hurst, Carl Chalmers, Yuanyuan Shen, Aaron Boddy
Department of Computer Science
Liverpool John Moores University
Byrom Street
Liverpool, L3 3AF, UK
Email: M.Mwansa@2017.ljmu.ac.uk, {W.Hurst, C.Chalmers, Y.Shen}@ljmu.ac.uk, A.Boddy@2011.ljmu.ac.uk

***Abstract*— A smart meter measures energy consumption with more granular detail than conventional analogue meters. The Advanced Metering Infrastructure (AMI) facilitates real-time two-way communication between the consumer and the rest of the energy grid. Information concerning electricity consumption, demand and response and home energy generation is communicated back to the local utility for monitoring and billing purposes. However, the detailed electricity usage patterns and trends can also be used to understand daily consumer habits and their routines. The collection and analysis of such data raises significant security and ethical concerns which must be adequately addressed. This paper focuses on the data collected from the residential smart grid using its default reading frequency of 30 minutes. The research demonstrates how the information can be exploited to remotely profile users and detect abnormal user behaviours using cloud-based analytics. Security implications are outlined and a case study is put forward as a demonstration of information that can be obtained through consumer profiling.**

Keywords- Smart meters; Profiling; Advanced Metering Infrastructure, Data Classification, Data Analysis.

I. INTRODUCTION

Smart meters are a core component of the smart grid, which is a complex dynamic network of interconnected devices. This infrastructure provides mechanisms for information exchange, decision making and actuation. Smart grid systems include producers, consumers and actors to ensure a resource saving and economically efficient electrical network. Typically, they reduce financial losses, operational costs and enable the suppliers to forecast their customers' demands [1]. As a result, smart meters are being implemented on a global scale. Many countries such as the UK, USA, Australia and Italy are already advanced in their smart meter implementation. The UK alone is aiming to install over 50 million gas and electricity smart meters by 2020 [2].

The smart grid represents a technological era of resilience, performance and efficiency across the entire power industry; from generation all the way to consumption. While the benefits of the smart grid are clear, it also introduces a number of different risks and challenges. The complexity and interoperability of the smart grid mean that it is exposed to a series of digital threats from invasion of privacy to a sabotage of a critical national infrastructure [3]. As the smart grid is highly interconnected, security attacks can originate from various points. As a result, electric industries invest heavily in cyber-threat mitigation [4]. In particular, the acquired data from the smart metering infrastructure poses a significant risk to both the grid and its stakeholders if a security breach occurs. The research composure in this paper demonstrates the type of sensitive information which can be constructed from smart meter data using its default reading frequency of 30 minutes. Here detailed routines of the occupant can be exposed using different profiling techniques and data analytics.

The data used in this paper is collated from a smart meter trial deployed in 75,000 homes. The remainder of the paper is as follows. A background research on smart grid systems and associated technologies is put forward in Section 2. Subsequently, typically section 3 presents a sample of the data collected from our smart meter case study. Both data visualisations and statistical analysis of the data is undertaken. Section 4 discusses the methodology and techniques used for profiling users and highlighting the benefits of cloud computing for data processing within the wider smart grid. The paper is concluded in Section 5 where a discussion of the results is presented. In particular, this paper focuses on the smart meter and investigates the novel approaches for consumer profiling and for the consumers to monitoring energy usage in real time.

II. BACKGROUND RESEARCH

A smart meter is an electronic device that records consumption of utility services (such as electricity and gas) at fixed intervals. It replaces existing analogue meters where energy usage readings are collected manually usually over a longer period. The system automatically communicates consumption information using a predefined schedule, to the Meter Data Management System (MDMS).

A. Smart Meters

Typically, the main aim of the smart meter is to facilitate real time energy usage readings at granular intervals, to both the consumer and smart grid stakeholders [5]. In order to achieve this aim, load information is obtained from consumer electrical devices while measuring the total aggregated energy consumption for the given property. Additional information, such as home generated electricity is provided to the utility company and/or system operator for enhanced monitoring and accurate billing. Some of these roles and benefits include:

 ISBN: 978-1-61208-607-1

- Accurate recording, transmitting and storing of information for defined time periods (to a minimum of 10 seconds). All UK smart meters must store energy usage readings for a maximum of 13 months providing a unique insight into energy consumption.
- Offer two way communications to and from the meter so that, for example, suppliers can read meters remotely [6], facilitate demand and response and upgrade tariff information.
- Support future management of energy supply to help distribution companies manage supply and demand across their networks [7]. This is achieved automatically through previously agreed Demand Response (DR) actions.

A significant amount of research exists on how the data collected from the smart meters can be used to detect energy usage patterns in residential homes via user profiling [8] However, investigations into the security behind the security of smart meter analytics, is still a challenging and prevalent area of research [9].

The collector of the device retrieves the data and may process it or simply pass it on for processing upstream. Data is transmitted via a Wide Area Network (WAN) to the utilities central collection point for processing and use by business applications. Since the communications path is two way, signals or commands can be sent directly to the meters, customer premise or distribution device. The combination of the electronic meters with two-way communications technology for information, monitor and control is commonly referred to as the AMI.

B. *The AMI*

The AMI facilitates the bidirectional communication between the consumer and the rest of the smart grid stake holders. It reduces the traditional need for energy usage readings to be collected manually [10]. The smart meter is able to communicate with a gateway through a Home Area Network (HAN), Wide Area Network (WAN) or a Neighbourhood Area Network (NAN), which is outlined as follows in Figure 1:

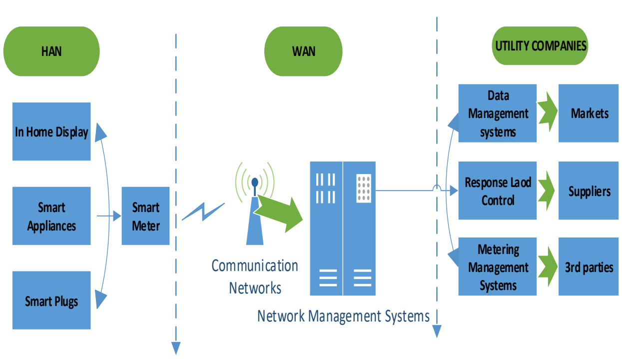

Figure 1. Advanced Metering Infrastructure

The HAN is housed inside the consumer premises and is made up of different devices e.g., Meters, Thermostats, Electric storage devices, ZigBee transmitters. The HAN contains both the electrical and gas smart meter, which generates detailed consumption data. The data generated is transmitted in NANs and WANs and eventually to the control station for power corrective measure [11]. The HAN is responsible for providing communication between electrical devices and the access points. The WAN handles the communication between the utility companies and the HAN. The WAN is responsible for sending all meter data to the utility, using a robust backhaul network, such as carrier Ethernet, GSM, CDMA or 3G [12].

All of the acquired data is sent to the Meter Data Management System (MDMS), which is responsible for storing, managing and analysing the data [13]. The MDMS sits within the data and communications layer of the AMI. This component is an advanced software platform, which deploys data analytics while facilitating the various AMI applications and objectives. These applications include: managing metered consumption data, outage management, demand and response, remote connect / disconnect, and smart meter events and billing [14]. This information can be shared with consumers, partners, market operators and regulators.

C. *Smart Grid Challenges concerning User Profiling*

The major concern for profiling users is the privacy of the consumer. As demonstrated in Section 4, smart meters enable detailed profiling of consumers, their energy usage and the household activities. This information can be used by others, either maliciously or inadvertently to ascertain an insight into an individual's home life. For example, activities or occupancies of a home for specific periods of time can be determined. In a general sense, analysis of granular smart meter energy data could result in 1) invasion of privacy; 2) unwanted publicity and embarrassment (e.g., public disclosure of private facts or the publication of facts which place a person in a false light); 3) endangering of the physical security of life. The security policies governing the reliability of the smart grid depend on appropriate connectivity protocols and the national institute of standards and technology being the reference model proposed. [15]. Behaviour profiling and confidentiality of security for the consumers as it is inextricably linked to their privacy. This research is devoted to presenting ways of ensuring confidentiality and privacy within the smart home/smart grid communication interoperability.

D. *Cloud Computing and the Smart Grid*

Cloud computing is currently used to mitigate the data processing challenges that are associated with the smart gird implementation. The data generated from the smart grid means that cloud processing platforms are now required to process and extract meaning from the acquired data while

 ISBN: 978-1-61208-607-1

ensuring a robust energy delivery network. There are numerous advantages that are associated with cloud computing platforms, which can be applied to the smart metering infrastructure to support its various objectives [16]. Cloud computing is an ever developing computational platform, which combines hardware, storage and high bandwidth networking to provide scalable solutions to third party organisations. The smart grid requires a fault tolerant, efficient data processing and communications infrastructure in order to deliver a reliable and affordable power distribution network [17]. The emergence of smart grids brings many benefits but also various challenges in terms of data management and integration. The smart grid by its very nature is a complex platform with vast storage, communication and computational requirements. To facilitate these requirements smart grids can leverage the following cloud computing benefits:

- Cloud computing is flexible and scalable. This ensures adequate resource allocation and provisioning [18]. As smart grid components are deployed on a large scale, cloud computing can be used to overcome scalability problems by provisioning additional resources as required.
- Cloud services maintain the underlying computational hardware and software. Smart grids are regarded as a critical infrastructure, which supplies essential utilities to the consumer. Any down time in services can have a detrimental impact on service users. As most cloud components are virtualised, guests can be migrated from one host to another while maintenance is undertaken. This removes the need for downtime and minimises service disruption [19].
- Many cloud providers are geographically distributed, which not only ensures low latency but also provides service replication. Essentially, services are mirrored in one or more additional data centres to prevent service disruption in the event of an outage.

III. CASE STUDY

The project outlined in this paper theorises that the detailed electricity usage patterns generated by smart meters describes the vast amount of data collected within the smart grid. In order to process and analyse large volumes of data, we propose the use of cloud computing because of its high performance computing resources and the high capacity storage devices [20].

The amount of data required to process transactions of two million customers reaches upwards of 22 gigabytes per day [21]. It is a significant challenge to manage this data; which may include the selection, deployment, monitoring, and analysis processes. A real-time information processing is usually required in the smart grid to meet the needs of smart grid condition monitoring based on the smart grid condition monitoring with cloud computing [22]. Any delay may cause a serious consequence in the whole system, which has to be avoided as much as possible. As such, the methodology put forward in this paper makes use of a cloud platform for data processing

By the end of 2020, the UK government plans to have smart meters installed in every household and commercial business. Providers are able to use this resource by integrating their own software frameworks through an agreed communication standard. Smart meters utilise the ZigBee Smart Energy profile, which can be used to pair Consumer Access Devices CAD's using the ZigBee protocol. ZigBee has an operating range up to 70 meters with a data transmission speed of 250kbs. In addition, the UK DECC have declared SMETS2, which cites the use of ZigBee Smart Energy 1.x. This facilities access to smart meter data for both consumers and other 3rd parties.

A. *Data Study*

In this sub-section, a demonstration of the data that is collected from smart meters and how it can be analysed to model user behaviour is demonstrated. Table 1 below demonstrates a sample of smart meter data collected over a period of one month (January) for a single home occupant. The general supply of energy used on a daily basis (the energy delivered) is measured in KWH and can be described as what is used to bill the customer. Figure II shows an example of energy reading of an individual household meter. Data is being collected over a 30 min time interval period and the "energy delivered" in KWH. The customer key is the primary key used to identify the consumer while the End Date Time highlights the time and date of the acquired reading. Both the general supply and off peak supply are recorded based on the specified tariff.

TABLE 1. INDIVIDUAL READING SHOWING A MONTH USAGE PERIOD.

1	CUSTOMER_KEY	End Datetime	General Supply KWH	Off Peak KWH	Year
2	8410148	1/1/2013 0:29	0.081	0	2013
3	8410148	1/1/2013 0:59	0.079	0	2013
4	8410148	1/1/2013 1:29	0.082	0	2013
5	8410148	1/1/2013 1:59	0.085	0	2013
6	8410148	1/1/2013 2:29	0.073	0	2013
7	8410148	1/1/2013 2:59	0.07	0	2013
8	8410148	1/1/2013 3:29	0.07	0	2013
9	8410148	1/1/2013 3:59	0.072	0	2013
10	8410148	1/1/2013 4:29	0.071	0	2013
11	8410148	1/1/2013 4:59	0.074	0	2013

In order to visualise and analyse the energy usage patterns the smart meter data was extracted, transformed and loaded into a data model. The software used for this task was Microsoft Power BI. The platform facilitates the aggregation of data from multiple sources including both on premise and cloud infrastructure. Fig. 2 presents an example of 70K household meter readings showing the energy usage and the

 ISBN: 978-1-61208-607-1

behaviour trend over a period of 12 months. Here we can see the general distribution of energy readings highlighting the energy requirements for different households. This type of data visualisation could give suggestion to the number of occupants living in a given premise. Houses with increased energy usage are more liking to have an increased number of occupants.

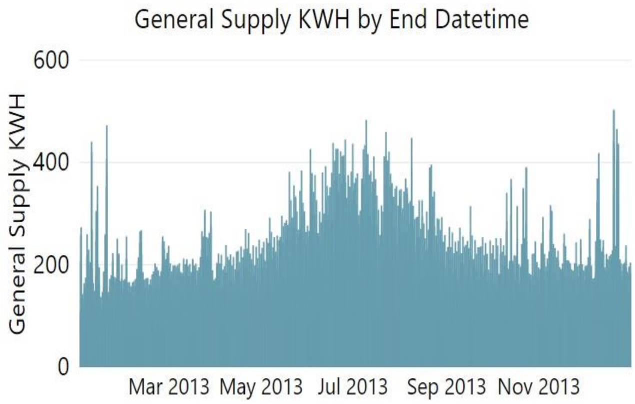

Figure 2. A Meter reading showing 12 months usage.

Fig. 3 highlights numerous meter readings taken over a 7 month period to demonstrate what effects a seasonal change has on energy usage.

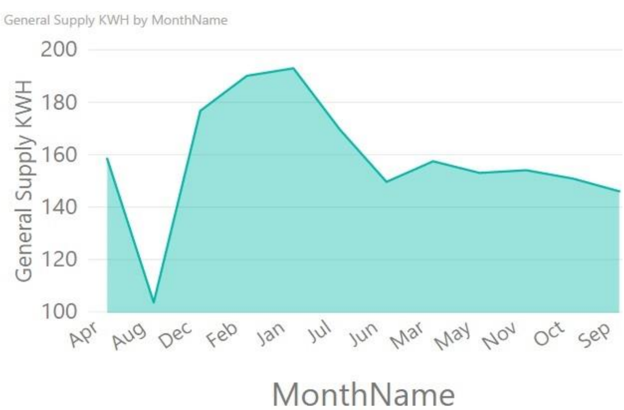

Figure 3. A Meter reading showing several months usage.

The data shows the trend of which period during the year consumers in residential homes use energy less. The high peaks indicate the energy is mostly used probably because of the winter when there is drastic drop in temperatures. As the data shows, it is very much possible to identify electrical device behaviour from smart meter data collection.

B. Stastical Analysis

A Smart-frame cloud computing, is a flexible, scalable, and secure information management framework for smart grids based on cloud computing technology. Our idea is to build the framework at three hierarchical levels: top, regional, and end user levels in which the first two levels consist of cloud computing centres while the last level contains end-user smart devices. The top cloud computing centre takes responsibility of managing general devices and accumulation of data across the regional cloud computing centres which are placed in the lower level in the hierarchy. The regional cloud computing centres are in turn responsible for managing intelligent devices, which have lower hierarchical level than the regional cloud computing centres in specific regions (e.g., with in a city), and processing data of these devices. The figures below show and compare two users one with normal behaviour and one with abnormal behaviour energy usage patterns. Figure 4 presents a scatter graph for one smart meter in the trail. Here the results show that the consumer does not use electricity with any repeatability in behaviour.

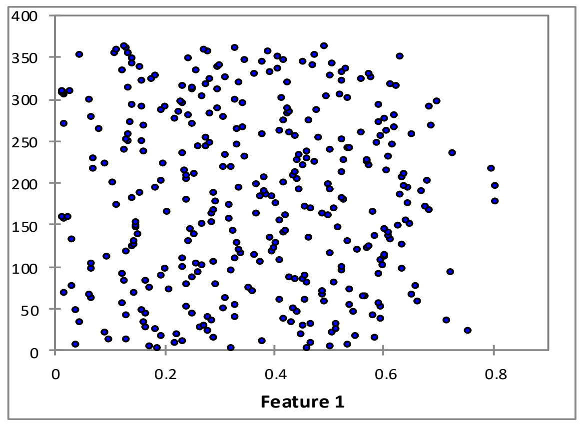

Figure 4. A Meter reading showing abnormal energy usage.

Fig. 5 presents a scatter graph for the second smart meter. Here, the consumer shows clear repeatability in patterns of behaviour while aspects of abnormal behaviour can be observed in the outliners.

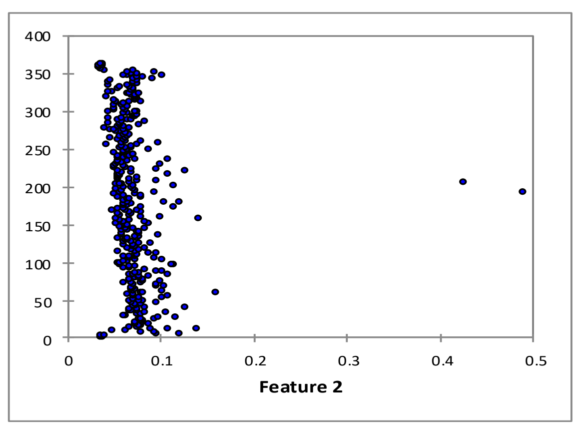

Figure 5. A Meter reading showing normal energy usage.

IV. METHODOLOGY

In the following section, we propose a novel system to facilitate the handling and analysis of smart meter data.

The solution proposed below will help network operators to identify, control and manage security risks in smart grid infrastructures and also to establish a detailed correlation between energy usage, weather conditions and other events as well as data management solutions.

Clear deviations can be seen from the above two energy users from their behaviour; the patterns reflect an individual's unique behavioural characteristics. Figure VI shows how much a meter can process a vast amount of data in a short period and if we have a number of such users then data storage is extreme and hence proposal of the cloud computing framework below.

A. System Design

A cloud computing based framework for big data information management in smart grids provides not only flexibility and scalability but also security. As displayed in Figure 6, the chosen topology to be implemented in this project is discussed as follows:

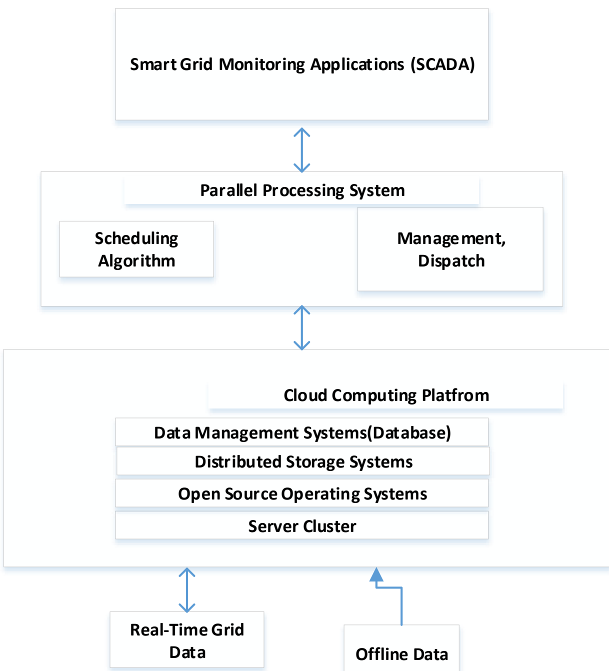

Figure 6. Cloud Computing framework for Smart Grid Monitoring Application

Smart grid monitoring application (SCADA): It is an application layer to monitor the status of the main power feeders, branch circuits, electrical equipment's, distributed generation units, energy storage units and the different parts of the smart grid to determine the real time online.

Parallel processing system: The amount of data collected on the smart grid is massive, distributed and complex. The parallel processing system enables the system to use a parallel computing to utilise the in demand computational and storage resource of cloud computing.

Cloud computing platform: In this architecture, cloud computing offers web based resizable computing capacity in the cloud, massive data management systems and distributed storage systems. Cloud platforms can play an important role in software architectures that allow more effective use of smart grid applications.

V. CONCLUSION

In this paper, the security implications of the smart meter installation are outlined. The electronic meters for electricity (smart meters) are undergoing an increasing deployment in private homes all over the world. As a consequence, an ever growing physical communication network, made up of millions of local meters, has been established. Likewise a complex data processing infrastructure has emerged which exploits numerous technologies and services to deliver an automated smart metering system.

The benefits of the smart meter implementation are vast. However, many of the considerable advantages are so far in favour primarily, if not solely, of the energy distributors. They provide a simplified, more efficient, and less costly transaction with customers, e.g., for meter reading, billing, and energy supply administration. The detail and granularity of the data collected can be used in many ways by utility companies. Some examples include forecasting energy usage, demand and response and consumer profiling. However there are future challenges facing the smart gird implementation.

Smart grids require an enormous pool of computing and massive data storage requirements are discussed. Cloud computing is proposed to overcome these demands by providing highly scalable computing resources to host smart grid applications. The details and granularity can be used to address many current and future challenges, which are faced by the grid. One of the main challenges is to meet the processing and storage of the vast data collected by the smart meter. The smart grid infrastructure will need information technology support to integrate data flow from numerous applications, to predict power usage and respond to events. Cloud computing services are ideally suited to support such data intensive always repeatedly. A significant issue surrounding security and data protection remain an ongoing challenge for smart grid operators. The data posed in this paper highlights how energy usage information can be used to profile both large numbers of households and individual consumers,

Our future work will be to implement coordinated fault protection mechanism with the help of cloud-based infrastructure in smart grid. In this infrastructure, different

 ISBN: 978-1-61208-607-1

pieces of equipment are able to perform together efficiently to implement privacy preserving data collection techniques such as billing, and real time monitoring. The use of machine learning (specifically anomaly detection) will be integrated into our approach to facilitate the real time detection of abnormal behaviour within the smart metering infrastructure.

REFERENCES

[1] S.Florian, "Smart grid security, Innovative solutions for a modernized smart grid," 2015.

[2] J.Zheng, D.W. Gao, and L.Lin, " Smart meters in smart grid: An overview. In Green Technologies," Conference IEEE, pp. 57-64, April, 2013.

[3] E.D. Knapp, J.T. Longil, " Industrial Network Security: Securing Critical Infrastructure Networks and Smart Grid, SCADA and other Industrial Control Systems," Chapter, 12, 2015.

[4] B.Obama, "Taking the cyberattack threat seriously," Wall Street Journal, 19, 2012.

[5] W. Lingfeng, V. Devabhaktuni, N. Gudi, "Smart Meters for Power Grid – Challenges, Issues, Advantages and Status," IEEE/PES Power Systems Conference and Exposition, PSCE, pp. 1-7, March , 2011.

[6] C. Bennett, D. Highfill, "Networking AMI Smart Meters," November 2008, IEEE Energy2030.

[7] M. Anas, N. Javaid, A. Mahmood, S. M. Raza, U. Qasim and Z. A. Khan, "Minimizing Theft using Smart Meters in AMI," Seventh International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, 2012.

[8] C. Chalmers, W. Hurst, W, M.Mackay, & P. Fergus, "Smart Meter Profiling For Health Applications. In the Proceedings of the International Joint Conference on Neural Networks," July, 2015

[9] K.K.R.Choo, "The cyber threat landscape: Challenges and future research directions: Computers & Security," vol. 30 no. 8, pp. 719-731, 2011.

[10] M.Popa, "Data Collecting from Smart Meters in an Advanced Metering Infrastructure, Proceedings of 15th International Conference on Intelligent Engineering Systems," 2011.

[11] R. Robinson, J. McDonald, B. Singletary, D. Highfill, N. Greenfield, M. Gilmore Advanced metering Security Threat Model .

[12] D. Niyato, P.Wang, "Cooperative transmission for meter data collection in smart grid," IEEE Communications Magazine, vol. 40, 2012.

[13] Y.T. Hoi, F.T. Kim , T.C. Kwok, C.T. Hoi, R.C. Hao, P.Gerhard, Hancke, F.M. Kim, "The Generic Design of a High-Traffic Advanced Metering Infrastructure Using ZigBee," vol. 10, pp. 836-844, 2014.

[14] B.Coalton, and H. Darren, "Networking AMI Smart Meters," IEEE Energy, vol. 2030, November, 2008.

[15] S.Bera, S.Misra, and J.J. Rodrigues, "Cloud computing applications for smart grid: A survey. IEEE Transactions on Parallel and Distributed Systems," vol .26, no. 5, pp.1477-1494, 2015.

[16] D.S.Markovic, D.Zivkovic, I. Branovic, R. Popovic, and D. Cvetkovic, "Smart power grid and cloud computing:Renewable and Sustainable Energy Reviews," vol. 24, pp.566-577, 2013

[17] A. Gupta, L.V. Kale, F. Gioachin, V. March, C.H. Suen, B.S. Lee, P. Faraboschi, R. Kaufmann, and D. Milojicic, "The who, what, why and how of high performance computing applications in the cloud. In Proceedings of the IEEE International Conference on Cloud Computing Technology and Science,", no.5,July,2013.

[18] M.Armbrust, A. Fox, A. Griffith, R. Joseph, A.D. Katz, R. Konwinski, A.Lee, G. Patterson, D. Rabkin, A. Stoica, and M.Zaharia, "A view of cloud computing. Communications of the ACM", vol.53, no.4, pp.50-58,2010.

[19] J.Baliga, R.W.Ayre, K. Hinton, and R.S.Tucker, "Green cloud computing: Balancing energy in processing, storage, and transport," Proceedings of the IEEE, vol. 99. No. 1, pp. 149-167,2011.

[20] B.Neenan, and R.C. Hemphill, "Societal benefits of smart metering investments. The electricity journal," vol. 21. No. 8. pp. 32-45,2008.

[21] J. Zheng, D.W. Gao, and L. Lin, "Smart meters in smart grid: An overview. In Green Technologies Conference," IEEE, pp. 57-64, April, 2013

[22] B.David "is a contributing editor at Scientific American, Is the U.S. grid better prepared to prevent a repeat of the 2003 blackout," Available from: [www.scientificamerican.com/article/us-electrical-grid-better-prepared-than-2003-blackout-ask-the-experts/], August, 2013.

 ISBN: 978-1-61208-607-1