EAI/Springer Innovations in Communication and Computing

Shalli Rani
Vyasa Sai
R. Maheswar *Editors*

IoT and WSN based Smart Cities: A Machine Learning Perspective

EAI
RESEARCH MEETS INNOVATION

Springer

EAI/Springer Innovations in Communication and Computing

Series editor

Imrich Chlamtac, European Alliance for Innovation, Ghent, Belgium

Editor's Note

The impact of information technologies is creating a new world yet not fully understood. The extent and speed of economic, life style and social changes already perceived in everyday life is hard to estimate without understanding the technological driving forces behind it. This series presents contributed volumes featuring the latest research and development in the various information engineering technologies that play a key role in this process.

The range of topics, focusing primarily on communications and computing engineering include, but are not limited to, wireless networks; mobile communication; design and learning; gaming; interaction; e-health and pervasive healthcare; energy management; smart grids; internet of things; cognitive radio networks; computation; cloud computing; ubiquitous connectivity, and in mode general smart living, smart cities, Internet of Things and more. The series publishes a combination of expanded papers selected from hosted and sponsored European Alliance for Innovation (EAI) conferences that present cutting edge, global research as well as provide new perspectives on traditional related engineering fields. This content, complemented with open calls for contribution of book titles and individual chapters, together maintain Springer's and EAI's high standards of academic excellence. The audience for the books consists of researchers, industry professionals, advanced level students as well as practitioners in related fields of activity include information and communication specialists, security experts, economists, urban planners, doctors, and in general representatives in all those walks of life affected ad contributing to the information revolution.

Indexing: This series is indexed in Scopus, Ei Compendex, and zbMATH

About EAI

EAI is a grassroots member organization initiated through cooperation between businesses, public, private and government organizations to address the global challenges of Europe's future competitiveness and link the European Research community with its counterparts around the globe. EAI reaches out to hundreds of thousands of individual subscribers on all continents and collaborates with an institutional member base including Fortune 500 companies, government organizations, and educational institutions, provide a free research and innovation platform.

Through its open free membership model EAI promotes a new research and innovation culture based on collaboration, connectivity and recognition of excellence by community.

Shalli Rani • Vyasa Sai • R. Maheswar
Editors

IoT and WSN based Smart Cities: A Machine Learning Perspective

Springer

EAI
RESEARCH MEETS INNOVATION

Editors
Shalli Rani
Chitkara University Institute
of Engineering and Technology
Chitkara University
Rajpura, Punjab, India

Vyasa Sai
Intel Corporation
Folsom, CA, USA

R. Maheswar
Department of ECE
KPR Institute of Engineering and
Technology
Coimbatore, India

ISSN 2522-8595 ISSN 2522-8609 (electronic)
EAI/Springer Innovations in Communication and Computing
ISBN 978-3-030-84181-2 ISBN 978-3-030-84182-9 (eBook)
https://doi.org/10.1007/978-3-030-84182-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

In the last decade, wireless sensor networks (WSNs) have covered an extensive range of applications in various areas, such as smart healthcare, smart industry, cloud-based applications, smart home scenarios, smart monitoring, and many other domains. The Internet of Things (IoT) has ensured the smart life via communications among objects, people, and devices. However when WSNs became the integral part of the Internet, the careful investigation of the design challenges in this area raised the new research directions.

The Internet of things (IoT) has changed the living scenarios in various ways. It has introduced connectivity among various objects such as transportation vehicles, portable devices, human beings, and buildings. It has changed the way of working by making cities smart, where data is available in the cloud. By due time, trillions of devices will be connected and all things/objects will be able to communicate remotely. IoT is being supported in both public and private networks where machine learning plays a huge role. Many research projects are being sanctioned by governments in India and other countries in the machine learning processes on IoT-WSN. In the upcoming times, this domain will have huge impact on society. Remote communication with the help of sensors and connectivity through IoT with ML analysis will become an integral part of daily activities.

Various categories of applications along with driving forces of the smart city phenomenon have increased the concerns of researchers. The contributing authors are expected to provide the insights into technologies used in IoT and WSN with data mining/machine learning. Smart cities framework is a people-centric system which revolves around real time applications and infrastructures. It requires the close integration of cyber and physical components for monitoring, understanding, and controlling the smart environment.

The editors have investigated in a top-down manner the driving forces in smart cities, by incorporating principal factors including “smart monitoring,” “public protection and strategic concerns,” and “science, technology, and Innovation.” Important areas and issues with roadblocks to implementation, use case scenarios, novel ideas and opportunities, and other factors leading to the development of practical scenarios for smart monitoring are discussed. The book has also focused on new algorithms, architectures, and platforms which can accelerate the growth of smart city concepts. Moreover, it has detailed about the set of real-world applications and case studies related to specific smart city infrastructures, huge data management, and prediction techniques by machine learning with privacy and security issues.

Rajpura, Punjab, India Shalli Rani

Folsom, CA, USA Vyasa Sai

Coimbatore, India R. Maheswar

Acknowledgment

First and foremost, I would like to thank my husband **Shvet Jain** for standing beside me throughout my research career so far and editing this book. He has been my motivation for continuing to improve my knowledge and move forward in my career. He is my lifetime achievement, and I dedicate this book of mine to him.

Moreover, without the prayers and best wishes of my parents, my sister, and my two sons, all of my achievements so far and forever would never be possible.

Here, I am also thankful to the Springer Editorial Team and my co-editors: **Dr. Vyasa Sai and Dr. R.Maheswar,** who really helped me in every aspect in the preparation of this book. Their prompt response and care to the literature and contribution is remarkable and sensational. They are great and responsible researchers indeed.

I would also like to thank each and every one behind the completion of this book, especially my **co-editors,** who actually helped a lot, and without their quick and efficient efforts, it would not be possible to get our book published. They motivated and gave me this platform to go ahead.

Last but not least, my gratitude is due to my inspiration, **Dr. Archana Mantri, Vice Chancellor, Chitkara University, Punjab,** for her trust in me and for making me confident that I can put efforts to get success in each and every field.

Contents

Protocol Design for Earthquake Alert and Evacuation in Smart Buildings

P. Sivakumar, R. S. Sandhya Devi, M. Ashwin, M. M. Rajan Singaravel, and A. D. Buvanesswaran

1 Introduction

Smart building is a major requirement when we want to make a city smart. Smart building provides reduced energy consumption and better resource management, saves life during a disaster, increases productivity etc. As the natural calamities nowadays are unpredictable due to climate changes, global warming etc., the buildings have to be smart enough to predict, warn before a disaster and provide safety and recovery measures after a disaster. Disasters could be floods, earthquakes etc. In specific earthquake poses a great threat for the lives inside a building. Here, we will discuss about the role of IoT in disaster management from various research papers and a protocol design for earthquake management in smart buildings.

P. Sivakumar (✉) · A. D. Buvanesswaran
Department of EEE, PSG College of Technology, Coimbatore, India
e-mail: psk.eee@psgtech.ac.in

R. S. Sandhya Devi
Department of EEE, Kumaraguru College of Technology, Coimbatore, India
e-mail: sandhyadevi.rs.eee@kct.ac.in

M. Ashwin
School of Computing Science and Engineering, VIT Bhopal University, Bhopal, Madhya Pradesh, India

M. M. Rajan Singaravel
Department of EEE, National Institute of Technology, Puducherry, India
e-mail: rajan.singaravel@nitpy.ac.in

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_1

2 Related Works

Early detection of a natural disaster is a challenging one, and researches are being carried out in this area. Early detection helps us to take safety measures as early as possible. The process of earthquake detection and alerting is discussed in detail in paper [1]. Event detection in natural disaster is discussed in paper [2]. Earthquake damage intensity scaling system based on Raspberry Pi [3] and Arduino Uno is clearly discussed using hardware in paper [4]. There are many works associated with the detection of earthquakes and alerting the concerned. There is also a need for automation in how to trigger safety measures as early as possible after a potentially dangerous earthquake is anticipated [5]. These safety measures are equally important as early detection for saving lives. For detection we use IoT platform for communicating the Richter scale reading to the concerned authority [6]. MQTT protocol is extensively used for this application. For carrying out safety measures, we use a microcontroller locally connected with the sensors and actuators [7]. SEWAS (Seismic Early Warning Alert System) is a software which will alert the people in an area about an earthquake as early as possible, so that people can prepare to face the situation and safeguard themselves [8]. Process chart of SEWAS is shown in Fig. 1.

3 Role of IoT in Disaster Management

Managing a disaster is the main part in implementing smart city, as the risk caused by a disaster is unpredictable [9]. Technological advancement is implemented more in other areas of IoT when compared to handling emergency situations. These advancements should be properly implemented after performing research in areas of disaster management. The IoT plays a major role in the step-by-step rescue operations starting from warning to the aftermath of a disaster [10]. Figure 2 shows the disaster management scheme for reducing the risk of disaster.

A disaster causes a catastrophic effect in the normal life. All the processes get affected, which leads to tremendous loss of life and materials. Disaster management is one of the challenging tasks for the government because the intensity cannot be predicted beforehand. Even though we have technologies for predicting and taking

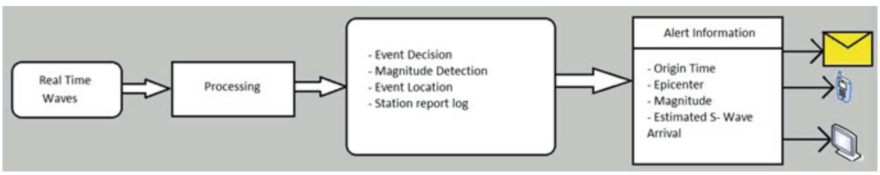


Fig. 1 Process chart of SEWAS

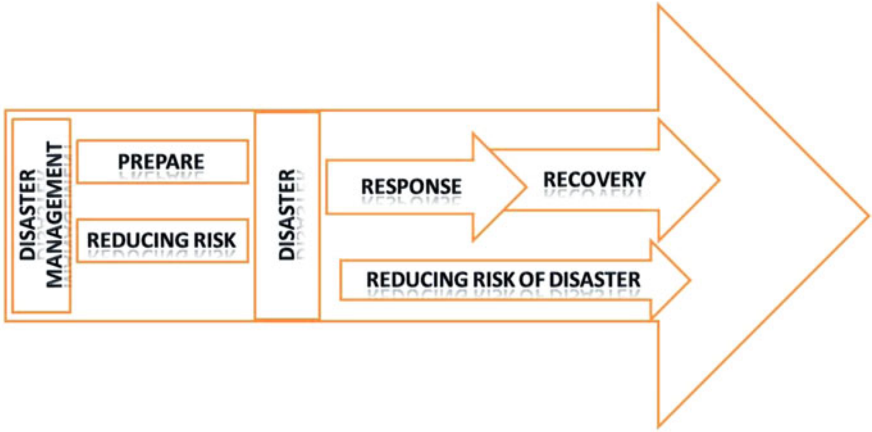


Fig. 2 Disaster management scheme for reducing the risk of disaster [10]

precautionary measures, the aftermath decides the recovery process. We can provide a step-by-step procedure to be followed during a disaster; we can call it as disaster management life cycle. The disaster management starts from preparedness after considering the risk of the terrain and height of buildings to recovering from the disaster [11].

The development of internet of things (IoT) technology makes it feasible for engaging with the headquarters and communicating with the people. Starting from the preparedness to emergency response, IoT plays a major role. Data handling is the main part of using IoT in any application starting from healthcare to on-road assistance in automotive. How this data is handled also differs from application to application. If we take on-road assistance period transfer of diagnostic data from the car to the cloud takes place, whereas healthcare is a critical application, so it should be error-free. Periodicity is important in data handling. Second thing is security; the data should be encrypted while transferring so that the data does not get corrupted in the transit and also interpreted by any middle man. Other factors include storing of data without data loss, accuracy etc. UDA-IoT is a technology which is used for information handling in these types of applications. This technology is used to boost the accessibility of IoT data. As we have said, encrypting the data, collecting it, interpretation and decision-making should be done appropriately [10].

The National Disaster Management Authority (NMDA) is a statistical body formed by the Home Ministry to manage the man-made and natural disasters in the country as a whole. This body also co-ordinates with the state body in managing a disaster. India is prone to more disasters, both man-made and natural – 58.6 % of the area is vulnerable to earthquakes of medium to high intensity; over 40 million hectares (12 % of land) is vulnerable to erosions and floods; in the 7,516 km coastal area, near to 5,700 km is vulnerable to tsunami and cyclone; 68 % of the fertile area is vulnerable to drought, and steep places are vulnerable to avalanches and landslides. Also, India is vulnerable to terrorist attack, where biological and chemical weapons are used as a means of destruction. We are in a position to consider this as a disaster also. The NDMA, as the apex body, has the responsibility of framing policies and guidelines in managing a disaster, directing the state government on how to be prepared and providing them necessary support in times of disaster. Towards this, it has the following responsibilities:

- Drafting policies in disaster management area
- Approving the national plan on disaster management
- Approving plans prepared by the departments of the government
- Laying guidelines to be followed by the state plan on disaster management
- Laying frameworks to be followed by the departments of the Government of India in terms of disaster management and mitigation plans and also supporting the development of safety infrastructures which minimises the effect of the disaster

The Red Cross and Red Crescent Movement is a global humanitarian network of 80 million people which helps those facing disaster and health and social problems [1]. Mankind, unprejudiced nature, unbiasedness, autonomy, intentional help, solidarity and comprehensiveness: these seven fundamental principles give a moral, operational and institutional system to be crafted by the Red Cross and Red Crescent Movement. They are at the centre of its way to deal with helping individuals deprived during intentional clash, catastrophic events and different crises. It consists of:

(a) Committee of the Red Cross
(b) Federation of the Red Cross and Red Crescent Societies
(c) National Red Cross and Red Crescent Societies

Although we have various authorities for managing a disaster, issue comes when they are not properly co-ordinated. This affects the preparedness and timely response. Wherever there is force, managing the force firstly is important. Only then that they can be properly deployed in needed areas. These are the areas where missing the time is like doing nothing.

Figure 3 gives the areas to be concentrated upon when a disaster is occurring and a step-by-step procedure for managing a disaster:

(i) Mobilizing the response team.
(ii) Alerting the medical team.
(iii) Rescue services.
(iv) Alerting the occupants and nearby people.
(v) Making the relief materials ready.
(vi) Rehabilitation and temporary stay should be arranged.

Fig. 3 Areas to be concentrated when a disaster is occurring [10]

3.1 IoT in Management

As we have seen already, constituting an authority or force is the first step process, but that is not sufficient. Co-ordination among them and with the public makes the response an effective one. For co-ordination among them, we are using IoT techniques. Figure 4 gives a co-ordinating strategy from the control station and the rescue personnel. Whenever it comes to managing something, there should be a control in the centre, which acts as a decision-making body and gives instructions. These instructions have utmost power, and the underlying bodies should act accordingly. If control is given to all the sub-sections, then there will be chaos, and efficiency will be drastically reduced. All the units should report to the control station then and there and get the approval.

For disaster management, there should be a proper managing structure. First and foremost, awareness about disasters should be created among the people and how they should respond and safeguard themselves when a disaster occurs. Then comes the early warning: a disaster should be priorly predicted, and it should be communicated with the stakeholders as soon as possible. This helps them to be prepared for early evacuation and staying safe. After this, when a disaster is about to occur, stakeholders should be alerted through loud speakers, mobile phones etc., simultaneously alerting the rescue team, medical team, and ambulances and regulating the traffic in the affected area. Food should be arranged as a first priority for the affected people, followed by sanitation facilities with proper hygiene.

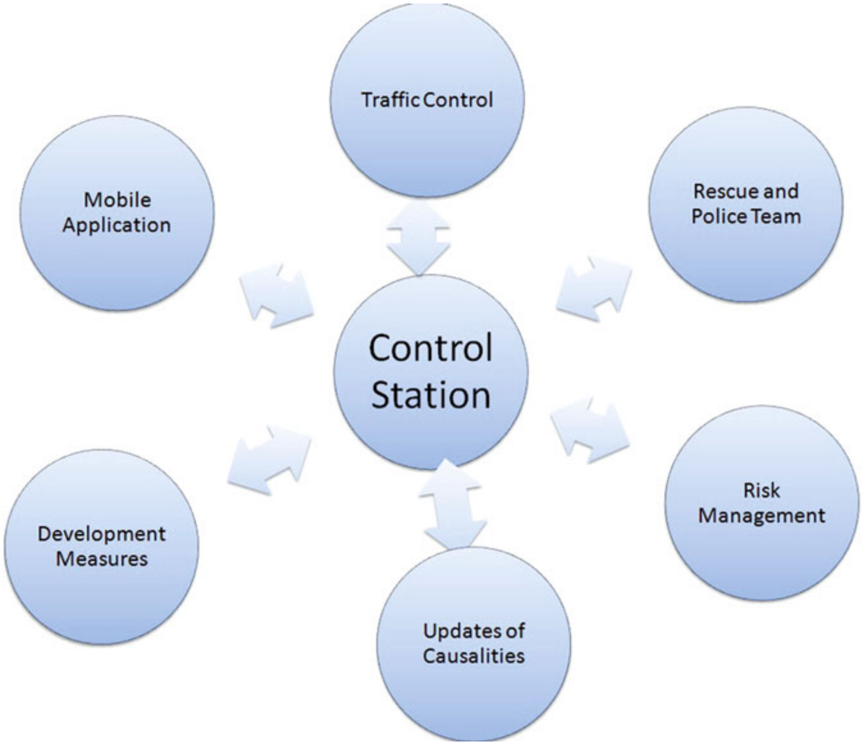


Fig. 4 Bidirectional communication model between the control centre and various departments functioning with disaster management [10]

These rehabilitations should be properly equipped with sufficient resources until the situation turns normal. During this time the control centre should be prepared for reducing the risk caused by the disaster by proper techniques and safeguarding the people quickly. Risk reduction is the main thing to be done in managing a disaster. Relief centres should be made ready correspondingly, to provide a safe stay for the affected people. Police personnel and rescue team should properly communicate with the control centre for updating the status and getting the efficient evacuation plan. Media personnel should be given proper update to let the people know about the status of the disaster. This is how a disaster should be properly communicated and responded by co-ordinating with the control centre. Figure 5 shows the flow of events that are arranged on their priority of occurrence when a disaster occurs.

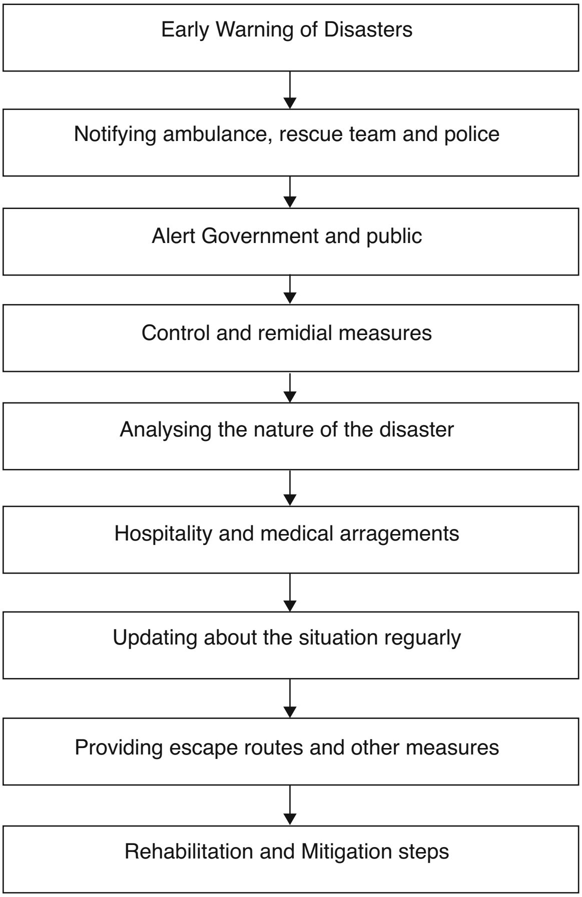


Fig. 5 Flow of events that are arranged on their priority of occurrence when a disaster occurs [10]

4 IoT Architecture for Disaster Management

IoT architecture should be effective enough to handle a disaster, as any man-made technique cannot stop a natural disaster from happening. But it can reduce it with proper strategy; this is based in the architecture with which the disaster is handled. Figure 6 shows the architecture for disaster management.

Let us take the model for recognition of fire in dense forest: sensors will be fixed on trees which will take readings and alert once a fire has broken out or if there is a vigorous danger, for example, temperature, dampness, gas emissions and CO levels. In the event that there is a change in the threshold, early warning, frameworks to alert the local people and their co-operation should be sought to overcome this. The firemen once they arrive should have definite data with respect to the circumstance and reason for the happenings. Distinctive IoT applications are being produced for different types of applications: microwave sensors, which will be

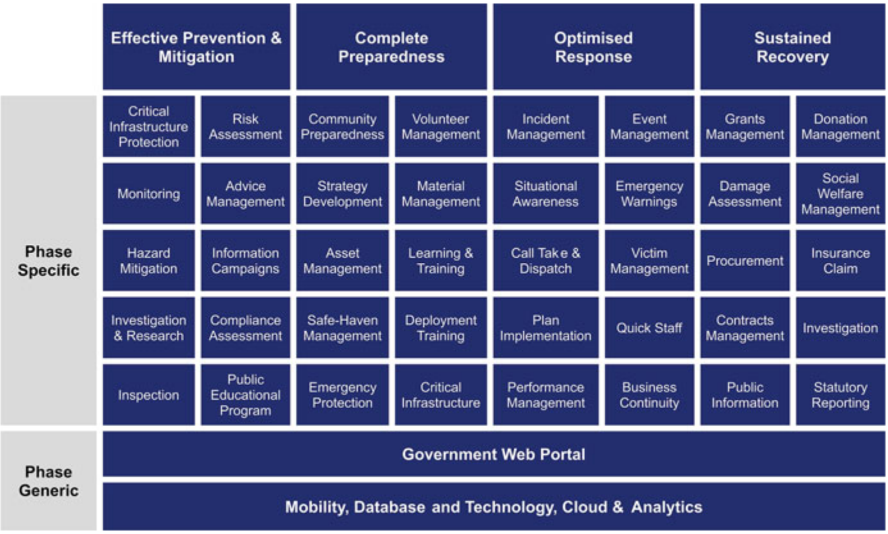


Fig. 6 Architecture for disaster management [11]

utilized to quantify earth movements previously and through tremors, for instance, infrared sensors which will sense the floods and mobility of people. The enormous usage of IoT-empowered gadgets by the people (regularly battery controlled and prepared to work and communicate remotely) might get favourable support in terms of information sharing and organizing the strength. IoT gadgets may pose some restricted communication problems (e.g. crisis small-scale message conveyance) just on the off chance that the standard communication technique is unavailable. To be prepared for basic disasters like avalanches and floods, sensors takes care of data transfer with respect to climate, traffic, police and clinical administrations in proper time. The evacuation procedures and isolation of the area help to speed up the recovery process. A climate and explanation program will anticipate crises 2 days before at least. In the event that such crisis occurs, people and networks are alarmed through online media, standard media channels and SMS. In high-hazard territories, alarms likewise are utilized to empty the disaster location.

5 Proposed Architecture for Earthquake Detection and Evacuation

Figure 7 is the proposed architecture for earthquake detection and evacuation.

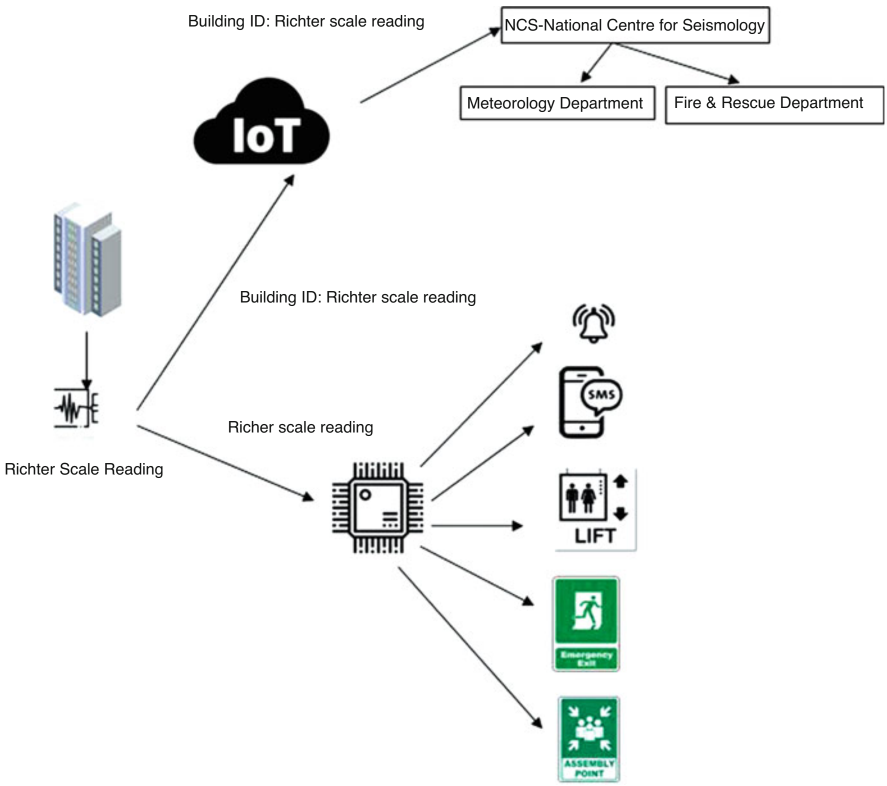


Fig. 7 Proposed architecture block diagram

5.1 Earthquake Detection

Earthquake affects high-rise building tremendously. So, earlier detection in those buildings saves life.

Such natural disasters deleteriously affect the environment, lives of human, etc. which indirectly affect the economy of the country [12]. Identifying earthquakes from background noise and obtaining reliable P-wave first arrival are key technology in earthquake early warning [13]. We have to have sensors in vulnerable points where earthquake may be first detected. This sensor node acts as a node for sensing or may be an embedded system which plays decision-making role also. The control centre will then receive similar data from different sensor systems placed at various locations, and by using the received data from at least three sensor systems, it could locate more accurate location of earthquake in real time [14]. Here, we use IoT platform for earthquake detection, so we will go with sensor nodes alone. These sensor nodes either continuously send signals to cloud or can be an event triggered [15]. Here, we go for continuous sending of data, as this may help the NDRF or the

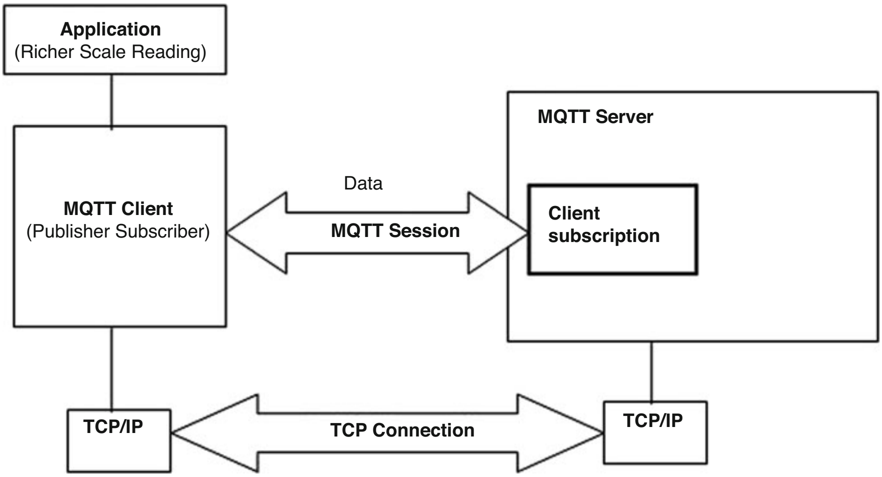


Fig. 8 MQTT session establishment

meteorological department to check the abnormal Richter scale reading. For sending this data, we use MQTT protocol, where readings will be sent as asynchronous messages to the cloud.

MQTT Protocol As discussed, it is extensively used in IoT applications for sending the messages from the sensor node to the actor node. Some of the features of MQTT protocol are as follows:

1. MQTT is a lightweight message queuing and transport protocol.
2. MQTT is suitable for the transport of telemetry data (usually sensor data).
3. MQTT is suitable for M2M (mobile to mobile), WSN (wireless sensor networks) and IoT (internet of things) applications where sensor and actor nodes communicate.
4. Asynchronous communication model.
5. Less complexity, less power and low foot print.
6. Runs on TCP connection.

Figure 8 shows the MQTT session establishment flow chart. It is clear how a MQTT session is established for sending the data – here it is Richter scale reading. First TCP/IP connection has to be established which continues with the MQTT session. After MQTT session is established, we have to discuss how the data flow occurs. Since this comes under the hard-real-time system, the readings are important, so there should be no chances of data loss. Hence, a stream of data flow should be there continuously.

The logical connection between the client and the server is called here as subscription. MQTT protocol supports two types of subscription, transient subscription and durable subscription.

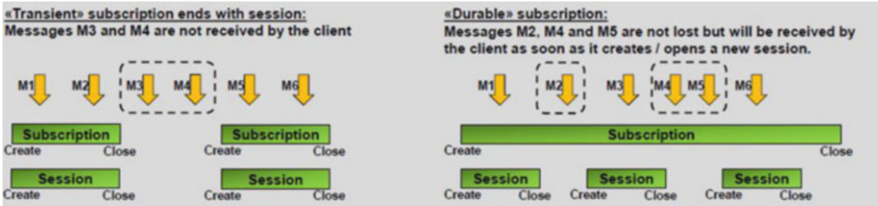


Fig. 9 Transient and durable sessions

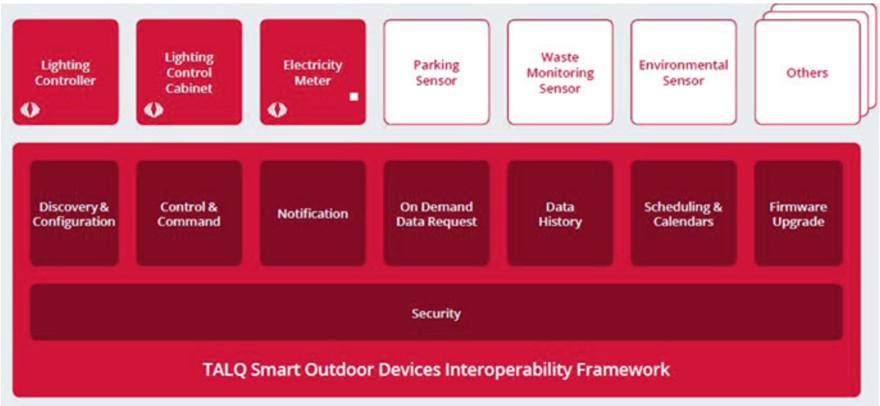


Fig. 10 TALQ protocol layers [16]

Figure 9 shows transient and durable sessions. It is clear that for this application, we have to go with durable subscription, as the message doesn't get lost when a session is closed. Messages has header and footer for transmission purpose, and the payload contains the actual sensor reading.

TALQ Smart City Protocol Specification defines a management interface for outdoor device networks, where a single central management system (CMS) can control different device networks for various applications in different parts of a city or region. It supports system monitoring and joint data collection, as well as simplified configuration and upgrades.

The TALQ Specification defines an application layer protocol that includes message types, data format, parameters and the behaviour of the application end points on the outdoor device network side (called TALQ Bridge) and at the CMS. Figure 10 shows TALQ protocol layers.

TALQ is a dynamic protocol designed for smart city applications, so this also can be used for earthquake detection system. TALQ protocol is not deeply discussed here; it is clearly explained at www.talq-consortium.org [15].

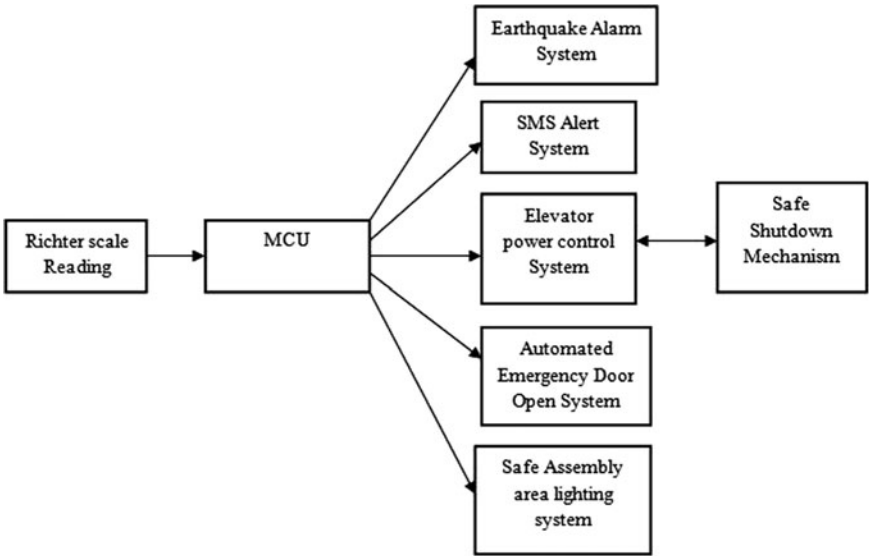


Fig. 11 Safety measures to be taken

5.2 Safety Measures for Evacuation

We will also have a separate controller; the sensors communicate with this controller for decision-making and to start safety procedures for building evacuation. The sensor values are constantly monitored at the microcontroller. This is taken as an input and checks with the threshold. Figure 11 shows the safety measures to be taken.

These safety procedures include:

1. Triggering an earthquake alarm.
2. Lift should go OFF, once the life becomes unoccupied.
3. Emergency exit doors should open.
4. Safe assembly area should be highlighted with visual lights.

The choice of selecting the microcontroller for safety evacuation depends on the building architecture and facilities available. And also, interaction between these systems and the microcontroller depends on the same, either it can be wired or wireless.

Alarm System is an audio-visual system which is used to alert the inhabitants; it can be mounted at places where the flasher will be clearly visible and the sound will be heard as possible. This system should have a single control for the entire building, and it should get activated once the control signal from the MCU is received [17].

SMS Alert System sends an earthquake warning message to all the registered building occupant's mobile phone. The message can also include the Richter scale reading.

Elevator Power Control System is a centralized system for all the elevator systems in the building; power-related controls should be handled by this system. All the other control mechanisms are mounted separately for each elevator in the building. This power system should get control signal from the MCU. Once this control signal is triggered, a separate safety shutdown procedure should be followed. The current elevator occupants should be safely dropped in the nearest floor, and the power should go off.

Automated Emergency Door Open System opens all the emergency exit doors in the buildings. These emergency paths should definitely lead to the safe assembly point.

Safely Assembly Area Lighting System switches ON the hazard lights in the safe assembly point, in order to lead the occupants to it. This will help in finding the easy way to reach the place as the occupants will be in a panic situation; running text can also be displayed [18].

Even if designing all the above-mentioned systems is easy, time plays a major role in this. As this is a hard-real-time system, everything should happen once the MCU triggers a signal. If time is missed out, the entire system goes to waste.

6 Conclusion

Here, we have analysed the possibility of implementing an earthquake alert and building evacuation system. Practical difficulties will arise while implementing the system; however, the precautionary measures will reduce the possibility of a disaster. Here, we have focused on early detection and warning and building evacuation using internet of things. The protocols are basically discussed for suitability. As discussed earlier, time plays a major role here. Detecting and processing the seismic waves to alerting the occupants should happen in considerable time, to make the system reliable. There are various other protocols that better suit smart applications. These protocols are specifically designed for these types of smart city applications. As future work, we can also add other safety measures for implementation when the need arises.

References

1. Allu Suresh Babu, G.T. Naidu, U. Meenakshi, Earth quake detection and alerting using IoT. Int. J. Eng. Sci. Invent. **7**(5 Ver. III), 14–18 (2018)

2. L. Greco, P. Ritrovato, T. Tiropanis, F. Xhafa, IoT and Semantic Web technologies for event-detection in natural disasters. Concurr. Comput. Pract. Exp **30**(21), e4789
3. Srihari, M. M., & Sivakumar, P. (2018). Implementation of multi-function printer for professional institutions, in *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 1–3. IEEE
4. P.N. Crisnapati, P.D. Wulaning, I.N.R. Hendrawan, A.A.K.B. Bandanagara, Earthquake damage intensity scaling system based on raspberry Pi and Arduino Uno, in *The 6th International Conference on Cyber and IT Service Management (CITSM 2018)*
5. F. Kiani, S. Nematzadehmiandoab, A. Seyyedabbasi, Designing a dynamic protocol for real-time industrial internet of things-based applications by efficient management of system resources. Adv. Mech. Eng. **11**(10) (2019). https://doi.org/10.1177/1687814019866062
6. J. Fox, A. Donnellan, L. Doumen, The deployment of an IoT network infrastructure, as a localised regional service, in *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*, (IEEE, Limerick, Ireland, 2019, April), pp. 319–324
7. I. Jawhar, N. Mohamed, J. Al-Jaroodi, Networking architectures and protocols for smart city systems. Journal of Internet Services and Applications **9**(1), 1–16 (2018).
8. R.D. Singh, P. Kumari, P. Singh, B.D.K. Patro, Seismic early warning alert system, in *2014 International Conference on Signal Processing and Integrated Networks (SPIN)*
9. P. Sivakumar, R. Nagaraju, D. Samanta, M. Sivaram, M.N. Hindia, I.S. Amiri, A novel free space communication system using nonlinear InGaAsP microsystem resonators for enabling power-control toward smart cities. Wirel Netw **26**(4), 2317–2328 (2020)
10. J.J. Wellington, P. Ramesh, Role of internet of things in disaster management, in *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, (IEEE, Coimbatore, India, 2017, March), pp. 1–4.
11. H.N. Saha, S. Auddy, S. Pal, S. Kumar, S. Pandey, R. Singh, et al., Disaster management using internet of things, in *2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON)*, (IEEE, Bangkok, Thailand, 2017, August), pp. 81–85.
12. V. Babu, V. Rajan, Flood and earthquake detection and rescue using IoT technology, in *Proceedings of the Fourth International Conference on Communication and Electronics Systems (ICCES 2019) IEEE Conference Record # 45898*; IEEE Xplore ISBN: 978-1-7281-1261-9
13. S.-Y. Li, J.-Q. Lu, J.-D. Song, Q. Ma, Z. Zhao, Study on new technology of earthquake early warning, in *2019 13th Symposium on Piezoelectrcity, Acoustic Waves and Device Applications (SPAWDA)*, (IEEE, Harbin, China, 2019, January), pp. 1–5.
14. Y. Sherki, N. Gaikwad, J. Chandle, A. Kulkarni, Design of real time sensor system for detection and processing of seismic waves for earthquake early warning system, in *2015 International Conference on Power and Advanced Control Engineering (ICPACE)*, (IEEE, Bengaluru, India, 2015, August), pp. 285–289.
15. P. Sethi, S.R. Sarangi, Internet of things: architectures, protocols, and applications, Hindawi J Electr. Comput. Eng. Article ID 9324035, 25 pages. (2017)
16. Talq, The Smart City Protocol, accessed on 30/11/2021 www.talq-consortium.org
17. Zheng, H., Shi, G., Zeng, T., & Li, B. (2011, April). Wireless earthquake alarm design based on MEMS accelerometer. In *2011* International Conference on Consumer Electronics, Communications and Networks (CECNet) (pp. 5481–5484). Xianning, China, IEEE.
18. E. Husni, F. Laumal, The development of an earthquake early warning system using an ADXL335 accelerometer, in *2018 21st Saudi Computer Society National Computer Conference (NCC)*, (IEEE, Riyadh, Saudi Arabia, 2018, April), pp. 1–5.

Impact of the Internet of Things and Clinical Decision Support System in Healthcare

Naveen Kumar, Rajesh Kumar Kaushal, Surya Narayan Panda, and Shanu Bhardwaj

1 Introduction

An Internet connection is a great thing which gives us all the benefits that were not available before. In earlier times, the IoT devices were not advanced. But now these IoT devices are so advanced that a person can read any book, watch any movie, play games, listen to any song, video call any other person or track someone's location. This is just because of connecting things to the Internet [1, 2, 5–15]. The IoT is much a simple concept. It defines connecting the things of world to the Internet. The term IoT was invented by Kelvin Ashton in 1999 [16]. The radio-frequency identification (RFID) group describes the IoT as worldwide network based on standard communication protocols. Some of the vital applications of the IoT are smart city, smart farming, smart homes, smart healthcare, smart grids, industrial internet, smart retail and smart retail.

Healthcare is one of the biggest challenges that everybody is facing in today's world. Smart medical and healthcare are two of the best applications of the IoT [13]. Some of the medical applications of the IoT are fitness programs, elderly care and real-time monitoring. Moreover, the Internet of Things can have various benefits in healthcare like online interaction with the patients, doctor's location, keeping track of patient's health report etc. The real-time monitoring of the patient helps doctors in diagnosing the disease earlier based on the symptoms. The healthcare providers provide treatment and medicines at home which can be seen as an important application through the Internet of Things. These healthcare services reduce the costs by upgrading the quality of our lives and ensuring the right kind of user

N. Kumar · R. K. Kaushal (✉) · S. N. Panda · S. Bhardwaj
Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India
e-mail: naveen.sharma@chitkara.edu.in; rajesh.kaushal@chitkara.edu.in; shanu.bhardwaj@chitkara.edu.in

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_2

experience [9]. This smart healthcare provides proactive healthcare treatment on continuous health monitoring. Depending upon a person's natural, conduct and social attributes, the combined practice of medical services, well-being and patient help is described as personalized healthcare.

This enables every single individual to follow the essential human services guideline 'the appropriate care for the right person at the opportune time' which results in great outcomes [15]. Information created through IoT gadgets helps in successful basic leadership as well as guarantees smooth human services tasks with decreased blunders, waste and framework costs.

Healthcare IoT enables the patients to invest more time with their specialist [11]. The IoT guarantees the personalization of medical services benefits by keeping up advanced character for every patient [2].

In the next few years, the medical sector is observed to be the broad appropriation of the IoT and advance in new e-Health Internet of Things gadgets and applications.

The IoT-enabled gadgets along with wireless sensor area network as shown in Fig. 1 allow to transmit the patient's data to the doctor at a remote location. On the other hand, there are some security issues associated with this technology; particularly when patient's data are transmitted to a remote location, it must be secure. Therefore, to accomplish secure administrations, there is a need to focus on the accompanying security necessities as follows:

1. Confidentiality
2. Self-healing
3. Integrity
4. Resiliency
5. Authentication
6. Availability
7. Authorization

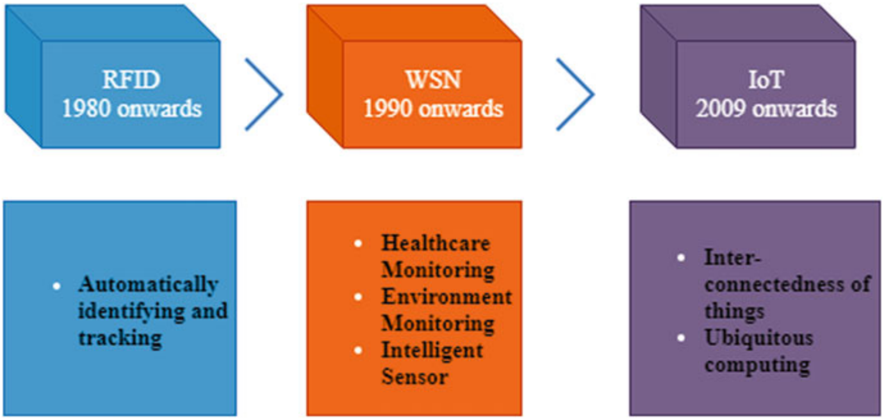


Fig. 1 Impact of IoT in healthcare

Table 1 CDSS and its significant effects in healthcare

Study	Type of preventive care reminders for outcome measures	Significant effect on healthcare practitioner
Barnett et al. [5]	For the measurement of the blood pressure	No change
McDowell et al. [21]	For the influenza medication in a hospital	Improved
McDowell et al. [22]	Reminders for the blood pressure measurement in a hospital	Improved
Burack et al. [7]	Generate reminders for the mammography	Improved
Burack et al. [7]	Reminder for the mammography	Improved
McPhee et al. [23]	Reminders for the cancer prevention	Improved
McPhee et al. [24]	For the cancer prevention	Improved
Rosser et al. [26]	Reminder for the preventive care manoeuvres	Improved

8. Fault toleration
9. Non-repudiation
10. Data freshness

CDSS is proposed to help medical services experts to take diagnostic decisions. The qualities of a patient are coordinated through an information-based framework; along these lines this framework creates explicit suggestions. The literature related to CDSS shows that it can possibly improve clinical exhibitions. Table 1 shows the significant effects of CDSS in healthcare.

2 Literature Review

Doukas and Maglogiannis [8] presented a stage that is dependent on cloud-based system for the board of versatile and wearable body sensors for medical services, showing information on IoT worldview which is connected and not able to avoid human services.

Kulkarni and Sathe [16] attempted to review the use of the IoT in medical services by achieving it at reasonable expenses. The authors clarified in short how IoT capacities and how it is utilized related to remote and detecting the strategies to actualize the ideal healthcare applications.

As far as the IoT and Zigbee technology are concerned, the authors [15] explained the usage of the IoT in medical services framework using Zigbee by using different IoT gadgets to total, investigate and convey the data to the cloud, making it easier to collect, store and break down the huge measure of information in several new forms. The paper proposed a computing-based methodology by collecting all the emergency clinical records from substantial size to small size to keep the patient's record under one roof. This methodology reduces the business model of corporate hospitals but guarantees the administration to run insurance companies for patient benefits.

A survey article on IoT-based advance technologies in healthcare system [13]. The study reviewed network architectures, platforms and applications in healthcare solution. As additional task, this article analysed IoT security along with privacy. The analysis includes security requirements, threat models and attack taxonomies.

In this paper the authors enhanced an existing health monitoring system based on the IoT with the help of fog computing method [10]. It provides a smart gateway of various advanced technologies such as distributed storage, data mining etc. They picked electrocardiogram (ECG) highlight extraction as the contextual analysis as it assumes a critical job in conclusion of numerous cardiovascular illnesses. The outcomes uncover that mist processing helps accomplishing over 90% data transfer capacity productivity and offers low-inertness constant reaction at the edge of the system.

Hossain and Muhammad [12] presented a health IoT-empowered checking structure, where data like ECG and other useful healthcare information are gathered through mobile phone.

Kalarthi [14] proposed a technique which provides the complete remote monitoring system of the patient by transferring the vital parameters. Sensors gather different parameters of the patient's body and exchange that information to a microcontroller known as Arduino Uno. Further Arduino Uno exchanges the information to cloud with the help of Wi-Fi module. The client can see this information with the assistance of Android App. The patient can check their restorative record. Hence, this system provides a quality of healthcare to patients.

Ukil, Bandyoapdhyay, Puri and Pal [29] discussed the healthcare analytics for cardiac disease prevention. The authors also presented the potential advantage of using cell phones in acknowledging moderate and inescapable m-health. They discussed the job of anomaly detection in human services frameworks and to discover the applications and role of anomaly detection to guarantee strong medical services biological system, better guess and survivability which is moderate and has more noteworthy effect.

Abu Naser and El-Najjar [3] had developed an expert system to diagnose nausea and vomiting among children. There is a lack of doctors in India especially in rural areas. Therefore, the clinical decision support system (CDSS) can play a vital role in this process.

Chao Li et al. [19] proposed an IoT-based coronary illness observing framework for inescapable social insurance benefit. This framework screens the patients; physical signs, for example, circulatory strain, ECG, SpO2, and also applicable natural pointers persistently, and gives four distinct information transmission modes that adjust the human services need and requests for correspondence and registering assets. The Internet of Things (IoT) systems have overpowering prevalence in taking care of the issue of heart ailments patients mind as they can change the administration mode into an inescapable way and trigger the social insurance benefit in light of patients; physical status instead of their sentiments. This observing framework satisfies the essential needs of inescapable medical services for heart sicknesses, additionally mulls over the cost to guarantee the unavoidable mode as practical as could reasonably be expected.

Abawajy and Hassan [1] presents IoT-based solution named as pervasive patient health monitoring (PPHM) system infrastructure. The proposed model was an adaptable, versatile and vitality proficient remote patient health monitoring framework.

Sethi and Sahoo [27] have designed an application on health monitoring system (HMS) based on WSNs in real time; the application is continuously monitoring patient's data. Further update and record the data for medical records.

Yang et al. [30] proposed a healthcare storage device. This IoT-based device is a privacy-preserving smart storage system. This study focused on the privacy of patients even in emergency. The device was designed for two approaches, i.e. password-based break-glass access in emergency situations and cross-domain attribute-based access in normal.

Subramaniyaswamy et al. [28] explained about an effective personalized recommendation of a travel recommender system (RS). The system is used to reduce traveling time and cost. ProTrip is entirely designed to be intelligent. The system will recommend whole food-diet plan during travelling. ProTrip is easy to use, and it has user-friendly interface.

Pace et al., also conducted a research on inter-health [25]. The study presents first the results of inter-health directly coming from application of the framework. The system provides less efforts as well as a new assisted healthcare system. It also provides more effective and responsive advice services.

Adhikary et al., conducted a survey on [4] IoT-based healthcare system. The study explained how major advantages of the IoT are increasing rapidly in healthcare system. It provided several applications of the IoT in order to control any healthcare system.

The authors have developed an IoT-based hybrid system which can be used for monitoring and medication purposes [17, 18]. The proposed model allows the doctor/paramedical staff to monitor the vital signs of the patient from a remote location. Remote drug delivery is another unique feature of the proposed model, which allows the doctor to deliver life-saving drugs when the patient is in transit mode.

3 Methodology

An IoT-based solution is developed which can be easily operated by the doctors and caretaker/family members of the patient for monitoring the vital signs of the patient from a remote location. The proposed solution is a three-layer model (hardware layer, application layer and cloud layer) as shown in Fig. 2. The application layer of the model interacts with the hardware layer and the cloud layer. The hardware

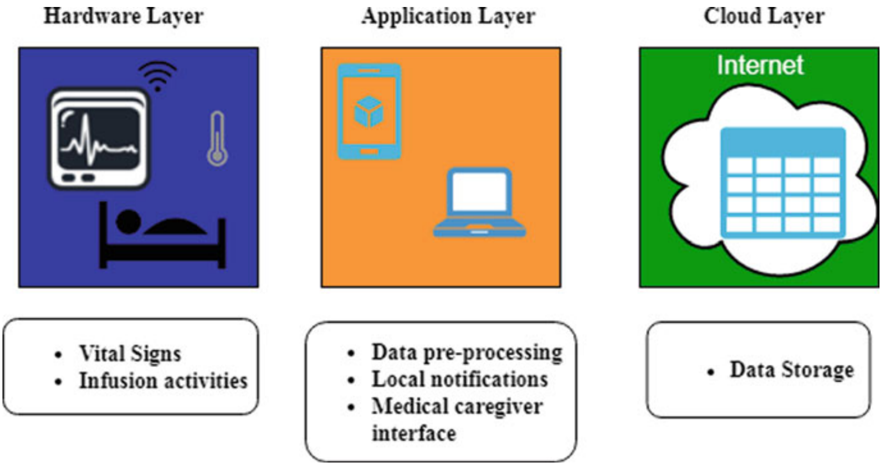


Fig. 2 Three-layer model for data transmission

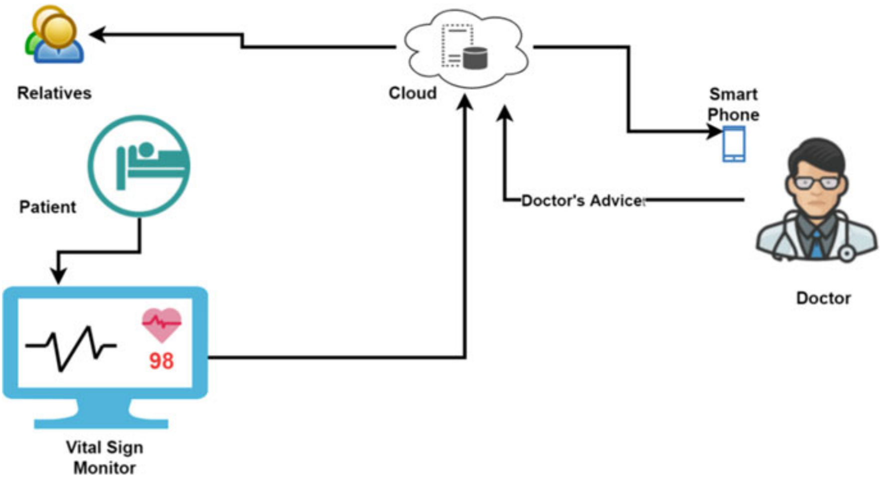


Fig. 3 Architecture model for the real-time data transmission of the patient

layer comprises vital sign monitor. The microprocessor located at the hardware layer pushes the data to the cloud which can be read by the application layer.

We have developed a mobile application to fetch the vital signs of the patient in real time. The architecture model of the proposed solution is given in Fig. 3. All the physiological data like heart rate (HR), body temperature, ECG, SPO2 and non-invasive blood pressure (NIBP) [6] are sent in JSON (Java Script Object Notation) form to the cloud.

We have integrated clinical decision support system for the healthcare professional that generates the alarms as per threshold value of the control variable. The threshold value of the control variable is set by the doctor. All the information is stored in the database for future reference, which can be used by the doctors as a case study. For experiment purpose dataset provided by the Royal Adelaide Hospital is used [20].

According to the provided dataset, patient's information is available in a .xls file. The proposed system generates the alarms if the value of the control variable is beyond the threshold value.

4 Results

IoT-based real-time data transmission helps the doctor to view the patient data anytime from anywhere. The doctor can see the vital signs of the patient in real time on his/her smart phone, which helps to take the diagnostic decision. Referring to Fig. 4, the proposed CDSS model reads the vital parameters of the patient and generates the alarms as per the threshold value of the control variable is set by the doctor. In the present dataset, all the physiological parameters are recorded after an interval of 1 second. A graphical user interface is developed to store the minimum and maximum values of the threshold as shown in Fig. 5.

All the results of the patient are stored in an Excel file which shows the status of the patient as normal/low/high, as well as the type of control variable, and presents status (low/high or normal) as shown in Table 2.

5 Discussion and Future Scope

The literature review suggests that the Internet of Things in the field of healthcare is playing a significant role [17, 18]. This technology is being used by the healthcare professionals to store the physiological parameters of the patient to the cloud. The patient data from the cloud can be accessed anytime from anywhere through the mobile or desktop applications. Thus, the IoT provides the real-time monitoring of the patient even from a remote location. Some studies reveal that IoT-based applications have been developed for the patients who are in transit mode, e.g. in ambulance. The IoT-based applications can be used to regularly monitor the health status of old-age and disabled persons which may suggest a reduction in their ability to live independently. The decision-making facilities in the field of healthcare are improved remarkably due to the significant contribution of the expert system. From the literature review, it has been found that the clinical decision support systems help in taking diagnostic decisions. A knowledge-based system helps to match the characteristics of a patient, and subsequently the system generates specific recommendations. Some rule-based systems have been developed which generate alarms in any unpleasant situation. Therefore, both and CDSS applications are playing a significant role in the field of healthcare.

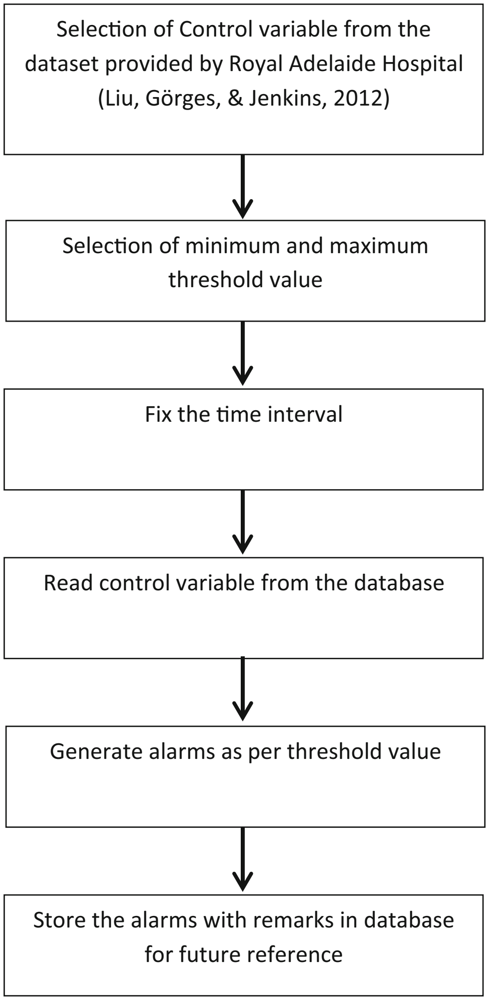


Fig. 4 Flow of the proposed model

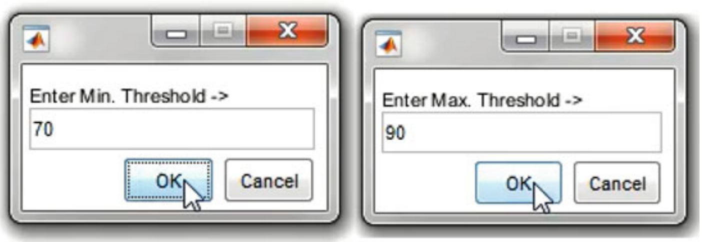


Fig. 5 Interface to set threshold value

Table 2 Status of selected control variable as per threshold value

NIBP(SYS)	NIBP(DIA)	Status
148	95	High
148	95	High
155	98	High
113	60	Low
113	60	Low
113	60	Low
113	60	Low
113	60	Low
113	60	Low
113	60	Low
113	60	Low
113	60	Low
113	60	Low
115	65	Low
115	65	Low
115	66	Low
115	66	Low
116	70	Low
116	70	Low
116	70	Low
116	73	Low
116	73	Low
119	76	Low
120	77	Low
120	77	Low
120	77	Low
121	80	Normal
121	80	Normal
121	80	Normal

The Internet of Things is providing numerous advantages to patients as well as to healthcare professionals. The IoT makes it possible to diagnose a patient from a remote location, and the doctor and relatives of the patient can view the patient's condition anytime from anywhere. Security of patient data is a big challenge in healthcare, so the integrity of the data must be maintained. In the future network technology like software-defined network (SDN) (Bhardwaj and Panda 2019) and virtual private network (VPN) may be integrated to provide a secure and fast network. The literature review recommended that a large amount of data is produced by the sensor which is a difficult assignment with the traditional software. Additionally, technology like big data may be collaborated to deal with the bulky data. Data mining is another technology which may be adopted to find out the hidden factors in the data.

6 Future Scope

The paramedical staff alongside patients and relatives of the patients have been benefited thanks to the use of the IoT in healthcare sector. The IoT also makes it possible to diagnose a patient from a remote location, and the doctor and relatives of the patient can view the patient's condition anytime from anywhere. Security of patient data is a big challenge in healthcare, so the integrity of the data must be maintained. In the future network technology like SDN (software-defined network) and VPN (virtual private network) may be integrated to provide a secure and fast network. The literature review suggested that a large amount of data is generated by the sensor which is a challenging task with the traditional software. Therefore, technology like big data may be collaborated to manage the bulky data. Data mining is another technology which may be adopted to find out the hidden factors in the data. In the future, blockchain technology can be used to deploy security in IoT applications. Blockchain implements security by introducing a distributed database with proof of concept (PoW) algorithm. This distributed ledger is immutable, and any effort to temper the ledger can be tracked through complex hashing algorithms and through consensus over the network.

7 Conclusion

IoT is a mixture of various advancements that connect a varying extent of machines, devices and things that allow people to relate as well as talk to each other using different networking technologies. Up till now, a great part of the data found on the Internet is given by people. In the case of IoT, brilliant objects generate the data. There exist a wide variety of applications that are reliant on the IoT, including healthcare or human services, which is the essential focal point of this work. Restorative healthcare systems use interconnected brilliant devices to develop an IoT network for health analysis by examining and observing the patient's conditions where a specialist's involvement is required. We have integrated CDSS which will help the healthcare staff for better diagnostic decisions.

References

1. J.H. Abawajy, M.M. Hassan, Federated internet of things and cloud computing pervasive patient health monitoring system. IEEE Commun. Mag. ***55***(1), 48–53 (2017)
2. W.A.N.W. Abdullah, N. Yaakob, M.E. Elobaid, M.N.M. Warip, S.A. Yah, Energy-efficient remote healthcare monitoring using IoT: a review of trends and challenges, in *Proceedings of the International Conference on Internet of Things and Cloud Computing, 29* (2016)
3. S.S. Abu Naser, A.E. El-Najjar, An expert system for nausea and vomiting problems in infants and children. International Journal of Medicine Research **1**(2), 114–117 (2016)

4. T. Adhikary, A.D. Jana, A. Chakrabarty, S.K. Jana, The internet of things (iot) augmentation in healthcare: An application analytics, in *International Conference on Intelligent Computing and Communication Technologies*, (Springer, Singapore, 2019, January), pp. 576–583
5. G.O. Barnett, R.N. Winickoff, M.M. Morgan, R.D. Zielstorff, A computer-based monitoring system for follow-up of elevated blood pressure. Med. Care **21**, 400–409 (1983)
6. R.C. Burack, P.A. Gimotty, J. George, W. Stengle, L. Warbasse, A. Moncrease, Promoting screening mammography in inner-city settings: a randomized controlled trial of computerized reminders as a component of a program to facilitate mammography. Med. Care ***32***(6), 609–624 (1994)
7. S. Bhardwaj, S.N. Panda, A study on noninvasive body wearable sensors, in *Intelligent Communication, Control and Devices*, (Springer, Singapore, 2020), pp. 345–351
8. C. Doukas, I. Maglogiannis, Bringing IoT and cloud computing towards pervasive healthcare, in *2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, (2012), pp. 922–926
9. B. Farahani, F. Firouzi, V. Chang, M. Badaroglu, N. Constant, K. Mankodiya, Towards fog-driven IoT eHealth: promises and challenges of IoT in medicine and healthcare. Future Gener. Comput. Syst. ***78***, 659–676 (2018)
10. T.N. Gia, M. Jiang, A.-M. Rahmani, T. Westerlund, P. Liljeberg, H. Tenhunen, Fog computing in healthcare internet of things: a case study on ecg feature extraction. in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, (2015) pp. 356–363
11. M.M. Hassan, H.S. Albakr, H. Al-Dossari, A cloud-assisted internet of things framework for pervasive healthcare in smart city environment. in *Proceedings of the 1st International Workshop on Emerging Multimedia Applications and Services for Smart Cities*, (2014) pp. 9–13
12. M.S. Hossain, G. Muhammad, Cloud-assisted industrial internet of things (iiot)–enabled framework for health monitoring. Comput. Netw. ***101***, 192–202 (2016)
13. S.M.R. Islam, D. Kwak, M.D.H. Kabir, M. Hossain, K.-S. Kwak, The internet of things for health care: a comprehensive survey. IEEE Access ***3***, 678–708 (2015)
14. Z.M. Kalarthi, A review paper on smart health care system using internet of things. Int. J. Res. Eng. Technol. ***5***(3), 8084 (2016)
15. R.K. Kodali, G. Swamy, B. Lakshmi, An implementation of IoT for healthcare. in *2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS)*, (2015) pp. 411–416
16. A. Kulkarni, S. Sathe, Healthcare applications of the Internet of Things: A Review. Int. J. Comput. Sci. Inf. Technol. ***5***(5), 6229–6232 (2014)
17. N. Kumar, S.N. Panda, P. Pradhan, R. Kaushal, IoT based E-critical care unit for patients in-294 transit. Indian J. Public Health Res. Dev. **10**(3), 46–50 (2019)
18. N. Kumar, S. Panda, P. Pradhan, R. Kaushal, IoT based hybrid system for patient monitoring 296 and medication. EAI Endorsed Trans. Pervasive Health Technology **5**(19), 162805 (2019)
19. C. Li, X. Hu, L. Zhang, The IoT-based heart disease monitoring system for pervasive healthcare service. Procedia computer science **112**, 2328–2334 (2017)
20. D. Liu, M. Görges, S.A. Jenkins, University of Queensland vital signs dataset: development of an accessible repository of anesthesia patient monitoring data for research. Anesth. Analg. ***114***(3), 584–589 (2012). https://doi.org/10.1213/ANE.0b013e318241f7c0
21. I. McDowell, C. Newell, W. Rosser, Comparison of three methods of recalling patients for influenza vaccination. Can. Med. Assoc. J **135**(9), 991 (1986)
22. I. McDowell, C. Newell, W. Rosser, A randomized trial of computerized reminders for blood pressure screening in primary care. Med. Care **27**, 297–305 (1989)
23. S.J. McPhee, J.A. Bird, D. Fordham, J.E. Rodnick, E.H. Osborn, Promoting cancer prevention activities by primary care physicians: results of a randomized, controlled trial. JAMA **266**(4), 538–544 (1991)
24. S.J. McPhee, J.A. Bird, C.N.H. Jenkins, D. Fordham, Promoting cancer screening: a randomized, controlled trial of three interventions. Arch. Intern. Med. **149**(8), 1866–1872 (1989)

25. P. Pace et al., INTER-health: An interoperable IoT solution for active and assisted living healthcare services, in *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*, (2019), pp. 81–86. https://doi.org/10.1109/WF-IoT.2019.8767332
26. W.W. Rosser, I. McDowell, C. Newell, Use of reminders for preventive procedures in family medicine. Can. Med. Assoc. J. ***145***(7), 807 (1991)
27. S. Sethi, R.K. Sahoo, Design of WSN in real time application of health monitoring system, in *Virtual and Mobile Healthcare: Breakthroughs in Research and Practice*, (IGI Global, 2020), pp. 643–658
28. V. Subramaniyaswamy, G. Manogaran, R. Logesh, V. Vijayakumar, N. Chilamkurti, D. Malathi, N. Senthilselvan, An ontology-driven personalized food recommendation in IoT-based healthcare system. The Journal of Supercomputing **75**(6), 3184–3216 (2019)
29. A. Ukil, S. Bandyoapdhyay, C. Puri, A. Pal, Iot healthcare analytics: the importance of anomaly detection, in *2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA)* (2016), pp. 994–997
30. Y. Yang, X. Zheng, W. Guo, X. Liu, V. Chang, Privacy-preserving smart IoT-based healthcare big data storage and self-adaptive access control system. Inf. Sci. **479**, 567–592 (2019)

Smart Healthcare Support Using Data Mining and Machine Learning

Theodora Chatzinikolaou, Eleni Vogiatzi, Anestis Kousis, and Christos Tjortjis

1 Introduction

The last few decades have observed an unprecedented trend of the people moving to live in metropolitan areas as cities accelerate time by compressing space [1]. The year 2008 was a milestone that, for the first time, the urban population has surpassed that of rural areas, and it is foreseen that until 2050, two-thirds of the world population will be urban inhabitants [2, 3]. This trend of people moving to cities is creating enormous pressure on the city infrastructure [4].

The smart city is an innovation of the physical city, aiming to provide real-time, interactive, and intelligent city services to end users with a high integration of advanced monitoring, sensing, communication, and control technologies, facilitated by the Internet of Things (IoT). A smart city rests on people, technology, and processes and how they connect with various domains including governance, healthcare, education, transportation, energy, environment, tourism, public safety, and buildings [5].

Admittedly healthcare is an essential part of our lives. With the contribution of technology, we are moving towards e-health and smart healthcare. Advancements in big data mining (DM) can give us insights into how people go about their lives [6]. Using DM techniques, raw information is extracted and used by several medical applications, such as decease prediction, treatment assistance, cost-effectiveness, hospital administration and management, etc. [7]. Doctors, patients, hospitals, research institutions, governments, and local authorities are some of the major

T. Chatzinikolaou · E. Vogiatzi · A. Kousis · C. Tjortjis (✉)
The Data Mining and Analytics research group, School of Science and Technology, International Hellenic University, Thermi, Greece
e-mail: t.chatzinikolaou@ihu.edu.gr; e.vogiatzi@ihu.edu.gr; a.kousis@ihu.edu.gr; c.tjortjis@ihu.edu.gr

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_3

stakeholders in smart healthcare. Alongside, information and communication technologies, such as IoT, mobile Internet, 5G, microelectronics, cloud computing, big data, and artificial intelligence (AI), together with modern biotechnology, constitute the cornerstone of smart healthcare [8].

A further consensus regarding the healthcare sector is that it is "information rich," as there is a considerable amount of digital data available [9]. Hence, with the use of all previously mentioned technologies, essential patterns and smart systems can be developed and integrated to support healthcare. The concept of smart healthcare involves an intelligent system supported by all the beforementioned technologies that gather constantly data, communicates information among all stakeholders, organizes resources and manpower among institutions, and finally responds to patient needs in a personalized manner [8]. It entails the automation of processes related to hospital administration, medical assistance, drug research, etc. that optimize "administration of health services, clinical care, medical analysis, and training" [10].

Based on the above, as Fig. 1 depicts, it becomes apparent that smart healthcare system requires participation and engagement of several layers of stakeholders to provide accurate and effective information as well as enhance public healthcare [11].

In this chapter, we analyze the basic elements of smart healthcare support using DM and machine learning (ML). We demonstrate the use of DM in smart healthcare applications covering all the abovementioned sectors: DM techniques and applications for smart healthcare. More specifically, we focus on two research questions:

Q1. Which are the major smart healthcare applications?
Q2. Which DM techniques are used with healthcare data?

The remaining of the chapter is organized as follows: Section 2 introduces key concepts and background on smart cities. Sections 3 and 4 review smart healthcare applications and DM techniques, respectively. Section 5 details a case study on diabetes, and Sect. 6 concludes the chapter with open research questions.

2 Smart Cities

The smart city concept implies an intelligent platform consisting of interconnected sensors, embedded devices, and decision-making systems that process real-time data [12]. According to Rodriguez et al. [13], the purpose of a smart city is "to achieve efficient management in all areas of the city while satisfying their needs and the one of its citizens. At the same time, it must be aligned with the principles of sustainable development and taking technological innovation and cooperation between economic and social agents as the main drivers of change."

Cities are like dynamic living organisms and constantly evolving [14]. To monitor these changes, hard (dedicated) sensing is the primary sensing paradigm, as it can be tailored to precisely meet the smart city application needs. On the

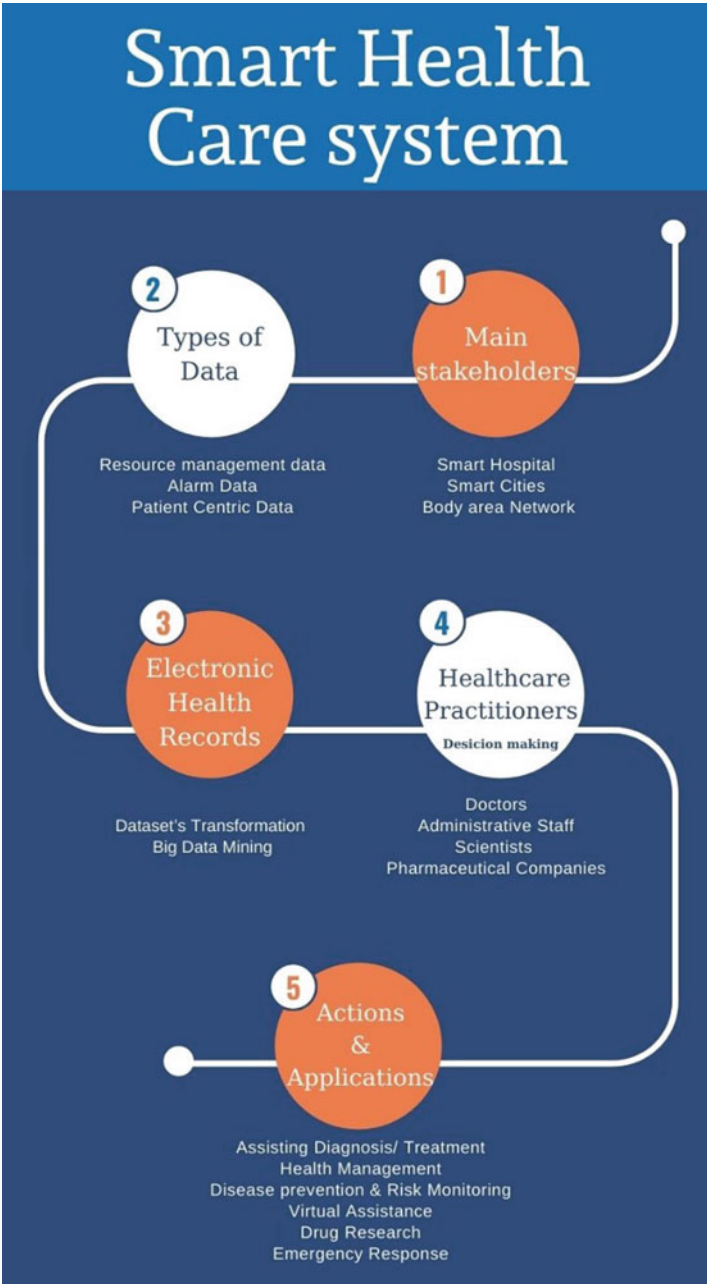


Fig. 1 Smart healthcare system, participants, and required procedures

other hand, soft (non-dedicated) sensing includes various smartphone capabilities and sensors [15]. The use of sensors can help identify "what" is occurring, but unable to detect "why" and "how" such an event happens [16]. As social networks or, alternatively, social media are becoming increasingly popular [17], Souza et al. [16] suggest using social media to capture the human perception of incidents. Big DM and ML utilization assists in solving several problems concerning citizens [18].

Smart cities apply to several aspects of urban life. Mohanty et al. [19] list nine components of a smart city: smart infrastructure, smart buildings, smart transportation, smart energy, smart healthcare, smart technology, smart governance, smart education, and smart citizens.

According to Townsend [1], the overall architecture of a smart city has a hierarchical structure of three layers, as shown in Fig. 2:

1. The "instrumentation" layer. The distributed sensor grid is embedded in infrastructure that measures conditions throughout the city for acquiring real-time environmental and social data including images, video, sound, temperature,

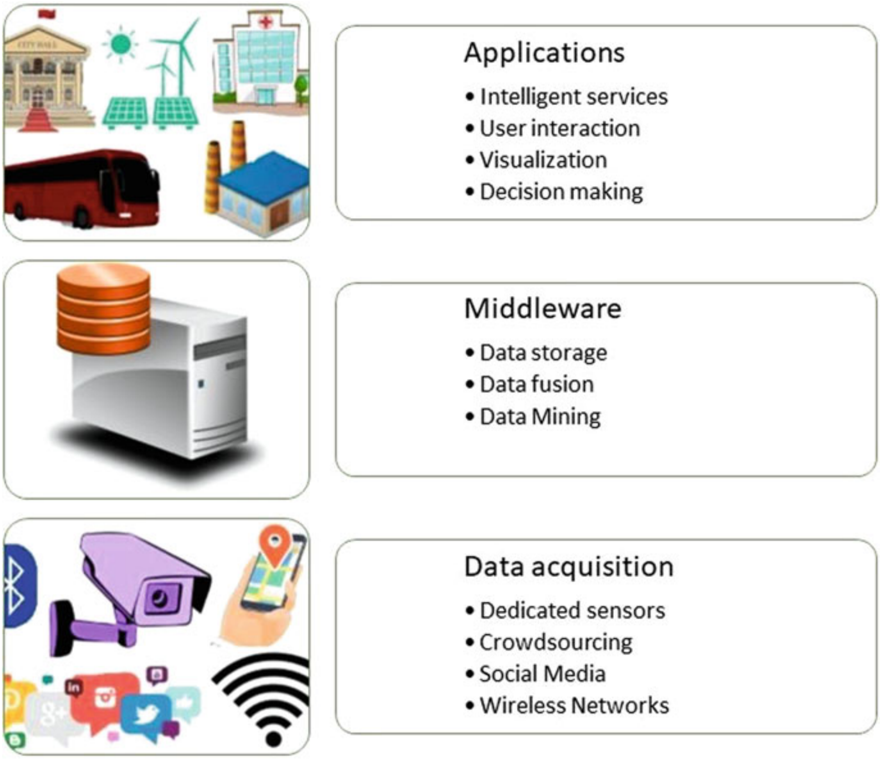


Fig. 2 The smart city layers

humidity, pressure, etc. [20]. This layer is responsible for collecting and locally storing external data.

2. The service-oriented middleware layer combines data-crunching hardware and software for massive data storage, real-time analysis, and processing [21]. It is based on cloud computing, DM, and highly efficient index services that let us discover patterns, visualize information, and make better decisions. The results can be used to support effective operation of smart city applications.
3. The application layer for end users applies a set of management practices and tailored intelligence services to different domains, and it is responsible for interacting directly with the user [22]. It facilitates the interaction with the system and provides the user with information in a comprehensible way, such as graphical form, tables, or other types of presentation assisting usage of the computation results.

3 Smart Healthcare Applications

Smart healthcare services can be grouped into three categories: clinical/scientific research institutions (e.g., hospitals), regional health decision-making institutions, and individual or family users. We focus on and analyze the most common services in which smart healthcare is applied.

3.1 Assisting Diagnosis and Treatment

One use of smart healthcare is to assist diagnosis and treatment of patients. Continuous patient monitoring using a body area network (BAN) [23] which integrates implanted or wearable sensors provides recognition of dangerous situations for people suffering of cognitive [24] or physical disabilities [25]. With the help of AI, robots, and mixed reality, diagnosis and treatment become more intelligent. Nowadays, illnesses like hepatitis, lung or breast cancer, and melanoma [26] can be diagnosed and treated in a more accurate way [27, 28] [29]. At this point, accuracy exceeds a doctor's accuracy and results, especially when the matter of discussion is related to “pathology or imaging” [30].

Systems that provide that kind of assistance are considered clinical decision support systems (DSS). IBM's Watson[1] is an outstanding and representative product in the field of clinical DSS. This is an intelligent cognitive system that provides an optimal solution through in-depth analysis of all clinical and literature data. This system helps doctors diagnose diabetes and cancer with more accuracy based on

[1] https://www.ibm.com/watson

certain algorithms. In this way, patients receive the right treatment for each case, and fault diagnoses or missed diagnoses are avoided.

Personalized treatment for each patient can be developed with smart diagnosis [31]. As every case is different requiring a different approach, smart diagnosis provides more precise treatment. For example, surgical robots have raised surgery to a new level. With the most common robot systems, such as the Da Vinci System, Sensei X Robotic Catheter System, and Flex® Robotic System [8], patients have better results and recover faster, and surgeons elaborate with equipment that provide greater flexibility and compatibility.

Another important aspect of diagnosis and treatment using DM is patient mood throughout the recovery period. Using DM, healthcare institutes can measure and analyze patient behavior and discover patterns, needs, preferences, and quality to establish a better relationship with them [6]. This, combined with the personalized treatment that DM can offer nowadays, results in better therapeutic conditions for a patient, which at the same time is cost-effective and appropriate. It is important for a patient to feel encouraged and satisfied and to receive the right messages at the right time.

Apart from more accurate diagnosis and personalized clinical as well as psychological treatment, DM techniques can effectively compare patients with the same diseases who take different drugs. Doctors can publicize practices they use and compare them with their colleagues'. Treatments for basic diseases can be standardized because side effects for each treatment can be more easily spotted [32, 33].

3.2 Health Management

The increasing number of patients having more or chronic diseases made traditional health management systems inadequate for supporting health systems and patients. Using DM, managing health systems has improved and become cost-effective. Now, health management is more focused on patient self-management with the use of wearable smart devices, smart homes, and smart health information platforms connected by IoT technology. Wearable/implantable devices can combine propelled sensors, microchips, and remote modules to constantly detect and monitor different physiological indicators of patients in a smart way while reducing power utilization, improving comfort, and permitting the information to be joined with health data from different channels. This kind of application reduced the associated risks caused by the disease while making it easier for medical institutions to monitor the prognosis [34].

An outstanding use of health management application is the "creation" of **smart homes**. A smart home is a residence where smart devices monitor space using sensors and actuators checking out the residents' physical signs and environment while performing operations that improve quality of life. Smart home services belong to two categories: **home automation** and **health monitoring**. These technologies provide some comfort while collecting health data, making people who need monitoring and health help more independent, but also safer; their lives also improve. For example, these kinds of applications are perfect for elderly people living on their own, because on one hand, they can have a simpler life and self-monitor their health but, on the other, doctors can be notified and act on time if needed [35, 36].

Alongside with smart homes, hospitals also need to be smart. Smart hospitals use ICT-based environments, IoT optimization, and automated processes, to improve existing patient care procedures and introduce new features [37]. As indicated in Fig. 1, there are three services for smart hospitals: services for medical staff, patients, and administrative staff. Hospital management decisions must take into consideration demands from all three groups of people. Digital devices, intelligent buildings, and personnel are connected through the smart hospital system. This system can also be used for:

- Patient identification and monitoring in hospitals, medical staff daily management, and instruments and biological specimens tracking [8]
- Tracking hospital stock in material (pills, masks, gloves, etc.) [38]
- Decision-making (resource allocation, quality analysis, and performance analysis, medical costs reduction, maximization of resource utilization, decisions regarding development) [39]
- Patient access to multiple functions (physical examination systems, online appointments, and doctor-patient interactions) [40]

In general integration, refinement, and automation are the future directions of smart hospitals.

Many institutes have tried to evolve smart health management over time. For example, the Arkansas Data Network, trying to develop better diagnosis and treatment protocols, focuses on readmission and resource utilization and compares its data with current scientific literature to determine the best treatment options, thus using evidence to support medical care. The Group Health Cooperative, in order to determine which patient groups use the most resources, classified its patient populations by demographic characteristics and medical conditions to develop programs to help educate these populations and prevent or manage their conditions [32]. The Lightweight Epidemiological Advanced Detection Emergency Response System (LEADERS) uses DM to analyze massive data volumes and statistics to search for patterns that might indicate an attack by bio-terrorists [41]. With this use of DM, automated early-warning systems for epidemic incidents can be created, like the one for the SARS virus in the past or COVID-19.

3.3 Disease Prevention and Risk Monitoring

Traditional disease prevention and risk techniques collect patient information, compare that with the guidelines of an authoritative organization, and finally release the prediction results. Such techniques have limitations in time and do not provide accurate advice to individuals. New and smarter disease risk techniques are dynamic and personalized. Doctors and patients participate, proactively monitor disease risk, and conduct targeted prevention based on results. Using wearable devices and smart apps, patient data are collected and stored in the cloud. Then the new disease prevention and risk system analyzes the results based on big data-based algorithms and feedback the predicted results to users in real time via short message service. These systems have been proven to be more effective than the traditional ones. They help doctors and patients to build their lifestyles according to real medical indicators and support decision-makers building regional health strategies to achieve the goal of reducing disease risk [42].

A well-known example of a disease prevention and risk monitoring system is the one that American Healthways (AH) has created. AH constructed a predictive system to identify high-risk patients for diabetes. AH's main concern was to provide a system in which diabetic patients could have improved health quality and lower treatment cost [43]. Brisimi et al. [44] focused on the two leading clusters of chronic diseases, heart disease and diabetes, to develop data-driven methods for hospitalization prediction based on patient medical history, as described in their electronic health records (EHR).

Mitigating health risks can be achieved using DM predictive methods over personal health-related and demographic data. Examples include, but are not limited to, childhood obesity prediction [45] and stroke prediction [46].

3.4 Virtual Assistant and Wearable Sensors

The concept of virtual assistance refers to algorithms that interact and communicate with users through different technologies. For instance, voice recognition with the use of big data communicates with users to address their needs [8].

Virtual assistant benefits the healthcare sector, as it provides on-time and easier communication among all parties involved. Patients contact doctors and communicate their health issues in a convenient manner, while doctors and hospitals have the advantage of managing patient and resource data more accurately.

On this premise, recently advanced technologies for wearable sensors combined with virtual assistants have been used as real enablers of human behaviors. Such technologies, using DM techniques, have been developed to gather data and extract meaningful information. The healthcare sector leverages these technologies to improve health monitoring in order to perform not only tasks like pattern recognition, but further provide personalized services to end users. Applications of wearable sensors involve, among others, health monitoring systems to aggregate medical informatics and reinforce ambient assisted living (AAL) [47].

Specifically, sensors can be used for patient and individual's home environments [6, 48, 49] as well as clinical environments. Patients can achieve independence and increase their sense of security with the use of sensors in their own environment.

However, hospitals and clinics leverage sensors and DM techniques for more precise decision-making. DM tasks such as prediction, anomaly detection, and diagnosis are applied to online and offline data, to facilitate accurate prediction and diagnosis that advance both medical services offered by clinical settings and ensure a healthy independent living. For instance, classification is applied to home monitoring systems to identify abnormal events and help medical practitioners intervene and make accurate and on-time diagnosis [47].

3.5 Drug Research

Traditional drug research and development involves time-consuming and thorough processes, such as manual target screening and clinical trials. Thus, smart and automated screening has become a prerequisite for direct and precise drug research and discovery. Big data and AI assist extensive virtual drug screening, to aggregate on timely and meaningful information. Scientists further analyze the aggregated information to make credible predictions of compounds and conclude on safe drug discovery [8].

Another major and costly drug research issue is the detection of adverse drug reactions (ADRs). As previously mentioned, DM was applied to healthcare data to detect relationships between drug and potential ADRs [50]. This application is of significant importance for pharmaceutical companies, medical practitioners, and scientists to construct credible drug recommendation systems and precisely predict and prevent patients from unexpected and menacing side effects.

4 Data Mining Techniques

Daily processes and transactions in healthcare systems, as medical records and administrative reports, generate massive collections of digitized raw data, which is an essential advantage for knowledge extraction [51]. Data collections involve "patient centric data, resource management data and transformed data" [9]. Patients, doctors, hospital administration, and pharmaceutical companies can leverage this information and support medical processes. Along this line, all parties involved can optimize decision-making regarding, among others, patient treatment, hospital management, and cost-saving [52].

DM, as previously mentioned, analyzes large amounts of data to extract patterns previously unknown [53]. Some specific DM benefits for healthcare are enhanced treatments, affordable treatment prices, accurate and early disease prediction, accurate drug prescription, prediction of duration of hospital stay, efficient task allocation among medical practitioners, etc. [54]. Hence, DM facilitates better medical services and provides a wide range of helpful applications to healthcare practitioners and stakeholders.

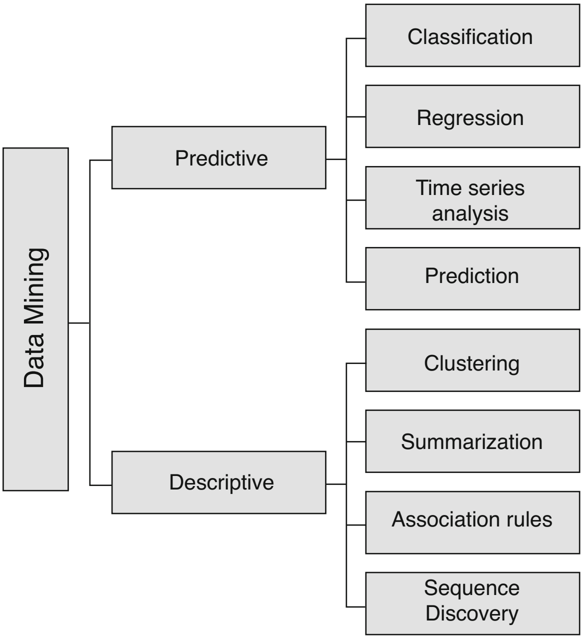


Fig. 3 Some of the most popular DM tasks [52]

As definitions state, DM “assists to identify, valid, novel, potentially useful, and understandable correlations and patterns in data” [55]. The highly referenced “CRISP-DM” methodology divides the DM process into six steps [56]:

- Business understanding
- Data preparation
- Modeling
- Evaluation
- Deployment

In the modeling phase, DM tasks and ML algorithms are implemented. There are two main models of DM, predictive and descriptive. Predictive models use supervised learning functions to predict previously unknown values and future incidents based on historical data [50], while descriptive models use unsupervised functions to explore datasets and create patterns with similar values [50].

A supervised learning algorithm creates a function that connects input points with outputs. It is designed to learn from a training dataset, which consists of predetermined classes. Thus, new data of the test set are classified based on the training set. On the other hand, in unsupervised learning there are not predefined class labels. However, the model forms groups of values based on observations, similarities, statistical properties, etc.

As Fig. 3 indicates, some of the most popular DM tasks are classification, clustering, and association rules [52]. These tasks and their application in healthcare system are briefly explained below.

4.1 Classification

Classification is the most widely used DM task in healthcare [54]. It is a supervised learning task that discovers a learning model to classify a record into one of the predefined classes, based on a function of values of other attributes [51]. Classifier outcome can be binary or multilevel. For instance, a patient, in relation to his/her age, weight, and smoking behavior, can be classified as "low-risk" or "high-risk" patient in the binary case, while in the multilevel classification there are more than two classes, such as "low," "medium," and "high." For instance, classification can be used to connect symptoms with recorded disease cases to enhance decision-making for patients with yet unknown diseases diagnosis [57].

Islam et al. [50] claim that classification techniques are the most dominant for analyzing health data and widely used in the literature for clinical decision support and healthcare administration. Some of the best known classification algorithms are K-nearest neighbor (k-NN) [58], decision trees (DT) [59], support vector machines (SVM) [23, 44], and Bayesian methods [48]. The literature aggregates several examples of classification techniques applied to healthcare data, for instance, the development of several disease diagnosis models, such as the widely cited experiment by Potter et al., trying to diagnose the highly fatal breast cancer disease, using Weka[2] [54]. S. Mohapatra et al. [60] aimed to produce a smart healthcare system for liver, kidney, and heart disease detection using K-NN. Pooja et al. [10] propose the "Smart E-Health Prediction System." In this application, the patient provides information about his/her health status, and with the use of a Bayesian classifier, the system predicts the disease and recommends a doctor in the vicinity. DT classifiers are also used in wearable sensors. Applications include the prediction of negative or positive emotions with the use of motion recognition or heat stress risk considering skin and environmental temperature [47].

4.2 Clustering

In contrast with classification, clustering is an unsupervised learning method, as it has no predefined classes [54]. Its goal is to group data points from large databases into clusters, based on some similarity/distance measure [61]. Data points are similar within the same cluster and are less similar with data points from different clusters [62]. Clustering can be used in healthcare data to form groups of diseases or patients with similar symptoms, to better profile them. Furthermore, the literature states that as clustering requires less information, it can effectively assist exploratory studies with large amounts of data, for instance, microarray analysis, because of the

[2] https://www.cs.waikato.ac.nz/ml/weka/

limited information about genes [54, 61]. Examples of clustering methods are K-means, K-medoid, hierarchical clustering, and density-based clustering.

Towards a smart healthcare perspective, Kelati et al. [63] make use of K-means to identify irregular behaviors that imply health problems. With the use of a smart meter, they gathered energy consumption data and tried to identify abnormal energy usage that indicates health problems. Specifically, in their experiment it is assumed that extended energy consumption in night hours implicate Alzheimer disease, while limited energy consumption relates to sickness and depression. Household data were grouped in five clusters, and clusters with the highest abnormalities were further analyzed [63].

4.3 Association Rules

Association rules is an essential DM method that has gained popularity after the inception of the Apriori algorithm [64]. It is used to find out frequent patterns and meaningful relationships in a data itemset [65, 66]. It is commonly referred as market basket analysis, as it creates probabilistic statements among purchased items [61]. In the case of healthcare, it is used to find relationships among diseases, medical condition, and symptoms. Its most common application is for identifying essential relationships between diseases, recommended drugs, and patient reactions [50, 54, 61]. Several studies have further used association rules for healthcare data to discover fraud and abuse. Such an example is the identification of fake prescriptions and medical patterns that patients, insurers, and care providers may take advantage of. Another well-known algorithm for association rules mining is FP growth [67], while several heuristic methods have been also proposed [68].

An association rules mining framework was proposed by Ji et al. in [69], based on a fuzzy recognition-prime decision (RFD) model, to identify on real-time electronic health data the relationship between a drug and its potential ADRs.

Table 1 summarizes key data mining tasks and their main applications in healthcare.

Table 1 Data mining tasks and their main applications in healthcare

Task	Techniques	Healthcare applications	Authors
Classification	k-NN, DT, SVM	Clinical Decision Support	[50]
Clustering	K-Means, Hierarchical & Density Based Clustering	Microarray analysis	[54, 61]
Association rules	Apriori, FP-Growth	Drug Research	[50, 54, 61]

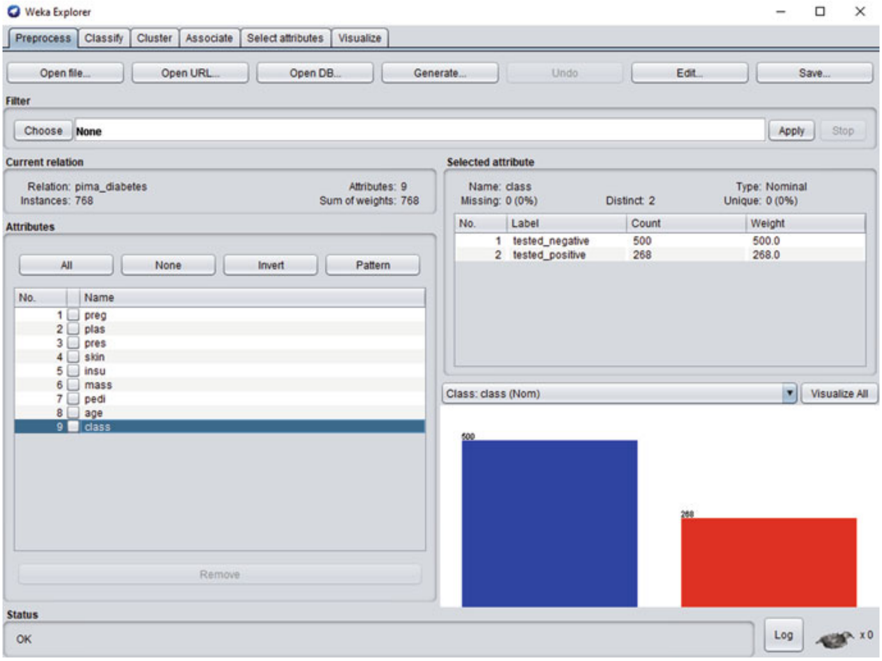


Fig. 4 Datase representation in Weka

5 A Case Study

For better interpretation and analysis of the use of DM in smart healthcare, we examined a publicly available UCI Repository dataset [70], which contains information related to diabetes. As depicted in Fig. 4, the dataset consists of 768 instances and 9 attributes. There are 8 numeric predicting attributes – “Preg,” “Plass,” “Pres,” “Skin,” “Insu,” “Mass,” “Pedi,” and “Age” (corresponding respectively to: (1) number of times pregnant, (2) plasma glucose concentration, (3) diastolic blood pressure, (4) triceps skinfold thickness, (5) 2-h serum insulin, (6) body mass index, (7) diabetes pedigree function, (8) age) – and the binary class attribute with two values: tested negative and tested positive.

Our aim was to apply DM and ML techniques, specifically classification and clustering, using Weka [71], to extract meaningful information from the abovementioned dataset.

As our dataset contains healthcare data, accuracy is of high importance. As we observed after pre-processing, like attribute removal and discretization, classifiers get more accurate and produce trees or rules easier to comprehend, like the tree in Fig. 5. Furthermore, all algorithms produce results in less than 1 s, which is also an important evaluation criterion, as healthcare experts need accurate and timely information. However, top classification accuracy was less than 80%.

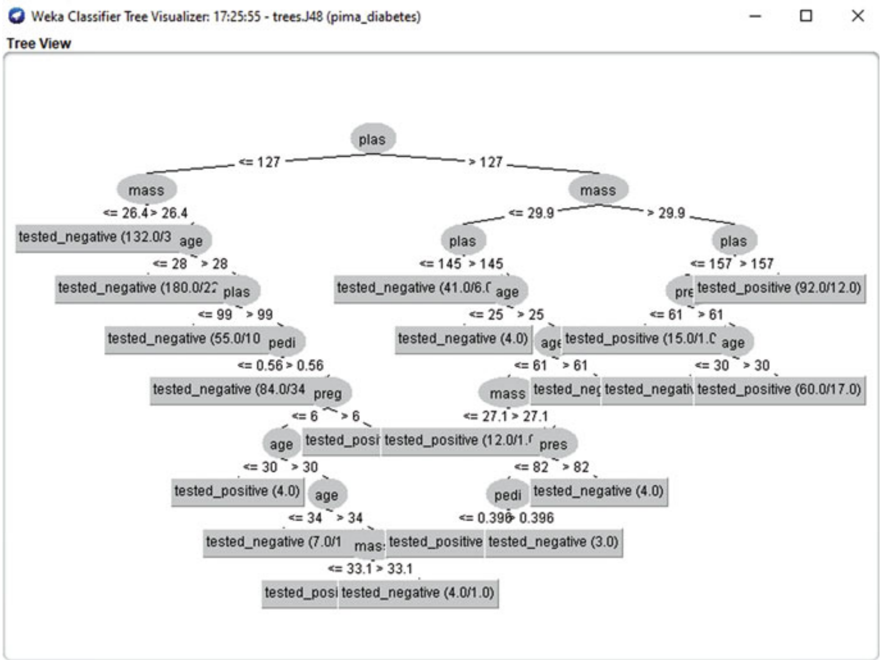


Fig. 5 Visualization of classifier tree with the use of trees. J48 classifier

As previously mentioned, the dataset is related to diabetes, and the preferable outcome is to predict as accurately as possible diabetes cases but also extract meaningful knowledge about patient characteristics for positively or negatively tested for diabetes.

The highest accuracy (77.99%, Fig. 6) was achieved by the *Multiclass Classifier Updatable*, a metaclassifier for handling multiclass datasets with 2-class classifiers. However, C4.5 (named *J48 in Weka*) generated an understandable decision tree shown in Fig. 5, alas with lower accuracy (73.83%). To provide useful and conducive information, we further examined the importance of attributes and their contribution to results.

As observed in Fig. 5, the decision tree does not include the "skin" and "Insu" attributes. Thus, these two attributes provide the minimum information gain. We removed those two attributes to reduce the volume of the dataset, as they do not significantly affect the analysis. As seen in Fig. 7, the result was a slightly improved accuracy at 75.13%.

Furthermore, as attributes are continuous, we performed discretization (dividing continuous values with equal frequency). After experimenting, we observed that if we discretize attributes with the use of Weka filters, to partition them into six bins with equal frequency, we can accomplish both an increase in accuracy and a decrease in tree size.

```
Time taken to build model: 0.06 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         599               77.9948 %
Incorrectly Classified Instances       169               22.0052 %
Kappa statistic                          0.4868
Mean absolute error                      0.2201
Root mean squared error                  0.4691
Relative absolute error                 48.4156 %
Root relative squared error             98.4167 %
Total Number of Instances              768

=== Detailed Accuracy By Class ===
```

Fig. 6 Meta-multiclass classifier updatable algorithm

```
Number of Leaves  :     20

Size of the tree :      39


Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         577               75.1302 %
Incorrectly Classified Instances       191               24.8698 %
Kappa statistic                          0.4435
Mean absolute error                      0.3069
Root mean squared error                  0.4298
Relative absolute error                 67.5185 %
Root relative squared error             90.1755 %
Total Number of Instances              768
```

Fig. 7 The results of the J48 classifier algorithm

We also ran experiments with J48 classification algorithm, deleting every time the root attribute with the intent to conclude to the list of most significant attributes. After conducting our seventh experiment, we identified that the least informative attributes that cannot produce a classification tree on their own are the "Press" and "skin" attributes.

Hence, we identified that the possible order of the most important attributes is "Plass," "Mass," "Age," "Preg," "2-Hour serum insulin," and "Diabetes pedigree function." From the previous steps, we observed that "Insu" attribute does not

```
J48 pruned tree
------------------

plas = '(-inf-99.5]': tested_negative (197.0/16.0)
plas = '(99.5-127.5]'
|   age = '(-inf-28.5]': tested_negative (143.0/17.0)
|   age = '(28.5-inf)'
|   |   mass = '(-inf-27.85]': tested_negative (41.0/10.0)
|   |   mass = '(27.85-inf)'
|   |   |   pedi = '(-inf-0.5275]': tested_negative (70.0/27.0)
|   |   |   pedi = '(0.5275-inf)': tested_positive (34.0/10.0)
plas = '(127.5-154.5]'
|   mass = '(-inf-27.85]': tested_negative (34.0/5.0)
|   mass = '(27.85-inf)'
|   |   age = '(-inf-28.5]': tested_negative (48.0/20.0)
|   |   age = '(28.5-inf)': tested_positive (79.0/28.0)
plas = '(154.5-inf)': tested_positive (122.0/24.0)

Number of Leaves  :     9

Size of the tree :      15
```

Fig. 8 The J48 classifier algorithm results after supervised discretization

contribute significantly to the classification tree, as it is not included in the first classification tree.

After removing the "Insu" attribute, ultimately, the most important attributes are the following:

- "Plass": Plasma glucose concentration
- "Mass": Body mass index
- "Age"
- "Preg": Number of times pregnant
- "Pedi": Diabetes pedigree function

We further experimented, applying supervised discretization filters to extract meaningful information. Initially, the accuracy of the J48 algorithm has slightly increased to 78.52%. Examining the classification tree depicted in Fig. 8, after discretization we observe that the Preg attribute is not included.

Consecutively, to make the decision process clearer for healthcare practitioners, we suggest to initially focus the diagnosis on the following four attributes:

- "Plass": Plasma glucose concentration
- "Mass": Body mass index
- "Age"
- "Pedi": Diabetes pedigree function

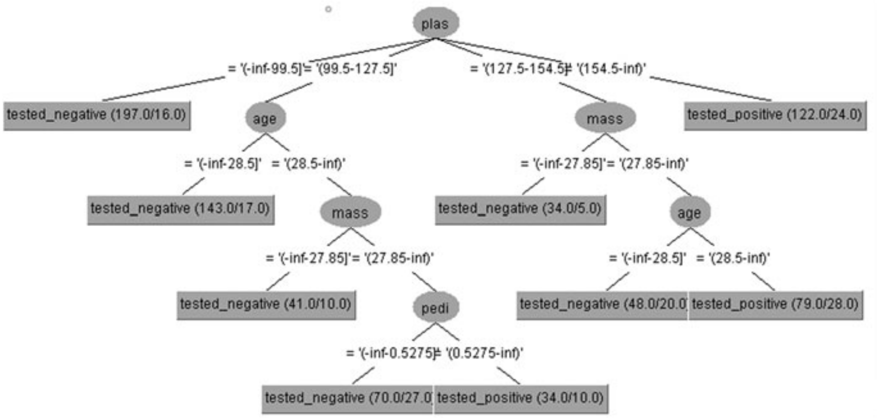
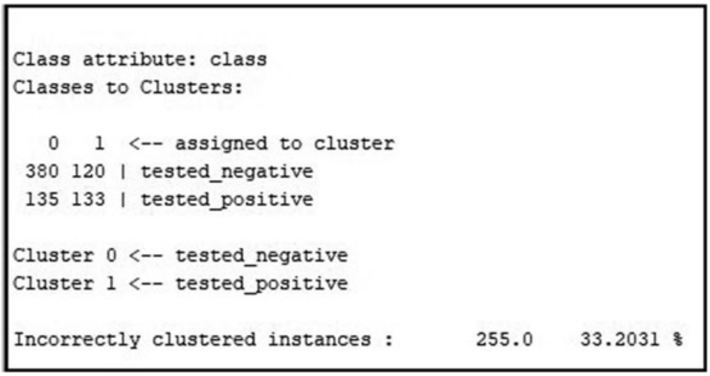


Fig. 9 The J48 classification tree considering four attributes



```
Class attribute: class
Classes to Clusters:

   0   1  <-- assigned to cluster
 380 120 | tested_negative
 135 133 | tested_positive

Cluster 0 <-- tested_negative
Cluster 1 <-- tested_positive

Incorrectly clustered instances :       255.0    33.2031 %
```

Fig. 10 Outcome of simple K-means cluster algorithm

Alongside, the above process resulted in a more comprehensive tree (Fig. 9).

Moreover, clustering algorithms created groups of patients with the same values, supporting healthcare practitioners to categorize future patients and provide early medical help. Clustering is unsupervised, i.e., there are no known clusters. Thus, K-means considers the class attribute as of equal importance with all the others. Specifically, K-means is a partitioning approach for which the most common evaluation measure is the sum of squared errors (SSE). On this premise, we clustered the dataset, ignoring the binary class attribute, into four groups (optimal number of classes given by the elbow method) [72]. We observed that SSE decreased, from SSE =119.18 to SSE= 95.23, and clusters were formed as Fig. 10 shows.

It is known that there is no single algorithm that performs better for every dataset. The level of accuracy highly depends on the nature and size of given data [46]. Thus, there is a need of iterative testing and experimenting to identify the most suitable algorithm or combination of algorithms that can lead to more accurate results for the specific objectives and dataset of the attempted analysis.

Doctors and hospital management could find the above information beneficial for disease diagnosis as well as health information provision to the general public in a comprehensive and clear manner. Furthermore, experimentation like the one discussed so far could help health practitioners to formulate and verify clinical hypotheses. The following proposals are mainly focused on classification outcomes.

Practitioners should leverage the information extracted for assistive diagnosis, but also to serve informational needs, as the outcomes is highly readable. An interesting application is to use the above three as health information leaflets to explain the diagnosis procedure to patients but also alert patients with similar symptoms.

To enhance assistive diagnosis the nature of attributes enables the use of wearable sensors. For instance, after examining clinically the plasma glucose concentration attribute, patient, with the help of a wearable sensor, could be alerted to visit his/her doctor when exceeding body mass index values based on the above tree (Fig. 5). Similarly, patients could be notified when he/she passes the 28 years old age threshold to visit the appropriate doctor as he/she is more likely to suffer from diabetes.

6 Conclusions

Technology has enhanced dramatically the way we live in cities and the healthcare sector. Healthcare providers leverage technology to optimize decision-making that leads to objective and personalized treatments, while traditional treatment methods tend to be substituted by algorithmic predictions. The IoT gave elderly people the opportunity to have self-assistance and self-monitoring in their own environments with the use of wearable devices and sensors; therefore, better life quality is achieved. Hospital administration operates with the use of smart systems to better manage supplies of hospital essentials and estimate costs.

To conclude, for answering the two research questions, we emphasized on DM techniques and their application on smart healthcare systems. At the same time, we focused and provided examples on major DM applications for smart health.

As a result, we found that several DM/ML techniques can be used for a cost-effective and accurate treatment, disease prediction with high accuracy, in-house self-assistance using wearable devices, hospital management, disease diagnosis, risk monitoring, drug research, and many others. In Table 1, we presented some prominent applications for each DM task, based on review of the literature.

However, there is no single algorithm that performs faultlessly for every dataset. Accuracy depends on the nature and size of datasets, while analyst and expert objectives need due consideration. Thus, the decision on which DM task and algorithm to apply to each case requires thorough examination, testing, and verification.

Using DM algorithms for classification, clustering, association rules etc., keeps healthcare getting better and ongoing research promises more to come in the near future. However, concerns regarding data privacy and effective, secure data sharing should be further examined to rectify limitations and achieve universal interoperation of smart healthcare systems.

References

1. A.M. Townsend, *Smart cities: big data, civic hackers, and the Quest for a New Utopia* (W.W. Norton & Company, New York, 2013)
2. M. Bermudez-Edo, P. Barnaghi, K. Moessner, Analysing real world data streams with spatio-temporal correlations: entropy vs. pearson correlation. *Automation in Construction* **88**, 87–100 (2018)
3. Q. Le-Dang, T. Le-Ngog, Internet of Things (IoT) Infrastructures for Smart Cities, in *Handbook of Smart Cities: Software Services and Cyber Infrastructure*, (Cham, Springer, 2018), pp. 1–30
4. P. Anatharam, P. Barnaghi, K. Thirunarayan, A. Sheth, Extracting city traffic events from social streams. ACM Trans. Intell. Syst. Technol. **6**(4), 43.:1-43:27, (2015)
5. A.K. Kar, S.Z. Mustafa, M.P. Gupta, P.V. Ilavarasan, Y.K. Dwivedi, Understanding Smart Cities: Inputs for Research and Practice, in *Advances in Smart Cities: Smarter People, Governence, and Solutions*, (CRC Press, Boca Ralton, 2017), p. 1
6. A. Yassine, S. Singh, A. Alamri, Mining human activity patterns from smart home big data for health care applications. IEEE Access **5**, 13131–13141 (2017)
7. B. Liu, K. He, G. Zhi, The impact of big data and artificial intelligence on the future medical model. Med. Philos. **39**(22), 1–4. (in Chinese), (2018)
8. S. Tian, W. Yang, J.M.L. Grange, P. Wang, W. Huang, Z. Ye, Smart healthcare: making medical care more intelligent. Glob. Health J. **3**, 62–65 (2019)
9. K. Kasikumar, M.M. Najumuddeen, R. Suresh, Applications of data mining techniques in healthcare and prediction of heart attacks. Int. J. Data Min. Tech. Appl. **7**, 172–176 (2018)
10. G.R. Pooja, M.B. Trinath, K. Vasanthi, K.S. Ramireddy, R.K. Tenali, Smart E-health prediction system using data mining. Int. J. Innov. Technol. Explor. Eng **8**(6), 787–791 (2019)
11. D.K. Singh, M. Ashraf, An experimental approach for prediction of disease in smart health system using data mining technique. Int. J. Adv. Sci. Technol. **27**, 112–119 (2019). Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/102
12. B. Kantarci, K.G. Carr, C.D. Pearsall, SONATA: Social Network Assisted Trustworthiness Assurance in Smart City Crowdsensing, in *The Internet of Things: Breakthroughs in Research and Practice*, (Hershey, IGI Global, 2017), pp. 278–299
13. J.A. Rodriguez, F.J. Fernadez, P. Arboleya, Study of the architecture of a smart city. Proceedings **2**, 1–5 (2018)
14. P. Lombardi, S. Giordano, Evaluating the Smart and Sustainable Built Environment in Urban Planning, in *Handbook of Research on Social, Economic, and Environmental Sustainability in the Development of Smart Cities*, (Hershey, IGI Global, 2015), pp. 44–59
15. H. Habibzadeh, A. Boggio-Dandry, Z. Qin, T. Soyata, B. Kantarci, H. Mouftah, Soft sensing in smart cities: handling 3Vs using recommender systems, machine intelligence, and data analytics. IEEE Commun. Mag. **56**, 78–86 (2018)
16. A. Souza, M. Figueredo, N. Cacho, D. Araujo, C.A. Prolo, Using big data and real-time analytics to support smart city initiatives. IFAC-Papers Online **49**(30), 257–262 (2016)
17. P. Koukaras, C. Tjortjis, D. Roussidis, Social media types: introducing a data driven taxonomy. Computing **102**(1), 295–340 (2020)

18. A. Mystakidis, C. Tjortjis, Big Data Mining for Smart Cities: Predicting Traffic Congestion using Classification, in The 11th International Conference on Information, Intelligence, Systems and Applications, Piraeus, 2020
19. S. Mohanty, U. Choppali, E. Kougianos, Everything you wanted to know about smart cities: the Internet of Things is the backbone. IEEE Consum. Electron. Mag. **5**(3), 60–70 (2016)
20. M.V. Moreno, F. Terroso-Saenz, A. Gonzalez-Vidal, M. Vldez-Vela, A. Skarmeta, M.A. Zamora, V. Chang, Applicability of big data techniques to smart cities deployments. IEEE Trans. Ind. Inf. **13**(2), 800–809 (April 2017)
21. J. Massana, C. Pous, L. Burgas, J. Melendez, J. Colomer, Identifying services for short-term load forecasting using data driven models in a Smart City platform. Sustain. Cities Soc. **28**, 108–117 (2017)
22. L. DeRen, C. JianJun, Y. Yuan, Big data in smart cities. Sci. China Inf. Sci. **58**(12) (2015)
23. D.J. Cook, G. Duncan, G. Sprint, R. Fritz, Using smart city technology to make healthcare smarter. Proc. IEEE **106**, 708–722 (April 2018)
24. K. Joo-Chang, C. Kyungyong, Depression index service using knowledge based crowdsourcing in smart health. Wirel. Pers. Commun. **93**, 255–268 (March 2017)
25. A. Copie, V.I. Munteanu, B. Manate, T.-F. Fortis, An Internet of Things Governance Architecture with Applications in Healthcare, in *The Internet of Things: Breakthroughs in Research and Practice*, (Hershey, IGI Global, 2017), pp. 112–136
26. A.A. Obinikpo, B. Kantarci, Big sensed data meets deep learning for smarter health care in smart cities. J. Sens. Actuator Netw. **6**, 1–22 (2017)
27. J. Dhar, A. Ranganathan, Machine learning capabilities in medical diagnosis applications: computational results for hepatitis disease. Int. J. Biomed. Eng. Technol. **17**, 330–340 (2015)
28. K. Polat, S. Gunes, Principles component analysis, fuzzy weighting pre-processing and artificial immune recognition system based diagnostic system for diagnosis of lung cancer, in *Expert Systems with Applications*, vol. 34, 1st edn., (2008), pp. 214–221
29. A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks. *Nature***542**, 115–118 (2017)
30. S. Wang, R. Summers, Mechine Learning and Radiology. Med. Image Anal. **16**(5), 933–951 (2012)
31. S.P. Somashekhar, M.-J. Sepúlveda, S. Puglielli, A.E.H. Shortliffe, C. Kumar, A. Rauthan, N. Kumar, P. Patil, K. Rhee, Y. Ramya, Watson for oncology and breast cancer treatment recommendations: agreement with an expert multidisciplinary tumor board, in *Annals of Oncology*, vol. 29, 2nd edn., (2018), pp. 418–423
32. K. Kincade, Data mining: digging for healthcare gold. Insur. Technol. ***23***(*2*), *IM2–IM7* (1998)
33. A. Milley, Healthcare and data mining. Health Manag. Technol. **21**(8), 44–47 (2000)
34. J. Andreu-Perez, D. Leff, H. Ip, G. Yang, From wearable sensors to smart Implants—toward pervasive and personalized healthcare. IEEE Trans. Biomed. Eng. **62**, 2750–2762 (2015)
35. M. Chan, E. Campo, D. Estève, J.-Y. Fourniols, Smart homes—current features and future perspectives. Maturitas **64**(2), 90–97 (2009)
36. L. Liu, E. Strouli, I. Nikolaidis, A. Miguel-Cruz, A.R. Rincon, Smart homes and home health monitoring technologies for older adults: a systematic review. Int. J. Med. Inf. **91**, 44–59 (July 2016)
37. J. Zhang, Y. Li, L. Cao, Y. Zhang, Research on the construction of smart hospitals at home and abroad. Chin. Hos. Manag, 64–66 (2018)
38. K. Li, J. Wang, T. Li, F. Dou, K. He, Application of internet of things in supplies logistics of intelligent hospital. Chin. Med. Equip., 172–176 (2018)
39. H. Demirkan, A Smart Healthcare Systems Framework, IT Prof. no. 5, pp. 38–45, Sept–Oct 2013
40. Q. Chen, Y. Lu, Construction, and application effect evaluation of integrated manage-ment platform of intelligent hospital based on big data analysis. Chin. Med. Herald., 161–164 (2018)
41. P. Piazza, Health alerts to fight bioterror, Secur. Manag. p. 40, 2002.
42. J. Redfern, Smart health and innovation: facilitating health-related behaviour change. Proc. Nutr. Soc., 328–332 (2017)

43. M. Ridinger, American healthways uses SAS to improve patient care, DM Rev. no. 12, p. 139, 2002.
44. T.S. Brisimi, T. Xu, T. Wang, W. Dai, W.G. Adams, I.C. Paschalidis, Predicting chronic disease hospitalizations from electronic health records: an interpretable classification approach. Proc. IEEE **106**(4), 690–707 (2018)
45. S. Zhang, C. Tjortjis, X. Zeng, H. Qiao, B. Iain, J. Keane, Comparing data mining methods with logistic regression in childhood obesity prediction. Inf. Syst. Front. J. Springer **11**(4), 449–460 (2009)
46. C. Tjortjis, M. Saraee, B. Theodoulidis, J. Keane, Using T3, an improved decision tree classifier, for mining stroke related medical data. Method Inf Med Schattauer GmbH **46**(5), 523–529 (2007)
47. H. Banaee, M.U. Ahmed, A. Loutfi, Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges. Sensors **13**, 17472–17500 (2013)
48. T. Nef, P. Urwuler, M. Buchler, I. Tarnanas, R. Stucki, D. Cazzoli, R. Muri, U. Mosimann, Evaluation of three state-of-the-art classifiers for recognition of activities of daily living from smart home ambient data. Sensors **15**, 11725–11740 (2015)
49. B. Lin, Y. Huangfu, N. Lima, B. Lobson, M. Kirk, P. O'Keeffe, S. Pressley, V. Walden, B. Lamb, D. Cook, Analyzing the relationship between human behavior and indoor air quality. J. Sens. Actuator Netw. **6**, 1–18 (2017)
50. M. Islam, M. Hasan, X. Wang, H. Germack, M. Noor-E-Alam, A systematic review on healthcare analytics: application and theoretical perspective of data mining. Healthcare **6**(2), 54 (2018)
51. N. Jothi, N.A. Rashid, W. Husain, Data Mining in Healthcare – A Review, in The Third Information Systems International Conference, 2015.
52. M. Durairaj, V. Ranjani, Data Mining Applications In Healthcare Sector: A Study. Int. J. Sci. Technol. Res. **2**(10), 29–35 (2013)
53. V. Tatsis, C. Tjortjis, P. Tzirakis, Evaluating data mining algorithms using molecular dynamics trajectories. Int. J. Data Min. Bioinf. Indersci. **8**(2), 169–187 (2013)
54. P. Ahmad, S. Qamar, S.Q.A. Rizvi, Techniques of Data Mining In Healthcare: A Review. Int. J. Comp. Appl. **120**(15), 38–50 (2015)
55. H. Chung, P. Gray, Data mining. J. Manag. Inf. Syst. **16**(1), 11–16 (1999)
56. M. Aggarwal, Medium.com, 7 January 2018. [Online]. Available: https://medium.com/@thecodingcookie/cross-industry-process-for-data-mining-286c407132d0
57. I. Parvathi, S. Rautaray, Survey on data mining techniques for the diagnosis of diseases in medical domain. Int. J. Comp. Sci. Inf. Technol. **5**(1), 838–846 (2014)
58. R. Martinez-Espana, A. Bueno-Crespo, I. Timon, J. Soto, A. Munoz, J.M. Cecilia, Air-pollution prediction in smart cities through machine learning methods: a case study in Murcia, Spain. J. Univer. Comp. Sci. **24**(3), 261–276 (2018)
59. P. Tzirakis and C. Tjortjis, "T3C: Improving a Decision Tree Classification Algorithm's Interval Splits on Continuous Attributes," *Advances in Data Analysis and Classification,* Springer, vol. 11, no. 2, pp. 353-370, 2017.
60. S. Mohapatra, P.K. Patra, S. Mohanty, B. Pati, Smart Health Care System using Data Mining, in International Conference on Information Technology, 2018.
61. D. Tomar, S. Agarwal, A survey on data mining approaches for Healthcare. Allahabad Int. J. Biosci. Biotechnol. **5**(5), 241–266 (2013)
62. Y. Kanellopoulos, P. Antonellis, C. Tjortjis, C. Makris, N. Tsirakis, k-Attractors: a partitional clustering algorithm for numeric data analysis. Appl. Artif. Intell. Taylor Francis **25**(2), 97–115 (2011)
63. A. Kelati, J. Plosila, H. Tenhunen, Smart Meter Load Profiling for e-Health Monitoring System, in 7th International Conference on Smart Energy Grid Engineering, 2019.
64. R. Agrawal and R. Srikant, "Apriori algorithm". 1994.
65. S.M. Ghafari, and C. Tjortjis, Association Rules Mining by improving the Imperialism Competitive Algorithm (ARMICA), in IFIP AICT Proceeding of 12th International Conference on Artificial Intelligence Applications and Innovations (AIAI 2016). Springer,, 2016.

66. S. Yakhchi, S. M. Ghafari, C. Tjortjis, M. Fazeli, ARMICA-Improved: A New Approach for Association Rule Mining, in Proceeding of 10th International Conference on Knowledge Science, Engineering and Management (KSEM 17), Springer *LNAI,* vol. 10412, pp. 296–306, 2017.
67. J. Han, H. Pei, Y. Yin, Mining Frequent Patterns without Candidate Generation, in *Proceeding of Conference on the Management of Data (SIGMOD'00, Dallas, TX)*, (ACM Press, New York, 2000)
68. S.M. Ghafari, C. Tjortjis, A survey on association rules mining using Heuristics. WIREs Data Min. Knowl. Discov. **9**(4) (2019)
69. Y. Ji, H. Ying, J. Tran, P. Dews, A. Mansour, M.R. Massanari, Mining Infrequent Causal Associations in Electronic Health Databases, in 2011 IEEE 11th Int'l Conf. on Data Mining Workshops, 2011.
70. A. Asuncion, D. Newman, UCI Machine Learning Repository, 2007. [Online]. Available: https://archive.ics.uci.edu/ml/datasets.php
71. H. Ian, E. Frank, A.H. Mark, J.P. Christopher, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd edn. (Morgan Kaufmann, San Francisco, 2011)
72. R.L. Thorndike, Who belongs in the family? *Psychometrika*18, 267–276 (1953)

Theodora Chatzinikolaou studied at Dpt. of Accounting and Finance, School of Business Administration, University of Macedonia and received her bachelor's degree in 2017. In 2018, she commenced studying towards an MSc in e-Business and Digital Marketing, School of Science and Technology, International Hellenic University, expected to graduate in May 2021. For her master's degree, she attained data science and data mining relevant courses.

Eleni Vogiatzi is a graduate of the business administration bachelor program in the University of Macedonia, Thessaloniki, Greece. She is currently an MSc student in e-business and Digital Marketing, at the School of Science and Technology, International Hellenic University, in Thessaloniki. Her dissertation concerns Search Engine Optimization (SEO) and Search Engine Results Page (SERP) ranking factors. During her participation in the MSc program, among others, she attained courses related to Data Science, Data Mining, ICT essentials, and Web analytics.

Anestis Kousis received a bachelor's degree in computer science from the Hellenic Open University Patras, Greece, in 2015 and the M.Ed. in ICT teaching and learning from Aristotle University of Thessaloniki, Greece, in 2017. He is currently pursuing a Ph.D. with the School of Science and Technology, International Hellenic University, Greece, working with Prof. C. Tjortjis on data technologies for smart cities. He is also working as a computer science educator. His current research focus is on developing and applying new data mining techniques in the smart cities' context.

Christos Tjortjis is Acting Dean and Assoc. Prof. in Knowledge Discovery and Software Engineering systems at the School of Science and Technology, International Hellenic University. He is Programme Director for the Erasmus Mundus MSc in Smart Cities and Communities, the MSc in Data Science, and the MSc in ICT systems. He holds a Deng (Hons) from Patras, Computer Eng. & Informatics, a BSc (Hons) from Democritus Law School, Greece, an MPhil in Computation from UMIST, and a PhD in Informatics from Manchester, U.K. He was Lecturer at UMIST, Computation, and the Schools of Informatics and Computer Science, at the University of Manchester, U.K., adj. Senior Lecturer at Eng. Informatics & Telecoms, W. Macedonia, and at Computer Science, Ioannina, Greece. His research interests are in data mining, smart cities, and software engineering. He worked in over 15 R&D projects, leading 4. He published over 75 papers in int'l referred journals and conf. and was PC member in over 90 int'l conf. He received over 1000 citations (h-index 18).

Window Functions for Phasor Signal Processing of Wide-Area Measurement in Smart Grid Internet of Things Communications

Mohammad Kamrul Hasan, Musse Mohamud Ahmed, Shayla Islam, Imran Memon, Riaz Ahmed Shaikh, and Hidayatullah Shaikh

1 Introduction

The smart grid represents the next generation of electrical grid aiming to resolve all the abovementioned problems with the use of advanced information and communication technologies (ICT). It is also foreseen that smart grids will ensure smart and real-time monitoring of the grid status through the deployment of advanced sensing capabilities. Using the phasor measurement technique, the wide-area measurement system was designed and developed using the IEEE C37.118 communication protocol [1–3]. These phasor measurement techniques are applying the digital signal processing methods in the base algorithm. Fast Fourier transform (FFT) is an execution of the DFT that delivers practically comparative outcomes as the DFT [3, 4]; however, it is more proficient and a lot quicker and regularly lessens the calculation time fundamentally. It is just a computational calculation utilized for quick and proficient calculation of the DFT. Different quick DFT

M. K. Hasan (✉)
Center for Cyber Security, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
e-mail: mkhasan@ukm.edu.my

M. M. Ahmed
Department of Electrical and Electronics Engineering, Faculty of Engineering, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia

S. Islam
Department of Computer Science, UCSI University Malaysia, Kuala Lumpur, Malaysia

I. Memon
Department of Computer Science, Bahria University, Karachi, Sindh, Pakistan

R. A. Shaikh · H. Shaikh
Department of Computer science, Shah Abdul Latif University, Khairpur, Sindh, Pakistan

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_4

calculation systems are referred to overall as the quick Fourier change or FFT. FFT has been broadly utilized for acoustic estimations in places of worship and show lobbies [5]. Other FFT applications remember otherworldly investigation for simple video estimations, sifting calculations, huge whole number and polynomial augmentation, ascertaining Fourier arrangement coefficients, figuring convolutions, registering isotopic dispersions, planning structures, performing thick organized grids, and picture preparing to produce low-recurrence clamor, and that's just the beginning.

The discrete Fourier transform (DFT) is one of the most significant apparatuses in advanced sign handling that computes the range of a limited span signal [2]. It is normal to encode the data in the sinusoids that structure a sign. In any case, in certain applications, the state of a period area waveform isn't an application for signals wherein case signal recurrence content turns out to be helpful in manners other than as advanced signs. The phasor digital signal representation in terms of its frequency component in a frequency domain is important. The algorithm that transforms the time domain signals to the frequency domain components is known as the discrete Fourier transform or DFT.

2 Window Functions

Digital signals are adequately huge that the dataset cannot be unguided overall. Adequately huge signs are additionally hard to examine measurably in light of the fact that factual computations require all focuses to be accessible for examination. A digital channel is one of the gadgets utilizing DFT. A channel is a gadget that transmits (or dismisses) a particular scope of frequencies. There are two kinds of channels, for example, finite impulse response (FIR) and infinite impulse response (IIR).

FIR channels can be intended to have a precisely direct stage, and it has likewise extraordinary adaptability in forming their greatness reaction. What's more, FIR channels are increasingly steady, and the impacts of quantization mistakes are not exactly IIR channels. While IIR channels require fewer coefficients than FIR channels for sharp cutoff recurrence reaction, and simple channels must be demonstrated utilizing IIR channels. FIR channel is the perfect channel when utilizing windows strategies. The investigation of DFT includes digitizing simple signals and afterward utilizing an FFT to break down their ghastly substance that causes ghostly spillage impacts [6]. The spillage impacts will confine the capacity to identify consonant signals in signal blends; then, it needs window capacities to decrease its spillage impacts [7–9]. Window capacities are a technique that is utilized to lessen otherworldly spillage impacts. The spillage impacts are the first sign that will be redirected from the predetermined recurrence band into the upper frequencies band. The convolution of the sign is changed utilizing an FFT and afterward duplicated by the comparable recurrence reaction, HD (Ω). A last converse change of the motivation reaction hD(Ω) which is $Y[k] = X[k]H[k]$ yields

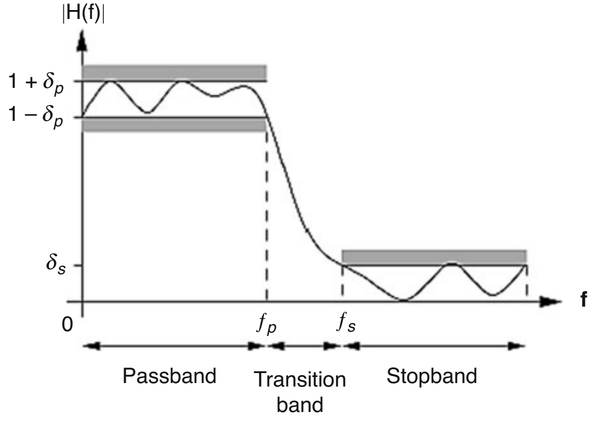


Fig. 1 Low-pass filter

the sign y[n]. The case of the determination recurrence band is, for example, Fig. 1 illustrated the instances.

In view of FFT on the window sine-wave information, the subsequent sin(x)/x bend shows lessening side projections and less spillage and improving affectability of ghostly estimations [10]. Other than that, the approach to conquer the loss of recurrence goals by examining signals quicker during the test period and multiplying the example rate by twofold the number of tests [11–16]. Next, various windows properties that will recognize recurrence goals or identify which sign is more fragile (side lobe level). The windows work by executing an FFT on the 32 sine-wave tests of unearthly information at that point contrasted with the hypothetical otherworldly information from the first unending sine wave [17, 18]. At that point, it duplicates the 32 sine-wave esteems by the relating window esteems. The windowing activity "packs" the beginning and consummation estimations of zero of the example interims. In this manner, the arrangement of 32 qualities never again has an unexpected beginning or end of advances [19–21]. The recurrence range of 32 sine-wave esteems is the convolution of the Fourier transform of the unending span sine wave (yields single an incentive at the sine wave's recurrence) and the Fourier transform of 32 1's. The convolution is a persistent sin(x)/x work that is subsequent the bend of DTFT. There are five categories of window capacities and its conditions; however, three kinds of window which are the HNNW, HMMW, and BW will be discussed.

2.1 *HNNW Technique*

This technique is mostly used to mitigate the noise level while frequency is filtering. The HNNW is a plane window that has the initial value and end value comfortably zero. It r=features a fast decay of the side lobes

$$0.5\text{–}0.5\ \cos\left(\frac{2\pi n}{N-1}\right) \tag{1}$$

where N represents the number of sine-wave samples you acquire and n equals the sample index 0, 1, 2, and so on up to N–1 data values

2.2 HMMW Technique

The HMMW has selected the best useable signal processing techniques of all. The presentation of HMMW appears similar to a HMMW just elevated on a pedestal. Nevertheless, the endpoints don't reach zero. The level of the side lobes is roughly constant.

$$0.54\text{–}0.46\ \cos\left(\frac{2\pi n}{N-1}\right) \tag{2}$$

2.3 HW Technique

This windowing technique is the most widely applicable window since it is very application-friendly and features a rapid deterioration of side lobes at the cost of a wide main lobe and low-frequency discrimination.

$$0.42\text{–}0.5\ \cos\left(\frac{2\pi n}{N-1}\right) + 0.08\cos\left(\frac{2\pi n}{N-1}\right) \tag{3}$$

3 Simulation Result Analysis

The coding for the simulation and testing is accomplished using Matlab Based. Figures 2, 3, 4, 5, 6, 7, and 8 summarized the HNNW that has a sinusoidal shape. As the number of samples (N) is increased, the peak of the bell-shaped waveform in the time domain gets closer to one (1). On the other hand, the number of side lobe per cycle increase as well. The peak of the first side lobe is approximately −30dB. The roll-off rate of the side lobe is approximately 60dB per decade.

From Figs. 8, 9, 10, 11, 12, and 13, the HMMW shows a similar characteristic waveform as HNNW. But the starting and ending point does not touch the 0 or known as the origin. The peak of the first side lobe is around −40dB, while the side lobe roll-off rate is approximately 20dB per decade.

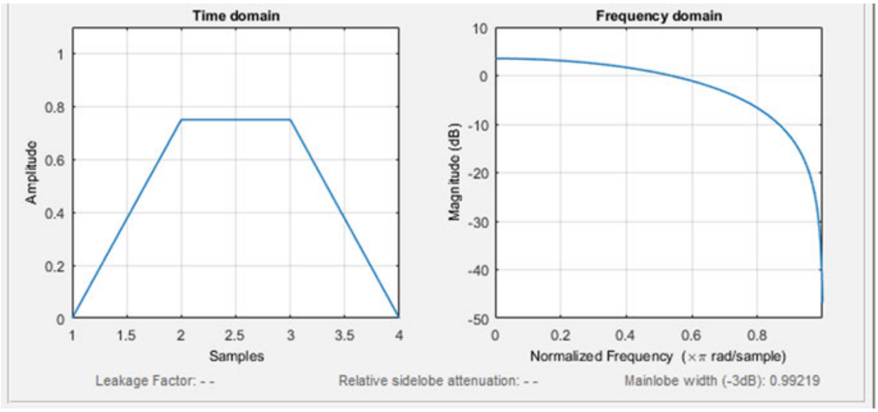


Fig. 2 HNNW for N=4

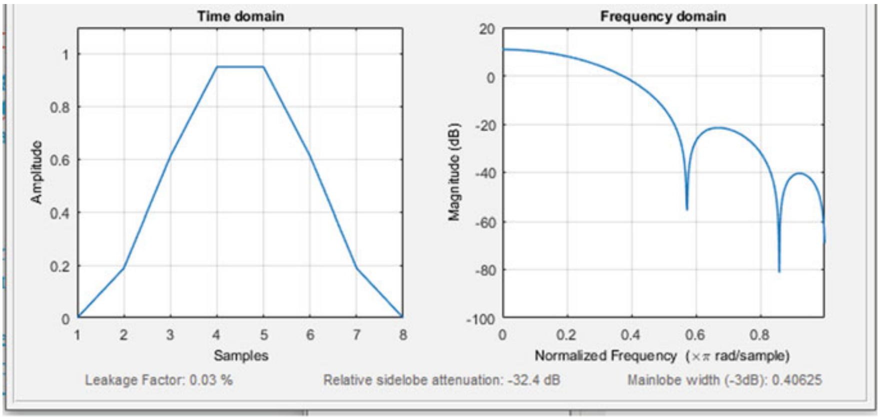


Fig. 3 HNNW for N=8

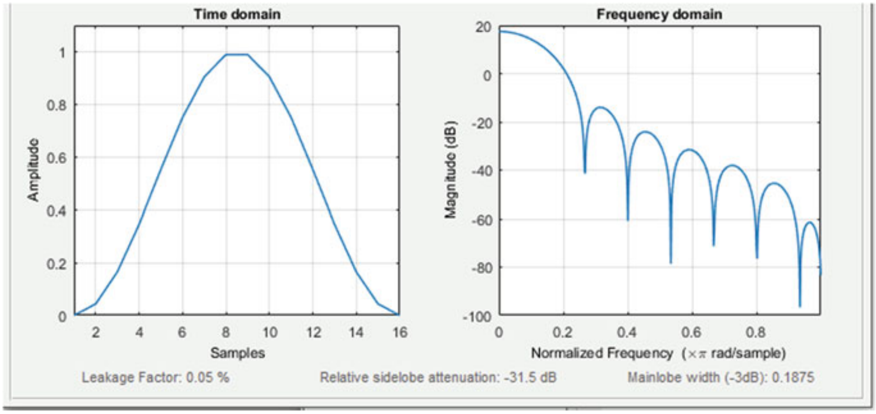


Fig. 4 HNNW for N=16

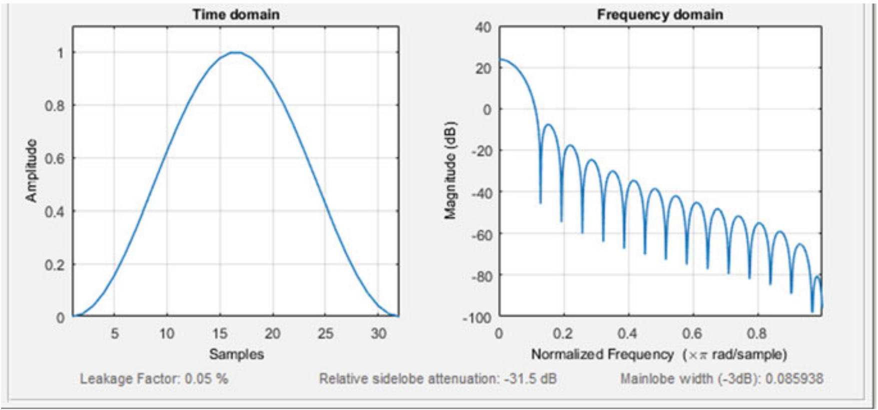


Fig. 5 HNNW for N=32

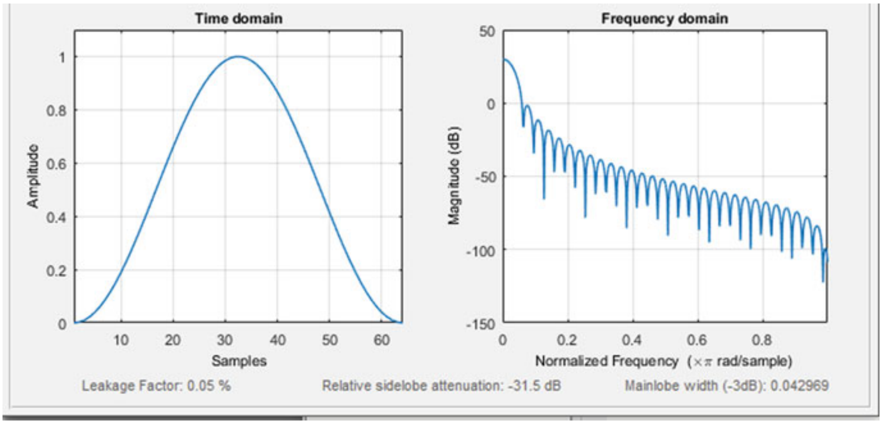


Fig. 6 HNNW for N = 64

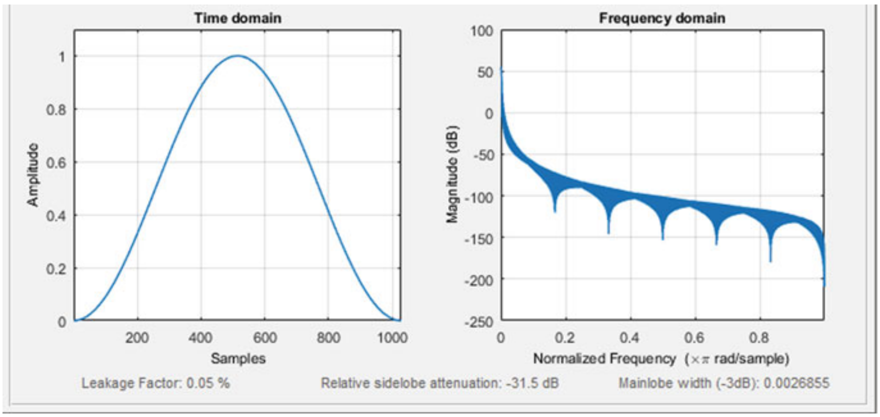


Fig. 7 HNNW for N=1028

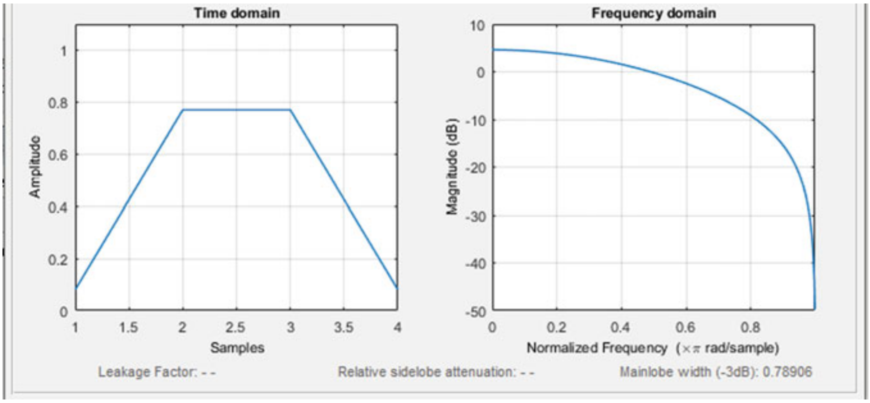


Fig. 8 HMMW for N=4

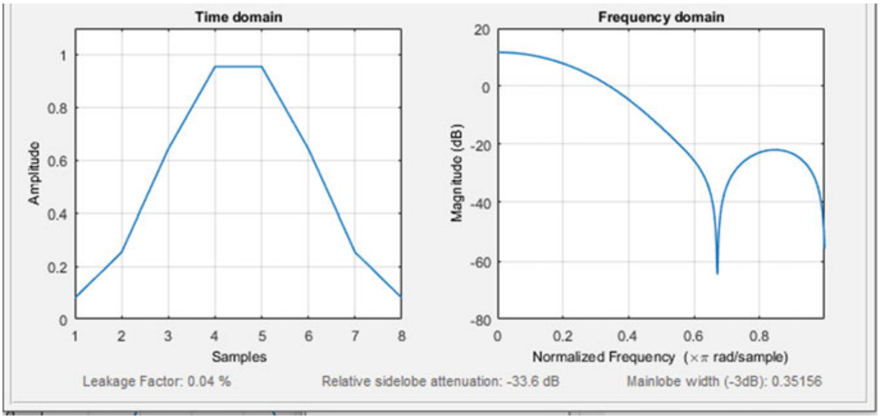


Fig. 9 HMMW for N=8

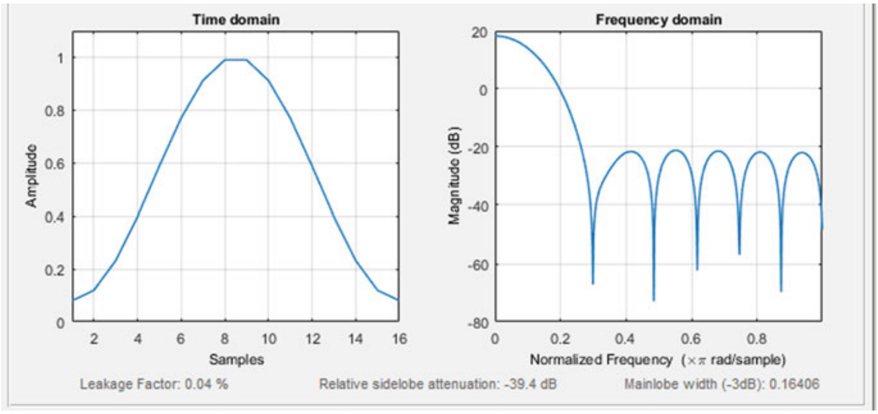


Fig. 10 HMMW for N=16

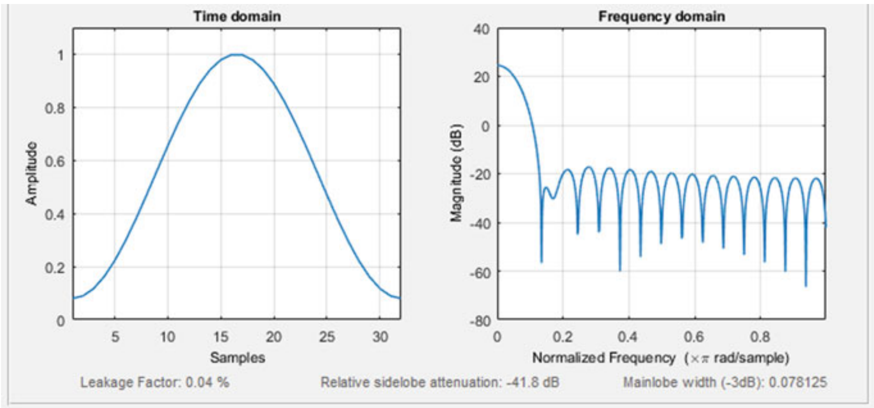


Fig. 11 HMMW for N=32

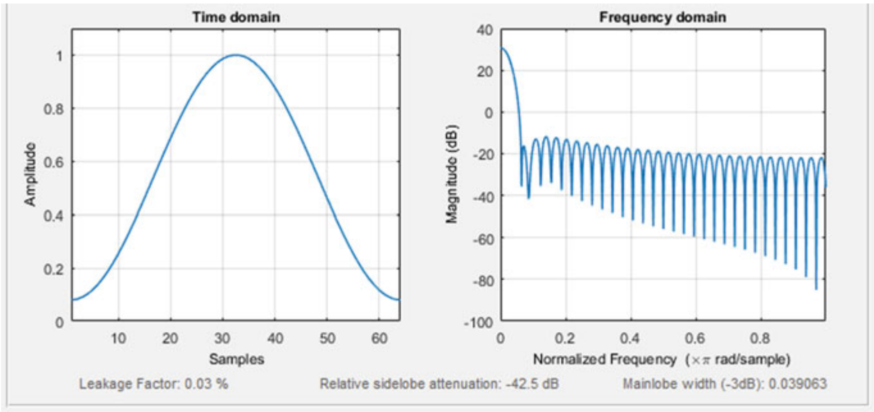


Fig. 12 HMMW for N=64

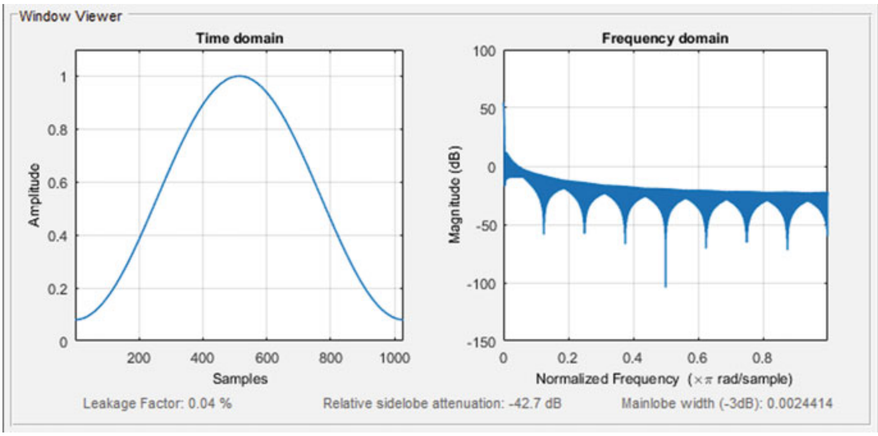


Fig. 13 HMMW for N=1028

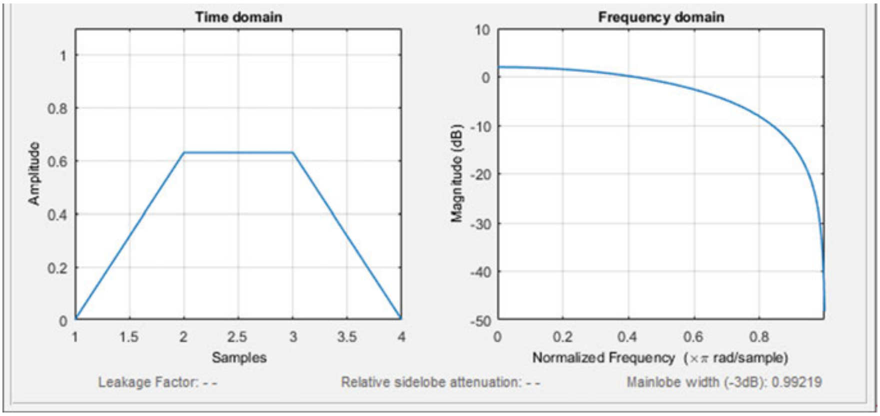


Fig. 14 Blackman window for N=4

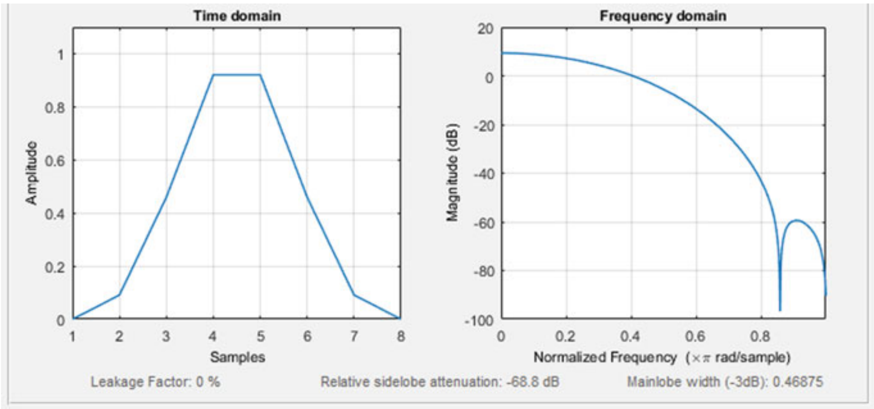


Fig. 15 Blackman window for N=8

The peak of the first side lobe is around -60dB, while the side lobe roll-off rate is approximately 60dB per decade. Figures 14, 15, 16, 17, 18 and 19 represents the Blackman window's general numerical simulation analysis; an extra cosine term is included which helps to reduce the side lobe.

Based on Fig. 20, the time-domain graph shows the comparison of three signals of Blackman (yellow line), Hanning (blue line), and Hamming (red line). It is observed that the Blackman signal touches the 0-axis, the Hanning signal slightly touches the 0-axis, and the Hamming signal does not touch the 0-axis at all. This means that Hamming is still having discontinuity in the signal, and thus it performs better in canceling the nearest side lobe but does a poorer job of canceling any others. Based on Fig. 21, the frequency-domain graph also shows the comparison of Blackman, Hanning, and Hamming signals. Blackman window has the widest main lobe width; it is stated that the narrower the main lobe, the smaller the range of

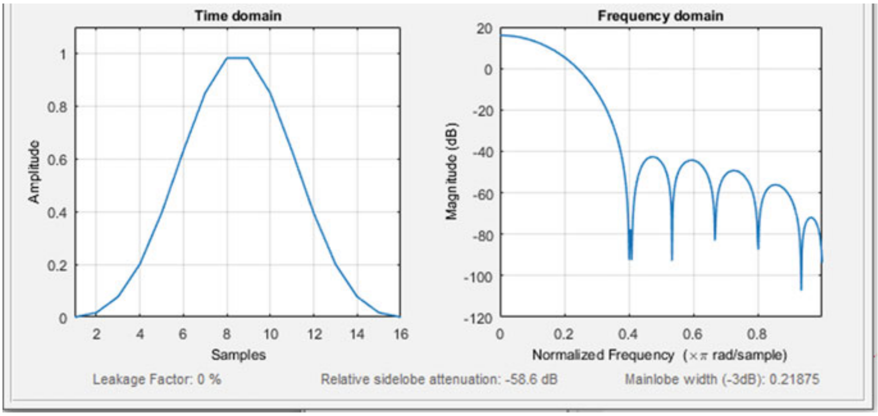


Fig. 16 Blackman window for N=16

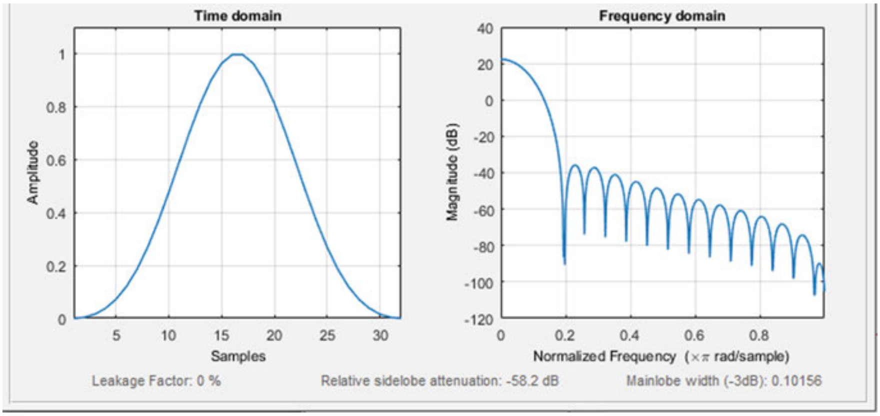


Fig. 17 Blackman window for N=32

frequency that can contribute to the output of any FFT filter. Besides, the accuracy of the FFT filter, in defining the frequency in a waveform, is improved by having a narrow main lobe. In short, the narrower the main lobe, the higher the spectral resolution and amplitude accuracy and the lower the spectral leakage. Moreover, based on the main lobe width of a window is associated with the transition width of the designed filter [22-23].

Next, the peak of the first side lobe of Hanning, Hamming, and Blackman windows is approximate to -35dB, -45dB, and -60dB, respectively. It is better if the peak of the first side lobe is more negative as possible, because the side lobe can overpower the main lobe response if the magnitude is too close [24].

Blackman has a lesser side lobe compare to Hamming and Hanning. Based on [24], the side lobe is the unwanted radiation from a different angle than the main lobe's direction. In addition, excessive side lobe radiation wastes energy and may

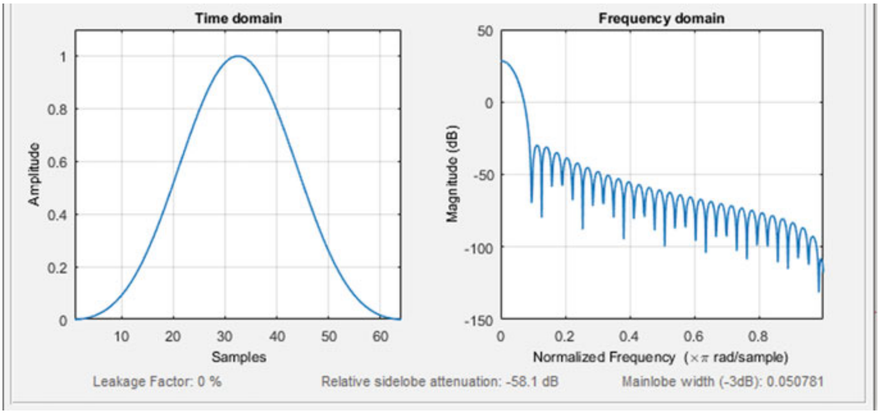


Fig. 18 Blackman window for N=64

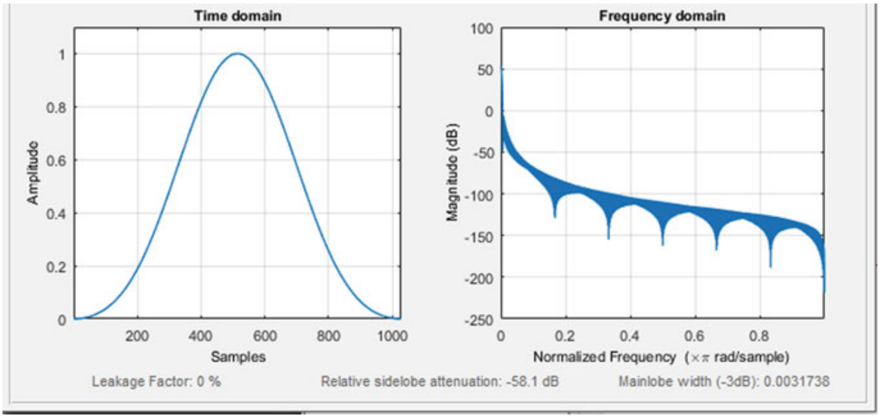


Fig. 19 Blackman window for N=1028

cause interference to other equipment to increase the noise level in the receiver. Blackman has an extra cosine term that reduces side lobe and thus increases the efficiency of the filter; also, less power is lost [9, 10].

Based on Fig. 21, Blackman window has the largest side lobe roll-off rate, while Hamming has the smallest rate. Typically, lower side lobes reduce leakage in the measured FFT but increase the bandwidth of the main lobe. The side lobe roll-off rate is the asymptotic decay rate of the side lobe peaks. By increasing the side lobe roll-off rate, you can reduce spectral leakage [4].

HNNW is useful for analyzing transients longer than the time duration of the window and for general-purpose applications [3, 6]. Based on the National Instruments (n.d), HNNW is satisfactory in 95% of cases, which simply means it is suitable to be used in most engineering cases due to good frequency resolution and reduced spectral leakage. One of the applications of HMMW is in noise measurements [1,

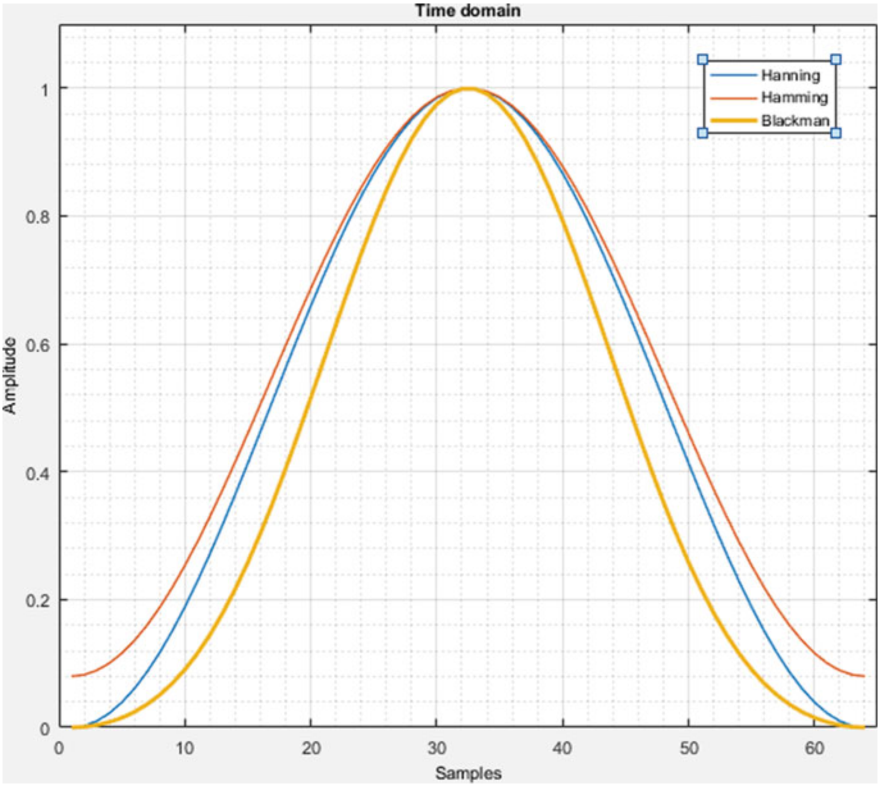


Fig. 20 Time domain for Blackman, Hanning, and Hamming

2]. The Blackman window is applicable for single-tone measurement because it has a low maximum side lobe level and a high side lobe roll-off rate.

Based on Figs. 22 and 23, there is a significant difference in using different values of zero paddings. When 20 zero paddings are applied, the shape of the graph is uneven. However, when 200 zero paddings are used, the shape of the graph is even and smooth. Therefore, increasing the number of zero paddings will smoothen the graph, producing desirable results.

4 Conclusion

This paper discusses the performances of the three windows (Blackman, Hamming, and Hanning) that were obtained from the gain responses observed in the FIR low pass, high pass, bandpass, and bandstop filter. It accomplishing in comparing with the responses from magnitude response, phase response, equivalent noise bandwidth, and side lobe width as well as response in time and frequency domain

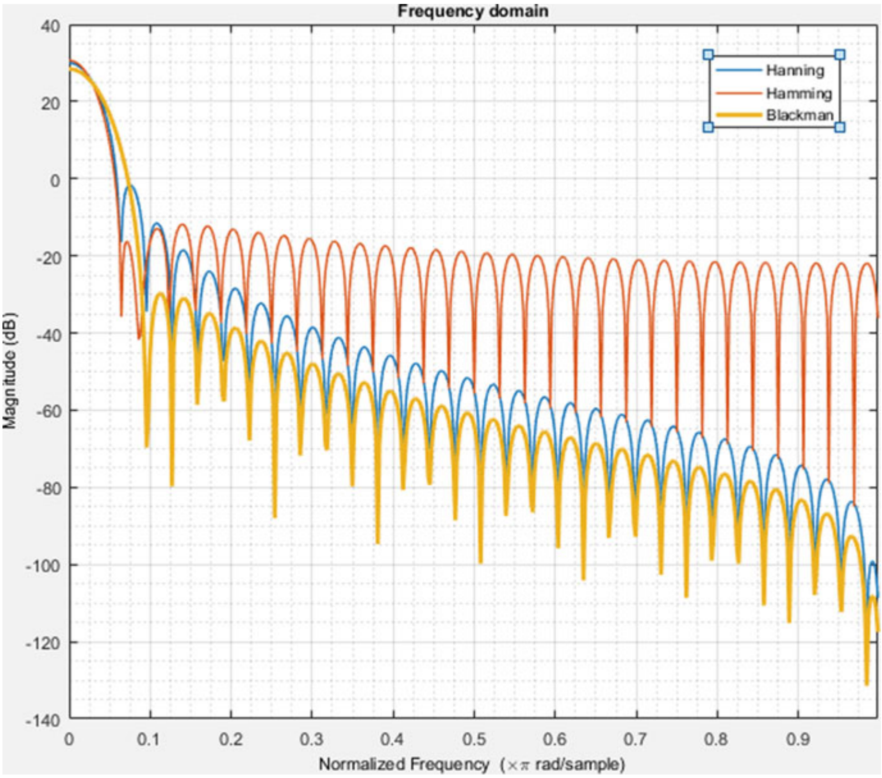


Fig. 21 Frequency domain for Blackman, Hanning, and Hamming

all done using Matlab simulation. HMMW does a better job at canceling the nearest side lobe, and it is useful in noise measurements. Then, the Blackman window results in the spectrum with a wide peak but good side lobe compression. In comparison with the other two, Hanning shows a better frequency resolution that can mitigate spectral leakage. On top of that, these three windows were used to design the FIR filter. Therefore, these three windows have their specifications, which are applied in different types of engineering cases. These windows have their pros and cons and are selected carefully based on their characteristics when undertaking any engineering problems.

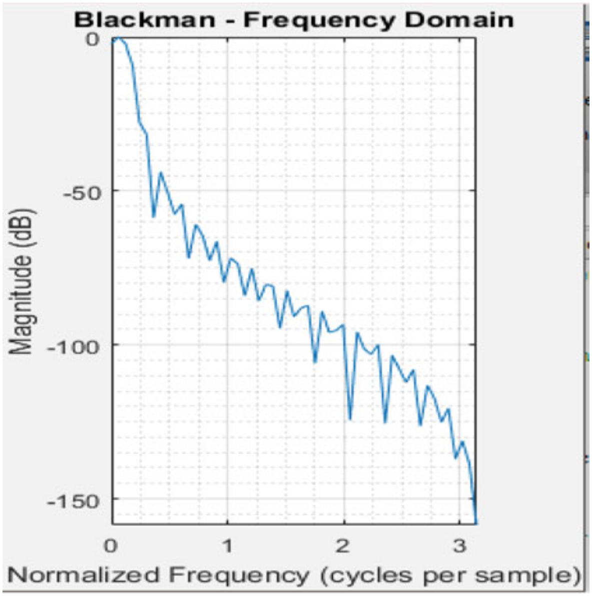


Fig. 22 Frequency domain of Blackman with 20 zero padding

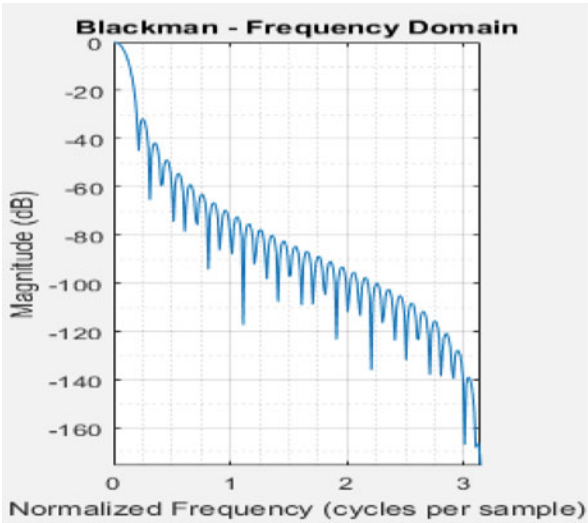


Fig. 23 Frequency domain of Blackman with 200 zero padding

References

1. M.K. Hasan, M.M. Ahmed, S.S. Musa, D. Lee, Calibration of dynamic line rating model for phasor measurement unit based wide area measurement system in smart grid application. Int. J. Recent Technol. Eng. **8**(4), 1879–1883 (2020)
2. M.A. Iwen, A. Gilbert, M. Strauss, Empirical evaluation of a sub-linear time sparse DFT algorithm. Commun. Math. Sci. **5**(4), 981–998 (2007)
3. Y.J. Lu, Z.H. Xi, J.B. Wang, Analysis and comparison between FFT and digital correlation theory in phase difference measurement [J]. Inf. Technol, **12** (2007)
4. V. Reyes-Rodríguez, M. Jiménez, D. Rodríguez, An error analysis model for floating-point DFT algorithms, in *2018 IEEE 61st International Midwest Sympo-sium on Circuits and Systems (MWSCAS),* pp. 532–535. IEEE (2018)
5. X. Ma, C.H. Lin, T. Jin, Phase difference algorithm and its FFT imple-mentation for high-accuracy power system frequency monitoring, in *2018 IEEE Interna-tional Conference on Consumer Electronics-Taiwan (ICCE-TW),* pp. 1–2. IEEE (2018)
6. S. Sascha. *1.1 The leakage effect*. Retrieved from https://dsp-nbsphinx.readthedocs.io/en/nbsphinxexperiment/spectral_analysis_deterministic_sig-nals/leakage_effect.html (2015)
7. L. Richard, Windowing functions improve FFT results, Part 1. Retrieved from www.edn.com: https://www.edn.com/electronics-news/4383713/Windowing-Functions-Improve-FFT-Results-Part-I (1998)
8. T.A. Shen, L. Chen, J. Guan, J. Duan, A new sliding discrete Fourier transform phase difference measurement method for extreme frequency signals. Rev. Sci. Instrum. ***91***(1), 015103 (2020)
9. J. Charonko, D. Fratantonio, J.M. Mayer, A. Bordoloi, K.P. Prestridge, Windowed Fourier transform and cross correlation algorithms for molecular tagging velocimetry. Meas. Sci. Technol. (2020)
10. S. Liu, Q. Wang, G. Liu, A versatile method of discrete convolution and FFT (DC-FFT) for contact analyses. Wear **243**(1–2), 101–111 (2000)
11. S. Habib, I. Khan, M. Islam, W. Albattah, S.M. Alyahya, S. Khan, M.K. Hassan, Wavelet frequency transformation for specific weeds recognition, in *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*, pp. 97–100. IEEE (2021)
12. M. K. Hasan, et al., Inter-cell interference coordination in LTE-A HetNets: a survey on self organizing approaches. in *2013 International Conference on Computing, Electrical And Electronic Engineering (ICCEEE)*. IEEE (2013)
13. M.K. Hasan et al., A novel HGBBDSA-CTI approach for subcarrier allocation in heteroge-neous network. Telecommun. Syst. **70**(2), 245–262 (2019)
14. M.K. Hasan et al., A novel artificial intelligence based timing synchronization scheme for smart grid applications. Wirel. Pers. Commun. **114**(2), 1067–1084 (2020)
15. M.K. Hasan et al., Femtocell network time synchronization protocols and schemes. Res. J. Appl. Sci. Eng. Technol **4**(23), 5136–5143 (2012)
16. M.K. Hasan et al., Phase offset analysis of asymmetric communications infrastructure in smart grid. *Elektronika ir Elektrotechnika***25**(2), 67–71 (2019)
17. Md Akhtaruzzaman, et al., HSIC Bottleneck based distributed deep learning model for load forecasting in smart grid with a comprehensive survey, in *IEEE Access* (2020)
18. M.K. Hasan et al., Throughput maximization for the cross-tier interference in heterogeneous network. Adv. Sci. Lett. **22**(10), 2785–2789 (2016)
19. M.K. Hasan et al., Cluster-based time synchronisation scheme for femtocell network. Int. J. Mobile Commun. **13**(6), 567–598 (2015)
20. S. Islam, A.-H. Abdalla, M.K. Hasan, Novel multihoming-based flow mobility scheme for proxy NEMO environment: a numerical approach to analyse handoff performance. ScienceA-sia **43**, 27–34 (2017)
21. M.K. Hasan et al., Cluster-based spectrum sensing scheme in heterogeneous network, in *Theory and Applications of Applied Electromagnetics*, (Springer, Cham, 2015), pp. 1–11

22. A.C. Bhatnagar, R. Sharma, R. Kumar, Analysis of hamming window using advance peak windowing method. Int. J. Sci. Res. Eng. Technol **1**(4), 15 (2012)
23. Y. Daren, B. Wen, X. Minqang, The measurement of phase angle difference between periodic signals using FFT and its error analysis [J]. J. Harbin Inst. Technol **3** (1995)
24. A. Wallenborg, A. Eriksson, *Excise Taxation to Protect our Planet: A Point of View from Students in the Global North* (A case from Södertörn University, Sweden, 2018)

Facens Smart Campus Integrated Dashboard: A Use Case Applied for Energy Efficiency

Roberto Silva Netto, Lester de Abreu Faria, Johannes Von Lochter, Emerson Roseiro Filho, and Guilhardo Augusto de Azevedo Rodrigues

1 Introduction and Contextualization

The purpose of this chapter is, in accordance with the concept of Smart Cities, and mainly with an IoT background, to describe how it is possible to develop and implement a unique dashboard, containing all data and systems, being fully integrated, not only for presenting all necessary information, but for generating predictive models focused on energy efficiency, considering the power factor as a function of meteorological conditions.

For the sake of understanding, in electrical engineering, power factor indicates how much power is actually being used to perform useful work by a load and how much power it is being "wasted." It is one of the major factors behind high electricity bills, power failures, and sometimes the imbalance in electrical networks. It is calculated based on resistive and reactive loads, where the resistive ones mean that all the power supplied to it are dissipated for useful work, due to the fact that the current is usually in phase with the voltage, while the reactive ones cause a drop in voltage and draw current from the source like resistive loads, dissipating no useful power, and can be capacitive or inductive. The objective of this chapter is not to go in detail and explain what power factor is, but to understand its impact on the energy network and its correlation with meteorological conditions, creating a model and predicting its behavior.

Nowadays, we can assure that a continuous transformation of energy network management leads to a growing necessity of integrated and interoperable systems to support the so-called smart management in a city. Thus, a more efficient,

R. S. Netto (✉) · L. A. Faria · J. V. Lochter · E. R. Filho · G. A. A. Rodrigues
Centro Universitario Facens, Sorocaba, Sao Paulo, Brazil
e-mail: roberto.netto@facens.br; lester.faria@facens.br; johannes.lochter@facens.br
http://www.facens.br

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_5

environmentalist and technological approach demands a significant research effort. Enhanced usage of communication technologies and the creation of data services enable a transformation of the operational methods, affecting the technical criteria of the system architectures, as well as operational constraints and integration requirements [1].

In this sense, different data sources will be accessed, used, and considered for our model generation, specially different information on weather conditions and weather forecast, which, after studied, showed a great correlation with the power factor measured by our system, leading to an excellent study case.

Thus, data from a living Lab in Brazil was collected and evaluated in a period of 50 months, in which weather conditions showed that it is possible to achieve a high accurate model for prediction of the power factor on energy consumption, allowing to achieve also the sensibility of each one of the parameters to the final results.

1.1 Understanding the Concept of Smart Cities and the Ecosystem for Energy Generation and Consumption

In the late 2000s, the concept of a Smart City has achieved the main focus in a multi-variable domain, in order to overcome the recent and rapidly growing challenges associated with urbanization. Through the use of multiple domains, and the IoT, it is possible to reduce most of drawbacks associated with population welfare, among others, the energy waste, just by allocating the resources more efficiently [2].

Smart, in a common sense, can be defined as “standardizing, monitoring, accounting, rethinking and transforming.” Each one of these elements is an important contribution for the future of cities, in order to achieve the expected challenging goals. While earlier cities could be considered the by-product for industrialization, now they are the node of economic growth [3].

In a planet where it is expected that more than nine billion people will live together until 2050, two-thirds of which in cities [4, 5], sharing and disputing the same resources, it is expected an increasing demand on services and infrastructures, specially in energy, demanding a higher production, or even a more efficient one. Therefore, in order to ensure the efficiency, sustainability, and safety of these urban communities, they should become smarter, integrating massive infrastructure components and services, providing not only more energy, but, on the other side, smart proposals for the best consumption patterns [6–8].

This way, future technologies for Smart City applications, as a whole, are expected to be massive economic sources in a close future, being expected to be worth a cumulative 3.3 trillion dollars just by 2025. Today, private and public companies are developing a number of solutions for Smart City Ecosystems, creating an expanding number of technologies and employment opportunities [2] (Fig. 1).

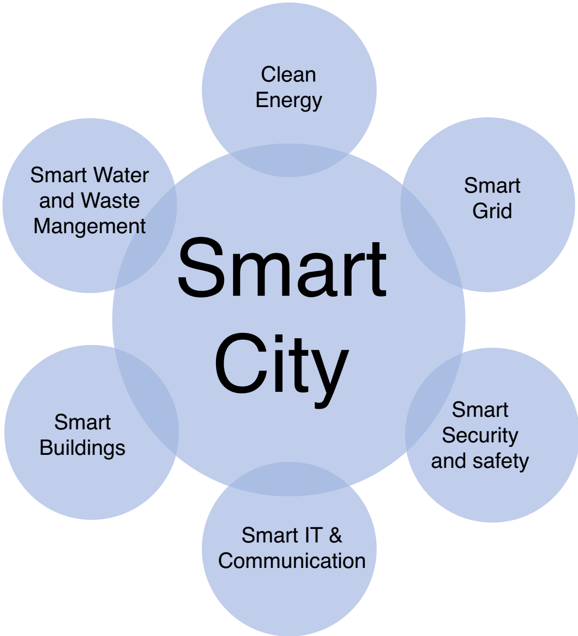


Fig. 1 Smart planning with smart infrastructure (Adapted from [31])

As a general idea of the size of the challenge that we are observing, we know that urban areas consume up to 80% of the global electricity generation and that, up to 2050, more than six billion people (70% of the population) will live in these cities, being also responsible for the greenhouse gas emission [9].

Therefore, when we think on Smart Cities, we shall combine, in a single urban model, solutions aimed at environmental protection, energy efficiency, and economic and social progress, not restricted to those. Inhabitants are, more and more, calling for new urban energy services, from domestic environment to private, challenges that must be addressed for the Academia and Government, both public and private solutions. The energy infrastructures, models, processes, and solutions are clearly the most important urban infrastructures to support the realization of sustainable Future City [9].

When we think the future cities, we arrive to complex systems with a different number of smart elements, all imbedded in their systems, and with a high level of Information and Communication Technologies (ICT) being used to connect them, sharing resources and information [10].

1.2 Internet of Things (IoT) and Artificial Intelligence as Tools

As previously reported, when we think on Smart City systems, we must consider that they demand a large number of sensors and massive datasets, being built with data analytics and machine learning techniques [11]. One of the tools that will enable this ecosystem is called Internet of Things (IoT). The other is the Artificial Intelligence.

The Internet of Things is one of the key technological existing tools, allowing we can realize the full potential of cities transformation. It is an infrastructure for the information age, enabling both physical and virtual interconnections [12, 13]. It is creating a revolution, in all fields. The idea of IoT supports an easy flow of communication between different devices finding applications in different fields like the ones related to, but not restricted to, energy efficiency, such as intelligent energy management, smart grids, and energy constraint, as it is based on batteries and many more [3]. That is, IoT benefits in the Smart City scenario are extensively acknowledged by countless real application deployments, which come as standalone or even domain wide activities within a city [14].

Considering the study case presented in this chapter, IoT is a communication paradigm that makes Internet immersive and pervasive, enabling easy access and interaction with a wide range of devices and parameters, like the ones we use for the weather measurement and weather forecast, as well as their correlation with the power factor of our institution [15].

Therefore, IoT data will be prevalent due to the dissemination of sensor devices in our institution, being strongly integrated for creating novel values, such the one we are proposing here: the influence of weather parameters on the power factor. We can see the critical of this issue to solve real-world problems, assuming that the integration of multiple data can lead to a better prediction and correlation with the actual power factor. The main purpose of this integration is to produce useful data about the surroundings, making them smarter, and providing environments access, based on a collection and analysis of past, present, and future data. Data allows optimal decisions and models, mainly in real time and for future predictions, based on weather forecast. This is only possible by the fusion of various types and forms of data, to enhance data quality and decision-making. Computational Artificial Intelligence plays a key role in this challenge, by working on the collection of various IoT data, and making them to be inter-connectable [13, 16].

Concerning to Artificial Intelligence (AI), it makes possible the study and the better understanding of the context awareness. Previously, this kind of approach and information was limited to the command and control in army operations, or robot manipulation in space, but today we can see it, more and more, applied to everyday services [17].

In this sense, AI can play a crucial role in the IoT paradigm, especially in those areas where we need some level of prediction or we need a decision support. Developing this kind of intelligence and acquiring a high level of accuracy is a challenging process that demands tools based on machine learning, deep learning, data fusion, Fuzzy Logic, besides others. It shows to be a high complex environment that allows us to achieve the best information and situational awareness [13].

Artificial neural networks (ANNs) show an extraordinary capacity to analyze and correlate complex and imprecise data, extracting patterns from complex and big datasets, making data more accurate and precise just by training and learning [13].

Finally, city and each one of its elements will start working as a multi- and interrelated system, with a minimum interaction of human, with a self-operation and fault diagnosis supported by AI models and predictions [18].

At this level of information, decision support and situational awareness are very critical and can be achieved only by data fusion. In the present study case, a multi-sensor data fusion was used to combine information from meteorological sensors and sources of weather forecast. This result enhances the understanding of the environment and of our ecosystem, providing the basis for planning, understanding and for the decision-making process of the energy consumption of our institution [13].

1.3 Data Fusion, Main Challenges, and Integration on a Dashboard

When referring to data fusion, we can find a number of opportunities and tools to support the decision-making process, as well as for prediction and analysis of a certain environment. It shows to be, day after day, the best alternative to analyze data and to correlate different sources of information to understand a specific phenomenon.

In a general way, we can say that, through the data fusion, information becomes more intelligent, decisive, sensible, and precise. Once it comes from multiple sensors, it is possible to understand its meaning as a whole, and not by each data itself, which can be useless. Therefore, it is able to clarify and to highlight information that could be hidden in a polarized (unique sensor) vision. But, as all computational tools, besides the advantages, we can find some drawbacks that must be overcome.

Considering the statistical approach, we achieved that data must be merged in an optimal manner. This is a challenge. But not only merging this data is necessary. It must be filtered and chosen, once sensors can be imprecise sometimes, inaccurate and uncertain. These defects, or outliers, must be filtered by a correct use of data fusion algorithms.

Other challenges show to be the ambiguities and inconsistencies, not only caused by the impreciseness of the sensors but also by the operating environment. The solution for those cases is always the detection, replacement, and data imputation.

Data correlation and alignment are present generally in wireless sensor networks affecting directly the confidence in a data fusion algorithm and are a result of the transformation of all sensor data to a common frame, prior to fusion.

When we talk about the IoT environment, applications may consist of several thousand sensors, detecting different parameters. Data fusion is not a clear and easy process in nature. A perfect data fusion algorithm does not exist, and regular refinement must be performed in every environment and data lakes [13].

But in order to have an adequate visualization of the data and correctly understand different correlations among them, it is necessary to provide a common dashboard where we can show different sensors indications, to agglutinate the data lake that is used to generate models and, at least, try to predict a behavior of the energy system. This data must be as embracing as possible, conjugating both open and private data, static and real-time ones. We have to study and to use several datasets and API sources generating data models, creating applications and models that can be useful for the prediction of energy consumption. Thus, what we did was to develop an application based on a dashboard that may assist the developers to understand the phenomena via visual queries on a graphical user interface, as well as to develop a methodology that facilitates to model the consumption standards and to predict the future, based on meteorological conditions and forecast [18].

It is clear that most of these predictions and models must cope with big data aspects dealing with data volume, variety, and veracity. Drawbacks come from real-time data, as weather conditions and its difficulty and uncertainty of prediction. Then the solution provided in this chapter took advantage of the huge amount of big data coming from several domains of the meteorological station, at different sampling frequencies, for exploiting and analyzing them, making predictions, detecting anomalies for early warning, and for producing suggestions and recommendations to our Smart Campus students and managers [18].

1.4 Challenges Concerning to Power Factor Prediction Based on Meteorological Conditions

Cities consume huge amounts of electricity, but not always we can say that their consumption profiles match the best shapes. Power factors are a problem that shows to be expensive and difficult to adjust, and more, people do not understand its impact on the energy network, not even how to deal with it in order to have better results. Thus, it is generally an inefficient and bad loss of energy that could be avoided if we had a good model and an adequate correlation with the meteorological conditions, as we propose here.

Based on everything that was previously described in this chapter, it is clear that treating with the prediction of the behavior of the power factor it is not an easy task. Data volume, variety, and veracity of meteorological conditions can be a complex and difficult element for creating prediction models. Maybe it is the reason for not existing an accurate and reliable model, in literature, until now. That is the main reason we addressed this subject in this chapter. It is essential to build accurate, real-time, and large-scale geospatially distributed sensing networks in future Smart Cities to monitor weather condition and focus on energy consumption, predicting what will be the best approach for reducing the power factor impact on a sustainable consumption, and using a massive and widely distributed sensor networks generating a massive volume of data, allowing us to adequately analyzing it. Integrating all these infrastructure components and services in a dashboard requires an efficient monitoring and an intelligent decision-making system to ensure reliability of the energy system, requiring, as well, a high-performance computing paradigm to support big data analysis with smart technologies and communications in IoT, providing location-awareness and latency sensitive computing near the data sources [2].

The motivation of everything that was developed here is not only to save money. More than that, it is related to independence, sustainability, reliability, security, power quality, and customer satisfaction. The developed model can coordinate the controlling and monitoring of the city's power demand and energy consumption to improve the power factor and/or schedule it during times when meteorological conditions favor it. Thus reliability, security, and efficiency of the electricity system based on a dynamic optimization of the system are of major importance.

2 The Integrated Dashboard

The challenge that was proposed was enormous, to create an efficient, sustainable, and human campus, which not only observes one dimension of the problem, but also the whole ecosystem, where each element has a relevant role and needs to be observed. Certainly, the use of IoT was a catalyst for this development, as it enabled a vision that was previously unexplored or often unknown.

We understand that the use of ICT is a driver for the development of Smart Cities and Smart Campus as mentioned by Pagliaro et al. [19] based on the six axis model that covers Governance, People, Economy, Environment, Living, and Mobility. But as the University Campus has all the characteristics of a small city [20] we need to pursuit the wellness of the citizen, and this make the technology as the medium and the people the goal of every initiative of digitization.

Starting with small steps, some devices were added and at that time, several solutions from different manufacturers were used. So the campus environment was gaining visibility, resulting in more detailed knowledge, creating a more detailed image on top of a known profile.

This sensor system was gaining complexity, since it was not possible to obtain IoT devices from a single manufacturer, nor using a single communication solution, which further increased the challenge of creating an integrated system, which would provide the necessary vision to build an efficient, sustainable, and human campus.

To help the development of an integrated view where all the data can be consolidated we need to overcome some obstacles. The first one is in the physical layer, where we have some protocols that we can use to transport information. When we climb the protocol stack, we can observe another number of protocols that are used to deliver information to an application (Fig. 2).

This generates a large number of possible solutions for the development of IoT and WSN devices, which makes it difficult to consolidate all this information. For example, in [22] we see that it is possible to use Bluetooth, RFID, Cloud Computing,

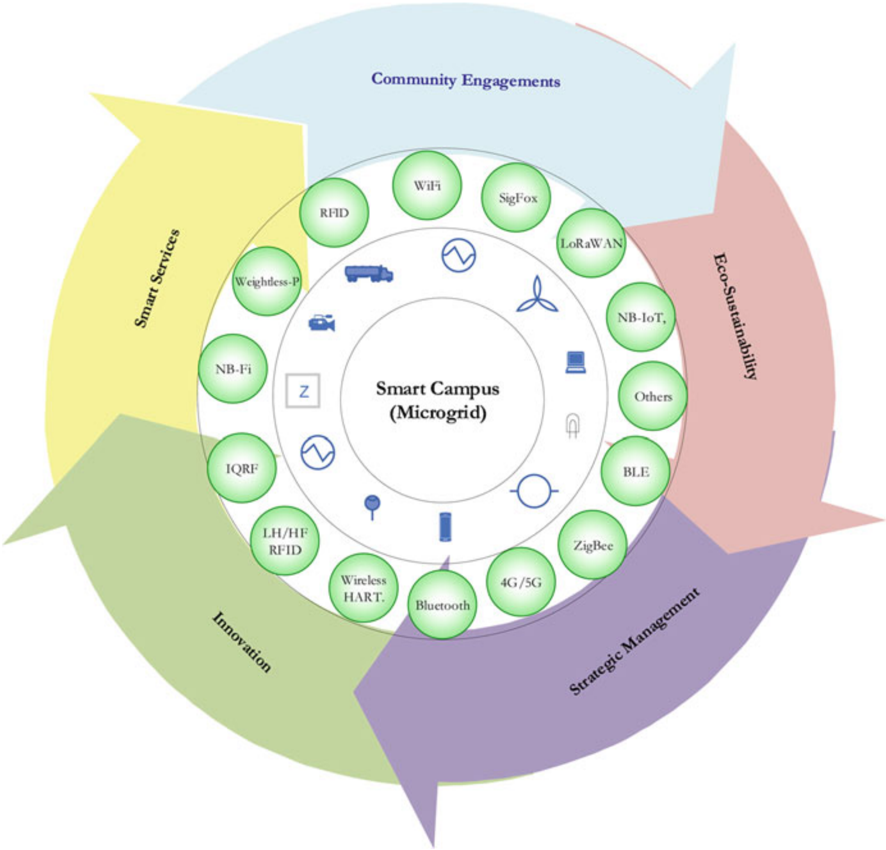


Fig. 2 Communication framework [21]

VR and AR, and NFC, to bring data from the field and insert it as an information in an integrated dashboard that will be used to construct a big picture of your campus and them create a smart environment for the citizens who live inside your city.

Consolidation is necessary, as the devices help to have a more macro-view of the city or the campus. When we look, for example, only at the electrical system, we have a micro-view of a system that makes up a system of systems. However, to have a greater awareness of a city or campus, it is necessary to have several systems, not only of energy, or water, or lighting or climate. All of these systems are necessary and important, but when separated, they cannot deliver a more critical view, nor can they arouse suggestions for correlations between data, because when we look at it, we only have one frame.

So the great challenge faced was the development of an integrated and consolidated dashboard and not a segregated and unrelated view that is obtained when we have several dashboards, each for its respective sensor.

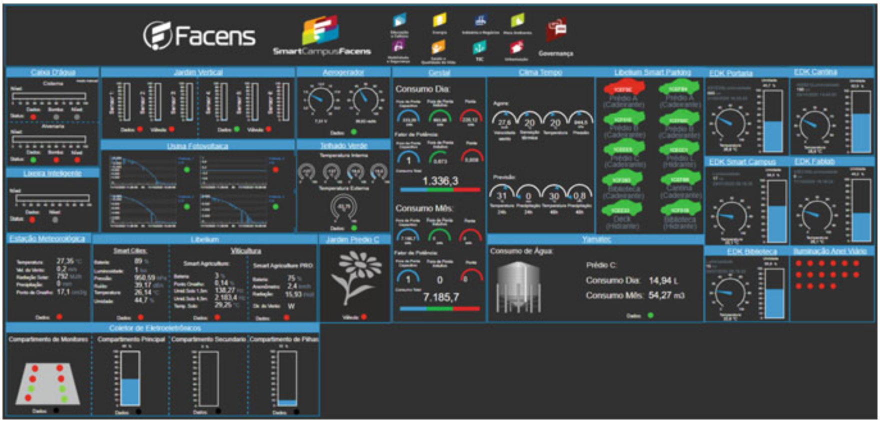


Fig. 3 Integrated dashboard powered by PI System®

Therefore, together with the partnership with OSIsoft through the PI System, a unique dashboard was developed, with all devices and their information, regardless of the physical layer or application protocol. The result of this integration can be viewed at Fig. 3.

2.1 *Data Architecture*

During the creation of the dashboard, it was necessary to develop some features that will ensure that besides the source of the data, they will arrive at the destination in the right format. For this, it was necessary to understand all the paths that the data could take to arrive at the PI System.

For the media that we can use to transmit data, we can mention SigFox, LoRa, NB-IoT, and WiFi.

So, was mapped each sensor, by the type of telecommunication protocol that was used, where the data will arrive, and which format it was sent, as we can see in Fig. 4. With all these information was possible the development of a middleware, capable of receiving data from any sensor and inserting it at a common database. This database is used to feed the PI System architecture and then create a unique vision of the campus.

We start from the sensor that we call "Data Source," and it is mapped to the technology that was used, the medium that data to be sent by the sensor, and which data protocol the sensor use to deliver the sample. As the data can be enveloped by many protocols, it is necessary to have a middleware that can understand it and insert the information in a common database. Then was started the architecture of the PI System that have a SQL Server, where all data from the middleware are stored, and a Data Archive that puts the data in a format that could be used by the PI Vision.

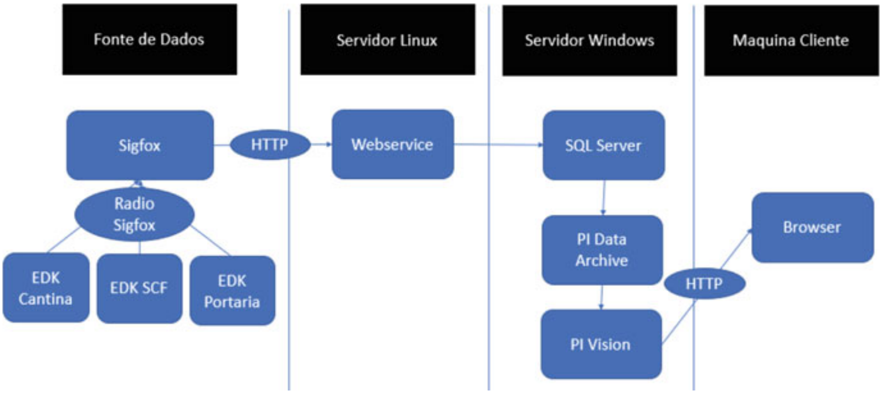


Fig. 4 Data architecture of the system

This last system is responsible for delivering the integrated view of all systems that we have inside the Smart Campus Facens.

2.2 The Ecosystem

Based on the architecture described, the number of sensors was increasing, sometimes through acquisition directed with the manufacture, and sometimes drove by projects involving students. Some challenges were set for the students, and they use the living lab to develop a solution that could attend the challenge. This initiative improves the learning from the students as they can observe the real problems in your neighborhood and can use technology and the Campus environment to develop device that will consolidate the knowledge acquired in the classroom.

One of the first initiatives to construct a Smart Campus at Facens was the insertion of a photovoltaic power source. This Renewable Energy Source (RES) has 60 kW of peak generation, with four inverters that send the data individually and the information is consolidated at the dashboard.

The data flow is modeled in Fig. 5, and it is possible to see the same architecture developed in Fig. 4. But to fit the needs of this system, some adjustment was made necessarily, since in the first interface from the data source as in the second interface, at the middleware.

With the data flow mapped and the integration with our system completed, was designed a single view of some data that the inverters could provide individually and then create a sense of whole RES. The result of this visualization is shown in Fig. 6.

In a second moment, was noticed that it was necessary to get the view of some weather variables and create the understanding of how the climate change could affect our environment. These information could support decision taken from the

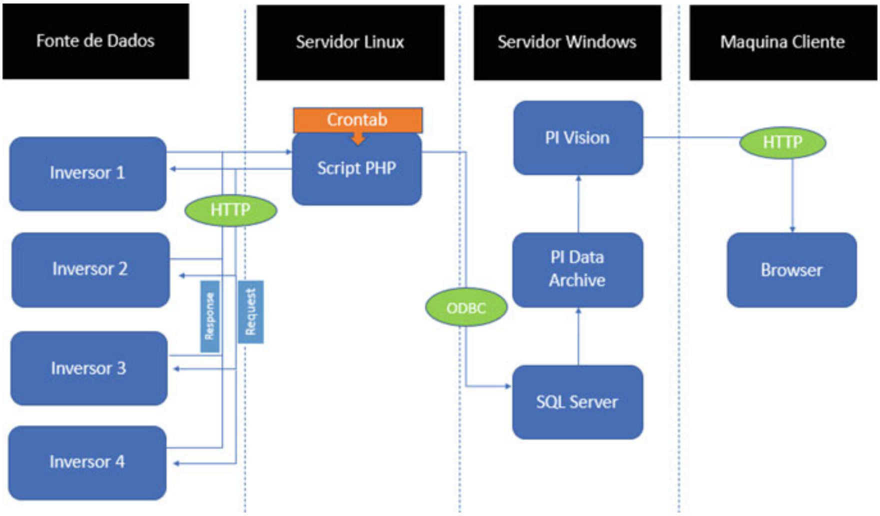


Fig. 5 Data flow of PV system

Fig. 6 Single view of the solution

governance of the Campus, to improve the thermal comfort in the classroom for the students during the hot period of the year.

Based on this necessity, we search for a solution provider that could deliver this type of information and allow to use some wireless technology to send the data to our dashboard. After it has been installed, the integration process starts and was mapped to all the data flows for this specific manufacture.

For this solution, another interface at data source level was developed, which receives data from a LPWAN (Low-Power Wide-Area Network) and should receive in a different format that was developed for other solutions. So, was created another

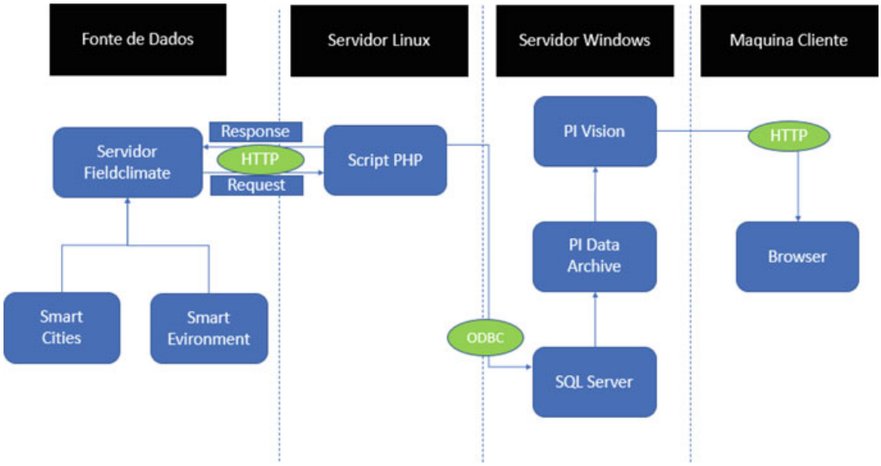


Fig. 7 Data flow of the Smart City sensor platform

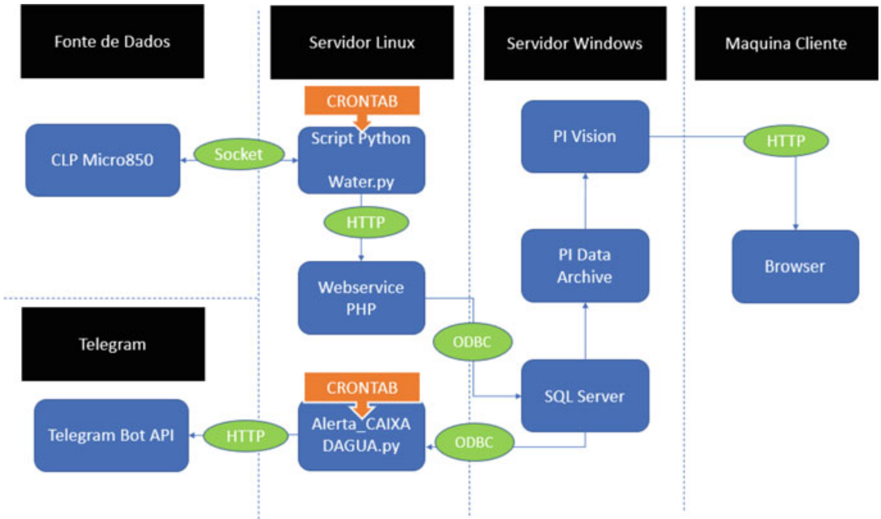


Fig. 8 Data flow of the water supply system

data translator at the middleware to support the data integration for this manufacture as shown in Fig. 7.

After inserting RES in our Smart City, Smart Campus, and the weather sensors, the next challenge is to monitor variables as utilities. The water supply was then chosen to be used as pilot project to understand the challenges in this type of initiative. In our case, it was necessary to monitor the state of water supply, install the sensors, and connect to a PLC that will control two pipe bombs. The dataflow for this system is shown in Fig. 8.

Fig. 9 Water supply monitor

For this solution was designed a view, Fig. 9, to guarantee the visibility of the state of the water supply and the pipe bomb used to get water from a artesian well, pump to another level inside the campus, so the water could be distributed to all buildings.

This ecosystem reflected by the integrated dashboard helps us to achieve a better awareness of the entire campus, allowing the view of power quality by another angle, where we do not look for smart meters in every power consumption device, that in your case do not exist. By the observation of whole picture of the campus, with all the data from many sensors, it was possible to get a insight that should possibly predict the power factor of the entire campus using other types of data, like the weather condition. It is widely used to forecast power consumption [23] and renewable energy resource as photovoltaic [24] and wind power plants [25], but for power factor could be a first initiative.

3 Machine Learning Applied to Power Factor Prediction

The disorderly urban growth and the rise of Smart Cities are struggling with the current infrastructure and energy supply. According to Smart Energy International,[1]

[1] Reshaping energy systems for future cities, available on https://bit.ly/2XQtZxm, accessed in September 20, 2020.

it has several effects for the development of a country, which will no longer be able to accommodate the demands of its society in a sustainable way. It is mandatory to revamp the archaic energy network infrastructures and break reliance on fossil fuels.

Brazil has hydroelectric power as the main source of electricity. The largest hydroelectric reservoirs are installed near the regions of greatest demand. At their maximum capacity, they were able to meet the demands for 3–4 years in the 1970s, shortening it to 4 months nowadays due to energy demand caused by population growth and climate changes [26].

Even though it is a renewable source of energy, the construction of a hydroelectric plant is linked to irreparable and invaluable environmental impacts, such as greenhouse gas emissions, fragmentation of the rivers, changes in water quality, relocating or displacing humans, interference in migration, and spawning fish [27].

Due to recent environmental impact and the concern with fossil fuels, Renewable Energy Sources, such as photovoltaic panels and wind power generators, arise as alternative to hydroelectric power plants. Among many renewable energy sources, the photovoltaic energy is one of the most popular because it is clean and inexhaustible energy, which does not produce noise or heat and requires low maintenance [24]. However, one of the disadvantages of this system is associated with the low efficiency of the current solar panels, which means that much of the sunlight is not converted into electricity.[2]

The most of the countries have a mixed energy matrix. The Brazilian energy matrix[3] is composed of hydroelectric plants (60.9%), fossil fuels (16.11%), biomass (8.74%), eolic (7.52%), import (4.91%), nuclear (1.20%), and solar (0.61%). The planning of which energy source is employed to meet the demand of society is an important matter, since it is possible to deliver electricity provided from expensive (import) and non-sustainable sources (fossil fuels) just when renewable energy source is unavailable. In this way, a prediction system for renewable energy would be helpful to ensure to meet the demand while balancing the consumption between clean energy and expensive or non-sustainable sources.

The meet of demand has some legal issues to be observed as well. The concessionary sells electricity on contracts with service level agreement (SLA) to companies across the world, which formalizes its availability and requirements. When the concessionary is unable to deliver the electricity for a time interval, known as downtime, it may occur a contract violation and a forfeit to the customer. For the customer as service user in the electricity business, to consume more than a previous demand contract it may also occur a contract violation. For these reasons, it is interesting to forecast the generation capacity in a plant.

Although there are many results in the literature about power generation forecasting, most of these works employ more complex features (e.g., radiation and cloud coverage) that are unavailable at meteorological Brazilian services. The remaining of this section outlines a prediction model for power factor using meteorological

[2] Photovoltaic Systems, available on https://bit.ly/2Ho62Yx, accessed in September 20, 2020.

[3] ANEEL - Energy Matrix, available on https://bit.ly/2wxxvD2, accessed in March 10, 2019.

features, like temperature and precipitation, which are commonly available through meteorological services around the world.

Solar energy production rate is closely related to climatic variations [28]. For example, cloudy and rainy days imply large fluctuations in the intensity of solar radiation, lowering the electricity generation. Due to limitations imposed by Brazilian climate centers and unavailability of complex information, our experiment was conducted using temperature and precipitation variables, which is publicly available to use around the world.

Similarly to this work, Yona [29] took solar radiation and temperature as features, using a neural network for its prediction. Their neural network architecture is shallow, and its parameters were defined after a grid search has been performed. The results obtained are outstanding, mainly due to the high correlation mentioned before between solar radiation and energy production rate in photovoltaic systems. Compared to this work, our approach does not include radiation, which is not always available.

Solar radiation was also employed as feature in the work published by Tao [30], but they proposed a different neural network architecture (NARX) from traditional multilayer perceptron (MLP). NARX is an autoregressive model that considers different inputs through the time and models it accordingly. Since the power generation task in photovoltaic systems can be seen as a time-series forecasting problem, it does make sense that the results of the authors showed that NARX performs better than MLP. However, a common problem known in time-series forecasting problem handled by neural networks is non-parallelism. It demands more computational resources to deliver the claimed results.

Shi [24] proposed a different approach combining a weather classification system and photovoltaic power generation forecasting. Both tasks are aligned to improve power generation forecasting task performance. Results are similar to previously published works, but they employed support-vector machines as learning method. It shows that a plethora of learning methods are suitable for this task.

We propose to predict factor power using meteorological data in a similar way many works were already proposed; however, factor power was not studied as output from a predictive model until now. It is important because an imbalanced factor power can produce damage into electrical network and generates forfeit in electrical bill due to Brazilian regulatory laws.

In order to model factor power as function of meteorological data input, we collected 6 months of data from energy consumption using specific sensors, meteorological data from a meteorological station inside our University Campus, and solar panel energy generation from our solar panel farm, because we understand the input of generated energy from solar panel can affect how factor power is seen.

3.1 Datasets

To elaborate this study case, we used three different sources: (1) consumed energy, (2) meteorological data, and (3) solar panel energy generated in the campus. The consumed energy dataset shows the tax period, the active and reactive potency kVa and kvar, respectively (kilovolt ampere and reactive kilovolt ampere), and the potency factor; the meteorological station data have the air temperature and dew point in Celsius, the solar radiation in W/m^2, vapor pressure deficit in kPa, air humidity percentage, precipitation in mm, and wind speed in m/s. The solar panel's produced energy data containing details of the every panel's instantaneous potency production are in kWh, and cumulative data produced in some periods (day, year, ever).

To merge these datasets, some challenges were faced. Each dataset has a different pattern with their own data time range and intervals. For example, the produced energy has more than nine billion rows with intervals close to 15 s for each panel starting at 2019/01/01 and ending at 2020/02/19, while the consumed energy has the same range, but a data interval of an hour between two measures, resulting in nearly 10,000 rows. While we analyzed the energy consumption data, a problem was found making all data before 2019/05/02 needed to be discarded due to failures in sensors when registering each data point. All datasets were standardized to the period between 2019/05/02 and 2019/11/27, sampling them in an hour interval, which enabled comparisons between data sources.

3.2 Exploratory Data Analysis

An exploratory data analysis was performed to have an overall understanding of the different data sources and their correlation. As our main goal is to understand and model a predictive function having power factor as target, we have interest in which a set of variables better describe power factor behavior.

A study was made using a pair plot (Fig. 10) about the behavior between energy consumption, the potency factor (PF) registered in the campus, and its period when data points were collected.

We have many activities across the day, mainly in evening and night periods, so we observed higher energy consumption in those periods. For potency factor, the pair plot had shown an unexpected behavior because potency factor would drastically decrease when energy consumption was low. To understand this behavior better, we investigated exactly when this phenomenon happened.

In Fig. 11 we compared PF against energy consumption in all available data points as time series. In order to enhance its visualization, PF data were multiplied by 100, and energy consumption was divided by 50. As it preserves their original distribution, it does not accumulate many points in bottom or top regions of chart.

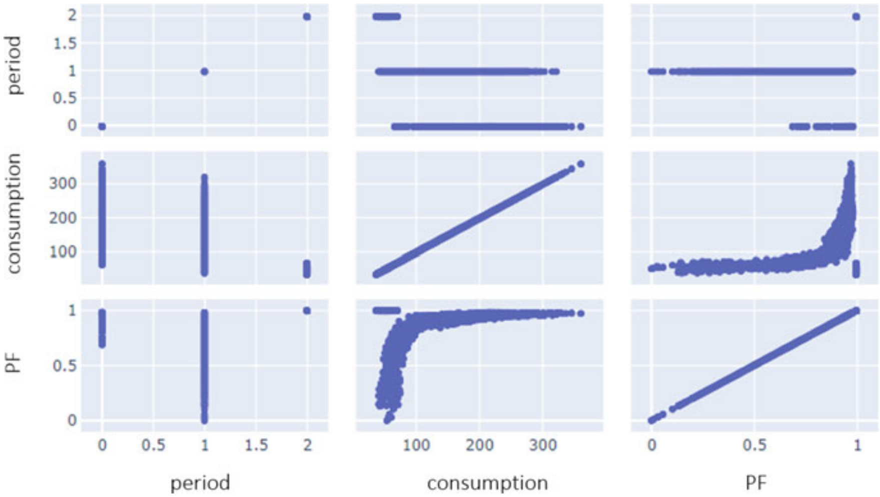


Fig. 10 Pair plot of energy consumption, potency factor, and period to understand behavior of those three variables

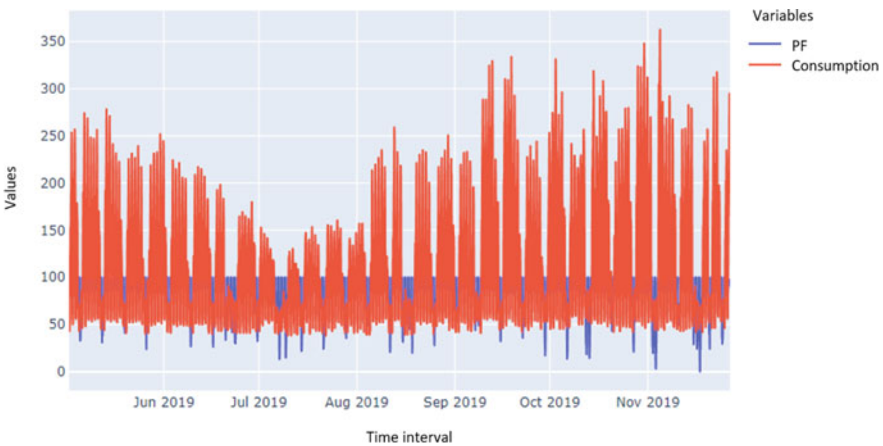


Fig. 11 Comparison of PF and energy consumption as time series

Almost every peak in PF was observed with a drastically low registered point in energy consumption. We also noted that it happened often in weekends, when campus is less crowded, and thus less energy is consumed as well. We raised a hypothesis that solar panel energy inserted in electrical network could have affect power factor in the same manner. As we have an energy source inside the campus in solar panels, we intended to observe its relationship to findings of PF and energy consumption discrepancy.

Figure 12 shows both time series for energy utilization in campus from solar panel energy generated and PF registered.

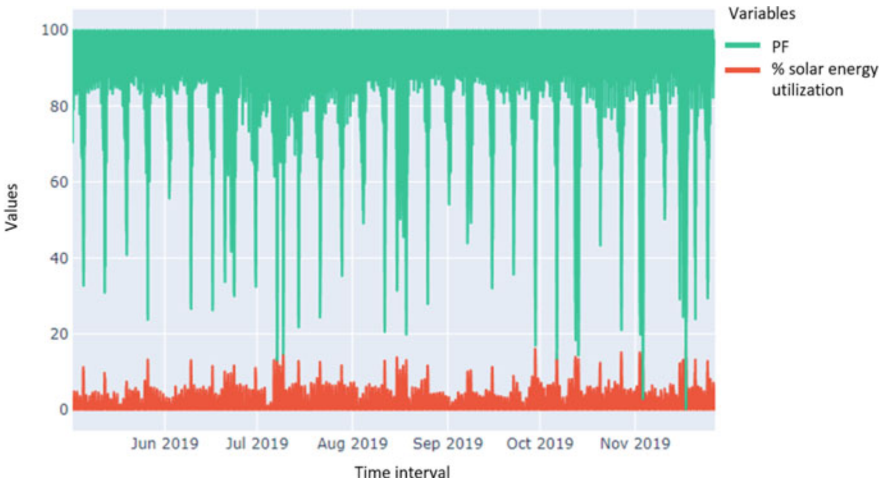


Fig. 12 Comparison of percentage of energy utilization from solar panel against PF as time series

Table 1 Observed correlation of variables to power factor

Variables	Correlation
Air temperature	−0.217
Dew point	0.083
Solar radiation	−0.318
Air pressure	−0.260
Air humidity	0.300
Precipitation	−0.001
Wind speed	−0.130
Energy consumption	0.247
% Solar energy utilization	−0.598
Potency	−0.339

It is possible to realize some pattern, like almost every time PF decreased, the percentage of solar energy utilization in campus drastically increased. This finding highlights that solar panel information has an important relationship to PF behavior. As solar panel is really affected by meteorological conditions, we investigated their correlation among other variables.

For correlation, every data point registered after 6 PM was dropped, since data points collected in period of night do not change on the perspective of solar panels (Table 1).

There is no strong correlation between meteorological features and power factor, as the absolute highest value observed in solar radiation is close to 0.3. It shows that meteorological features, which could be related to power generation, have no correlation with power factor. It is an important finding, because air conditioning machines are high likely to impact power factor, which leads to more investigation.

For such scenario, once the temperature was not correlated to power factor, it was relevant to analyze other variables beyond meteorological features, like those from energy consumption and solar panel energy utilization. As expected, because we already had seen the behavior among PF and percentage of solar energy utilization, this feature has a strong correlation to PF with an absolute value close to 0.6.

3.3 Experimental Methodology and Results

After having a better understanding on how data sources are related, and which variables explain the PF behavior better, we describe in this section details to make results reproducible. As many works in the literature suggest the use of neural networks as learning algorithm to model functions with power generation forecast as target, we extrapolate this to power factor as target and evaluate it using different sets of variables.

Since power factor is a continuous variable, the model should be a regression function like $f(x) = y$, where y is the function answer to an input x, usually a sample containing data that will be multiplied by coefficients learned in a process called machine learning. After exposed to many samples collected in different circumstances, a learning algorithm is able to generate a model that has the capability to infer a result using its input and coefficients learned through the process.

The implementation of neural network was accomplished with programming language `Python 3.6`, using `scikit-learn` library to make it easier this implementation. From this specific library, we used MLPRegressor method, which implements a multilayer perceptron neural network for regression.

A neural network has some important components on its composition, like the number of neurons, the solver chosen to optimize neurons weights, and the learning rate. The number of neurons is important to make complex networks or simpler ones. As the number of neurons increases, more connections are generated, making functions more complex. The solver is an optimization algorithm to change how connections between neurons affect the output of network, which can affect better or worse depending on the task to be handled by neural network. The learning rate defines how much the neural network will learn each step of optimization, where high values can lead to minima local and low values can make the optimization process prohibitively slow.

As neural networks are highly dependent on parameters tuning, we evaluated some configurations to obtain one to enhance the obtained results. For a number of neurons, we evaluated 25, 50, 100, and 200 neurons. For solver, we evaluated three different ones: LBFGS, SGD, and ADAM. The learning rate was evaluated using two different values only: 10^{-5} and 10^{-4}.

Firstly, all data were split into two sets: training (80%) and test (20%). In training set, each configuration was evaluated in a fivefold cross validation using r^2, which measures variance of predicted power factor and true values for power factor from

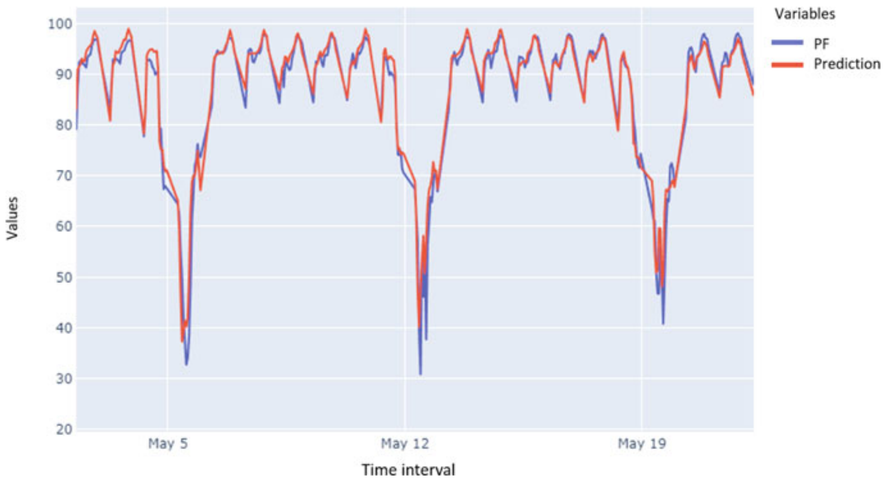


Fig. 13 Comparison of power factor (PF) against its prediction output from a neural network

training set. The best configuration found had a number of neurons set to 100, a learning rate set to 10^{-5}, a solver set to ADAM and achieved $r^2 = 0.93$.

The best model obtained with best configuration was used to predict output for test set. The power factor known for test set and the power factor prediction from model are compared in Fig. 13.

As the model had $r^2 = 0.93$ in training set, it is expected low variance when comparing real power factor to its prediction using same data for each point. That is observed in the comparison, when both lines are close to each other, showing high capabilities to predict even drastically decreases, the main interest in this work. With this model, operating at low error rate as shown, it is possible to conclude that power factor can be predicted using information from energy consumption and solar panel energy utilization.

References

1. T. Atasoy, H.E. Akınç, Ö. Erçin, An analysis on smart grid applications and grid integration of renewable energy systems in smart cities, in *2015 International Conference on Renewable Energy Research and Applications (ICRERA)* (IEEE, Piscataway, 2015), pp. 547–550
2. B. Tang, Z. Chen, G. Hefferman, S. Pei, T. Wei, H. He, Q. Yang, Incorporating intelligence in fog computing for big data analysis in smart cities. IEEE Trans. Ind. Inf. **13**(5), 2140–2150 (2017)
3. M. Ghosal, A. Bobade, P. Verma, Digitized city with IoT, in *2018 Second International Conference on Computing Methodologies and Communication (ICCMC)* (IEEE, Piscataway, 2018), pp. 877–880
4. M. R. Montgomery, "The urban transformation of the developing world," *Science*, vol. 319, no. 5864, pp. 761–764, 2008.

5. K. Davis, The urbanization of the human population. Sci. Am. **213**(3), 40–53 (1965)
6. V.C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, G.P. Hancke, Smart grid technologies: Communication technologies and standards. IEEE Trans. Ind. Inf. **7**(4), 529–539 (2011)
7. W. Su, H. Eichi, W. Zeng, M.-Y. Chow, A survey on the electrification of transportation in a smart grid environment. IEEE Trans. Ind. Inf. **8**(1), 1–10 (2011)
8. H. Chourabi, T. Nam, S. Walker, J.R. Gil-Garcia, S. Mellouli, K. Nahon, T.A. Pardo, H.J. Scholl, Understanding smart cities: An integrative framework, in *2012 45th Hawaii International Conference on System Sciences* (IEEE, Piscataway, 2012), pp. 2289–2297
9. M. Curiale, From smart grids to smart city, in *2014 Saudi Arabia Smart Grid Conference (SASG)* (IEEE, Piscataway, 2014), pp. 1–9
10. M. Postránecký, M. Svítek, Assessment method to measure smartness of cities, in *2017 Smart City Symposium Prague (SCSP)* (IEEE, Piscataway, 2017), pp. 1–5
11. A. Fensel, Keynote: Building smart cities with knowledge graphs, in *2019 International Conference on Computer, Control, Informatics and its Applications (IC3INA)* (IEEE, Piscataway, 2019), p. 1
12. M.C. Domingo, An overview of the internet of things for people with disabilities. J. Netw. Comput. Appl. **35**(2), 584–596 (2012)
13. F. Alam, R. Mehmood, I. Katib, N.N. Albogami, A. Albeshri, Data fusion and IoT for smart ubiquitous environments: A survey. IEEE Access **5**, 9533–9554 (2017)
14. R. Petrolo, V. Loscri, N. Mitton, Towards a smart city based on cloud of things, a survey on the smart city vision and paradigms. Trans. Emerg. Telecommun. Technol. **28**(1), e2931 (2017)
15. M.A. Pradhan, S. Patankar, A. Shinde, V. Shivarkar, P. Phadatare, IoT for smart city: Improvising smart environment, in *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)* (IEEE, Piscataway, 2017), pp. 2003–2006
16. R. Lee, R.-y. Jang, M. Park, G.-y. Jeon, J.-k. Kim, and S.-h. Lee, Making IoT data ready for smart city applications, in *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)* (IEEE, Piscataway, 2020), pp. 605–608
17. K. Hashimoto, K. Yamada, K. Tabata, M. Oda, T. Suganuma, A. Rahim, P. Vlacheas, V. Stavroulaki, D. Kelaidonis, A. Georgakopoulos, iKaaS data modeling: a data model for community services and environment monitoring in smart city, in *2015 IEEE International Conference on Autonomic Computing* (IEEE, Piscataway, 2015), pp. 301–306
18. C. Badii, P. Bellini, P. Nesi, M. Paolucci, A smart city development kit for designing web and mobile apps, in *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)* (IEEE, Piscataway, 2017), pp. 1–8
19. F. Pagliaro, B. Mattoni, F. Gugliermenti, F. Bisegna, B. Azzaro, F. Tomei, S. Catucci, A roadmap toward the development of Sapienza Smart Campus, in *EEEIC 2016 - International Conference on Environment and Electrical Engineering* (2016)
20. J. Torres-Sospedra, J. Avariento, D. Rambla, R. Montoliu, S. Casteleyn, M. Benedito-Bordonau, M. Gould, J. Huerta, Enhancing integrated indoor/outdoor mobility in a smart campus. Int. J. Geogr. Inf. Sci. **29**(11), 1955–1968 (2015)
21. S. Alrashed, Key performance indicators for Smart Campus and Microgrid. Sustain. Cities Soc. **60**(December 2019), 102264 (2020). [Online]. Available: https://doi.org/10.1016/j.scs.2020.102264
22. W. Muhamad, N.B. Kurniawan, S. Yazid, Smart campus features, technologies, and applications: A systematic literature review, in *International Conference on Information Technology Systems and Innovation* (2017), pp. 384–391
23. M.A.M. Daut, M.Y. Hassan, H. Abdullah, H.A. Rahman, M.P. Abdullah, F. Hussin, Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review. Renew. Sust. Energ. Rev. **70**, 1108–1118 (2017)

24. J. Shi, W. Lee, Y. Fellow, Liu, Y. Yang, P. Wang, Forecasting power output of photovoltaic systems based on weather classification and support vector machines. IEEE Trans. Ind. Appl. **48**, 1064 (2012)
25. X. Wang, P. Guo, X. Huang, A review of wind power forecasting models. Energy Procedia **12**, 770–778 (2011)
26. V. de Souza Dias, M. Pereira da Luz, G. Medero, D.T. Nascimento, An overview of hydropower reservoirs in Brazil: Current situation, future perspectives and impacts of climate change. Water **10**, 592 (2018)
27. P.M. Fearnside, Greenhouse-gas emissions from Amazonian hydroelectric reservoirs: the example of Brazil's Tucuruí Dam as compared to fossil fuel alternatives. Environ. Conserv. **24**(1), 64–75 (1997)
28. M. Campos, L. Alcantra, Interpretação dos efeitos de tempo nublado e chuvoso sobre a radiação solar em belém/pa para uso em sistemas fotovoltaicos. Rev. Bras. Meteorol. **31**, 570 (2016)
29. A. Yona, T. Senjyu, A. Saber, T. Funabashi, C. Sekine, H. Kim, Application of neural network to 24-hour-ahead generating power forecasting for PV system, in *IEEE* (2008)
30. C. Tao, D. Shanxu, C.C., Forecasting power output for grid-connected photovoltaic power system without using solar radiation measurement, in *2nd IEEE International Symposium on Power Electronics for Distributed Generation Systems* (2010)
31. Blogs, https://collaboration.worldbank.org/content/sites/collaboration-for-development/en/groups/the-smart-cities/blogs.entry.html/2016/03/14/how_gis_supportspla-ey3X.html. Accessed 19 March 2021

Cloud Internet of Things in Medical and Smart Healthcare Applications

Shayla Islam, Mohammad Kamrul Hasan, Ahmad Fadzil Ismail, and Imran Memon

1 Introduction

Healthcare is the act of maintaining or improving a person's health through various methods, the most common ones being through the examination of body health and improving health with medication such as therapy or medicines. Especially for older adults, having access to good healthcare makes sure that people are getting treatments that cannot be done at home. Life expectancy of adults 65 years old and above increases by 2–2.5 years compared to those without access to healthcare [14]. Establishing high-quality healthcare facilities is difficult as it is costly and cannot be done unless there are enough people in the area to make sure it is still able to stay in business. As such, oftentimes denser places such as cities will have better healthcare facilities, while areas such as the countryside will only offer basic treatments. People working in the cities also generally have higher earnings and thus will be able to afford better healthcare and thus further widening the gap of the healthcare quality received by the society. Cloud computing is one of the technologies that are more and more implemented as time goes on. It is a model

S. Islam
Department of Computer Science, UCSI University Malaysia, Kuala Lumpur, Malaysia

M. K. Hasan (✉)
Center for Cyber Security, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
e-mail: mkhasan@ukm.edu.my

A. F. Ismail
Department of Electrical and Computer Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia

I. Memon
Department of Computer Science, Bahria University, Karachi, Sindh, Pakistan

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_6

in which a host offers computing powers to the consumers through the Internet. Consumers will then pay to the host for the service and convenience of not having to host and maintain the server hardware on their own.

The types of services can vary between the service provider, and it is common for a provider to offer many highly used features in hopes that consumers will not need to search for other services and stay inside the provider's ecosystem. Most providers offer two types of cloud service, storage services and virtual servers. Storage services commonly act as a method for the consumers to store a large amount of data either to be accessed daily or as a backup in case other storage locations lose their data. Virtual servers will give consumers a virtualized system consisted of a certain number of allocated hardware (CPU, RAM, storage) of which the users are free to use and to handle the server traffic or task, and if the consumers find the given hardware is not enough, they can pay more to be allocated more hardware to handle the task well.

One of the major advantages cloud computing offers is scalability and pay-as-you-go service, which means that users that have low workload will only pay a small amount and if the users see that they'll need much more computing power they can pay more temporarily to handle much higher workloads. This convenience of rapidly scaling up and down means that the cost of doing the same computing load will be lower for the consumers. Consumers also now have easier access to analytical data that is only possible with higher processing power. This can be achieved as cloud computing is made to handle and process a lot of data. Reliability is also one of the major features consumers are switching to cloud computing. If the computing machines are self-hosted, then the business will also need to handle the maintenance and repair the hardware in the instance that some of them broke down. With cloud computing, the servers are set up to run 24/7, and in the case of an error, it will be resolved as soon as possible. Storages are also handled in a way that the server can handle several broken storage disks without losing the users' data.

The Internet of Things (IoT) has been predicted to rise in usage, and in the past few years, there has been an increase in IoT devices, while the cost of the devices is going down. One of the most prominent and popular IoT devices is a fitness tracker as it is very affordable and offers relatively accurate body tracking such as exercises, heart rate, and sleep cycle. Existing devices with IoT are starting to emerge as well as Internet accessibility is high enough in some countries and the device itself is priced not much more expensive compared to the non-IoT versions while offering conveniences. The most basic smart speakers that are IoT compatible are most of the time only slightly more expensive or just as cheap as normal speakers yet offer virtual assistant services so users could access information from the Internet anywhere from their house without ever needing to touch their phone.

One of the new uses of IoT is smart devices that track users' health through various data. As mentioned previously there are smart wearables, but there are also other devices such as smart scales that will integrate the data into the application which will then be shown to the users. There have been discussions to implement smart toilets to help warn users early of any potential health risk or disease [15]. This combined with multiple other data if combined and processed can give accurate representations of the users' health without needing them to go to hospitals to get basic health examinations.

The Internet of Things can also be implemented for healthcare. As it is connected to the Internet devices that can measure users' health such as a blood pressure monitor can send the data to the cloud. With multiple devices measuring various data, there is also the need for cloud computing to make sure all the data is connected and processed to give the correct diagnosis. This paper will discuss how the inclusion of the Internet of things and cloud computing can improve healthcare quality.

2 Background

Cloud computing is a unique advanced technique that mainly deals with big data to enhance the operations of the store and retrieve, update, and process the data. Cloud computing uses the Internet to provide its services to the users and developers of unlimited storage space; it's used in most applications and organizations to enhance the user experience and for more advanced innovative applications in the world. Cloud computing is involved in many modern techniques and fields like government applications, educational applications, and many more.

The IoT is a technique that enables the ability to connect multiple objects wirelessly using the Internet. It's one of the most thrilling techniques in modern days and the research field of networking. The Internet of Things depends mainly on the Internet to connect "things" that deal with sensors and incredible amounts of data that need to be stored, which indicates the need for cloud computing in IoT applications; IoT is providing the ability to develop major smart applications and is opening the way to the future of fully smart cities.

Health has been one of the most important components of life. The majority of people consider health as more important than money, and it is often the number one priority in life; hence, there are many advanced methods and techniques to implement smart health applications and equipment to enhance the overall medical field which will positively affect the world economy and personal health.

All experts in both medical and information technology fields are encouraging and supporting the use of the Internet of Things and cloud computing in the medical section in multiple applications and devices, as the IoT is responsible for connecting millions of medical objects that might be a smartphone or any machine in a form of large frame network to enhance the medical section; however, it requires exchanging a huge amount of data; therefore, the use of cloud computing greatly improves the operations and the data security as well as the user experience.

3 Legacy Work and Literature Review

In this section of the paper, a detailed intense study of multiple legacy review papers and research is thoroughly analyzed to precisely manage and generate a comprehensive study of the contributions and positive effects of those papers; in contrast the limitations of the papers are discussed to enhance and improve this paper.

A review paper titled "Cloud Computing in Healthcare – A Literature Review on Current State of Research" [1] has been published in 2016 majorly having cloud computing as its topic and discussing its services and implementations in healthcare. This paper is considered an impactful recourse consisting of 36 articles on the use of cloud computing in the healthcare; the literature review has covered a large amount of useful data of multiple aspects of the topic by following a four-phase categorical method of the cloud computing and the architecture, cloud computing implementations on the medical devices and applications, the privacy and security, and finally the benefits of the cloud computing on the health sector. The research has shown the importance of cloud computing by analyzing its implementations on multiple medical services; an example is the emergency ward applications. In 2015 [2] a comprehensive survey is conducted of the IoT implementations as of the majorly used IoT applications in the health sector and the multiple devices that are connected through the network; this paper is considered a highly influencing resource as it analyzed the different types of IoT application that are used in the health sector, as well as how the big data is being processed and handled, and this paper has contributed in analyzing must challenges of the security and the protection of data. The result of this survey is an IoT model that is highly secure with advanced security features.

In 2018 [3] another systematic review paper is published covering major articles in a period of 6 years from 2000 to 2016 on different topics related to the use of IoT in medical care. The purpose of this study is to answer the research questions of this paper: what are the main areas applications of medical health using IoT, a detailed study of the IoT architecture components, the major Internet of Things techniques and services used in the health sector, the behaviors and features of cloud computing as a hybrid system with IoT and its contribution in the health sector, and the security challenges of the IoT architecture. This paper highly contributed to the IoT architecture security challenges as it was discovered that there were a relatively small number of research papers related to this topic. "A Survey on Internet of Things and Cloud Computing for Healthcare" [4]: in this research paper, the IoT and cloud computing uses were intensely objected to a high level by implementing a comparison table of multiple papers with a duration of 4 years from 2015 to 2019 (Table 1).

The paper contributed in understanding the satisfaction level of both patient and medical personnel in using the IoT medical applications, therefore improving the user experience and highly focusing on the data side and the implementations

Table 1 Review study of legacy work

Paper title	Year of publication	Topic	Advantages	Disadvantages
"Application of IoT in Predictive Health Analysis-A Review of Literature"	2020	IoT	1. A review study to enhance overall knowledge 2. Study the positive effect of IoT medical applications on the economy 3. Provides a detailed study of the IoT medical solutions	1. No comparison table or graph 2. Limitation of reviewing only the medical application used
"A Survey on Internet of Things and Cloud Computing for Healthcare"	2019	IoT, cloud computing, fog computing	1. Comparing between cloud computing and fog computing 2. Contributing to investigating the privacy of the applications 3. Deep analyzing security models	1. The limited number of review papers
"Internet of Things for Healthcare Using Effects of Mobile Computing: A Systematic Literature Review"	2019	IoT	1. Understanding the advantages and features the mobile computing could provide for the IoT medical applications 2. Contributing to analyzing the medical applications of IoT in the use of M-health 3. A detailed review study of 116 articles 4. Security challenges novel study	1. Limitation of cloud computing and dealing with big data topics

(continued)

Table 1 (continued)

Paper title	Year of publication	Topic	Advantages	Disadvantages
"The application of the internet of things in healthcare: a systematic literature review and classification"	2018	Mainly IoT and characteristics of cloud computing	1. Covering 60 articles from equally high-rank journals 2. Understanding the IoT technologies and main use in medical applications aspect 3. Contributing to the security and confidentiality of e-health applications	1. The low number of security articles for the review
"The Internet of Things for Health Care: A Comprehensive Survey"	2015	IoT and cloud computing	1. Analyzing the IoT techniques, architecture, and its positive effect on the medical section 2. Proposing an intelligent IoT security model for enhanced innovative features	1. Only the applications review 2. No comparison table of the review legacy study
"Cloud Computing in Healthcare – A Literature Review on the Current State of Research"	2013	Cloud computing	1. A heavy analysis using 36 articles on cloud computing and big data management 2. A major variety of applications in emergency wards, appointment and communication applications, smart hospitals, and image processing	1. Limitation of security and privacy analysis due to the small number of related papers 2. No comparison table

of both cloud computing and fog computing and the difference between them in services security and architecture.

In this [5] review paper, the use of mobile computing in the IoT medical applications was analyzed. The study is based on over 100 resources to establish and get the answer to 4 research questions which are the effect of mobile computing in IoT mainly on the medical sector, how the security and privacy elements in IoT are enhancing with the use of mobile computing, what are the main applications and medical devices that are developed in the use of IoT and mobile computing, and lastly the use of cloud computing to secure big data with IoT. The paper used a systematic review methodology protocol of six phases starting from the defined protocol to the quantitative data analysis.

Lastly to manage an advanced detailed literature review, a paper from 2020 [6] is analyzed. This paper initially highlighted the importance of IoT in medical health as a live or dead situation to enhance the patient health. IoT is a network to connect people and devices and innovatively exchange data. The literature review was conducted based on peer-review papers from journals or web articles. This review study only analyzed the different applications of IoT and cloud computing in the medical section. To have a more understandable analysis of the legacy works, a comparison table was created containing the review works as well as its advantages and disadvantages. From the table above and the detailed analysis, the conclusion is that although there are many contributions and advantages in the review papers, negative points are found as well that should be solved and highlighted in the paper study.

4 The Health Applications and Equipment

The IoT is becoming the key to the 4.0 industry evolution as it is used in highly innovative applications to rise to the future of smart cities and convenient life; thus, it is used in many applications, with the most important and significant use in the field of medical health. However, as the Internet of Things creates a wireless network of sensors and devices, it requires a high level of privacy and confidentiality of the data in the network. To complete a system of IoT, cloud computing is essential for completing the system data life cycle as the data in the network needs to be stored and processed. A simple application scenario of the IoT and cloud computing is a heart rate monitor application. In this type of application, there are two key points of users, the patients and the medical personals. Furthermore, the data are collected from the patients as there will be sensors to detect and read the patient heart rate. This data will be exchanged through the IoT network to the main manager which is the processor (cloud computing). The next step requires the information to be sent to the medical personal for analysis, and if there is an emergency like an abnormal situation of the patient, the application within the medical personals will send an alert code to save the patient as seen in Fig. 1.

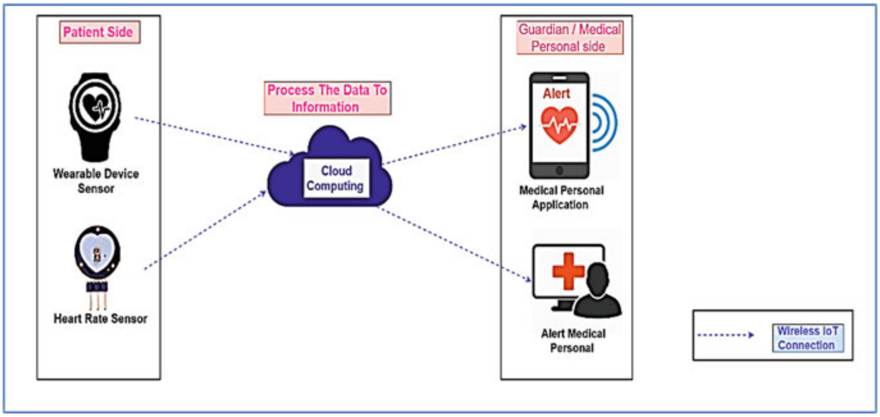


Fig. 1 Scenario of how the IoT and cloud computing work in an application

As seen in the scenario above, the architecture is almost similar to all of the emergency applications of heart monitoring, body temperature, and blood pressure, and much more life-saving technique will possibly reduce many financial costs, therefore improving the economic situation as well as continuously impacting the evaluation of people which was made possible thanks to the IoT and cloud computing.

Besides the abovementioned application of saving lives, IoT can connect people using medical wearable devices to monitor the heart conditions and fitness activities. These types of applications are used by patients and people with a family or genetic history of chronic diseases to provide a health risk-free life [20–26].

In the medical section and the use of IoT with cloud computing-based system, many applications will shape the future and is made to provide a higher healthier personal and improving the overall finance of a country; there are numerous applications to cover, and many more are still in the developing and testing area [20–25]:

1. **Telemedicine applications**: These types of applications provide a communication network to exchange data and resources between personnel; there are many types of applications under this category. It might be used to connect the patients with the hospital medical advisors and the emergency center. Those data collected from patients could also be shared with family members or to create a historical record of the patients. Another type of communication is to connect people wanting to lead a healthy lifestyle or having a chronic disease like asthma to smart hospitals for consultation and emergency, using devices like smartphones or wearable watches alike.
2. **Booking applications**: This is a subcategory of telemedicine applications with advanced features of connecting outside patients to the smart hospital administrator office for booking purposes to provide a convenient smart user experience.

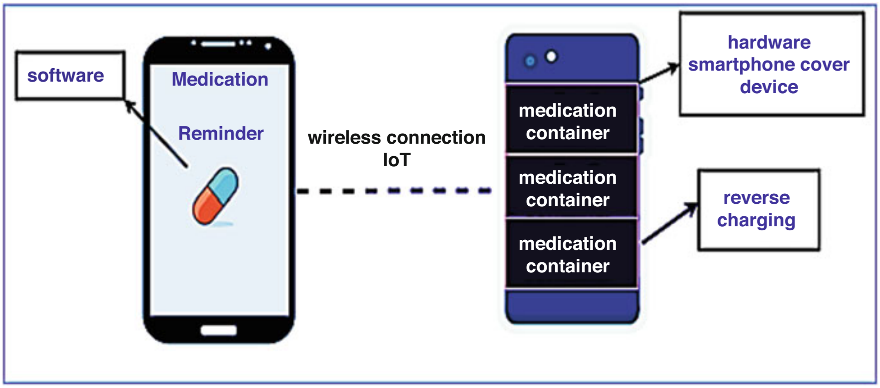


Fig. 2 Novel medication reminder application architecture using IoT [2]

3. **Medications reminder**: Based on a recent study, it was revealed that almost 50% of the population forgets to take their medications [7], which goes back to the reason of forgetting and taking wrong pills at incorrect timing. This nonadherence issue has a major effect on the health and the recovery speed; furthermore, it could lead to death for people with chronic diseases. IoT and the use of cloud computing significantly decrease the nonadherence percentage to none and therefore dramatically increase the health and save lives by the medication reminder applications, which works by connecting the application using IoT to a smart container that holds the medicine as illustrated in Fig. 2.

Based on the multiple uses of IoT and cloud computing in the medical industry and the applications, a comprehensive study of the various types of applications is illustrated in the table below with the positive and negative effects (Table 2).

5 Discussion and Analysis

Telemedicine applications are a combination of expertise and technology that helps to deliver medical services and information over distances. In fact, telemedicine has a variety of applications, for instance, education, research, videoconferencing, diagnosis, consultation, treatment, nursing call center, and monitoring remotely. The rapid growth of technology results in millions of physicians and healthcare providers accessing the web, transforming traditional service to remote consultation and distant learning.

The Internet of Things (IoT) is defined as numerous physical devices around the world that are connected to the Internet where all of these devices collect and share data. As the development of technology grows, the computer chips are becoming

Table 2 Comprehensive analysis of the medical applications using IoT and cloud computing

No.	Title	Application and method	Advantages and contributions	Disadvantages
1.	"Remote Heart Rate Monitoring System Using IoT [8]"	A medical monitoring system for the heart rate using Raspberry Pi3 sensor and controller with the IoT and cloud use	1. Decrease the cost and hospital visitation fee 2. Monitor the heart rate remotely using a wearable device	1. The large hardware and the configurations 2. Limitation of giving the user instructions on how to handle the abnormal rate
2.	"The IoT-based heart disease monitoring system for pervasive healthcare service [9]"	The use of IoT technique to provide a monitoring service of multiple tests (blood pressure, EGG, blood fat, a heart pulse) and send the data to the medical applications for consultation	1. Provide a complete package of the different tests based on one 2. Small hardware implementation 3. Eliminate any risks and save patients by sending alert data	1. Limitations of the early warning to users 2. Limitations of the data management, as this system revolves around the data and its availability; no measurements or methods are introduced
3.	"IoT for Telemedicine Practices enabled by an Android™ Application with Cloud System Integration [10]"	The use of IoT and cloud computing is a monitoring device for communication between the anesthesia patient and the doctor, as the patient is insensitive to the pain; the required data collected by the Raspberry Pi will be sent to the doctor tablet, and if an error or emergency is accrued the alert will show in the doctor wearable watch	1. Provide a consistent monitoring application to reduce any risks by sending all the patient data to the doctor in charge remotely 2. Significantly contributing to a smart hospital and saving time for doctors therefore continuously saving other patients	1. Limited to android mobile users

4.	"A Smart Pill Box to Remind of Consumption using IoT [11]"	In this system, the hardware part is the pillbox designed with a real-time clock to remind users by sensing an SMS or email	1. Significantly decreasing the nonadherence percentage	1. Limited to only send a message to the user or guardian, which makes it complicated and doesn't solve the issue because of the need to count the pills
5.	"IoT based Advanced Medicine Dispenser Integrated with an Interactive Web Application [12]"	It's a reminder application that works by using IoT to connect the dispenser to the application, and it uses cloud computing to manage and process the data and using machine learning	1. The test included has shown a positive dramatic increase in the overall health and lifestyle of the users 2. The use of machine learning to establish a smart detector	The test conducted in this paper revealed a few limitations: 1. As the dispenser is of large size, users tend to not take it outside, thus resulting in not taking medication on time 2. Limited to pills only
6.	"Application of mobile cloud computing in emergency health care [13]"	This application mainly focuses on the use of cloud computing and the correct management of big data; the emergency application provides an efficient way to easily detect the emergency patient location and allow the user to choose the nearest hospital	1. An effective convenient way to report an emergency patient by just clicking the app which will detect the location 2. Providing a registration option to the hospital before arriving to decrease the amount of time the patient must wait	1. Limited to only providing the database search of cloud computing; thus a video call option with early diagnosis is required

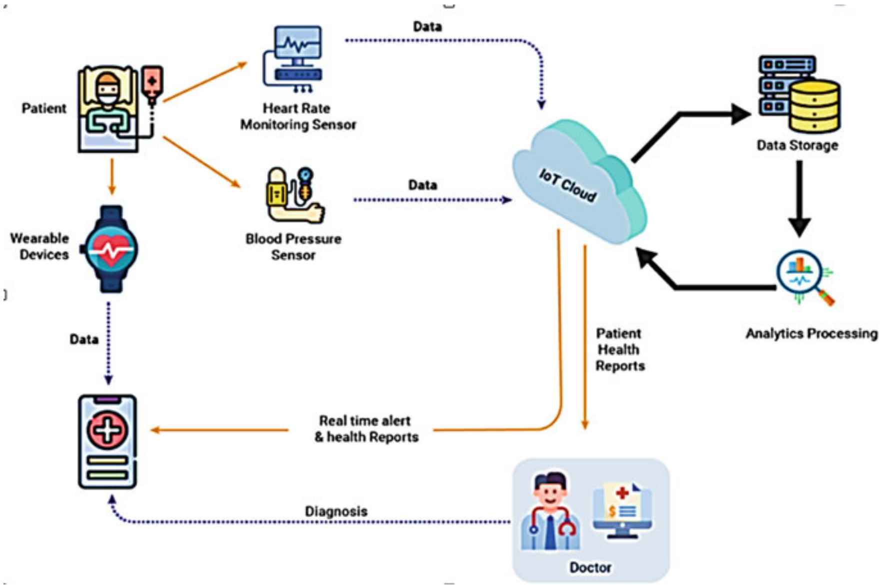


Fig. 3 Digital transformation of healthcare applications [14]

cheaper and smaller year by year. Therefore, the devices that can connect to the Internet are from the size as small as a pill to as big as an aircraft.

IoT provides solutions for different applications that include many aspects of life, for example, hot topics like smart cities, smart traffic management, security, emergency services, retail, supply chain, health monitoring, and healthcare. Based on a report by CISCO [14], there are a total of 500 billion physical devices that will be connected by 2030; it means that each person on our planet owns 58 smart devices equivalently. Figure 3 shows how IoT and cloud computing helps in transforming the healthcare sector.

On the other side, cloud computing is the delivery of various computing services such as servers, data storage, databases, networking, and software. Cloud computing is becoming popular over these years for people and businesses; this is because of several benefits that include cost-effectiveness, enhancement in productivity, speed and efficiency, performance, and also security [15]. With the integration of the Internet of Things and cloud computing, all aspects in the healthcare sector can gain a better improvement. Generally, the related professionals in the healthcare sector are able to provide better and more efficient services to patients.

5.1 Concepts

Before the integration of the Internet of Thing and cloud computing, healthcare services were only limited in physical visit or telecommunications. The greatest changes after the implementation of IoT and cloud computing are enabling healthcare providers to monitor patients' health conditions remotely and making these medical services available over distance. Moreover, IoT and cloud computing also help to improve patient engagement and satisfaction as communications have become more efficient and accessible. As mentioned in the earlier part, one of the benefits of implementing IoT and cloud computing is cost-saving as it reduces the duration of hospital stay and avoids hospital readmissions.

5.1.1 Ambient Assisted Living (AAL)

Ambient assisted living refers to the usage of information and communication technologies in a person's daily life and working environment to help them to stay active longer, remain socially connected, and live independently especially for elderly people. AAL aims to make lives easier and also self-dependent. The related fields of requirement to study include health, safety, independence, mobility, and social connectivity.

The advantages of using AAL are plenty. First, it helps to reduce the number and effort of caregivers. This is because AAL devices give sufficient support to users in order to live independently. Elder people can control simple things like switching on or off the television and lights easily through the use of AAL. Other than that, AAL technology equipped with an emergency response system helps to maintain a high-level safety through the usage of smartphone or other smart devices. For example, features like fall detection or video surveillance help other users to track and take quick action in case of any emergency.

The other advantage is pill dispensers. As people grow, it is common we take medicines to maintain our normal health condition. Elderly people often miss taking medicine due to poor memory. As such, the pill dispenser plays an important role to remind them to take their medicine on time.

5.1.2 Wearable Devices and Smartphone

This can be an innovative solution for healthcare problems in this modern era. By wearing these smart devices, wearable technologies are able to monitor human physical activities and behaviors and even physiological and biochemical parameters continuously during daily activities. Generally, the devices are able to detect and measure data such as body temperature, blood oxygen saturation, calories burnt, posture, physical activities through the use of electrocardiogram (ECG), and ballistocardiogram (BCG). Currently, wearable devices can be smartwatches,

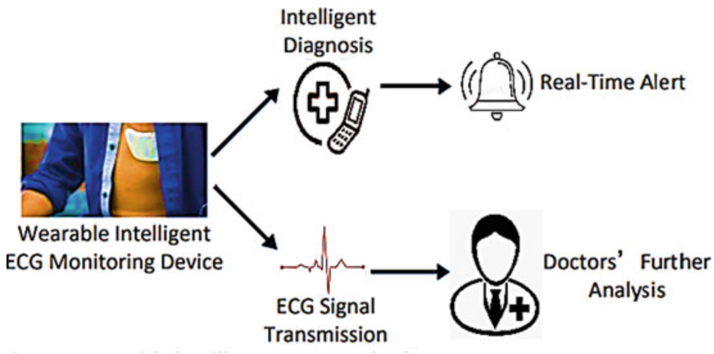


Fig. 4 The integration of IoT in ECG monitoring [16]

shoes, eyeglasses, necklaces, etc. In the future, it may evolve or transform into skin-attachable devices. Besides, sensors or chips also can be embedded into the physical device or environment such as wheelchairs, car seats, mattresses, or others. Basically, these wearable devices help to analyze and transmit data to a cloud computing-based architecture for storing, processing, and analyzing the data. Then, through another health application on smartphones, users can view their data.

5.2 Healthcare Applications

5.2.1 Electrocardiogram Monitoring

An electrocardiogram test helps to monitor human heart electrical activity. The patterns of ECG can screen and diagnose different cardiac problems. In other words, it helps to check if the heart is working properly or is experiencing some abnormal problem. The information that ECG test gives is valuable in early prediction of heart diseases (see in Fig. 4).

The integration of IoT in ECG monitoring aids to alert potential patients about their abnormal heart rate, which can lead to other heart disease. Recently, there are few smartwatches invented with the implementation of IoT to track heart rate. In [16], an energy-saving framework is introduced; it performs ECG compression for wearable gadgets in real time. Optical heart rate sensor is another excellent tool to collect data like on-the-spot readings or resting heart rate data, which can indicate current health situation.

5.2.2 Body Temperature Monitoring

Another important variable to diagnose a patient's health condition is body temperature. The change of body temperature, either too high or too low, is also a vital sign for some illness. The traditional approach in measuring body temperature is always using a thermometer that is attached to the human body; however, it is not so convenient to track body temperature in that manner in our daily life. The latest invention in this aspect is one kind of hearing aid [17] that consists of an infrared sensor, printed by intelligent 3D. This can be worn comfortably on human ears. Besides tracking the human ear temperature, it is also equipped with a microphone and amplifier that helps in hearing.

5.2.3 Blood Pressure Monitoring

Blood pressure monitoring is important as it helps with early diagnosis. Daily blood pressure monitoring allows people to create a baseline that can provide useful information regarding treatment outcomes. According to authors in [18], a wearable cuffless gadget is introduced to monitor blood pressure. Then, the recorded blood pressure is saved in cloud storage. In this study, the proposed gadget gains accurate blood pressure data with no mistake or error in systolic and diastolic blood pressures. Other than that, in [19], there is a prototype of a monitoring system that evaluates blood pressure based on fingertips. The signals are collected by obtaining ECG and photoplethysmogram (PPG) and then transmitted to a micro-controller; thus, blood pressure was computed.

5.2.4 Asthma Monitoring

Asthma is a condition where the airways get inflamed and narrow, and it causes difficulty in breathing. This is a long-term lung disease. Patients who suffer from asthma often experience temporary shrinking of airways, which further causes coughing, wheezing, shortness of breath, and chest pain. An IoT asthma monitoring system is introduced in [20–26]. This system is equipped with a heart pulse sensor; then, the data is collected from the sensor and sent to a micro-controller wirelessly in real time and then transferred to a server lastly [27, 28]. At the server, a database is deployed to gather and store the data. Therefore, the healthcare provider can access this database to observe the latest update of patient health conditions.

6 Conclusion

The study has dived into the topic of Internet of Things (IoT) and cloud computing in the healthcare industry. We have determined the architectures of IoT, cloud computing, and how can these architectures improve the existing healthcare systems. For deeper understanding, the brief background of IoT and cloud computing was studied as well as their effects on the medical healthcare sector. Review of legacy work and literature review are also included in the study, with a comparison table to summarize the pros and cons of what past research papers had discussed in their respective research. As IoT is becoming the key to the 4.0 industry evolution and brings significant improvements to healthcare, there are health applications and equipment under development and testing, such as telemedicine application, booking application, and medication reminder. Again, these applications come with their own advantages and disadvantages. The following comparison table has highlighted the key positive and negative aspects based on the proposed applications by past research. Lastly, in the discussion and analysis section, concepts of IoT and cloud computing integration were discussed, given how the integration helps to improve patient engagement and satisfaction, for example, ambient assisted living (AAL) and wearable/portable devices. Life-changing healthcare applications like electrocardiogram monitoring, body temperature monitoring, blood pressure monitoring, and asthma monitoring can definitely improve the quality of healthcare services. The summary of the findings in this study will help future researchers to focus on specific key topics to improve the healthcare system as well as their implementations when the technology is ready to make the integration, such as the 5G network that is crucial for future IoT developments. In a long-term perspective, integration and implementation of IoT and cloud computing in the medical and healthcare sector can increase productivity and better quality of life, and health inequities can be solved.

Acknowledgment This work has supported by the Institute of Computer Science and Digital Innovation, UCSI, University of Malaysia.

References

1. T. Ermakova, J. Huenges, K. Erek, R. Zarnekow, Cloud Computing in Healthcare– A Literature Review on Current State of Research, in Americas Conference on Information Systems, Chicago, 2013
2. S.M. Riazul Islam, D. Kwak, M.H. Kabir, M. Hossain, K. Kwak, The Internet of Things for health care: a comprehensive survey. IEEE Access **3**, 678–708 (2015)
3. H. Ahmadi, G. Arji, L. Shahmoradi, R. Safdari, M. Nilashi, M. Alizadeh, The application of internet of things in healthcare: a systematic literature review and classification. Univers. Access Infor. Soc. **18**(4), 837–869 (2018)
4. L.M. Dang, K. Min, D. Han, M.J. Piran, H. Moon, A survey on Internet of Things and cloud computing for healthcare. Electronics **8**(7), 1–49 (2019)
5. S. Nazir, Y. Ali, N. Ullah, I.G. Magariño, Internet of Things for healthcare using effects of mobile computing: a systematic literature review. Mobile Comp. Netw. **1**, 1–20 (2019)
6. K. Geetha Poornima, K.P. Karani, Application of IoT in predictive health analysis-a review of literature. IJMTS **5**(1), 185–214 (2020)
7. R.H. Thomas, M.I. Abdul Rahim, Gamification of medication adherence in epilepsy. Seizure Eur. J. Epilepsy **52**, 11–14 (2017)
8. S.S. Kazi, G. Bajantri, T. Thite, Remote heart rate monitoring system using IoT. IRJET **5**(4), 2956–2963 (2018)

9. C. Li, X. Hu, L. Zhang, The IoT-based heart disease monitoring system for pervasive healthcare service, in International Conference on Knowledge Based and Intelligent Information and Engineering, Marseille, 2017.
10. F. Stradolini, N. Tamburrano, T. Modoux, A. Tuoheti, D. Demarchi, S. Carrara, IoT for Telemedicine Practices enabled by an Android™ Application with Cloud System Integration, in 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, 2018.
11. S. Khedkar, S. Deshpande, M. Choudhari, D. Charles, A smart pill box to re-mind of consumption using IoT. Int. J. Comp. Appl. **128**(1), 38–40 (2018)
12. N.K. Panda, S. Bhardwaj, H. Bharadwaj, R. Singhvi, IOT based Advanced Medicine Dispenser Integrated with an Interactive Web Application, in *International Conference on Advances in Mathematical Sciences*, (Cambridge, UK, 2018), pp. 46–48
13. S.A. Hameed, A. Nirabi, M.H. Habaebi, A. Haddad, Application of mobile cloud computing in emergency health care. Bull. Electr. Eng. Infor. **8**(3), 1088–1095 (2019)
14. L. Hao, X. Xu, M.E. Dupre, A. Guo, X. Zhang, L. Qiu, Y. Zhao, D. Gu, Adequate access to healthcare and added life expectancy among older adults in China. BMC Geriatr **20**, 1–15 (2020)
15. X.J. Wang, M. Camilleri, A smart toilet for personalized health monitoring. Nat. Rev. Gastroenterol. Hepatol., 453–454 (2020)
16. J.Z. Ning Wang, Energy-efficient intelligent ECG monitoring for wearable device. IEEE Trans. Biomed. Circuits Syst. **13**(5), 1112–1121 (2019)
17. H. Ota, M. Chao, Y. Gao, E. Wu, L.-C. Tai, K. Chen, Y. Matsuoka, K. Iwai, H.M. Fahad, W. Gao, H.Y.Y. Nyein, L. Lin, A. Javey, 3D printed "Earable" smart devices for real-time detection of core body temperature. ACS Sens. **2**(7), 990–997 (2017)
18. Q. Xin, W. Jianping, A novel wearable device for continuous, non-invasion blood pressure measurement. Comput. Biol. Chem **69**, 134–137 (2017)
19. E.H.A.-H. Tabarek, G. Al-Jaf, Internet of Things Based Cloud Smart Monitoring for Asthma Patient, in 1st International Conference on Information Technology, p. 380, 2017
20. Z.E. Ahmed, M.K. Hasan, R.A. Saeed, R. Hassan, S. Islam, R.A. Mokhtar, S. Khan, M. Akhtaruzzaman, Optimizing energy consumption for cloud Internet of Things. Front. Phys. **8**, *358* (2020). https://doi.org/10.3389/fphy
21. R. Hassan, Q. Faizan, M.K. Hasan, A.H.M. Aman, A.S. Ahmed, Internet of Things and its applications: a comprehensive survey. Symmetry **12**(10), 1674 (2020)
22. M.K. Hasan, M.M. Ahmed, S.S. Musa, Measurement and Modeling of DTCR Software Parameters Based on Intranet Wide Area Measurement System for Smart Grid Applications, in *International Conference on Innovative Computing and Communications*, (Springer, Singapore, 2020), pp. 1139–1150
23. M.K. Hasan, A.F. Ismail, S. Islam, W. Hashim, M.M. Ahmed, I. Memon, A novel HGBBDSA-CTI approach for subcarrier allocation in heterogeneous network. Telecommun. Syst. **70**(2), 245–262 (2019)
24. S. Islam, A.H. Abdalla, M.K. Hasan, Novel multihoming-based flow mobility scheme for proxy NEMO environment: a numerical approach to analyse handoff performance. ScienceAsia **43**, 27–34 (2017)
25. I. Memon, R.A. Shaikh, M.K. Hasan, R. Hassan, A.U. Haq, K.A. Zainol, Protect mobile travelers information in sensitive region based on fuzzy logic in IoT technology. Secur. Commun. Netw. **2020**, 1–12 (2020)
26. T.M. Ghazal, M.K. Hasan, R. Hassan, S. Islam, S.N.H.S. Abdullah, M.A. Afifi, D. Kalra, Security vulnerabilities, attacks, threats and the proposed countermeasures for the Internet of Things applications. Solid State Technol. **63**(1s), 2513–2525 (2020)
27. M.K. Hasan, M.M. Ahmed, S.S. Musa, S. Islam, S.N. Abdullah, E. Hossain, N.S. Nafi, N. Vo, An improved dynamic thermal current rating model for PMU-based wide area measurement framework for reliability analysis utilizing sensor cloud system. IEEE Access. **9**, 14446–14458 (2021)
28. M.K. Hasan, M. Shafiq, S. Islam, B. Pandey, Y.A. Baker El-Ebiary, N.S. Nafi, R. Ciro Rodriguez, D.E. Vargas, Lightweight cryptographic algorithms for guessing attack protection in complex Internet of Things applications. Complexity **5**, 2021 (2021)

Tornado Forecast Visualization for Effective Rescue Planning

Abhinav Dayal, Sumit Gupta, Sreenu Ponnada, and D. Jude Hemanth

1 Introduction and Related Work

Disaster management is one of the key focuses of any society. Tornadoes have been one of the difficult problems to handle as they are very unpredictable, short and have potential to do great damage to life and property. Authors in [21] highlight and analyze the importance and implementation of warning systems—"Warning systems detect impending disaster, give that information to people at risk, and enable those in danger to make decisions and take action."

The National Weather Services (NWS) of the United States model tornadoes and generate specific alerting trapezoids, but the verbal communication of these areas eventually breaks down into political polygons. NWS issue warnings and broadcast the probable path the tornado is heading and its time frame as well. When a warning is issued, people are advised to take shelter and/or move away. There are general guidelines provided to people via various educational initiatives; however, there needs to be a better systematic way to predict which way to move or what is the best place to take shelter as the strength and direction of tornado may vary. For example, a recent tornado warning [9], and its map is shown in Fig. 1.

A historical analysis of tornado directions [16] elucidates upon the most common geographic directions tornadoes assume in North America. The article suggests splitting the country into tornado warning zones that incorporate the "general runway of storm movement." It highlights a prospective tornado warning grid

A. Dayal (✉) · S. Gupta · S. Ponnada
Vishnu Institute of Technology, Bhimavaram, India
e-mail: abhinav.dayal@vishnu.edu.in; sumit.gupta@vishnu.edu.in; sreenu.p@vishnu.edu.in

D. J. Hemanth
Karunya Institute of Technology and Sciences, Coimbatore, India
e-mail: judehemanth@karunya.edu

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_7

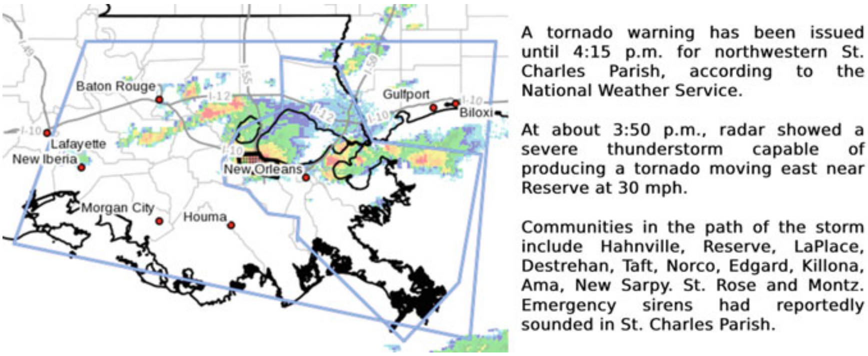


Fig. 1 National Weather Service Tornado Watch—an example

system[22] to pre-generate generic alerting zones as a way of communicating about an approaching storm. It suggests to split the region up into columns angled to match the average tornado travel direction and into rows by how far the average tornado travels in, say, 10 min. A visualization can then be made by shading the rectangular cells by the probability that the tornado may travel through it. FACETs [19] is a proposed next generation weather forecasting system that produces very fine-grained and user-specific actionable alerts, with high-resolution visualization. They divide the regions as tiny heat map styled color-coded rectangular cells that define the probability of a weather event occurring within a spatial and temporal range (Fig. 2). Our work is to support this effort and offer additional ways to determine the probability, and alert for tornadoes.

A popular opinion on public reaction to tornado warnings and the lead time given suggest that if enough lead time is there (1–2 h) people prefer to flee rather than take shelter [10]. Researchers have also studied plausibility of tornado prediction at different time scales, and the corresponding warning system as well as public acceptance [4]. Work toward a probability-based convective-scale forecast-driven warning system has been recommended [23]. This has a potential to improve the evacuation process and give people a better way to respond to the warning for it can suggest possible routes to move so as to evade the tornado, etc. However, caution is required as the visibility hazards due to an imminent tornado can be fatal [2]. But the plausibility to predict a path that successively makes a traveler away from the tornado can be explored, to find a nearby shelter, or in scenarios when that may not be easily possible.

1.1 Tornado Event and Path Prediction

First step to any disaster handling is to detect if there is an imminent disaster. In case of tornadoes, this is called tornado WARNING (or even emergency at some

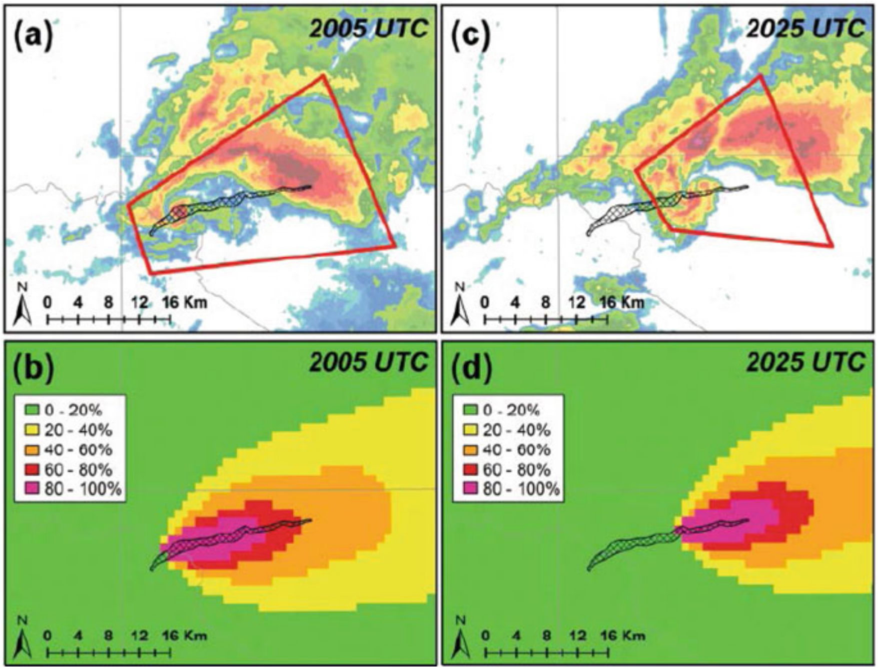


Fig. 2 Deterministic polygon warnings (**a** and **c**) and probabilistic warnings (**b** and **d**) for 20 May 2013 Moore tornado. PHI is updated every 1–2 min (image taken from the original FACETs research)

places). Usually this happens when a tornado is actually spotted and reported by someone. With the rise of social networking, such information can be made available sooner than what could have been possible in the past. In particular Twitter data has become one of the fastest first-hand information providers for event detection. TEDAS [13]: A Twitter-based Event Detection and Analysis System that analyzes tweets for tornado spotting and tornado spatial/temporal pattern discovery. Meteorological Phenomena Identification Near the Ground (mPING) [17] is NSSL's mobile application to crowd-source weather reports. Use of such platforms can give quick and accurate insights into tornado spotting and direction.

There has been significant work in predicting if a tornado may occur at a location. Due to the uncertainty in tornado behavior, it is often difficult to predict with sufficient accuracy. NSSL (National Severe Storms Laboratory) and researchers at several other places have generated various predictive measures to determine if a thunderstorm can lead to a possible tornado. Marzban [14] presents a neural network-based learning algorithm to classify whether the current radar data is conducive for tornado. Trafalis et al. [26] present dynamic selection of time-sensitive radar data to promote active learning with Support Vector Machines (SVM) leading to a more efficient means of tornado prediction.

Our hexagonal grid-based dynamic visualization system can benefit from these and any other predictor determined in future.

1.2 Using Historical Data

Benefits of Probabilistic Hazard Information (PHI) have been well studied [11]. The authors in [3] use space- and time-bound Gaussian smoothers applied to historical data to generate daily predictive maps at any location within 80×80 km grid. They also analyze the annual cycle and repeatability as per the season.

The authors in [18] suggest the limitation of warning issued by NOAA in that they lack the necessary detail and flexibility. Hierarchical Tornado Watch Scale (TWS) [15] presents a way to react to predictive tornado intensity in the forecast area over 24 h period by adjusting the audio alerts issued by NOAA to a clearer action-driven statement. They additionally include shelter options and what to do in various indoor or outdoor scenarios. However, they do not give a predictive route for evacuating the tornado watch polygon to safety.

The authors in [1] discuss a spatial and temporal analysis of tornado fatalities in the United States from 1880 to 2005 in order to improve future warning dissemination and mitigation techniques. Thom [24] presents a way to determine the probability of a tornado striking a point for a group of Iowa counties using historical data from 1937 to 1962. He did that by applying statistical techniques to model historical data of tornado path length and widths. According to Thom, the probability of a tornado striking a point in any year in a 1 degree square centered at a given latitude and longitude is $P = z \times t/A$, where $z =$ mean path area of a tornado in sq. miles (calculated by him as 2.8209) and t is the mean number of tornadoes per year. $t \times z =$ avg area covered by tornadoes per year in 1 degree square. $A =$area of the 1 degree square in sq miles centered at the desired point. This is approximately 60×60 miles region, which is quite large when thinking of rescue planning. Also this does not take into account the regional directions.

In our approach, we find the tornadoes in past in a neighborhood and weigh the directions. The weighting factor can be determined by research outcomes like affect of weather conditions, etc. This gives a historic data-based indicator of a probable direction the tornado may take.

1.3 Forecasting Visualization and Warning System

As soon as a tornado is detected a warning is issued. At this point depending on the strength of the tornado, public is alerted to take proper response, which may be taking a safe shelter or moving away from the area, given sufficient lead time. Several surveys have been conducted to measure the public response to the warnings issued and measure to take to improve that response, e.g., [6, 7, 20]. Warning systems are usually complex because they have to ensure effective communication to people who may be affected, ensure proper response from them to the warning issued, and link to government and private organizations who offer support to successfully carry out the rescue measures. Sorenson [21] and Doswell III et al. [8] have reviewed storm spotting and public awareness measures including public education, tornado spotting by government officials and civilian volunteer organizations. According to Doswell III et al. [8], the users of weather forecasting information must hear the forecasts, must interpret them in their own terms in order to make decisions, and must know what to do in order to achieve some desired result, if the forecasts are to be successful in having a positive societal impact.

Therefore, offering people a detailed and easy visualization of the warning in real time and updating them on the current status and possible means of rescue cannot be neglected in light of improving the public response to the warnings.

2 Hexagonal Warning Grid System

We envision an alternative method of pre-generating generic alerting zones as a means of communicating about an approaching storm. FACETs [19] use a rectangular grid system. Given the dynamic and unpredictable nature of tornado movement, the shapes of regions with various probability bands keep changing. The elongated shapes and concavities give an idea of the direction of current or probable movements (Fig. 2). We present an alternative method of pre-generating generic alerting zones as a means of communicating about an approaching storm. We use a hexagonal grid of appropriately sized geographic hexagonal boundaries connected to one another. Each hexagonal cell is assigned a probability of a tornado entering into it. This can be calculated by taking into account several factors that predict tornado movement like NWS forecasts, crowd-sourced data, tornado spotters, historical bias, and so on. Figure 3 presents a sample grid where a tornado is spotted in the middle cell and the probability of its moving in adjoining cells is calculated.

We believe that, such a visualization, when extended to further adjacent hexagonal areas, will lead to an effective visualization of where and how resources should

Fig. 3 Proposed hexagonal grid warning system

be rescued in a timely manner. Using the average forward speed of a tornado, which is 30 mph (calculated from over 60 years historical U.S. tornado data), we divide the local region up into hexagonal cells, each with a diameter that average tornado can cross in approximately 10 min, which gives us a 10 km radius. This gives a balance between general tornado speeds and practically sized alerting regions. Using this arrangement, we create a warning grid that, based on a hypothetical tornado occurring in each cell, would populate neighboring cells with the probability that the tornado will travel into it.

Tornado speed, location, and direction can be collected from various sources—weather reports, public response (Tweet analysis also possible), storm chasers, government officials, etc. Once this is determined, the rest of the process follows. For example, a tornado is sighted within the center cell. Probabilistic estimates are made for neighboring cells and appropriate alerts are issued. Notifications go out to people within neighboring cells telling them something like, "A tornado has been spotted to the southwest. There is an 18% chance it will travel nearby your location." Or something similar. (More study is required here.)

The probability can be a combination of various factors, some of which are listed below:

1. **Current Status**: This is the current reported location and direction of movement of the tornado.
2. **Historical Data**: This may indicate a general direction of tornado travel in and around that neighborhood. Historical analysis depicts a general trend of tornadoes to move north-east. However, it would have some regional variability. A mathematical model can be devised to estimate local direction of travel based on historical data. It may be a hierarchical model, or a near neighbor model.
3. **Proximity**: A tornado may change course drastically, so the proximity of a tornado location to a cell can be a factor too.
4. **Atmospheric changes**: Pressure tends to decrease, temperature tends to decrease, and moisture tends to increase in the vicinity of a tornado. For example, if pressure is decreasing in a certain cell over time. Is there difference in pressure in neighboring cells and so on? Some model may be made based on inter-cell differences and possibility of tornado moving in that cell can be evaluated. Researchers have used S-bad phased-array radar (PAR) data to find mesocyclone intensification and abrupt short-term changes in tornado motion [12]. The authors in [5] fuse data from various sources like temporal trends in convection, storm hydrometeors size, extent and motion, etc., to a statistical model for predicting early onset of a thunderstorm or tornado, in order to increase lead time for safety. Multi-Radar Multi-Sensor (MRMS) systems also give very high-spatio-temporal-resolution radar data to be integrated with various other sources like satellite imagery, rain/lightening gauge data, etc. [27]. The authors in [25] have successfully correlated historical tornado incidents to convective mode mesocyclone strength from full volumetric radar data.
5. **Public Reaction**: People often tweet about such things [13]. We can notice tweets about tornado and form a map of tweets in time and location, form clusters

of nearby points and fit a line along those points to get an average direction of motion based on public reaction. How much credence to be given to such data requires further study.

The simplest way to combine several factors is by linearly combining them; however, more complex means of combination cannot be ruled out and are subject to research. The coefficients of linear combination can be determined experimentally. A simple data mining strategy using currently available historical data can suggest more suitable coefficients to use. The alerting system will not only give a warning of an imminent tornado, it would also provide detailed visualization for aid in:

1. **Rescue planning**: If people need to move out, what directions are safer relatively.
2. **Traffic control**: If we need people out of an area, the traffic on roads must be cleared so as to avoid congestions.
3. **Shelter locations**: If people need to take shelter, help guide them to nearest shelter location.
4. **Estimation**: Estimate damage and recovery, etc.

3 Proposed Method

In a hexagonal grid, each region has six neighbors. A Graph $G = V, E$ can easily represent such a grid, where V is the set of regions and E is a set of edges, where an edge represents "is neighbor of" relationship between two regions. Since there are maximum six neighbors to any region, an adjacency list may be a suitable data structure to represent such a graph. The adjacency matrix will take significantly more storage space.

We propose another representation using a 2-D array to construct this hexagonal grid that not only makes it easier to construct a grid, but also locates neighbors of any region directly. Figure 4 depicts the construction clearly. Notice that for a grid cell i, j, when j is even, the neighbors are $\{(i-1, j), (i+1, j), (i-1, j-1), (i-1, j+1), (i, j-1), (i, j+1)\}$, whereas when j is odd, the neighbors are $\{(i-1, j), (i+1, j), (i+1, j-1), (i+1, j+1), (i, j-1), (i, j+1)\}$.

Since in this chapter, our main goal is to present a visualization system that can adapt to several inputs easily, we are not experimenting with all known means through which tornado direction and location are spotted. Instead we take an example of historical data that is easily available to us and use that only to find the probability of tornado entering a neighboring region.

For each hexagonal region, we determine a probable angle the tornado may travel through the region based on the historical information. Using the center location of the region (x, y), we fetch from history data all recorded tornadoes in the past within 100 km radius of the region center. We then calculate the probable angle of travel, $A_{i,jhistory}$, as a weighted mean that is proportional to the strength and inversely proportional to the age and distance as shown below:

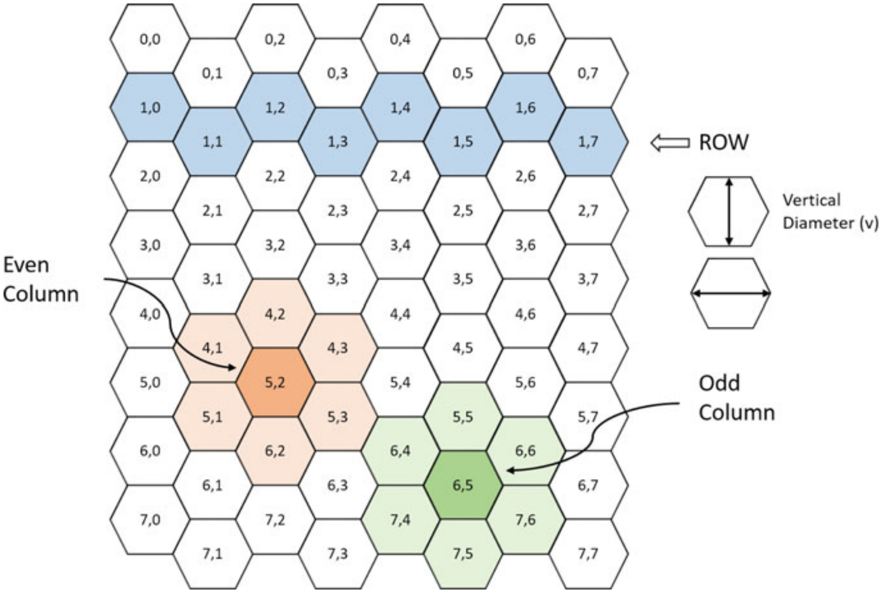


Fig. 4 Construction of hexagonal grid

$$A_{i,jhistory} = \frac{\sum_{k=1}^{n} \frac{F_k+1}{a_k \times d_k}\theta_k}{\sum_{k=1}^{n} \frac{F_k+1}{a_k \times d_k}}, \tag{1}$$

where n is the number of tornadoes within 100 km radius of the region center, F_k is the F-scale (from 0 to 5) of the kth tornado, a_k is its age in days, d_k is the distance in meters from region center, and θ_k is the direction angle of that tornado.

For each measure, we can similarly calculate the probable angle based on that measure and then combine the results to give a single predicted angle of travel in that region. For example, we can calculate $A_{i,jtwitter}$, $A_{i,jradar}$, $A_{i,jweather}$, etc. Based on crowd sourcing, current location/direction, weather conditions, etc., we may then calculate the final angle $A_{i,j}$ as

$$A_{i,j} = \alpha A_{i,jhistory} + \beta A_{i,jtwitter} + \gamma A_{i,jradar} + \delta A_{i,jweather} \ldots, \tag{2}$$

where α, β, etc., are the experimentally determined constants.

Once we have evaluated a number, which tells us the possible prominent angle or direction of motion of the tornado in a region, we can determine the probabilities of tornado entering a neighboring cell from a given cell by a linear combination of dot products of each determining vector to the vector depicting the direction of the neighboring cell. So given a tornado is spotted in a region, we need to run a breadth-first traversal of the grid centered at this region and propagate the probabilities of the tornado travel in the outward region, using the flow information already calculated

Table 1 Hexagon region neighbor and angles for a cell (i, j)

Neighbor	Angle	j odd	j even
Top	0	$i-1, j$	$i-1, j$
Top right	60	$i, j+1$	$i-1, j+1$
Bottom right	120	$i+1, j+1$	$i, j+1$
Bottom	180	$i+1, j$	$i+1, j$
Bottom left	240	$i+1, j-1$	$i, j-1$
Top left	300	$i, j-1$	$i-1, j-1$

Algorithm 1 Probability propagation using breadth-first traversal

```
Initialize a Queue, Q
let c be the center hex region
Q.Enqueue(c)
c.TornadoProbability = 1.0; // tornado spotted here
Set c.visited = true
while Q is not empty do
    let v = Q.Dequeue()
    for EACH Neighbor w of v [Figure 4] do
        if w.visited = FALSE then
            SET w.visited = TRUE
            Q.Enqueue(w)
        end if
        calculate p using Equation 3
        w.TornadoProbability =
                MAX(v.TornadoProbability × p,
                      w.TornadoProbability)
    end for
end while
```

in each region in terms of the prominent angle of travel of the tornado in every region.

Table 1 displays the angle of each neighbor. Given neighbor angle $\theta_{neighbor}$, probable tornado angle in region $A_{i,j}$, the probability that the tornado will travel to the given neighbor can be easily determined by how close the angle $\theta_{neighbor}$ is to the angle of travel $A_{i,j}$, i.e.,

$$p = \frac{360.0 - \left|A_{i,j} - \theta_{neighbor}\right|}{360.0}. \quad (3)$$

This analysis can be continued for distant neighbors and so on within a predetermined radius, and thus we can estimate a probable plume or region where the tornado may travel in near future. Based on this analysis, rescue measure can be made. People very close to the tornado may seek shelter. People having opportunity to move away can be suggested appropriate routes and direction to move. And people not going to be affected by in the vicinity can be advised to make way for people who will be moving. Algorithm 1 depicts the process clearly of propagating the probabilities of tornado travel to neighboring cells. As the tornado moves, the visualization can be dynamically updated every few seconds to get real-time updates.

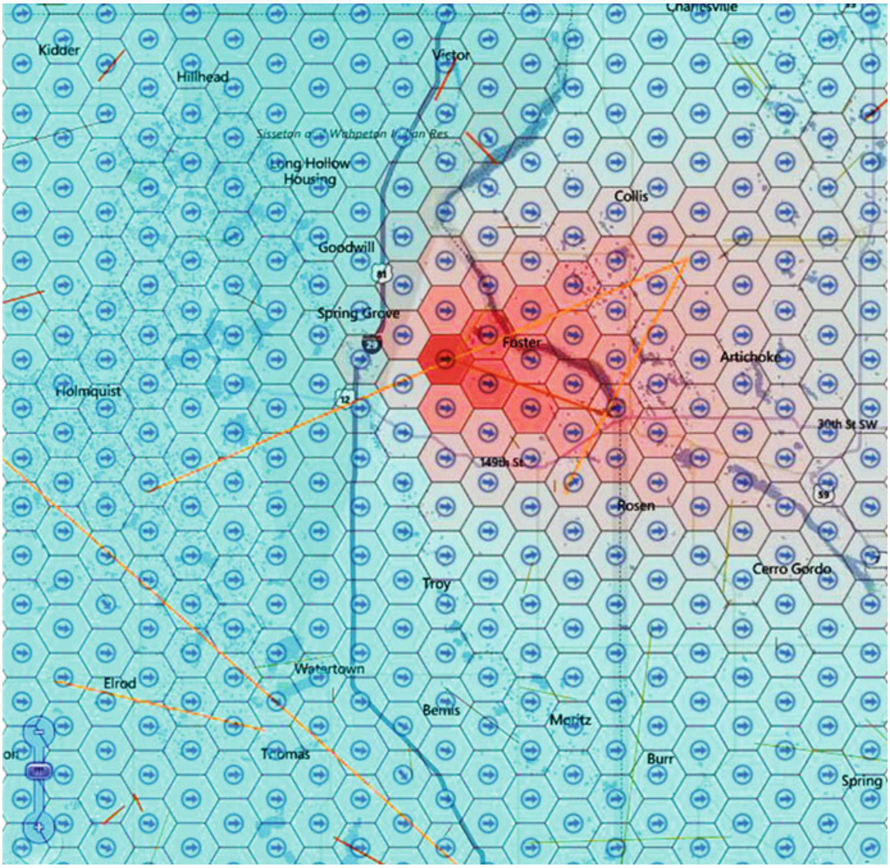


Fig. 5 Hexagonal grid visualization using weighted history data. The orange lines indicate historical tornado data in the U.S. in past 60 years. They are color coded from red for F-scale of 5 to green for F-scale of 0. The blue hexagonal cells indicate no current tornado event. The arrow in each cell indicates the predicted direction of tornado travel. For blue cell, it is the average from history. The bright red cell is where the tornado is currently spotted and the arrow suggests its current direction of travel. The various reddish tinted cells indicate the probability (by intensity) that the tornado may come to that cell. The data is taken from one of the tornado in U.S. Tornado History Data

4 Results

Figure 5 shows the results. Blue means no or little chance for tornado and red means a higher chance. Notice how, for each cell, the visualization indicates not only the chance of tornado but also from where it may come and to where it may go. The dark red region at center is where the tornado is initially spotted. The lines in the picture depict historical data color scaled according to F-scale. The color scale changes from red for F-scale of 5 to green for F-scale of 0. Rest are interpolated

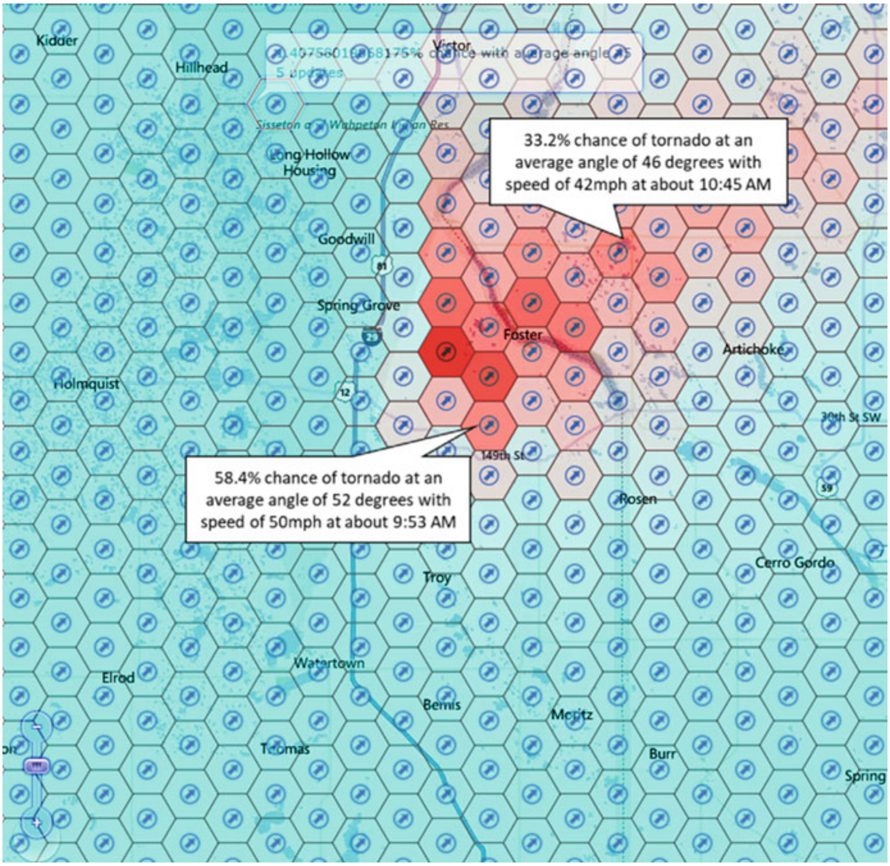


Fig. 6 Hexagonal grid visualization using 45° global direction. The colors etc. have same meaning as in the previous figure

values in between. In the example figures, the orange color of lines indicates some value in between. Here the weighted average directions calculated from 100 km neighborhood history for each point at the center of the hex cell are also shown. Compare this to results shown in Fig. 6 that show visualization based on global north-east direction of 45° at center of each hexagonal cell. The local dynamic adaptability is the strength of this visualization. As more parameters are added to determine the average angle of movement in a region, the visualization can easily adapt and help aid rescue planning.

5 Conclusion and Future Work

In conclusion, this chapter presents a novel way to aggregate tornado track data within a hexagonal cell to get a weighted average angle of travel. Based on each cell's average travel angle (and either a local average tornado track length or maybe a global length variable), one can assign a percent probability of incursion into neighboring cells. However, the source tornado track data is not dense enough to give us meaningful aggregation at this 10 km scale. So we look into a 100 km neighborhood. If there is no data in this larger region, the history predictor can either give equal probability in each direction, or point to a global average. We chose the former and argue that this combined with other predictors like social data analysis, atmospheric factors, etc., can give us much accurate data and updates in real time. Note that, as current research shows, the predictors we determine here can also help correlate with other factual scientific data.

As future work, we would like to incorporate some more predictors like TEDAS given by Li et al. [13] etc. mentioned in this chapter. We would also like to implement manual perturbation in real time where an officer who is witnessing a tornado can directly adjust certain factors to bias the visualization giving much accurate decision support. We believe that such a system can be very helpful for disaster management, not only for a tornado, but also in other disaster scenarios like other weather-related events, or even criminal investigations, etc.

Acknowledgments Acknowledgments are due to Robert Staskowski who initiated this work by his blog [22] after the disaster in Moore, Oklahoma in 2013. John Nelson (now working at ESRI) and Josh Stevens (now GIScientist at NASA's Earth Observatory) helped brainstorm the idea and encouraged to continue working and develop the idea more. Acknowledgments are due also to IDV Solutions to allow use of their Visual Fusion Visualization engine to help visualize and implement the ideas presented in this article.

References

1. W.S. Ashley, Spatial and temporal analysis of tornado fatalities in the United States: 1880–2005. Weather Forecast. **22**, 1214–1228 (2007)
2. W.S. Ashley, S. Strader, D.C. Dziubla, A. Haberlie, Driving blind: weather-related vision hazards and fatal motor vehicle crashes. Bull. Am. Meteorol. Soc. **96**(5), 755–778 (2015). https://doi.org/10.1175/BAMS-D-14-00026.1
3. H.E. Brooks, C.A. Doswell, M.P. Kay, Climatological estimates of local daily tornado probability for the United States. Weather Forecast. **18**(4), 626–640 (2003). https://doi.org/10.1175/1520-0434(2003)018<0626:CEOLDT>2.0.CO;2
4. J. Brotzge, W. Donner, The tornado warning process: a review of current research, challenges, and opportunities. Bull. Am. Meteorol. Soc. **94**(11), 1715–1733 (2013). https://doi.org/10.1175/BAMS-D-12-00147.1
5. J.L. Cintineo, M.J. Pavolonis, J.M. Sieglaff, D.T. Lindsey, An empirical model for assessing the severe weather potential of developing convection. Weather Forecast. **29**(3), 639–653 (2014). https://doi.org/10.1175/WAF-D-13-00113.1

6. N.W. Croom III, Providing public notification of tornadic events in the town of Castle Rock (2009). https://www.hsdl.org/?abstract&did=8501. Accessed January 2020
7. W.R. Donner, H. Rodriguez, W. Diaz, Public warning response following tornadoes in New Orleans, LA, and Springfield, MO: A sociological analysis, in *87th AMS Annual Meeting* (2007)
8. C.A. Doswell III, A.R. Moller, H.E. Brooks, Storm spotting and public awareness since the first tornado forecasts of 1948. Weather Forecast. **14**, 544–557 (1998)
9. B. Estes, Tornado warning for northwestern St. Charles Parish (2015). http://www.nola.com/weather/index.ssf/2015/04/tornado_warning_for_northweste.html. Accessed November 2015
10. S. Hoekstra, K. Klockow, R. Riley, J. Brotzge, H. Brooks, S. Erickson, A preliminary look at the social perspective of warn-on-forecast: preferred tornado warning lead time and the general public's perceptions of weather risks. Weather Climate Soc. **3**(2), 128–140 (2011). https://doi.org/10.1175/2011WCAS1076.1
11. C.D. Karstens, G. Stumpf, C. Ling, L. Hua, D. Kingfield, T.M. Smith, J. Correia, K. Calhoun, K. Ortega, C. Melick, L.P. Rothfusz, Evaluation of a probabilistic forecasting methodology for severe convective weather in the 2014 hazardous weather testbed. Weather Forecast. **30**(6), 1551–1570 (2015). https://doi.org/10.1175/WAF-D-14-00163.1
12. C.M. Kuster, P.L. Heinselman, M. Austin, 31 May 2013 El Reno Tornadoes: Advantages of rapid-scan phased-array radar data from a warning forecaster's perspective*. Weather Forecast. **30**(4), 933–956 (2015). https://doi.org/10.1175/WAF-D-14-00142.1
13. R. Li, K.H. Lei, R. Khadiwala, K.C.C. Chang, Tedas: A Twitter-based event detection and analysis system, in *ICDE*, ed. by A. Kementsietsidis, M.A.V. Salles (IEEE Computer Society, 2012), pp. 1273–1276. http://dblp.uni-trier.de/db/conf/icde/icde2012.html#LiLKC12
14. C. Marzban, A neural network for tornado prediction based on Doppler radar-derived attributes. J. Appl. Meteorol. **35**, 617–626 (1995)
15. J.B. Mason, J.C. Senkbeil, A tornado watch scale to improve public response*. Weather Climate Soc. **7**(2), 146–158 (2015). https://doi.org/10.1175/WCAS-D-14-00035.1
16. J. Nelson, When tornados strike, which way do they travel? (2013). https://io9.gizmodo.com/when-tornados-strike-which-way-do-they-travel-513093207. Accessed November 2020
17. NOAA.gov: Meteorological phenomena identification near the ground - crowdsourcing weather reports (2015). http://www.nssl.noaa.gov/projects/ping/. Accessed November 2015
18. L.P. Rothfusz, C. Karstens, D. Hilderband, Next-generation severe weather forecasting and communication. EOS Trans. Am. Geophys. Union **95**(36), 325–326 (2014). https://doi.org/10.1002/2014EO360001, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014EO360001
19. L.P. Rothfusz, R. Schneider, D. Novak, K. Klockow-McClain, A.E. Gerard, C. Karstens, G.J. Stumpf, T.M. Smith, FACETs: a proposed next-generation paradigm for high-impact weather forecasting. Bull. Am. Meteorol. Soc. **99**(10), 2025–2043 (2018). https://doi.org/10.1175/BAMS-D-16-0100.1
20. S. Seddig, Improving citizen response to tornado warnings (2009). https://www.hsdl.org/?view&did=689972. Accessed January 2020
21. J.H. Sorenson, Hazard warning systems: Review of 20 years of progress. Nat. Hazards Rev. **1**(2), 119–125 (2000). https://doi.org/10.1061/(ASCE)1527-6988(2000)1:2(119)
22. R. Staskowski, Tornado warning Grid System (2013). http://www.poolonthenet.com/tornado_grid.htm. Accessed November 2013
23. D.J. Stensrud, M. Xue, L.J. Wicker, K.E. Kelleher, M.P. Foster, J.T. Schaefer, R.S. Schneider, S.G. Benjamin, S.S. Weygandt, J.T. Ferree, J.P. Tuell, Convective-scale warn-on-forecast system: a vision for 2020. Bull. Am. Meteorol. Soc. **90**(10), 1487–1500 (2009). https://doi.org/10.1175/2009BAMS2795.1
24. H.C.S. Thom, Tornado probabilities. Mon. Weather Rev. **91**, 730–736 (1963)

25. R.L. Thompson, B.T. Smith, J.S. Grams, A.R. Dean, C. Broyles, Convective modes for significant severe thunderstorms in the contiguous United States. Part II: Supercell and QLCS tornado environments. Weather Forecast. **27**(5), 1136–1154 (2012). https://doi.org/10.1175/WAF-D-11-00116.1
26. T.B. Trafalis, I. Adrianto, M.B. Richman, Active learning with support vector machines for tornado prediction, in *Proceedings of the 7th International Conference on Computational Science, Part I: ICCS 2007*. ICCS '07 (Springer, Berlin, 2007), pp. 1130–1137
27. J. Zhang, K. Howard, C. Langston, B. Kaney, Y. Qi, L. Tang, H. Grams, Y. Wang, S. Cocks, S. Martinaitis, A. Arthur, K. Cooper, J. Brogden, D. Kitzmiller, Multi-radar multi-sensor (MRMS) quantitative precipitation estimation: initial operating capabilities. Bull. Am. Meteorol. Soc. **97**(4), 621–638 (05 2016). https://doi.org/10.1175/BAMS-D-14-00174.1

Situational Awareness for Law Enforcement and Public Safety Agencies Operating in Smart Cities – Part 1: Technologies

Daniel Minoli, Andrzej Koltun, and Benedict Occhiogrosso

1 Introduction

Situational awareness[1] (SA) provides the capability to efficiently comprehend one's physical environment. As discussed elsewhere in this book, SA is thought of as knowing what is going on around us; a useful definition from [1] is "The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future." More specifically, SA involves the timely acquisition of knowledge about real-world events, distillation of those events into higher-level conceptual constructs, and their synthesis into a coherent context-sensitive view [2]. In practical terms SA can be defined as being in a state of awareness. For public safety and law enforcement teams, this translates into real-time visibility on the status of prospective threats – and possible targets. SA deals with capturing incoming data, analyzing the data, and generating actionable intelligence. Recent technology trends in video sensing, edge computing, big data storage, and machine learning (ML) can be aligned to create a shared real-time information system for SA in a number of environments including physical security, law enforcement, and vehicular systems. Ultimately, many smart city applications including livability, infrastructure management, traffic management, utility management, policing, and physical security collectively support key aspects of city operations, incorporate

[1] Some researchers have used the phrase "situation awareness"; in this chapter we consider both terms describing the same discipline.

D. Minoli (✉) · A. Koltun · B. Occhiogrosso
DVI Communications, New York, NY, USA
e-mail: daniel.minoli@dvicomm.com; andrzej@dvicomm.com; ben@dvicomm.com

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_8

situational awareness, which in turn entail big data analytics, image processing, and cloud computing, all of which can benefit from ML and deep learning (DL) mechanisms.

2 Overview of Smart City/IoT Applications

Preceding chapters in this book have described the challenges faced by large cities worldwide, which of late have included the management of the COVID-19 pandemic in general, and delivery of healthcare services in particular, and described the solution opportunity afforded by IoT platforms and ecosystems. The extensive use of IoT-based solutions has given rise to the smart city concept (e.g., [3–13], which is but a mere fraction of the numerous papers published on this topic). A smart city is also comprised of smart campuses, smart buildings, and smart homes. Traditional smart city applications have included the following:

- Physical security: security in streets, parks, stations, tunnels, bridges, trains, buses, ferries, etc. Networked sensors (fixed, portable, and mobile [e.g., drones]) support multiple, related applications including surveillance video, license plate reading, gun-shot detection, biohazard and radiological contamination monitoring, face recognition, and crowd monitoring and control.
- Traffic, transportation, and mobility: optimizing traffic flow and reducing congestion and latency as well as enhancing safety high expediency, lowering noise, and reducing fuel consumption and CO_2 emissions. Networked sensors to support traffic flow, driverless vehicles including driverless bus transit, and multimodal transportation systems. For driverless vehicles, sensors such as high-resolution mapping, coupled with real-time telemetry data, facilitate traffic and hazard avoidance.
- Power and other utilities: reliable electric energy, gas, and water; optimized waste-management and sewer; safe storage of fuel.
- Electric and other utility manhole monitoring: electric power manholes require monitoring to avoid and/or prevent dangerous situations.
- Logistics: supplying city inhabitants with fresh food, supplies, goods, and other materials in a timely fashion.
- Livability: quality of life, expeditious access to services, efficient transportation, low delays, and safety.
- Infrastructure and real-estate management: monitor status and occupancy of spaces, buildings, roads, bridges, tunnels, railroad crossings, and street signals. Networked sensors (possibly including drones) provide real-time and historical trending data allowing city agencies to provide enhanced visibility into the performance of resources, facilitating environmental and safety sensing, smart parking and smart parking meters, smart electric meters, and smart building functionality.

- Environmental monitoring: monitor outdoor temperature, humidity, air quality, and the presence of contaminants or other environmental gases.
- Pollution monitoring: real-time and historical monitoring of a vast array of pollutants generated by factories, incinerators, and urban crematoria.
- Flood and storm drainage control.
- Lighting: using LED technology as well as IoT control adapting to traffic conditions, weather events, lunar cycles and seasonal variations, etc.

Almost all of these applications rely to a greater or lesser extent on SA capabilities. In addition, ML has become an important component to address the voluminous data collected from multitude of sensors deployed in the environment and to draw meaningful on-point conclusions and actionable intelligence from the data.

3 General Concept of Situational Awareness (Examples)

3.1 General Concept

As implied above, SA is defined in a number of ways, including the following "up-to-the-minute cognizance or awareness required to move about, operate equipment, or maintain a system" [14]. More formally, SA is described as knowing what is "going on" around a reference entity and, within that knowledge of the entity's surroundings, knowing what is important [15] – the process of perceiving the elements in the environment, understanding the elements in the environment, and the projection of their status into the near future [16]. SA originates from human factor and cognitive studies [17]. It has been extensively studied and applied to several disciplines, including psychology, aviation, and military operations, starting with the important work of M. R. Endsley, e.g., [16] and numerous other publications. Figure 1 is adapted from Endsley's situational awareness reference model; it defines three levels of situational awareness: perception, comprehension, and projection. McGuinness and Foy [18] provided extension of Endsley's situational awareness model that includes resolution of the perceived state as the level four situational awareness component.

Many classical definitions of the term still provide insight [19]; these include:

- "Perception of the elements in the environment within a volume of time and space; comprehension of their meaning; projection of their status in the near future," Endsley (1987, 1990, 1995) [1, 20, 21];
- "Estimate of the purpose of activities in the observed situation; understanding of the roles of participants in these activities; Inference about completed or ongoing activities that cannot be directly observed; Inference about future activities," Noble (1989) [22];

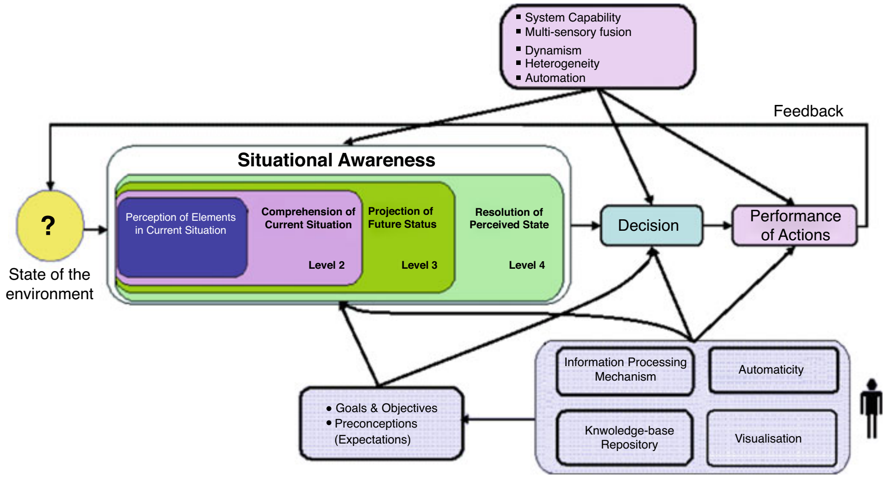


Fig. 1 Endsley's situational awareness reference model (simplified)

- "Just a label for a variety of cognitive processing activities that are critical to dynamic, event-driven, and multitask fields of practice," Sarter and Woods (1991, 1995) [23, 24];
- "Continuous extraction of environmental information, integration of this knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events," Dominguez (1994) [25];
- "Spatial awareness; Mission/goal awareness; System awareness; Resource awareness; Crew awareness," Pew (1995) [26];
- "Perceive the information; Interpret the meaning with respect to task goals; Anticipate consequences to respond appropriately," Flach (1995) [27];
- "The ability to envision the current and near-term disposition of both friendly and enemy forces," Stiffler (1988) [28].

SA spans many domains and applications. A short list of illustrative examples include military applications, behavioral science, emergency medical callouts, vehicle driving, search and rescue, law enforcement, and cybersecurity threat operations/security incident response. ML and fuzzy logic are two of the many underlying technologies used in SA.

3.2 *Video/Security Framework*

Another definition of SA is the ability to develop and deploy traditional or AI-based mechanisms to assess, recognize, anticipate, and intercept events specific to the use case of interest. Some underlying trends of relevance that impact SA include the following, among others [2]:

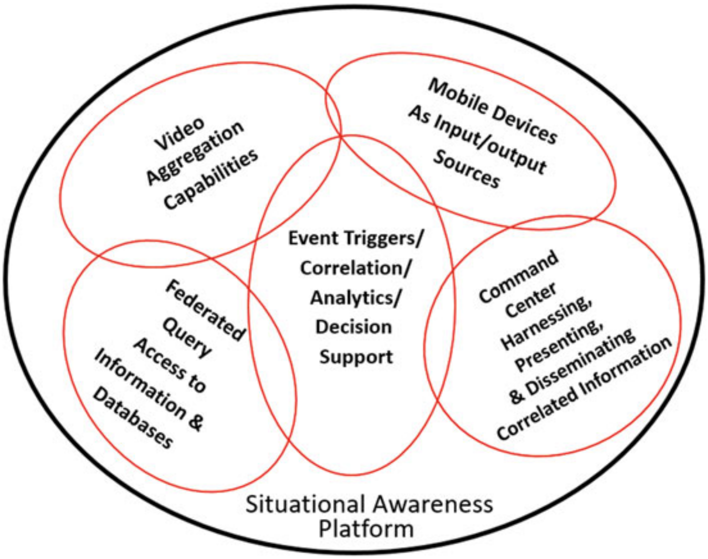


Fig. 2 A situational awareness ecosystem

(i) The growth of always-on video surveillance in public spaces. Some have predicted that, "It will soon be possible to find a camera on every human body, in every room, on every street, and in every vehicle" [29].
(ii) Real-time video analytics using edge computing [13, 30]: to facilitate scalability, such analytics are performed close to the event's occurrence to reduce bandwidth required to transmit video to the cloud from many cameras via the ingress networks; to mitigate the bandwidth demand, video analytics are performed via edge gateways connected using wired or wireless LANs to associated cameras.

SA helps decision-makers spanning a large set of environments to have the distilled information and focused understanding to make effective decisions in the course of their work; clearly this is desirable in the law enforcement context. Figures 2, 3, 4, and 5 depict the broad ecosystem for the family of use cases that are part of the discussion in this chapter.

When properly designed, configured, and deployed, these systems are harnessed law enforcement and public safety environment to provide incident commanders as well as analysts the ability to assess threats with the stated goal of anticipating an event in advance of their actual occurrence (i.e., providing timely and accurate intelligence). SA provides ongoing value during and after events' occurrence, but admittedly depending on the situation this may offer less value than advance warning.

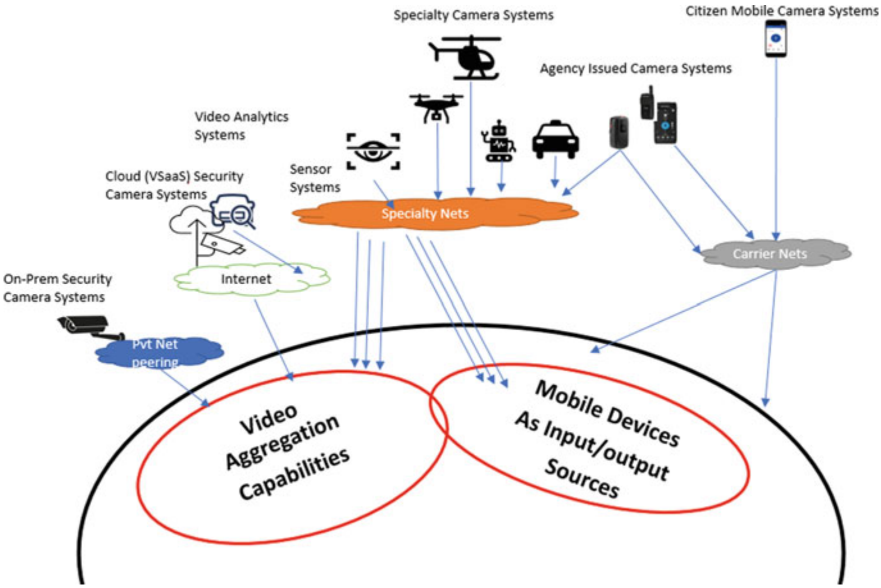


Fig. 3 Plethora of input devices

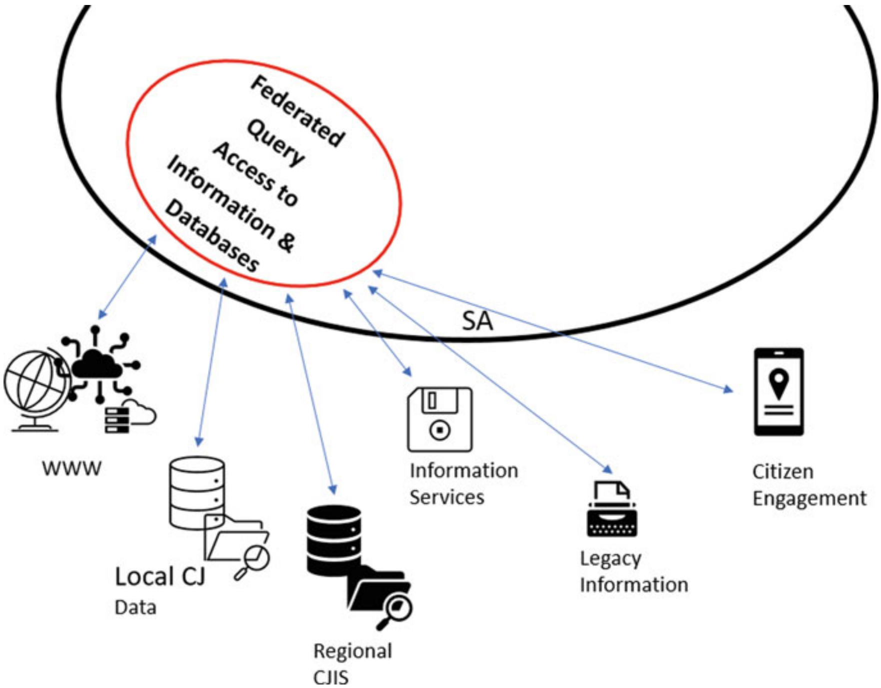


Fig. 4 Federated queries

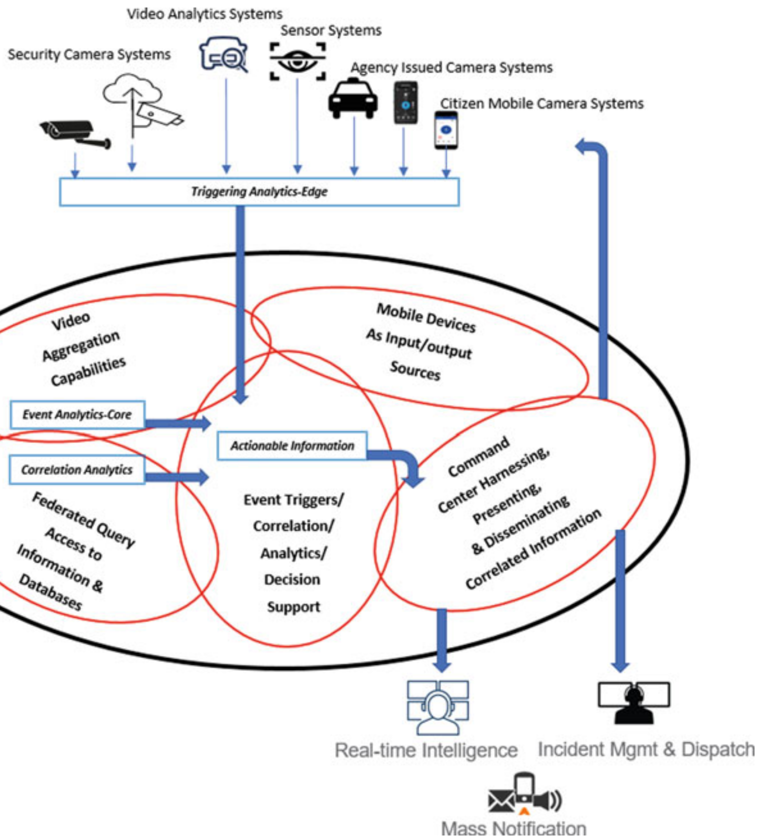


Fig. 5 Harnessing information overload for actionable decisioning

4 Use of Machine Learning in Some General SA Use Cases

As covered in other chapters in this book, ML predicts and classifies data using various algorithms optimized to the dataset in question. ML is an example of artificial intelligence (AI), which itself is a subfield of computer science that focuses on the development of computer-based systems, applications, and algorithms that mimic cognitive processes intrinsic to human intelligence. ML is a mechanism used to implement AI concepts. See Fig. 6. It entails (complex) algorithms that parse data, learn from that data, and then apply what they have learned to make informed decisions. ML techniques are increasingly utilized to analyze, cluster, associate, classify, and apply regression methods to SA data (as well as to many other IoT environments). ML techniques endeavor to examine and establish the internal relationships of a set of data collected from the plethora of input devices collecting visual, audio, and signal data from the field. Among other applications,

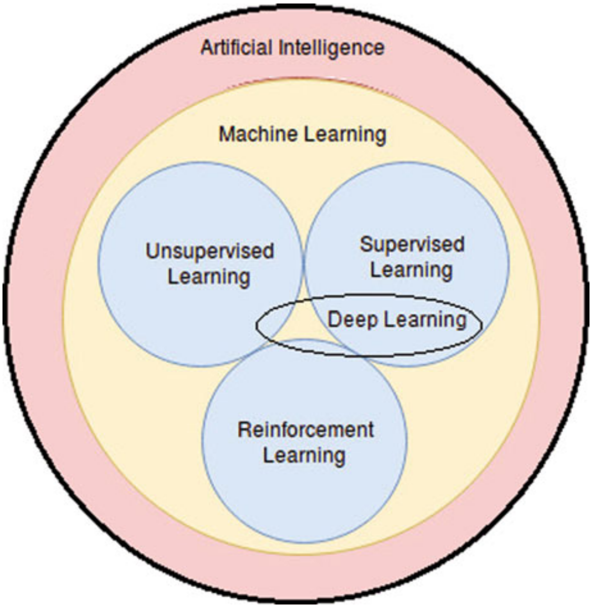


Fig. 6 Positioning of various AI systems usable in situational awareness

ML techniques are applicable to image processing and analysis, computer vision, speech recognition, and natural language understanding.[2]

A relatively large number of ML methods and algorithms are currently in common use:[3]

- Unsupervised learning methods (e.g., [31, 32]): analysis is undertaken (by the "machine") on unlabeled data in order to find latent patterns in the data, after which the patterns are used to cluster populations into distinct groups.
- Semi-supervised learning methods (e.g., [33, 34]): analysis is performed (by the "machine") using a small set of tagged data and larger (large) set of unlabeled data to refine the cluster classification. This offers higher accuracy with less human effort than traditional ML approaches to classification.
- Supervised learning methods (e.g., [35, 36]): the system in question (the "machine") undertakes the process of learning utilizing the best set of labeled examples and one of a number of data-driven learning algorithm; the goal is to

[2] Other advanced AI fields (not further discussed in this chapter, but with theoretical applicability to situational awareness) include *cognitive computing* (systems that endeavor to understand and emulate human behavior, while also providing more natural and intuitive interface to the machine) and *natural language processing* (NLP) systems that allow machines to understand written language or voice commands (also including natural language generation that enables the machine to communicate in "spoken conversation").

[3] Google Scholar identifies over 3,100,000 papers on the topic of "machine learning."

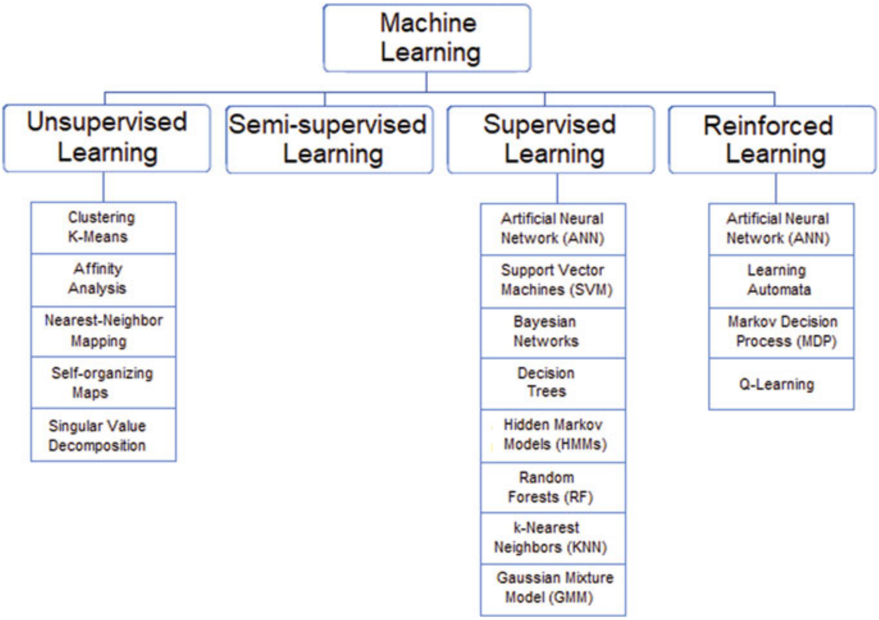


Fig. 7 Machine learning methods and algorithms in common use

produce an inference model where discrete data are categorized in an appropriate manner, and continuous data are related in the form of a regression function.

- Reinforcement learning methods (also known as boost learning) (e.g., [37, 38]): the system in question (the "machine") assesses the environment and selects and executes certain activities (decisions) within its action domain, and depending on how it is "rewarded" or "punished," it then learns which strategy is best for future, comparable decisions.

Figure 7 depicts ML methods and algorithms in common use and which are applicable to the general field of SA.

Some of the algorithms in use are briefly summarized in the following [39]:

- Neural network algorithm (e.g., [40, 41]): a widely deployed methodology that enjoys good performance for estimation and approximation. Neural networks have origin in biology where the network is made up of biological neurons; an artificial neural network is composed of nodes that represent artificial neurons that can store information. Such an arrangement works very well for solving AI problems such as predictive modeling or adaptive control. The connections of the nodes are modeled as weights; inputs are modified by a weight and summed, and an activation function controls the amplitude of the output; training occurs via a dataset.
- Support vector machines (SVM) algorithm (e.g., [42, 43, 44]): a linear model for classification and regression assessment. The purpose of these algorithms

is to identify and differentiate complex patterns in the data: it uses learning algorithms to classify and analyze data regression. The fundamental concept of SVM is straightforward: the algorithm creates a line (also called a hyperplane) that separates the data into classes, seeking a separation between the two classes that is as wide as possible. It takes the data as an input and then outputs a line that separates the data classes to the degree possible.

- Bayesian network (BN) algorithm (e.g., [45, 46, 47]) (also known as a belief network or unidirectional graph model): a probabilistic representation of a set of random variables and their conditional independence. A BN is a probabilistic model that uses Bayesian inference for computation of probabilities. Such a network aims at modeling conditional dependence, which highlights causation, representing conditional dependence by edges in a directed graph.
- Decision tree algorithm (e.g., [48, 49]): a predictive mechanism used for solving regression and classification problems that utilizes observations that represent graphs. Predictions are correlated with the traits studied by leaves and branches: it utilizes a decision tree (as a predictive model) to progress from observations about a phenomenon (represented in the branches) to conclusions about the item's target value (represented in the leaves). The goal is to create a training model that can be used to predict the class or value of the target variable by learning simple decision rules inferred from prior data (which is the training data). The decision tree can be designed with different ML algorithms discussed in the cited literature.
- Hidden Markov model (HMM) algorithm (e.g., [50, 51, 52]): a statistical approach where the system is modeled as a Markov chain which is assumed to have unseen (specifically, hidden) states or modes. The chain has a few "states," and the probability of being in a state j depends only on the previous state. Markov chains are often described by a graph with transition probabilities. In an HMM, one has an unobservable Markov chain in which each state randomly generates one out of k observations, which are, in fact, visible. In practical settings a series of observations are available from which the most probable corresponding hidden states are to be found. In ML, the process entails parameter learning, namely, one has some dataset, and one wants to find the parameters that best fit the HMM model.
- Random forest (RF) algorithm (e.g., [53, 54]): a structure consisting of a number of decision trees for classification and regression that uses a tree pattern to make a decision. The "forest" is an ensemble of decision trees, usually trained with the "bagging" method where a combination of learning models is able to increase the overall quality of the result. Using and merging multiple decision trees yields more accurate and stable predictions. RF algorithms are used for both classification and regression problems, which comprise the majority of ML systems; it is often used for clustering information.
- K-nearest neighbor (KNN) algorithm (e.g., [55, 56]): an algorithm where a sample is categorized by the majority of its nearest k-neighbors (k is generally a small integer).

- K-means algorithm (e.g., [57, 58]): an algorithm where samples are divided into categories whose members are similar – the k members are randomly distributed among n members selected as cluster centers; the remaining n-k members are then allocated to the nearest cluster. This algorithm is often used for simple clustering of large datasets consisting of high-dimensional numerical data and is very efficient when classifying similar data into the same cluster.
- Gaussian mixture model (GMM) algorithm (e.g., [59]): probabilistically represent subpopulations (according to a normal distribution) within an overall population. An advantage of GMMs is that the subpopulation which a data point belongs to can be unknown. So effectively, the model can learn the subpopulations automatically – a form of unsupervised learning; the methodology is also usable in supervised learning. GMMs have been used in many applications including feature extraction from speech data and object tracking of multiple objects at each frame in a video sequence, where the number of mixture components and their means predict object locations.

Some examples of application in SA in general, and computer vision in particular, using neural network algorithms are described in references [60–63] among others; using support vector machine algorithms, references [64–65]; using Bayesian network algorithms, reference [66] among others; using decision tree algorithms, reference [67]; using hidden Markov model algorithms, references [68–70]; using random forest algorithms, references [71–73]; using K-nearest neighbor algorithms, references [72, 73]; using K-means algorithms, reference [74]; and using Gaussian mixture model algorithms, reference s [75].

Although "machine learning" and "deep learning" are occasionally used synonymously, there are differences. DL as a discipline is a subcategory of ML – and functions in a similar manner; however, DL uses artificial neural networks (ANNs) which autonomously allow machines to make better decisions without any help from humans. While ML models evolve and become better with more experience performing their function, they still need some guidance – for example, if an ML algorithm returns an inaccurate prediction, then, a human intervention is required to make appropriate adjustments. With a DL model, the algorithm automatically assesses prediction accuracy relying on its own neural network, continually analyzing data using logic similar to human would draw conclusions. DL systems use a layered structure of algorithms based on ANNs, using multi-layered processing [76]. DL models are significantly more complex than traditional neural networks, and they also ingest vastly larger amounts of data than their predecessors [77]. From the point of view of a discipline, DL is considered to be an elaboration of ML.

Similar to ML, deep learning also utilizes supervised, unsupervised, and reinforcement learning principles and algorithms – some of the deep learning methods discussed in the cited literature are convolutional neural network, recurrent neural network, long short-term memory, and auto and generative adversarial network. In summary, ML uses algorithms to parse data, learn from that data, and make informed decisions based on what it has learned; DL structures algorithms in layers to create an ANN that can learn and make intelligent decisions on its own.

Among other applications, DL can be used for object recognition (a classification and a regression function), classification (finding patterns and assigning entities to classes), feature extraction, time series prediction, and regression functions such as localization, all of which have SA applications. In particular, DL is assuming an increasingly prominent role in computer vision analytics, which are relied upon by law enforcement agencies.

5 Conclusion

This chapter reviewed practical aspects of SA technologies when used for public safety and law enforcement applications, to provide real-time visibility on the status of prospective threats – and possible targets. SA deals with capturing incoming data, analyzing the data, and generating actionable intelligence. The large amount of data generated by these cameras as well as by the fairly-dense set of stationary cameras found in smart cities, and other sensors, gives rise to the need for advanced situational awareness platforms (SAPs) that can undertake advanced analytics in real time or near real time, likely using ML and DL techniques; many of the systems are now cloud-based. The following chapter focuses on practical tools used for SA by law enforcement agencies; these tools include, among others, the following: in-field policing with body-worn cameras, license plate recognition, object leave behind detection, crowd formation, aberrant behavior, weapons detection/armed intruder, and tracking a target from multiple cameras' views not just a single camera.

References

1. M.R. Endsley, Toward a theory of situation awareness in dynamic systems. Hum. Factors **37**(1), 32–64 (1995)
2. M. Satyanarayanan, Edge computing for situational awareness, in *2017 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN)*, (Osaka, 2017), pp. 1–6. https://doi.org/10.1109/LANMAN.2017.7972129. Available online on December 1, 2020 at https://www.cs.cmu.edu/~satya/docdir/satya-lanman2017.pdf (cmu.edu)
3. S. Rani, R. Maheswar, et al., *Integration of WSN and IoT for Smart Cities* (Springer, 2020) ISBN-10:3030385159, ISBN-13:9783030385156
4. D. Minoli, B. Occhiogrosso, Chapter 12: Internet of things applications for smart cities, in *Internet of Things A to Z: Technologies and Applications*, ed. by Q. Hassan, (IEEE Press/Wiley, June 2018) ISBN-13: 978-1119456742
5. O. Bates, A. Friday, Beyond data in the smart city: Learning from a case study of re-purposing existing campus IoT, in *IEEE Pervasive: Special Issue on Smart Buildings and Cities*, (Jan. 2017)
6. A. Ramaswami, A.G. Russell, et al., Meta-principles for developing smart, sustainable, and healthy cities. Science (AAAS) **352** (6288), 940–943 (20 May 2016). https://doi.org/10.1126/science.aaf7160

7. D. Martínez, T. Gracia, et al., Smart cities' challenge: How to improve coordination in the supply chain, in *Sustainable Smart Cities, Part of the series Innovation, Technology, and Knowledge Management*, (06 October 2016), pp. 129–142. https://doi.org/10.1007/978-3-319-40895-8_10. Print ISBN: 978-3-319-40894-1, Online ISBN: 978-3-319-40895-8
8. N. Mali, A review on smart city through Internet of Things (IoT). Int. J. Adv. Res. Sci. Manage. Technol. **2**(6) (June 2016)
9. S. Srivastava, N. Pal, Smart cities: The support for Internet of Things (IoT). Int. J. Comput. Appl. Eng. Sci. Jorhat **6.1**, 5–7 (Mar 2016)
10. D. Minoli, B. Occhiogrosso, Practical aspects for the integration of 5G networks and IoT applications in smart cities environments, in *Special Issue Titled "Integration of 5G Networks and Internet of Things for Future Smart City," Wireless Communications and Mobile Computing*. vol. 2019 (Hindawi/John Wiley & Sons, August 2019), Article ID 5710834, 30 pp. https://doi.org/10.1155/2019/5710834
11. D. Minoli, B. Occhiogrosso, Ultrawideband (UWB) technology for smart cities IoT applications, in *2018 IEEE International Smart Cities Conference (ISC2) – IEEE ISC2 2018-Buildings, Infrastructure, Environment Track, Kansas City*, (16–19 September, 2018)
12. D. Minoli, B. Occhiogrosso, Mobile IPv6 protocols and high efficiency video coding for smart city IoT applications, in *CEWIT2017*, (2017)
13. D. Minoli, B. Occhiogrosso, Blockchain-enabled fog and edge computing: Concepts, architectures and smart city applications, in *Blockchain-Enabled Fog and Edge Computing: Concepts, Architectures and Applications*, ed. by M. H. Rehmani, M. M. Rehan, (CRC Press, Taylor & Francis Group, Boca Raton, 2020)
14. R. Pew, A. Mavor (eds.), *Modeling Human and Organizational Behavior: Application to Military Simulations* (National Academy Press, 1998)
15. A. D'Amico, M. Kocka, Information assurance visualization for specific stages of situational awareness and intended uses: Lessons learned, in *Workshop on Visualization for Computer Security, USA*, (2005)
16. M. R. Endsley, D. J. Garland (eds.), *Situation Awareness Analysis and Measurement* (Lawrence Erlbaum Associates, Mahwah. ISBN: 0-8058-2133-3, 2000)
17. C. Onwubiko, Designing information systems and network components for situational awareness, in *Situational Awareness in Computer Network Defense*, (January 2012), pp. 104–123. https://doi.org/10.4018/978-1-4666-0104-8.ch007. Available online on December 1, 2020 at https://www.c-mric.com/wp-content/uploads/2019/12/CyberSA_Design_Requirements.pdf
18. B. McGuinness, L. Foy, A subjective measure of SA: The crew awareness rating scale (CARS), in *Proceedings of the First Human Performance, Situation Awareness, and Automation Conference*, (Savannah, Georgia, 2000)
19. National Research Council, *Modeling Human and Organizational Behavior: Application to Military Simulations* (The National Academies of Sciences, Engineering, and Medicine, Washington, DC, 1998) ISBN: 0309060966, ISBN13: 9780309060967
20. M. Endsley, SAGAT: A methodology for the measurement of situation awareness, in *NOR DC 87–83*, (Northrop Corporation, Los Angeles, CA, 1987)
21. M. Endsley, Predictive utility of an objective measure of situation awareness, in *Proceedings of the Human Factors Society 34th Annual Meeting*, (Santa Monica, CA, 1990), pp. 41–45
22. P.A. Nobel, Schema-based knowledge elicitation for planning and situation assessment aids. IEEE Trans. Syst. Man Cybern. **19**(3), 473–482 (May–June 1989)
23. N.B. Sarter, D.D. Noble, Situation awareness: A critical but ill-defined phenomenon. Int. J. Aviat. Psychol. **1**, 45–47 (1989)
24. N.B. Sarter, D.D. Noble, How in the world did we ever get into that mode? Mode error and awareness in supervisory control. Hum. Factors **37**(1), 5–19 (1995)
25. C. Dominguez, Can SA be defined? in *Situation Awareness: Papers and Annotated Bibliography*, ed. by M. Vidulich, C. Dominguez, et al., (Vol AL/CF-TR-1994-0085, Armstrong Laboratory, Wright-Patterson Air Force Base, OH, 1994)

26. R.W. Pew, The state of situation awareness measurement, in *Experimental Analysis and Measurement of Situation Awareness*, ed. by D. J. Garland, M. R. Endsley, (Embry-Riddle Aeronautical University, Daytona Beach, FL, 1995), pp. 7–15
27. J.M. Flach, Situation awareness: Proceed with caution. Hum. Factors **37**(1), 149–157 (1995)
28. D.R. Stiffer, Graduate level situation awareness, in *USAF Fighter Weapons Review*, (Summer 1988), pp. 115–120
29. D.O.W. Banerjee, *Final Report from the NSF Workshop on Future Directions in Wireless Networking* (National Science Foundation, November 2013)
30. P. Simoens, Y. Xiao, et al., Edge analytics in the Internet of Things. IEEE Pervasive Comput. **14**(2), 24–31 (April–June 2015)
31. T. Shi, S. Horvath, Unsupervised learning with random forest predictors. J. Comput. Graph. Stat. **15**(1), 118–138 (2006). https://doi.org/10.1198/106186006X94072
32. M. Caron, P. Bojanowski, et al., Deep clustering for unsupervised learning of visual features, in *Proceedings of the European Conference on Computer Vision (ECCV)*, (2018), pp. 132–149
33. X. Zhai, A. Oliver, et al., Self-supervised semi-supervised learning, in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, (2019), pp. 1476–1485
34. N.N. Pise, P. Kulkarni, A survey of semi-supervised learning methods, in *2008 International Conference on Computational Intelligence and Security, Suzhou*, (2008), pp. 30–34. https://doi.org/10.1109/CIS.2008.204
35. L. Collingwood, J. Wilkerson, Tradeoffs in accuracy and efficiency in supervised learning methods. J. Inform. Tech. Polit. **9**(3), 298–318 (2012). https://doi.org/10.1080/19331681.2012.669191
36. R.A. Caruana, A. Niculescu-Mizil, An empirical comparison of supervised learning algorithms, in *ICML '06: Proceedings of the 23rd International Conference on Machine Learning*, (June 2006), pp. 161–168. https://doi.org/10.1145/1143844.1143865
37. V. Mnih, A. Puigdomènech, et al., Asynchronous methods for deep reinforcement learning, in *Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA*, (JMLR: W&CP Volume, 2016)
38. C. Wirth, R. Akrour, et al., A survey of preference-based reinforcement learning methods. J. Mach. Learn. Res. **18**(1), 1–46 (January 2017)
39. D. Djenouri, R. Laidi, et al., Machine learning for smart building applications: Review and taxonomy. ACM Comput. Surv. **52**(2), Article No. 24 (March 2019). https://doi.org/10.1145/3311950
40. F. Wu, K. Fu, et al., A spatial-temporal-semantic neural network algorithm for location prediction on moving objects. Algorithms **10**, 37 (2017)
41. A. Wang, W. Yuan, et al., A novel pattern recognition algorithm: Combining ART network with SVM to reconstruct a multi-class classifier. Comput. Math. Appl. **57**(11–12), 1908–1914., ISSN 0898-1221 (2009). https://doi.org/10.1016/j.camwa.2008.10.052
42. R. Pupale, Support Vector Machines (SVM) – An overview (16 June 2018). Towards Data Science. Available online on December 1, 2020 at https://towardsdatascience.com/https-medium-com-pupalerushikesh-svm-f4b42800e989
43. İ. Güven, F. Şimşir, Demand forecasting with color parameter in retail apparel industry using Artificial Neural Networks (ANN) and Support Vector Machines (SVM) methods. Comput. Ind. Eng. **147**, 106678., ISSN 0360-8352 (2020). https://doi.org/10.1016/j.cie.2020.106678
44. J.L. Awange, B. Paláncz, et al., Support vector machines (SVM), in *Mathematical Geosciences*, (Springer, Cham, 2018). https://doi.org/10.1007/978-3-319-67371-4_10
45. D. Soni, Introduction to Bayesian networks (8 June 2018). Towards Data Science. Available online on December 1, 2020 at https://towardsdatascience.com/introduction-to-bayesian-networks-81031eeed94e
46. A.L. Madsen, F. Jensen, et al., A parallel algorithm for Bayesian network structure learning from large data sets. Knowl.-Based Syst. **117**, 46–55., ISSN 0950-7051 (2017). https://doi.org/10.1016/j.knosys.2016.07.031
47. X. Sun, C. Chen, et al., Hybrid optimization algorithm for Bayesian network structure learning. Inform **10**, 294 (2019)

48. M. Somvanshi, P. Chavan, et al., A review of machine learning techniques using decision tree and support vector machine, in *2016 International Conference on Computing Communication Control and Automation (ICCUBEA)*, (Pune, 2016), pp. 1–7. https://doi.org/10.1109/ICCUBEA.2016.7860040
49. Y. Mu, X. Liu, et al., A parallel C4.5 decision tree algorithm based on MapReduce, in *Concurrency and Computation Practice and Experience*, (Wiley Online, 21 February 2017). https://doi.org/10.1002/cpe.4015
50. T. Amit, Introduction to hidden Markov models (7 June 2019). Towards Data Science. Available online on December 1, 2020 at https://towardsdatascience.com/introduction-to-hidden-markov-models-cd2c93e6b781
51. S. Srinivasan, G. Gordon, B. Boots, Learning hidden quantum Markov models, in *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research (PMLR) 84:1979–1987*, (2018)
52. Y. Zhou, R. Arghandeh, C.J. Spanos, Online learning of contextual hidden Markov models for temporal-spatial data analysis, in *2016 IEEE 55th Conference on Decision and Control (CDC)*, (Las Vegas, NV, 2016), pp. 6335–6341. https://doi.org/10.1109/CDC.2016.7799244
53. M. Schonlau, R. Yuyan Zou, The random forest algorithm for statistical learning. Stata J. Promot. Commun. Stat. Stata (24 March 2020). https://doi.org/10.1177/1536867X20909688
54. J. Abellán, C.J. Mantas, et al., Increasing diversity in random forest learning algorithm via imprecise probabilities. Expert Syst. Appl. **97**, 228–243., ISSN 0957-4174 (2018). https://doi.org/10.1016/j.eswa.2017.12.029
55. T. Cover, P. Hart, Nearest neighbor pattern classification. IEEE Trans. Inf. Theory **IT-13**(1), 21–27 (Jan. 1967)
56. Y. Wang, Z. Pan, Y. Pan, A training data set cleaning method by classification ability ranking for the *k*-nearest neighbor classifier. IEEE Trans. Neural Netw. Learn. Syst. **31**(5), 1544–1556 (May 2020). https://doi.org/10.1109/TNNLS.2019.2920864
57. S.-S. Yu, S.-W. Chu, et al., Two improved K-means algorithms. Appl. Soft Comput. **68**, 747–755., ISSN 1568-4946 (2018). https://doi.org/10.1016/j.asoc.2017.08.032
58. M. Ahmed, R. Seraj, S.M.S. Islam, The k-means algorithm: A comprehensive survey and performance evaluation. Electronics **9**(8), 1295 (2020). https://doi.org/10.3390/electronics9081295
59. J. McGonagle, G. Pilling, A. Dobre, Gaussian mixture model. Available online on December 1, 2020 at https://brilliant.org/wiki/gaussian-mixture-model/
60. T. Gandhi, M.M. Trivedi, Computer vision and machine learning for enhancing pedestrian safety, in *Computational Intelligence in Automotive Applications. Studies in Computational Intelligence*, ed. by D. Prokhorov, vol. 132, (Springer, Berlin/Heidelberg, 2008). https://doi.org/10.1007/978-3-540-79257-4_4
61. S.R. Khan, H. Rahmani, et al., A guide to convolutional neural networks for computer vision. Synth. Lect. Comput. Vision **8**(1), 1–207 (February 2018). https://doi.org/10.2200/S00822ED1V01Y201712COV015
62. E. Lygouras, N. Santavas, et al., Unsupervised human detection with an embedded vision system on a fully autonomous UAV for search and rescue operations. Sensors (Basel) **19**(16), 3542. Published 2019 Aug 14 (2019). https://doi.org/10.3390/s19163542
63. M.S. Veillette, E.P. Hassey, et al., Creating synthetic radar imagery using convolutional neural networks. J. Atmos. Ocean. Technol. **35**(12), 2323–2338 (2018). https://doi.org/10.1175/JTECH-D-18-0010.1
64. Y. Ma, W. Chen, et al., EasySVM: A visual analysis approach for open-box support vector machines. Comp. Visual Media **3**, 161–175 (2017). https://doi.org/10.1007/s41095-017-0077-5
65. F. Alam, F. Ofli, M. Imran, Processing social media images by combining human and machine computing during crises. Int. J. Hum. Comput. Stud. **2018** (2018). https://doi.org/10.1080/10447318.2018.1427831

66. B. Mertens, L. Rothkrantz, P. Wiggers, Dynamic Bayesian networks for situational awareness in the presence of Noisy data, in *ACM International Conference Proceeding Series*, (2011). https://doi.org/10.1145/2023607.2023676
67. S.C. Satapathy, K.S. Raju, et al., Advances in decision sciences, image processing, security and computer vision, in *International Conference on Emerging Trends in Engineering (ICETE)*, vol. 1, (2020) ISBN: 978-3-030-24321-0;978-3-030-24322-7
68. F. LeGland, L. Mevel, Recursive estimation in hidden Markov models, in *The 36th IEEE Conference On Decision And Control*, vol. 4, (San Diego, USA, 1997), pp. 3468–3473
69. V.E. Balas, L.C. Jain, B. Kovačević (eds.), Soft computing applications, in *Proceedings of the 6th International Workshop Soft Computing Applications (SOFA 2014)*, vol. 2 ISSN 2194-5357 ISSN 2194-5365 (electronic). Part of Advances in Intelligent Systems and Computing, ISBN 978-3-319-18415-9 ISBN 978-3-319-18416-6 (eBook). https://doi.org/10.1007/978-3-319-18416-6
70. M. Moussa, G. Beltrame, On the robustness of consensus-based behaviors for robot swarms. Swarm Intell. **14**, 205–231 (2020). https://doi.org/10.1007/s11721-020-00183-1
71. N. Kausar, A. Majid, Random forest-based scheme using feature and decision levels information for multi-focus image fusion. Pattern. Anal. Applic. **19**(1) (February 2015). https://doi.org/10.1007/s10044-015-0448-4
72. B. Arshad, R.I. Ogie, et al., Computer vision and IoT-based sensors in flood monitoring and mapping: A systematic review. Sensors **19**(22), 5012 (2019). https://doi.org/10.3390/s19225012
73. Y. Wang, H. Xia, et al., Distributed defect recognition on steel surfaces using an improved random forest algorithm with optimal multi-feature-set fusion. Multimed. Tools Appl. **77**(13), 16741–16770 (July 2018). https://doi.org/10.1007/s11042-017-5238-0
74. N. Naikal, Towards autonomous situation awareness. Technical Report No. UCB/EECS-2014-124 (21 May 2014). Electrical Engineering and Computer Sciences University of California at Berkeley. Available online on December 1, 2020 at https://www2.eecs.berkeley.edu/Pubs/TechRpts/2014/EECS-2014-124.pdf
75. H.H. Santoch, P. Venkatesh, et al., Tracking multiple moving objects using Gaussian mixture model. Int. J. Soft Comput. **3** (2), 113–119 (May 2013). ISSN: 2231-2307
76. B. Grossfeld, Deep learning vs machine learning: A simple way to understand the difference (23 January 2020). Available online on December 1, 2020 at https://www.zendesk.com/blog/machine-learning-and-deep-learning
77. K. Nevala, *The Machine Learning Primer* (SAS Best Practices e-book, SAS Institute Inc, Cary, 2017)

Daniel Minoli, Principal Consultant, DVI Communications, New York, graduate of New York University, has published 60 technical books and 350 papers and made 90 conference presentations. He has many years of technical and managerial experience in planning, designing, deploying, and operating secure IP/IPv6, VoIP, telecom, wireless, satellite, and video networks for carriers and financial companies. He has published and lectured in the area of M2M/IoT, network security, satellite systems, wireless networks, IP/IPv6/Metro Ethernet and has taught at New York University, Stevens Institute of Technology, and Rutgers University. He also has served as a testifying expert witness.

Andrzej Koltun, Director of Network and Systems Architecture, DVI Communications, has technical responsibility for many of DVI's projects, including aspects of: strategic planning, network and system design, and network implementation management. Mr. Koltun has designed and deployed large-scale networks for over 35 years supporting multiple missions – critical applications including financial services, security, and telemetry. Mr. Koltun possesses an extensive background in real-time video surveillance applications using IP.

Benedict Occhiogrosso, is a Co-founder of DVI Communications, New York. He is a graduate of New York University Polytechnic School of Engineering. His experience encompasses technical and managerial disciplines. He also served as a testifying expert witness in numerous cases encompassing patent infringements and other legal matters.

Situational Awareness for Law Enforcement and Public Safety Agencies Operating in Smart Cities – Part 2: Platforms

Daniel Minoli, Andrzej Koltun, and Benedict Occhiogrosso

1 Situational Awareness for Law Enforcement Agencies

In Part 1 of this two-chapter set, it was highlighted that law enforcement and public safety agencies require full, real-time visibility on the status of prospective threats – and possible targets. Situational awareness (SA) deals with capturing incoming data, analyzing the data, and generating actionable intelligence. Recent technology advancements in video sensing, edge computing, big data storage, and machine learning (ML) can be leveraged to generate powerful real-time information – managed by situational awareness platforms (SAPs) – that optimize real-time decision-making. These concepts are explored in full in this chapter.

SA for law enforcement agencies includes the following applications (use cases), among others:

- In-field policing with body-worn cameras (BWCs) (typically for after-the-fact video analysis, but occasionally for real-time SA).
- Identification of threats via facial recognition.
- License plate recognition.
- Object "leave behind" detection.
- Real-time robotic survey of crime environments (e.g., suspected bombs).
- Crowd formation.
- Aberrant behavior – e.g., turns left when should turn right.
- Speedy approach – e.g., car accelerating.
- Tracking a target from multiple camera view not just single camera.
- Weapons detection/armed intruder.

D. Minoli (✉) · A. Koltun · B. Occhiogrosso
DVI Communications, New York, NY, USA
e-mail: daniel.minoli@dvicomm.com; andrzej@dvicomm.com; ben@dvicomm.com

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_9

These applications can also be seen as expressing requirements.

The first item in this list identifies a key use case. However, the video processing is typically done after the fact, rather than in real time to sustain or achieve a SA goal. Many community stakeholders and criminal justice leaders have suggested that placing BWCs on police officers improves the civility of police-citizen encounters and enhances citizen perceptions of police transparency and legitimacy. BWCs are recording devices police officers wear as part of their uniforms to document what they see as they perform their duties. In the USA, 34 states and the District of Columbia had created laws for body cameras by early 2018. As of mid-2017, over one-third of the 18,000 or so law enforcement agencies in the USA have begun using this technology, and the number has gradually increased since then [1]. A key study on body-worn cameras, a field experiment conducted by the Rialto, CA., Police Department, found that when police were equipped with cameras during the test period, use of force incidents and citizen complaints against officers were reduced by 50% and 90% respectively. Their findings spurred national interest in the benefits body cameras could potentially provide [1–6].

Successful implementation not only requires the deployment of an entire technical ecosystem, but it also often requires substantial changes to a police organization and its partners (e.g., the prosecutor's office, defense attorneys, the judiciary, in policy, training, staffing, investigations, and technology) [1]. BWCs may de-escalate aggression or have a "civilizing" effect on the nature of police-citizen encounters. The complaint and use of force reductions associated with placing BWCs on police officers may be particularly important for improving police-community relations in impoverished neighborhoods.

It has been estimated that a police department typically could result in an average of 25% fewer community complaints when using BWCs. Thus, the benefits derive from the estimated decrease of complaints, as well as the reduced cost to investigate each complaint (because of the available BWC video evidence), and the reduced amount of time it takes to resolve a complaint when video evidence is available [1]. BWC approaches have a cost implication: the installation, training, operation, and maintenance costs; these costs include both one-time (e.g., facilities and infrastructure upgrades) and recurring costs (e.g., licenses and storage); many of the costs are up-front investments in assets that have useful lives exceeding 1 year. Estimates show that BWCs cost between $828 and $1097 per user per year. But they have a cost-saving implication: namely, the number of complaints can be reduced driving a measurable administrative/processing cost-displacement. Studies have shown that BWCs can save over $6200 in officer time spent investigating an average complaint, compared to complaint investigations for officers without BWCs [1]. BWCs generate savings mainly through significantly faster investigation of complaints. The cited study (a randomized controlled 2017 trial with BWCs in the Las Vegas Metropolitan Police Department that included over 400 officers) assumed, based on extensive analysis, that there would be 0.84 complaints per officer each year in the absence of BWCs. The "break-even" level of complaints occurs between 0.23 and 0.27 complaints per officer per year. At the breakeven level, the costs avoided by BWCs would just offset the costs to implement BWCs.

Most notably, further applying the cost-benefit estimates to all 1400 patrol officers suggests BWC net annual savings of $4.1 million to $4.4 million department-wide.

In the USA, as of early 2018, five states have enacted laws that require at least some officers to use body-worn cameras. South Carolina's law requires, contingent on state funding, that every police department implement a body camera program. Nevada requires that uniformed peace officers who work for law enforcement agencies that routinely interact with the public must wear body cameras. In California certain members of their state highway patrol are required to wear body cameras. Connecticut's law requires their division of state police and special police forces, as well as municipal police officers receiving grant funds, to wear body cameras while interacting with the public [2]. Furthermore, 13 states and the District of Columbia have legislated funding opportunities for state and local police departments to purchase body-worn camera equipment, hire new support staff, and operate or purchase data systems. Twenty-three states and the District of Columbia have legislated how body-worn camera data is addressed under open record laws. In statute, states specify procedures for the public to request footage and which footage is and is not to be released to the public. The goal of these provisions is to be transparent in law enforcement without unnecessarily infringing on privacy. Nineteen states and the District of Columbia require written policies in order for law enforcement to use or receive funding for body-worn cameras. Legislation sets statutory minimum standards for policies, requires state entities to write or coordinate the development of policies, and charges individual departments with creating their own policies.

SA is part and parcel of a BWC environment. BWCs represent another data input (consider it a mobile video sensor) that has to be properly processed. In that context, a complex information technology (IT) infrastructure that spans in-the-field wireless networking, core transport, and cloud-based analytics, especially as software-as-a-service (SaaS) applications, is needed.

1.1 Characterization of a Situational Awareness Platform

All of the law enforcement use cases listed above warrant the utilization of analytics to alert and mobilize personnel of an event of interest (forthcoming or ongoing). Advanced automation has become a clear requirement to assimilate the data generated by the myriad of appropriate sensors and cameras for this plethora of applications, especially for large municipalities. In the discussion that follows, however, a distinction is made between constituent technology sub-elements (some or all of which can be combined in a number of specific or parochial ways) and a complete platform. A number of distinct subsystems comprise the overall SAP ecosystem, as listed below.

- A plethora of static and mobile input systems, sensors, and cameras collect city-wide real-time data of all sorts.
- SA video aggregation subsystem: harnessing the video ecosystem.
- SA federated query subsystem: harnessing the big data ecosystem.
- SA analytics subsystem: data mining for actionable information.
- SA event correlation subsystem: analysis also using ML techniques.
- SA analytics/presentation subsystem: harnessing information overload for actionable decisioning.

A SA ecosystem entails a large number of mobile and static input systems such as cameras and various other sensors; in addition, there will be video aggregation subsystems, advanced analytics engines (including event correlation techniques typically utilizing ML techniques), and data storage and mining facilities, as well as an effective query capability, that can be used in conjunction with display/presentation tools that make the delivery of actionable decisioning information to the intended enforcement agencies an intuitive and expeditious process.

SAPs integrate these capabilities into one overall system umbrella. SAPs are the commercially available technology systems that *help* facilitate a state of awareness of a certain environment, for example, a city, a park, a shopping area, a train station, an airport, a stadium, or a (large) indoor space. There are many variations among these SAPs, with each one providing some enhancement of one's situational awareness. There are many products on the market that are portrayed as SAPs, but these products provide those capabilities within a narrow operational area, for example, video optimization, or further refined as object recognition, or even further narrowed as facial recognition. "Wholistic platforms" provide mechanisms that enhance one's SA across a broad range of input mediums that include visualization, traditional data, and sensors of many types and, most significantly, enable the correlation of information across all of these content providers.

A SAP ideally fulfills the specific function of being a "tripwire" and enhancing real-time event response by the authorities by providing timely and effective decision support services. Unfortunately, SAPs are often deployed in a much broader role or capacity, which contributes to what can be considered as the dilution of their brand recognition: because of the capabilities of these platforms, they are also used across a much broader role spectrum, thus overlapping with Enhanced 911 (E-911) response functions, video evidence management, crime analysts' research tools, and so on. Consequently, many products are correctly branded as "situational awareness systems," but incorrectly marketed as SAPs. Such misuse dilutes the critical need of the SAPs in law enforcement for speed and accuracy.

Where the analyst's support software needs to be complete and accurate, the SAP needs to be as complete/accurate as possible within the available timeframe which could be real time or near real time. In that role, the SAP closely resembles the operating environment of the autonomous vehicle: the road path must be selected before striking the divider. Video coverage must immediately focus on a zone of interest, for example, triggered by gunshot detection, whether by sensor, phone call, or social media. Background information and known associates are needed while in the act of responding, much more so than during the subsequent interrogation. However, because the tools can sometimes do both, they are tasked as multi-tools and are biased towards completeness more so than timeliness.

While SAPs' ability to access information is the typical selling point of SAPs, their ability to filter out insignificant content is emerging as an important functional capability. It was only in the not-so-distant past that law enforcement had very limited access to video surveillance images and then only after lengthy delays. At this time, video imagery exists almost everywhere and is accessible in real time; the present-day challenge, therefore, is how to simultaneously watch and/or analyze hundreds of cameras. Many municipalities currently have E-911, 311, 411, 611 mechanisms, coupled with an endless array of hotlines and nonstop social media. Therefore, it tends to be the case that there is an overabundance of information but not enough real-time actionable summarization. In many municipalities, one sees a trend for adding more content sensors, but they are not adding more content sensor analysis/monitoring tools at a matching pace; therefore, the content sensors become merely data that is not acted upon. In fact, not only are these sensors not helping, but they are actually hurting the public image of law enforcement through the discovery that the information existed, if only one looked in the right place. SAPs are being considered for deployment as a tool that corrals the large arrays of cameras, sensors, data, and so on into some form of manageable scale.

Application of artificial intelligence (AI) for SA within public safety, and more specifically law enforcement, places an extra burden on accuracy and speed, where neither can be sacrificed for the benefit of the other. In that context the decision paradigm more closely approximates the real-time field conditions of autonomous vehicles than those of the crime analyst's/investigator's environment: their decision support guidance needs to be fast and accurate.

1.2 Broad Functional Requirements of a Situational Awareness Platform for Law Enforcement

Functional requirements of a SAP to be utilized by law enforcement agencies seeking support for the use cases described above address two somewhat distinct categories, although there is a technology overlap in the mechanisms:

1. Historical non-real-time analysis of large amounts of sensor data, typically video-based, but also comprised of other types of information and/or data (including Big Data).
2. Real-time or near-real-time analysis of environmental (ecosystem, theater) data to support intervention decisioning on the part of law enforcement to deal with an ongoing safety matter.

The discussion that follows focuses mostly on the latter. A well-designed SAP provides a single-pane enterprise-wide view culled from the disparate systems and technologies. A SAP typically consists of a broad range of underlying, yet, disparate technologies. The systems and technologies providing public safety and law enforcement agencies input content are many and vary depending on the agency and the environment in which it operates. The technologies continuously capturing data and are leveraged within the platform include visual input, sensor data, content analysis, informational records, and real-time dispatch activities.

These technologies represent a broad range of applications and, within each application category, a very broad range of capabilities. At press time the vendor community was adding capabilities to their SAP products at an accelerated pace. While specific applications and their underlying technologies may be replaced or enhanced over time, their general function category will remain a viable requirement.

Many SA environments require computer vision capabilities. This field of computer vision is currently a very active area of research particularly in the context of using AI and deep learning (DL) tools. Work in this arena started in the mid-1960s and continued through the late 1970s, but with modest success due to limits on AI technology at the time. The use of convolutional neural networks (CNNs) was advanced in the 1980s, although the volume of data and computing resources requires were also limiting. A CNN comprises multiple layers of artificial neurons. When the CNN processes an image, each of its layers extracts specific features from the pixels. The first layer detects very basic things, such as vertical and horizontal edges. As one moves deeper into the neural network, the layers detect more-complex features, including corners and shapes. The final layers of the CNN detect specific things such as faces, doors, and cars. The output layer of the CNN provides a table of numerical values representing the probability that a specific object was discovered in the image [7]. It was not until the 2010s that CNNs would resurface and trigger a revolution in deep learning. Recent advances in CNN and DL technology for computer vision have become prevalent and include but are not limited to facial recognition (a short list of references with emphasis on SA include [8–13]; a large set of literature exists).

The general public's perception of a SAP focuses on its visualization components; however, the most significant functionality is derived from the applications/methods that access informational records; these records present a wealth of historical knowledge about the ecosystem and events within that ecosystem. There is significant innovation underway in this field: not long ago information was so siloed that mere access to it, as cumbersome as it was, was viewed as an accomplishment; investigatory platforms now available strip away that cumbersome access mechanism, presenting many siloed result sets. At this time, SAPs are beginning to apply correlation analysis logic to help facilitate decision-making by digesting the accessed documents and highlighting the common elements.

The analytic correlation capability represents a key function of a SAP: while multiple systems used by crime analysts have this capability, the unique requirement for the SAP is timeliness and presentation. The operator staffing the display console of SAP has such requirements. Because the key role of a SAP is to support effective decision-making within the real-time duration of an event, the SAP needs the capability to access multiple disparate records, to analyze the correlations among the result sets, and to display all the potential result sets in some meaningful manner. Often enough the least probable become reality; thus the low probability result must not be hidden from the operator's eyes and mind. Computers do things faster, but often humans do things better. For the SAP operator, it is important to provide them the opportunity to display this capability. As such, because they all differ, how a candidate SAP handles the access to records, the correlation among those records, and the display of the result sets are all meaningful differentiators in the platform evaluation/selection process.

911/E-911 dispatchers now use computer-aided dispatch (CAD) systems and records management systems (RMS) to log these emergency calls and take the appropriate steps to dispatch police to the scene. CAD systems allow for the capture of critical data related to each call: phone number, address, call type, and call-related times, such as call received, dispatch time, unit responses. and unit arrivals. The system can also provide recommendations related to stations or responding units as well as access to pre-plan information. They serve as essential information hubs for dispatchers, police officers, and the community alike. The CAD system is also the point of origin for all types of data that provides demographics and research materials and feeds the departmental demand analysis [14]. Access to real-time dispatch information as provided by CAD/RMS platforms is often considered of key value. Sometimes that value is assigned based on the personal and historical perspective of the individuals guiding the selection process. Timely and integrated CAD/RMS information is important, but only as an input source to the overall SA view. The SAP operator does not displace the CAD function nor the scene commander. The SAP operator's role is to merely add a hundred focused eyes and handfuls of savant brains with instant recollection to the tool sets of the responding individuals. The extent of the integration of CAD/RMS is a consideration in SAP selection, but it should not be a deciding factor.

Both visualization input and sensor data can generate overwhelming content input to decisions-makers that if not effectively managed tends to obscure meaningful information. A mosaic view of hundreds of cameras is great if not presented concurrently. Dozens of beeping sensor alarms simply confuse the mind. Instead of adding to the sensory overload, what is most needed is relevant content from the universe of input devices. Pedestrians clutter the scene when one is looking for a car. When looking for "The Car or The License Plate," all other vehicles add to the clutter. One of the values of a SAP is its ability to start with the universe of input and quickly focus on the singular object of interest. That focus can be objects of interest, physical areas of interest, or in fact any entity grouping that becomes a need at any point in time. The capabilities of a SAP to achieve these needs is a key evaluation criterion in the selection process. Another value of a SAP is to alert and focus attention on what is not being actively looked at. Sensors and analytics provide this key functionality, which is an area of major differentiation in product offerings that highlights the difference in how vendors approach assembling the SAP.

1.3 Input Devices and Data Acquisition Technologies

To provide actionable information, SAP processed content must be made available to the decision-makers. Such content is made available through a myriad of technologies that are in place based on their own merit but when accessible by the SAP serve an enabling function. In theory, the SAP would merely tap into the underlying technology ecosystem, though pragmatically it sometimes drives the deployment of these underlying technologies to fill coverage gaps. In that context it is beneficial to address some of the areas where gaps may exist and the characteristics of best-of-breed approaches to their fulfillment.

Input devices typically include sensors, municipal cameras, and other devices (e.g., commercial cameras operated by businesses or end-user video – these are typically used for after-the-fact video analysis, but occasionally for real-time SA).

Sensors continually sample the environment and signal when they sense the event under observation. Within any sensor category, there is a significant difference in the mechanics of how the sensor functions and in the accuracy of its detection capability. Sensors provide the most accurate method of detecting occurrence and thus should be extensively deployed as is appropriate to the operating environment. Gunshot detection sensors illustrate some of the issues. All gunshot detection sensors operate on the basis of detecting environmental changes that result from a gunshot. There are acoustical, visual, chemical indicators and sensors that operate independently or in concert with each other. Some sensors report to on-premises servers, while others report to cloud-based servers. Some sensors monitor a small coverage area, while others monitor a larger area. Some gunshot indicators behave differently indoors than outdoors, introducing specialty products. Lastly, as all sensors are victim to some degree of inaccuracy, some sensor products are marketed as a service that attempts to validate the accuracy before dispatching the event signal to the SAP. The key point here is that while gunshot detection sensors are an important component, product selection should be based on the significance/frequency of this event category in each jurisdiction. Because the best indoor gunshot detectors are very different from the best outdoor gunshot detectors, and there is a 20-fold (20x) price difference between products, sensor selection should very much be based on the degree of need for this type of sensor within the operating environment of the agency. Most sensor systems, however varied, are provided as independent products that are integrated into the SAP, which is different than how analytics are incorporated in the SAPs. All SAPs provide some imbedded analytics capabilities. These analytics come as correlation capabilities for informational records data and as image analytics for video streams. Bundling these analytics capabilities as software in the SAP presents issues. Because of the public safety procurement preference for the tried and true of a complete solution, inclusion of the analytics package serves as a check-off box, which defers assessment of their capabilities. Those capabilities span a broad range of functionality with significant qualitative and performance variations in both accuracy and speed.

The four areas in which one currently finds the greatest voids are in (i) access to informational data records, (ii) monitoring social media activity, (iii) acquiring visual surveillance, and (iv) the analytics of visual imagery. Attempts to remedy such voids often result in the deployment of the underlying technology, which leads us to this discussion of the issues that best-of-breed approaches attempt to fulfill or avoid.

1.3.1 Access to Informational Data Records

Data record information exists in countless database repositories, many of which relate to criminal justice, which are accessed via a structured methodology that controls access through strict credentialing and audit processes. These access control mechanisms were developed individually when each was an independent system; consequently, their credentialing, identification, and access audit techniques differ. Furthermore, the fee structures associated with access to these records differ based on method of access and the business contract negotiated with each jurisdiction. For example, access to records through a CAD/RMS platform can be governed by a single fee master agreement, but access to the same content by the same individual through a client application may be billed on a per-access basis or vice versa. Because of the fluidity of these agreed-upon methods, implementation of SAPs includes the task of negotiating the least costly records access method, which often results in convoluted configurations. As a result, the best practices approach is to customize the processes to the unique conditions of each agency.

1.3.2 Acquiring Visual Surveillance

Visual imagery is considered a critical element of SAPs, but access to that video content is its Achilles heel and main burden. There are many video landscapes today. Municipalities have stationary cameras deployed by many entities for numerous purposes. One finds mobile cameras deployed in numerous fashions, and one cannot overlook the wealth of content produced by the citizenry's cellphone cameras. Some say that regardless of what happened, one is likely to have a video of it. This may be true today, or it may be a goal, but the reality of these challenges to locate events quickly and gain access to that video content is inescapable. In this context, the need of the SAP's access to video is unique in that it must be quickly available. Ideally it can trigger an event alarm, but if the video imagery is not available as the events unfold, it loses its value to the SAP and reverts to being an investigator's support tool.

Today's situations find SAPs being added to existing camera environments. In such conditions, best practices take a back seat to enabling practices. Since enabling practices are governed by the notion of worth, many localities have evolved to the pragmatic decision that they will use whatever approach is available to enable access to the camera resource deemed valuable. While field conditions govern the immediate options, the short lifetime of these video resources avails us the opportunity to prepare for their planned obsolescence with best practice approaches.

Any best practice discussion needs to start with an understanding of the core function of the SAP. Although it is capable of performing many functions, its core role is to provide real-time SA to be applied by its operator to unfolding events. In that role, the SAP needs access to observe live video and to retrieve (playback) recently recorded video. In that context the SAP is a voyeur on someone else's video system. Because there are many "someone else's," accessing each particular video system presents major technical and bureaucratic challenges.

1.3.3 Monitoring Social Media Activity

The phenomenon of social media can significantly enhance SA through access to the content being shared and the quantity of activity. Law enforcement has been using a range of stand-alone applications for these purposes. The ideal condition is to integrate this social media information into the SAP and to use the platform's correlation capabilities as one more trip wire to alert the operator of possible event activity. Recent events have raised attention to law enforcement's use of its access to social media information, which introduced it as a privacy concern. This public debate is causing the administrators of these social media applications to constrain the sharing of content. Consequently, law enforcement needs to periodically change the applications it uses for its access to critical social media information. The growing frequency for this need to alter applications reintroduces the debate about whether a "product solution" or an "amalgamation of products" is the best practice approach to a SAP solution. Regardless of outcome SAPs will be well served if they embrace the capability to easily "plug in" modules (applications) that provide an effective point solution at a specific time and to easily "unplug" and replace them with succeeding products as the conditions dictate.

1.3.4 Analytics Software

SA analytics software in law enforcement settings is, in some sense, bleeding-edge technology. Modern analytics software employs DL techniques to attain its AI capabilities. As discussed earlier, DL can be described as teaching a computer to do one thing very well and very fast. For example, facial recognition, license plate recognition, and long-gun recognition are all specializations within the concept of object recognition; however, they are each very different form each other. Even the subtle differences represent huge differences in capabilities. In this context, DL is not knowledge, nor is it the skill to recognize. DL is the ability of software to recognize a specific type of object that it learned through a reiterative process of observing a million instances of many incarnations of that type of object and eliminating all objects that are not that object. In effect it is a process of recognizing it by learning all the things that it is not. Because of the specificity of these learning processes, each recognition module is effectively its own application. Thus, one has facial-recognition systems, license plate recognition systems, gun recognition systems, and object recognition systems that are able to search on a match for a selected object. Under the perception that DL is too slow for the real-time recognition needs of certain security-surveillance applications, one has started to see evolution in recognition analytics to the mechanisms used in autonomous vehicles.

In addition to object specificity, video analytics also has a major divide between real-time triggering and forensic search. Under real-time triggering, the live video stream is continuously analyzed for any number of specific conditions, which if encountered would trigger a system event. Forensic search is video playback technology that analyzes each stream of recorded video. There are numerous technologies and techniques used for this function, though generically it is a technique to expedite locating an object of interest in a recorded video archive. As with all things from the either-or category, there are hybrid approaches. In video analytics the hybrid approach comes in the form of edge analytics. Namely, it is the process of executing the analytics logic in or local to the camera, as opposed to streaming the video to the remote recorder and executing the analytics at the central server. There are benefits and drawbacks to each approach, but for each specific scenario, there is an absolutely best approach.

This rapid rate of change, plus the specificity in system capabilities, is the reason why acquiring the SAP with an imbedded video analytics solution could be a suboptimal strategy. It is less of a suboptimal strategy in regard to forensic search, as this technology changes at a slower pace; however, it is an absolute mistake when it comes to specialized object recognition, which changes monthly. One can attain much better capabilities by “plugging in” specialty analytics engines into the SAP to address a specific recognition need and “unplugging it” when a better solution evolves. To do so, however, requires adopting the notion of application modules in lieu of and partnership with the solution platform preferences of public safety procurement.

1.4 Commercial Platforms Classes

There are six major types of technology platform categories that are used to present real-time information to law enforcement/public safety. This information incorporates visual images acquired from numerous cameras (both real time and historical); data contained in countless databases/systems; and event triggers provided by a range of sources including 911 calls, sensors, and video analytics. The technologies used to deliver these capabilities are specific to their mission, but also overlap in some capabilities, which complicates what could otherwise be a straightforward correlation of needs to capabilities. These six technologies are:

(a) Video management system (VMS).
(b) Physical security information management (PSIM).

(c) CAD/RMS.
(d) Correlation/data mining/analytics.
(e) Visualization/presentation.
(f) Hybrid SA platform.

1.4.1 VMS

There are three parts to a modern VMS, which are (i) cameras, (ii) recorders/access-managers, and (iii) viewers.

The modern VMS has evolved from its closed-circuit television (CCTV) physical security roots, though many still see the VMS as a simplistic video recording platform for security-related video. One need only look at the solutions from the leading vendors to understand that these VMS platforms offer a wealth of functions that go far towards meeting the needs for real-time information and the deep analysis to support decision-making. However, the single overwhelming limitation of these VMS offerings is that they must manage the connection to the camera, either directly or through some federation process. Unfortunately, the capability to federate VMS platforms is limited to federation within a specific vendor's platform. Thus, Vendor A can only federate with Vendor A, Vendor B strictly with Vendor B, Vendor C exclusively with Vendor C, and so on, but none can federate with others. Because all public/private partnerships entail accessing video content from a multitude of different VMS platforms, the advantages presented by VMS federation cannot be exercised.

The camera's imager captures the scene and streams it to a recorder that is controlled by an access manager. A viewer connects through the access manager to observe the video either live from the camera or through the recorder. A typical VMS has anywhere from three elements to tens of thousands of elements, all of which are connected by IP networks using wired and wireless technologies.

The SAP is also situated on an IP network. Thus, to view a camera's video, it is necessary to have a connection between the SAP's network and the camera's/recorder's network. It is also necessary for the SAP to have a viewer that is compatible with the remote camera's technology and to have the credentials to authorize access to that camera's video content. These elements must be repeated for each camera source participating in these partnerships.

Under almost all scenarios, the SAP is remote from the camera/recorder, necessitating a network connection between the two. The two alternatives for connectivity today are either a private dedicated line or a public Internet connection. Private dedicated lines are rarely used because they are cumbersome to install/maintain and incur a recurring fee. Public Internet connections leverage existing connections but introduce greater security concerns than those presented by the private circuits.

Though greatly oversimplified, the challenge of connecting the SAP's network to the camera's network is that it connects the two networks, thereby exposing each of these networks to the risks which may exist on the other. Today's environments of malware, viruses, ransomware, and general hacking dictate that every network connection presents a risk that must be mitigated. Because networks are bidirectional, the risks must be mitigated on both sides. Because firewalls are the risk mitigation platforms of choice, specialized rules must be written to allow each such connection. Because all such rules effectively open another network access “hole,” there is reluctance on the part of the cybersecurity team to facilitate connections.

Given that without connections there is no video, one best practices goal is to reduce the quantity of such connections and the variability of these connections. One addresses these goals by aspiring to connect to the group of cameras collectively through their access manager (i.e., the VMS) instead of to each individual camera. Also, taking advantage of recent popularity of cloud-based security, when the cameras of interest happen to be associated with a cloud-based security service that also supports multi-tenancy, it is possible to connect to many different groups of cameras utilizing a single connection from the SAP to the cloud security service provider.

The SaaS popularity introduces another emerging trend that furthers the goals of SAPs. Because high-definition video streams consume significant amounts of bandwidth, as occurs with traditional cloud-based recording, new technologies are being deployed to divide the recording burden between on-premise appliances that are local to the cameras and cloud-based storage for event-driven video. Professional configurations for security video define camera recording to occur 24×7 at about 17 frames per second and at a relatively high resolution. These are experience-borne settings which should be adhered to. It is also known that only a small fraction of that recorded stream will be retrieved for review. It is further known that the small portions of retrieved video have characteristics that could be classified as “events” to the camera system. The concept of edge or hybrid recording evolved by merging the desire to conserve bandwidth with the knowledge that one can tag the video most likely to be replayed.

The edge recording technology deploys a small appliance local to a group of cameras that performs as the primary video recorder and as the gateway for that camera group to the cloud-based master controller. In such deployments the cameras send their 24×7 video stream to be recorded on the local appliance and in parallel send any event video (video characterized by embedded analytic software as meeting pre-defined conditions) to the cloud-based controller/recorder. This methodology safeguards the likely video of value in the cloud while delegating the quantitative volume to the local recorder. A viewer can access both video repositories through the cloud in a transparent manner.

It would seem that the edge recording concept would be of no relevance to the SAP, as it has a voyeur relationship with the parent VMS, which is always remote to it. However, innovative adaptations of this technology are providing major benefits to SAP deployments by introducing significant capabilities and removing significant burdens.

The capabilities come in the form of increased analytics, which will be addressed in a following subject category. Burden relief comes from the methodology used by the edge appliances to connect the cameras to the cloud controller. In the traditional approach, the camera generates one to three video streams and connects to the VMS controller/recorder. The SAP, in turn, would connect to either the VMS controller or the camera through a labyrinth of network devices. To view the camera's video, the SAP operator would activate the appropriate viewer on the SAP, which would pull into the viewer the camera's video stream. Should the camera's video stream be selected for active video analytics, the camera's video image would be continually streamed to the analytics engine of the SAP.

The edge recording appliance provides an alternative approach. Instead of defining the SAP's network connection to the local VMS, the edge appliance is introduced into the local environment. The edge appliance connects to the applicable cameras (this is controlled by the camera credentials provided by the participating partner whose cameras are being accessed) and makes use of the reverse dynamic DNS mechanism to establish a connection with the viewer on the SAP. The edge appliance is recording the video streams of the defined cameras and retaining it for a short duration (a few days in general). Should the SAP operator desire access to that camera's content, either in live view or recent replay, they merely invoke the viewer on the SAP. The presence of that edge appliance and the method by which it connects through the Internet to the SAP make it significantly simpler to configure and can be secured much tighter than the traditional methods of allowing external connections to the VMS.

1.4.2 PSIM

The PSIM system concept has its roots in the private security practice. These capabilities fit well into the needs of public safety to access video content from different VMS platforms and to CAD's needs to script response protocols. However, as the quantity of input sources grew, these platforms were found lacking in their ability to apply AI mechanisms to video analytics as well as in their integration with federated search or event/information correlation capabilities present in several data mining and analysis tools used by law enforcement analysts.

The potential of this platform was recognized however and formed the basis for the SAP solutions. Some vendors enhanced the capabilities of their core software, others cobbled together multiple systems into a platform solution, and others reacted to deficiencies in the early integrations and recreated the wheel from ground up.

1.4.3 CAD/RMS: Correlation/Data Mining/Analytics and Visualization/Presentation

Combining the above three platform categories into one discussion, we summarize them as providing enhanced capabilities within a specific mission definition. Though each has some capabilities that encroach on the communities' needs, they do not justify selection as the platform of choice.

CAD/RMS systems, briefly described above, are well understood by stakeholders. Correlation analytics represents multiple data mining platforms. While they are excellent tools, they are very focused on a subset of the overall functional needs. There is a trend afoot of combining these platforms with PSIM platforms to create powerful hybrid SA solutions. The visualization platforms entail technologies that help the visualization of events with a global perspective. These are very popular with large/multinational corporate risk management teams.

1.4.4 Hybrid SA Platform

In the past few years, a new category of use evolved for the traditional PSIM. Public safety and law enforcement defined a need for real-time SA and for event correlation among their technology platforms. Their expectations included the ability to view cameras from numerous sources through a single pane of glass. They expected to be able to enter search criteria once and to launch a federated query across multiple databases using that single entry. They expect complex analytics that run behind the scene to identify objects of interest and behavior anomalies from live and recorded video input across many cameras, as well as provide real-time triggers when an object of interest passes the field of view of a camera.

The response to those needs was "solutions" from several vendors, all of whom employed a hybrid approach. Whether embedded under the label of a single system or as an integration between multiple systems, these solutions strive to aggregate the information, analyze it to correlate possible links, and embed the information with AI to increase the speed and accuracy of assessments.

Each solution excels in certain aspects over others, both within its own platform and as compared to the capabilities of other platforms. Some imbed most functions within their "unified solution," while others integrate specialty functions as "plug-ins" from independent providers.

Given these attitudinal differences, some of these platforms present a bias towards correlations capabilities, while others focus on the unified presentation of accessed information. On the extreme opposite end of this scale, one finds the platforms that retain their footing in the traditional PSIM environment, as exemplified by a select number of vendors.

While raw capabilities present one profile, their portability presents another. Historically, these systems were provided through on-premises dedicated software and hardware solutions. They are now being introduced under the SaaS model, which makes them affordable to a huge range of public safety consumers. Where previously either breadth of function or quantity/capacity were considered to define the cost-size of the platform, the adoption of SaaS provided an opportunity to separate this relationship and to enable service providers to market these elements independently. The current trend is to provide a single solution of broad functionality with quantity defining the price structure. Thus, every client has access to the full suite of functions, but cost is calculated based on the number of integrations in the solution and by the number of users of the solution (users are defined as a combination of physical consoles and individualized login credentials). For smaller jurisdictions, this is the singular element that takes the idea from being an aspirational goal to one of reality.

2 Challenges in Delivering a SAP to Law Enforcement

This section focuses on the practical deployment process of SAP systems into an active operating environment of a law enforcement agency. Challenges associated with delivering a complete, integrated system for law enforcement include all those typically associated with any complex system integration project. However, there are additional areas where the SAP presents specific challenges and introduces potential concerns:

- Reluctance of the stakeholder law enforcement agency to entertain technology-introduction risks.
- Misunderstanding IT's role at the law enforcement agency, from reluctance to engage through maintenance responsibility.
- Misaligning the SAP's role within the agency.
- Underestimating the public's capacity for misperception.
- Setting public/private partnership policy without sufficient regard to the technical complexities.

2.1 Reluctance to Entertain Risk

Risk aversion in public safety's agencies when it comes to the deployment of advanced new technologies is fairly common. The risk aversion, particularly in the procurement process, is the single largest cause for the underwhelming potential that can be seen in deployed SAPs. As previously described, the key functions of the SAP are to connect to the large plethora of content sources available to the agency and to aggregate these individual content sources into a mechanism that provides actionable information within the life of an event. These two functions are accomplished through numerous task-specific application modules. The applications that perform the connection functions are typically viewed as the core of the SAP, while the applications that perform the aggregation/correlation functions are typically viewed as a value-added feature set. Failing to differentiate core components from value-added components leads the planner to categorizing the availability of both as being critical. Because tried and true usage over a long duration is the best predictor for system availability, this singular categorization may inject bias into the evaluation of applications that fall on the correlation side of the SAP's functionality. Whereas core connectivity components need to be reliably available, the value-added correlation components can tolerate outages. While no outage is desirable, the potential of an outage may be warranted based on its risk/reward ratio.

The applications providing the best and most innovative correlation capabilities tend to be based on newly developed leading (bleeding) edge technologies. These applications may not have a demonstrated record of performance, but they offer the promise of significantly enhanced capabilities. As these innovations mature to the degree of demonstrated stability, they are superseded by the next batch of innovative technologies.

What the public safety procurement process occasionally fails to recognize are the consequences of a restrictive approach to new technologies. To protect the SAP operator from loss of functionality resulting from an unstable application, the agency procurement team may preemptively deny the operator access to such new functionality. In the category of these correlation-focused Applications, the differentiating performance criteria are their sophistication, speed, and accuracy. It is not uncommon to see a threefold improvement in capabilities between a generation of applications that are less than a year apart. For SAPs that are goaled to provide actionable information within the life of an event, waiting for demonstrated stability can be the difference between interdicting the person committing the act and documenting the person who committed the act.

2.2 Misunderstanding IT's Role, from Reluctance to Engage Through Maintenance Responsibility

The IT department of an organization may find itself in a challenging position in its interactions with the business units of an organization, as it may be perceived to be laborious in its activities and resistant to new initiatives. While some of the appearances are often correct, the underlying causes are not readily understood. The structure of the IT department differs by organization, and often even among similar organizations, based on how the function is budgeted. In the public sector IT departments create an annual budget, which includes known operational responsibilities and an allowance for expected enhancements. As is true of any departmental budget, the actual approved version falls short from the proposed. On the technology side of the equation, the often-cited adage of IT professionals is that things break when one endeavors to change them and that the environment is much more complicated at the detailed level than at the conceptual level.

As a business unit ventures into this IT environment to address a technology introduction initiative, they may be confronted with some challenges: while the business unit is in anticipation for the prospect of the new system, the IT department may be less enthusiastic. While the business unit has been thinking for some time about the potential of this new system, the IT department perceives it as added work and added responsibilities. A number of things can be done to better align these conflicting perceptions. In principle, the IT department's role is to support the technology-based needs of the business or line units. In practice many IT departments see themselves as an operational unit tasked with the responsibility to support systems as an end in itself. Also, many business units assess the needs of their new systems only from the perspective of the directly attached components, failing to understand the potential impact to the utility services being provided. Frequent omissions include underestimating the increased load on the network, introducing new cybersecurity exposures, not understanding the authentication infrastructure, and introducing different mechanisms for utility functions than those currently in use, for example, network-based storage.

2.3 Misaligning the SAP's Role Within the Agency

The broad capabilities of the SAP often expose it to possible misuse. Misuse on one extreme originates from isolating its functionality to support the narrow mission of a specific entity within the agency (in effect continuing the practice of creating silos that do not communicate with each other). The misuse on the other extreme is to deploy the SAP to serve the broad information acquisition needs of an agency. The overly narrow focus is a tendency within large agencies; such focus does not reduce the capabilities of a SAP as much as it reduces the usefulness of its product. Dilution of mission through a multi-tool acquisition logic is a tendency seen in smaller agencies. The risks of the multi-tool approach, when one tool can serve multiple purposes, is to have the incorrect element be defined as the function leader, thereby impacting the speed with which quality information is produced. The tendencies with multi-tool deployments are to focus on accuracy and completeness at the cost of timeliness. It is useful to reiterate that the goal of the SAP is to generate actionable information within the duration of the event, more so than to provide the complete investigator's report in support of a prosecution.

2.4 Underestimating the Public's Capacity for Misperception

While the goal of the SAP is to collect and aggregate all accessible information to produce timely actionable information to support response activities during the life of the event, there is some vocal public pushback to the availability of information. Some of these misperceptions have their roots in the earlier notions of barring the use of evidence acquired through illegal means or through coercion. It is true that many items of information can be misused, so that by denying access to such information one protects the owners of that information, and one prevents the potential for its misuse. However, SAPs are used to attain information within the life of an event, which is a much narrower focus than intelligence collection, proactive or reactive. Within that context, the goal of the SAP is to harness all the meaningful information of potential relevance to an evolving event that is possible within the timeline of the evolution of that event; SAPs are not goaled to enhance the investigator's toolbox nor as a means to attain general intelligence – SAPs are a tool to support the effectiveness of the response to a triggered event, which may be a 911 call or activation of a monitoring sensor.

Within the context of social media, atypical increases in activity levels may represent a triggering sensor of a potential event of interest. Keyword searches are an enhancement of such triggering capability. SAPs do not monitor conversations to build a database of known associates; they are platforms used to access those preexisting databases efficiently. The unintentional consequence of denying SAP access to social media content is that such action reduces the quality of information available to support the response to an active event, which contributes to the problems of inappropriate response but without addressing the underlying concern about collection of inappropriate information.

2.5 Setting Public/Private Partnership Policy Without Sufficient Regard to the Technical Complexities

Often SAP technology is deployed via public/private partnerships (PPPs). The mechanics for establishing a PPP connection between two entities, where each partner is protected from the risks emanating from the other member, are often labyrinthine. The goal of a PPP is to provide public safety with real-time access to security video resources existing at a private entity's location. In concept, this is a one-way connection to provide public safety with a view of the existing video. In actuality, the current state of technologies cause this to be a bidirectional network connection, which if not constrained provides each member with full access of every member's network assets and problems. Many PPPs for sharing security video content with public safety originated back in the era CCTV security; in that environment, public safety's access to the camera video was as simple as provisioning a telephone circuit connection to the partner's site, attaching a wire to the camera's video splitter and watching the video content flow. The evolution from analog CCTV to the present-day IP-based VMS platforms has facilitated an entirely new era of technological capability; however, this has resulted in an increase in complexity.

It is easy to appreciate the risks that arise from unfettered access and exposure to camera video, both live and recorded. In both multicast and unicast implementations, some form of a logical session needs to be set up between the communicating pairs; to establish communications, a bidirectional exchange of protocol data is needed, and such packet exchange must be enabled while blocking all other types of packet flow. This requirement creates the technical configuration complexity and associated costs; in particular, adequate firewall systems and technology-converting mechanisms need to be deployed to limit the content flow between the partnering entities to just the camera video. Given that neither party in many of these partnerships is in the position to fully trust the other, these firewall setups tend to be deployed in pairs, with each partner administering their own. To make this arrangement viable, both of the firewall pairs must be configured in a compatible manner; the configuration process, however, can be error-prone and inefficient. Recognizing that the complexity of the technology challenges cannot be simplified or avoided, new initiatives should endeavor to reduce their quantity. One way to do this is by taking advantage of the cloud-based video security trend.

The popularity of this cloud-oriented trend, especially for smaller enterprises, presents the opportunity to attain many camera views through a single connection. Two features of cloud-based video security present the enabling capabilities for simplifying the complexity and reducing the quantity of those complex connections. True to its name, in cloud-based video security solutions, the camera access controller resides in the cloud.

In the modern IP-based VMS environment, the connections between cameras, recorder, and viewers are not direct but go through the intermediary element, the camera access controller. To view a video stream, a viewer connects/authenticates to the access controller and is passed to the video stream from either directly at the camera or through a media server or through the video recorder associated with that camera. These components tend to be physically separate devices and can be placed anywhere along the network path. When that controller is located at the partner's site, the network path from the viewer must extend to the partner's site. When the controller is in the cloud, the network path from the viewer must extend only to the cloud, with the controller's path to the camera/recorder used to facilitate the second leg of the connection to carry the video from the participant's site. Because most of the convoluted complexity exists at that link to the participant's site, by limiting the extent of the logical connection to only the cloud camera access controller, one greatly simplifies the configuration task – the connection from the SAP viewer to the cloud still needs to be secured, but that presents a much more structured requirement than extending that connection to the partner's site. In addition, many of the vendors use what is known as multi-tenant capability. Multi-tenant originated with residential security services that used a central monitoring service to manage many small instances of security systems. In the new cloud-based VMS service, this same technical capability provides the means for a single public safety SAP to access numerous partner sites that happen to use the same cloud-based service provider. While each participant is screened from all others, the public safety connection can access all permitted cameras.

3 Ethical Questions

This section just scratches the surface on the matter of SA ethics.

Recently a number of scientists and human-rights activists have raised ethical concerns about a number of AI-based technologies, including face recognition technology in particular. A relatively large research investment of the computer-vision community has focused in recent years on advancing and improving facial-recognition methods. An increasing number of scientists are now urging researchers to avoid working with firms or universities linked to unethical projects, to re-evaluate how they collect and distribute facial-recognition data sets, and to rethink the ethics of their own studies. Some institutions are already taking steps in this direction, and several journals and an academic conference have announced that they would undertake additional ethics checks on studies on these topics [15]. There are acceptable applications of face and biometric recognition: technology that recognizes or analyzes human faces has a number of useful applications, such as to assist in the process of finding lost children; of tracking criminals and terrorists by law enforcement agencies; of accessing smartphones and cash machines; and, in telemedicine, to help diagnose or remotely track consenting participants. But there must be an ethical recognition that technologies that can remotely identify or classify people without their knowledge is fundamentally perilous.

For facial-recognition algorithms to work effectively, they must be trained and tested on large data sets of images, ideally captured many times under different lighting conditions and at different angles. The issue is that it is difficult to train accurate facial-recognition algorithms without vast data sets of photos. Earlier in time, researchers used volunteers to pose for these photos, but now most researchers collect facial images without obtaining any consent, e.g., taken at university campuses or taken from the image-sharing site Flickr. Some of these data sets have then been used by companies that worked on military projects in China, and large online image collections such as MSCeleb are still distributed among researchers who continue to use them [15, 16]. In many states in the USA (but not all), it is illegal for commercial firms to use a person's biometric data without their consent, and some have argued that the European Union's General Data Protection Regulation (GDPR) also precludes researchers from doing so (although additional clarifications of these regulations are still outstanding as of press-time). Recently Facebook agreed to settle a class-action lawsuit related to the use of non-public photographic data for facial recognition; other companies have also been sued including IBM, Google, Microsoft, and Amazon.

A particular controversy deals with facial-recognition research on vulnerable populations (e.g., refugees or minority groups). For example, a university recently announced (and then retracted) that they had developed facial-recognition software "capable of predicting whether someone is likely going to be a criminal," with "80 percent accuracy and no racial bias." To get a sense of academic views on facial-recognition ethics, *Nature* surveyed 480 researchers in 2020 who have published papers on facial recognition, AI, and computer science. When asked for their opinions on studies that apply facial-recognition methods to recognize or predict personal characteristics (such as gender, sexual identity, age, or ethnicity) from appearance, around two-thirds said that such studies should be done only with the informed consent of those whose faces were used [15].

It is expected that ethical considerations will continue to be an issue of grave concern in the future as all these technologies become more powerful and more sophisticated.

4 Research Directions and Expected Developments

The disciplines discussed in this chapter, SA, computer vision, AI/ML/DL, data management, data fusion, and IoT (to list just a few), are currently enjoying a high degree of academic research, with the expectation that such research will continue for the foreseeable future. Some of the areas of research include but are not limited to:

- Research in computer vision, including information extracted from a digital image, image segmentation, and classification; image transformation using GANs (generative adversarial networks); learning and inference; efficient video streaming; pattern recognition; mathematical modeling; modeling and capturing 3D shape and motion of people; image-based rendering of faces; detection of face morphing; tracking of deformable objects and non-rigid structure; optical flow estimation; shape analysis; scene flow estimation; deblurring; multiresolution methods; and augmented reality.
- Research in SA including algorithms for perceiving, understanding, and predicting; exploiting use of IoT data for SA applications; situation-aware network/context-aware network; social network/social media intelligence; web analytics and incident response; sensor fusion; and data correlation.

5 Conclusion

This two-chapter set reviewed practical aspects of SA technologies when used for public safety and law enforcement applications, to provide real-time visibility on the status of prospective threats – and possible targets. SA for law enforcement agencies include, among others, the following: in-field policing with body-worn cameras, license plate recognition, object leave behind detection, crowd formation, aberrant behavior, weapons detection/armed intruder, and tracking a target from multiple cameras' views not just a single camera. As discussed in this chapter, the use of body-worn cameras by police officers is experiencing widespread deployment. The large amount of data generated by these cameras as well as by the fairly dense set of stationary cameras found in smart cities, and other sensors, gives rise to the need for advanced SAPs that can undertake advanced analytics in real time or near real time, likely using ML and DL techniques; many of the systems are now cloud-based.

References

1. A. Braga, J.R. Coldren Jr., et al., The benefits of body-worn cameras: New findings from a randomized controlled trial at the Las Vegas Metropolitan Police Department, in *Final Report to the National Institute of Justice*, (September 2017)

2. NCSL (National Conference of State Legislatures), *Body-Worn Camara Laws Database*, 2/28/2018. Available online on December 1, 2020 at https://www.ncsl.org/research/civil-and-criminal-justice/body-worn-cameras-interactive-graphic.aspx#/
3. D. Yokum, A. Ravishanakar, A. Coppock, *Evaluating the Effects of Police Body Worn Cameras: A Randomized Controlled Trial* (The LAB @ DC, 20 Oct 2017)
4. D. McClure, N. La Vigne, et al., *How Body Cameras Affect Community Members' Perceptions of Police* (Urban Institute, August 2017)
5. Citizen Perceptions of Body-Worn Cameras: A Randomized Controlled Trial. Police Executive Research Forum, April 2017
6. J.E. Gaub, N. Todak, et al., Beyond patrol: Exploring the perceptions of police body-worn cameras among officers in specialized unit, in *Arizona State University Center for Violence Prevention and Community Safety*, (February 2017)
7. B. Dickson, What is computer vision? PC Magazine, 9 February 2020. Available online on December 1 at https://www.pcmag.com/news/what-is-computer-vision
8. S. Khan, H. Rahmani, et al., A guide to convolutional neural networks for computer vision, in *Synthesis Lectures on Computer Vision*, (February 2018), 207 pp. https://doi.org/10.2200/S00822ED1V01Y201712COV015
9. S. Liu, Z. Liu, Multi-channel CNN-based object detection for enhanced situation awareness, in *Sensors & Electronics Technology (SET) panel Symposium SET-241 on 9th NATO Military Sensing Symposium* (30 Nov 2017)
10. S. Liu, H. Liu, et al., Enhanced situation awareness through CNN-based deep multi-modal image fusion. Opt. Eng. **59** (5), 053103 (19 May 2020). https://doi.org/10.1117/1.OE.59.5.053103
11. J.A.D. Cameron, P. Savoie, et al., Design considerations for the processing system of a CNN-based automated surveillance system. Expert Syst. Appl. **136**(1), 105–114 (December 2019)
12. Q. Zhu, Research on road traffic situation awareness system based on image big data.IEEE Intell. Syst. **35** (1), 18–26 (1 Jan.-Feb. 2020). https://doi.org/10.1109/MIS.2019.2942836
13. E. Maltezos, L. Karagiannidis, et al., Preliminary design of a multipurpose UAV situational awareness platform based on novel computer vision and machine learning techniques, in *2020 5th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM)*, (Corfu, Greece, 2020), pp. 1–8. https://doi.org/10.1109/SEEDA-CECNSM49515.2020.9221786
14. Police1 Staff, How to buy computer-aided dispatch (CAD) systems and records management systems (RMS), Whitepaper, 5 Nov 2020. Available online on January 5, 2021 at https://www.police1.com/police-products/police-technology/software/cad/articles/how-to-buy-computer-aided-dispatch-systems-and-records-management-systems-ebook-Vs39ElTpAfRWcNxl/
15. R. Van Noorden, The ethical questions that haunt facial-recognition research. Nature, 18 November 2020. Available online on January 5, 2021 at https://www.nature.com/articles/d41586-020-03187-3
16. M. Murgia, Who's using your face? The ugly truth about facial recognition. Financial Times, 18 September 2019. Available online on January 5, 2021 at https://www.ft.com/content/cf19b956-60a2-11e9-b285-3acd5d43599e

Daniel Minoli, Principal Consultant, DVI Communications, New York, graduate of New York University, has published 60 technical books, 350 papers and made 90 conference presentations. He has many years of technical and managerial experience in planning, designing, deploying, and operating secure IP/IPv6-, VoIP-, telecom-, wireless-, satellite- and video networks for carriers and financial companies. He has published and lectured in the area of M2M/IoT, network security, satellite systems, wireless networks, IP/IPv6/Metro Ethernet and has taught at New York University, Stevens Institute of Technology, and Rutgers University. He also has served as a testifying expert witness.

Andrzej Koltun, Director of Network & Systems Architecture, DVI Communications, has technical responsibility for many of DVI's projects, including aspects of: strategic planning, network and system design, and network implementation management. Mr. Koltun has designed and deployed large-scale Networks for over 35 years supporting multiple missions–critical applications including Financial services, security and telemetry. Mr. Koltun possesses an extensive background in real-time video surveillance applications using IP.

Benedict Occhiogrosso, is a Co-Founder of DVI Communications, New York. He is a graduate of New York University Polytechnic School of Engineering. His experience encompasses technical and managerial disciplines. He also served as a testifying expert witness in numerous cases encompassing patent infringement, and other legal matters.

A Wireless Sensor Architecture for Efficient Water Quality Measurement and Monitoring Using the IoT

N. Kavitha and P. Madhumathy

1 Introduction

The surface groundwater quality has been influenced by humans mainly and by nature. The natural factors that influence a system are dust and salt deposition by wind, evapotranspiration by atmospheric processes, organic matter and nutrients naturally leaching in soil, biological factors, and hydrological factors. These factors change the chemical composition of water deciding the water quality without human influence [1]. Globally eutrophication is one of the most prevalent water quality problems due to high-nutrient loads. Microbial pollution is the main source of domestic sewage; fossil fuel burning and bush fires are atmospheric inputs and industrial effluents that pollute the water [2]. Polluted water has many adverse effects such as unsafe drinking water and drastic reduction in the amount of useable water due to the pollutants in useable water within a given area.

Water is one of the basic requirements in our everyday life, and 70% of the earth is covered by water, but still there is an acute shortage of water for living beings [3, 4]. Water pollution is one of the primary reasons behind the acute shortage of water. There are many regions in India known for natural resources, Karnataka being one among them in the south part of India. The main reason for knowing it as the main source of natural resources is because it includes canals, backwaters, rivers, lakes, oceans, wells, etc. [5–7]. Here only 29.5% of rural households have access to safe drinking water as per the survey conducted by the Central Pollution

N. Kavitha
Dayananda Sagar Academy of Technology and Management, Bangalore, India

P. Madhumathy (✉)
RV Institute of Technology and Management, Bangalore, India

Dayananda Sagar Academy of Technology and Management, Bangalore, India

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_10

Control Board (CPCB). The major sources of water contamination in Karnataka are the waste being disposed of in rivers passing near temples, churches, mosques, and agricultural areas [8, 9]. For continuous monitoring of water pollution for a long time duration, a wireless sensor network [10] has been considered since it is one of the efficient technology in continuous monitoring. Several aspects of the environmental condition are monitored by wireless sensor networks. A wide area network is advantageous because it consumes low power by continuous monitoring through wireless sensor nodes. It is economical and feasible in event tracking in the Kaveri River which is one of the most polluted holy rivers of Karnataka and Tamil Nadu [11]. The rivers are contaminated majorly due to the dilution of soapy water into the river and unfiltered water waste from temples and nearby hotels. It is presumed that river water is a holy water, and it is used in the daily routine of millions of people [12, 13]. Hence, there is an increased interest in developing a system which continuously monitors change in water contamination level and warns the citizens in real time [14, 15]. To accomplish this task, the research work designs a wireless sensor network that monitors the water quality in natural water bodies in a real-time scenario.

A system that has an inbuilt function of a mechanical or electrical system is often used for real-time computation. It forms an absolute device embedded with hardware and mechanical parts.

Worldwide maximum percentage of microprocessors produced as embedded systems components. When a typical embedded system is compared with general-purpose counterparts, it is observed that it consumes less power, can be miniaturized, and has a wide operating range and less per-unit cost; it is economical coming under within the bounded processing resources. However, an intelligent mechanism is built on top of the hardware comprising of embedded units and existing sensors which can optimally manage available resources at the network unit and provides elevated results [16]. Henceforth, embedded systems power consumption has been managed by designing intelligent techniques. As we are aware, most of the existing embedded systems are mainly based on ordinary microprocessors, wherein it has memory and peripheral interface circuits as the external unit, and on microcontrollers, wherein it has memory and peripheral interface circuits as the internal unit which is one of the essential units in a complicated system [17]. In several cases, the processor used for certain calculations would range from common purpose to specific level in both the cases and also somewhere even custom designed for application level.

A usual level of class for fanatical processors is the digital signal processor (DSP). We know that embedded systems are designed to perform a specific task and miniaturization of the system [18] is carried out by design engineers which in turn yields cost reduction and consistency in the performance growth. Embedded systems range stretches from small portable devices to a large complex system inclusive of small devices, like MP3 players and digital watches, and to the large complex systems in avionics and automobiles, such as hybrid vehicles; in the medical field, it includes magnetic resonance imaging (MRI). It has been observed that the level of intricacy changes from minimal level to peak level, wherein the minimal level is achieved by a single microcontroller chip and the peak level is achieved by integrating several peripheral units and networks [19] inside a framework.

2 Literature Survey

Mamun [1] proposed a water quality monitoring system implemented in Fijian locations. Solar-based renewable energy source (REs) is used to interface monitoring system with GIS. Here, different parameters are quantified such as temperature, potential of hydrogen (pH), oxidation reduction potential (ORP), and conductivity. Finally, the obtained data were tested and analyzed using statistical methods and verified comparing with the FNDWQS. The findings demonstrated that the system is capable of delivering an accurate and consistent measurement of water quality in real time. The system model efficiency decreases in other locations.

Alexander [2] proposed a water quality monitoring system applied to remote rivers, lakes, and coastal areas. It mainly detects water temperature, dissolved oxygen, and pH in a pre-programmed time interval, consisting of off-the-shelf electrochemical sensors, a microcontroller, a wireless communication system, and the customized buoy. The developed prototype disseminates the gathered information in graphical and tabular formats through a customized web-based portal and preregistered mobile phones to better serve relevant end users. The system lags in constant monitoring for implementation in large lakes.

Viani [3] has proposed a methodology for irrigation management in agriculture using fuzzy logic. Here it provides context awareness through numerical soil output and crop models which are one of the irrigation best practices through fuzzy rule set; in addition to that, it also provides enhanced irrigation schedule. This in turn enhances the crop yield in accordance with the weather conditions which will also downgrade the water wastage. Here, wireless decision support system is embedded in the network gateway. The experimental results are being validated by establishing experimental setup in the north of Italy which showed a significant fall in water wastage compared to the existing methods on the basis of threshold level and an improved manipulation in water tanks used for irrigation which in turn reduces the percolation phenomenon which is not showing any impact on crop quality.

Unnikrishna Menon [5] has proposed wireless sensor network-based system for monitoring the river water quality, which assists in remote monitoring of the water quality data in India continuously. Here, the pH level of water has been monitored by designing a system with wireless sensor node. There are many modules such as wireless communication module, signal conditioning module, processing module, and power module being embedded in the proposed sensor node. Through Zigbee communication mode, the pH value information that has been sensed will be transmitted to the base station through wireless mode after the processing techniques such as signal conditioning. Here, the circuit has been designed for the sensor node, and a hardware prototype is created using the appropriate components for simulation which results in a cost-effective platform and less power consumption for water quality monitoring.

Rohini Krishna [4] has proposed a methodology for monitoring the water quality of the Pampa River in Kerala. The river is one of the most contaminated rivers in Kerala. Based on few chemical parameters, we determine the quality of water in the Pampa River, Chengannur segment. The chemical parameters include nitrate, phosphate, dissolved oxygen, and pH. Here the experiment has been carried out by taking the mean with the standard deviation followed by considering three study sites. Further, analysis of variance (ANOVA) was conducted in two different ways. In addition to that, they have also calculated the water quality index (WqI), which showed chemical parameter levels in pre-monsoon and summer, wherein the biochemical oxygen demand (BOD) content is 81 and 59, pH content is 68 and 80, dissolved oxygen content is 5 and 96, nitrate content is 99 and 96, and phosphate content is 31 and 43. These experimental results inferred that the water quality of the Pampa River is moderate.

João Matos [6] has proposed a technique for air and water quality measurement using sensing module which is built on multichannel and Raspberry Pi platform, wherein continuous monitoring of water and air quality is carried out. To implement it they have used underwater acoustic signals for measuring relative humidity, temperature, conductivity, and gas concentration. The measured data are being stored in the drone's computational platform memory which is also being coordinated with a remote server database. Here on the server side, advanced data processing algorithms are implemented. In addition to that, for data visualization and statistical analysis a mobile application was also designed so that it would be helpful for the people whose area of interest is in the same field.

Hari Kumar [7] has proposed a methodology for traditional methods of aquaculture and hydroponics for groundbreaking food production known as aquaponics; using this they were able to develop a single integrated system for growing both crops and fish in a single system. Here, the essential nutrients for the plants are provided through aquaponics. In turn symbiotic relationship is maintained, wherein for fish a biofilter is served by plants. For constructing a global infrastructure, an open standard of WSN is used in the system. Sensor devices are used for sensing the water quality in various dimensions by designing an aquaponic system, and finally in the cloud database, data is being stored. It has been observed that compared to the traditional methods of aquaponics, this methodology shows there is an automatic database updation with very less human intervention. Additionally in their aquaponics system, they have used next-gen telco technologies which have high bandwidth and low latency resulting in automatic detection of infected fishes in the right time which in turn maintains balance in aquaponics ecosystem. Here, wireless sensor network (WSN) and next-gen telco are used for providing "connected aquaponics" between the system and also provide organic sustainable food to the world community and increase the crop yield.

3 Motivation

One of the global problems that exist for decades is water scarcity in addition to water pollution, which majorly requires increased interest in developing a guiding principle for water resource available at international levels to individual wells. It is verified that the major cause of disease spreading worldwide is mainly water pollution. In the survey it was found that everyday nearly 10000 people die worldwide, among them 500+ people die daily because of the problems related to water pollution. The extensive research has reported that the amount of useful water level will reduce drastically to the minimum level after few years. In most places the water that is used for drinking is contaminated which is not safe for drinking and results in major health problems. The major reason is that the public is unaware about the extremities caused due to this water pollution, and there is no water quality monitoring system, which is one of the serious global issues. In addition to that, earthquakes, volcanoes, and algae tints are also naturally affecting the ecological status of water. There are two extremities: one extreme was lots of inventions, and the other extreme is pollution, global warming, and so on. Due to these extremities, safe drinking water is very minimal due to pollution. Presently real-time water quality monitoring faces major challenges due to limited water resources that resulted from global warming, growing population, etc. Hence, there is an immediate requirement in developing a system with improved methodologies for monitoring the real-time water quality parameters.

4 Block Diagram

The water quality is measured based on pH concentration of hydrogen ions which indicates whether the water is towards acidic concentration mode or alkaline concentration mode. Generally, pH concentration of pure water is 7; if pH concentration is less than 7, then it is more towards the acidic type, and if pH concentration is more than 7, it is towards the alkaline type. The pH range for drinking purpose is 6.5–8.5pH. The major suspended particles in water that are invisible are measured by turbidity. Increase in turbidity level has an adverse effect on human health, i.e., there will be more chance of getting diarrhea and cholera when the turbidity level is higher. If turbidity level in water is lower, then we say that water is safe for drinking. The hotness or coldness of water is measured by a temperature sensor. The flow of water is measured through flow sensor. In the existing system, water samples are collected manually from various locations in most of the traditional methods.

The water purity is checked through pH sensors. Depending upon the pH level, our system will control the valve for pH level changer tanker. If the pH level changer fails, the system will switch on the valve for another valve to empty the tank. This system also checks the water level; if the water tank reaches the maximum level, then the sensor control will stop the water motor. The IoT system monitors the pH and the temperature level of water. This system only indicates the purity of the water and displays the value in LCD and sends the message. The purity of the water is checked through turbidity sensor. This system will only check the purity and doesn't control any motor or flow. To overcome these disadvantages, a system has been proposed which provides an environment for safe drinking water by doing several checks on the water that has been polluted. Depending upon the pH level, our system will control the valve for changing the pH level in the tank. If the modification of the pH level fails, then the system will switch on to another valve to empty the tank. The IoT system monitors the pH and the temperature level of water.

4.1 Monitoring System

Liquid Crystal Display It is an electronic visual flat panel display abbreviated as LCD where liquid crystals have the properties of not emitting light directly and light modulation; all these properties are used for video display, such as a general-purpose computer display for displaying arbitrary images or stationary images like digital clock seven-segment displays. Concerning the display part, both stationary images and other displays will be using the same basic technology. The only difference is that a large number of small pixels are used in the case of stationary images, whereas larger elements are used for other displays. LCD is economical as well as minute display element. It is flexible and reliable due to the embedded controller, allowing it to be easily interfaced with a microcontroller. This is one of the universal standard controllers inclusive of Arduino with respect to many displays such as HD 44780 because of the inbuilt libraries, which uses a single line of code to display messages.

In many applications, LCDs used such as instrument panels, aircraft cockpit displays and in our daily routine such as computer monitors, televisions, etc. They are very commonly used in watches, calculators, telephones, and video players, which have completely substituted cathode ray tube (CRT) displays with LCD in many applications. These are available in wider screen sizes than plasma displays and CRT, and they do not undergo image burn-in because phosphors are used and are susceptible to image persistence.

Temperature Sensor (LM35) LM35 is an integrated circuit sensor where the ratio of electrical output and temperature in °C is considered for measuring the temperature. This temperature sensor is mainly preferred against thermocouples because it generates a higher output voltage and does not require any additional amplification of output voltage.

In LM35 series, the output is voltage, which is directly proportional to the Celsius temperature. It is one of the temperature sensors with an integrated circuit. One of the major advantages of getting directly Celsius temperature compared to other sensor where linear temperature obtained in the form of ° Kelvin we can directly know the temperature from centigrade whereas in ° Kelvin a large constant voltage needs to be subtracted from obtained output voltage for scaling it into Centigrade. In addition to that, compared to other temperature sensors, here we need not do any calibration externally for a wide temperature range of −55 to +150 °C. In LM35 we have inbuilt calibration allowing easy control of the circuitry; it has a very low output impedance and produces linear output. Only 60 μA current is drawn from its supply, and it has less than 0.1 °C of self-heating even in air.

Water Level Sensors To detect the level of substance flow, we use this water level sensor. Here, the substance flow not only includes water; it comprises granular materials, liquids, slurries, and powders. Whatever the level detection or the measurements can take place in river or Lake Place or it can also be inside containers. Mainly these measurements will provide us the information about water flow, or it will give us the amount of materials that are present inside a closed container.

IoT IoT is the abbreviation of the Internet of Things which is one of the major trending topics in today's world. Kevin Ashton's groundwork has led to the Internet of Things (IoT) in today's world at MIT's AutoID labs. Ashton examined the ways through which business can be improved by connecting RFID information to the Internet. The concept which was thought seems to be very simple but it is most powerful because daily life equipped objects if they are connected through wireless means these devices can be managed by computers when they communicate among each other. Ashton wrote an article in RFID Journal stating that: "If we had computers that knew everything there was to know about things—using data they gathered without any help from us—we would be able to track and count everything, and greatly reduce waste, loss and cost. We would know when things needed replacing, repairing or recalling, and whether they were fresh or past their best. We need to empower computers with their own means of gathering information, so they can see, hear and smell the world for themselves, in all its random glory. RFID and sensor technology enable computers to observe, identify and understand the world—without the limitations of human-entered data." At this time, the major challenges are how everything can be connected on the planet, on what type of wireless technology devices could be built, and how existing Internet infrastructure could be modified to make billions of new devices to communicate in a cost-effective manner during early 1999. Today, for most of these questions, a solution has been found, and major obstacles have been overcome. Nowadays for designing a wireless radios cost effectiveness and miniaturization is been majorly taken care. Billions of devices need to establish communications between devices using IPv6 addressing. Mainly cellular wireless connectivity and Wi-Fi connectivity have been established among a wide range of devices. The high broadband speeds offered by numerous networks result in significant increase in coverage of mobile data. Many devices have inbuilt solar recharging which has slightly improved battery technology. In the future, there will be a network that connects billions of objects. Through wired or wireless Internet, numerous items such as sensors in the physical world will be linked to transmit information to the users directly through IoT technology. To transmit the information between devices and users, we can use Bluetooth, RFID, NFC, and Zigbee types of local area connections. These sensors not only have local area connections but also have connectivity such as 3G, LTE, GSM, and GPRS, which are of wide area connectivity. To connect living things and nonliving things, the IoT is used, as well as in interconnecting industrial equipment. The wide scope of improvement in today's IoT vision resulted in interconnecting everyday objects with industrial equipment. The range of connection exists between living and nonliving things, wherein living things such as animals, people and plants further nonliving things such as gas turbines to automobiles to utility meters. Due to the advancement in IoT, the definition has been slightly modified to a higher level – as IoT to the Internet of Everything (IoE) – where everything corresponds to objects, people, places, and things. It is mainly specifying that sensor can be attached to anything; we can participate in the new connected ecosystems if connectivity is established, where data is collected through sensors. Any physical objects that have been used with IoT to collect information will be performing many tasks such as identifying temperature, location, motion, vibration, etc.; to get all these data, multiple sensors will be connected to each data we get. Mainly these sensors will get present information from data feeds based on which people will get new information content by interconnecting sensors together.

This Internet of Things is covering various primary sectors in the economy field by facilitating towards subsequent life-enhancing services. The evolution of smart connected devices has been encouraged by mobile networks due to the rise in Internet of Things; it mainly encompasses to give flawless connectivity, which will yield in keeping opportunity doors wide open, and also productivity for the industries will be boosted, and also for consumers life-enhancing services will be provided. Figs. 1, 2, 3 and 4 shows the major 13 industry sectors tha have adopted IoT services.

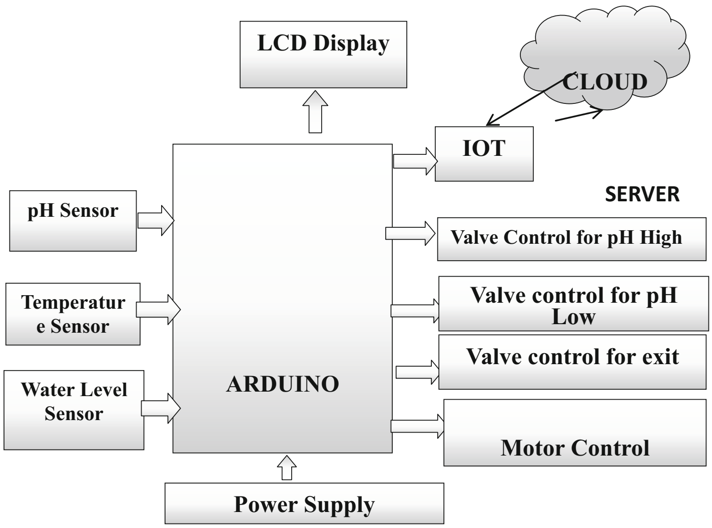


Fig. 1 Block diagram of the working system

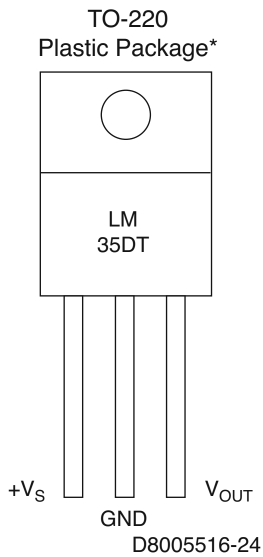


Fig. 2 PIN diagram of the temperature sensor

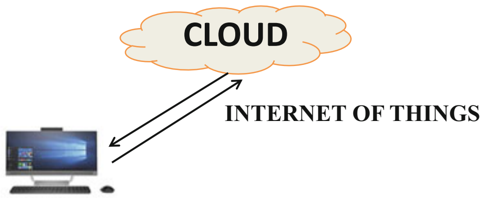


Fig. 3 Monitoring system

Fig. 4 IoT ecosystem

Water Pump Motor In the early days, to get one bucket of water, scooping was done on one hand which was a very tedious procedure, so practically to make it easier, motor was used to pump water upwards. Using a water pump was helpful for irrigation, washing, removing water from an undesirable location, and moving fresh source of water to the desired location. Irrespective of the water requirement, it is tedious and extremely energy demanding to pump water for various requirements of water consumption. Most of the processes are intransitive and helpful when water comes from a higher height or by using plumbing system.

The ancient concept of channelizing the elevated range of water for longer and far distances is one of the articulate advantages. Due to this articulate advantage, far distance water movement is possible, and by spending small portions of the energy, a larger component of its potential energy is retained. Fresh water at a higher elevation is a major source for articulate system.

5 Circuit Diagram

The main source of electrical power is power supply. Power supply unit is mainly composed of electrical supplies or any energy suppliers to an output load or group of loads. The term power supply is mainly related to electrical energy supplies, rarely to mechanical ones and to others.

There are mainly two categories of power supplies; they are linear supply and switched-mode power supply. Simple design exists in linear supply, which makes high-current devices to be heavier and bulkier which results in lower efficiency, whereas switched-mode supply design is more complex making lighter linear supply and is more efficient (Fig. 5).

There are mainly two types of voltage regulators: one is fixed voltage regulator of the range 5V, 12V, and 15V, and the other one is variable output voltage regulator. Many other voltage regulators are also available based on passing the maximum current, and for dual supplies negative voltage regulators are used. Automatic protection exists in most of the regulators due to excessive current and overheating.

Fixed output voltages are available in LM78XX series, which is a three-terminal regulator used in a wide range of applications. Some of the applications such as removing the distribution problems associated with single point regulation, card regulation. Most of the solid-state electronic equipment uses the available voltages

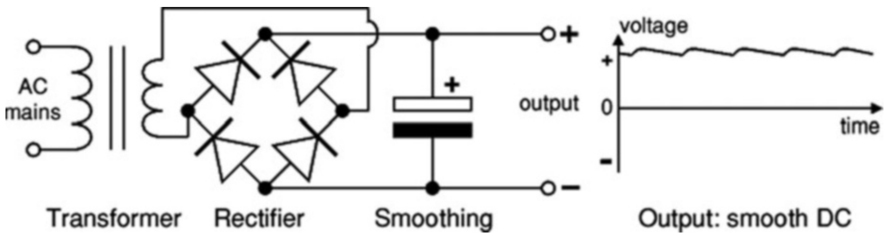


Fig. 5 Power supply circuit

in logic systems, instrumentation, etc. To get variable voltages and current fixed voltage regulators primarily designed.

6 Hardware and Software Used

The hardware we use here is Atmel 8-bit RISC-based microcontroller that has the following features: a flash memory of 32KB with read-write capabilities, 23 general-purpose I/O lines, 1KB EEPROM, 2KB SRAM, 32 general-purpose working registers, 3 timer/counters, internal and external interrupts signals, SPI serial port, 6-channel 10-bit A/D converter, and 5 software selectable power saving modes. The operating range of microcontroller is between 1.8 and 5.5 volts. Single clock cycle is used to execute complex instruction set; 1 MIPS per MHz throughputs has been achieved by the device; it has a higher processing speed and a balanced power consumption. The Atmel megaAVR series has single-chip microcontroller called ATmega328.

The Atmel AVR series is an 8-bit RISC-based microcontroller which combines 32 KB ISP flash memory with read-while-write capabilities. In recent days, ATmega328 is most commonly used in most of the autonomous systems because it is economical and because of its simplicity and low-power consumption. Most commonly, this chip is implemented in Arduino board such as Arduino Uno and Arduino Nano models. Perhaps we have faster data transfer rates and more memory being present in Arduino Uno R3 an ATmega16U2 which in turn does not require any drivers for Linux or Mac and has the capability for having Uno such as a keyboard, mouse, joystick, etc. Few can be interchangeably used but are functionally dependent to an Arduino. In many school-level education applications, basic Arduino will be enhanced and used by using additional output drivers for making small robots, whereas other variants use various processors which are compatible and electrically equivalent to changing factors.

6.1 Arduino IDE

Arduino is one of the open-source cross-platforms which is used for building digital devices by designing such microcontroller kits which can even sense the object as well as control the objects in the physical world. The designed products are licensed and categorized into two categories, namely, GNU Lesser General Public License (LGPL) and the GNU General Public License (GPL), which helps to distribute software as well as Arduino boards to anyone by the manufacturer. The programming languages C and C++ are used to program microcontrollers in which the dialect of features is used. Arduino not only provides tool chains for compiling; it also provides an integrated development environment (IDE) by keeping language project processing as the basis. Initially this Arduino project was introduced in Ivrea, Italy, which mainly aimed in providing platform to create devices at a lower cost which is reliable and uses sensors and actuators to interact with their environment such as motion detectors, simple robots, etc.

The main functionality of Arduino IDE is to generate text file in the encoded hexadecimal format converted from the executable code; the Arduino board will be loaded with hexadecimal encoding in the board's firmware by a program loader. The Arduino IDE programmer performs mainly two functions, namely, setup and loop.

In the setup phase, initialization takes place where input and output pin modes, variables, and other libraries needed in the sketch, after reset when a sketch is started function once.

In the loop phase, Arduino board is controlled until the board is reset, and it also executes the function loop repeatedly in the main program after initialization is performed in the setup phase. For the convenience of testing, the program functions in most of the Arduino boards between pin 13 and ground, and a load resistor is connected, and it also contains a light-emitting diode (LED).

6.2 Proteus

Though assembly level language is understood by microcontroller, it has to be compiled into a language of binary form. To write a program for the microcontroller, we follow a set of rules called assembly language program, and to convert assembly language statements into the language of binary form, a program on a personal computer is used called assembler do not have the same meaning. This assembler called machine code is a compiled program. In machine code a 14-bit array of zeros and ones is used to represent the same command which is understandable by the microcontroller. There is an "executive file," i.e., "HEX data file," which stores the compiled program. The data file is called by the name HEX due to its hexadecimal representation like "probe.hex." Programmers used to load the data file into the microcontroller after generating the HEX data file. Assembly language programs written in text editor ASCII data files will be created either on a hard disc or in MPLAB which is one of the specialized work environments.

6.3 Embedded C

In our day-to-day life, we use many varieties of embedded system in our daily routine such as mobile phone, digital camera, washing machine, etc. where all these embedded devices have inbuilt functioning processor associated with embedded software to perform a specific task. Each embedded system is composed of two main things, namely, hardware and software, where the body of an embedded system is hardware, the brain is the embedded processor, and the soul is the embedded software. Here, the soul of an embedded system, i.e., embedded software, plays a vital role in regulating the embedded systems function.

During the early years, assembly level languages were used to program microprocessor-based systems, and EPROMs were fused into it. To observe the working of the program, there was no separate mechanism; so to ensure accurate execution of the program, they used to make use LEDs, switches, etc. Alternatively in-circuit simulators (ICEs) were also used by few developers but this was unreliable as well as costly. Hence, embedded C was introduced which has to be downloaded to microcontrollers/microprocessors to run the application and to create files compilers are required. In desktop computer applications, access to all resources has not been provided, but access to all resources has been provided by embedded compilers.

7 Results and Discussion

For a successful construction and implementation of a water quality monitoring system using wireless sensors. The implementation is appropriate for ecosystem monitoring, environment monitoring, etc. The monitored information can be accessed from anywhere in the world. This system endows safe drinking water by performing close monitoring on the source of pollutant for water resources, so that we can ensure that all human beings are being provided with nontoxic drinking water.

7.1 Circuit Diagram

Most of the voltage regulator ICs are very much similar to power transistors like 7805 having three leads with a +5V 1A regulator. There is a provision for attaching a heat sink if required by drilling a hole. Usually to convert the voltage from the mains to a lower voltage, a transformer is used in an AC-powered linear power supply. A rectifier is used to produce DC; further, to get pulsating DC, a capacitor is used in the rectifier. After using a capacitor, we don't get a complete smooth direct current; still, we will be having a component called as ripple. These pulsating DCs having AC components will be ranging from a multiple of 50 or 60 Hz (Fig. 6).

7.2 Proteus Result

Labcenter Electronics has developed design software called Proteus for electronic circuit simulation, PCB design, and schematic capture. Most of the electronics engineers opt for this design software because of its user-friendly nature and its simplicity. When any digital simulations are to be carried out using microprocessors

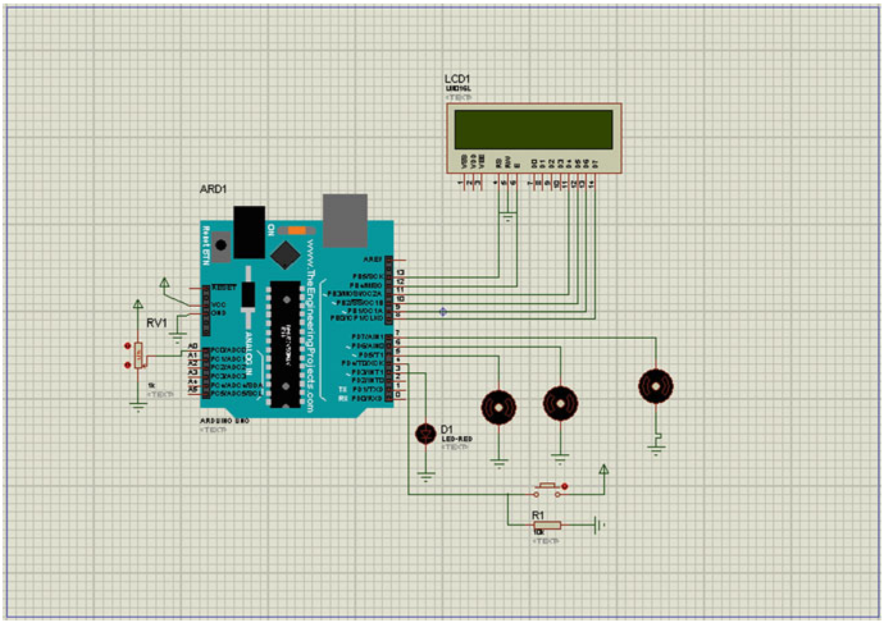


Fig. 6 Circuit diagram of Arduino system

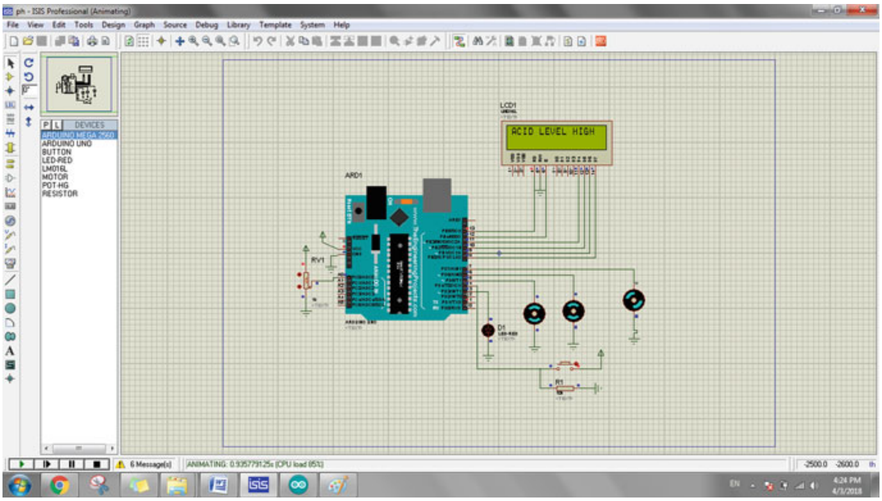


Fig. 7 Test result for acid level high

and microcontrollers, mainly Proteus is preferred as it can do simplified simulations of USB Communication, LED, LDR, etc. (Fig. 7).

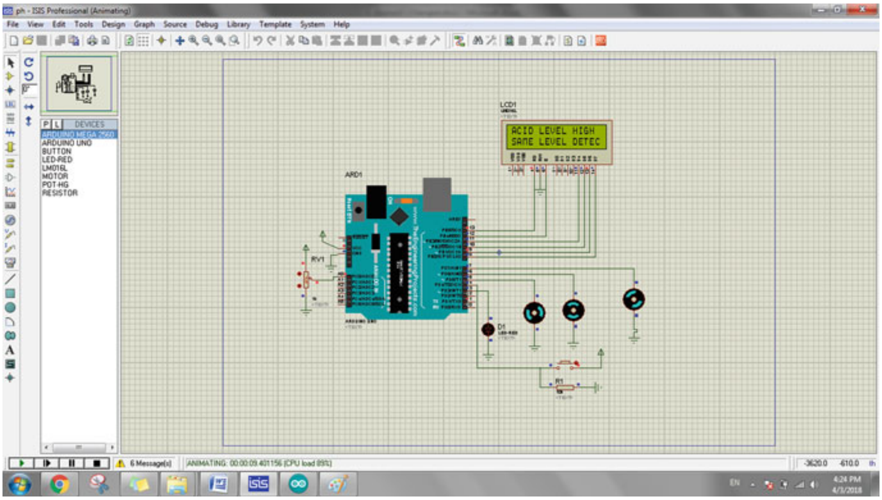


Fig. 8 Test result for checking acid level again

The implementation is appropriate for monitoring the ecosystem, environment, etc. Here, the information will be made to access worldwide. This system strictly monitors the main source of pollutants for water resources by testing acid level.

7.3 Acid Level Checking Again

Here, the acid level is being rechecked for which double assurance of drinking water purity is being checked and the system endows an environment for safe drinking water. Depending upon the pH level, our system will control the valve for changing the pH level in the tank. If the modification of the pH level fails, then the system will switch on to another valve to empty the tank. The IoT system monitors the pH and temperature level of water (Fig. 8).

7.4 Alkaline Level High

Depending upon the pH level, our system will control the valve for changing the pH level in the tank by testing the alkaline level. The alkaline nature in water is very beneficial because it is going to neutralize excess acidity in the body which in turn prevents various ailments in the body. The benefits of alkaline nature have increased the interest in changing the pH levels of water towards alkaline. Hence, normal tap water having a neutral pH level can be changed to alkaline nature: it can be done in

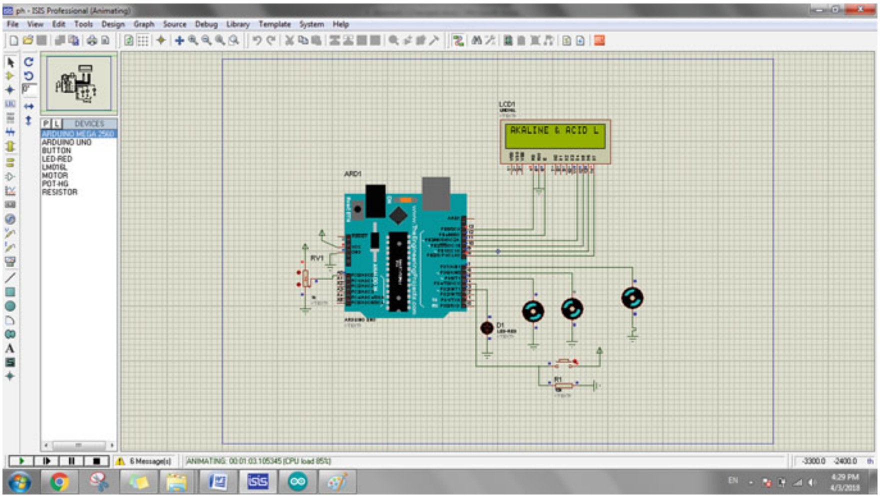


Fig. 9 Test result for alkaline level high

many ways such as faucet attachments, by increasing the additives in the pH levels, and also by the usage of special filters (Fig. 9).

7.5 Alkaline Same Level Detected Again

For double assurance level, the alkaline level is further checked, wherein an increase in water pH level indicates that the water is alkaline and it is ionized water. The water pH level determines whether the water is acidic or alkaline in a 0–14 scale, wherein if water is very acidic, the pH level will be 1; if water is very alkaline, the pH level will be 13; but for neutral case, the water pH level will be 7 like normal tap water (Fig. 10).

7.6 Simulation for Normal Water

Finally the normal water final test result will be tested and displayed which will ensure that the system endows safe drinking water to all the human beings (Fig. 11).

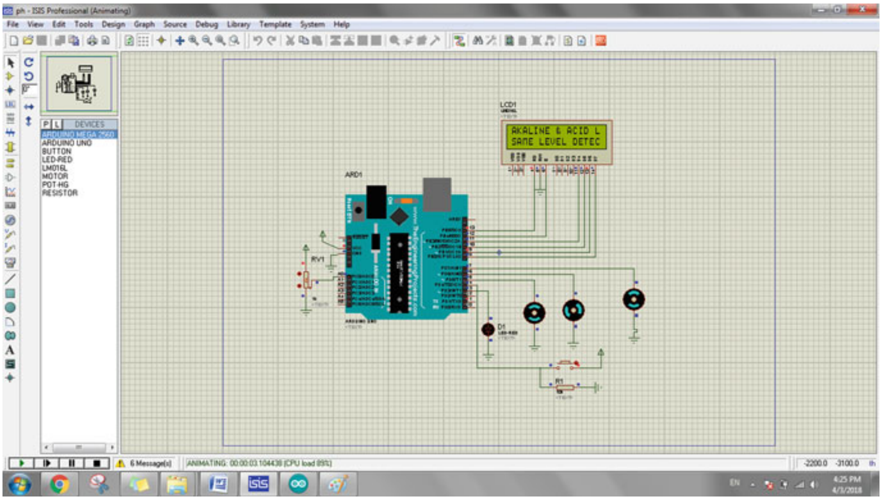


Fig. 10 Test result for checking alkaline level again

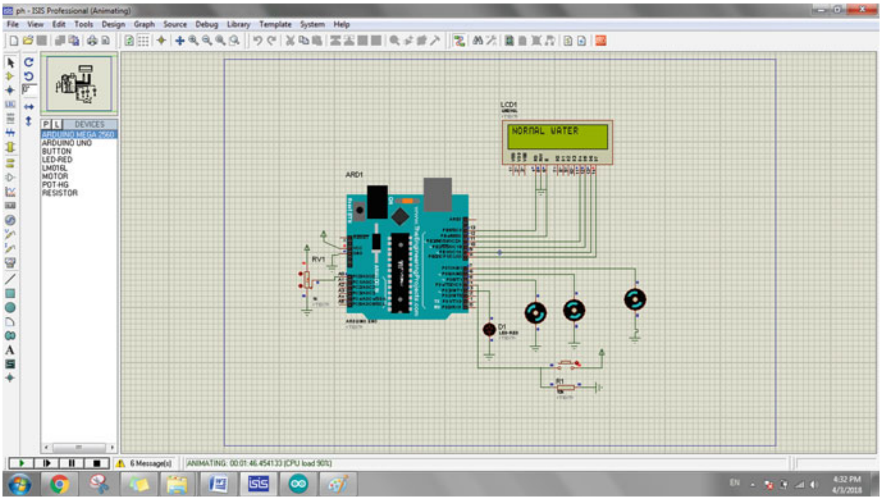


Fig. 11 Test result for normal water

8 Conclusion

This paper develops a real-time system for water quality monitoring in IoT environment. The proposed system consists of several water quality parameter sensors such as Arduino IDE and an IoT module. These devices are less expensive, have a higher efficiency, and have the ability to send and view the data that have been processed and analyzed to mobile device through Wi-Fi on cloud. The implementation is appropriate for ecosystem monitoring, environment monitoring, etc. The monitored information can be accessed from anywhere in the world. This system endows safe drinking water by performing close monitoring on the source of pollutant for water resources, so that we can ensure that all human beings are being provided with nontoxic drinking water. Depending upon the pH level, our system will control the valve for changing the pH level in the tank. If the modification of the pH level fails, then the system will switch on to another valve to empty the tank. The IoT system monitors the pH and temperature level of water.

References

1. K.A. Mamun, F.R. Islam, R. Haque, Smart water quality monitoring system design and KPIs analysis: case sites of Fiji surface water. Sustainability **11**, 7110 (2019). https://doi.org/10.3390/su11247110. MDPI journal 2019
2. A.T. Demetillo, M.V. Japitana, E.B. Taboada, A system for monitoring water quality in a large aquatic area using wireless sensor network technology. Sustain. Environ. Res. **29**, Article number: 12. (Springer Nature 2019)
3. H. Navarro-Hellin, R. Torres-Sanchez, F. Soto-Valles, C. Albaladejo Perez, J.A. Lopez-Riquelme, R. Domingo-Miguel, A wireless sensor architecture for efficient irrigation water management. Agric. Water Manag. **151**, 64–74 (2015)
4. C.T. Kone, A. Hafid, M. Boushaba, Performance management of IEEE 802.15.4 wireless sensor network for precision agriculture. IEEE Sens. J. **15**(10), 5734–5747 (2015)
5. N. Harris, A. Cranny, M. Rivers, K. Smettem, E.G. Barret-Lennard, Application of distributed wireless chloride sensors to environmental monitoring: initial results. IEEE Trans. Instrum. Meas. **65**(4), 736–744 (Apr. 2016)
6. M. Rivers, N. Coles, H. Zia, N.R. Harris, R. Yates, How could sensor networks help with agricultural water management issues?, in *IEEE 2015 Sensors Applications Symp., Zadar, Croatia, April 13–15* (2015)
7. J. Gutierrez, J.F. Villa-Medina, A. Nieto-Garibay, M.A. Porta Gandara, Automated irrigation system using a wireless sensor network and GPRS module. IEEE Trans. Instrum. Meas. **63**(1), 166–177 (2014)
8. F. Viani, A. Polo, M. Donelli, E. Giarola, A relocable and resilient distributed measurement system for electromagnetic exposure assessment. IEEE Sens. J. **16**(11), 4595–4604 (2016)
9. F. Viani, F. Robol, A. Polo, E. Giarola, A. Massa, Localization strategies in WSNs as applied to landslide monitoring, in *American Geophysical Union Fall Meeting*, San Francisco, USA, December 9–13 (2013)
10. F. Viani, F. Robol, A. Polo, P. Rocca, G. Oliveri, A. Massa, Wireless architectures for heterogeneous sensing in smart home applications, concepts and real implementations. Proc. IEEE **101**(11), 2381–2396 (2013)
11. F. Viani, A decision support system based on wireless power metering for energy saving in smart buildings. Microw. Opt. Technol. Lett. **57**(12), 2750–2752 (2015)
12. F. Viani, M. Bertolli, A. Polo, Low-cost wireless system for agrochemical dosage reduction in precision farming. IEEE Sens. J. **17**(1), 5–6 (2017)
13. F. Viani, Experimental validation of a wireless system for the irrigation management in smart farming applications. Microw. Opt. Technol. Lett. **58**(9), 2186–2189 (2016)
14. T. Wark, P. Korke, P. Sikka, L. Klingbeil, Y. Guo, C. Crossman, P. Valencia, D. Swain, G. Bishop-Hurley, Transforming agriculture through pervasive wireless sensor network. Pervasive Comput. **6**(2), 50–57 (2007)

15. B. Majone, F. Viani, E. Filippi, A. Bellin, A. Massa, G. Toller, F. Robol, M. Salucci, Wireless sensor network deployment for monitoring soil moisture dynamics at the field scale. Proc. Environ. Sci. **19**, 426–435 (2013)
16. O. Kaiwartya, A.H. Abdullah, Y. Cao, R.S. Raw, S. Kumar, D.K. Lobiyal, I.F. Isnin, X. Liu, R.R. Shah, T-MQM: testbed-based multi-metric quality measurement of sensor deployment for precision agriculture – a case study. IEEE Sens. J. **16**(23), 8649–8664 (2016)
17. H.P. Gupta, P.K. Tyagi, M.P. Singh, Regular node deployment for k-coverage in m-connected wireless networks. IEEE Sens. J. **15**(12), 7126–7134 (2015)
18. J.G. Jaguey, J.F. Villa-Medina, A. Lopez-Guzman, M.A. Porta Gandara, Smartphone irrigation sensor. IEEE Sens. J. **15**(9), 5122–5127 (2015)
19. P. Gimenez-Gomez, R. Escude-Pujol, C. Jimenez-Jorquera, M. Gutierrez-Capitan, Multisensor portable meter for environmental applications. IEEE Sens. J. **15**(11), 6517–6523 (2015)

Design of a WSN Platform for Internet of Things Applications

P. Madhumathy and Shweta Babu Prasad

1 Introduction

Drainage is a natural or man-made channel to drain the land from overload of unutilized water, rainwater, and wastewater. In developed cities, drainage channels are constructed to control the flow of water to prevent any damage to land or infrastructure and assets in the city.

Drainage are ought to be observed so as to maintain the appropriate capacity. Also all the places are not privileged to have a drainage monitoring squad which causes unreliability in the servicing and maintenance of the drainage. This in turn causes the drains to clog which leads to siltation bringing about floods in the adjoining and low-lying areas. Manual monitoring is additionally ineffective requiring multiple committed personnel to document all the changes. Hence, these shortcomings cause inefficient handling of any issues in the drainage.

Wireless sensor network (WSN) is a tracking technology composed of node sensors which are dispersed and organized by utilizing a wireless network system. Every hub is able to handle information using a microcontroller and GSM receiver and comprises a power supply system (battery or solar cell) along with at least one sensor. WSN frameworks have a more significant level of effectiveness when compared to wired networks as far as cost, adaptability, and reliability are concerned and might replace hybrid technology (wired or wireless) in the future. Additionally this technology can be enforced in several areas which regularly require control of data.

P. Madhumathy (✉) · S. B. Prasad
RV Institute of Technology and Management, Bangalore, India

Dayananda Sagar Academy of Technology and Management, Bangalore, India

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_11

In the course of recent decades, WSNs and their applications have been the subject of numerous studies. WSN is a system liable for gathering, preparing, and appropriating remote information to the proposed database. Since these sensors are generally introduced at remote places, regardless of the ongoing advances in the WSN innovation, its applications despite everything face critical difficulties. Among these network security threats, network framework, data compilation, deployment, and network coverage are the biggest concerns.

For instance, the environment can be completely monitored by sensor networks to avert any danger, or examination of mountain areas having volcanic activity by utilizing remote monitoring of seismic values. In metropolitan areas, WSNs can screen air pollution, strength building, flood hazards, noise pollution, and the video feed, and all this is entirely done through GSM.

The studies conducted place emphases on the control simulation in these examinations and accentuate control reenactment of sewerage framework for checking sensors and apparatus for screening in various conditions. The analysis concerning the application of a WSN in the waste management is not yet completed. In this chapter, discussion of the architecture of waste management systems to screen parameters using WSN has been done. Few of the hub sensors are set up to transmit the information of the current state of the drainage to the server. These values which are screened are the height of water accumulated in drainage, the way in which the water is let out and rainfall conditions in the vicinity of the drainage area.

1.1 Problem Statement

The structure currently used for drainage is not found to be technologically advanced. Hence, in case of a blockage, there is difficulty in narrowing down the precise location of the blockage. Additionally, immediate alerts of the blockage are not communicated which in turn causes the detection and repair of blocks to be tedious. The situation also turns disastrous when the pipes are entirely blocked. This leads to a lot of problematic conditions as a result of failed drainages.

Hence the proposed system claims to:

1. Identify the area.
2. Oversee the flow of waste through the channels by making use of the framework.
3. Utilize the flow sensors to identify the fluctuations in the flow.
4. Receive immediate alerts when channels are blocked, and these areas are identified by making use of IoT. GSM is used to send SMS.

1.2 Objectives

1. Cleaner urban areas and inventive management of waste.
2. Detection of levels of wastewater and blockages in the channels.
3. The rate of flow is constantly monitored, and the parameters are displayed as well as sent as an email when the parameters reach unexpected levels.
4. The environment is constantly checked for leakage in gas and also abnormal temperature which is then updated via IoT.
5. To achieve an efficient economical and flexible solution to monitor different parameters and also infrastructure management.

1.3 Chapter Organization

This chapter is divided into seven sections.

The chapter begins with the introduction which includes the background information and the problem statement.

Section 2 presents the literature survey performed for the chapter which is the study of different papers related to the topic.

Section 3 describes the proposed system and working of the module.

Section 4 gives a detailed description of the various hardware and software components used to build the system.

Section 5 covers the results obtained and also comprises discussions made in this regard and details of the user interface.

Section 6 gives the applications and advantages of the system.

Section 7 describes the summary of the chapter including the conclusions drawn and also suggests several ideas for future work.

2 Literature Survey

Prof. S.A. Shaikh et al. propose a system for overseeing smart city applications by employing Raspberry PI and IoT [1]. The purpose is to oversee the condition of resource in the city to upgrade administration and assist in growth and progress of the city. The smart city theory can be realized by adopting IoT so that wireless communication can be implemented. The system comprises of sensors from which several varieties of data can be collected and transmitted to the Raspberry Pi3 controller. These results obtained via the controller is communicated to the control room via an email and furthermore displayed on the user's PC.

Muragesh S. K. et al. propose manhole monitoring system by using automated IoT [2]. IoT consists of different organisms, communication equipment connected to sensor networks to assist in providing communication, and mechanized processes within the living world and information areas. Sensor network has a considerable function in IoT. A framework of an underground drainage and manhole monitoring system (UDMS) for IoT functionality which can be used to check the water level and the temperature pressure within the manhole and to check if the lid of the sewer vent is open. This design has advantages like being inexpensive, low efforts towards upkeep, quicker implementation, having a large number of sensors, extended life span, and increased QoS. Additionally, the power lines which are connected below the ground are also screened. UDMS can be used to remotely screen the different parameters of the sewer vents.

Chun Ho Wu et al. propose a framework for a WSN screening of biological and pharmaceutical (B&P) products [3]. To ensure that there is quality and integrity and no corrupt elements, the present screening structure requires upgradation. The authors suggest the framework of a WSN monitoring system for B&P product supply chain management consisting of three processes: (1) WSN is formed to provide complete data transmission, (2) to decide the ideal position of the sink, and (3) extra relay nodes are supplied to assist in fault tolerance in the network. The system provides a moderately basic arrangement for manufacturing corporations to follow and enforce while a WSN system is being designed.

Mohammad Dahman Alshehri et al. put forth the dispersed trust management architecture for the IoT (DTM-IoT) in [4]. In order to do this, the authors provide understanding of trust management and also negotiations by adopting clustering techniques in WSNs. In order to do this, the authors additionally establish the new approach to support their claims and administration in the IoT.

Bao F. et al. offer an effective protocol for administering trust in IoT systems to work with fluctuating nodes in which parameters might alter constantly [5]. An IoT system is set up in a smart community in which every node individually carries out trust evaluation. The convergence, accuracy, and resilience values are evaluated and validated through simulation.

Yichuan Wang et al. offer an energy-aware efficient multi-hop routing infrastructure for IoT [6]. In order to stabilize energy consumption, routing from source to destination needs to be examined for trust evaluation, instead of just the end hubs. An accurate spanning tree algorithm is used to stabilize power utilization and minimize trust cost. The outcome from the simulation claims that the approach put forth can minimize energy consumption largely by utilizing the trustworthy transmission.

3 Proposed Methodology

The majority of the urban communities embraced the underground drainage system for the maintenance and upkeep of cleanliness in the cities. In case the sewage system is not maintained well, there are chances that clean or drinking water gets contaminated with the seepage from the drain water. This could lead to rapid spreading of infectious diseases. During the rainy season, the stormwater drains might get blocked due to the accumulation of waste which in turn creates havoc in the lives of people. The vehicular traffic may get disrupted, the environment becomes messy, and the entire society is agitated. If the lid of the sewer vent is not closed properly, people and animals may accidently fall inside the drains. This problem can be efficiently resolved, by implementing an isolated screening system to screen various parameters of sewage system.

In order to find a solution to the issue of open sewerage system, the underground drainage system was maintained in multiple cities for health and sanitation reasons. The architecture included controlling manholes via the Internet. This existing system failed to cater with reduced cost and maintenance and rapid distribution and also included multiple sensors, but did not have a long lifetime, nor did it provide increased QoS. Here, the design included three-way alert systems and consisted of a route map, LCD display, and speakers.

3.1 Block Diagram

Figure 1 is a representation the flow of control among different components of the module.

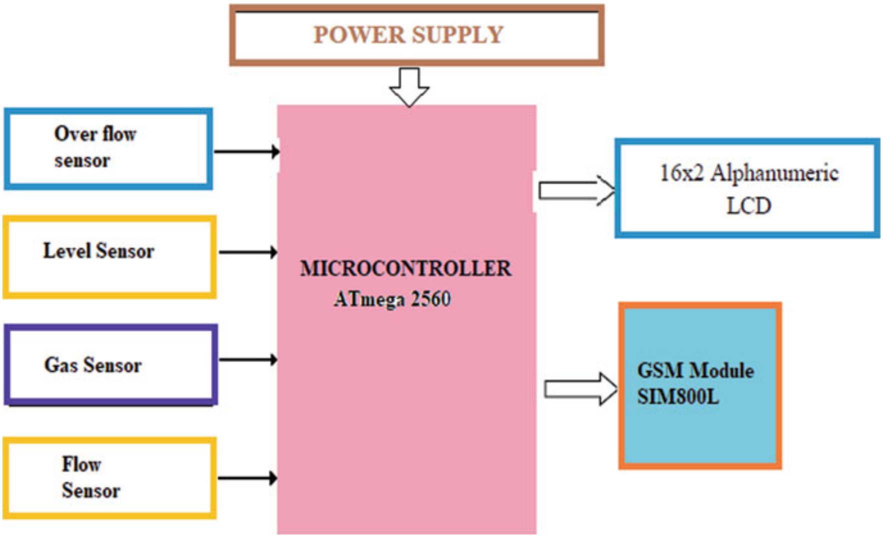


Fig. 1 Block diagram

3.2 Working

An underground sewerage monitoring system assists to keep up the well-being of the society in addition to security in the best way, thereby reducing the work crafted for the government work force.

Assortments of sensors having varied functionality (movement, level, overflow, gas sensors) are incorporated with the microcontroller so as to enable the framework to be "smart." At the point when the individual sensors reach their limits, the notification and the sensor value is transmitted to the microcontroller. Microcontroller then transmits the signal and position of the sewer vent to the authority in charge via GSM and GPS so that the responsible personnel are able to locate the issue without difficulty and take necessary action towards solving the problem.

Additionally, the microcontroller refreshes the current values of all the sensors in the sewer vents of a particular locality utilizing the IoT. A note on the status is displayed on the 16*2 LCD.

4 Details of Hardware and Software Used

In this section the different hardware components used are mentioned along with their functionality.

1. **Microcontroller**: The microcontroller used is ATmega 2560 which is 8-bit AVR RISC-based and operates at a voltage of 5 V and 256 KB of flash memory, 8 KB of SRAM, and 4 KB of EEPROM.
2. **Ultrasonic sensor**: Ultrasonic sensors can recognize mobility of objects and figure out the range of separation. They can also identify the edge of material in web guiding system.
3. **GSM module**: SIM800L is a Quad band to send data as SMS and voice using minimal power, GSM/GPRS system in a SMT type, capable of being planted in user applications.
4. **Flow sensor**: YF-S201 is used as flow sensor which makes use of Hall effect to sense the flow of water. A pinwheel sensor measures the volume of liquid moved through it.
5. **MQ-4 gas sensor**: Detects leakage in gas, and its sensitive material is SnO2. It is used to detect a variety of combustible gases.
6. **MQ-7 gas sensor**: It is a CO sensor which is highly sensitive to CO concentrations in air. The sensor is stable and has a long life span.
7. **LCD display (JDH162A)**: A 16X2 LCD (liquid crystal display) screen is the simplest and widely used electronic display unit used for different operations.
8. **Rain sensor module**: It is a tool for rain detection and measures the intensity of rain by the raindrops falling through the board.

The model is programmed as given below:

Arduino Mega2560 may be customized with the Arduino software, and a boot loader already exists on it to assist in uploading a different code to it. The communication takes place via the STK500 protocol.

It is possible to bypass the boot loader and manipulate the microcontroller via the ICSP header.

Rather than manually utilizing the reset button, Arduino Mega2560 is uses via a software active link. The control line (DTR) of the ATmega8U2 is linked to the reset line of the ATmega2560 via a 100 nF capacitor. The chip can be reset when this line is low and this assists to upload the code via the upload button.

This leads to multiple complications. In case the Mega2560 is linked to a PC having Mac OS X or Linux, a reset is performed every time there is a connection to it via USB. The boot loader runs on the Mega2560 for the next 30 s or so. While it is customized to disregard distorted information, the initial few bytes of information are received by the board over the connection. In case that a sketch on the board gets a single-time setup or other data, be assured the product it communicates and wait for a second once the connection is opened and before transmitting the data.

The Arduino Mega comprises a trace capable of being cut in order that the auto-reset can be disabled. To re-enable, the pads labeled "RESET-EN" available on the sides of the trace can be soldered. The auto-reset can be disabled by forming a connection between a 110-ohm resistor and the reset line.

To protect the USB ports from damage due to leaking of current or short circuit, the Arduino Mega consists of a polyfuse which can be reset. This fuse supplies with an added layer of protection in spite of there being internal protection.

If the USB port receives an overflow of 500 mA, the line eventually loses connection by the time the short is expelled.

The PCB has a great length of 4 inches and a width of 2.1 inches, and the USB connector and power jack extend beyond the earlier dimensions. Three screw gaps are provided so that the board is secured to any other surface.

4.1 Flowchart

The flowchart given below depicts how the various components of the model are connected, and also the flow of control is shown in Fig. 2.

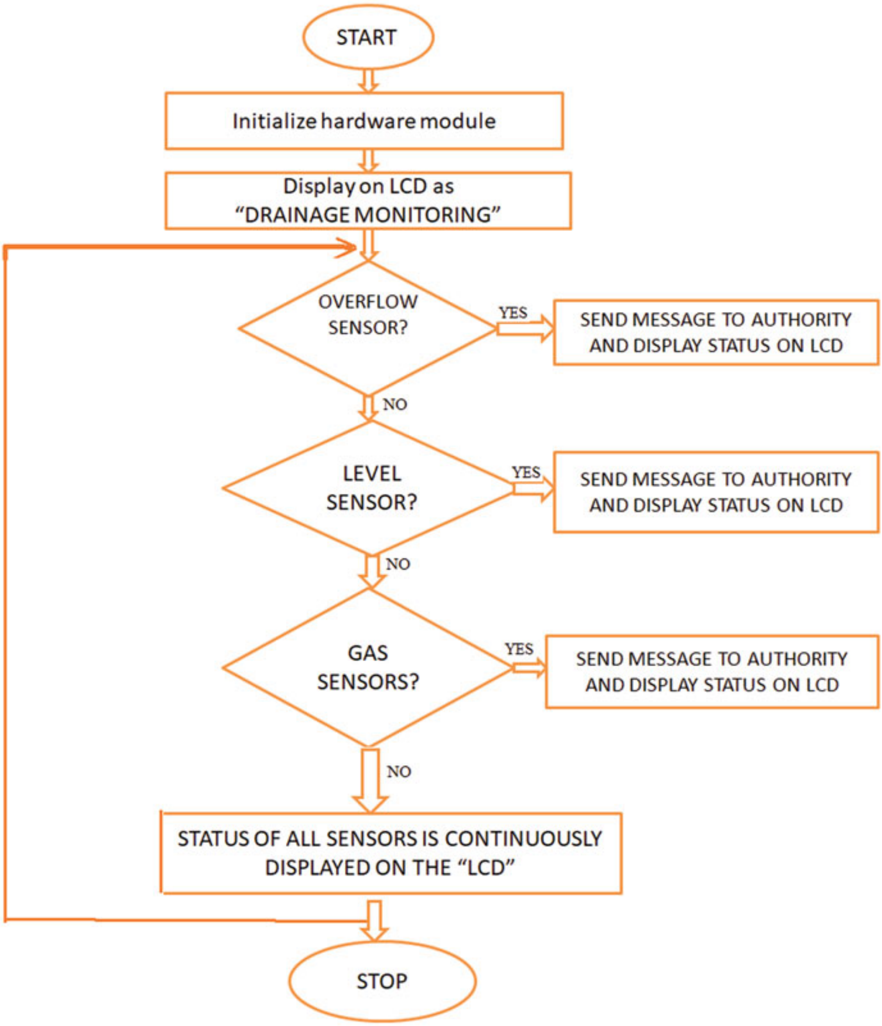


Fig. 2 Flow chart of the proposed system

4.2 Algorithm

1. START
2. Switch on the power supply to the hardware.
3. The hardware module is booted up.
4. "**DRAINAGE MONITORING SYSTEM**" is displayed on the LCD.
5. Microcontroller senses sensor value.
6. Temperature sensor senses the thermal conditions, and LCD displays it.

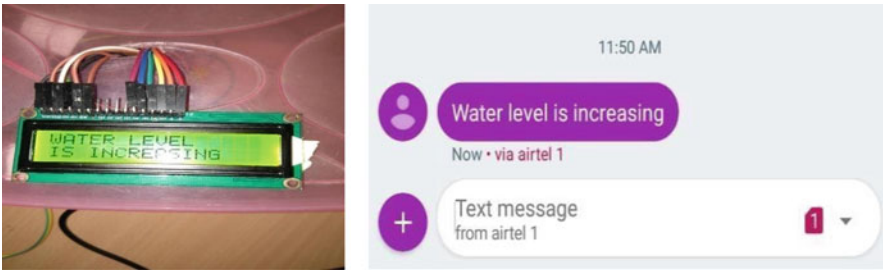


Fig. 3 Increase of water level displayed on LCD and corresponding SMS alert

7. MQ-4 sensor checks for extent of methane in the surroundings.
8. Sensor checks for the extent of in the surroundings.
9. With the increase in the level, there is a decrease in the flow.
10. GSM is used to send the location.
11. If the value on any sensor exceeds, then the value is set, and a message is sent through GSM.
12. Value obtained from the sensors is updated to the web server via IoT.
13. The measured parameters are displayed on the LCD.
14. STOP

5 Results and Discussions

5.1 Detection of Increasing Water Level in Manhole

The model is tested for water levels in the manhole. The message is generated to inform the user of the increasing water levels in the manhole and as shown in Fig. 3. This message serves as a warning of the possibility of overflow in the locality.

5.2 Detection of Leakage of Gas

The model to be tested is exposed to gas leakage. In this case the model is exposed to methane gas present in the sewage system. The model detects the presence of the gas and displays a message on the LCD. The value displayed alongside suggests the dangerous levels and suggests if it is safe or not for someone to enter the manhole for upkeep and maintenance. This information is also sent to the user in the form of an SMS as shown in Fig. 4.

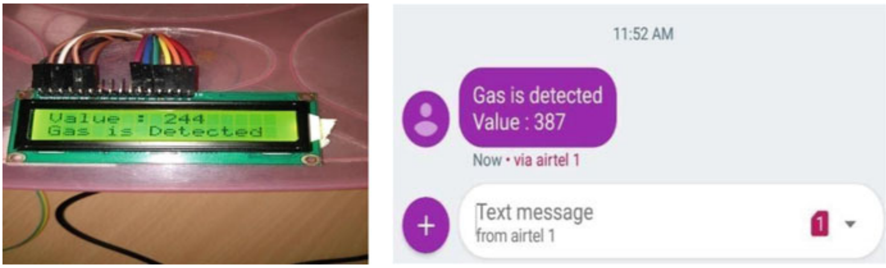


Fig. 4 Gas detection in LCD display and SMS alert

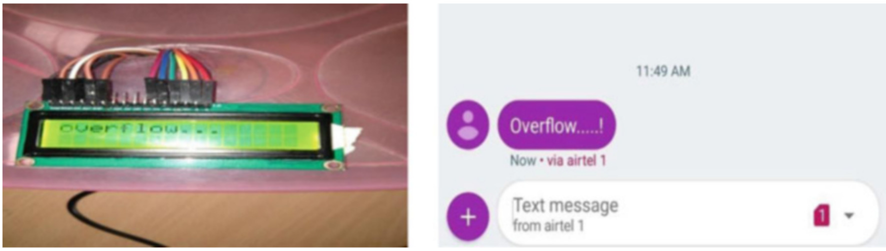


Fig. 5 Detection of overflow of manhole shown in LCD display and SMS alert

5.3 Detection of Overflow in Manhole

The model to be tested is subjected to an overflow. In case of an overflow in the manhole due to a blockage, the LCD display shows "Overflow." This message is also conveyed to the user via an SMS, and the user can thereby undertake removal of waste material from the manhole causing the blockage leading to an overflow as indicated in Fig. 5.

5.4 Status of the System

The generic status of the manhole can be displayed on the LCD which provides a summary of water flow, blockage, and gases generated within a manhole, and the status is displayed in the display unit as shown in Fig. 6.

6 Advantages and Applications

The advantages of the model are that it helps in preventing accumulation of water and hence any damage to property. The model can be used to detect any leakages

Fig. 6 Status of the system displayed on the LCD

via its sensors. The model is constructed in a way such that its various parameters assist to reduce soil erosion. Alerts will be generated in the initial stages of any detection of blockages. Also the model helps in the prevention of accumulation of toxic materials in addition to the prevention of the spread of diseases through contaminated water.

As far as the applications of the model are concerned, there are multiple uses of it. The system can be deployed in all metropolitan and developing cities. Industries which include equipment to carry liquids can implement the system. The government (civil department) has several uses of it. It can be used in water plants and also gas-supplying agencies.

7 Conclusion and Future Work

Underground drainage systems can face several difficulties including blockage. The proposed project introduces several ways to monitor and supervise the underground sewerage system. Multiple applications are analyzed in order to identify issues in real time which occur in the underground sewerage system. The values of thermal conditions, poisonous gases, movement, and elevation of water are regularly checked, and they are saved onto the server via IoT. As these values are easily available on the server, it helps the individual in charge to take any necessary actions, thereby minimizing manual intervention and inspection of the sewer vents which can be done only when very crucial. Likewise, these values are instantly updated over the web page; this further assists in the regular upkeep of the drainage, and any alerts received about the abnormal values help to prepare in advance. In this way the constructed model helps in monitoring the drainage remotely and also helps in updating the values of the different parameters to prevent any mishaps further assisting in maintenance.

Sensor networks are crucial empowering agents for the IoT. Due to the large variety in the utilizations and applications, it is difficult to characterize the basic necessities for the WSN hubs. Aforementioned strategy may be utilized to control the particular parameters, enhancements, and improvement in the field of WSN for various IoT operational areas. The strategy employed in the construction of the model can be further developed to incorporate in multiple fields where it is necessary to reduce human intervention for basic jobs of upkeep and maintenance or even to read current values. Systems can be developed along the same lines to assist in remote monitoring and instantaneous updates via WSN and the IoT.

References

1. H. Landaluce, A. Perallos, L. Arjona, A review of IoT sensing applications and challenges using RFID and wireless sensor networks, in *April 2020 RFID and Wireless Sensor Networks: Recent Advances in the Design of IoT Applications*, (2020)
2. R. Hassan, F. Qamar, M.K. Hasan, A.H.M. Aman, A.S. Ahmed, Internet of Things and Its Applications: A Comprehensive Survey, in *Symmetry*, (October 2020) (ISSN 2073-8994; CODEN: SYMMAM)
3. J. Sengupta, S. Ruj, S.D. Bit, A comprehensive survey on attacks, security issues and blockchain solutions for IoT and IIoT. J. Netw. Comput. Appl. **149**, 102481 (2020)
4. H. Ahmadi, G. Arji, L. Shahmoradi, R. Safdari, M. Nilashi, M. Alizadeh, The application of internet of things in healthcare: A systematic literature review and classification. Univ. Access Inf. Soc. **18**, 837–869 (2019)
5. S.A. Shaikh, S.A. Sonawane, Monitoring smart city application using raspberry PI based on IoT. Int. J. Adv. Res. Ideas Innov. Technol. **4**(1), 42–48 (2018)
6. S.K. Muragesh, S. Rao, Automated internet of things for underground drainage and manhole monitoring systems for metropolitan cities. Int. J. Innov. Sci. Eng. Technol. **2**(4), 211–216 (June 2015)

Performance Analysis of Modulation Techniques over a Smart City Optical Communication Channel Under Weak Atmospheric Turbulence

Sazid Hasan, Md Musfiqul Islam, Firoz Ahmed, S. M. Mustahsin Saabit, Nazmus Shaker Nafi, Mohammad Kamrul Hasan, Shayla Islam, and Imran Memon

1 Introduction

Free space optical (FSO) communication is a technology which is line of sight (LOS) that utilizes light-emitting diodes (LED) or lasers to propagate a light signal through free space, which acts as a communication channel between transceivers. FSO links providing 2.5 Gbps [1] of video, voice, and data transmission have been implemented. FSO is generally implemented using transmission wavelengths between 780 nm and 1600 nm. FSO systems have the following characteristics:

S. Hasan · M. M. Islam · F. Ahmed
Military Institute of Science and Technology, Dhaka, Bangladesh

S. M. M. Saabit
Swinburne University of Technology, Melbourne, Australia

N. S. Nafi
Melbourne Institute of Technology, Melbourne, Australia
e-mail: nnafi@mit.edu.au

M. K. Hasan (✉)
Center for Cyber Security, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
e-mail: mkhasan@ukm.edu.my

S. Islam
Department of Computer Science, UCSI University Malaysia, Kuala Lumpur, Malaysia
e-mail: shayla@ucsiuniversity.edu.my

I. Memon
Department of Computer Science, Bahria University, Karachi, Sindh, Pakistan
e-mail: Imranmemon.bukc@bahria.edu.pk

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_12

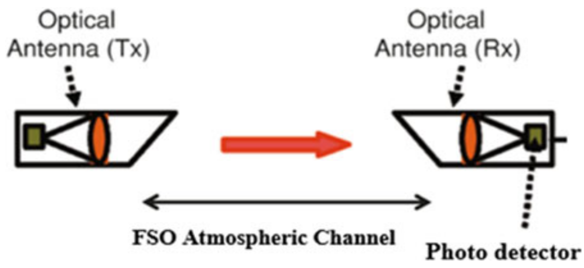


Fig. 1 Simple FSO communication model

- Capable of operating at very higher power levels
- High-speed data modulation
- Low power consumption
- Minimal performance degradation with varying temperatures

The elements of FSO are typically described in three stages. By following the Beer-Lambert law [2], the transmitter converts digital information into optical radiation that is transmitted across the free space, where turbulent eddies exist that include smoke, gas, fog, cloud, or rain. The signal at the receiver is processed to return the digital information.

In Fig. 1 an electrical message signal is converted into light using a media converter before the light is modulated and transmitted using an optical transmitter. The antenna of receiver gets the signal which is optical and demodulates it before conversion to an electrical message signal. However, signal distortion can occur when the transmission of signal occurs through freeways due to turbulence.

One of the major challenges for FSO is atmospheric turbulence, which creates signal distortion and possible signal loss. Light waves can be attenuated, distorted, or deflected by atmospheric turbulence. The light signal at the receiver can be detrimentally affected by the attenuation that occurs in free space due to absorption and scattering. Gaining an understanding of the FSO communication's performance under weak nature of conditions of turbulence is important and will help in the design of systems that minimize bit error rates (BER).

In central business district (CBD), making connection via cable is very complex and costly. FSO is the future concept for communication in short distance. To establish FSO ground-to-ground connection, some factors that obstruct the LOS needs to be improved. Ground-to-ground FSO communication is the focus of this paper. An analysis is presented of the BER performance of selected modulation techniques when weak turbulence conditions are experienced. Also, the power penalty has been derived for selected modulation techniques and for various link distances. The analysis provides the optimum FSO system parameters.

The organization of this paper is as follows. Methodology is elaborated in Sect. 2. Results and analysis are presented in Sect. 3. Section 4 concludes the paper and highlights future research.

2 Methodology

FSO transmits optical wave of data like voice, video, etc., using air as the medium of transmission. FSO transmission technology is relatively simpler than optical fiber technology. It includes two structures. One is a transceiver of optical nature that comprises a transmitter which has laser, and another one is a receiver to offer capability of being full-duplex (bidirectional). There is a telescope in the receiver end which collects the information. The telescope at the receiving end links to a receiver having a high sensitivity that is connected by optical fiber. Unlike frequencies of radio, this technology needs not require licenses of spectrum. Its interfaces which are open support tools from various vendors and are easily upgradeable.

In the field of FSO, extensive research is ongoing. There are different types of modulation techniques for FSO transmission. ASK, FM, OOK, PSK, PWM, OFDM, and PPM are some modulation techniques that are considered for FSO transmission in a different research project. This paper focused on PSK and considered the optimum link distance to transmit with less error.

In this paper, a weak atmospheric turbulent channel is considered to simulate the weak atmospheric turbulence; additive white Gaussian noise model, atmospheric turbulence model, log normal model, gamma PDF model, and atmospheric turbulence channel with pointing errors are used. Using these models in the probability of bit error for M-PSK (BPSK, Q-PSK, 8-PSK, and 16-PSK) modulation is simulated over 500 m, 1000 m, 1500 m, and 2000 m length respectively in MATLAB simulation. In the simulation the probability of BER for M-PSK is calculated and plotted against SNR for the lengths 500 m, 1000 m, 1500 m, and 2000 m. After modulating the optical signal with M-PSK modulation and transmitted through media containing weak atmospheric turbulence particles, the modulated optical signal at the receiver ends with the errors which are calculated with standard Gaussian approximation. Then, power penalty over link distances is plotted which denotes the optimum modulation technique and performance of the optical communication.

2.1 Error Rate Performance on M-PSK Modulation

After modulation, what we get is the signal with the errors. So to get the exact signal performance, it is needed to measure the rate of bit error. The mean rate of error of a subcarrier system over channels of turbulence can be shown as

$$P_e = \int_0^\infty P_e(I) f_I(I) dI \tag{1}$$

where p_x presents the probability of conditional error and f_I (I) presents the channel gain's PDF.

2.2 Bit Error Rate (BER)

The rate of error of bit that represents the digital link's quality is measured from the error bits number obtained divided by the quantity of transmitted bits

$$\text{BER} = (\text{bits in error}) / (\text{number of bits obtained in total})$$

BER can also be known as per the probability of error (POE) [3] and shown by Eq.

$$POE = \frac{1}{2}(1 - \text{erf})\sqrt{\frac{E_b}{N_0}} \tag{2}$$

Here, erf = error function, E_b = energy in one bit, and N_0 = noise power spectral. For links of FSO with a modulation scheme of an on-off keying, BER can be written as

$$BER = \frac{\exp(-SNR/2)}{(2\pi SNR)^{0.5}} \tag{3}$$

2.3 Signal-to-Noise Ratio (SNR)

SNR is the total obtained signal power's ratio over the strength of noise in the range of frequency of an operation. Noise strength includes the environment's noise and other signals which are unwanted. BER is in reverse to SNR. The relation between BER and SNR is very difficult to be determined in the environment of multi-channel. Signal-to-noise ratio (SNR) is being calculated in decibels and shown by Eq. (4)

$$SNR = 10\log_{10}\left(\frac{signal\ power}{noise\ power}\right) \tag{4}$$

Both BER and SNR are utilized for evaluating the excellence of communication systems. In presence of turbulence, the SNR is represented as follows [4]:

$$SNR = \left(0.31{C_n}^2 k^{7/6} l^{11/6}\right)^{-1} \tag{5}$$

The photo detector's total surface area is large enough so that the effective SNR comprises the effect of the spreading of the beam. The effective SNR is stated as [5–7]

$$SNR_{eff} = \frac{SNR}{1 + 1.33{\sigma_i}^2 \left[\frac{2l}{k\omega(l)^2} \right]^{5/6}} \tag{6}$$

2.4 BER and SNR on M-PSK Modulation

The mean value of BER for modulation of BPSK can be estimated by a finite series as [8]

$$P_2 \approx \frac{\Delta\,(\alpha,\beta,\varphi)}{2\sqrt{\pi}} \sum_{p=0}^{K} \left[\frac{\Gamma\left(\frac{p+\beta+1}{2}\right) a_p\,(\alpha,\beta,\varphi,A_0)}{(p+\beta)\left(\varphi^2-p-\beta\right)} \overline{\gamma}^{-\frac{p+\beta}{2}} - \frac{\Gamma\left(\frac{p+\alpha+1}{2}\right) a_p\,(\beta,\alpha,\varphi,A_0)}{(p+\beta)\left(\varphi^2-p-\alpha\right)} \overline{\gamma}^{-\frac{p+\alpha}{2}} \right] \tag{7}$$

where K = $\lfloor \varphi^2 - \alpha \rfloor > 0$ and $\lfloor . \rfloor$ is the floor function. The BER for the modulation of BPSK in high regimes of SNR can be estimated by [9]

$$P_2^{\infty} = \frac{\Delta\,(\alpha,\beta,\varphi)\left(\frac{\alpha\beta}{A_0}\right)^{\beta} \Gamma\left(\frac{\beta+1}{2}\right)}{2\sqrt{\pi}\,\Gamma\,(\beta+1-\alpha)\left(\varphi^2-\beta\right)\beta} \overline{\gamma}^{-\frac{\beta}{2}} \tag{8}$$

where $\alpha < \varphi^2$. Since $\alpha > \beta$, the inequality $\varphi^2 > \alpha$ also implies $\varphi^2 > \beta$, which will ensure the rate of error as positive.

2.5 Probability of Bit Error for M-PSK Modulation

For a modulation of M-PSK, the bits amount in each symbol of constellation is

$$k = \log_2(M) \tag{9}$$

Since each symbol consists of bits of k number, the symbol-to-noise ratio E_s/N_o is K times the ratio of bit to noise E_b/N_o, i.e.,

$$\frac{E_s}{N_o} = K \frac{E_b}{N_0} \tag{10}$$

The bit error probability for a general M-PSK modulation is [10],

$$P_b = \frac{1}{k}\left[\sqrt{\frac{kE_b}{N_o}}Sin\left(\frac{\pi}{M}\right)\right] \quad (11)$$

2.6 Binary Phase-Shift Keying (BPSK)

BPSK is a digital system of modulation that carries data by modulating or changing a reference signal phase. It practices two phases, which are opposite by 180°. The BER equation of modulation of BPSK is as given below [11, 17–19]:

The signal of BPSK can be represented as

$$S_n(t) = \sqrt{\frac{2E_b}{T_b}}\cos\left(2\pi f_c t + \pi\ (1 - n)\right. \quad (12)$$

The transmitted power of BPSK signal is $P_t = (E_b/T_b)$. Received signal power of BPSK signal is $S = P_r(d) = P_t.e^{\alpha d}$, where α is the attenuation constant and d the distance between receiver and transmitter. The signal-to-noise ratio is

$$SNR = \frac{s^2}{\sigma_{th}^2 + \sigma_{shot}^2} \quad (13)$$

The conditional BER is given by

$$BER = \frac{1}{2}\text{erfc}\left(\sqrt{SNR}\right) \quad (14)$$

The mean value of BER of the BPSK system over weak turbulence due to pointing errors and path loss can be expressed from Eq. (1) by [12]

$$P_{BPSK} = \int_{-\infty}^{+\infty} P_e(I)\ f_I(I)\ dI \quad (15)$$

2.7 M-Phase Shift Keying (M-PSK)

In M-PSK ($M = 4,8,16$), the signal equation is like BPSK. The only difference is the phase shifts. With increasing M value, the phase is changing more. The signal of M-PSK can be represented as

$$S_n(t) = \sqrt{\frac{2E_b}{T_b}} \cos\left(2\pi f_c t + (2n - 1)\frac{\pi}{M}\right) \tag{16}$$

The transmitted power of M-PSK signal is $P_t = E_b/T_b$. Received signal power of M-PSK signal is $S = P_r(d) = P_t.e^{\alpha d}$, where α is the attenuation constant and d the distance between receiver and transmitter. The signal-to-noise ratio is

$$SNR = \frac{s^2}{\sigma_{th}^2 + \sigma_{shot}^2} \tag{17}$$

The equation of rate of bit error of M-PSK modulation is given below:

$$BER = \frac{M}{4} \text{erfc}\left(\sqrt{SNR}\right) \tag{18}$$

The mean value of BER of the M-PSK system over weak turbulence due to pointing errors and path loss can be expressed from Eq. (1)

$$P_{M-PSK} = \int_{-\infty}^{+\infty} P_e(I)\ f_I(I)\ dI \tag{19}$$

3 Results and Discussions

Following the analytical approach, the free space optical communication's performance with different techniques of modulation over different lengths is evaluated. The standard Gaussian approximation (SGA) evaluates the SNR and BER performance in different modulations (Table 1).

3.1 Performance Analysis of Different PSK Modulations

Plots of BER vs SNR for BPSK, QPSK, 8-PSK, and 16-PSK modulation considering link distance of 500 m are shown in Fig. 2. In Fig. 2, it is clear that increasing the symbol bits for modulation SNR increases irrespective with BER. The result showed

Table 1 System parameters [11–19]

Parameter	Symbol	Value
Link distance	L	500 m < L < 2500 m
Boresight	A_0	0.20 m
Wavelength	λ	1550 m
Large-scale distribution	α	2.04 < α < 4.03
Small-scale distribution	β	1.1 < β < 3.45
Effective radius of beam at the transmitter	W_0	0.2 m
Receiver radius of beam at the receiver	W	0.5 m
Distance from the mean beam center on the transverse plane	r	0.05 m
Optical wave number	K	2~16
Strength of atmospheric turbulence	${c_n}^2$	10^{-14}

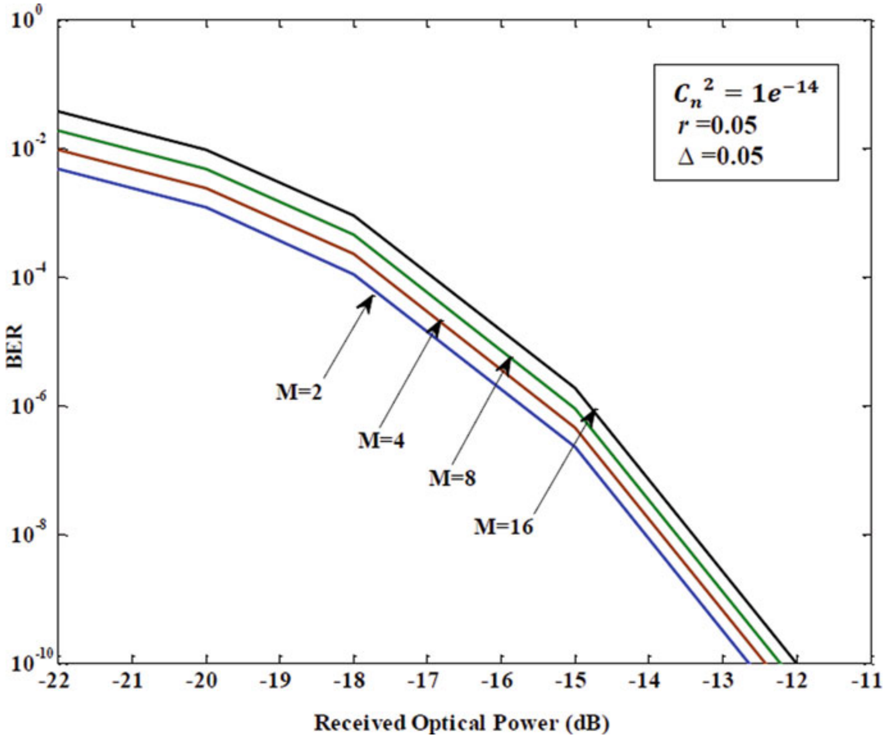


Fig. 2 BER vs SNR of BPSK, QPSK, 8-PSK, and 16-PSK modulation, where distance of link L = 500 m, and taking the parameters ${C_n}^2 = 1e^{-14}$, $r = 0.05$, $\Delta = 0.05$

that 16-PSk has many bit symbols, which indicates better performance than others. Taking a specific SNR, we can evaluate that 16 PSK has less BER than others, 2 PSK has higher BER, and 4, 8 PSK have BER in between 2 and 16 PSK.

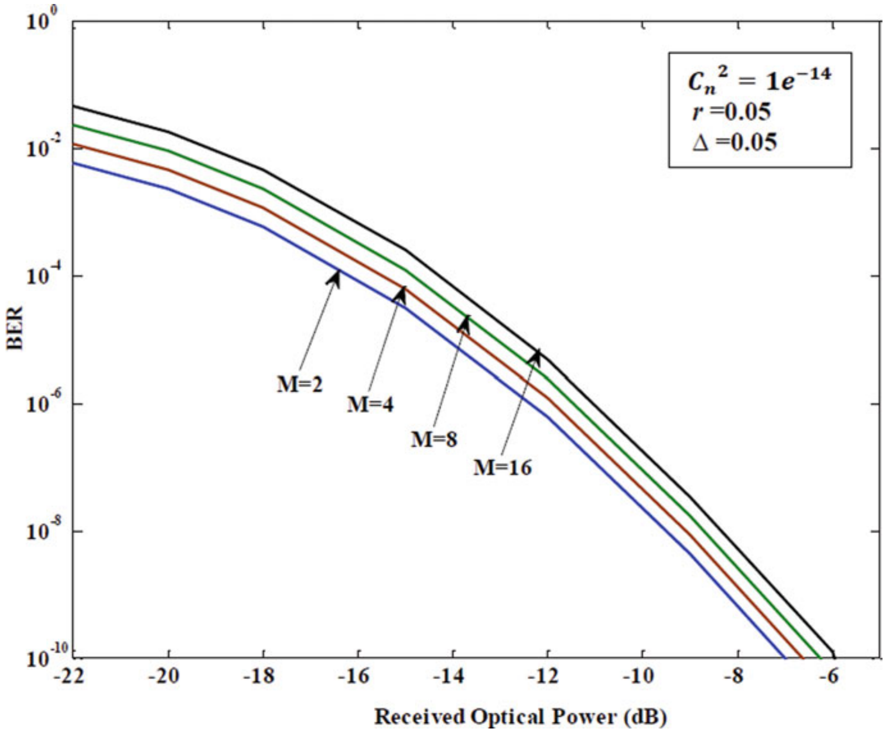


Fig. 3 BER vs SNR of BPSK, QPSK, 8-PSK, and 16-PSK modulation, where link distance L = 1000 m

Plots BER vs SNR of BPSK, QPSK, 8-PSK, and 16-PSK modulation for length of 1000 m are shown in Fig. 3. In Fig. 3, we can see that increasing the symbol bits for modulation SNR increases irrespective with BER. Similarly, with increasing modulation array, the BER performance is much more better.

Plots BER vs. SNR of BPSK, QPSK, 8-PSK, and 16-PSK modulation for the length of 1500 m are shown in Fig. 4. It can be noticed from Fig. 4 that the symbol bits for modulation SNR increases irrespective of BER. Similarly, 16- PSK has better performance.

Plots BER vs SNR of BPSK, QPSK, 8-PSK, and 16-PSK modulation for length of 2000 m are depicted in Fig. 5. Figure 5 depicts that the symbol bits for modulation SNR increases irrespective with BER.

Plots BER vs SNR of BPSK, QPSK, 8-PSK, and 16-PSK modulation for length of 2500 m are shown in Fig. 6.

From the above figures (Figs. 2 until 6), it is cleared that with the increase in distance from the transmitter to the receiver, BER increases, and consecutively SNR decreases.

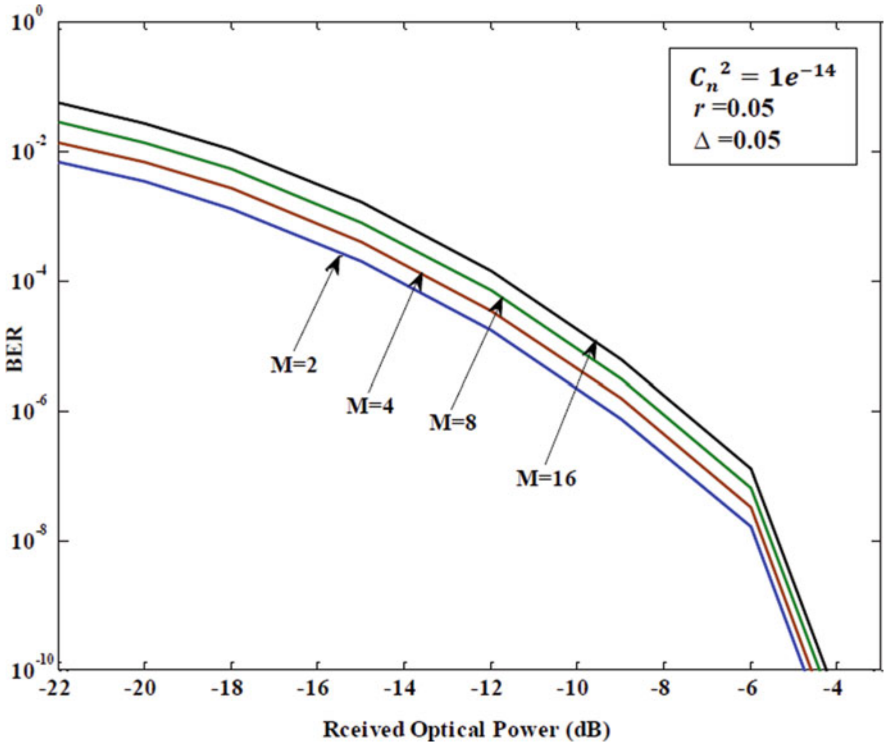


Fig. 4 BER vs. SNR of BPSK, QPSK, 8-PSK, and 16-PSK modulation, where link distance L = 1500 m

Figure 7 shows that for each M-PSK, the power penalty curve is increased over increasing link distance. Increasing link distance, power penalty curve behaves differently. L = 1500 m shows less power variation for all types of modulation as it is less steeper. Hence, considering each M-PSK (BPSK, QPSK, 8-PSK, 16-PSK) at 1500 m link distance, power penalty is more acceptable than other link distances.

3.2 Analysis of Different Modulations for Constant Link Distance

In Fig. 8, the result of simulation depicts the BPSK modulation's performance using different link distances. By increasing the link distance, BER also increases. The figure shows that at distance of link L = 500 m, the SNR is approximately −13.5, and at distance of link L = 2500 m, the SNR is approximately −4.50. Taking 500 m length as a reference, with increased length from 500 m to 1000 m required power to be increased by 5.3 dB approximately, for 1500m to 500m increased amount

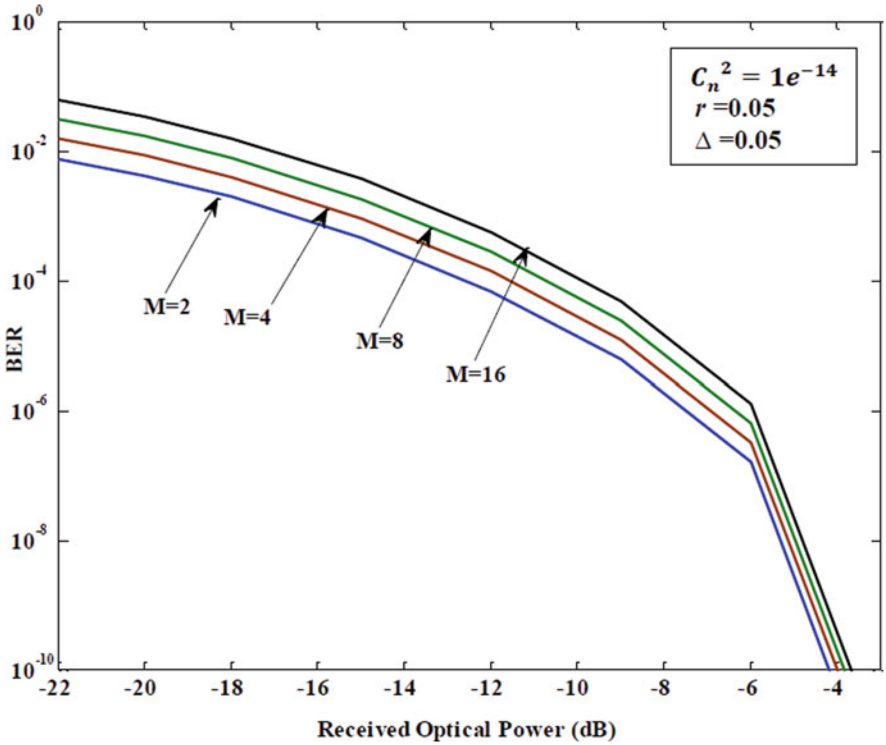


Fig. 5 BER vs SNR of BPSK, QPSK, 8-PSK, and 16-PSK modulation, where link distance L = 2000 m

of power is around 8.1 dB. With an increasing length from 500 m to 2000 m, required power is increased by 8.7 dB, and from 500 m to 2500 m, required power is increased by 9 dB. So with the enhance in length, the required power also gets enhanced. Again at −13.5 dBm received power (as reference) for 500 m length, BER is 10^{-9}; for 1000 m length, BER is 10^{-6}. Accordingly, for 1500 m, 2000 m, and 2500 m, BER > 10^{-6}. So it is noticed that for 1500 m, 2000 m, and 2500 m, bit error rate is higher.

In Fig. 9, QPSK modulation is done using different link distances. Increasing the link distance, BER increases. At distance of link L = 500 m, the SNR is approximately −13.1 dB, distance of link L = 2500 m where the SNR is approximately −4.2. Taking 500 m length as a reference, with an increase in length from 500 m to 1000 m, the required power is increased by 5.2 dB accordingly; from 500 m to 1500 m, it is 8.6 dB; and for 2000 m the required power is 9.3dB. With an increasing length from 500 m to 2500 m, the required power is increased by 9.5 dB. Hence, with the enhance in length, the required power also gets enhanced.

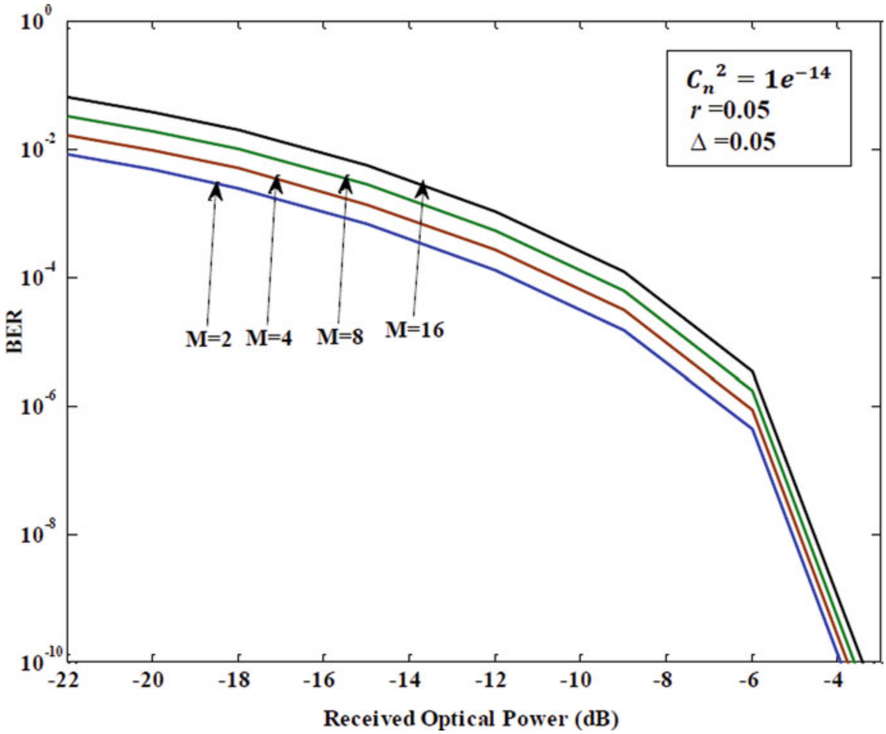


Fig. 6 BER vs SNR of BPSK, QPSK, 8-PSK, and 16-PSK modulation, where link distance L = 2500 m

In Fig. 10, evaluation is done for the performance of 8-PSK modulation using different link distances. Figure 10 shows that at link distance L = 500 m, SNR is around −13 dB, and at link distance L = 2500 m, SNR is around −4.1 dB.

Taking 500 m length as a reference, with an increase in length from 500 m to 1000 m, the required power is enhanced by 6.5 dB; for length from 500 m to 1500 m, the required power is enhanced by 9 dB. Similarly, for link distance from 500 m to 2000 m, the required power is increased by 9.5 dB; for length from 500 m to 2500 m, the required power is increased by 9.7 dB. Hence, with the enhance in length, the required power is also enhanced. Again at −13 dBm received power (as reference) for 500 m length, BER is 10ˆ(−9); for 1000 m length, BER is 10ˆ(−5). Accordingly, for 1500 m, 2000 m, and 2500 m, BER > 10ˆ(−6). So it is justified that for 1500 m, 2000 m, and 2500 m, bit error rate is higher(Fig. 11).

In the case of 16-PSK modulation using different link distances, from the figure, at distance of link L = 500 m, the SNR is about −12.9, and at distance of link L = 2500 m, the SNR is approximately −4 dB. Taking 500 m length as a reference, with increased length from 500 m to 1000 m, the required power is increased by 5.8 dB. Similarly, for length 500 m–1500 mt, the required power is increased by 8 dB. For link distance from 500 m 2000 m, the required power is increased by 8.7

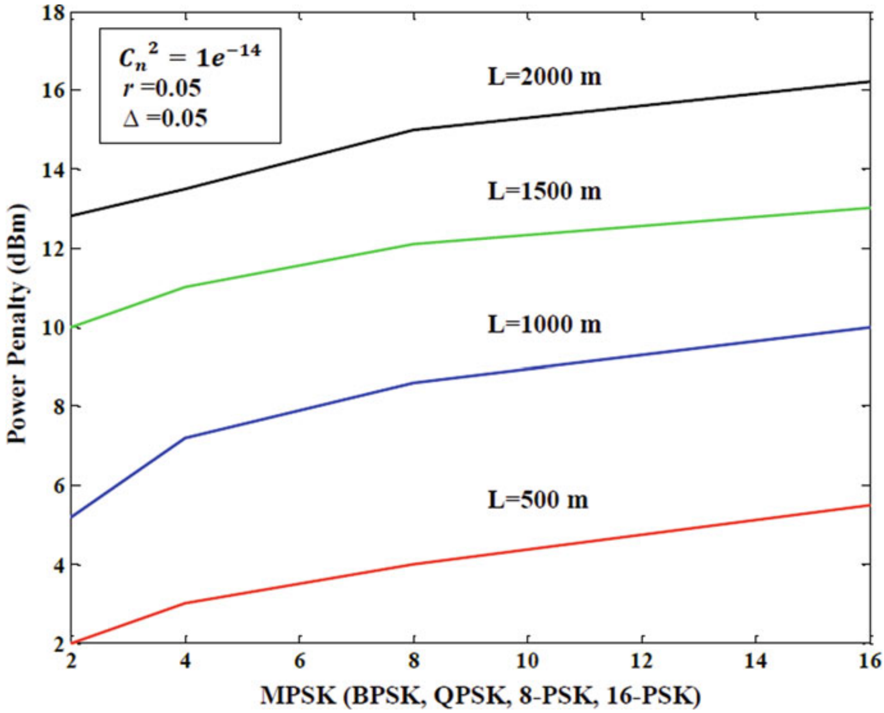


Fig. 7 Power penalty vs. M-PSK (BPSK, QPSK, 8-PSK, 16-PSK), where L varies from 500 m to 2500 m

dB. For length from 500 m to 2500 m, the required power is increased by 8.9 dB. Hence, with the enhance in length, the required power also gets enhanced. Again at −12.9 dBm received power (as reference) for 500 m length, BER is 10^(−9); for 1000 m length, BER is 10^(−5). Accordingly, for 1500 m, 2000 m, and 2500 m, BER > 10^(−6). So it is noticed that for 1500 m, 2000 m, and 2500 m, bit error rate is higher.

From Figures 8 until 11, it is confirmed that the increase in link distance power of 16-PSK is much more than that of BPSK; hence, the power penalty of 16-PSK is larger than other PSK. Here for a specific link distance (1000 m), the required SNR (power) for BPSK, QPSK, 8-PSK, and 16-PSK are 5.1 dB, 5.5 dB, 5.9 dB, and 6.3 dB, respectively. So the required power for BPSK is lower than other modulations. But if we consider both the performance analysis and power penalty, then 8-PSK has power less than 6dB and gives better BER than 16-PSK. So 8-PSK can be considered as optimum modulation.

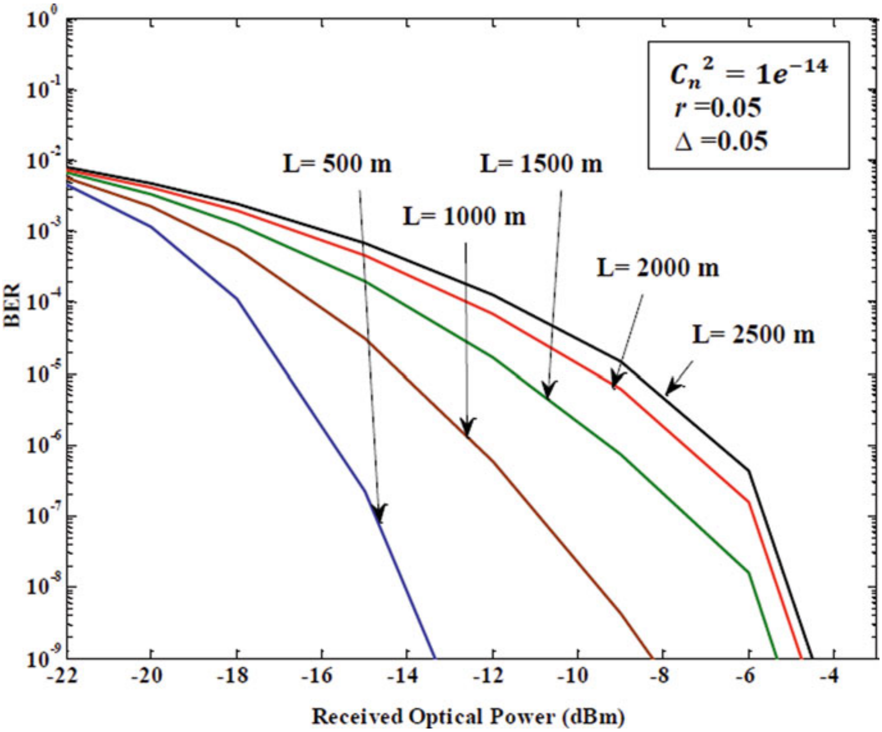


Fig. 8 BER vs. SNR of BPSK over variant distances of link (L), where L varies from 500 m to 2500 m

4 Conclusion

In the paper, at first, we introduced the model of FSO system and then the various multiple modulation techniques and finally discussed the M-PSK technique briefly. We analyzed different modulation techniques for different link distances using MATLAB by simulating data due to weak turbulence. Considering different link distances, we have found different values of received power at the same BER. The effect of BER on different link distances has been calculated and simulated also. Link distance from 500 m to 2500 m and BPSK, QPSK, 8-PSK, and 16-PSK modulation techniques are considered for the analysis. In the simulation, two power penalty curves are finally driven. One is the power penalty for different M-PSK modulations (M = 2, 4, 8, 16); another is the power penalty for different link distances. Here we observe that BPSK and QPSK BER are comparatively higher than others, whereas power loss is comparatively less at different link distances. For 8-PSK, 16-PSK BER is comparatively less, whereas power loss is comparatively high at different link distances. Comparing all the power loss for different modulation techniques, 8-PSK gives the optimum parameter. At link

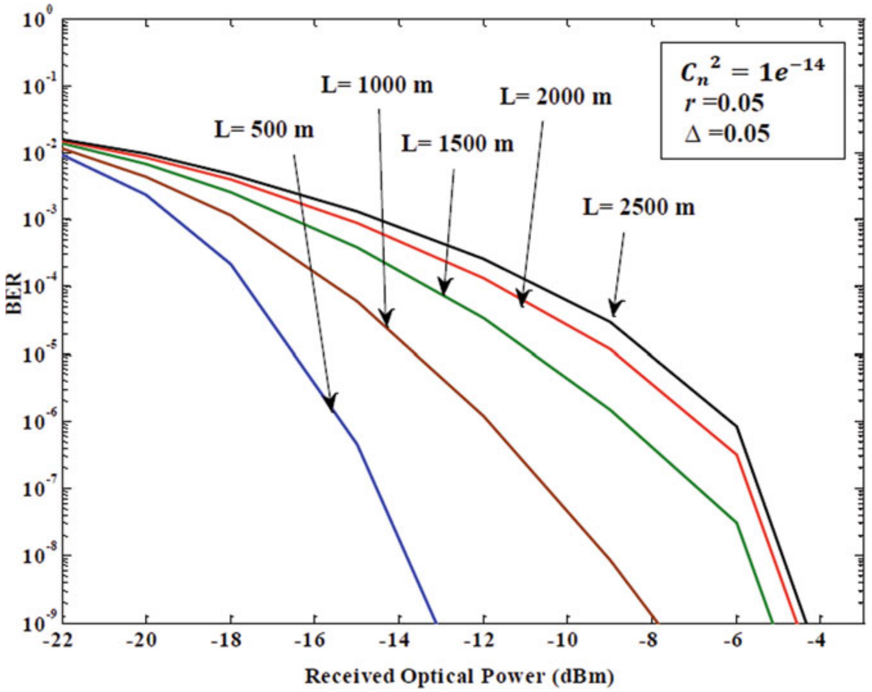


Fig. 9 BER vs. SNR of QPSK over variant distances of link (L), where L varies from 500 m to 2500 m

distance 1500 m power penalty becomes less steeper and stable, it represents less power deviation for all the modulation technique. From the above analysis, we can say that 1500 m is the optimum link distance. So, by analyzing all the parameters (BER, power loss, link distance) and Figures 7 and 12, we can consider 8-PSK as the optimum modulation technique, and 1500m is considered the optimum link distance.

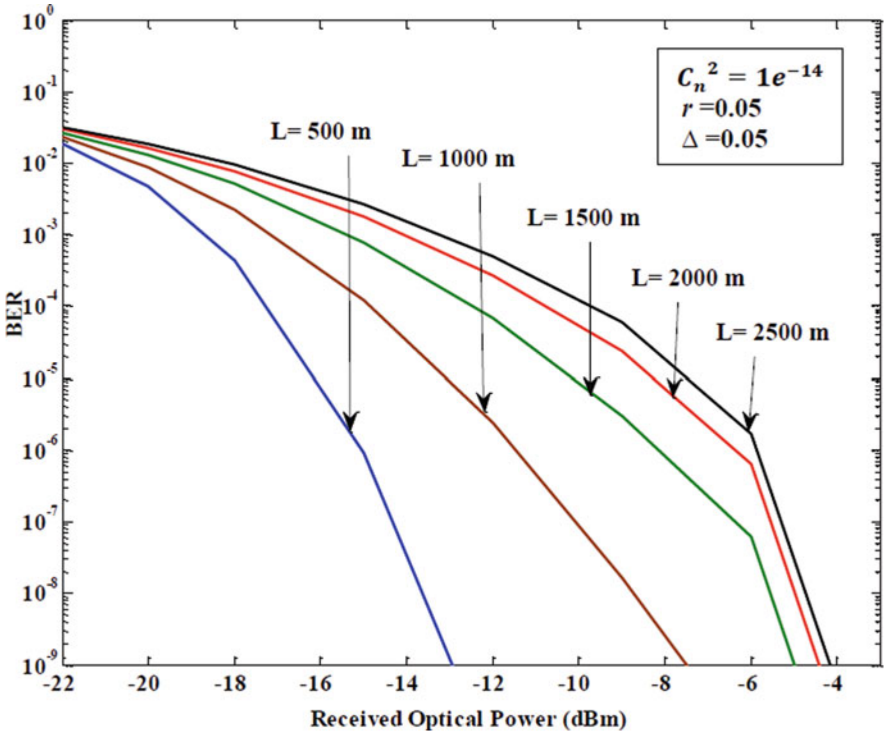


Fig. 10 BER vs. SNR of 8-PSK over variant link distances (L), where L varies from 500 m to 2500 m

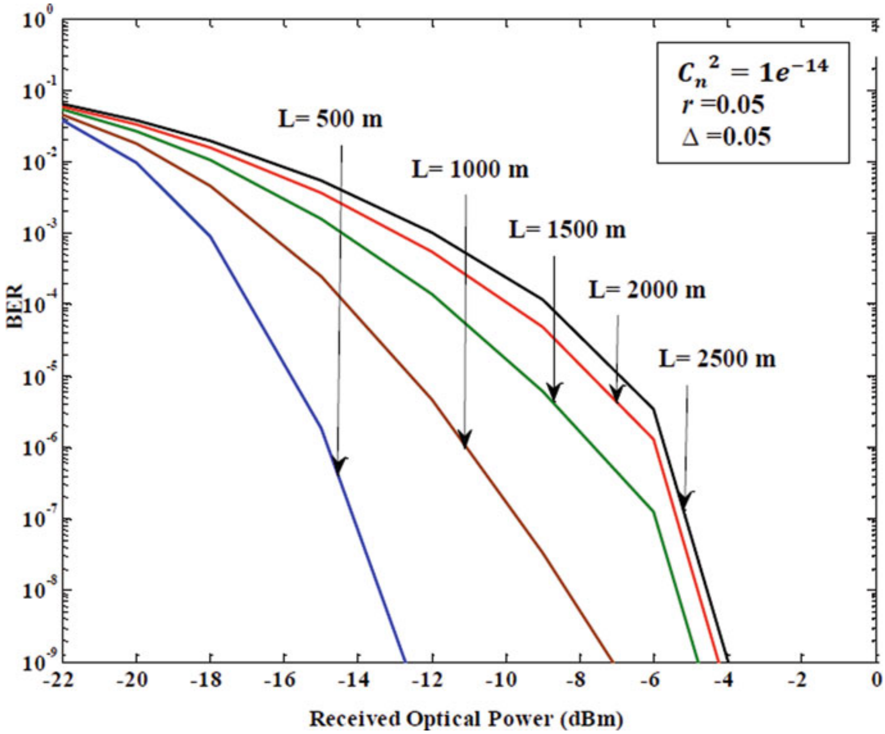


Fig. 11 BER vs. SNR of 16-PSK over different link distances (L), where L varies from 500 m to 2500 m

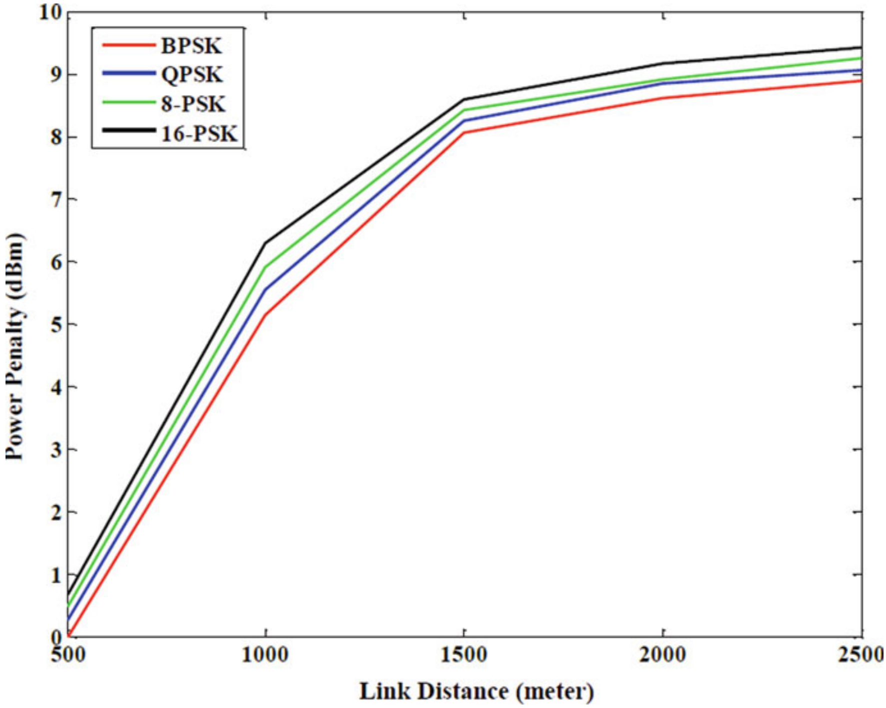


Fig. 12 Power Penalty vs. Link Distance, where L is from 500 m to 2500 m

References

1. M. A. Khalighi, M. Uysal, Survey on free space optical communication: a communication theory perspective. IEEE Commun. Surv. Tutor. **16**(4), 2231–2258, Fourthquarter 2014, https://doi.org/10.1109/COMST.2014.2329501
2. H. Hemmati, *Near-Earth Laser Communication*, CRC Press; ISBN-13: 978-0-8247-5381-8 (2008)
3. C. Deepak Kumar, Md Khaliluzzaman, Comparative performance of BER in the simulation of digital communication systems using raised cosine filter. in *Third International Conference on Advances in Computing, Electronics and Electrical Technology - CEET 2015*, https://doi.org/10.15224/978-1-63248-056-9-25, pp. 29–33 (2015)
4. A.C. Motlagh, V. Ahmadi, Z. Ghassemlooy, K. Abedi, The effect of atmospheric turbulence on the performance of the free space optical communications, in *Proceedings of the 6th International Symposium on Communication Systems, Networks and Digital Signal Processing*, pp. 540–543 (2008)
5. L.C. Andrews, R.L. Phillips, Laser beam propagation through random media, in *Washington: SPIE Optical Engineering Press* (1998)
6. L.C. Andrews, R.L. Phillips, C.Y. Hopen, *Laser beam scintillation with applications* (SPIE Optical Engineering Press, Washington, 2001)
7. H. Yuksel, Studies of the effects of atmospheric turbulence on free space optical communications, Dissertation for the Doctoral Degree. University of Maryland, pp. 76–78 (2005)
8. X. Song, M. Niu, J. Cheng, Error rate of subcarrier intensity modulations for wireless optical communications. IEEE Commun. Lett. **16**, 540–543 (2012)
9. X. Song, F. Yang, J. Cheng, Subcarrier intensity modulated optical wireless communications in atmospheric turbulence with pointing errors. J. Opt. Commun. Netw. **5**(4), 349–358 (2013)
10. S. Krishna, BER for BPSK in OFDM with Rayleigh multipath channel. Int. J. Sci. Adv. Technol. (ISSN 2221-8386), August 26 (2008)
11. T. Singh Hanzra, G. Singh, Performance of free space optical communication system with BPSK and QPSK modulation. J. Electron. Commun. Eng. (IOSRJECE) **1**, 38–43 (2012)
12. R. Ramirez-Iniguez, M. Sevia Idrus, Z. Sun, Optical wireless communications IR for wireless connectivity, (Taylor & Francis Group, Book, CRC Press, 2007) ISBN-13: 978-0-8493-7209-4
13. N.I. Othman, A.F. Ismail, M.K. Hasan, W. Hashim, K. Badron, Dynamic spectrum allocation scheme for heterogeneous network: BER analysis. J. Telecommun. Electron. Comput. Eng. (JTEC) **9**(3–10), 99–104 (2017)
14. S. Islam, A.H. Abdalla, M.K. Hasan, Novel multihoming-based flow mobility scheme for proxy NEMO environment: a numerical approach to analyse handoff performance. SCIENCEASIA **43**, 27–34 (2017)
15. M.K. Hasan, S.H. Yousoff, M.M. Ahmed, A.H. Hashim, A.F. Ismail, S. Islam, Phase offset analysis of asymmetric communications infrastructure in smart grid. Elektronika ir Elektrotechnika **25**(2), 67–71 (2019)
16. M.K. Hasan, A.F. Ismail, S. Islam, W. Hashim, M.M. Ahmed, I. Memon, A novel HGBBDSA-CTI approach for subcarrier allocation in heterogeneous network. Telecommun. Syst. **70**(2), 245–262 (2019)
17. S. Islam, O.O. Khalifa, A.H. Hashim, M.K. Hasan, M.A. Razzaque, B. Pandey. Design and evaluation of a multihoming-based mobility management scheme to support inter technology handoff in PNEMO. in *Wireless Personal Communications.* 2020 May 1.
18. M. Park, H. Jun, J. Cho, N. Cho, D. Hong, C. Kang. PAPR reduction in OFDM transmission using Hadamard transform. in *2000 IEEE International Conference on Communications. ICC 2000. Global Convergence Through Communications.* Conference Record 2000 June 18 (Vol. 1, pp. 430–433). IEEE
19. S. Sarowa, N. Kumar, S. Agrawal, B.S. Sohi, Evolution of PAPR reduction techniques: a wavelet based OFDM approach. Wirel. Pers. Commun. **115**(2), 1565–1588 (2020)

Toward Secure Cyber Infrastructure for Smart Cities: Learning-Based Intelligent Solutions

Shafkat Islam, Arpan Bhattacharjee, and Shahriar Badsha

1 Introduction

Smart city concept has revolutionized the traditional city settings through its multiple offerings in terms of mobility, electrification, producibility, among others. Data is considered as the backbone of this revolution as data-driven approaches have automated the traditional approaches in every sector. However, this also raises concern regarding data privacy and security that may become a pivotal factor in fast adopting such approaches in our daily lives.

Smart transportation, being an integral component of smart cities, is equipped with multi-verse sensors, e.g., GPS, dashcam, lidar, sonar, radar, etc., which can create a huge amount of data in every second of its operation. According to the study of automotive insights, it is expected that an autonomous vehicle (in the era of smart transportation) will generate more than 100 terabytes of data in every 8 h.[1] This huge amount of data possesses enormous insight regarding the transportation behavior of our cities that will be an asset for the smart transportation industry to improve different data-driven services. To understand the complex relationship among different factors of the transportation eco-system, the data will be required for further processing. Cloud computation is considered as a possible solution in this regard. However, offloading this huge data to the cloud will create gigantic traffic in the vehicular network, which may increase latency as well as traffic drop

[1] https://www.dxc.technology/auto/insights.

S. Islam (✉) · A. Bhattacharjee · S. Badsha
University of Nevada, Reno, Reno, NV, USA
e-mail: shafkat@nevada.unr.edu; abhattacharjee@nevada.unr.edu; sbadsha@unr.edu
https://www.unr.edu/cse

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_13

in the network. Moreover, the third-party cloud vendors may become curious about any sensitive information in the offloaded data, which can intercept the privacy of the vehicular users, thus demotivating vehicles to engage in smart transportation facilities.

In recent years, edge computing has become popular in smart transportation due to its manifold offerings, e.g., faster computation, low-latency service processing, data preprocessing, etc. Multiple research works [1, 2] have incorporated vehicular edge computing for providing faster and smarter services in the smart transportation industry. But the vehicular edge infrastructure is also subject to issues related to governance (as edge infrastructure is distributed in nature), security, and privacy.

The incorporation of artificial intelligence (AI) techniques, in the power grid, transforms it into an intelligent smart grid that brings powerful smart technologies for designing, simulating, and controlling the states of the grid in real time with an expert system for fault diagnostic and detection of cyberattacks in the modern smart grid (SG) equipped with distributed renewable energy sources (DREs).

The manufacturing industry with the incorporation of cutting edge technologies, i.e., artificial intelligence (AI), especially machine learning (ML) has introduced highly sophisticated tools to analyze the extensive amounts of production data originated from the interconnected cyber-physical system and detect countless system anomalies and provide real-time resiliency against those cyber as well as system anomalies.

2 Threats in Smart Transportation

To develop smart solutions for transportation, it is inevitable to understand the threats as well as drawbacks of transportation industry. Table 1 illustrates a typical network and data attacks related to the intelligent transportation system.

2.1 Byzantine Activity

In the era of smart transportation, vehicles as well as roadside units (RSUs) or other edge entities are expected to communicate with each other at a faster rate. This type of communication involves in message transfer to important information/data

Table 1 Cyber attacks on smart transportation network

Network attacks	Data attacks
DoS	EM poisoning
GPS spoofing	Location inference
Malware application	Location tracking
ECU malfunction	Image manipulation

transfer in a distributed manner without the presence of any centralized governance. In decentralized systems, byzantine attacks is a trivial type of vulnerability that means to disseminate false/forged information in the system to make the system collapse or dysfunctional. In the future smart transportation industry, it will be a great challenge to defend such byzantine attacks.

2.2 Inference Attack

As data propagation in the network increases, the privacy concerns regarding the data also increase. In vehicular network, sensitive information/data, e.g., credit card information, vehicle trajectory, vehicle functional information, propagates through the network and multiple entities can access the data. This raises the privacy concerns related to the owner (such as vehicle users or ride users) of such data. If an adversary can get access to this sensitive information (either by authorization or by violating authorization), he/she can breach the personal identifiable information (owner's location, taste, daily activity, etc.) related to the data owner.

2.3 Intelligent Solutions for Smart Transportation

Multiple research works have utilized AI for innovating smart solutions in transportation [3–5]. In this section, we describe the impact of artificial intelligence (AI) in overcoming the security vulnerabilities of smart transportation.

2.4 Artificial Intelligence (AI)-Enabled Vehicular Edge Resource Allocation

The resources of edge infrastructure in vehicular network are limited compared to its demand. This limitation raises concern regarding the fair utilization of the edge resources. In different words, it is of paramount importance to ensure fairness as well as transparency in allocating edge resources to the vehicular network. However, resource allocation in a dynamic environment (such as vehicular network) is not a trivial task, and it becomes intractable for humans to engage in such complex process.

In this regard, artificial intelligence (AI)-based automated process can become a viable solution. As AI algorithms can automate any repeating process without the need for any human intervention, the incorporation of AI in edge resource allocation can become a viable solution. Multiple AI algorithms such as reinforcement learning, deep reinforcement learning have been utilized for solving resource allocation issues in edge network. Using reinforcement learning, the edge devices allocate resources in a continuous feedback loop-based learning process. In deep reinforcement learning method, deep neural network is used as a function approximation method. Using these AI techniques, optimal allocation policy is adopted to maximize the resources in the edge network. This eliminates the need for any complex formulation of edge network as well as human intervention.

2.5 Federated Learning for Vehicular Edge Computing

In traditional machine learning model, data is accumulated in a central server that may become a source of data breaching for the adversary. Moreover, the multiple entities may become disinterested to share data to any central server due to such possibility. To avoid this type of vulnerability, federated machine learning has been explored by the research community. In federated machine learning, each entity trains a local model with its personal dataset and sends the model parameters to a central server for training the global model. In this way, each entity can avoid the possibility of sensitive data breaching while model training.

2.6 Transparency in Vehicular Edge Computing

Edge computing works like a black box for the end users. The end user is not aware of the computation or resource allocation process in the edge server. This raises the issue of transparency in the edge computing environment [6]. To solve this transparency issue in distributed systems, blockchain technology has emerged with the advantage of verifiability through a distributed ledger. In blockchain, each entity holds a copy of the ledger from which it can verify any transaction. Although blockchain provides the advantage of verifiability, non-repudiation, irreversibility, etc., it suffers from scalability, transaction verification process, and efficiency. Scalability issue arises from the consensus process in blockchain. Most of the traditional blockchain consensus mechanisms, i.e., proof of work (PoW), practical byzantine fault tolerance (PBFT), etc., suffer from either computation or communication overheads, which makes it unpluggable in the vehicular applications as most of the vehicular edge entities are resource-constrained devices. In this regard, AI-based mining node selection can become a viable solution. The AI model will predict the reputation value for each mining candidate and output the predicted mining node. Moreover, machine learning models can be used for setting up dynamic hyperparameters to optimize the transaction throughput.

2.7 Anomaly Detection in Vehicular Edge Computing

Adversaries try to intercept the edge network for multiple adversarial purposes ranging from collecting sensitive information to breaking the edge network. In recent years, anomaly detection in edge infrastructure has become an indispensable part to detect abnormal behavior of the functional system. Multiple machine learning methods have been deployed for detecting anomalies, i.e., deep neural network, random forest, support vector machines, ensemble learning, etc. However, designing an appropriate anomaly detector is still a challenge for researchers.

2.8 Cryptography-Based Security Solutions

Location privacy has become a pivotal factor in motivating end user vehicles in engaging data sharing applications. In the literature, multiple cryptographic solutions have been proposed for enabling vehicular data sharing in a privacy-preserving manner, i.e., secret sharing, homomorphic encryption [7], etc. However, these techniques impose additional overhead in the system, which makes the existing methods unpluggable in smart transportation industry in which the data sharing frequency is much higher. In this regard, low-cost privacy-preserving techniques, such as differential privacy, can be adopted. Although these techniques provide privacy in data sharing, the system utility reduces by significant level for using such techniques. Hence, a dynamic optimization technique is required to find the optimal trade-off point between the privacy cost and the system utility.

3 Artificial Intelligence (AI) to Enhance Security and Privacy of Advanced Manufacturing

Advanced manufacturing [8] is the trend toward converting the industries into an intelligent and interconnected networking of business processes with the aid of present-day information and communication technologies like the IIoT, sensors, CPS, data analytics, machine learning, and artificial intelligence. This transformation allows scalability, efficiency, interoperability, and automation to exceed industries productivity. However, introducing modern technologies in the manufacturing paradigm also raises security and privacy challenges like (1) cyber-physical malicious attacks on the IoT sensors and data to compromise the integrity of manufacturing process, (2) launch DDoS/Sybil attacks on the central controller-dependent data centers to cause service unavailability, and (3) data interception attacks on the supply chain system to compromise the credibility and confidentiality of the whole manufacturing process. Moreover, security and privacy concerns in smart manufacturing and its supply chain are mainly caused by internal attacks like

Table 2 Cyber threats in advanced manufacturing (AM)

AM components	Attack types
IoT sensors	Data fabrication
Central unit	DDoS/Sybil
Supply chain unit	Data leakage

about 60% of the attacks are launched by employees and third-party contractors through malicious data analytics to learn/steal confidential intellectual property and gain a financial incentive. So, the interconnected manufacturing approach while ensures major value in the production and distribution tasks; it also raises the exposure of the system to many security and privacy threats, with critical social and financial affects. Table 2 illustrates typical attack vectors in different components of advanced manufacturing system.

To tackle these challenges, researchers already conducted some work, for example, O. Novo [9] proposes a blockchain-based access control system to manage IoT devices. However, their reliance on a central management hub causes a single point of failure issue. Z. Li [10] exploits the consortium blockchain technology, but it cannot guarantee the confidentiality of sensitive data. To overcome the computational requirements of IoT devices, Z. Xiong [11] leverages edge computing for mobile applications to model an efficient edge resource management system. However, these existing solutions are not enough to solve the modern intelligent cyber-physical manufacturing industries security as well as privacy issues.

So, to safeguard the industrial IIoT sensor-enabled manufacturing systems and its data from unauthorized intrusion and ensure immutability and resiliency against data and command tampering, researchers are currently adopting artificial intelligence-based novel privacy-preserving defense mechanisms for the smart manufacturing paradigm. This AI-incorporated defense tool will empower the advanced manufacturing system by enabling the ability to (a) intelligently eliminate the need for trusted third parties and guarantee security against the single point of failure issues, (b) designing a privacy-preserving private data management mechanism to restrict the access to sensitive sensor data, (c) an intelligent supply chain system that decreases delivery time by eliminating third parties and increases the efficiency of stock management, (d) artificial intelligence-based data analytics and sharing platform among the involved parties of the manufacturing industry, which will help to predict future demands of the market as well as solving imminent problems in the value chain without compromising sensitive data.

4 AI for Smart Grid Security

With the unification of advanced telecommunication technologies (ICT) and IoT sensors in the complex cyber-physical smart grid [12, 13] paradigm leads to new security and privacy challenges such as (1) maliciously manipulating the power generation data and sensors to interrupt the grids optimal operation, (2) running query analytics on the aggregated sensitive data to identify the exact location, operation principle, and power generation capacity of different DERs to compromise their load-balancing capability, (3) identifying power consumption patterns of its associated consumers to maliciously use it for achieving financial incentive, and (4) injecting anomalies such as false power data to compromise the real-time situational awareness of the grid. To counter these security issues, current smart grid operators require an integrated and automated cyber-physical defense tool that can provide them (a) efficient intelligent privacy-preserving data aggregation ability so that data and its sources cannot be compromised by active/passive cyber attacks, (b) protect sensitive smart grid data from malicious query analytics, (c) run simultaneous anomaly detection on the confidential time series data to pinpoint the anomalies of the cyber-physical layer. To this end, we argue that a novel artificial intelligence-enabled, efficient, and privacy-preserving framework is needed for smart grid data sensing, aggregation, and analytics. This integrated defense tool will empower the smart grid paradigm by enabling the ability to (1) protect the highly confidential grid data from integrity compromising data and query analytic attacks, (2) improve the situational awareness by identifying anomalies in real time using AI for the utilities as well as (3) ensure preventive maintenance equipment at the cyber-physical layer of the grid.

In the presence of an untrusted fog/cloud aggregator in the cyber-physical smart grid, an artificial intelligence-enabled privacy-preserving data aggregation mechanism is required for both the utilities and their consumers to protect the sensitive data and its sources. However, the existing researchers only focused on the privacy protection of consumer side smart meter data and completely neglect the security importance of the power-generating grids [14, 15]. Moreover, their over-reliance on cryptographic techniques that are computationally intensive for resource-constrained grid devices raises security vs. budget trade-off [16]. Alongside that modern power grid also requires a cyber-physical-aware anomaly detection mechanism to identify the anomalies without revealing sensitive information about the training data. Over the last decades, several anomaly detection researches have been conducted based on edge computing, differential privacy (DP); however, they are limited to provide a comprehensive solution to protect sensitive information as well as identify attacks that cause anomalies [17, 18]. To this end, artificial intelligence gained some reputation to overcome the limitations of the above-mentioned mechanisms.

4.1 Privacy-Preserving Anomaly Detection Using AI

In this part, we focus on the importance for an AI-enabled privacy-preserving mechanism (as shown in Fig. 1) to facilitate efficient real-time anomaly identification on confidential time-variant power data by dividing the anomaly detection tasks between the cloud-based cyberspace and edge-level devices. Normally, when the

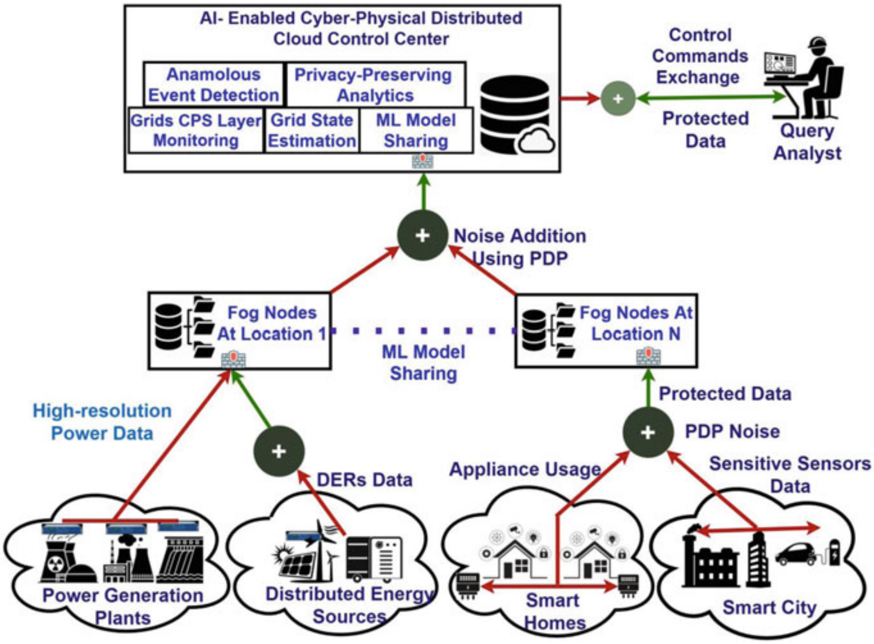


Fig. 1 AI-enabled privacy-preserving anomaly detection in smart grid

data is encrypted before transferring to the cloud, the edge devices are inadequate to process this cryptographed data effectively because of their limited computational power. To mitigate this limitation and achieve efficient and semantically secure protection, researchers can investigate artificial intelligence- and machine learning-based model sharing techniques that can be deployed at the central cloud server as well as its edge fog servers to perform the anomaly detection efficiently without revealing any privacy of the data. This solution can be achieved by incorporating personalized differential privacy to ensure the edge-level sensory devices' sensitive data privacy by incorporating privacy assurance based on each devices choice and an encrypted deep learning-based anomaly detection approach at the cloud layer. The high computation power of the central cloud server will help the AI models to run the highly secured privacy-preserving cryptography mechanism efficiently and ensure the highest possible privacy on the sensor data while performing anomaly detection.

5 Conclusion

In this chapter, we present the security vulnerabilities in three most inevitable components of smart cities such as intelligent transportation, smart grid, and smart manufacturing. We discuss different cyber attacks for each of these three sectors as well as the potential data-driven solution directions that can leverage the advanced AI algorithms. This chapter provides a substantial knowledge to the reader regarding the cyber-security challenges in smart city infrastructure.

References

1. S. Islam, S. Badsha, S. Sengupta, A light-weight blockchain architecture for v2v knowledge sharing at vehicular edges, in *2020 IEEE International Smart Cities Conference (ISC2)* (2020), pp. 1–8
2. S. Maskey, S. Badsha, S. Sengupta, I. Khalil, Reputation-based miner node selection in blockchain-based vehicular edge computing. IEEE Consumer Electron. Mag. **10**(5), 14–22 (2021)
3. S. Islam, A. Razi, A path planning algorithm for collective monitoring using autonomous drones, in *2019 53rd Annual Conference on Information Sciences and Systems (CISS)* (2019), pp. 1–6
4. S. Islam, Q. Huang, F. Afghah, P. Fule, A. Razi, Fire frontline monitoring by enabling UAV-based virtual reality with adaptive imaging rate, in *2019 53rd Asilomar Conference on Signals, Systems, and Computers* (2019), pp. 368–372
5. S.R. Maskey, S. Badsha, S. Sengupta, I. Khalil, Bits: Blockchain based intelligent transportation system with outlier detection for smart city, in *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)* (2020), pp. 1–6
6. S. Islam, S. Badsha, S. Sengupta, H. La, I. Khalil, M. Atiquzzaman, Blockchain-enabled intelligent vehicular edge computing. IEEE Network **35**(3) , 125–131 (2021)
   IEEE Netw. Mag. **35**, 125 (2021)
7. S. Badsha, X. Yi, I. Khalil, E. Bertino, Privacy preserving user-based recommender system, in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)* (2017), pp. 1074–1083
8. A. Bhattacharjee, S. Badsha, S. Sengupta, Blockchain-based secure and reliable manufacturing system, in *2020 International Conferences 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) and IEEE Congress on Cybermatics (Cybermatics)* (2020), pp. 228–233
9. O. Novo, Blockchain meets IoT: An architecture for scalable access management in IoT. IEEE Internet Things J. **5**(2), 1184–1195 (2018)
10. Z. Li, J. Kang, R. Yu, D. Ye, Q. Deng, Y. Zhang, Consortium blockchain for secure energy trading in industrial internet of things. IEEE Trans. Ind. Inf. **14**(8), 3690–3700 (2017)
11. Z. Xiong, Y. Zhang, D. Niyato, P. Wang, Z. Han, When mobile blockchain meets edge computing. IEEE Commun. Mag. **56**(8), 33–39 (2018)
12. M.T. Hossain, S. Badsha, H. Shen, PoRCH: a novel consensus mechanism for blockchain-enabled future SCADA systems in smart grids and industry 4.0, in *2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)* (2020), pp. 1–7
13. A. Bhattacharjee, S. Badsha, A.R. Shahid, H. Livani, S. Sengupta, Block-phasor: A decentralized blockchain framework to enhance security of synchrophasor, in *2020 IEEE Kansas Power and Energy Conference (KPEC)* (2020), pp. 1–6
14. S. Li, K. Xue, Q. Yang, P. Hong, PPMA: Privacy-preserving multisubset data aggregation in smart grid. IEEE Trans. Ind. Inf. **14**(2), 462–471 (2017)
15. D. He, N. Kumar, S. Zeadally, A. Vinel, L.T. Yang, Efficient and privacy-preserving data aggregation scheme for smart grid against internal adversaries. IEEE Trans. Smart Grid **8**(5), 2411–2419 (2017)

16. L. Lyu, K. Nandakumar, B. Rubinstein, J. Jin, J. Bedo, M. Palaniswami, PPFA: Privacy preserving fog-enabled aggregation in smart grid. IEEE Trans. Ind. Inf. **14**(8), 3733–3744 (2018)
17. S. Mehnaz, E. Bertino, Privacy-preserving real-time anomaly detection using edge computing, in *2020 IEEE 36th International Conference on Data Engineering (ICDE)* (2020), pp. 469–480
18. M. Keshk, E. Sitnikova, N. Moustafa, J. Hu, I. Khalil, An integrated framework for privacy-preserving based anomaly detection for cyber-physical systems. IEEE Trans. Sustain. Comput. **6**, 66 (2019)

Utilizing ICN Caching for IoT Big Data Management in WSN-Based Vehicular Networks

Divya Gupta, Shalli Rani, and Syed Hassan Ahmed Shah

1 Introduction

The rapid development in the automobile technology has changed the vision towards road vehicles. Nowadays, along with the basic property of vehicles to manage safe travelling between different stations, they are expected to support infotainment applications (i.e., information as well as entertainment), while self-driving is also a requirement of autonomous vehicles [19]. The Wireless Sensor Network (WSN) is a promising solution to achieve expectation based on the requirement of data. The WSN-based vehicular networks enable communication between different network entities where vehicles can communicate with each other through vehicle-to-vehicle (V2V) communication and with network through vehicle-to-infrastructure (V2I) communication. In addition, the communication between different network entities (such as Road Side Units (RSUs), Base Stations (BS), central server) can be achieved through infrastructure-to-infrastructure (I2I) communication [26]. Today, the vehicular network has proved its support for a variety of applications including road safety, mobile data services, traffic management, real-time emergency services, autonomous vehicles assistance, etc. Perhaps, with the increasing demand of mobile data services by passengers and road management services along with the rapid development of various entertainment sources, the amount of data generated, collected, and transmitted by vehicular network has shown an exponential growth.

D. Gupta · S. Rani (✉)
Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India
e-mail: divya.gupta@chitkara.edu.in

S. H. Ahmed Shah
Product Specialist, JMA Wireless, Liverpool, NY, USA
e-mail: sh.ahmed@ieee.org

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_14

Moreover, the network entities such as vehicles, RSU, BS have limited resource constraint in terms of storage and processing capability. Hence, the management of this big data generated by several IoT applications (such as social media, Youtube, safety applications, etc.) is a challenge in vehicular network environment [42]. In addition, all the requests in network have been given unique identification in the form of IP addresses. Providing and managing IP address to each request in huge data volume will ultimately suffer from IP address exhaustion and NAT problem. Based on this big data, current vehicular networks are in the need of some future network architecture that could efficiently handle all these network requests while maintaining system performance. Information-Centric Networking (ICN) is one such project proposed by Van Jacobson to solve future internet issues and is considered as a promising solution for all internet-related problems [3]. ICN supports content-centric communication where requests travel inside the network using content names instead of IP addresses. ICN supports several inherent features such as scalability, robustness, content security, in-network caching, name-based routing, and so on. The in-network caching feature of ICN supports caching of content by all network nodes. The content caching by network nodes is beneficial to support better resource utilization with fast content delivery time. Due to several advantages offered by ICN caching terminology, this feature can be integrated into vehicular networks where vehicles on the network can cache content and serve requests coming from their neighbour vehicles. Utilizing ICN caching in vehicular networks for management of IoT big data traffic will reduce burden on network server while enhancing the overall system performance. The following are the contributions made in this chapter:

– First, the background related to various technologies such as IoT, WSN and their integration has been discussed.
– Second, various challenges offered in the implementation of WSN-IoT environment are presented in brief with its possible solution.
– Further, the benefit of introducing ICN in WSN-IoT environment with its system architecture is presented in detail. In addition, various applications utilizing ICN in IoT have been discussed to provide insight about them.
– Next, caching in ICN has been utilized in vehicular environment to manage big data traffic produced by different IoT applications. The different caching strategies that are considered as ICN benchmark schemes have been mentioned to present working of these schemes.
– Lastly, the different metrics used for evaluating the performance of caching strategies in vehicular environment have been presented in detail with their computing equations.

2 Background

2.1 IoT: An Overview

Can we imagine the world without internet? The answer to this question is clearly No. Nowadays, the internet has become more and more important for each of us in our daily life [33]. All the day activities that either belong to personal tasks or professional work cannot be completed without internet support. Various devices such as laptops, smart phones, desktop, sensors, and other smart objects are essential components being used to accomplish these tasks. Such internet-related technologies have majorly affected new Information and Communication Technologies (ICT) [13]. During the birth of internet, it was known to be "internet of computers", in its early evolution referred as "internet of Persons", and now during rapid advancements in technology, it is termed as "internet of things (IoT)". The evolution in the internet since its birth till today has been illustrated in Fig. 1.

The IoT environment mostly comprises different smart objects capable of processing and transferring information and is uniquely identified in the network. The connection with internet in IoT has changed from any-one, any-person, any-where to any-thing, any-time, any-where. With the rapid growth in the recent times in ICT technology, a significant attention has been devoted to the emerging IoT due to its worldwide acceptance. Today, IoT is considered as one of the most important technologies for promising future demands. The main goal of IoT is to connect everything on the internet by establishing interaction between physical world and cyber world [34]. Without any doubt, IoT has its strength in establishing advance communication among different objects, smart devices, system, and services. IoT as a strong pillar for future internet has brought a revolution in the current architecture. This new IoT paradigm tries to connect everything starting from small tooth brush to large aeroplane and that too by assigning unique identification for each object.

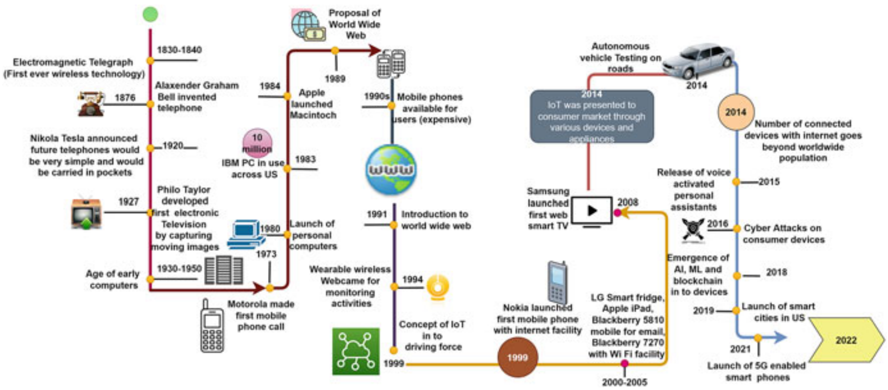


Fig. 1 Internet technology evolution

The unique identification would help in providing address for each object while connection management and exchange of information. Due to its wide support and advance features, no standard definition exists for IoT. IoT is being defined by different researchers in different way based on their perspective in use of IoT. The authors in [24] defined IoT as base for information transmission in internet and telecommunication network. According to them, IoT is used to connect intelligent objects with unique addresses and thereby perform services. However, the work in [10] mentioned IoT as interconnection of worldwide objects based on the semantics and protocols used in establishing the communication and for information retrieval. In actual, the ability of IoT lies in connecting various heterogeneous devices of different domains that include but are not limited to only sensors, everyday objects, smart devices, communication network, protocols, computations, and services that may differ based on its size, cost, design, application, vendor, etc. These entities in integration are able to communicate, sense, gather, process, and transfer information through different applications on the main data centres such as network clouds. This integration of small devices to form a computation network will ease the complex computations to be performed in collaboration and to make independent decisions.

2.2 WSN in IoT

The various IoT applications are using a million of devices for its implementation in real life. These devices in connection with each other and internet perform several computations and consume resources available on internet. Smart city is one of the most popular IoT applications that include various domains such as intelligent transportation, e-healthcare, smart education, smart grid, supply chain, smart industries, and so on [22].

IoT basically deals with three aspects to perform computation: One is to use sensors and RFID tags to sense, monitor, and gather information anytime at anyplace. Second is real-time gathering and transmission of information utilizing a communication network. Third is to incorporate some advance technologies such as cloud computing, fog computing, or edge computing to scrutinize the gathered information.

The key behind the IoT's large information database is WSN. In past few years, the research in microelectronics domain has become more advanced, which leads to the development of low-cost sensors with limited resources and small size. These tiny nodes with storage and processing capabilities are main building blocks of any IoT application. This integration of WSN in IoT (WSN-IoT) has opened doors for new technological advancements in different fields. WSN-IoT supports a broad range of applications such as health sector monitoring, industrial wireless networks, security monitoring at borders, home surveillance, environment monitoring, etc. [2, 36]. Different devices such as sensors, detectors, actuators, RFID tags, network components such as gateways, routers, Base Stations and various network support services in combination result into formation of WSN-assisted IoT network to

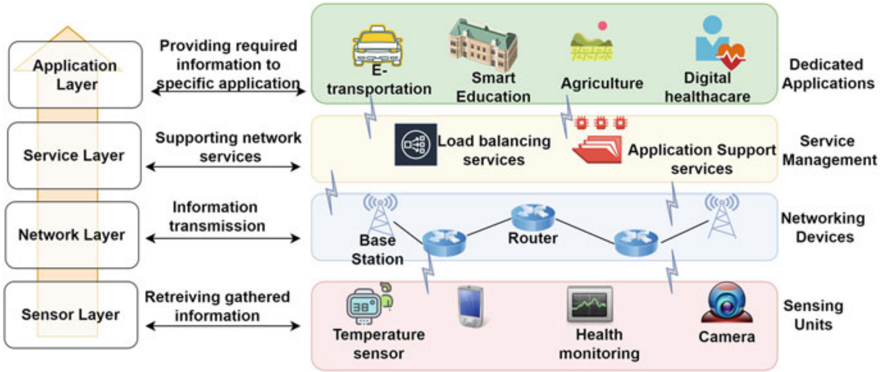


Fig. 2 Architecture of WSN in IoT

accomplish some common goals, which is then referred as WSN-IoT architecture (refer to Fig. 2).

The sensor layer in this architecture utilizes various physical components regarding attributes of deployed environment such as monitoring, hearing, sensing and then triggers an event while interacting with other such devices [16]. The collected information from different sensors is transmitted to Base Station via different routers using a defined routing protocol. The flow of information transmission among different routers in the network layer depends upon the routing and forwarding strategies being used in the network architecture. Most of the sensing devices in WSN-IoT network run on limited battery and storage constraints. In addition, for WSN-IoT applications deployed in dense and rugged environment, providing external storage space and computation resources is not feasible [32]. Therefore, the service layer in this architecture provides support to network services in the form of maintaining load balancing on all the nodes, adjusting the transmission flow, managing energy distribution, better utilizing network bandwidth, etc. The top layer of this architecture mainly comprises all the applications where end user is present to fetch the required information from network.

Due to frequent topology change and a large number of devices deployed in WSN-IoT environment, the existing WSN-based solutions cannot be directly applied to this architecture. Therefore, several challenges still exist in the implementation of WSN-assisted IoT environment in real life with dense conditions. The next section will focus on listing all such issues with their possible solutions.

3 Challenges in WSN-Assisted IoT Environment

In recent times, the tremendous growth in the IoT environment with its data collection through WSN network has opened new doors for various applications.

The integration of IoT and WSN supports automatic data processing even at remote stations. Perhaps, along with the several benefits offered by this integration, new challenges are being faced in their implementation due to add-on responsibility. The most important challenges that need to be addressed are:

- **Security:** The initial design of WSN has full potential to maintain confidentiality, integrity, and authentication of information being sensed and transformed through it even in the absence of internet connection. To attack WSN environment either in the form of inserting malicious nodes, jamming the network, or capturing the sensed information, the physical presence of attacker is mandate. The WSN ensures protection from such attacks by providing closed-loop network where no entry point exists for attacker to attack in. However, with the integration of IoT in WSN, the attackers can perform their malicious activities from everywhere due to opening of WSN into internet [39]. The issues such as malware and others offered due to connection with the internet need to be definitely addressed by WSN. The protection of the shared as well as sensed information within WSN can be ensured by providing a unique central gateway. All the information floating between IoT and WSN should be passed through this secured gateway only. In addition, the sensor nodes in WSN always possess limited resource constraints such as battery lifetime, storage space, processing capability, and so on [35]. The existing security solutions cannot be directly applied to WSN nodes due to their resource limitation. Therefore, to secure WSN from different kinds of internet possible attacks, there is a need to develop efficient security mechanism that can better work in these existing conditions.
- **Quality of Service (QoS):** The amount of data generated through these WSN-IoT-assisted environments must meet some quality parameters before their transmission to end user [7]. To ensure QoS, all the sensor nodes being utilized in WSN network need to be load balanced with their available resources and computations. Nevertheless, the dynamic network topology and intermittent connectivity make the existing QoS solutions inadequate in IoT environment. Hence, new approaches providing QoS solutions must be designed to support fast content delivery with minimal retransmission.
- **Network Configuration:** In addition to abovementioned requirements for efficient WSN-IoT implementation, managing network configuration is also very important task [30]. All the sensor nodes in WSN environment must be capable of managing their network configuration as several new nodes would be joining the network rapidly. Further, at the same time, several nodes exhausting their lifetime power would be leaving the network. The sensor nodes should handle such diverse conditions in addition to detecting and removing faulty nodes thereby ensuring scalable and reliable network configuration. As self-healing and configuration management is not the inbuilt feature of available sensor nodes, some add-ons need to be installed by WSN-IoT application user for smooth operation of this network environment.
- **Architecture:** The combination of WSN and IoT has enlarged the traffic on the internet. Each data being sensed by the sensor nodes travels towards IoT with

unique address. The WSN-IoT architecture supports assigning IP addresses for every request generated through WSN network [1]. With these large volumes of uniquely identifying requests, the network will soon suffer from these three issues in near future: address exhaustion, address management, and network address traversal problem. The current TCP-IP support utilized by WSN-IoT will no longer be sufficient for future requests. Therefore, either modification in the current architecture or some future internet architectures are required to support these emerging requests.

4 Leveraging ICN in WSN-IoT

The rapid growth in the IoT technology with its support in connecting various heterogeneous devices on the internet has also come up with various implementation challenges. The sensing units in the WSN-IoT architecture have limited storage and processing capability. Moreover, combining all the sensors, detectors, actuators and allowing all the small devices to establish connection to internet have increased the magnitude of the number of internet connected devices. Providing IP addresses to all the requests made by such IoT devices as well as managing these addresses is not feasible in current scenario. Several solutions have been proposed by different research groups to deal with these emerging technology issues. Among all the solutions, Information-Centric Networking (ICN) is proposed as one of the architectures for content-based networking [5, 6]. Unlike IP-based host-centric approach, ICN works on data-centric communication. The consumers in ICN network are concerned only about getting data irrespective of the location of data, i.e., from where data is being retrieved. ICN offers several inherent features such as scalability, reliability, content security, in network caching, stateful routing and forwarding, and light configuration with easy management operations [4]. The various ICN support features as per the requirements of WSN-IoT have been listed in Table 1.

Table 1 Mapping WSN-IoT requirements with ICN features

Sr. No.	WSN-IoT requirement	ICN support
1	Scalability	Hierarchical application-specific names
2	Robustness	In-network storage, anycasting, interest aggregation
3	Security	Data authentication, content integrity, per packet signature, encryption possibility
4	Reliability	Multi-path routing, retransmission of interest from actual consumer, intermediate node retry
5	Heterogeneity	Customized forwarding and caching strategies, unbounded namespace
6	Mobility	Receiver-driven connectionless communication, location-independent names, any node data retrieval
7	Energy efficient	Aggregation, in-network storage, anycasting

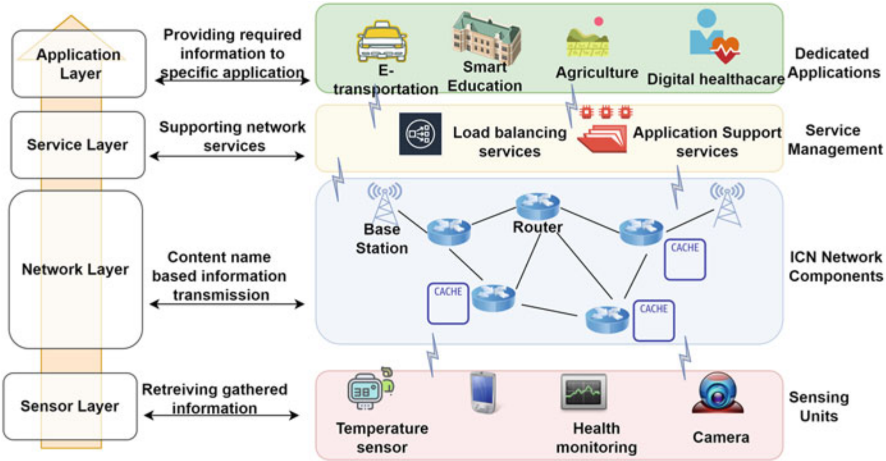


Fig. 3 Introduction of ICN in WSN-IoT

The integration of ICN in WSN-IoT architecture is shown in Fig. 3. ICN works on the network layer of IoT architecture. Each node in NDN carries limited storage capacity to cache passing content. The request for any content by any consumer travels inside the network in the form of hierarchical names. Each node in ICN manages three data structures to perform content routing and forwarding. The request and response messages are contained inside interest packet and data packet, respectively. Any intermediate node satisfying the requested content in its cache will deliver the content to consumer and therefore is known as the provider of that content. This feature makes ICN an efficient future network architecture with improved system performance and better resources utilization.

4.1 Applications of ICN-Enabled WSN-IoT

Nowadays, the IoT is being used by various applications [8, 15]. The huge amount of data sensed by WSN components gets processed and analyzed to meet IoT application demand. Enabling ICN facility within WSN-IoT helps in efficient and fast processing of content requests made by any IoT application. The different IoT application areas where ICN has proved its effectiveness have been listed in Fig. 4. The detailed description of all these applications is mentioned below:

– **Smart Home:** The evolution in IoT technology has turned traditional houses to intelligent homes [21]. Various home appliances such as refrigerator, washing machine, microwave, smart phone, lights, etc. are now connected to internet with communication and processing ability. With the introduction of new technologies such as edge computing, fog computing, etc., the monitoring of home can be

Fig. 4 Various applications of ICN-enabled WSN-IoT

performed remotely by the owner even though sitting at far distance from its house location.

- **Intelligent Transportation:** The traffic on the road is increasing each day. Therefore, managing the huge volume of traffic is need of the hour. The ICN in IoT supports traffic management using intelligent transportation system where Road Side Units (RSUs) can cache the passenger requests regarding road traffic, journey time, and parking availability [27]. The works proposed various strategies for efficient content delivery in vehicular networks.
- **E-Healthcare:** The advancements in IoT have granted real-time monitoring of patient's health using sensors and edge-computing techniques. The various smart devices associated in health sector have to go through authorization and authentication phase due to security and privacy concerns of patient's records [14].
- **Smart Agriculture:** IoT is one of the promising solutions for the agriculture domain. The sensing units deployed at different places in any field are capable of monitoring moisture value, water level, soil quality, and crop size [25]. The remote monitoring in agriculture domain saves a lot of efforts, money, and time of a farmer with improved results than traditional farming.
- **Smart Grid:** The power generation and distribution in grid system have been made automatic to be implemented in real time using IoT. As the amount of

energy consumed from these power grids is payment based, therefore, security and privacy to user transactions must be provided [29]. The use of blockchain technology in ICN-enabled IoT network for providing security is best solution in this regards.

5 ICN Caching for Traffic Management in Vehicular Networks

With the advancements in technology and its adoption in various applications, the interests have now been emerged to deliver content efficiently in vehicular network environment [23]. The vehicular networks mainly comprise components such as Road Side Units (RSUs), Base Stations (BS), intelligent vehicles, and On-Board Units (OBUs). The vehicles are equipped with OBUs that are capable of processing the information and help in transferring messages between vehicles and different RSUs [18]. The distribution of a huge volume of information among heterogeneous users (mainly passengers and drivers) based on their request under high vehicular mobility is a challenge. Moreover, the limited transmission range of each vehicle and RSU along with the intermittent connectivity in highly dynamic environment imposes further challenges in its implementation. The conventional host-centric IP-based approaches barely work in this scenario due to their end-to-end channel connectivity requirement [12]. The vehicular network with ICN caching facility is represented in Fig. 5.

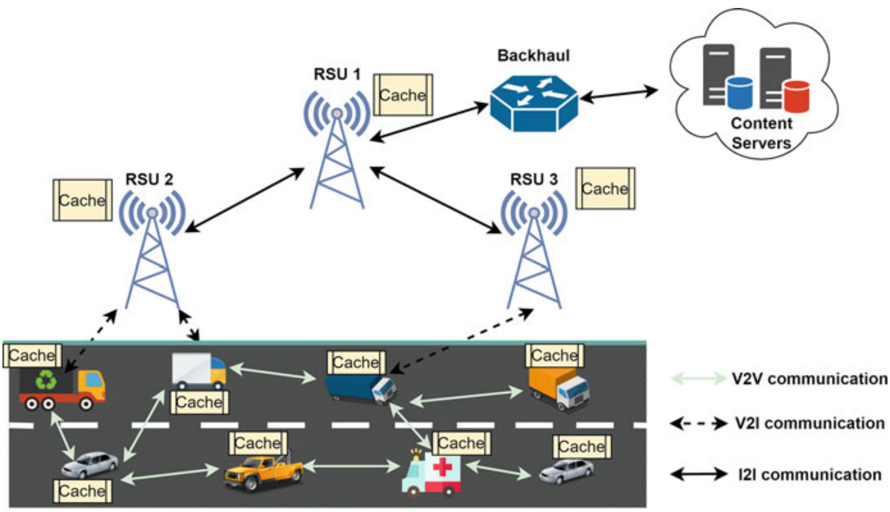


Fig. 5 ICN caching in vehicular environment

The inherent feature of ICN architecture that supports in-network caching of content by network nodes is best suited for achieving this goal. Here, each vehicle can cache a limited amount of content in its unused cache space. Any vehicle in demand of same content can get its replica from its neighbouring vehicle within its transmission range. This will reduce the content retrieval delay for user request as packet need not to be fetched from the original provider. The maximum benefit of caching feature can be availed at time when multiple requests for same content arrive on network and vehicles requesting contents are out of transmission range of BS (the original provider). In this case, caching content on network vehicles significantly improves the whole network performance by reducing burden on server, utilizing less bandwidth requirement, better resource utilization, and fast delivery of content [37]. The caching in ICN is beneficial not only for vehicular networks but provides its assistance to different domains with varying applications. The various caching strategies proposed for ICN irregardless of its use in any specific applications are listed in below subsection.

5.1 Different Caching Strategies

The various caching strategies proposed recently for ICN communication can be divided based on their properties: i.e. either to cache content chunk on the delivery path or to cache anywhere in the network [40, 41]. Based on these properties, the categorizations are: (1) on-path caching and (2) off-path caching.

The off-path caching is similar to Content Delivery Networks (CDN) approach where Name Resolution System (NRS) is queried upon receiving a request [38]. On the other hand, in on-path caching, the requested content is delivered by the node locating in the path of requested content without informing NRS. The off-path caching schemes are less preferred in the ICN architecture due to their limitations such as high content search time and poor utilization of network node's cache capacity. The most preferred on-path caching schemes that are known as benchmark caching strategies for ICN network are discussed below. The summary of these strategies is represented in Table 2.

– **Leave Copy Everywhere (LCE):**
  This scheme involves multi-level caches. For any node requesting a content, whether hit for that request occur at original server or any intermediate node, all the nodes in the path from requester to provider will cache the replica of content. LCE involves caching of same content by every node in the path. Therefore, whenever request for same content arises again, the content need not be fetched from the original server, but any node with its replica will help by providing content [9]. However, caching by each and every node creates a lot of redundancy in the network. Due to increased redundancy, all the caches will soon become full and there will be no space for new packets. The increased redundancy will

Table 2 Summary of the existing caching strategies

Cache strategy	Strength	Limitation
LCE	Increased content availability, reduced content access time	Enhanced content redundancy, poor resource utilization, Low cache diversity
LCD	Higher cache diversity, improved cache hit ratio	Less content availability, maximum bandwidth utilization
ProbCache	Enhanced content diversity, efficient resource utilization, lower content redundancy	Complex computation, additional packet overhead
Random	Efficient resource utilization, single replica on one path, reduced redundancy	Poor calculation to select a node, maximum load imbalance
CL4M	Minimal content redundancy, efficient node selection criteria, single replica maintenance	Difficult implementation in dynamic environment, hard centrality calculation

ultimately have reduced cache content diversity. Moreover, caching at every node offers poor utilization of available network cache space.

- **Leave Copy Down (LCD):**
  The LCD works better than LCE in terms of offered content redundancy. LCD ensures caching the content at exactly one node below the node on the path where actually hit for content request occurred. Each request for same content allows content replica to be get stored nearer the content requester. Unlike LCE, LCD ensures multiple requests for same content in order to store its replica at edge [17]. The level-wise caching of replicas by LCD significantly improves the cache space utilization as well as enhances the whole network's cache hit ratio. However, due to multiple rerequests for same content at same time, all the nodes may be caching that content within its cache as per LCD operation. By performing such action, all the caches will be occupied to its maximum possible limit, leaving no space for new upcoming contents. The LCD in this case may use Least Recently Used (LRU), Least Frequently Used (LFU), or any other cache replacement policies to create space for new data contents.
- **ProbCache:**
  The ProbCache is one of the famous caching policies for caching content based on probability. This scheme mainly deals with two fields such as Time Since Inception (TSI) and Time Since Birth (TSB). The former is the field used in the request message header during packet transmission, while the latter is being used in the response message header during its propagation towards requester [28]. These fields are basically used in this scheme to support efficient resource management of network caching capacity. In addition, this scheme works in the direction to provide maximum cached content diversity that in turn reduces content redundancy. Therefore, the probability of finding the requested content along the path becomes higher. However, due to frequent replacements of cached content inside nodes near to edge, sometimes the requested content needs to be

fetched from the original source as it was evicted from edge nodes. The increased hop count for packet fetching will lead to high content retrieval delay.

– **Random(Choice):**
  In this scheme, the requested content is cached randomly at any node on the path from requester to provider. This content caching scheme randomly chooses any node to store replica without knowing its cache availability. In case of no leftover space in the cache buffer of chosen node, the scheme applies any cache replacement algorithms either LRU, LFU, LRFU, etc. to accommodate this new data content [20]. The LRU ensures reduced redundancy of the content by selecting only one node to act as cache for a specific content. However, multiple requests for same content may lead to increase in network content redundancy as each request will allow caching of content on its path.
– **Cache Less for More (CL4M):**
  This scheme allows content to be cached at exactly one node along the path. The node selected for caching the content is one with highest Betweenness Centrality (BC) value. The BC of any node is computed as the maximum number of shortest routing paths going through it [11]. For the case where multiple nodes with same BC value are located, the node that is closer to requester is selected as caching node. As this scheme purely works on betweennness centrality computation of each node, it may be challenging to use this scheme in mobile ad hoc networks due to highly dynamic network topology. In addition, this scheme offers poor utilization of available cache space of each network node due to its caching at selected routers based on BC. Further to this, popular content needs to be evicted when cache of selected node becomes full. This will increase bandwidth utilization as popular evicted content now has to be fetched from the original source.

6 Evaluation Metrics of ICN Caching Strategies

The different metrics affecting the content retrieval time, user experience, and network performance are listed below [31]. These metrics have been used for a vehicular environment with total V vehicles.

6.1 Cache Space Utilization

It represents the total amount of data being cached in the whole network. This can be computed by combining the data stored at each vehicle v. To maintain uniformity among network nodes, each vehicle is provided with equal cache space c_v. The total of cache space of each vehicle represents the total network cache capacity and can be computed using equation:

$$Cache_{total} = \sum_{v=1}^{V} c_v, \tag{1}$$

where $Cache_{total}$ is the total network cache capacity, V is the total number of network vehicles, and c_v is the cache space provided to any vehicle v. The total cache space utilization can be computed as

$$Cache_{utilize} = \frac{\sum_{v=1}^{V} \frac{\sum_{r=1}^{R} c_{v,r}}{R}}{V}. \tag{2}$$

Here, r is the number of requests made by vehicle v from the total R available requests.

6.2 Network Delay

It denotes the total amount of time taken, i.e., t to satisfy all the content requests r and then to receive the data. For any vehicle v requesting content c while sending I interest packets, the network delay is computed as the total time taken from sending first interest packet to the reception of last content chunk. The following equation computes the delay:

$$Delay = \frac{\sum_{v=1}^{V} \frac{\sum_{r=1}^{R} t_{v,r}}{R}}{V}. \tag{3}$$

6.3 Hop Reduction

It represents the total number of hops used to fetch the content from cache stores of intermediate vehicles instead of being fetched from the original provider. For each request r made by vehicle v to obtain content c, the reduction in hop is computed as the ratio of the number of hops utilized to get c from intermediate vehicle to the total number of hops used while retrieving c from the original source. The computation of hop reduction ratio can be done using the below equation:

$$Hop_{reduction} = 1 - \frac{\sum_{v=1}^{V} \frac{\sum_{r=1}^{R} \frac{h_{v,i}}{h_{v,P}}}{R}}{V}. \tag{4}$$

Here $h_{v,i}$ is the number of hops used to get content from intermediate vehicle i and $h_{v,P}$ is the number of hops used to get content from the original provider P.

7 Conclusion

This chapter utilizes ICN caching facility to manage huge traffic produced by various multimedia and security applications in WSN-based vehicular network. To provide insights about the technologies used in accomplishment of this objective, a brief overview of both IoT and WSN is presented along with their integration. Further, the need of incorporating ICN in WSN-IoT-based vehicular environment is presented in detail. In the last, handling of big data traffic in the vehicular environment is facilitated through use of caching in ICN. The complete study has proved ICN caching as a best candidate for managing big data traffic issues of vehicular networks.

References

1. S.A. Al-Qaseemi, H.A. Almulhim, M.F. Almulhim, S.R. Chaudhry, IoT architecture challenges and issues: Lack of standardization, in *2016 Future Technologies Conference (FTC)* (IEEE, Piscataway, 2016), pp. 731–738
2. T. Alhmiedat et al., A survey on environmental monitoring systems using wireless sensor networks. J. Netw. **10**(11), 606–615 (2015)
3. M. Amadeo, C. Campolo, J. Quevedo, D. Corujo, A. Molinaro, A. Iera, R.L. Aguiar, A.V. Vasilakos, Information-centric networking for the internet of things: challenges and opportunities. IEEE Netw. **30**(2), 92–100 (2016)
4. D. An, D. Kim, ICN-based light-weighted mobility support in IoT, in *2018 27th International Conference on Computer Communication and Networks (ICCCN)* (IEEE, Piscataway, 2018), pp. 1–2
5. S. Arshad, M.A. Azam, M.H. Rehmani, J. Loo, Recent advances in information-centric networking-based internet of things (ICN-IoT). IEEE Internet Things J. **6**(2), 2128–2158 (2018)
6. E. Baccelli, C. Mehlis, O. Hahm, T.C. Schmidt, M. Wählisch, Information centric networking in the IoT: Experiments with NDN in the wild, in *Proceedings of the 1st ACM Conference on Information-Centric Networking* (2014), pp. 77–86
7. K. Bajaj, B. Sharma, R. Singh, Integration of WSN with IoT applications: A vision, architecture, and future challenges, in *Integration of WSN and IoT for Smart Cities* (Springer, 2020), pp. 79–102
8. S. Balaji, K. Nathani, R. Santhakumar, IoT technology, applications and challenges: a contemporary survey. Wirel. Pers. Commun. **108**(1), 363–388 (2019)
9. B. Banerjee, A. Kulkarni, A. Seetharam, Greedy caching: An optimized content placement strategy for information-centric networks. Comput. Netw. **140**, 78–91 (2018)
10. M. Botterman, Internet of things: an early reality of the future internet, in *Workshop Report*, European Commission Information Society and Media, vol. 15. sn (2009)
11. W.K. Chai, D. He, I. Psaras, G. Pavlou, Cache “less for more” in information-centric networks, in *International Conference on Research in Networking* (Springer, 2012), pp. 27–40
12. C. Chen, C. Wang, T. Qiu, M. Atiquzzaman, D.O. Wu, Caching in vehicular named data networking: Architecture, schemes and future directions. IEEE Commun. Surv. Tutorials **22**(4), 2378–2407 (2020)
13. L. Da Xu, W. He, S. Li, Internet of things in industries: A survey. IEEE Trans. Ind. Inf. **10**(4), 2233–2243 (2014)

14. B. Farahani, F. Firouzi, V. Chang, M. Badaroglu, N. Constant, K. Mankodiya, Towards fog-driven IoT ehealth: Promises and challenges of IoT in medicine and healthcare. Futur. Gener. Comput. Syst. **78**, 659–676 (2018)
15. A. Gaur, B. Scotney, G. Parr, S. McClean, Smart city architecture and its applications based on IoT. Procedia Comput. Sci. **52**, 1089–1094 (2015)
16. K. Guleria, A.K. Verma, Comprehensive review for energy efficient hierarchical routing protocols on wireless sensor networks. Wirel. Netw. **25**(3), 1159–1183 (2019)
17. D. Gupta, S. Rani, S.H. Ahmed, R. Hussain, Caching policies in NDN-IoT architecture (2020). Integration of WSN and IoT for Smart Cities, pp. 43–64
18. D. Gupta, S. Rani, S.H. Ahmed, S. Garg, M.J. Piran, M. Alrashoud, ICN-based enhanced cooperative caching for multimedia streaming in resource constrained vehicular environment. IEEE Trans. Intell. Transp. Syst. **22**(7), (2021)
19. J. Gwak, J. Jung, R. Oh, M. Park, M.A.K. Rakhimov, J. Ahn, A review of intelligent self-driving vehicle software research. KSII Trans. Internet Inf. Syst. **13**(11), 5299–5320 (2019)
20. Y. Han, R. Wang, J. Wu, Random caching optimization in large-scale cache-enabled internet of things networks. IEEE Trans. Netw. Sci. Eng. **7**(1), 385–397 (2019)
21. Y. Jie, J.Y. Pei, L. Jun, G. Yun, X. Wei, Smart home system based on IoT technologies, in *2013 International Conference on Computational and Information Sciences* (IEEE, Piscataway, 2013), pp. 1789–1791
22. Q. Jing, A.V. Vasilakos, J. Wan, J. Lu, D. Qiu, Security of the internet of things: perspectives and challenges. Wirel. Netw. **20**(8), 2481–2501 (2014)
23. H. Khelifi, S. Luo, B. Nour, H. Moungla, In-network caching in ICN-based vehicular networks: Effectiveness & performance evaluation, in *ICC 2020–2020 IEEE International Conference on Communications (ICC)* (IEEE, Piscataway, 2020), pp. 1–6
24. H.D. Ma, Internet of things: Objectives and scientific challenges. J. Comput. Sci. Technol. **26**(6), 919–924 (2011)
25. V.N. Malavade, P.K. Akulwar, Role of IoT in agriculture. IOSR J. Comput. Eng. (IOSR-JCE) **1**(13), 56–57 (2016)
26. J.P. Muñoz-Gea, P. Manzanares-Lopez, J. Malgosa-Sanahuja, J. Garcia-Haro, Design and implementation of a p2p communication infrastructure for WSN-based vehicular traffic control applications. J. Syst. Architect. **59**(10), 923–930 (2013)
27. S. Muthuramalingam, A. Bharathi, N. Gayathri, R. Sathiyaraj, B. Balamurugan et al., IoT based intelligent transportation system (IoT-ITS) for global perspective: A case study, in *Internet of Things and Big Data Analytics for Smart Generation* (Springer, Cham, 2019), pp. 279–300
28. I. Psaras, W.K. Chai, G. Pavlou, Probabilistic in-network caching for information-centric networks, in *Proceedings of the Second Edition of the ICN Workshop on Information-Centric Networking* (2012), pp. 55–60
29. A.I. Sarwat, A. Sundararajan, I. Parvez, Trends and future directions of research for smart grid IoT sensor networks, in *International Symposium on Sensor Networks, Systems and Security* (Springer, Cham, 2017), pp. 45–61
30. S.H. Shah, I. Yaqoob, A survey: Internet of things (IoT) technologies, applications and challenges, in *2016 IEEE Smart Energy Grid Engineering (SEGE)* (IEEE, Piscataway, 2016), pp. 381–385
31. S. Shailendra, S. Sengottuvelan, H.K. Rath, B. Panigrahi, A. Simha, Performance evaluation of caching policies in NDN—an ICN architecture, in *2016 IEEE Region 10 Conference (TENCON)* (IEEE, Piscataway, 2016), pp. 1117–1121
32. R. Sharma, S. Prakash, P. Roy, Methodology, applications, and challenges of WSN-IoT, in *2020 International Conference on Electrical and Electronics Engineering (ICE3)* (IEEE, Piscataway, 2020), pp. 502–507
33. D. Singh, G. Tripathi, A.J. Jara, A survey of internet-of-things: Future vision, architecture, challenges and services, in *2014 IEEE World Forum on Internet of Things (WF-IoT)* (IEEE, Piscataway, 2014), pp. 287–292

34. M. Stoyanova, Y. Nikoloudakis, S. Panagiotakis, E. Pallis, E.K. Markakis, A survey on the internet of things (IoT) forensics: challenges, approaches, and open issues. IEEE Commun. Surv. Tutorials **22**(2), 1191–1221 (2020)
35. M. Tellez, S. El-Tawab, M.H. Heydari, IoT security attacks using reverse engineering methods on WSN applications, in *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)* (IEEE, Piscataway, 2016), pp. 182–187
36. C.W. Tsai, C.F. Lai, A.V. Vasilakos, Future internet of things: open issues and challenges. Wirel. Netw. **20**(8), 2201–2217 (2014)
37. R. Ullah, M.A.U. Rehman, M.A. Naeem, B.S. Kim, S. Mastorakis, ICN with edge for 5g: Exploiting in-network caching in ICN-based edge computing for 5g networks. Futur. Gener. Comput. Syst. **111**, 159–174 (2020)
38. A. Vakali, G. Pallis, Content delivery networks: Status and trends. IEEE Internet Comput. **7**(6), 68–74 (2003)
39. A. Viswanathan, N.S. Shibu, S.N. Rao, M.V. Ramesh, Security challenges in the integration of IoT with WSN for smart grid applications, in *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)* (IEEE, Piscataway, 2017), pp. 1–4
40. G. Zhang, Y. Li, T. Lin, Caching in information centric networking: A survey. Comput. Netw. **57**(16), 3128–3141 (2013)
41. M. Zhang, H. Luo, H. Zhang, A survey of caching mechanisms in information-centric networking. IEEE Commun. Surv. Tutorials **17**(3), 1473–1499 (2015)
42. K. Zheng, Q. Zheng, P. Chatzimisios, W. Xiang, Y. Zhou, Heterogeneous vehicular networking: A survey on architecture, challenges, and solutions. IEEE Commun. Surv. Tutorials **17**(4), 2377–2396 (2015)

Integration of WSN and IoT: Its Applications and Technologies

Roopali Dogra, Himanshi Babbar, and Shalli Rani

1 Introduction to Wireless Sensor Networks (WSN)

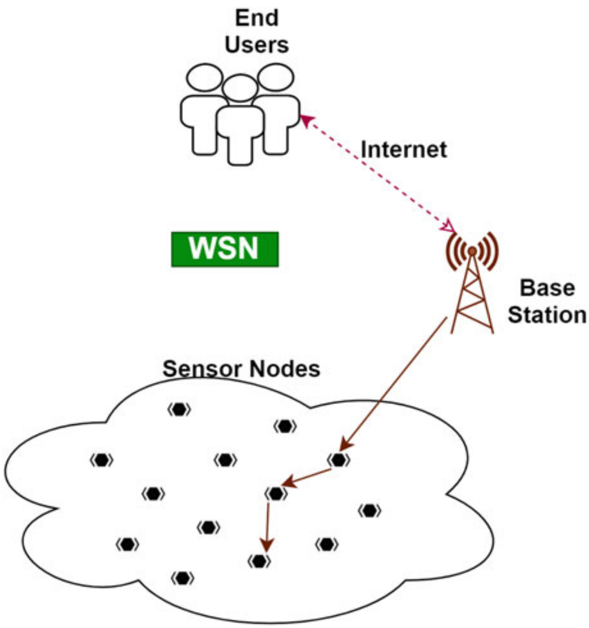
As the name suggests, wireless sensor networks are a class of networks where the nodes are sensor nodes, the nodes that sense and have the capability of sensing the massive number of sensor nodes where every node is furnished with a sensor to identify the physical phenomenon that occurs around them [15]. The particular sensor node might be the vibration, temperature that would be able to sense the pressure if there is any object or when there is any sound around the sensors that is moving around them, respectively [18]. Therefore, the sensor nodes have one of the components as a sensor, and these sensor nodes collectively form a network that is known as wireless sensor networks. These wireless sensor networks are accessible because of their numerous types of applications that can be used to monitor an object in a specific environment such as for medical purposes, healthcare, agriculture, etc [23]. WSNs are recognized as a groundbreaking information collection method for constructing an information and communication system that will significantly enhance infrastructure system performance and efficiency. WSNs have a simpler implementation process and more platform versatility than wired networks. WSNs would consider a crucial technology for IoT due to the growing technological growth of sensors and are considered important components for building smart cities. Figure 1 shows the WSN topology, which consists of sensor nodes for

Supported by Chitkara University.

R. Dogra · H. Babbar · S. Rani (✉)
Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India
e-mail: himanshi.babbar@chitkara.edu.in; shalli.rani@chitkara.edu.in

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_15

Fig. 1 Example of WSN



monitoring systems. The sensor nodes coordinate with one another and transmit the processed information to the base station through wireless communication. All of the nodes transmit information to the base station, which is then distributed to end users via the internet.

"Wireless Sensor Network (WSN) is an infrastructure-less wireless network that is deployed in a large number of wireless sensors in an ad hoc manner that is used to monitor the system, physical or environmental conditions."

1.1 Evolution of WSN and IoT

Since the 1980s, wireless sensor networks (WSN) are evolved, whenever the US Defense Advanced Research Projects Agency (DARPA) was in charge of the distributed sensor networks (DSNs) system for the army. The Advanced Research Projects Agency Network (ARPANET) had already been operational for many decades at the moment with around 230 hosts, but this is only since 2001 that they have sparked the attention of industry and academia [13, 15]. That was due to the extensive development of low-power tiny components such as computers, antennas, and sensors, which were often combined on a single chip (system on a chip (SoC)).

The concept of the IoT arose concurrently with the creation of WSNs. Kevin Ashton developed the word "internet of things" in 1999 that describes the distinctly recognizable objects and their digital representations in an "internet-like" frame-

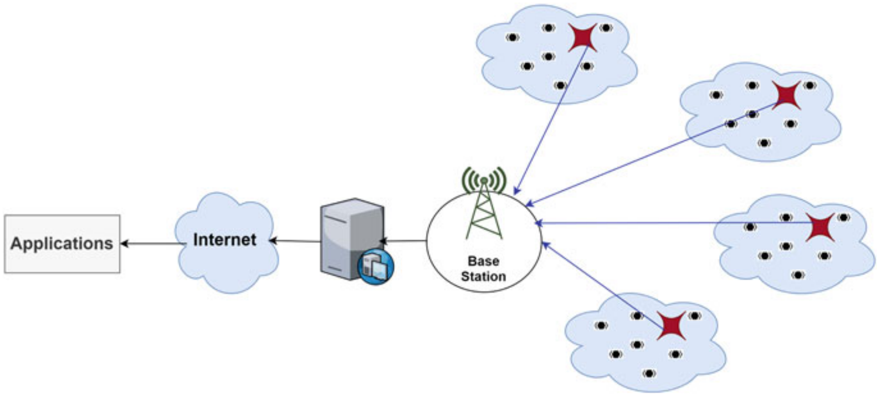


Fig. 2 Architecture of WSN-IoT

Table 1 Comparing WSN and IoT

WSN	IoT
Although nodes are not directly linked to the internet, traffic is routed through them to reach the sink node	Since they have an internet connection, sensors transmit their data straightly to the internet
Sensors are one type of device that collects information in the WSN	Sensors, people, cameras, and phones are all examples of things. These devices will all upload their data over the internet so that other people can access it

work [19]. Large houses, manufacturing facilities, aircraft, vehicles, computers, all types of goods, individual parts of a broader structure human beings, animals, and plants, and individual body parts of them are all examples of these artifacts.

In the given Fig. 2 the WSN uses sensor nodes in conjunction with an onboard processor to control and track the environment in a specific region. They are linked to the Base Station, which serves as the WSN System's processing center. A WSN system's base station is linked to the internet to exchange data.

WSN is considered as the subset of IoT. IoT exists higher than the WSN; therefore, WSN is often the technology used within the IoT system. Table 1 shows the contrast between WSN and IoT.

1.2 Growth of Wireless Sensor Networks

From 2020 to 2027, the global manufacturing wireless sensor network market is anticipated to achieve a CAGR of 25.2%, between Dollars 4560.2 trillion in 2020 and Dollars 9780.5 trillion in 2027. WSN is a network infrastructure that allows sensor nodes and gateways to communicate without the use of fiber cables. It also leads to improved communication according to radio nodes organized in the right topologies. As a result, the appetite for WSN has already been rising in recent years, but this growth is projected to grow in the future.

Due to rising network infrastructure growth and developments in Artificial Intelligence (AI), Machine Learning (ML), and big data analytics, the WSN market is anticipated to increase dramatically soon. Organizations have been forced to examine vast quantities of data gathered from a variety of sensors, including temperature, motion, pressure, gas, flow, and chemical, among many others, as per these technologies.

1.3 Components of WSN and IoT

A wireless sensor network is a collection of different sensor nodes, and these sensor nodes wireless networks are comprised of various components:

1. **Sensor Node:** This is the prominent component of WSN, in which the sensor node is a tiny device with low power [20]. Though it has a sufficient amount of resources, it has the feature of processing and has the least cost. Data collection, information collection, and communication with other network nodes are all capabilities of this node. A unique sensor node with a 4×10^{-6} GB RAM, 256×10^{-6} GB flash, and ideally 1054 MHz radio frequency has a capacity of 8–16 MHz.
2. **Relay Node:** It is a node in the middle that communicates with the node next to it. It aims to improve the reliability of the network. A relay node is a field research device that lacks process sensors and control devices and therefore does not communicate with the seven processes. The speed of the processor for the relay nodes depends upon the node processor which is around 8 MHz, with 8×10^{-6} GB of RAM, 128×10^{-6} GB of flash, and probably 916 MHz of radio signals.
3. **Actor Node:** It is an elevated node that performs and builds a decision based on the application's specifications. These nodes are usually resource-rich machines with huge computational capacities, higher transmitting powers, and longer battery life. A prominent actor node processor with a clock speed of 8 MHz, 16×10^{-6} GB of RAM, 128×10^{-6} GB of flash, as well as radio signals of 916 MHz is preferred.
4. **Cluster Head:** In a WSN, it is an increased sensing node that performs data fusion and consolidation. There may be several cluster heads within the cluster, depending on the device specifications and applications. A prominent cluster head processor has a clock speed of 4–8 GHz, 512×10^{-6} GB of RAM, 8 MB of flash, and a radio signal of approx. 3.4 MHz. This node is presumed to be extremely reliable, stable, and accepted by all of the sensor network's nodes.
5. **Gateway:** A gateway is a device that connects sensor networks to external networks. In terms of main memory, CPU used, transmitter range, and the ability

to expand via external memory, the gateway node outperforms the sensor node and cluster head. A unique gateway processor has a clock speed of around 16 MHz, 512×10^{-6} GB RAM, 64 MB of flash memory, radio signals of approx. 3.4 GHz.

6. **Base Station:** It is a specialized form of a node with a lot of computing power and processing capacity.

2 Applications of WSN in IoT

The term wireless sensor network (WSN) is referred to as the collection of specialized sensors for tracking and measuring environmental conditions and storing the captured data in a centralized location [16]. The measurements of WSN are namely: temperature, sound, humidity, wind, and other environmental factors. There are numerous applications in WSN as depicted in Fig. 3:

- **Home Control Application:** Control, conservation, comfort, and protection are all provided by home control applications:
  - Sensor networks allow for versatile illumination, warming, and conditioning system control from any place in the house.

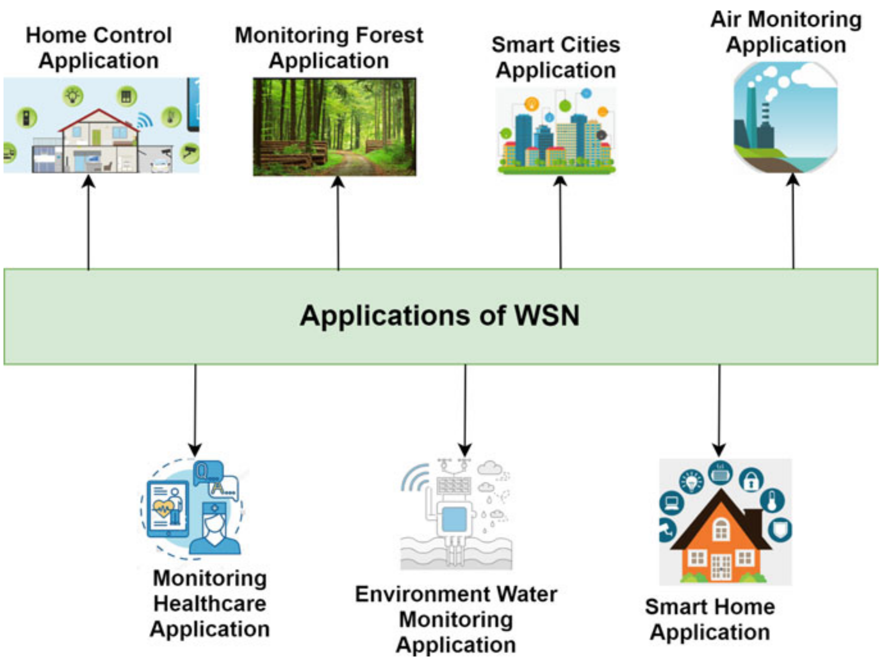


Fig. 3 Applications of WSN-IoT

- Sensor networks optimize access to multiple home devices, resulting in greater efficiency, comfort, and protection.
- Utility consumption data is captured in great detail by sensing applications for power, water, and gas utilities [11].
- Sensor networks incorporate information to maximize natural resource use.
- Sensor networks make it possible to mount, upgrade, and network a home control device without the need for wires.
- Sensor networks allow you to monitor numerous devices from a unified remote.

• **Monitoring Health Care Application:** For medical applications, there are different categories of sensor networks implanted, wearable, and environment-embedded. Implantable medical equipment is those that have been embedded into the human body [3, 14]. Wearable devices are placed on a human's body surface or in close vicinity to the consumer [4]. Sensors incorporated in the atmosphere are used in ecosystem systems. Body position assessment, individual location, and comprehensive monitoring of ill patients in the hospital and at homes are all application areas. Data from a network of deep cameras, a sensing floor, or other similar devices is used as input by devices incorporated in the environment to monitor a person's physical state for consistent health diagnosis. Body-area networks can gather information on a person's health, fitness, and energy use. The safety and integrity of user data are critical in healthcare applications. With IoT, the authentication process becomes much more difficult, particularly with the integration of sensor networks.
Monitoring the Forest Application: The role of the forest to human life and the earth's natural cycle of the atmosphere are inextricably linked. Afforestation has become increasingly important as a result of urbanization. Supervision of the forest is also expected, which necessitates monitoring. WSNs are often used in the aforementioned region to ensure protection shown in Fig. 4.
Given the expected climate change, observing the microclimate in the forest has become critically important [17]. Individuals with an in-depth understanding of the natural, biochemical, and biological status of soil in tree root systems contribute significantly to the intensive observation of forest ecosystems. Such long-term research is needed for the development of forest management steps. The wireless transmission of sensor readings has no impact on forest operations. Rather, as opposed to a wired sensor installation, wireless transmission guarantees robustness.
Environment Water Monitoring Application: Water, whether for consumption or use in the ocean, is a vital component of human life, so water surveillance is of strong academic importance to researchers.
The authors created a cyber-physical system (CPS) that is an in-pipe system for water surveillance that uses RFID (Radio Frequency Identification)-WSN for such a WSN application to assess the condition of fresh drinking water [11]. The network will facilitate data on water demand and quality, as well as information on numerous maintenance issues namely weak spots and leakage of pipe. The

Fig. 4 Forest monitoring

RFID sensors of in-pipe gather data from the device and transmit it to the data servers, in which algorithms help make decisions. An application of WSN used for managing the marine environment is described. To avoid disruption of a fish farm's flora and fauna origin of feed and feces, the authors created an Underwater WSN (UWSN) with ground-based wireless sensor nodes having the capability of detecting the farm's emissions. To evaluate a larger region, the sensor nodes are portable in a space that is restricted.

Smart Cities Application: The control of different parameters in metropolitan areas is critical for citizens' optimal living conditions. WSNs, therefore, provide a wide range of applications for providing real-time data to authorities to ensure a city's optimum operation. Growing people transportation, in particular, causes difficulties and wastes time when a wide number of vehicles are traveling to the same place. WSNs may be used to track traffic to minimize congestion, show car parking spaces, and so on.

- **Smart Homes Application:** Numerous systems can enhance human lives in the age of informatics. WSN can be used in the indoor world, such as in smart homes, to communicate with machines. Indoor localization and motion control, as well as indoor air quality monitoring, are two common examples [9]. With the increasing population phase that is occurring, surveillance of indoor air quality (IAQ), which is a concept that contributes to the quality of air in the house, is one of the most relevant aspects of urban applications in WSN. Since urban people invest so much of their time inside buildings, air quality is crucial to their health, protection, and comfort.

- **Air Monitoring Application:** Numerous cities (Stockholm, London, and Brisbane) have installed wireless sensor networks to detect hazardous gas concentrations for people. These may use ad hoc wireless connections instead of wired connections, making them more mobile for checking readings in various locations. Air is a critical component of human life, and today's air pollution is a product of numerous newer human activities. WSNs may be used to regulate air quality in populated areas to avoid infectious pathogens and contaminations from spreading and endangering people's health.

3 Technologies in WSN-IoT

3.1 Software-Defined Networking (SDN) in WSN

The WSN comprises a base station (BS) and tens or hundreds of sensor nodes. Temperature sensors, snowstorm sensors, and other sensor nodes have different types and specifications depending on their use. In most cases, a massive network cannot function effectively without some form of organization. As a result, we suggest clustering the network. Every cluster in the network has one cluster head while all the other nodes in the cluster are the basic sensor nodes. The CH requires managing the activity of the sensor nodes in each cluster. The information gathered regarding the ecosystem on the cluster by nodes will be transmitted to the CH using this clustering method.

A logically centralized controller is required for a software-defined sensor network. The controller may not have to be a standalone node from the perspective of the network. The control logic should be incorporated based on the base station, according to our proposal. The sensor nodes in a software-defined sensor network do not want to undertake routing decisions. Rather, they transmit packets to another cluster head or base station based on the routing table that the base station generates [2]. Figure 5 shows to put it another way, the controller (in the base station) determines which routes are indeed the greatest based on application-specific parameters. The main objective of integrating the SDN [6] and WSN is that the SDN controller manages the CH nodes and forwarding strategies wholly rely on the controller, link between the nodes is controlled, and the forwarding path can be changed and adjusted by the nodes of the controller. There are two main stages carried: Network Building and Data Gathering [5].

– **Network Building:** In the first stage, since the controller does not know the network, the nodes must transmit their state information to the controller namely their name, location, residual energy, and email address, and enroll themselves with the controller.

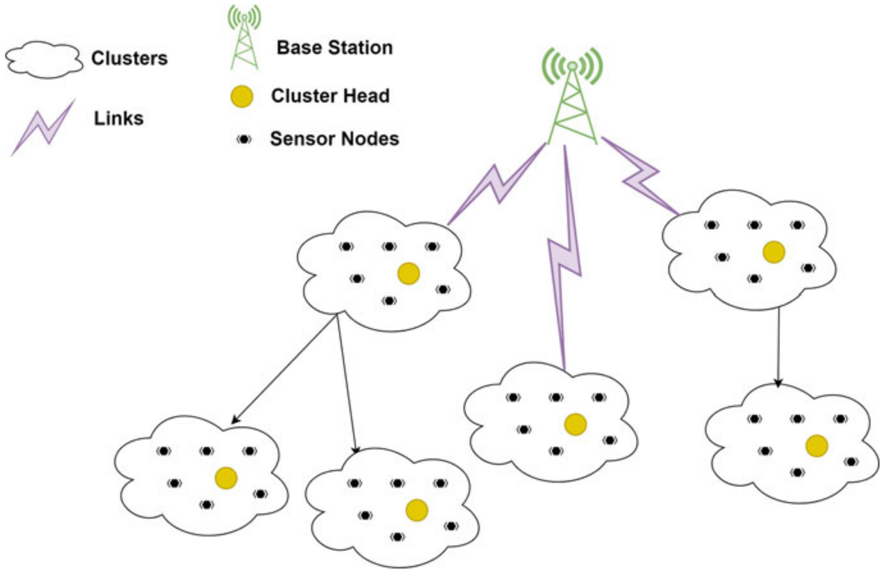


Fig. 5 Software-defined wireless sensor networks

In the second stage, the controller performs three actions based on the information obtained. The controller firstly splits the network into the uniform clusters; secondly, the controller chooses a CH for every cluster based on the node's location; thirdly, the controller produces the routing table for all of the CHs, informing the CHs of the next path for transmitting data to the BS.

In the third stage, the cluster outcomes and routing tables are forwarded to the CHs by the controller. The id of CH and the id of members in the cluster are included in the cluster outcomes. Lastly, the CH declares its own identity to the sensor nodes in a cluster, and the sensor nodes must react to the CH for the CH to recognize the cluster's state and respond to the controller afterward [7, 22].

– **Data Gathering:**
  Firstly, as per distinct needs, users set unique trigger conditions for distinct sensor nodes. Secondly, the BS sends the user's configurations among all sensor nodes, and each sensor node, therefore, configures itself. The sensor node begins sending data to its CH once the condition is met. The CH gathers and merges the data before passing it on to the following hop. Since the data transmitted by the sensor node is appropriate for the user's needs, the CH assigned by the controller's routing table does not need to align the sensor node data with the flow table. We inject the FlowVisor into the BS to achieve the target because the user can use unique controllers for various applications. The data from sensor nodes to CH contains not only the collected data but also the residual energy of the sensor nodes themselves.

3.2 Cloud Computing in Wireless Sensor Networks

WSN was configured to gather information in the actual world, while cloud computing (CC) is configured to be data-centric and provides powerful interaction with the external world [10]. When the gathered information is no longer useful, many companies face a problem. At a certain stage in the future, these data could be used for statistical, advanced analysis, or other purposes. The architecture's goal is to facilitate data transmission from the WSN to the cloud computing system so that critical data can be completely exploited. Using WSN and cloud computing technologies, a realistic, accessible, and dependable Agricultural Environment Management System (AEMS) was designed.

The **benefits** of using WSN and cloud computing are quite obvious that these two technologies can be merged to permit for sharing and examining the sensor data in real time. Furthermore, over the internet [1], this merging permits for facilitating the sensor data so that the data sensed can be examined properly not locally but anywhere in the world. Therefore, merging of these two technologies offers some benefits with the massive number of distinct applications, namely:

1. **Transport Surveillance:** For traffic management, electrically operated plate recognition, tollway management, emergency vehicle warning, adaptive traffic signal management, and other applications. Sensor data gathered by the cloud service will enable users to create an overall traffic analysis that could be dynamically modified [12]. These data may be used for a variety of purposes, including vehicle classification, braking systems, and automated toll gates, among others.
2. **Weather Prediction:** Environmental sensor data is an illustration of a so-called big data problem that cannot be easily handled utilizing conventional database methods rather can be viably solved utilizing the cloud's limited vast computing capacity.
3. **HealthCare:** Sensor networks have been commonly used in a variety of healthcare applications, not just in hospitals [8]. In general, there is still growing progress in promoting personal and portable devices for ongoing health surveillance and, as a result, improved life quality.

3.3 Named Data Networking (NDN) in Wireless Sensor Networking

The recently developed Named Data Networking (NDN) framework, which is a content-centric networking framework, is considered one of the most important in the scientific community. NDN communicates several principles with data strategies for WSNs, which have been suggested since the initial 2000s the basic principle is to separate sensed data from nodes' identity [21].

The advantages of using NDN in WSN include:

WSNs can find NDN to be a viable solution. Likewise, its capabilities are well suited to the use case scenarios and applications that are built on the top of sensors, with their respective possible limitations.

- **Easy to access the data:** In large-scale WSNs, structured naming enables information exploring and extracting can prepare the data consolidation easier.
- **Expandability:** The framework is especially useful for retrieving data from various nodes in a surveillance environment. NDN avoids data quality deterioration as the number of concerned nodes grows by exploiting the broadcast wireless medium and implementing lightweight techniques of forwarding.
- **Caching:** The content cached at multiple nodes (depending on their storage capacity) can be accessed even when communication is unreliable (e.g., due to low-power operation).
- **Development of application:** Applications can requisite for the data in a content-centric manner irrespective of the information's physical location, making it accessible to them.

4 Research Issues in the Integration of WSN and IoT

Although WSNs are better suited for use as an incorporating technology for IoT-based applications, there are a few problems to recognize that are outlined below:

1. **Protection:** There are a few safety concerns with sensor nodes. To avoid physical attacks on configured sensors especially for military applications, the sensors must be inconspicuous to attackers, which is difficult to accomplish because the only alternative is to minimize the size of the sensors, limiting their resources. The self-organizing structure of sensor nodes can enable some of them to act intentionally, causing WSNs to malfunction. Currently, the alternative to safety-related problems is to adopt cryptographic solutions such as encryption, but the challenge is difficult due to the sensor's low energy and storage.
2. **Efficiency of Energy:** As previously stated, the design of the area where the sensors are mounted (e.g., no-land, man's border areas, etc.) can prevent the batteries from being recharged. As a result, developing energy-efficient algorithms for WSNs is indeed an area under research.

   Quality of Service (QoS) specifications: As with Web or conventional wired architecture, enforcing QoS requirements for WSNs used by several applications is still a work in progress. Sensor nodes may not have been able to meet the rigid QoS specifications associated with many actual IoT applications due to their flexibility and changing network topology.

5 Conclusion

The internet of things (IoT) is a new concept that seeks to integrate the smart physical devices in such a way they can deliver intelligent services to the users. Wireless sensor networks (WSNs), which are linked to the internet of things (IoT), are beneficial networks for tracking, detecting, and sensing various environmental activities. Sensors are crucial in the design and implementation of any WSN. This chapter aims to develop the integration of WSN and IoT in a single domain. Discussed the evolution and growth of WSN and IoT in the real world, later we described the applications of WSN and IoT in different areas, and then the technologies used in the integration namely SDN with WSN, Cloud with WSN, and NDN with WSN explain how these different technologies integrated to achieve the basic objective. Lastly, the research issues/challenges are discussed arose in the integration of WSN and IoT.

Key Points

1. Wireless sensor networks (WSN) are a class of networks where the nodes are sensor nodes, the nodes that sense and have the capability of sensing the massive number of sensor nodes where every node is furnished with a sensor to identify the physical phenomenon that occurs around them.
2. Wireless sensor networks (WSNs) have been there since 1980s, whenever the US Defense Advanced Research Projects Agency (DARPA) was in charge of the distributed sensor networks (DSNs) system for the army.
3. The concept of the IoT arose concurrently with the creation of WSNs. Kevin Ashton developed the word "internet of things" in 1999 that describes the distinctly recognizable objects and their digital representations in an "internet-like" framework.
4. WSN is considered as the subset of IoT. IoT exists higher than the WSN; therefore, WSN is often the technology used within the IoT system.
5. From 2020 to 2027, the global manufacturing wireless sensor network market is anticipated to achieve a CAGR of 25.2%, between Dollars 4560.2 trillion in 2020 and Dollars 9780.5 trillion in 2027.
6. A wireless sensor network is a collection of different sensor nodes, and these sensor nodes wireless networks are comprised of various components.
7. For medical applications, there are different categories of sensor networks implanted, wearable, and environment-embedded. Implantable medical equipment is those that have been embedded into the human body.
8. The role of the forest to human life and the earth's natural cycle of the atmosphere are inextricably linked. Afforestation has become increasingly important as a result of urbanization.

9. A logically centralized controller is required for a software-defined sensor network. The controller may not have to be a standalone node from the perspective of the network.
10. The main objective of integrating the SDN and WSN is that the SDN controller manages the CH nodes and forwarding strategies wholly rely on the controller, link between the nodes is controlled, and the forwarding path can be changed and adjusted by the nodes of the controller.
11. The benefits of using WSN and cloud computing are quite obvious that these two technologies can be merged to permit for sharing and examining the sensor data in real time. Furthermore, over the internet this merging permits for facilitating the sensor data so that the data sensed can be examined properly not locally but anywhere in the world.
12. NDN communicates several principles with data strategies for WSNs, which have been suggested since the initial 2000s the basic principle is to separate sensed data from nodes identity.
13. WSNs are better suited for use as an incorporating technology for IoT-based applications, and there are a few problems to recognize namely protection, energy efficiency, and quality of service.

References

1. K. Ahmed, M. Gregory, Integrating wireless sensor networks with cloud computing, in *2011 Seventh International Conference on Mobile Ad-hoc and Sensor Networks* (IEEE, Piscataway, 2011, December), pp. 364–366
2. A. Ali, Y. Ming, S. Chakraborty, S. Iram, A comprehensive survey on real-time applications of WSN. Future Internet **9**(4), 77 (2017)
3. A.A.A. Alkhatib, G.S. Baicher, Wireless sensor network architecture, in *International Conference on Computer Networks and Communication Systems (ICNCS 2012)*, vol. 35 (2012, April), pp. 11–15
4. T. Arampatzis, J. Lygeros, S. Manesis, A survey of applications of wireless sensors and wireless sensor networks, in *Proceedings of the 2005 IEEE International Symposium on, Mediterranean Conference on Control and Automation Intelligent Control* (IEEE, Piscataway, 2005, June), pp. 719–724
5. H. Babbar, S. Rani, Software-defined networking framework securing internet of things, in *Integration of WSN and IoT for Smart Cities* (Springer, Cham, 2020), pp. 1–14
6. H. Babbar, S. Rani, Performance evaluation of QoS metrics in software defined networking using Ryu controller, in *IOP Conference Series: Materials Science and Engineering*, vol. 1022, no. 1 (IOP Publishing, 2021), p. 012024
7. H. Babbar, S. Rani, Software-defined networking based on load balancing using Mininet, in *Proceedings of the Second International Conference on Information Management and Machine Intelligence* (Springer, Singapore, 2021), pp. 69–76
8. H. Babbar, S. Rani, M. Masud, S. Verma, D. Anand, N. Jhanjhi, Load balancing algorithm for migrating switches in software-defined vehicular networks. CMC-Comput. Mater. Continua **67**(1), 1301–1316 (2021)
9. O. Cetinkaya, O.B. Akan, Use of wireless sensor networks in smart homes, in *Emerging Communication Technologies Based on Wireless Sensor Networks: Current Research and Future Applications* (2016), pp. 233–258

10. W.Y. Chung, P.S. Yu, C.J. Huang, Cloud computing system based on wireless sensor network, in *2013 Federated Conference on Computer Science and Information Systems* (IEEE, Piscataway, 2013, September), pp. 877–880
11. M.P. Durišić, Z. Tafa, G. Dimić, V. Milutinović, A survey of military applications of wireless sensor networks, in *2012 Mediterranean Conference on Embedded Computing (MECO)* (IEEE, Piscataway, 2012, June), pp. 196–199
12. A. Flammini, E. Sisinni, Wireless sensor networking in the internet of things and cloud computing era. Procedia Eng. **87**, 672–679 (2014)
13. O.I. Khalaf, B.M. Sabbar, An overview on wireless sensor networks and finding the optimal location of nodes. Periodicals Eng. Nat. Sci. **7**(3), 1096–1101 (2019)
14. J. Ko, C. Lu, M.B. Srivastava, J.A. Stankovic, A. Terzis, M. Welsh, Wireless sensor networks for healthcare. Proc. IEEE **98**(11), 1947–1960 (2010)
15. M. Kocakulak, I. Butun, An overview of Wireless Sensor Networks towards the internet of things, in *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)* (IEEE, Piscataway, 2017, January), pp. 1–6
16. S.J. Ramson, D.J. Moni, Applications of wireless sensor networks—A survey, in *2017 International Conference on Innovations in Electrical, Electronics, Instrumentation and Media Technology (ICEEIMT)* (IEEE, Piscataway, 2017, February), pp. 325–329
17. B. Rashid, M.H. Rehmani, Applications of wireless sensor networks for urban areas: A survey. J. Netw. Comput. Appl. **60**, 192–219 (2016)
18. C.C. Sobin, A survey on architecture, protocols and challenges in IoT. Wirel. Pers. Commun. **112**(3), 1383–1429 (2020)
19. S.G.H. Soumyalatha, Study of IoT: understanding IoT architecture, applications, issues and challenges, in *1st International Conference on Innovations in Computing & Networking (ICICN16), CSE, RRCE. International Journal of Advanced Networking & Applications (No. 478)* (2016, May)
20. A.J. Swati, R. Priyanka, Wireless sensor network (WSN): Architectural design issues and challenges. Int. J. Comput. Sci. Eng. **2**(9), 3089–3094 (2010)
21. K. Tetsuro, S. Tetsuya, A Study on Implementation of NDN to WSN, in *2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA)* (IEEE, Piscataway, 2017, March), pp. 392–398
22. S. Wadhwa, H. Babbar, S. Rani, A survey on emerging software-defined networking and blockchain in smart health care, in *IOP Conference Series: Materials Science and Engineering*, vol. 1022, no. 1 (IOP Publishing, 2021), p. 012056
23. X. Xu, H. Zhang, T. Li, L. Zhang, Achieving resilient data availability in wireless sensor networks, in 2018 *IEEE International Conference on Communications Workshops (ICC Workshops)* (IEEE, Piscataway, 2018, May), pp. 1–6

Choice-Based Recreation Facility for Smart Cities

Ramalatha Marimuthu, Bindu A. Thomas, Namratha DCruz, and Aawatif Hayar

1 Introduction

Smart cities are the ones which use technology to provide services according to the activities and tastes of the citizens to improve the quality of life. This is done through making informed and better decisions by acquiring data from numerous sensors and developing systems to analyse them in various aspects to provide a comprehensive knowledge on the behaviour of people. Businesses flourish, quality of life improves and, consequently, the people are happier. But these do not come without compromises and challenges. The diversity in the nature of residents, their differing needs and requirements and available technology all play an important role in deciding the feasibility of service implementation in a smart city.

According to the McKinsey Global report, cities are occupied by more than 50% of the population now, with a steady increase towards an additional 2.5 billion new residents by 2050. These increasing populations increase environmental pressures and infrastructure needs, and to provide a better quality of life, it is essential to have real-time data [1]. Smart cities are the only solution since they provide better living conditions and are more responsive.

R. Marimuthu (Retired) (✉)
Department of ECE, Kumaraguru College of Technology, Coimbatore, Tamil Nadu, India

B. A. Thomas · N. DCruz
Vidya Vikas Institute of Engineering and Technology, Mysore, Karnataka, India

A. Hayar
University Hassan II of Casablanca, Casablanca, Morocco

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9_16

The essential areas where smart cities are expected to provide better performance are:

1. Healthcare
2. Transportation and mobility
3. Security
4. Services and utilities
5. Economic development and citizen engagement

Out of this, citizen engagement covers the facilities provided by the city administration to engage the citizens, whether it is education or entertainment. Smart cities nowaday come with packaged recreation centres and serve a diverse set of citizens. The engagement required for a child will be different than that required for an adult or an elderly person. The same can be applied on all aspects of developing a smart city service. The difference also can be found in the requirements of diverse groups based on gender, economy, education and culture. This chapter discusses the various types of recreation to be built in a smart city environment and provides a technology for on-demand, choice-based recreation facility.

2 What Is Recreation?

Recreation has been considered an essential part of human life from the stone age to current age. Even walking, which is a day-to-day activity, becomes a recreation if it is carried out without a destination in mind and with different outcomes than just reaching a place. The recreation can be physical or cognitive but both bring out similar benefits. Recreation is a great stress buster, and activities like hiking, trekking, jogging and other games not only help in increasing the metabolism and physical health but also increase mental wellness. They relieve the mental fatigue which sets in through our regular responsibilities and works and rejuvenates us, thus creating happiness. Without recreation, life would be dull and miserable.

Humans are social beings and they live in communities. Recreation provides a great way of increasing social relationships, introducing inclusivity and empathy and, thus, paving the way for cultural tolerance. In an increasingly complex world, more and more importance is placed for recreational activities, which can be perceived by the highly successful theatre industry, gaming industry, severely competitive international games and races and greatly enjoyable theme parks. When we think about the outcomes, in addition to the customers, we should also consider the livelihood these industries provide to the millions of people working for them.

If we look at the history of recreation activities, mobility has been one of the most chosen recreation activity. People who go for excursions visiting attractive destinations use various modes of mobility, such as railways, bicycle, automobile, aeroplane, etc. This recreation activity had a direct impact on the economic growth of the geographical location and leads to the growth of the tourism and hospitality industry.

It was reported that there were British rail companies that could exclusively put special trains for excursions in earlier days. The surge in travel activities had a great impact on the increasing demand for consumer products as people would travel to places as part of recreation and shop for souvenirs or go to restaurants for delicacies unique or popular to the destination.

The above points show that recreation is something which can be a central component to ensure the health of individuals as well as a community. We find that gated communities often house a recreation centre with indoor and outdoor games for children and adults alike. So from being a nice-to-have amenity, it has become a must-have in communities. For this reason, instead of simple non-technical games, technology is now creating a plethora of games for children and adults alike.

2.1 Types of Recreation

Recreation falls into two broad categories –active and passive recreation. Active recreation involves our direct participation, like playing games, walking, hiking, trekking, reading, etc. Passive recreation is something we do not take part in but watch to be entertained, like movies, plays, nature, etc. This chapter discusses the various types of recreations, impact of technology on the old recreation activities and the recent increase in technology-based games and recreations and its impact on the economy and society and the challenges in developing and using them.

Apart from being broadly categorized into active and passive recreation, the activities can further be divided based on various factors:

- *Nature of outcome:* Relaxing, learning, keeping physically fit, product making
  Sample activities: Reading, watching plays and movies, visiting museums, hobbies like collecting antiques, stamps and coins, walking, trekking, swimming, painting, candle making, wood working
- *Number of people involved:* Individual, family and friends, teams from various places
  Sample activities: Skiing, surfing, ice skating, dancing, family games, sports activities
- *Nature of activity:* Indoor and outdoor, organized and unorganized
  Sample activities: Playing with children and pets, gardening, parties, banquets, picnics
- *Requirement of tools:* Specific tool requirements, technology requirements, no tools or technology
  Sample activities: Digital games using phone apps; team sports like cricket, hockey, etc.

Figure 1 shows the types of recreation with sample activities and their interconnections. Though there are many overlaps in the above categories, these cover the activities that society has understood as recreation. The general outcomes of all these activities can be towards the benefit of the individual's physical and mental fitness,

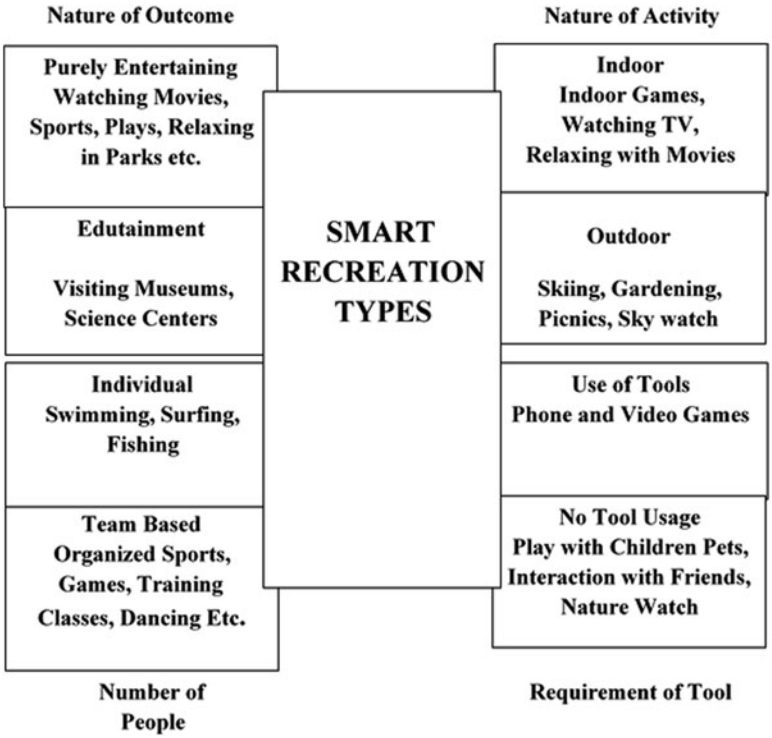


Fig. 1 Types of recreation in a smart city environment

but they also contribute towards societal well-being and economic development. While recreation activities reduce stress and improve confidence and creativity, they also provide commercial value. For example, hobbies like art and craft making can produce commercial value products, thus providing a business for the person making them. Sports activities and outdoor recreations are definitely increasing the economy of the country by attracting tourists and visitors to the area. Movie making and theme parks are becoming high investment businesses with excellent returns.

While a smart city is expected to have monitored pay-and-use operations like smart waste management, smart transportation and smart information management systems to cover the requirements of the residents, it is also important to plan the city with pay-and-use recreational facilities like smart parks, smart science and technology museums and smart outdoor activity areas. This requires infrastructure – not only physical infrastructure but also social, economic and institutional infrastructure. The residents are to be educated on the usage, maintenance and safety measures of the smart systems, which is essential to sustain the systems.

2.2 Evolution of Technology in Recreation

Smart city recreations can be classified into home entertainment and outdoor entertainment – both individual and team based. Recreation which first impacted the home leisure is the musical entertainment. In the 1920s, when radio came to the commercial market, it was brought as a home entertainment system. It was monitored and controlled by a household person, usually the head of the family, and the recreation was limited to specific timings. But this type of homogenous experience was not relished by the younger generation who wished to have control of their own recreation. The same thing happened with television sets which brought about the audiovisual experience.

But on demand, entertainment was the need of the hour in a world starting to move from joint family structure to nuclear family structure, owing to the movement of people from their native places to pastures new in search of jobs. Technology also has grown so much that on demand, choice-based and point-to-point entertainment is possible. Digital technologies have transformed home-based leisure into digital leisure and have changed the interactions between household members and other friends in the immediate circle [1]. In addition to increasing the exposure to different experiences, the homogeneity has decreased and control-based heterogeneous on-demand experiences are possible [2].

Reading is an important leisure activity, especially among the older generation. Technological advances have brought about ereading tools like Amazon Kindle, which took the world by storm when it was introduced in 2007. Similarly, the parallel advancements in hardware development has increased the use of these tools. Touch screen tablets and e-readers have changed the way the reading activity was perceived by people for both education and leisure purposes [3].

Cheng Cheng et al. prove that online leisure activities lead to marketing opportunities based on the behaviour of the customers – those who shop online without initial intention. This proves that online leisure activities help companies increase their sales [4]. Laura Rojas Francisco suggests that a study on these leisure activities can lead to the development of new digital technologies for social leisure activities and new devices for running these devices since the nature of the demand and the leisure differs from person to person [5].

Höjer and Wangel explain two different approaches to build a smart city using the technologies available to sustain the smart city services –top-down approach and bottom-up approach. The top down approach is the one where we take the available technologies and provide services, and the bottom-up approach is more customer oriented where the requirements are first thought out and the technologies are applied and developed as required. Their view is that the interconnectivity plays an important role in making a city smart [6].

3 Business Opportunities: Smart Tourism

Tourism is an important business deciding the economy of various destinations. Smart tourism has become an accepted concept in this digital era since the users now expect to learn about the place they are visiting and try to extract as much experience as possible within a short time. Connectivity is the core of smart city, and smart tourism draws the attention of people to the different attractions of the city through connectivity-based information communication technologies (ICT). The technology-based availability of the real-time information increases the possibility of business opportunities which rely on tourism, like services and goods, transportation and entertainment. The various businesses that can be enhanced through new opportunities created by a smart city are sales of household and individual goods, taxi and other means of public transportation, restaurants and hotels and edutainment centres like museums, zoos, theme parks, amusement parks and a whole other gamut of related small businesses. They also provide job opportunities for thousands of people, increasing real estate and production businesses.

Tourists have heterogeneous preferences, and many of the smart cities which concentrate on tourism have been trying to provide cultural immersion for the tourists by creating customized travel experience with comfort [7]. For example, Anaheim is identified with Disney World, and Dubai and Singapore are preferred as shopping destinations. Governments of countries like China, Japan and Australia are heavily funding the rapid development of new technologies to showcase their iconic tourist attractions and thus create an enriched tourism experience. Ulrike Gretzel et al. say, this leads to competitiveness and collaboration among the technology developers, service providers and infrastructure developers [8].

Technology-based recreation: To provide for the increasing urban population and visitors to the city, a city has to manage and integrate various city services and improve the quality of life. People from rural areas migrate to cities in search of jobs and income. Though most of the cities have e-tourism as a service, smart tourism is yet to evolve. E-tourism requires only the information and communication technologies and the connectivity, while smart tourism requires sensor-based real-time optimization to provide the required immersion experiences for the visitors. Here the infrastructure has to be developed with both hardware and software. Here the difference is providing value additions during the visit rather than the pre-visit and post-visit experiences, which is important for e-tourism.

According to the International Telecommunication Union of the United Nations, globally almost 85% of the population has access to 4G networks while 93% of the population has access to a mobile broadband network in 2020 [9]. But even now e-tourism has been flourishing all over the world only in developed countries. The important difference between e-tourism and smart tourism is that smart tourism uses real-time data from the sensor to give the best immersion experience for the user while e-tourism uses only the database and the connectivity to create the experience.

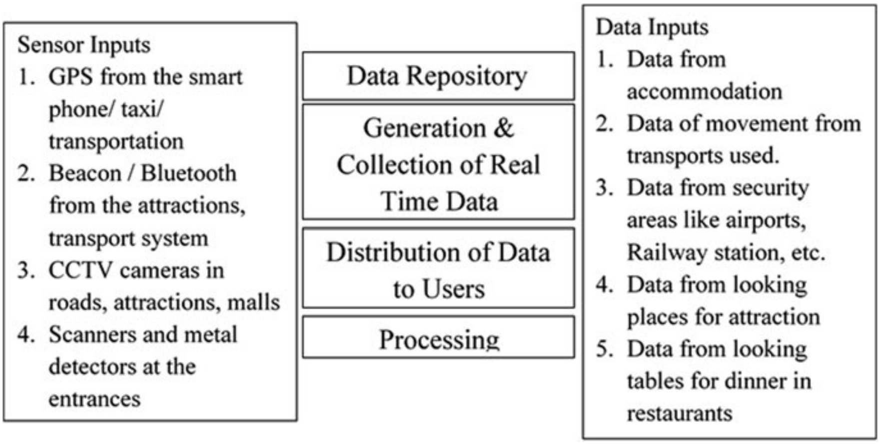


Fig. 2 Smart tourism process

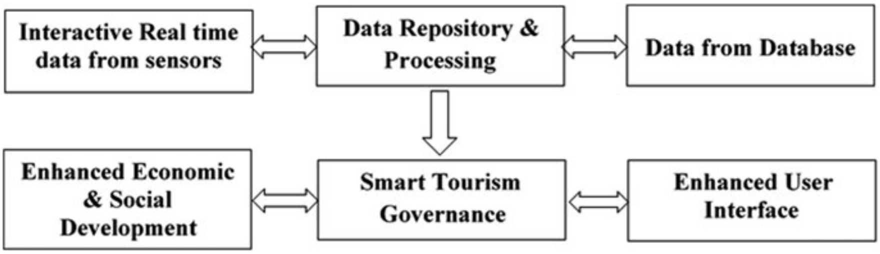


Fig. 3 Smart Tourism Outcomes

Figure 2 shows the smart tourism process, where the sensor-based data goes in to the data repository. This is the real-time data retrieved from sensors like GPS, camera, scanners, robots, etc. For example, the cameras fixed in the junctions provide real-time data on the traffic in that particular place, thus helping the people select an alternate route if they want to. Leo Foster et al. describe the inner city recreational facilities developed in Lowfield Park in Sheffield, UK, a City Council–maintained Athletes stadium called FieldLabs, which contains the ICT facilities with hardware infrastructure. Automatically timed sprint and running tracks along with the gait analyser provides the immersion facilities necessary to impact the inmates to take up physical activity [10].

Figure 3 shows the smart tourism outcomes. This gives an overall picture of how the data are essential for smart tourism to operate

Figure 4 shows the smart recreation ecosystem, which provides the complete picture of the hardware and software technologies as well as the people, businesses and economy involved. According to Boes et al., the smartness comes with two aspects – hard and soft, where hard smartness is the use of information technologies while soft smartness is the interaction of businesses, people and government,

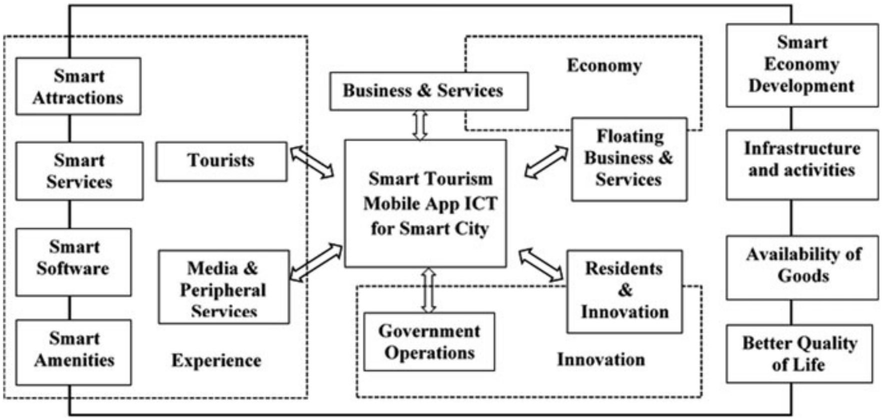


Fig. 4 Smart tourism ecosystem

utilising these information technologies and sustaining the smartness of the place [11]. The consumers or tourists are the deciding factor of the smart destination level of the city, and it is important to improve the quality of life for the residents too. This is what the ecosystem indicates. All are interwoven, and the data gathered from the tourists not only should be used to increase the satisfaction level of the tourists but to improve the amenities, businesses and the general quality of life [12, 13].

The mobile app is the user interface, and it provides the connection and data about the preferences of the users. During the planning phase of the visit, the user chooses the recreation activity. First, the user must ensure the particular mobile application to be installed on the mobile. A good network connectivity is essential for the mobile app to work. The user shall enter all the personal details and contact details, further moving to the next page, where the user shall enter the details of the visit, the dates of the visit and the number of persons, including their name, age and other details. Following this, the user will be directed to any page where the essentials like transportation, food or accommodation can be selected.

In the subsequent pages, the application can capture all the interested areas of recreation of the user, like the interest to visit various sightseeing locations in the area and nearby locations, shopping areas, exhibition and fairs, restaurants with choice of food types and type of nutrition.

Further, the user can engage in various relaxation and wellness facilities, like meditation, yoga, spa and rejuvenation therapies, as these are some of the very popular recreation activities.

Next, the application shall update the user with some of the special attractions, year-round attractions or even customized attractions, thereby enhancing the experience of the user. The app can connect with the official pages of attractions, which will allow the user to book the date, time and number of persons; ticket required and possible discounts or offers, including the payment gateway, making hassle-free recreations for the user at the destination. Further, the possible modes of transport, both public and private, for visiting the attractions can be listed along with the connections to the respective official pages.

An example of the data that will be collected while the user is planning to visit an attraction and wishing to book a ticket is given here:

Page I: About the user:

Name:
Age:
Contact details:

Page II: Details of visit

Dates of visit:
Number of persons:
Age of accompanying persons:
If accommodation is required:
If transport is required:

Page III: Interests of the user:

1. Attractions
2. Shopping
3. Food
4. Relaxation
5. Others

Page IV: Attractions available for the dates:

1. Special attractions
2. Year-round attractions
3. Customised attractions

Page V:

Booking required for attractions?
Connect with the official pages of attractions
Date, time and no. of tickets
Any special offers

Page VI: Transport for visiting the attractions required

The weblinks to public and private transports are available and the booking can be done by connecting to the webpages of the transport.

Page VI:

List of recreation centres with accessibility, for children and old people

To strengthen the user-friendly aspect and the optimization of response time and in order to have more information on activities, preferences as well as problems of the users, the development of a chatbot tourism proves necessary.

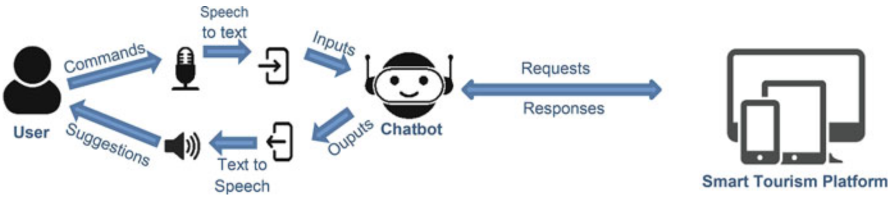


Fig. 5 Proposed architecture of SmartTourismBot

In the recent scenarios, the chatbots based on artificial intelligence (AI) are playing a very important role in modelling a virtual assistant by text, also by voice. Based on the the user's preferences as well as the tourist database of each destination, the chatbot will provide suggestions on measures, allowing better performances and a good decision-making for the user and help them enjoy every moment of their journey.

The SmartTourismBOT is developed with natural language processing (NLP), which is a major pillar of artificial intelligence and will be an ideal tool for understanding the context of discussion and self-assessment and improvement over time. And to overcome the barriers of languages, the proposed chatbot will be multi-lingual (Fig. 5).

The chatbot is based on NLP (Natural Language Processing) and NLU (Natural Language Understanding) to understand the user's command and generate an adequate response. NLP in turn makes it easier to read, decode and understand the language.

Firstly, the user's command is received through the microphone and then translated into text via the speech-to-text module. Secondly, the text is processed via NLP according to a syntactic analysis and a semantic analysis.

- Syntax analysis involves identifying the grammatical rules in a sentence in order to decipher its meaning.
- Semantic analysis, based on artificial intelligence and deep learning, directly deciphers the meaning of a text using algorithms to analyse words and sentence structures.

Once the processing is complete, the result is compared with the platform's database to collect the information that serves as a response from the user's command. This text response is translated to sound through text-to-speech conversion and transmitted to the user.

The smart tourism ecosystem is based on three elements – innovation, experience and economy. These three elements are closely related to one another and provide the basis for the objective of smart cities. The core of this ecosystem is the data developed and the connectivity available. But the data sharing has its pros and cons, which might affect the outcomes. While the data sharing is essential for understanding the needs of the customers to enhance innovation, this also provides concerns on the security aspects of the people and property [14]. Continuous monitoring of the movement of people is considered to be a breach of privacy, and it also creates a basis for tracing the weak points of security by the anti-social elements. But the planners of the smart city models also plan for security systems to increase the safety and security and encourage innovations for the same [15].

The interconnection between the satisfaction level of the tourist and the development of businesses and economy is a straight one. Smart tourists always go for smart destinations, and the experience is considered to be a smart experience [16]. This is done with the help of technologies which can personalize and provide information aggregation through ubiquitous connectedness and real-time synchronization of the data [17].

Smart businesses: The experience of the tourist for attractions includes purchasing some souvenirs from the smart destination which might portray something related to an attraction or the iconic symbol of the place or some special indigenous products exclusively made in the area [18]. All kinds of small businesses with innovation behind the objective have been relished by tourists and have enhanced their experience [19, 20]. For example, the experience shared through photographs, feedbacks, reviews and social media feeds have greatly increased the awareness on this as a smart information technology business. This innovative mechanism has been successfully applied to supply businesses to enhance the shopping experiences of the tourists.

The increase in shopping also increases all peripheral businesses and attached ancillary businesses, and many small-time sellers and service providers in the local communities are the beneficiaries. The cultural immersion provided by the shopping experience is important for increasing the economy as well as the visibility of the place as a "Smart Destination". The other businesses that will be benefited by the ICT in tourism are the restaurants, photography, small and exclusive souvenir businesses.

4 Innovative Technologies in Smart Tourism

Technologies used in earlier days for scientific or military purposes have now been used to provide recreation. Some of the technologies are even customised for these recreation activities, i.e. basically the user can select the recreation activity. This generally happens when we do online shopping. The user downloads an application for shopping. Once the user starts to browse through the various catalogues and pause for a certain item, the artificial intelligence (AI) technology captures the user's preferences and machine learning (ML) happens. Next time the user opens the same application or any similar application, the preferred choices of the user pop up with offers, availability of sizes, colour options, price range, brands, etc. in the apparel, footwear section or any previously visited items. The use of AI and ML has been a boon to the retail industry and hence has also been a recreation activity for many users, especially in the case of virtual fitting rooms, which use augmented reality technology, and hence online shopping is considered as a recreational activity with all versatile options open for the users.

Smart cities around the world boast a plethora of technologies to cater to the various needs of both tourists and locals. In public transportation, the websites provide the timing and destination of the transports, which is very common. One other earlier known recreation activity was photography. Here simple camera technology was used and thus photography became a very popular recreational activity and in turn led people to go globetrotting to capture the tourist attractions, natural landscapes, hotels and resorts, cultural experiences, outdoor adventures and wildlife.

Aerial tramway, or more popularly known as cable car, which started as an alternate mode for conventional urban transport, has now become a recreational activity, the technology of electrically propelled steel cables has also led to other recreational activities like riding on roller coasters. But innovation in some of the smart cities provides additional attraction. For example, interactive bus shelters, where the tourists can enter their destination and the shelter kiosk can provide the available buses which will stop at that particular shelter and their timings of arrival, the availability of seats in them, whether passes are allowed in them and the cost of the ticket. This will help the tourist get the exact change for the ticket or the pass and decide whether the wait will be too long or short. The tourist can also ask for alternate routes/means of transport if buses are not available.

Many eco-friendly cities provide bicycles to anyone with a credit card to pay for them. Apps are developed and the bicycles are available in many part of the cities where the user can use the kiosk to rent a bicycle and use it. The return of the bicycle can be done in any of the places where the bicycle rent centre is available. This provides a low cost and convenient method of touring the city at their leisure, and convenience is the key.

Information exchange is an important aspect of smart tourism from the business perspective and the management perspective. As long as all the stakeholders are well connected, the information flow will be smooth and will cater to an effective and efficient smart tourism management. Dynamic linking of all stakeholders and processes will require a strong infrastructure and a better information exchange mechanism. According to Tomas Gajdosik et al., sharing of data will lead to communication costs and increase the opportunity for better economic development [21].

The buying capacity of people has increased in recent years, leading to the mushrooming of shopping complexes, housed within huge buildings or areas. Navigating within these complexes is really complex, and nowadays, most of the complexes have digital navigation tools available for the visitors. When visitors enter the complex they can connect to the app and get information on the businesses available, location of the shops, services and goods available and the directions to any of these. Smart apps can find the buying habits of repeating visitors and provide suggestions on related businesses, discount offers available and review of

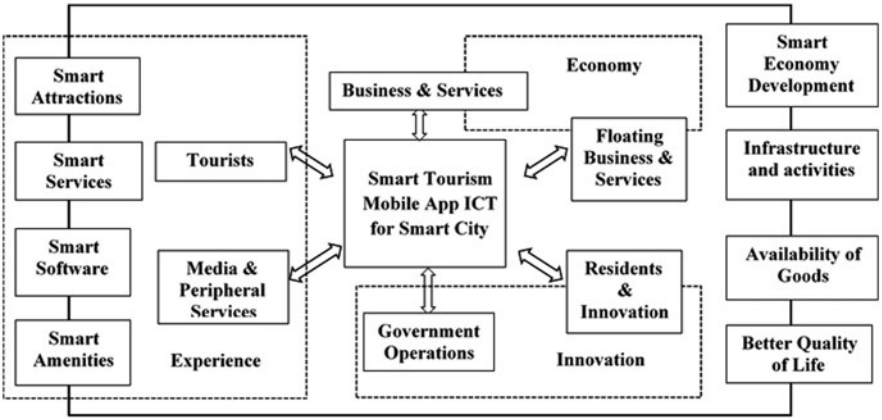


Fig. 6 Components of smart city recreation

the products. This information will be used for further refining the efficiency of the app [22].

4.1 Accessible Technologies

Accessibility is a worldwide requirement since language diversity poses a problem in communication between communities. Amsterdam provides translation for the roadside signs and other directions to help tourists. Most of the smart cities will have charging points for mobiles at the attractions, and wheel chairs are available for the mobility of impaired people. Attractions also have personal translator devices available for a cost which will talk about the history and the development of the place in English or some other language popular in that country. Singapore has an online platform called "Virtual Singapore" which helps tourists by educating them about the government policies and potential risk management [23, 24] (Fig. 6).

Sauro vicini et al. show using the development of City of the Future Living Lab how the user-driven open innovation ecosystems can evolve new concepts in a smart city environment and enhance the user experience [25]. According to their argument, interactions between the services and users are the potential stimulant for the development of new technologies. Hence, for the rehabilitation centres catering to old age and the differently abled, a smart city environment will be the best because of the evolving technologies. The speech recognition systems can be enhanced to provide noise-filtered user experience for people who are hard of hearing.

5 Challenges in Building These Systems

There are many innovations required for building these systems, which also bring in many challenges. The first and foremost challenge any smart city project faces is raising enough funds to create all the facilities required. A smart city can be functioning as a smart one by providing the basic and essential technologies for mobility, operations and interconnectivity. So how smart is called smart? It is not enough to have a basic infrastructure for providing sustainability for the inmates as well as the city operations. The smartness of the city can be sustained only by providing a continuous inflow of tourists and businesses, which will ensure the economic development. This requires an all-round development of technology, infrastructure and management and the necessary strategy to raise the funds [26, 27].

Apart from funding, the next challenge might be in the planning and development. A smart city will not be developed from scratch, which means that the existing infrastructure has to be somehow utilized or retrofitted to create smartness. The most important is to determine the weak areas, like the availability of basic amenities like water supply, waste management and sanitation. The existing structures are to be modernized to achieve efficiency and smartness using technologie, redesigned to accommodate the increasing population and businesses and maintained to attract the steady flow of floating population [28].

One more challenge might be drawing a layout or a masterplan which provides a list of developmental activities and enhancement activities based on the availability of facilities and amenities. The master plan will be necessary to keep track of the city development and determine where the speeding up needs to be done.

The next challenge is the effective coordination of the various institutions providing amenities both on the local level and central level. Government agencies should ensure smooth operations by providing clearances in a timely manner, and other institutions should coordinate and share the best practices and service delivery processes. Technical constraints like non-availability of technology to cover the connectivity of all regions in the city, cost-effectiveness of the technologies available and coordination of software technologies supplied by the different vendors are some other challenges. The final challenge may be educating the residents and the visitors about the safety measures and building a reliable service utility for water, electricity and connectivity services like mobile networks and broadband services. Shifting towards renewable resources and developing green infrastructure might mitigate part of the challenge [29].

6 Conclusion

Building smart city and maintaining it using the technologies provide economic, social and cultural development for a country and community. The sustainability of a smart city depends on the availability and maintenance of technology, services and operations. Building a city development plan and an inclusive innovative model at the beginning will help the development of the smart city without mishaps. The model will bring out the basis for smart destination requirements, and for sustainability, an internal coordination of all stakeholders is very much essential.

References

1. Smart cities: Digital solutions for a more livable future, Report by McKinsey Global Institute, June 5, 2018
2. J. López-Sintas, L. Rojas-DeFrancisco, E. García-Álvarez, D. Bara (Reviewing Editor) (2017) Home-based digital leisure: doing the same leisure activities, but digital, Cogent Soc. Sci., Taylor and Francis online, 3:1, https://doi.org/10.1080/23311886.2017.1309741
3. G. Biancarosa, G.G. Griffiths, Technology tools to support reading in the digital age. Future Child. **22**(2), 139–160 (2012). https://doi.org/10.1353/foc.2012.0014. PMID: 23057135
4. C. Cheng, X. Yang, Study of online leisure activities—exemplified by shopping online without initial intention, in *2010 Second International Conference on Communication Systems, Networks and Applications*, 2010, pp. 160–163, https://doi.org/10.1109/ICCSNA.2010.5588674
5. L. Rojas-DeFrancisco, Digital leisure study: the consumption experience, habits and social uses, in Thesis for: Doctor in humanities – Production and consumption of culture, February 2015, https://doi.org/10.13140/rg.2.1.1198.2808
6. M. Höjer, J. Wangel, Smart sustainable cities: definition and challenges, in *ICT Innovations for Sustainability, Advances in Intelligent Systems and Computing*, pp. 333–349 (2015)
7. P. Lee, W.C. Hunter, N. Chung, Smart tourism city: developments and transformations. Sustainability **12**(10), 3958 (2020) 10.3390/su12103958
8. U. Gretzel, M. Sigala, Z. Xiang, et al., Smart tourism: foundations and developments. Electron. Mark. **25**, 179–188 (2015). https://doi.org/10.1007/s12525-015-0196-8
9. D.B. Martin, *Measuring Digital Development Facts and Figures 2020* (ITU Publications, September 2020)
10. L. Foster, B. Heller, A. Williams, M. Dunn, D. Curtis, S. Goodwill, *Development of Smart Inner City Recreational Facilities to Encourage Active Living, Lecture Notes in Computer Science, II (10070).* (Springer, 2016), pp. 458-468
11. K. Boes, D. Buhalis, A. Inversini, Smart tourism destinations: ecosystems for tourism destination competitiveness. Int. J. Tour. Cities **2**(2), 108–124 (2016). https://doi.org/10.1108/IJTC-12-2015-0032
12. Z. Xiang, I. Tussyadiah, D. Buhalis, Smart destinations: foundations, analytics, and applications. J. Destin. Mark. Manag **4**(3), 143–144 (2015). https://doi.org/10.1016/j.jdmm.2015.07.001
13. Z. Xiang, D.R. Fesenmaier, Big data analytics, tourism design and smart tourism, in *Analytics in Smart Tourism Design, Concepts and Methods*, ed. by Z. Xiang, D. R. Fesenmaier, (Springer, Cham, 2017), pp. 299–307. https://doi.org/10.1007/978-3-319-44263-1
14. U. Gretzel, Intelligent systems in tourism: a social science perspective. Ann. Tour. Res. **38**, 757–779 (2011). https://doi.org/10.1016/j.annals.2011.04.014
15. S. Zygiaris, Smart city reference model: assisting planners to conceptualize the building of smart city innovation ecosystems. J. Knowl. Econ **4** (2012). https://doi.org/10.1007/s13132-012-0089-4
16. A. Lopez de Avila, Smart destinations: XXI century tourism. in *Presented at the ENTER2015 Conference on Information and Communication Technologies in Tourism, Lugano, Switzerland*, February 4–6, 2015

17. D. Buhalis, A. Amaranggana, Smart tourism destinations: enhancing tourism experience through personalisation of services, in *Information and Communication Technologies in Tourism 2015*, ed. by I. Tussyadiah, A. Inversini, (Springer, Heidelberg, 2015), pp. 377–389
18. M. Callon, F. Muniesa, Economic markets as calculative collective devices. Organ. Stud. ***26***(8), 1229–1250 (2005)
19. J.A. Coca-Stefaniak, Place branding and city centre management – exploring international parallels in research and practice. J. Urban Regen Renew **7**, 367–373 (2014)
20. J.A. Coca-Stefaniak, R. Powell, A.M. Morrison, D. Paulauskaite, Living like a local: authentic tourism experiences and the sharing economy. Int. J. Tour. Res **19**. https://doi.org/10.1002/jtr.2134 (2017)
21. T. Gajdošík, Smart tourism: concepts and insights from central Europe. Czech J Tour **7**(1), 25–44 (2018). https://doi.org/10.1515/cjot-2018-0002
22. R. Elgazzar, R. El-Gazzar, Smart cities, sustainable cities, or both? in *A Critical Review and Synthesis of Success and Failure Factors. SMARTGREENS 2017 Conference* (2017)
23. L. Foster, B. Heller, A. Williams, M. Dunn, D. Curtis, S. Goodwill, Development of smart inner city recreational facilities to encourage active living, in *Ubiquitous Computing and Ambient Intelligence. IWAAL 2016, AmIHEALTH 2016, UCAmI 2016*, Lecture Notes in Computer Science, ed. by C. García, P. Caballero-Gil, M. Burmester, A. Quesada-Arencibia, vol. 10070, (Springer, Cham, 2016). https://doi.org/10.1007/978-3-319-48799-1_50
24. U. Majeed, et al., Blockchain for IoT-based smart cities: recent advances, requirements, and future challenges. J. Netw. Comput. Appl. (2021) p.103007
25. S. Vicini, S. Bellini, A. Sanna, How to co-create internet of things-enabled services for smarter cities. in *SMART 2012, The First International* .... pp. 55–61 (2012)
26. R. Elgazzar, R. El-Gazzar. Smart cities, sustainable cities, or both? in *A Critical Review and Synthesis of Success and Failure Factors. SMARTGREENS 2017 Conference* (2017)
27. E. Sigalat-Signes, R. Calvo-Palomares, B. Roig-Merino, I. García-Adán, Transition towards a tourist innovation model: the smart tourism destination: Reality or territorial marketing? J. Innov. Knowl. **5**, 96–104 (2020)
28. H. Allahar, What are the challenges of building a smart city?, Technol. Innov. Manag. Rev., September 2020
29. P. Padode, The top 10 implementation challenges for smart cities in India, *The Economic Times*, Realty Check, July 2015

Dr. Ramalatha Marimuthu has retired as Professor in ECE from Kumaraguru College of Technology, Coimbatore, India. She has published 6 technical books and over 70 research papers in international conferences and journals in the areas of Vedic Mathematics, Data Sciences and VLSI. Her interest in building assistive systems for the health care and differently abled people has culminated into developing unique solutions for dyslexia, speech impaired, drop foot and autism. This earned her invitations from Google and various universities all over the world to deliver special lectures. She has won many international awards from organisations like IEEE, USA, Lions Club and Anita Borg Institute for Women and Technology, California for her work towards the community.

Dr. Bindu Thomas is IEEE Senior Member and Chair of IEEE Mysore Sub Section. With 24 years of teaching experience, she is now Professor and Head Department of Electronics and Communication Engineering at Vidya Vikas Institute of Engineering and Technology, Mysore, Karnataka, India. Her research areas are Pattern Recognition, Embedded Systems, Computational Intelligence applications in Robotics and Micro Electro Mechanical Systems MEMS. She has published over 85 peer-reviewed research publications in various international Journals and conferences.

Mrs. Namratha Dcruz is currently working as Assistant Professor in the Department of Electronics and Communication Engineering at Vidya Vikas Institute of Engineering and Technology, Mysore, Karnataka, India. She has a BE and MTech in Signal Processing from Visvesvaraya Technological University, Karnataka. She is having 5 years of teaching experience. Her interests include Signal Processing and MEMS. She is an active member of IEEE and is currently the EXECOM member of IEEE Mysore Subsection.

Prof. Aawatif Hayar is IEEE Senior Member, Chair of Casablanca IEEE Core Smart City project, President of University Hassan II Casablanca Morocco-GreenTIC-ENSEM and Co-founder of E-madina. She was selected in 2015 by the African Innovation Foundation as one of the Top-10 innovative African women. She is currently leading R&D projects in smart cities fields with the City of Casablanca, IRESEN, CNRST, Lydec, GIZ and HBS Foundation.

Index

S. Rani et al. (eds.), *IoT and WSN based Smart Cities: A Machine Learning Perspective*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-84182-9