Internet of Things and Data Mining: From Applications to Techniques and Systems

*Mohamed Medhat Gaber†, Adel Aneiba‡, Shadi Basurra§, Oliver Batty¶Ahmed M. Elmisery‖Yevgeniya Kovalchuk**, Muhammad Habib Ur Rehman††

Article Type:

Advanced Review

Abstract

The Internet of Things (IoT) is the result of the convergence of sensing, computing, and networking technologies, allowing devices of varying sizes and computational capabilities (things) to intercommunicate. This communication can be achieved locally enabling what is known as *edge and fog computing*, or through the well-established Internet infrastructure, exploiting the computational resources in the cloud. The IoT paradigm enables a new breed of applications in various areas including health care, energy management and smart cities. This paper starts off with reviewing these applications and their potential benefits. Challenges facing the realisation of such applications are then discussed. The sheer amount of data stemmed from devices forming the IoT requires new data mining systems and techniques that are discussed and categorised later in this paper. Finally, the paper is concluded with future research directions.

*All authors contributed equally to the paper with the first author coordinating the overall project.
†School of Computing & Digital Technology, Birmingham City University, Birmingham, United Kingdom
‡School of Computing & Digital Technology, Birmingham City University, Birmingham, United Kingdom
§School of Computing & Digital Technology, Birmingham City University, Birmingham, United Kingdom
¶School of Computing & Digital Technology, Birmingham City University, Birmingham, United Kingdom
‖Department of Electronic Engineering, Universidad Tecnica Federico Santa Maria, Valparaiso, Chile
**School of Computing & Digital Technology, Birmingham City University, Birmingham, United Kingdom
††Department of Computer Science, National University of Computer and Emerging Sciences, Lahore, Pakistan

Introduction

The massive adoption of Internet of Things (IoT) opens a plethora of new use cases, applications, frameworks, and data processing architectures. A new ecosystem of supporting technologies is being developed in parallel with IoT to enable resource provisioning for resource-constrained devices and systems (F. Wang, Hu, Hu, Zhou, & Zhao, 2017; Baktir, Ozgovde, & Ersoy, 2017; Mao, You, Zhang, Huang, & Letaief, 2017). The core of future IoT systems will be designed by integrating mobile edge computing systems, software-defined networks, 5G, augmented reality, and data mining (including machine learning and artificial intelligence) to name a few (Baktir et al., 2017; Mao et al., 2017). Data mining is the process of discovering hidden knowledge patterns from raw data; therefore, the execution of knowledge discovery processes in IoT environments will leverage the utility of IoT systems. In essence, data mining will play a vital role in highly interactive and intelligent IoT systems.

The adoption of IoT systems at every level from small and medium organisations to large-scale multinational enterprises creates unlimited opportunities (Hsu & Yeh, 2017). In addition, governmental and non-governmental non-profit organisations are willing to adopt IoT systems to improve their services. The continuous growth of data streams in these IoT systems will help in developing new business models, improved and massively customised products, and real-time personalised services (Rehman, Chang, Batool, & Wah, 2016). The integration of data mining and knowledge discovery processes in IoT applications will facilitate development of highly intelligent IoT systems considering operational efficiency and performance of businesses, governmental, and non-governmental organisations (M. M. Gaber, Gomes, & Stahl, 2014).

Data mining methods in IoT systems are integrated in order to discover a variety of knowledge patterns using well established supervised, unsupervised, semi-supervised, and statistical methods (Cao, Wachowicz, & Cha, 2017). These data mining methods enable classification, clustering, frequent pattern mining, and regression of incoming streaming data in order to visualise the knowledge and activate the actuators in IoT systems (Patel, Ali, & Sheth, 2017). Since the data mining methods vary in terms of operations (such as data cleaning, shaping, ingestion, preprocessing, model training, testing, visualisation, and actua-

tion) and computational complexity (such as memory and CPU utilisation), IoT applications could be scaled up from IoT devices to edge and cloud servers.

Numerous review articles were presented for data mining and IoT in existing systems whereby researchers focused on ubiquitous environments (Alam, Mehmood, Katib, Albogami, & Albeshri, 2017), mobile data stream mining applications (Rehman, Sun, Wah, & Khan, 2016), and big IoT data analytics (E. Ahmed et al., 2017). However, to the best of our knowledge, existing articles significantly lack in terms of data mining applications, techniques and systems. The main objective of this article is to present a review of IoT applications from data mining perspective considering healthcare, energy, and smart cities use cases and highlight some important research challenges. To this end, we present a detailed literature review and discuss the related issues for privacy preservation, networking and deployment of IoT applications in the edge computing environments. Finally, we present a detailed review of data mining methods, which were adopted to be deployed on the edge servers that reside at single-hop distance from IoT devices.

IoT Applications

IoT has opened the door for a sheer number of applications that were otherwise unreleasable. Out of these applications, healthcare, energy management and smart cities are identified to be the areas to benefit the most from this rising technological advancement. This is due to the use of wearables and smart sensing ambiently, enabling applications in these areas. At the same time, IoT technologies are expected to be adopted in a wide range of other applications. For example, a system proposed in (Mehmood et al., 2017) uses IoT for enhancing learning experience through monitoring of the learners' activities using a variety of devices (things) including smart watches and smart phones.

IoT in Healthcare Applications

"When a patient is treated, their care is informed not only by their own health data - their medical history, test results, imaging and so on - but also by the health data of thousands of other people." (Deepmind, 2017). This additional information provides essential knowledge

and insight, allowing for increases in both access to care and quality of care, as well as reducing costs (Islam, Kwak, Kabir, Hossain, & Kwak, 2015). The recent rise in IoT-driven healthcare systems and applications has revolutionised this process, providing the necessary platform for innovation across all healthcare domains (Niewolny, 2013). Data collected in this manner may pave the way to an entirely new healthcare paradigm, allowing a focus aimed more at prevention and early intervention than the responsive methods of present day.

Mobile devices are now ubiquitous with most people carrying between one and two devices at any time (Konstantinidis, Bamparapoulos, Billis, & Bamidis, 2015). These devices along with multiple other physiological signal sensor technologies allow for the constant real-time capture of personal biomedical data. Using IoT, this data may then be leveraged for use in a broad spectrum of healthcare applications. Data collected through this approach enables both faster medical intervention during emergencies (Mohammed et al., 2014), as well as offering potential preventative measures in some cases (Hii, Lee, Kwon, & Chung, 2011), whilst simultaneously reducing equipment costs and other resource usage. These applications may be split into two classes: those that look to improve health and those that raise productivity (Manyika et al., 2015). Some of the most prominent uses of IoT within both categories are reviewed in the following subsections.

Remote Patient Monitoring

Remote patient monitoring provides a prime example of the potential of IoT based healthcare applications. A wide range of non-intrusive sensors for monitoring various health parameters have been developed for an array of applications (Islam et al., 2015), allowing for the real-time monitoring and optimisation of patient care (Cognizant, 2016). By capturing such data and securely storing it in the cloud, real-time access of a users physiological data may be accessed by all involved in their care, from the patient and the patients family to medical professionals (Mohammed et al., 2014).

Bringing monitoring to a patient rather than vice-versa provides a multitude of benefits. Without constant monitoring, particularly during stages of rehabilitation and recovery (Hii et al., 2011), patient relapse is shown to be far more common (Fong & Chung, 2013). Pressures

on important medical resources in both equipment and medical professionals are reduced and significantly increased access to care means patients are given the freedom of being able to live much more independently, greatly improving their quality of life (Niewolny, 2013). By applying a diverse range of complex analysis algorithms to collected data in real-time, patients deemed most at risk, such as the elderly and those suffering from chronic conditions, are far less likely to develop any serious complications, as this constant observation often allows diagnosis to be made substantially earlier.

However, it should be noted that benefits of such applications are not reserved solely for the elderly and incapacitated. There is also great scope for such technologies to be utilised by more able-bodied users. Through monitoring a users general wellbeing whilst undertaking daily activities, it is possible for diagnosis of potential health concerns to be made significantly earlier than they otherwise would have been. This allows for a far more preventative approach to healthcare, where early intervention is critical.

Condition Monitoring and Management

In a similar vein to that of remote patient monitoring, IoT-based applications allow for the monitoring and management of chronic health conditions. Whilst conventional methods often involve visiting a hospital or clinic and are therefore costly and time consuming, these applications bring the potential for in-home monitoring of such conditions (Manyika et al., 2015). This comes with the additional benefit of much more frequent health parameter readings over the episodic readings of more traditional methods (for example, when blood is taken). By using wearable devices, or even devices used intermittently, there has already been evidence of reduced healthcare costs and improved health in those suffering from COPD (chronic obstructive pulmonary disease), diabetes and chronic heart failure (Manyika et al., 2015). Various other research looks at monitoring glucose levels (Istepanian, Hu, Philip, & Sungoor, 2011), oxygen saturation levels (Larson et al., 2011), blood pressure (Puustjarvi & Puustjarvi, 2011), electrocardiogram (EKG) (Jara, Zamora-Izquierdo, & Skarmeta, 2013) and body temperature (Jian, Zhanli, & Zhuang, 2012) through use of a combination of non-invasive sensors and smart phones.

Assistive Technologies

The integration of IoT into assistive technologies has proven to be a life changing advancement for the disabled and elderly, enhancing both quality of life and allowing for much improved independent living (H. Lee, 2016). Interconnected devices creating automated homes and other smart environments (G3ict, 2015) are often used in conjunction with brain computer interfaces (BCIs), equipping usually disadvantaged users to better navigate daily life. By utilising connected sensors and cameras, especially within the users home, and coupling these with an accessible and intuitive smart phone interface, users may perform tasks, which would otherwise be impossible or extremely demanding. Other assistive technologies, such as smart wheelchairs (Abhishek, Manjunatha, Sudarshan, & Reddy, 2016) and wheelchair management systems (Islam et al., 2015), have also been developed, using IoT to monitor the status of the user and collect data on their location and surroundings.

Smart home platforms have been adapted to benefit less abled users suffering from a range of conditions. Home automation applications developed for simple control of appliances and other home devices, such as security systems and thermostats combined with a smart phone with compatible screen reader, have proved beneficial for blind and low vision users. Users with poor mobility have found advantageous applications to manage systems physically difficult to reach, such as door locking and lighting. Sensors used for these purposes also have the potential to gather information over a period of time, defining a users typical daily routine and allowing for remote monitoring by care givers.

Hospital Workflows

Hospital workflows present another area set to greatly benefit with advances in both interoperability and digitalisation. The DeepMind Health Streams application, currently being trialled in a number of the UK NHS hospitals for identifying and treating acute kidney injury (AKI) (Deepmind, 2017), gives a prime example of this. Analysing blood test results as soon as they become available and automatically escalating to the relevant medical professional whenever there appears to be a cause for concern ensures warning signs are picked up and promptly acted upon. Urgent alerts sent to a clinicians mobile device also contain addi-

tional important patient data, including personal information, a linear record of all previous work since admission and medical history (Connell et al., 2017) allowing for an immediate diagnosis and prioritised treatment to be made.

Current workflows frequently involve communication between multiple technicians, nurses and doctors using outdated, often paper based methods and requiring access to a desktop computer to review results. Alleviating these delays allows for both better coordination of patient care and ensures those patients most at risk receive the care they need in adequate time. Applications like this may be developed to cover a wide range of other conditions, as well as being used to simplify everyday events, such as shift changes and patient handovers.

Issues and Challenges in IoT-enabled Healthcare Applications

Despite the many notable benefits of the healthcare applications discussed, many issues and challenges are also present, especially regarding the security and privacy of often sensitive gathered medical data and interference with other medical devices. These issues are often made more challenging by the diversity of IoT components (Abouzakhar, Jones, & Angelopouloui, 2017). In particular, IoT based healthcare applications rely on interconnected devices exchanging potentially sensitive information (Abouzakhar et al., 2017) and, as such, are prone to a range of different security and privacy attacks (Al-mawee, 2012). Multiple questions have been raised regarding gathered personal medical data, specifically about ownership of the data, who has access rights to it and where it is stored, questions that must be answered to ensure the integrity of the data and the privacy of the individual it belongs to. These concerns are often magnified in the case of commercial organisations being involved (Deepmind, 2017), as more and more private companies gain access to often intensely personal health data.

Healthcare Applications: a Leap in the Future

The real vision for the future is that the discussed various smaller applications will converge to form a whole. Emerging applications have the potential to transform a wide range of health-care therapies and enable remote surgery. Ingestibles and injectables (smart pills and nanobots) have the potential eventually to replace many surgeries with less invasive

procedures that could offer faster recovery, reduced risk of complications, and lower cost. IoT solutions can have a real impact on one of the most vexing problems in healthcare today: human behaviour. Using IoT systems to convince healthy people to change their living habits and to help sick patients adhere to doctors prescriptions would be a true breakthrough. As technology evolves, costs will continue to fall, enabling broader adoption and use by a wider range of patients (Manyika et al., 2015).

Energy Applications

Smart Grid – Energy Consumption Monitoring and Management

Electrical grids continuously grow to meet the increasing power demands, hence, monitoring and controlling such grids become complicated and far from efficient. Moreover, the growing shift in Europe and USA to integrate more distributed and renewable energy, originated from wind, solar and biomass, introduces unpredictability, variability and intermittent power generation. All these undesirable operational scenarios combined with aging infrastructure cause electrical grids to constantly operate at their maximum limits. This in turn reduces their life expectancy, and causes high power disturbances (Gungor et al., 2011). Electrical grids operate under unpredicted circumstances, as they have to cope with rapid changes in seasonal loads and variation of weather conditions which has increased due to climate change. These operational circumstances raise challenges in terms of reliability and stability of the grid. For example, dealing with sudden regional energy demand for cooling due unpredicted heat can cause energy congestions, rapid change of power patterns, and inefficacy in emergency loading, etc. (Gungor et al., 2011). To cope with these challenges, a Supervisory Control And Data Acquisition (SCADA) (Sallam & Malik, 2011) is a widely used situational awareness system to efficiently manage energy in grids. However, SCADA offers a steady state analysis, which fails to fully monitor and predict the dynamically changing power system. Due to the recent advances in IoT in terms of communication and its capability in capturing huge amount of life sensor data, IoT can increase situation awareness through monitoring of the gird status, which can lead to balancing energy load on transmission lines, controlling technological derangements, reducing power disturbances, and fine-tuning emergency and

protective automation (Sallam & Malik, 2011).

As an example, Figure 1 shows a community scale energy supply from renewable energy, where energy is produced and controlled locally. In this scenario, energy delivery at community level utilises IoT and data mining to optimise energy distribution and energy allowance for each property based on occupants usage, tenants habits, property size and house energy profile. Principles of Multi-Agent System can also be applied here by allowing autonomous agents to act on behalf of the households. Agents can monitor and perform data mining to establish adaptive models for the household they represent, as well as use techniques from Game Theory to efficiently share energy when a household's energy demand is high at specific times (e.g., to perform household activities, such as running a washing machine, ironing, lawn cutting, etc.). Intelligent agents play important role in sustaining the community self-regulation by monitoring users, activates or situations to prevent selfish behaviour, when a household asks for more energy than is actually needed or consumes more energy without being considerate to the entire community. The job of these intelligent agents is to acquire just enough resources for their household. If household agents are energy literate and decide to save energy from their daily allowance, then they can feed this saved energy back to the community grid, or sell it to a house with higher energy demand. Hence, virtual money can be generated and spent in energy form or other forms of community related activities. This management system uses storage facilities at household level to store and distribute energy efficiently without the need for a central storage system that is expensive to maintain and prone to single point of failure problem.

Despite many benefits, such IoT-based smart grid also raises concerns that were not present in the classical power grid. Smart grid can be subject to spoofing attacks. This is when an identity of occupant's smart meter is stolen and used to pay for the attacker energy consumption. Data exchanged between a smart meter and energy supplier is subject to tampering attacks, where an attacker sends wrong information about its tariff or by pretending that most energy has been consumed during off peak periods. Such attacks increase household consumptions and overload the grid. Since smart grids are connectivity-enabled, they are at risk of cyber-attacks, but unlike software attacks that normally damage users software or data, cyber-attacks against smart grids can also cause physical damage to

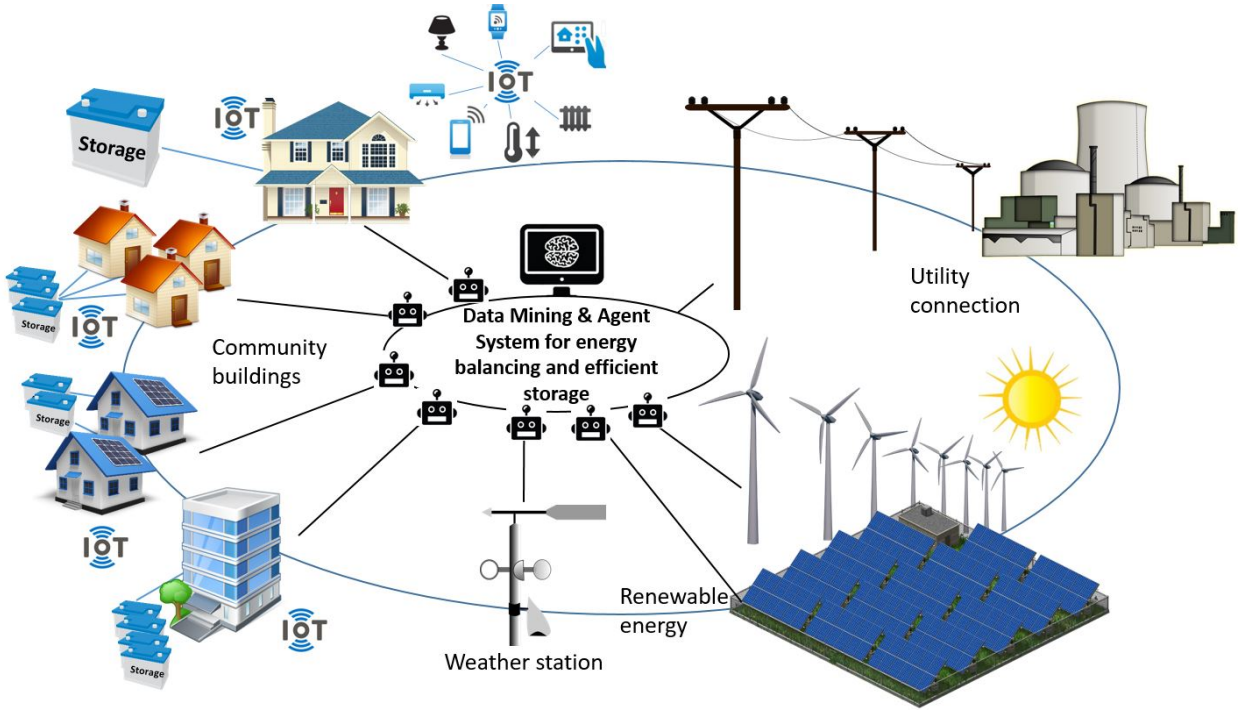

Figure 1: Using IoT and data mining, energy generated from renewable sources can be distributed fairly; energy can be stored and shared through utilisation of private energy storage systems (batteries) in houses

users, as grids are connected to transformers, circuit breakers, cables and smart meters. The second challenge concerns with interoperability between IoT devices and legacy devices that are not compatible with standard TCP/IP (e.g., Zigbee v1 and HART) that allow IP-based devices to communicate (E.-K. Lee, 2016).

In the UK, the interoperability issue is extended by adoption of modern smart meters for gas and electricity. In an attempt to reduce energy consumption from buildings, the UK government has introduced legislation that requires all energy suppliers to install smart meters in every home in England, Wales and Scotland by 2020 (Department for Business, prices, & bills, 2013), so that meter reads could be sent to suppliers for more accurate energy billing. Such smart meters provide consumers with more control of their energy use, allowing to adopt energy efficiency measures that can help the users to save money. The interoperability issue resides in the fact that smart meters use various communication mediums and protocols to communicate only with their suppliers (Erlinghagen, Lichtensteiger, & Markard,

2015). Hence, the absence of a standard communication platform for smart meters makes it difficult to maintain a comprehensive view required by smart grids to predict operational problems. Consequently, common IoT solutions can also offer a cheap and rapid solution to the interoperability issues between meters and smart grids (Bekara, 2014; Erlinghagen et al., 2015).

Smart Cities Applications

IoT-based Smart Buildings

40% of the total energy consumption is related to residential and commercial buildings. This has led many governments to introduce new policies in order to improve energy efficiency in buildings. For example, the directive 2010/31/EU of the European Parliament and of the Council was first introduced in May 2010 on the energy performance of buildings (Commission, 2010). This directive proposed measures to improve the performance of lighting, heating and ventilation systems in an attempt to reduce the carbon footprint at a global scale.

With the increasing demand for occupants' thermal comfort in the form of heating and cooling, growth in population, expansion of building sector and increase of time spent inside buildings, building energy consumption will continue to increase by 0.5% - 1.5% per annum in the UK and Europe (Prez-Lombard, Ortiz, & Pout, 2008) and by 7.80% in China (Zhang, He, Tang, & Wei, 2015). Personal initiatives together with technological solutions can be used to decrease energy consumption in buildings. Various measurements are already used to increase energy efficiency in buildings:

1. **The use of renewable energy from wind and solar power via the use of solar panels and wind turbines**. These technologies can be deployed at a large-scale in the form of wind farms that generate energy to cover entire cities, or at a small-scale, when households install standalone renewable energy systems that can cover some of their energy needs (M. A. Ahmed, Kang, & Kim, 2015). This is normally accompanied with energy storage batteries and energy management systems. These systems are designed to allow customers to feedback unused energy into the grid or store it to

be used when needed (Figure 1). However, managing and controlling the operation of small-scale energy systems raise various challenges (Bouhafs, Mackay, & Merabti, 2012); to achieve reliable and cost efficient operations in small-scale energy systems, IoT and communication infrastructure are necessary to make smart grids a reality. For example, it is important to first analyse how energy is currently consumed in buildings. If the use of renewable energy is desirable, then IoT sensor data regarding weather condition, including average wind speeds and sunlight peaks, would enable determining the size and number of wind turbines, number of solar panels and size of battery required to store energy for a house (M. A. Ahmed et al., 2015).

2. **Retrofitting of buildings**. This involves the retrofit of old inefficient buildings that experience heat loss due to air leakage, aging of windows and wall degradation. Retrofitting of such buildings includes putting external and internal insulation panels, fixing double glazed windows and sealing any damage to floor, roof or walls to prevent any air leaks. To identify the best retrofit package for any house, various parameter values need to be identified to allow the highest possible performance of a building in terms of energy consumption, cost and thermal comfort, as well as how such packages would cope with various weather conditions. These parameters can include the depth and type of insulation panels, for example. Rapid deployment of sensor networks is needed to collect the necessary data that help to identify these parameters. Building simulation can then enable comparison of various energy conservation measures in a form of theoretical extensions or refinements to the input model to reduce energy consumption in a building, as well as assess various performance optimisation measures during the operational stage (Basurra & Jankovic, 2016).

All the aforementioned approaches that aim to reduce energy consumption in houses share a common challenge that makes energy efficiency hard to achieve, and this is largely due to behaviour of the occupants and their poor energy awareness and training, especially, in the context of increased demands by occupants on heating, ventilation and air conditioning to provide the highest thermal comfort. The findings in (Zhao, McCoy, Du, Agee, & Lu, 2017) suggest that even owners of new buildings, which are built in accordance with the

energy efficiency requirements and standards and are usually equipped with energy efficient appliances and modern thermostatic controls for heating and cooling, are still responsible for high energy wastage. Hence, ongoing monitoring of home appliances, users' locations, motion and habits using a large number of sensors may help predict human behaviour in the house resulting in better energy management via predictive control and providing humans with actionable messages to better inform individuals about their consumptions and behaviour (Abdallah, Basurra, & Gaber, 2017).

Figure 2 depicts a typical household with a family of five members, where various IoT sensors are placed around the house to measure thermal comfort of the individuals. The term thermal comfort defines a persons state of mind of whether they feel too hot or too cold. If thermal comfort is quantified in real time, the thermostat system can be automatically adjusted to constantly produce the heating/cooling that suits most occupants in a building, while also focusing on effective use of energy to reduce electricity waste. To achieve this, various IoT sensors can be used to measure the external temperature, internal temperatures in separate rooms, as well as motion to identify occupants location (i.e., which rooms are occupied, and whether the house is empty or not). Thermal cameras can be fitted to measure the thickness of the clothing occupants are wearing while indoors. Furthermore, body temperature and levels of activity (e.g., sleeping, sitting or moving around) can be measured using personal activity trackers such as Fitbit and internal tracking systems. These measurements can be used to calculate the metabolic heat generated by occupants, which is an important factor of the thermal comfort.

Generating models based on a combination of real IoT data and predictive patterns identified using data mining techniques that represent evolution of the parameters affecting energy consumption, can help to develop intelligent building management systems. For example, research in (Hagras, Callaghan, Colley, & Clarke, 2003), demonstrates a smart system capable of managing the main comfort services provided in the context of a smart building, i.e., HVAC and lighting, while user preferences concerning comfort conditions are established according to the occupants' locations. In (Moreno, beda, Skarmeta, & Zamora, 2014), the authors studied the main parameters affecting energy consumption of buildings considering different contexts. Such analysis permits to propose an optimum prediction

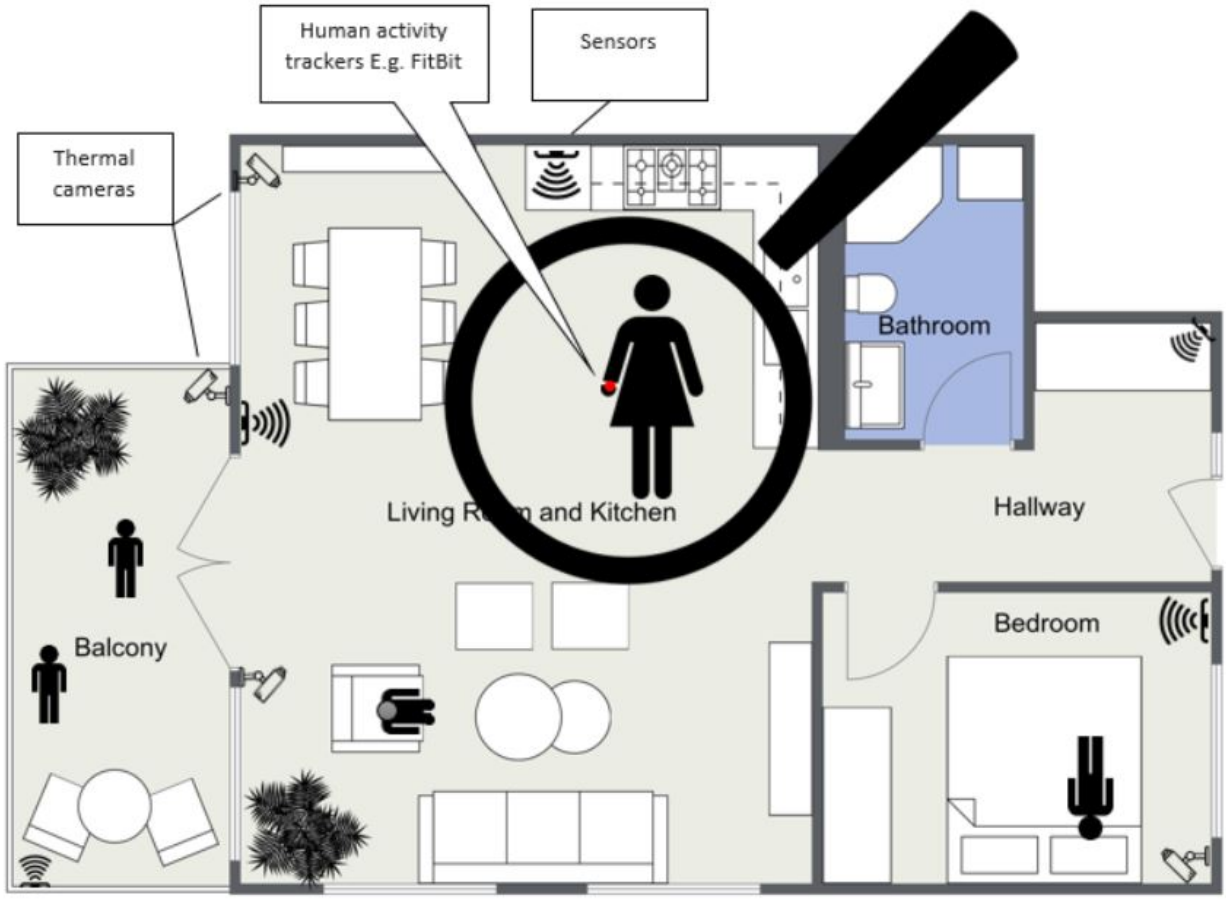

Figure 2: Using IoT to predict human behaviour in buildings for optimal energy management control leading to energy savings in buildings

concerning the daily energy consumption in buildings by integrating the most relevant input data into such models. Once energy usage profiles have been extracted, actions can be designed to save energy, by proposing strategies to adjust operation times and configuration of the involved appliances or devices, selecting the optimal distribution of energy to maximise the use of alternative renewable energies, etc.

IoT for Smart Traffic

Traffic congestion is a major problem in many developing cities. According to TomTom traffic index (*TomTom Traffic Index 2017*, 2017) – which is a global leader in navigation, traffic and map production – the major cities, such as Bangkok, Moscow, Jakarta and London,

experienced an increase in traffic by an average of 5% in the year 2016-2017. One of the contributors to congestions is the behaviour of people in cities and their way of life. According to (Isa, Yusoff, & Mohamed, 2014), traffic congestions are caused by people travelling to and from work at the same time each day, which introduces delays, impacts motorists productivity (as a result of arriving late), and consequently, affects the city economic health. Traffic congestions also increase fuel consumption due to vehicle frequent acceleration and breaking, resulting in increased air pollution and more repairs due to rapid wear and tear. Moreover, congested traffic can increase the response time of emergency vehicles, and may even increase their chances of being involved in an accident (Nellore & Hancke, 2016).

In addition to human behaviour, congestions are caused by the increasing population of vehicles and poor road management (Bretzke, 2013). Constructing alternative routes in big cities to ease congestion is not possible due to the following (Zavitsas, Kaparias, & Bell, 2010): 1) planning restriction in constructing newer roads on green belt, existing infrastructure, historical buildings; 2) lack of funding; and 3) that the land value is extremely high in big cities, hence, local authorities would rather use these lands to build new houses to accommodate the rapid population growth, rather than using them for constructing roads. Therefore, many major cities have no option but to work with the existing roads.

One solution to improving management of road capacities is to use IoT technology via installation of fixed road sensors and vehicle to vehicle sensors to obtain live traffic data (e.g., wireless sensor networks (WSN), Radio Frequency Identification (RFID), vehicular ad hoc network (VANET) and GPS data from drivers and passengers mobile phones). IoT can help optimising live traffic using load balancing mechanisms to reduce travel time and ensure the steady traffic flow to prevent the frequent acceleration and breaking in vehicles (Jabbarpour, Nabaei, & Zarrabi, 2016). In addition, historical traffic data obtained from IoT devices can be used to predict the locations of traffic congestions and their density, so that vehicles nearby can be rerouted towards less congested roads while maintaining reasonable traffic delays (Masek et al., 2016). In (Mittal & Bhandari, 2013), a green wave system is proposed that uses IoT devices to provide traffic clearance to emergency vehicles. Green wave works by turning all red lights into green along the path of the emergency vehicle. The system proposed by (Sundar, Hebbar, & Golla, 2015) uses IoT devices at traffic lights to track

stolen vehicles and slow the flow down by turning traffic lights along their mobility route into red. (Arfat et al., 2017) propose to use IoT, social media, and other data sources for a more accurate awareness of the road conditions along the route for all the users, including vehicles, bicycles, scooters and pedestrians crossing streets.

Optimising traffic can also lead to challenges. Providing priority measures for emergency vehicles and public transport can make the situation drastically worse to other vehicles. Linking all traffic lights into the cloud, which can be configured and controlled in real time, introduces vulnerability of the system against cyber attacks that could direct traffic with the purpose of harming the users, or shut down the traffic system for the entire city (Jin, Zhang, Walton, Jiang, & Singh, 2013).

Challenges/Issues with Current Solutions

From the above overview, it is clear that while applications of IoT in healthcare, energy management and smart cities can provide many benefits, realisation of such applications is faced by many challenges. Issues associated with privacy, communication and edge computing in particular are discussed in this section, along with some of the current solutions to these issues.

Privacy Issues in IoT

IoT is a promising paradigm, with expected applications in different domains, as discussed in the previous section. However, this new paradigm faces a new kind of security and privacy challenges in addition to those inherited from the traditional Internet and have not yet been properly addressed so far. IoT systems spanning across many vertical domains, with applications in areas such as remote patient monitoring, smart meters, waste management, smart cities, intelligent surveillance, and remote and industrial controllers. Most of these applications demand the privacy preservation of the end-users data, which is something intuitive, with the current spread of privacy violations (Martin & Palmatier, 2017). Privacy is an essential requirement for users' personal data that may contain their preferences, habits, living patterns, movements, and social relationships. Some research that has been carried

out to address the issues associated with user privacy in the context of IoT applications is outlined below.

In (Rutledge & Swire, 2014), the authors provide an abstract framework for analysing security and privacy concerns for IoT devices.This framework can aid in providing guidance for where security and privacy analysis may need to be supplemented with other research fields. In this framework, the authors examined IoT devices that accept inputs for security concerns and IoT devices that produce outputs for privacy concerns. Using a five-stage general policy framework for evaluating privacy and security, concerns raised from IoT devices were differentiated from concerns of other contexts such as Big Data, cloud computing, robotics, and ubiquitous computing. Understanding these contexts is essential to resolve or mitigate various security and privacy concerns for deployed systems. The research work in (Evans & Eyers, 2012) suggested the utilisation of data tagging to enhance data privacy in IoT devices. Various techniques based on information flow control were presented to assign privacy properties as tags to the data representing network events. These tags can enable the IoT system to preserve privacy of individuals and to better understand data flows. However, this solution may not fit all types of IoT devices, since the utilisation of tags in a resource-constrained IoT devices is an expensive solution in terms of storage, processing and communication. The inserted tags can be too large in comparison with originally collected data, which can generate an additional overhead in processing and handling such data. Consequently, this solution may not be appropriate for the IoT domain.

In (Appavoo & Chang, 2016), a lightweight functional encryption for privacy-preservation based on trust model was proposed. The proposed approach aims to minimise privacy loss and secondary uses in the presence of untrusted parties. A uniformisation solution was projected, which utilises device aliasing to conceal the identity of the sensing-source and a precomputation initialisation vector that is provided to extract trigger information only to the relevant services of the appropriate trusted parties. Untrusted parties would not gain any additional information about the end-users when the published sensor readings and the trigger information are not known. Through the use of aliases, the service provider cannot identify the sensor from which it receives data, and through the uniformisation scheme, it cannot deduce whether the trigger has been activated or not. Finally, the authors implemented the

proposed scheme to demonstrate feasibility of the proposed scheme on resource-constrained IoT devices.

In (Otgonbayar & Dahal, 2016), the authors proposed an anonymisation algorithm for publishing data streams generated from various IoT devices using the k-anonymity privacy model. The proposed scheme examines the similarity of the input tuples when performing a clustering process, then uses a time-based sliding window technique to anonymise tuples with similar description into a single cluster under a specific delay constraint. Later, it has to check if the cluster has enough tuples satisfying the k-anonymity requirement. This preliminary step facilitates rapid construction of clusters by localising tuples and supports the cluster merging process. Various clusters can be merged based on their similarity that is being measured using Jaccards similarity coefficients. Additionally, the proposed scheme implements resolutions to anonymise tuples with missing values by utilising their representative values. Finally, the authors presented experimental results on a real dataset, which indicate that the proposed scheme can successfully publish data streams with less information loss and attains faster execution time when compared to conventional anonymisation approaches.

The authors in (Tso & Hossain, 2017) studied the use of public key cryptographic techniques to preclude data leakage in healthcare systems. The expensive nature of these techniques makes them impractical for a real implementation in the domain of web-enabled medical IoT systems. Their research proposed a practical approach based on a secure multi-party framework fairplay to inhibit data disclosure from insider attacks. Their proposed solution enables software developers to easily implement security protocols in a distributed IoT systems with multiple participants. Within this approach, the initial setup demands each IoT node to store only one secret key before communicating with external data servers, which is appropriate for a small memory size of IoT nodes. This approach is an improved version of the previously proposed scheme in (J. W. Yi Xun & Nait-Abdesselam, 2013), which demands each sensor node to store three secret keys before connecting with three data servers.

The research work in (Prez & Gigante., 2017) studied the need for ensuring privacy of sensitive and private information in healthcare and automation systems. A solution was

presented to deal with these issues and to promote acceptance of IoT services to the end-users. Attribute-based cryptography techniques and AES symmetric encryption scheme were utilised to construct a novel architecture combining flexibility and expressiveness of the first techniques with the efficiency of the second scheme. The usage of this architecture facilitates the execution of secure data exchange and preservation of privacy of participating parties, since it satisfies security and privacy requirements during the full lifecycle of sensitive information. Within this solution, data sources delegate the execution of attribute-based cryptography operations to a trusted proxy, the IoT cloud platform is responsible for managing encryption keys and handling of protected data.

A framework for modelling and assessing security of IoT systems was recently proposed in (Ge & Kim., 2017), which addresses new security issues emerging from this new paradigm. The framework incorporates five different phases: 1) data processing, 2) security model generation, 3) security visualisation, 4) security analysis, and 5) model updates. The authors presented a prototype for three nodes in their framework, an IoT generator, a security model generator and a security evaluator. The IoT generator assembles an IoT network from different subnets based on the information gathered from network reachability options and node vulnerability; the security model generator forms an extended hierarchical attack representation model (Hong & Kim, 2016) based on the previously assembled IoT network; the security evaluator inspects security of the IoT network using various security metrics, which later can be used as an input to the security analysis phase. The presented framework can be utilised to discover potential attack scenarios in the IoT network, to measure the overall security of IoT network using purpose-specific security metrics, and to evaluate the impact of various defense procedures. The authors concluded their paper by evaluating the proposed framework in three different scenarios to show the capabilities of their framework in predicting and mapping different attack paths and the suggested strategies to mitigate their impacts.

A privacy protection mechanism for computerised numerical control information in IoT was proposed in (Li & Li, 2017) that incorporates a lightweight authentication method for wireless sensor networks and an Internet data privacy protection strategy based on organisational characteristics of numerical control information in IoT. The numerical control machine

tool can be envisioned as a complex network system that consolidates sensor networks, Internet, and mobile communication networks. The lightweight authentication protocol consists of five parts, namely, system setup phase, sensor node registration phase, user node registration phase, login phase, and authentication and session key agreement phases. The Internet data privacy protection strategy involves central controllers, computerised numerical control machines, local monitoring centres, cloud server, and the end-user. The authors have performed a series of analysis to demonstrate the efficiency and safety of the proposed solution. The results presented illustrate that utilising the proposed solution in the cloud computing setting can guarantee security of control information and privacy of the numerical control machine tool. However, employing double encryption in the Internet data privacy protection strategy increases the length of the key, which poses certain restrictions on data transmission and storage.

The authors in (Elmisery & Aborizka, 2017) proposed a new framework for collective privacy protection, which utilises the personal gateways at the end-users side to act as intermediate nodes between the IoT devices and cloud services. A lightweight middleware is envisioned to be hosted on these intermediate nodes for an efficient aggregation of end-users data while maintaining privacy and confidentiality of their collective profiles. The proposed middleware executes a two-stage concealment process that utilises the hierarchical nature of the IoT devices. The concealment process utilises a hierarchical topology for data collection, where different IoT devices and their corresponding intermediate nodes are organised into a coalition for aggregating their data in specific profiles. This could help to unburden the constrained IoT devices by performing computationally intensive privacy-preserving operations. The end-users are empowered with a tool to control privacy of their health data by enabling them to release their data only in a concealed form. Further processing of the cloud service continues over the concealed version of the data by applying customized secure multiparty computation protocols. Additionally, the cloud service uses privacy policies for specifying their data usage practices. The end-users can describe their privacy constraints in a dynamically updateable fashion using privacy preferences specification language. Trust-based concealment mechanism was also utilised in the course of producing different copies of data based on the various trust levels with different cloud services. The proposed solution

was integrated into a scenario related to preserving privacy of patients health data when utilised by a cloud healthcare recommender service to generate health insights. The authors tested the performance of the proposed solution on a real dataset to measure the overall accuracy of the results based on various parameters of the two-stage concealment process. The experimental and analytical results show that privacy increases under the proposed solution without hampering the accuracy of the results. The proposed approach presents a straightforward solution with accurate results, which are beneficial to both the end-users and service providers.

Finally, the authors in (Elmisery & Botvich, 2016) presented an attempt to develop an innovative approach for handling privacy in the current service oriented model. The holistic privacy framework was developed in complying with the OECD privacy principle. The proposed framework was implemented as a middleware that was entitled EMCP "enhanced middleware for collaborative privacy". The authors presented a novel concealment process, which provides a complete privacy control to the end-users when sharing their data with external third-party services. The proposed framework permits a fine-grained enforcement of privacy policies by allowing the end-users to ensure extracted data for specific queries do not violate their privacy. This is automatically done by checking whether there is an APPEL preference corresponding to any given P3P policy. The Fog nodes were utilised to aggregate multiple end-users' data obtained from the underlying IoT devices, to encapsulate them in a group profile, and then to send to an external recommender service. Performance of the proposed framework was measured by the authors on a case study for a healthcare recommender service using a real dataset. The presented results depict that privacy increases under the proposed framework based on off-the-shelf recommendation techniques without hampering the accuracy of results.

Networking Issues in IoT

This section will address the issue of connectivity and how data are delivered from the source to their final destination in the context of IoT domain. IoT connectivity is a key element in the entire IoT ecosystem (Bröring et al., 2017) and it is responsible to link up the physical and digital world together.

Providing the fact that different devices with different capabilities are responsible for creating data in different ways, sensing for example, this heterogeneity requires different connectivity capabilities in order to achieve the interoperability factor which is a major component of any distributed system, such as smart city systems. There are many communication technologies that are in use to meet these requirements. Selecting which communication technologies (wired and wireless) are appropriate depends on many factors such as interoperability, availability, cost, reliability, scalability, coverage, power consumption, data rate, and range (Al-Sarawi, Anbar, Alieyan, & Alzubaidi, 2017). A combination of wired and wireless connection is the way forward for many scenarios and configurations. Peripheral devices are normally connected via wireless technologies that are used for short range data communications, then collected data can be aggregated and delivered by long range wireless communications, such as Cellular and LPWAN networks (Bardyn, Melly, Seller, & Sornin, 2016; Margelis, Piechocki, Kaleshi, & Thomas, 2015). Creating a scalable and robust communication system for loT is essential for its success (Gubbi, Buyya, Marusic, & Palaniswami, 2013; Marchiori, 2017).

The diversity of applications that demand use of the IoT communication infrastructure, make the selection of the appropriate wireless technology option a big challenge (Battle & Gaster, 2017). Across the entire IoT stack, there are several communication links to be established and maintained. However, in order to establish a link, many wireless and wired technologies can be adopted, depending on the IoT network segments requirements. For example, there are three major segments (Smart Things Devices, Last Mile Connectivity Technologies and Backhaul Links Technologies) in IoT connectivity systems available to IoT developers to select when developing IoT solutions, as shown in Figure 3.

At segment one, where things (these can be anything from machine to appliance or vehicle) can be connected to the gateway in many different ways using long or short range, licensed or unlicensed wireless technologies. Gateway devices can be smart phones, stationary WiFi access points (AP), LPWAN gateway (e.g., LoRa Gateway) or even cellular networks. Low-power communication between the IoT device and gateway device using several standards such as Bluetooth, Zigbee, IEEE 802.15.4, or IEEE 802.11 (WiFi) achieves a maximum communication range typically less than 100 meters. However, wide area networks

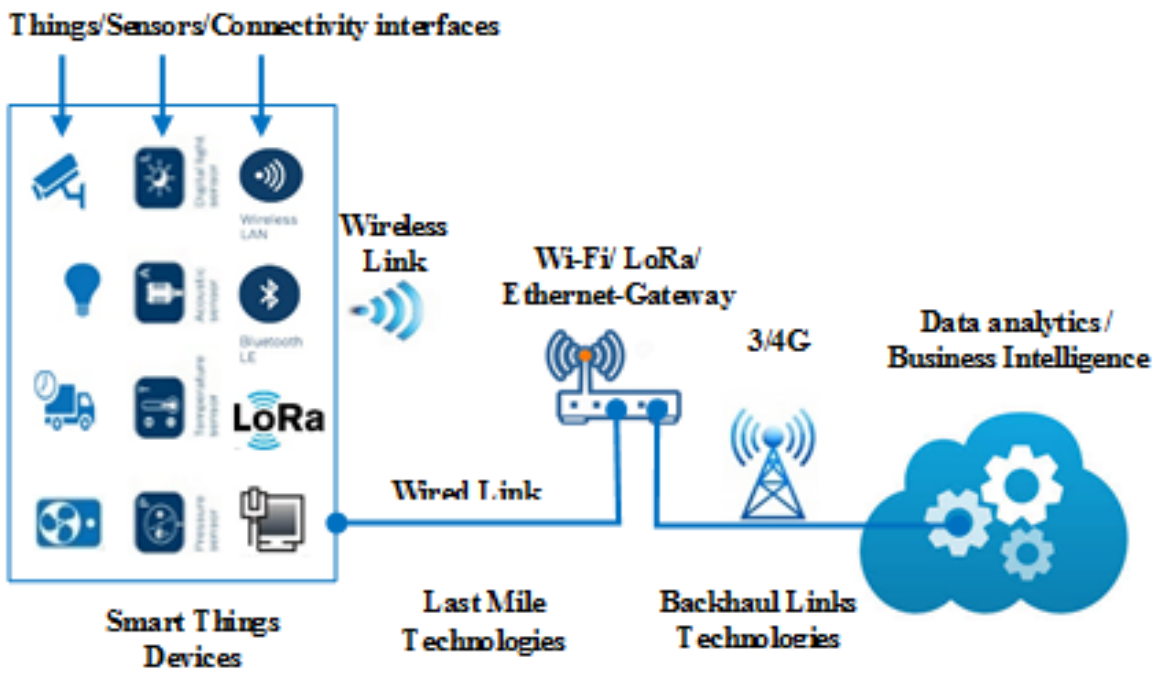

Figure 3: IoT Networks Connectivity Options

tend to be less complex than mesh networks since the endpoints can be connected directly to a gateway or a base station, rather than relying on a relay system to transmit messages 'last mile' connectivity for the IoT connected objects as seen in segment two. At segment three, which connects the gateway to the backend systems, many backhaul connectivity options can be used from wireless cellular networks to high speed wired connectivity, such as fibre optic links.

Unlicensed Short Range Communication Protocols

These protocols are mainly used to connect sensors with IoT boards (sensor nodes) or connecting these sensor nodes directly to the gateway; it all depends on the given scenario and topology. However, these wireless technologies have low power and can be utilised for short range connectivity scenarios. The choice of a particular technology for a particular application can be made by examining required data rates, power consumption, range and sensor

nodes communication capability, according to a recent whitepaper by Ericsson [1]. The majority of these communication technologies use licensed free ISM bands with different radio frequency rates.

One of the most common short-range wireless technologies with less power consumption compared to others including the basic Bluetooth is Bluetooth Low Energy (BLE). Bluetooth and BLE are used for different purposes (Noreen, Bounceur, & Clavier, 2017). Conventional Bluetooth can handle almost all the variety of data, but it consumes more power and cost. BLE, on the other hand, is used for low data rate applications, and can, therefore, have longer battery life time. Like classic Bluetooth, BLE also utilises licensed free ISM band and offers 40 different channels.

Another solution is the IEEE standard 802.15.4 commonly known as ZigBee, which is the most popular choice in Low Rate Wireless Personal Area Networks (LR-WPAN) and WSNs. For example, ZigBee has been proposed as the communication solution for a telemonitoring health care solution in (Corchado, Bajo, Tapia, & Abraham, 2010). The IEEE standard 802.15.4 has only defined the characteristics of physical (PHY) layer and Medium Access Control (MAC) layer. Although Bluetooth and ZigBee are low power and low complexity wireless sensor technologies, they have some limitations, such as low data rate, short range, and less penetration across obstacles (Noreen et al., 2017).

Table 1 shows a comparison of some basic features related to the above mentioned technologies used in IoT connectivity at segment one. As can be seen from Table 1, Bluetooth BLE in particular supports much reduced power consumption at a comparable range, but at a significantly reduced maximum data rate of 1 Mbps, which is well-suited to the most of the IoT solutions. However, no single technology or solution is perfectly suited to all the different potential IoT applications. In many cases, Zigbee can be a better choice providing the underlying scenario. Ideally, the optimal short range IoT connectivity option should have the right combination of power, coverage, data rate and cost.

[1] Cellular networks for massive IoT. https://www.ericsson.com/res/docs/whitepapers/wp_iot.pdf

Table 1: Comparison of IoT Sort Range Wireless Communication Protocols

	Bluetooth	**ZigBee**	**Wi-Fi**
Max. end-devices	255 (2 Billion in BLE)	more than 64000	Depends on number of IP addresses
Peak Current Consumption	10 mA	30 mA	100 mA
Range	10 m	100 m	100 m
Data Rate	1 Mbps up to 250	kbps	11 Mbps and 54 Mbps
Relative Cost	Low	Low	Medium
Topology	Star and Mesh	Mesh only	Star and point to point Transmission
Technique	FHSS (Frequency Hopping Spread Spectrum)	DSSS (Direct Spread Spectrum Sequence)	OFDM (Orthogonal Frequency Division Multiplexing)

Unlicensed Long Range Communication Protocols

The wireless network industry is gradually changing their interest from the traditional cellular networks to Low Power Wide Area Networks (LPWAN). There are many LPWAN technologies on the market today (SigFox, LoRa, Weightless-W, WiSUN, etc.), but the most common ones are LoRa and SigFox. They successfully propose wide area connectivity from a few to tens of kilometres for low data rate, low power and low throughput applications (Noreen et al., 2017). They facilitate a large scale of connectivity, especially in large scale developments such as smart city applications.

LoRa and SigFox are both long range and provide the ability to connect devices that are very low-power for longer geographic distance. They operate at different radio frequencies across the world. For instance, in Europe, they operate between 867-869 MHz, while in the US – between 902-928 MHz. LoRa, the physical layer of LoRaWAN (Augustin, Yi, Clausen, & Townsley, 2016) [2], supports very low power devices (aiming for longer battery life) offering

[2] A technical overview of LoRa R and LoRaWANTM, Published by the LoRa alliance, accessed 2017-03-27.

a maximum date rate of just 27 kbps in exchange for long range, a 2-5 km range in urban areas and in excess of 10 km in suburban areas (Battle & Gaster, 2017).

Neither LoRa nor SigFox has good indoor signal penetration, but are well suited to IoT sensors nodes in outdoor scenarios. A key feature of these standards is that they are low-bandwidth, often each message is limited to 100 or so bytes, and for SigFox, it is even less. This suits IoT data requirements. Each standard has its drawbacks, but SigFox requires more expensive chipsets in its gateways, which is not the case for LoRa. When selecting hardware, it is important to choose the correct frequency for the required region.

Licensed Long Range Wireless Communication Protocols

To achieve longer range communication in the IoT domain, traditional cellular networks are a visible option to many mobile operators due to the well-established infrastructure and coverage. However, this solution is relatively expensive and power hungry as current cellular networks were not designed to support IoT devices (Flore, 2016). Emerging network standards such as 3GPP aimed at optimised cellular networks for IoT devices are in various development stages (e.g., LTE-M, NB-IoT, EC-GSM-IoT, etc.), however, these are not yet widely deployed (Bardyn et al., 2016).

The NB-IoT in particular is designed to meet the IoT device requirements related with low cost, low power consumption and long battery life-time. This initiative was taken by 3GPP after attempting to modify the current cellular infrastructure to meet the IoT requirements (Marchiori, 2017). However, it did not succeeded due to complexity and cost associated with the cellular networks, as these networks are not designed to meet the IoT applications requirements in their initial creation.

After looking at various IoT connectivity options, the essential question that the community should ask is that, which option should be adopted? The answer really depends on many factors, as mentioned earlier, such as cost, scalability, power consumption, availability, interoperability, data rate requirements. LPWA IoT networks in the unlicensed bands are limited in efficiency due to the fact that current duty cycle regulations limit the downlink capacity, and hence the network to perform, e.g., power control. On the other hand, LPWAN

[Online]. Available: https://www.lora-alliance.org/portals/0/documents/whitepapers/LoRaWAN101.pdf

approach and technology is the way forward for many parties due to the fact that it provides control over the infrastructure such as LoRa. However, if the national coverage is required and important, then it probably needs to be LTE-M or NB-IoT in the future (Dama, Sathya, Kuchi, & Pasca, 2017). This recommendation is based on the fact that the Sigfox option leaves too many aspects outside of the developer control with no back-up plan.

Perhaps the most controversial part of the deployment cost for cellular networks is the data plan and the cost to transfer data from the IoT device to the cloud. This cost is not applied to LPWAN networks. Frequently, analysis of the market stops with the modem cost, but hardware is not the only thing that should be considered. The next element of deployment cost is the provisioning cost. Technologies like GSM and LTE provide an enormous range that aims to provide higher and higher network throughput at the cost of power consumption. On the other hand, LPWAN, such as LoRa, provides lower throughput at a much lesser level of power consumption compared to its counterpart in licensed bands (Saravanan, Das, & Iyer, 2017).

Current cellular networks require more planning and optimisation when it comes to handling IoT traffic, as it requires complex protocols to support hand-off from one IoT gateway to another where things are mobile (Ozyilmaz & Yurdakul, 2017). By comparison, LoRa networks can be "chaotic" with minimal planning required. Furthermore, LPWAN networks do not have to support anything like a phone-call or guarantees about streamed data; messages are small and intermittent.

Many players tend to put the above solutions in competition to LPWAN operated in unlicensed bands. In fact, the operation in licensed bands is more valuable for selected professional services, while unlicensed bands provide generally better coverage, lower power and lower cost, at the expense of a lower QoS and no guaranteed latency. Most telecommunication operators understood this complementarity and are taking advantage of the earlier time to market through existing LPWAN solutions in unlicensed bands (Bardyn et al., 2016). Table 2 summarises the differences among the most common licensed and unlicensed, low power and long range wireless communication protocols. Finally, the authors strongly argue that low-power wide-area (LPWA) technologies like LoRa and cellular-based (NB-IoT) will be the great enablers for mass deployment of low-power end-devices.

Table 2: Wireless Communication Protocols

Characteristics	**SigFox**	**LoRa**	**NB-IoT**
Standards	SigFox	IEEE 802.15.4g	3GPP, UMTS/HSPA(3G)/ LTE (4G)
Frequency bands	868MHz (EU) 902 MHz (USA)	ISM band 868 MHz and 915 MHz	Common Cellular bands
Power	10 mW-100 mW	+20 dBm at 100 mW constant RF output	High Power Consumption
Data rate	100 bps (UL), 600 bps (DL)	290 bps-50 Kbps (DL/UL)	DL:234.7 kbps; UL:204.8 kbps
Range	10km (Urban) 50km (Rural)	2-5 Km in dense urban and 15 Km in suburban areas	Several km
Security	Partially addressed	Embedded end-to-end AES128 encryption	RC4
Modulation	UNB DBPSK(UL) GFSK (DL)	spread spectrum	BPSK/OFDM
Spreading	DSSS	Chirp Spread Spectrum modulation (CSS)	DSSS
Features	Long battery life	Long battery life	Long Range

Edge Computing Constraints

Edge computing refers to bringing the core cloud services, computations, and data at the close proximity of data sources in IoT, i.e., sensing devices and systems (Mao et al., 2017; Chiang & Zhang, 2016; Hu, Patel, Sabella, Sprecher, & Young, 2015; Garcia Lopez et al., 2015). Edge computing benefits by enabling local intelligence and distributed data processing at the edge of the Internet, hence it reduces backhaul traffic. Since the cost of computing on the edge is reducing and energy efficiency is increasing, data mining on the edge is a feasible choice. At the same time, data stream is massively evolving in terms of volume and velocity, and it contains great value for citizens and organisations.

Data mining applications in IoT devices and systems could be fully deployed on the edge, but they always require additional support from large-scale cloud infrastructures. The edge computing network is based on multiple layers including edge devices (i.e., sensors, data generators, and sensing systems), edge nodes (i.e., smart phones, smart vehicles, or any other devices or systems having enough computational capabilities to support data processing requirements of edge devices), edge gateway (which provides interface between edge devices and backend network) (Chiang & Zhang, 2016), and edge computing servers that are based on micro data centres, cloudlets, edge servers, and RAN servers. Edge computing systems also encompass SCADA systems, cyber physical systems, software defined networking, network function virtualisation, and robotic technologies. Edge computing systems enable massively distributed, resilient, and scalable IoT systems for low latency real-time distributed data mining systems. In addition, these systems provide access to spatio-temporal data streams, local and geographically distributed intelligence, and distributed security and privacy controls of data and systems (Montero et al., 2015). Edge computing systems are facilitated by D2D Communication at the edge, location-awareness, proximal computing, and context-aware networking (Amento et al., 2016).

Despite immense benefits and high utility, several challenges must be addressed before full adoption of edge computing technologies. New standards, benchmarks, and marketplaces are needed considering the massive heterogeneity in terms of data, devices, communication systems and protocols, computing technologies, and business and application

models (Dastjerdi, Gupta, Calheiros, Ghosh, & Buyya, 2016; Varghese, Wang, Barbhuiya, Kilpatrick, & Nikolopoulos, 2016; García, Fernández, Ruiz-Cortes, Dustdar, & Toro, 2017). In addition, new frameworks, programming models, and languages are needed (Varghese et al., 2016). The edge-first approach in edge computing systems require lightweight data mining libraries, APIs, assemblies, and algorithms (Varghese et al., 2016). Considering the different features of IoT devices, new micro operating systems and data mining specific micro-services could be adopted. Also, containerisation and virtualisation technologies can enable fast data processing across the edge networks (Varghese et al., 2016; Farahmandpour, Versteeg, Han, & Kameswaran, 2017).

Mobility is the primary characteristic of most of IoT devices, however, inconsistent intermittent Internet connectivity and low bandwidth can easily lead towards application failures which may degrade the quality of service and quality of experience (Liang, 2017). Therefore, data management and data governance strategies are needed in order to ensure seamless execution of data mining applications in edge computing systems (Rehman, Sun, Wah, & Khan, 2016). In addition, the continuous data transfer in cloud servers increases the cost of data communication, hence application processing within edge computing systems can decrease this cost. Efficient resource management across edge computing systems and backend cloud is required in order to maximise the utilisation of IoT systems (L. Wang, Jiao, Kliazovich, & Bouvry, 2016). Data mining methods can help in monitoring, detection, and predicting resource-intensive operations in IoT applications in order to develop efficient resource management schemes. IoT devices has bounded computing, networking, and battery power resources, therefore, it is quite challenging to orchestrate general purpose computing services on the edge (Rehman, Sun, Wah, & Khan, 2016). Data mining applications are implemented in the form of complete data pipeline that converts raw data streams into knowledge patterns. The computational complexities and resource consumption of each data conversion operation vary among different applications, therefore, developing application partitioning and computation offloading strategies in edge computing systems becomes very hard (Orsini, Bade, & Lamersdorf, 2015; Rehman, Sun, Wah, & Khan, 2016; Rehman, Sun, Wah, Iqbal, & Jayaraman, 2016; Samie et al., 2016).

The continuous data transfer in edge servers can quickly overload the MEC systems,

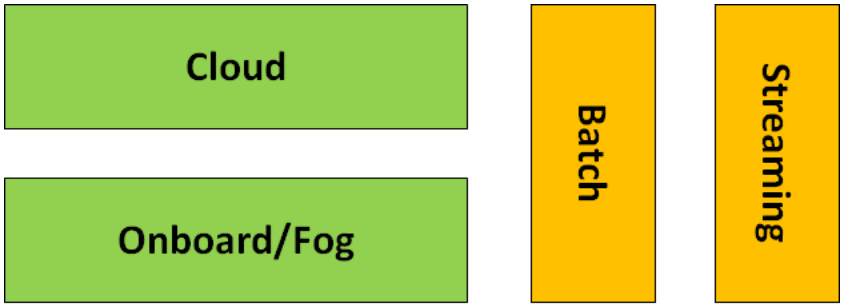

Figure 4: Categorisation of Data Mining Techniques for IoT

which can significantly impact the quality of experience. Although researchers have proposed different recovery schemes for overloaded MEC servers (Satria, Park, & Jo, 2017), the schemes work by offloading the workloads to nearest MEC servers. However, efficient and intelligent recovery schemes are needed in order to execute computation-intensive processes (such as virtual reality and computer vision applications) in MEC servers. Alternately, proactive service replications help in handling overloaded servers but the replicated services increase the resource consumption and bandwidth utilisation across edge computing system.

Having discussed the major issues and solutions for realising IoT applications including issue pertaining to privacy-preserving, networking and resource constraints at the edge, the following section discusses data mining methods and systems suited for IoT applications.

Data Mining Methods and Systems in IoT

Methods

Data mining techniques used in IoT applications can be broadly categorised according to their execution platform into (1) onboard and (2) cloud-based. They can also be categorised according to the mode of operation to (1) batch and (2) streaming. Figure 4 shows these two classifications of the adopted data mining techniques in IoT. In the following, we provide a discussion of each category, and the suitability to various IoT applications.

Onboard data mining methods have the capability to run in resource-constrained environments. Onboard data mining algorithms have been developed for wireless sensor networks and small computational devices over the last two decades (Gama & Gaber, 2007). A number

of motives make such methods particularly suitable for IoT applications including addressing privacy issues and networking constraints discussed in the previous section. Early work in the area of *onboard data mining* can be traced back to lightweight data mining methods developed using *granularity-based approach* (M. Gaber, Krishnaswamy, & Zaslavsky, 2005; M. M. Gaber & Yu, 2006; M. Gaber, 2009; M. M. Gaber & Philip, 2006). It is worth noting that the onboard methods may run out of the 'thing(s)' at a computational facility closer to the data source, forming what is usually referred to as *fog computing* (S. Yi, Li, & Li, 2015). Onboard data mining enables what has been termed as *smart objects* in (Kortuem, Kawsar, Sundramoorthy, & Fitton, 2010) to be realised. Smart objects are *things* that are not only able to sense the environment they operate in, but also to interpret events and react to them.

Cloud-based data mining methods are designed for scalability through parallelisation and distribution of processes and data sets for large volumes of data. These methods are best fit for IoT applications that run at a national level, or more generally over a large geographical area. When data are collected and possibly aggregated from various 'things', they can be used for longer term data mining tasks. Examples of such IoT applications can be weather forecasting, environmental monitoring, and large scale healthcare applications. A number of platforms have been developed by giant vendors including Google[3], IBM[4] and Microsoft[5]. For example, the authors in (Loai, Mehmood, Benkhlifa, & Song, 2016) discuss the use of data mining methods in the cloud to enable efficient healthcare services, utilising advances in networking capabilities. The use of *cloudlet* as a hardware infrastructure between the cloud and mobile devices used by patients is proposed to reduce latency.

The aforementioned two categories organise the data mining methods applied to IoT applications according to the executing location. However, regardless to where the data mining process is done, there are two modes of operating these methods, namely, batch and streaming.

Batch data mining methods operate on stored data, as the methods are inherently iterative. These methods suit IoT applications that operate on historical data at different

[3]https://cloud.google.com/iot-core/

[4]https://www.ibm.com/internet-of-things

[5]https://azure.microsoft.com/en-gb/suites/iot-suite/

levels of granularity. Thus, such methods are naturally cloud-based. However, in applications with sufficient storage capacity at the edge, batch methods are still valid solutions, including healthcare applications for home monitoring of an elderly or a patient with a chronic condition. Vital signs monitored continuously can be stored locally, and suitable data mining algorithms can be applied accordingly. A large number of methods lie under this category, including classification and regression methods for predictive data mining, and clustering and link/affinity analysis methods for descriptive data mining. More recently, deep learning methods have been used to mine IoT sourced data (e.g., (Kim & Kim, 2017) and (Lane, Bhattacharya, Georgiev, Forlivesi, & Kawsar, 2016)). The authors in (Alam, Mehmood, Katib, & Albeshri, 2016) analysed the use of 8 data mining techniques on IoT data sets. These techniques include Support Vector Machine, K-Nearest Neighbours, Linear Discriminant Analysis, Naive Bayes, C4.5 decision trees, C5.0 decision trees, Artificial Neural Networks, and Deep Neural Networks. The experimental work in this paper concluded that decision trees are favourable in terms of accuracy and computational overhead.

Streaming data mining methods are applied on live data, and best suit IoT applications when the data velocity is high, and there is a real-time necessitation in acting to the modelling process. Streaming methods can be used in the cloud, or more typically at the edge. A large number of streaming data mining methods have been proposed over the last couple of decades. In (M. M. Gaber, 2012), these methods are categorised into: (1) two-phase methods, where an online phase of processing is used to feed in a batch phase; (2) Hoeffding bound-based methods, where a statistical measure is used to determine a sample size used in a variety of ways according to the adopted technique; (3) symbolic approximation-based methods, where a times series is converted to a compact symbolic representation; and (4) granularity-based techniques, where the techniques adapt to the availability of computational resources.

Edge Data Mining Systems for IoTs

Edge computing is an alternate solution for data mining targeted for IoT applications (Salman, Elhajj, Kayssi, & Chehab, 2015; Satyanarayanan et al., 2015). This section will discuss various systems that adopted edge computing.

Data Aggregation on the Edge-servers

Researchers in (Rahmani et al., 2017) developed smart edge servers for full deployment of healthcare applications. The proposed system (named as UT-GATE) enables multiple services for local data storage, embedded data mining, and real-time data processing. To this end, an intermediary layer between edge nodes (i.e., sensors and sensing systems) and centralised cloud servers was designed to cope with mobility, security, energy-efficiency, reliability, and scalability related issues. Local data processing in UT-GATE is performed using multiple application components for data filtering, data compression, data fusion, and data analysis. In addition, the UT-GATE provides components to control data rate between sensors and edge servers and local data stores for intermediate storage.

Data Analytics process in U-GATE is executed by collecting raw data streams from different biomedical sensors and other relevant IoT devices. U-GATE performs data filtering and fusion operations followed by execution of data mining and machine learning processes for data analysis. Considering the outcomes of data analysis, U-GATE notifies users or instantiates different actuators. Further, the system adjusts the data streams for adaptation purposes and compresses the knowledge patterns for local storage in edge servers. Finally, the obtained data patterns are securely and opportunistically transferred in the cloud servers.

Trading Timeliness and Accuracy in Geo-distributed Streaming Analytics

Despite massively distributed and availability of onboard computational resources, efficient bandwidth utilisation is the primary concern in edge analytics systems (Rehman, Batool, Liew, Teh, et al., 2017). Therefore, data mining algorithms on the edge need to trade-off accuracy and timeliness in order to be efficiently executed in mobile-edge environments. Researchers in (Heintz, Chandra, & Sitaraman, 2016) focused on windowed-grouped aggregation method to study the tradeoff between timeliness and accuracy, and proposed a set of offline optimal algorithms to minimize latency in the case of acceptable accuracies and vice versa. Additionally, the real implementation of online algorithms was made in order to find the optimal points on accuracy-timeliness curve. The research outcomes reveal that the proposed online algorithms outperformed offline optimal algorithms on several datasets.

Researchers in (Harth, Anagnostopoulos, & Pezaros, 2017) proposed a lightweight and distributed prediction method for efficient data aggregation on the edge. The scheme also enables efficient predictive modelling within the distributed edge computing system. The scheme works by monitoring the changes in data streams and predicts if the collected data stream should be transferred to edge servers or not. In addition, it reconstructs the remaining data stream in order to minimise the bandwidth communication. Experimental results reveal that the proposed prediction schemes reduce the bandwidth utilisation and error rates in massively distributed edge networks. The authors proposed that communication efficiency in edge networks could be further increased using intelligent delay-tolerant mechanisms in edge computing networks.

Edge Data Mining Services for IoTs

Data aggregation in centralised cloud increases privacy concerns and reduce responsiveness of IoT applications. Researchers in (Xu et al., 2017; Rehman et al., 2018) proposed scalable approaches for deploying real-time services for IoT analytics. Moreover, a rule-based unified analytic model was proposed in order to improve responsiveness of IoT applications. The proposed model was implemented as an extension of IBM Watson and released as part of IBM Bluemix. The proposed model was tested with an industrial use case and the results show that the proposed engine works well with edge servers, however, it should be further improved to be deployed in highly resource-constrained devices and systems.

Large scale enterprises collect high-speed and continuous data streams from geographically distributed data sources across multiple continents. Researchers proposed WANalytics which is a highly distributed analytic system and pushes data to edge servers (Vulimiri, Curino, Godfrey, Karanasos, & Varghese, 2015). Additionally, the system enables optimised workflow execution and opportunistically replicates data whenever needed at the edge. The experimental evaluation was performed using Microsoft workloads and the results reveal that WANalytics reduces bandwidth consumption around 257 times.

Data Reduction on the Edge

Edge servers can help in reducing data streams from geographically dispersed data sources (Yang, 2017; Rehman, Jayaraman, Medhat Gaber, et al., 2017). Researchers in (Dubey et al., 2015) proposed an embedded computation server to run data mining and data analytics operations on raw data streams. The resultant knowledge patterns are stored at edge servers and unique patterns are transmitted to the cloud. The proposed model was tested with tele-health big data use-case and the results show the significant data reduction and improvement in overall system efficiency. Researchers in (Gupta, Vahid Dastjerdi, Ghosh, & Buyya, 2017) proposed a simulator, named iFogSim, for fog cloud computing environment. iFogSim supports multiple modules for traditional data processing in cloud and reduces data using edge devices. The data reduction on the edge is achieved by simulating data mining and machine learning algorithms inside mobile devices. The experimental evaluation exhibits significant reduction in data streams and latency on the edge of the network. Considering the mobility and limited computational resources, the partial execution of dataflow in mobile IoT devices is an alternate approach for data reduction on the edge. Researchers in (Alturki, Reiff-Marganiec, & Perera, 2017; Rehman, Sun, Wah, Iqbal, & Jayaraman, 2016) created transient datasets by performing preprocessing, feature extraction, and data fusion operations on the edge while executing data mining and machine learning operations at cloud data centers. The research reveals about 98% data reduction on the edge of the network.

Distributed Analytics in Edge Cloud Environments

Event analytics is one of the core functionalities in IoT based data mining systems. Researchers in (Ghosh & Simmhan, 2016) mapped events on direct acyclic graph (DAG) whereby the data streams were placed on the nodes and the queries were represented on the vertices. In order to perform event analytics across edge and cloud computing environments, the DAG optimisation was performed using optimal brute force algorithms and Genetic Algorithm meta-heuristics. The proposed scheme was evaluated using 17 use cases and applied multiple queries in order to find the efficiency and effectiveness. The proposed scheme returns optimal or near-optimal results as compared to a traditional brute force ap-

proach. The experiments were performed without considering virtual machine migrations and device and data level heterogeneities. Therefore, this experimentation may not be generalised for all types of IoT systems, especially in the case, when IoT devices are deployed in large scale mobile environments.

Latency-reduction on the Edge Servers

The execution of computation-intensive applications in traditional cloud environments and even in the highly-dense edge servers increases the latency. Researchers in (S. Yi et al., 2017) deployed video analytics applications on the edge servers and applied multiple optimisation methods to offload the workloads in different edge servers. The optimisation methods help in prioritising the workloads at edge servers in order to minimise the response time. In addition, a latency-aware inter-edge collaboration scheme was used that would potentially reduce the latency across edge and cloud computing environments.

Final Remarks

Data mining methods and systems are set to play an important role in realising the full potential of IoT systems. This advanced review paper has been an application driven that highlighted main applications of IoT, challenges facing the area, and data mining methods and systems addressing these challenges. We can identify three advancements that are likely to shape the future of data mining in IoT as follows.

1. **Increasing computational power at the edge**: as cloud is still the main source of computation in the majority of system architectures, the ever increasing power of small computational devices such as smart phones is set to balance out the reliance on cloud computing and edge computing in the foreseen future. Consequently, edge analytics will flourish, and a new breed of distributed data stream mining algorithms will be developed to serve the variety of IoT applications.

2. **The increase in communication capabilities with 5G technologies**: this would enable new system architectures that bring both cloud and edge computing working

together rather seamlessly. This, in turn, will enable data and models to move at extremely high speeds between cloud and edge devices.

3. **Improvement in battery technologies**: the development of long lasting batteries of IoT devices along with lightweight methods will result in reaching an equilibrium where nowadays power intense computation becomes much more energy efficient processes. This will consequently enhance the quality of experience in using IoT applications.

References

Abdallah, F., Basurra, S., & Gaber, M. (2017, 7). A hybrid agent-based and probabilistic model for fine-grained behavioural energy waste simulation. *29 IEEE International Conference on Tools with Artificial Intelligence (ICTAI), Boston, USA*.

Abhishek, P. V., Manjunatha, H. G., Sudarshan, P. B., & Reddy, K. P. V. (2016). Iot operated wheel chair. *International Journal of Engineering Research*, *5*(4), 1089 1091. doi: 10.5958/2319-6890

Abouzakhar, N. S., Jones, A., & Angelopouloui, O. (2017). Internet of things security: A review of risks and threats to healthcare sector. In *Proceedings of ieee international conference on internet of things* (p. 373 - 378).

Ahmed, E., Yaqoob, I., Hashem, I. A. T., Khan, I., Ahmed, A. I. A., Imran, M., & Vasilakos, A. V. (2017). The role of big data analytics in internet of things. *Computer Networks*, *129*, 459–471.

Ahmed, M. A., Kang, Y. C., & Kim, Y.-C. (2015). Communication network architectures for smart-house with renewable energy resources. *Energies*, *8*(8), 8716–8735. Retrieved from http://www.mdpi.com/1996-1073/8/8/8716 doi: 10.3390/en8088716

Alam, F., Mehmood, R., Katib, I., & Albeshri, A. (2016). Analysis of eight data mining algorithms for smarter internet of things (iot). *Procedia Computer Science*, *98*, 437–442.

Alam, F., Mehmood, R., Katib, I., Albogami, N., & Albeshri, A. (2017). Data fusion and iot for smart ubiquitous environments: A survey. *IEEE Access*.

Al-mawee, W. (2012). *Privacy and security issues in iot healthcare applications for the disabled users a survey* (Unpublished doctoral dissertation). Western Michigan University.

Al-Sarawi, S., Anbar, M., Alieyan, K., & Alzubaidi, M. (2017). Internet of things (iot) communication protocols. In *Information technology (icit), 2017 8th international conference on* (pp. 685–690).

Alturki, B., Reiff-Marganiec, S., & Perera, C. (2017). A hybrid approach for data analytics for internet of things. *arXiv preprint arXiv:1708.06441*.

Amento, B., Balasubramanian, B., Hall, R. J., Joshi, K., Jung, G., & Purdy, K. H. (2016). Focusstack: Orchestrating edge clouds using location-based focus of attention. In *Edge computing (sec), ieee/acm symposium on* (pp. 179–191).

Appavoo, M. C. C. A. B., Paramasiven, & Chang, E.-C. (2016). *Efficient and privacy-preserving access to sensor data for internet of things (iot) based services.* [Conference Paper]. IEEE.

Arfat, Y., Aqib, M., Mehmood, R., Albeshri, A., Katib, I., Albogami, N., & Alzahrani, A. (2017). Enabling smarter societies through mobile big data fogs and clouds. *Procedia Computer Science*, *109*, 1128–1133.

Augustin, A., Yi, J., Clausen, T., & Townsley, W. M. (2016). A study of lora: Long range & low power networks for the internet of things. *Sensors*, *16*(9), 1466.

Baktir, A. C., Ozgovde, A., & Ersoy, C. (2017). How can edge computing benefit from software-defined networking: A survey, use cases, and future directions. *IEEE Communications Surveys & Tutorials*, *19*(4), 2359–2391.

Bardyn, J.-P., Melly, T., Seller, O., & Sornin, N. (2016). Iot: The era of lpwan is starting now. In *European solid-state circuits conference, esscirc conference 2016: 42nd* (pp. 25–30).

Basurra, S., & Jankovic, L. (2016). Performance comparison between knn and nsga-ii algorithms as calibration approaches for building simulation models. In N. Hamza (Ed.), *Proceedings of the 3rd ibpsa-england conference bso 2016, great north museum, newcastle* (p. 1093). ibpsa.org. (An optional note)

Battle, S., & Gaster, B. (2017). Lorawan bristol. In *Proceedings of the 21st international*

database engineering & applications symposium (pp. 287–290).

Bekara, C. (2014). Security issues and challenges for the iot-based smart grid. *Procedia Computer Science*, *34*(Supplement C), 532 - 537. Retrieved from http://www.sciencedirect.com/science/article/pii/S1877050914009193 (The 9th International Conference on Future Networks and Communications (FNC'14)/The 11th International Conference on Mobile Systems and Pervasive Computing (MobiSPC'14)/Affiliated Workshops) doi: https://doi.org/10.1016/j.procs.2014.07.064

Bouhafs, F., Mackay, M., & Merabti, M. (2012, Jan). Links to the future: Communication requirements and challenges in the smart grid. *IEEE Power and Energy Magazine*, *10*(1), 24-32. doi: 10.1109/MPE.2011.943134

Bretzke, W.-r. (2013, 06). Global urbanization: a major challenge for logistics. *Logistics Research*, *6*(2-3), 57-62. Retrieved from https://search.proquest.com/docview/1355885969?accountid=10749 (Copyright - Springer-Verlag Berlin Heidelberg 2013; Document feature - ; Last updated - 2015-09-05)

Bröring, A., Schmid, S., Schindhelm, C.-K., Khelil, A., Käbisch, S., Kramer, D., ... Teniente, E. (2017). Enabling iot ecosystems through platform interoperability. *IEEE software*, *34*(1), 54–61.

Cao, H., Wachowicz, M., & Cha, S. (2017). Developing an edge analytics platform for analyzing real-time transit data streams. *arXiv preprint arXiv:1705.08449*.

Chiang, M., & Zhang, T. (2016). Fog and iot: An overview of research opportunities. *IEEE Internet of Things Journal*, *3*(6), 854–864.

Cognizant. (2016). *How the internet of things is transforming medical devices.* Retrieved 2017-12-01, from ⌀https://www.cognizant.com/whitepapers/how-the-internet-of-things-is-transforming-medical-devices-codex1945.pdf

Commission, E. (2010). *European commission directive 2010/31/eu on 19 may 2010 on the energy performance of buildings* (Tech. Rep. No. 18). European Parliament and of the Council.

Connell, A., Montgomery, H., Morris, S., Nightingale, C., Stanley, S., Emerson, M., ... Laing, C. (2017). Service evaluation of the implementation of a digitally-enabled care pathway for the recognition and management of acute kidney injury. *F1000Research*,

1(1), 47 65. doi: 10.12688/f1000research.11637.1

Corchado, J. M., Bajo, J., Tapia, D. I., & Abraham, A. (2010). Using heterogeneous wireless sensor networks in a telemonitoring system for healthcare. *IEEE transactions on information technology in biomedicine*, *14*(2), 234–240.

Dama, S., Sathya, V., Kuchi, K., & Pasca, T. V. (2017). A feasible cellular internet of things: Enabling edge computing and the iot in dense futuristic cellular networks. *IEEE Consumer Electronics Magazine*, *6*(1), 66–72.

Dastjerdi, A. V., Gupta, H., Calheiros, R. N., Ghosh, S. K., & Buyya, R. (2016). Fog computing: Principles, architectures, and applications. *arXiv preprint arXiv:1601.02752*.

Deepmind. (2017). *Deepmind health independent review panel annual report.* Retrieved 2017-12-01, from œhttps://deepmind.com/documents/85/DeepMind%20Health%20Independent%20Review%20Annual%20Report%202017.pdf

Department for Business, E. e., Energy & Industrial Strategy Part of: Household energy, prices, E., & bills. (2013). *Smart meters: a guide.* Retrieved 2017-11-09, from https://www.gov.uk/guidance/smart-meters-how-they-work

Dubey, H., Yang, J., Constant, N., Amiri, A. M., Yang, Q., & Makodiya, K. (2015). Fog data: Enhancing telehealth big data through fog computing. In *Proceedings of the ase bigdata & socialinformatics 2015* (p. 14).

Elmisery, S. R., Ahmed M., & Aborizka, M. (2017). A new computing environment for collective privacy protection from constrained healthcare devices to iot cloud services [Journal Article]. *Cluster Computing*, 1-28. doi: 10.1007/s10586-017-1298-1

Elmisery, S. R., Ahmed M., & Botvich, D. (2016). A fog based middleware for automated compliance with oecd privacy principles in internet of healthcare things. [Journal Article]. *IEEE Access*, *4*, 8418-8441. doi: 10.1109/ACCESS.2016.2631546

Erlinghagen, S., Lichtensteiger, B., & Markard, J. (2015). Smart meter communication standards in europe a comparison. *Renewable and Sustainable Energy Reviews*, *43*(Supplement C), 1249 - 1262. Retrieved from http://www.sciencedirect.com/science/article/pii/S1364032114010041 doi: https://doi.org/10.1016/j.rser.2014.11.065

Evans, D., & Eyers, D. M. (2012). *Efficient data tagging for managing privacy in the internet*

of things [Conference Paper]. IEEE.

Farahmandpour, Z., Versteeg, S., Han, J., & Kameswaran, A. (2017). Service virtualisation of internet-of-things devices: techniques and challenges. In *Proceedings of the 3rd international workshop on rapid continuous software engineering* (pp. 32–35).

Flore, D. (2016). 3gpp standards for the internet-of-things. *Recuperado el*, *25*.

Fong, E., & Chung, W. (2013). Mobile cloud-computing-based healthcare service by non-contact ecg monitoring. *Sensors*, *13*(1), 678 708. doi: 10.3390/s130100001

G3ict. (2015). *Internet of things: New promises for persons with disabilitie.* Retrieved 2017-12-01, from ☙http://g3ict.org/download/p/fileId_1025/productId_335

Gaber, M. (2009). Data stream mining using granularity-based approach. *Foundations of Computational, IntelligenceVolume 6*, 47–66.

Gaber, M., Krishnaswamy, S., & Zaslavsky, A. (2005). On-board mining of data streams in sensor networks. *Advanced methods for knowledge discovery from complex data*, 307–335.

Gaber, M. M. (2012). Advances in data stream mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(1), 79–85.

Gaber, M. M., Gomes, J. B., & Stahl, F. (2014). Pocket data mining. *Big Data on Small Devices. Series: Studies in Big Data*.

Gaber, M. M., & Philip, S. Y. (2006). A holistic approach for resource-aware adaptive data stream mining. *New Generation Computing*, *25*(1), 95–115.

Gaber, M. M., & Yu, P. S. (2006). A framework for resource-aware knowledge discovery in data streams: a holistic approach with its application to clustering. In *Proceedings of the 2006 acm symposium on applied computing* (pp. 649–656).

Gama, J., & Gaber, M. M. (2007). *Learning from data streams: processing techniques in sensor networks*. Springer.

García, J. M., Fernández, P., Ruiz-Cortes, A., Dustdar, S., & Toro, M. (2017). Edge and cloud pricing for the sharing economy. *IEEE Internet Computing*, *21*(2), 78–84.

Garcia Lopez, P., Montresor, A., Epema, D., Datta, A., Higashino, T., Iamnitchi, A., ... Riviere, E. (2015). Edge-centric computing: Vision and challenges. *ACM SIGCOMM Computer Communication Review*, *45*(5), 37–42.

Ge, J. B. H. W. G., Mengmeng, & Kim., D. S. (2017). A framework for automating security analysis of the internet of things. [Journal Article]. *Journal of Network and Computer Applications*, *83*, 12-27.

Ghosh, R., & Simmhan, Y. (2016). Distributed scheduling of event analytics across edge and cloud. *arXiv preprint arXiv:1608.01537*.

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.

Gungor, V. C., Sahin, D., Kocak, T., Ergut, S., Buccella, C., Cecati, C., & Hancke, G. P. (2011, Nov). Smart grid technologies: Communication technologies and standards. *IEEE Transactions on Industrial Informatics*, *7*(4), 529-539. doi: 10.1109/TII.2011.2166794

Gupta, H., Vahid Dastjerdi, A., Ghosh, S. K., & Buyya, R. (2017). ifogsim: A toolkit for modeling and simulation of resource management techniques in the internet of things, edge and fog computing environments. *Software: Practice and Experience*, *47*(9), 1275–1296.

Hagras, H., Callaghan, V., Colley, M., & Clarke, G. (2003). A hierarchical fuzzygenetic multi-agent architecture for intelligent buildings online learning, adaptation and control. *Information Sciences*, *150*(1), 33 - 57. Retrieved from http://www.sciencedirect.com/science/article/pii/S0020025502003687 (Recent Advances in Soft Computing) doi: https://doi.org/10.1016/S0020-0255(02)00368-7

Harth, N., Anagnostopoulos, C., & Pezaros, D. (2017). Predictive intelligence to the edge: impact on edge analytics. *Evolving Systems*, 1–24.

Heintz, B., Chandra, A., & Sitaraman, R. K. (2016). Trading timeliness and accuracy in geo-distributed streaming analytics. In *Socc* (pp. 361–373).

Hii, P., Lee, S., Kwon, T., & Chung, W. (2011). Smart phone based patient-centered remote health monitoring application in wireless sensor network. *Sensor Letters*, *9*(2), 791-796.

Hong, J. B., & Kim, D. S. (2016). Assessing the effectiveness of moving target defenses using security models [Journal Article]. *IEEE Transactions on Dependable and Secure*

Computing, *13*(2), 163-177.

Hsu, C.-W., & Yeh, C.-C. (2017). Understanding the factors affecting the adoption of the internet of things. *Technology Analysis & Strategic Management*, *29*(9), 1089–1102.

Hu, Y. C., Patel, M., Sabella, D., Sprecher, N., & Young, V. (2015). Mobile edge computinga key technology towards 5g. *ETSI White Paper*, *11*(11), 1–16.

Isa, N., Yusoff, M., & Mohamed, A. (2014, Dec). A review on recent traffic congestion relief approaches. In *2014 4th international conference on artificial intelligence with applications in engineering and technology* (p. 121-126). doi: 10.1109/ICAIET.2014.29

Islam, S., Kwak, D., Kabir, H., Hossain, M., & Kwak, K. (2015). The internet of things for health care: A comprehensive survey. *IEEE Access*, *3*(1), 678 708. doi: 10.1109/access.2015.2437951

Istepanian, R. S. H., Hu, S., Philip, N. Y., & Sungoor, A. (2011). The potential of internet of m-health things m-iot for non-invasive glucose level sensing. In *Proceedings ieee annual international conference on engineering in medicine and biology society (embc)* (p. 5264 - 5266).

Jabbarpour, M. R., Nabaei, A., & Zarrabi, H. (2016, Dec). Intelligent guardrails: An iot application for vehicle traffic congestion reduction in smart city. In *2016 ieee international conference on internet of things (ithings) and ieee green computing and communications (greencom) and ieee cyber, physical and social computing (cpscom) and ieee smart data (smartdata)* (p. 7-13). doi: 10.1109/iThings-GreenCom-CPSCom-SmartData.2016.29

Jara, A. J., Zamora-Izquierdo, M. A., & Skarmeta, A. F. (2013). Interconnection framework for mhealth and remote monitoring based on the internet of things. *IEEE Journal on Selected Areas in Communications*, *31*(9), 47 65. doi: 10.1109/JSAC.2013.SUP.0513005

Jian, Z., Zhanli, W., & Zhuang, M. (2012). *Temperature measurement system and method based on home gateway.* Google Patents. Retrieved from https://www.google.com/patents/CN102811185A?cl=en (CN Patent App. CN 201,110,148,247)

Jin, P. J., Zhang, G., Walton, C. M., Jiang, X., & Singh, A. (2013, Dec). Analyzing the impact of false-accident cyber attacks on traffic flow stability in connected vehicle

environment. In *2013 international conference on connected vehicles and expo (iccve)* (p. 616-621). doi: 10.1109/ICCVE.2013.6799866

Kim, H.-Y., & Kim, J.-M. (2017). A load balancing scheme based on deep-learning in iot. *Cluster Computing*, *20*(1), 873–878.

Konstantinidis, E. I., Bamparapoulos, G., Billis, A., & Bamidis, P. D. (2015). Internet of things for an age-friendly healthcare. *Digital Healthcare Empowering Europeans*, *210*(1), 587 591. doi: 10.3233/978-1-61499-512-8-587

Kortuem, G., Kawsar, F., Sundramoorthy, V., & Fitton, D. (2010). Smart objects as building blocks for the internet of things. *IEEE Internet Computing*, *14*(1), 44–51.

Lane, N. D., Bhattacharya, S., Georgiev, P., Forlivesi, C., & Kawsar, F. (2016). Accelerated deep learning inference for embedded and wearable devices using deepx. In *Proceedings of the 14th annual international conference on mobile systems, applications, and services companion* (pp. 109–109).

Larson, E. C., Goel, M., Boriello, G., Heltshe, S., Rosenfeld, M., & Patel, S. N. (2011). Spirosmart: using a microphone to measure lung function on a mobile phone. In *Proceedings of the 2012 acm conference on ubiquitous computing* (p. 280 - 289).

Lee, E.-K. (2016, January). Advancing building energy management system to enable smart grid interoperation. *Int. J. Distrib. Sen. Netw.*, *2016*, 1:1–1:1. Retrieved from https://doi.org/10.1155/2016/3295346 doi: 10.1155/2016/3295346

Lee, H. (2016). The internet of things and assistive technologies for people with disabilities: Applications, trends, and issues. In *Internet of things and advanced application in healthcare* (p. 32 - 65). IGI Global. Retrieved from "https://www.igi-global.com/chapter/the-internet-of-things-and-assistive-technologies-for-people-with-disabilities/170236" doi: 10.4018/978-1-5225-1820-4.ch002

Li, Y., & Li, M. (2017). A privacy protection mechanism for numerical control information in internet of things [Journal Article]. *International Journal of Distributed Sensor Networks*, *13*(8), 1550147717726312.

Liang, B. (2017). *Mobile edge computing*. Cambridge University Press.

Loai, A. T., Mehmood, R., Benkhlifa, E., & Song, H. (2016). Mobile cloud computing model

and big data analysis for healthcare applications. *IEEE Access*, *4*, 6171–6180.

Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., & Aharon, D. (2015). *The internet of things: Mapping the value beyond the hype.* McKinsey Global Institute. Retrieved 2017-12-01, from ☙https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/The%20Internet%20of%20Things%20The%20value%20of%20digitizing%20the%20physical%20world/The-Internet-of-things-Mapping-the-value-beyond-the-hype.ashx

Mao, Y., You, C., Zhang, J., Huang, K., & Letaief, K. B. (2017). A survey on mobile edge computing: The communication perspective. *IEEE Communications Surveys & Tutorials*.

Marchiori, A. (2017). Maximizing coverage in low-power wide-area iot networks. In *Pervasive computing and communications workshops (percom workshops), 2017 ieee international conference on* (pp. 467–472).

Margelis, G., Piechocki, R., Kaleshi, D., & Thomas, P. (2015). Low throughput networks for the iot: Lessons learned from industrial implementations. In *Internet of things (wf-iot), 2015 ieee 2nd world forum on* (pp. 181–186).

Martin, A. B., Kelly D., & Palmatier, R. W. (2017). Data privacy: Effects on customer and firm performance [Journal Article]. *Journal of Marketing*, *81*(1), 36-58.

Masek, P., Masek, J., Frantik, P., Fujdiak, R., Ometov, A., Hosek, J., ... Misurec, J. (2016). A harmonized perspective on transportation management in smart cities: The novel iot-driven environment for road traffic modeling. *Sensors*, *16*(11).

Mehmood, R., Alam, F., Albogami, N. N., Katib, I., Albeshri, A., & Altowaijri, S. M. (2017). Utilearn: A personalised ubiquitous teaching and learning system for smart societies. *IEEE Access*, *5*, 2615–2635.

Mittal, A. K., & Bhandari, D. (2013, Feb). A novel approach to implement green wave system and detection of stolen vehicles. In *2013 3rd ieee international advance computing conference (iacc)* (p. 1055-1059). doi: 10.1109/IAdCC.2013.6514372

Mohammed, J., Thakral, A., Ocneanu, A. F., Jones, C., Lung, C., & Adler, A. (2014). Internet of things: Remote patient monitoring using web services and cloud computing. In *Proceedings of ieee international conference on internet of things (ithings)* (p. 256-

263).

Montero, D., Yannuzzi, M., Shaw, A., Jacquin, L., Pastor, A., Serral-Gracia, R., ... others (2015). Virtualized security at the network edge: a user-centric approach. *IEEE Communications Magazine*, *53*(4), 176–186.

Moreno, M. V., beda, B., Skarmeta, A. F., & Zamora, M. A. (2014, 7). How can we tackle energy efficiency in iot based smart buildings? *Sensors (Basel, Switzerland)*, *10*(6), 95829614. (An optional note)

Nellore, K., & Hancke, G. (2016, 11). Traffic management for emergency vehicle priority based on visual sensing. *Sensors*, *16*(11). doi: 10.3390/s16111892

Niewolny, D. (2013). *How the internet of things is revolutionizing healthcare.* Retrieved 2017-12-01, from œhttps://www.nxp.com/docs/en/white-paper/IOTREVHEALCARWP.pdf

Noreen, U., Bounceur, A., & Clavier, L. (2017). A study of lora low power and wide area network technology. In *3rd ieee international conference on advanced technologies for signal and image processing (atsip'2017).*

Orsini, G., Bade, D., & Lamersdorf, W. (2015). Computing at the mobile edge: designing elastic android applications for computation offloading. In *Ifip wireless and mobile networking conference (wmnc), 2015 8th* (pp. 112–119).

Otgonbayar, Z. P., Ankhbayar, & Dahal, K. (2016). *Toward anonymizing iot data streams via partitioning.* [Conference Paper]. IEEE.

Ozyilmaz, K. R., & Yurdakul, A. (2017). Integrating low-power iot devices to a blockchain-based infrastructure: Work-in-progress. In *Proceedings of the thirteenth acm international conference on embedded software 2017 companion* (pp. 13:1–13:2). New York, NY, USA: ACM.

Patel, P., Ali, M. I., & Sheth, A. (2017). On using the intelligent edge for iot analytics. *IEEE Intelligent Systems*, *32*(5), 64–69.

Puustjarvi, J., & Puustjarvi, L. (2011). Automating remote monitoring and information therapy: An opportunity to practice telemedicine in developing countries. In *Proceedings of 2011 ist-africa conference* (p. 1 - 9).

Prez, D. R. D. P. L. S. M. J. N., Salvador, & Gigante., F. (2017). *Towards the cp-abe application for privacy-preserving secure data sharing in iot contexts* [Conference Paper].

Springer.

Prez-Lombard, L., Ortiz, J., & Pout, C. (2008). A review on buildings energy consumption information. *Energy and Buildings*, *40*(3), 394 - 398. Retrieved from http://www.sciencedirect.com/science/article/pii/S0378778807001016 doi: https://doi.org/10.1016/j.enbuild.2007.03.007

Rahmani, A., Gia, T., Negash, B., Anzanpour, A., Azimi, I., Jiang, M., & Liljeberg, P. (2017). Exploiting smart e-health gateways at the edge of healthcare internet-of-things: a fog computing approach. *Future Generation Computer Systems*.

Rehman, M. H., Ahmed, E., Yaqoob, I., Hashem, I. A. T., Imran, M., & Ahmad, S. (2018). Big data analytics in industrial iot using a concentric computing model. *IEEE Communications Magazine*, *56*(2), 37–43.

Rehman, M. H., Batool, A., Liew, C. S., Teh, Y.-W., et al. (2017). Execution models for mobile data analytics. *IT Professional*, *19*(3), 24–30.

Rehman, M. H., Chang, V., Batool, A., & Wah, T. Y. (2016). Big data reduction framework for value creation in sustainable enterprises. *International Journal of Information Management*, *36*(6), 917–928.

Rehman, M. H., Jayaraman, P. P., Medhat Gaber, M., et al. (2017). Rededge: A novel architecture for big data processing in mobile edge computing environments. *Journal of Sensor and Actuator Networks*, *6*(3), 17.

Rehman, M. H., Sun, C., Wah, T. Y., Iqbal, A., & Jayaraman, P. P. (2016). Opportunistic computation offloading in mobile edge cloud computing environments. In *Mobile data management (mdm), 2016 17th ieee international conference on* (Vol. 1, pp. 208–213).

Rehman, M. H., Sun, L. C., Wah, T. Y., & Khan, M. K. (2016). Towards next-generation heterogeneous mobile data stream mining applications: Opportunities, challenges, and future research directions. *Journal of Network and Computer Applications*.

Rutledge, A. K. M. A. I. A., Richard L., & Swire, P. (2014). *Defining the internet of devices: Privacy and security implications.* [Manuscript].

Sallam, A. A., & Malik, O. P. (2011). Scada systems and smart grid vision. In *Electric distribution systems* (p. 469-493). Wiley-IEEE Press. Retrieved from http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=5733135 doi: 10.1002/

9780470943854.ch13

Salman, O., Elhajj, I., Kayssi, A., & Chehab, A. (2015). Edge computing enabling the internet of things. In *Internet of things (wf-iot), 2015 ieee 2nd world forum on* (pp. 603–608).

Samie, F., Tsoutsouras, V., Bauer, L., Xydis, S., Soudris, D., & Henkel, J. (2016). Computation offloading and resource allocation for low-power iot edge devices. In *Internet of things (wf-iot), 2016 ieee 3rd world forum on* (pp. 7–12).

Saravanan, M., Das, A., & Iyer, V. (2017). Smart water grid management using lpwan iot technology. In *Global internet of things summit (giots), 2017* (pp. 1–6).

Satria, D., Park, D., & Jo, M. (2017). Recovery for overloaded mobile edge computing. *Future Generation Computer Systems*, *70*, 138–147.

Satyanarayanan, M., Simoens, P., Xiao, Y., Pillai, P., Chen, Z., Ha, K., ... Amos, B. (2015). Edge analytics in the internet of things. *IEEE Pervasive Computing*, *14*(2), 24–31.

Sundar, R., Hebbar, S., & Golla, V. (2015, Feb). Implementing intelligent traffic control system for congestion control, ambulance clearance, and stolen vehicle detection. *IEEE Sensors Journal*, *15*(2), 1109-1113. doi: 10.1109/JSEN.2014.2360288

Tomtom traffic index 2017. (2017). Retrieved 2017-11-09, from http://corporate.tomtom.com/releasedetail.cfm?releaseid=1012517

Tso, A. A. S. M. M. R. M.-E. W., Raylin, & Hossain, M. S. (2017). Privacy-preserving data communication through secure multi-party computation in healthcare sensor cloud." journal of signal processing systems [Journal Article]. , *89*(1), 51-59.

Varghese, B., Wang, N., Barbhuiya, S., Kilpatrick, P., & Nikolopoulos, D. S. (2016). Challenges and opportunities in edge computing. In *Smart cloud (smartcloud), ieee international conference on* (pp. 20–26).

Vulimiri, A., Curino, C., Godfrey, B., Karanasos, K., & Varghese, G. (2015). Wanalytics: Analytics for a geo-distributed data-intensive world. In *Conference on innovative data systems research (cidr).*

Wang, F., Hu, L., Hu, J., Zhou, J., & Zhao, K. (2017). Recent advances in the internet of things: Multiple perspectives. *IETE Technical Review*, *34*(2), 122-132.

Wang, L., Jiao, L., Kliazovich, D., & Bouvry, P. (2016). Reconciling task assignment

and scheduling in mobile edge clouds. In *Network protocols (icnp), 2016 ieee 24th international conference on* (pp. 1–6).

Xu, X., Huang, S., Feagan, L., Chen, Y., Qiu, Y., & Wang, Y. (2017). Eaaas: Edge analytics as a service. In *Web services (icws), 2017 ieee international conference on* (pp. 349–356).

Yang, S. (2017). Iot stream processing and analytics in the fog. *arXiv preprint arXiv:1705.05988*.

Yi, J. W., Xun, & Nait-Abdesselam, F. (2013). *Privacy-preserving wireless medical sensor network* [Conference Paper]. IEEE.

Yi, S., Hao, Z., Zhang, Q., Zhang, Q., Shi, W., & Li, Q. (2017). Lavea: Latency-aware video analytics on edge computing platform. In *Distributed computing systems (icdcs), 2017 ieee 37th international conference on* (pp. 2573–2574).

Yi, S., Li, C., & Li, Q. (2015). A survey of fog computing: concepts, applications and issues. In *Proceedings of the 2015 workshop on mobile big data* (pp. 37–42).

Zavitsas, K., Kaparias, I., & Bell, M. (2010, 7). *Conduits, coordination of network descriptors for urban intelligent transport systems* (Tech. Rep. No. 11). The address of the publisher: Imperial College London. (An optional note)

Zhang, Y., He, C.-Q., Tang, B.-J., & Wei, Y.-M. (2015). China's energy consumption in the building sector: A life cycle approach. *Energy and Buildings*, *94*(Supplement C), 240 - 251. Retrieved from http://www.sciencedirect.com/science/article/pii/S0378778815002030 doi: https://doi.org/10.1016/j.enbuild.2015.03.011

Zhao, D., McCoy, A. P., Du, J., Agee, P., & Lu, Y. (2017). Interaction effects of building technology and resident behavior on energy consumption in residential buildings. *Energy and Buildings*, *134*(Supplement C), 223 - 233. Retrieved from http://www.sciencedirect.com/science/article/pii/S0378778816313718 doi: https://doi.org/10.1016/j.enbuild.2016.10.049