ANALYSIS OF DATA STREAM EVENTS TO ENHANCE CONTEXT AWARENESS

Ms. Shilpa Sachin Bhojne[1], Dr. Amol D. Potgantwar[2], Mrs. Neha Rohan Hiray[3]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Even in muddy, dark underwater habitats, killer whales have learned to recognise a variety of different marine life. Recent research has shown that killer whales can locate and identify different species of marine life even in waters with poor visibility. This is due to the fact that killer whales send out sound waves to other aquatic life and listen for their echoes. Right now, the standard is what we are focusing on in order to recognise the Internet of Things-based issue. We all coexist underwater, and there is a core criterion for producing sound waves and detecting echoes that is used to investigate different living groups. Thus, it is essential to play a central function and a reference role among numerous sensors for scenario detection in order to raise the amount of condition information by utilising a broad range of sensors in an ideal IoT system.

In order to establish such standards, this study suggests looking at additional sensor data streams based on one data stream among several other sensor data streams that are streaming into a network. With this, it is able to deduce the correlation between distinct sensor data and to see how numerous sensor types carry out sensing tasks as one circumstance develops, which may significantly improve situational awareness. In this study, different sensors were used to choose one of the many sensor data streams. A different sensor data stream was then searched around the reference sensor data stream to choose the data section required for analysis, and the data stream was used to demonstrate better situation recognition.

Keywords: Data Stream, Internet of Things, Event Data, Multi-Sensor, Context Awareness

[1]Assistant Professor, Department of Computer Science & Engineering, School of Engineering and Technology, Sandip University, Nashik, Maharashtra 422213

[2]Professor, Department of Computer Science & Engineering, Sandip Institute of Technology & Research Centre, Nashik, Maharashtra 422213

[3]Assistant Professor, Department of Computer Science & Engineering, Sandip Institute of Engineering and Management, Nashik, Maharashtra 422213

Email: [1]shilpa.bhojne@sandipuniversity.edu.in, [2]amol.potgantwar@sitrc.org, [3]neha.hiray@siem.org.in

DOI: 10.31838/ecb/2023.12.s3.145

1. Introduction

Many application initiatives and research on real-world situational awareness based on the Internet of Things are now underway. Artificial intelligence is being employed in a variety of sectors thanks to the Internet of Things, and previously unusable domains can now benefit from intelligence capabilities. The most commonly stated objective for artificial intelligence built on the Internet of Things is "situation awareness." The Internet of Things protocol scenario recognition systems have two features. There are several different kinds of sensors being employed, and each sensor's data is being continually reported as a host throughout time, giving the data a stream shape. If so, it can be claimed that processing different sensor data and continuous data processing are the difficulties that the IoT-based scenario recognition system has to overcome.

Studies already conducted on the analysis and processing of sensor detection data that is collected by sensors and sent to hosts through networks frequently centre on minimising the quantity of data processing. On the other hand, as IoT application systems frequently aim at "situation identification," it can be claimed that it is important to both lessen the volume of data processed and to raise the standard of situational data.

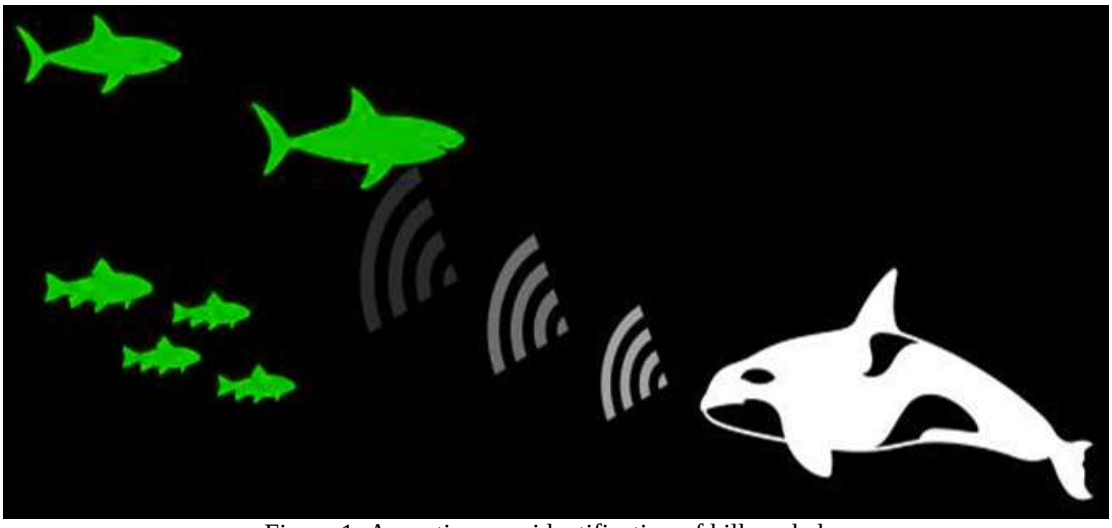

Figure 1: Acoustic wave identification of killer whales

Hence, this study suggests a strategy to lessen the quantity of data processing in current IoT-based systems while also increasing the accuracy of information in detecting real-world circumstances. Due to the nature of the ideal IoT system, data reported from sensors continues to flow over time. As a result, it is challenging to handle the vast volumes of data that are introduced over time by storing and analysing a set quantity as before. A relevant data group must be chosen when evaluating data groups that are in continuous flow across time. In other words, it entails establishing time intervals gradually, discarding data that is unnecessary for the majority of the study, and carefully choosing the data that is required for analysis. If so, the study's objective will be to come up with a strategy for choosing a data stream interval in order to decide which data to choose from the incoming data. It is problematic to choose the criteria, time intervals, and data sections that are picked.

On the other hand, many researchers who are interested in finding solutions to engineering issues have adopted a strategy that imitates how solutions appear in the natural world. Bats resemble ultrasonic sensors, and radar mimics a dragonfly's eye structure. It is known as biomimicry to solve issues by copying particular processes or algorithms found in the natural world in order to attain functional goals for various manmade systems in the actual world. This study also aims to draw conclusions from how killer whales detect a variety of different marine life in the ocean's undersea habitat and apply it to challenges. According to new research, killer whales utilise their capacity to produce sound waves and hear their echoes to locate and identify the species of marine life in waters with poor visibility. For now, the standard is what we are focusing on. We all coexist underwater, and there is a core criterion for producing sound waves and detecting echoes that is used to investigate different living groups. In order to raise the degree of scenario information by using a broad range of sensors in an ideal IoT system, it is important to play a central role and a reference role among numerous sensors for situation recognition.

This paper suggests a way of evaluating additional sensor data streams based on one data stream among several sensor data presented utilising a network in order to establish such standards. This makes it feasible to discern how different types of sensors are detecting a given scenario and infer the relationships between diverse sensor data, which may greatly improve situational awareness. In this study, different sensors may be utilised to choose one of the different sensor data streams as a reference and look for other sensor data streams nearby to demonstrate superior scenario identification utilising the acquired data stream.

These are the elements of this study. The related investigations are summarised in Chapter 2, and the data stream analysis methods suggested in this study are introduced and explained in Chapter 3. The suggested approach is tested and evaluated in Chapter 4, which also finishes Chapter 5.

2. Literature Survey

The term "Internet of Things" (IoT) refers to the information and communication infrastructure that links any thing to a network, enabling anytime, everywhere communication between people and objects. By identifying and adjusting to the immediate physical environment, situation recognition aims to deliver material that is appropriate for users (Kiris et al., 2008). For situational awareness, sensor data about the surrounding physical environment should be gathered. To identify the sensor network and process the sensor data gathered, abstraction or inference procedures are then used. Image recognition techniques, Bayesian networks, and probability-based reasoning are all used in artificial intelligence (Knappmeyer et al., 2009).

With regard to the possibility of an increase in traffic accidents brought on by a driver's cognitive decline in the ageing population, Yoo Je-wook and four other researchers studied the situation recognition module. To help users quickly recognise and react to unexpected situations, LEDs and speakers were attached (Vavoulas et al., 2014). In response to a rise in information volume brought on by the growth of the surveillance, reconnaissance, and communication domains, Jin So-yeon and four other individuals are requesting a difficult scenario detection service. There is a requirement for a system that can detect difficult scenarios through spontaneous inquiries away from the rule base based on combat expertise. It delivers rule-based knowledge that analyses, saves acquired information, and assesses situations according to instructional norms. As a consequence, a method was proposed (Wang et al., 2009) for automatically producing connected queries and obtaining sophisticated inference results from the battlefield knowledge base. Robots might see, pick up on, and react to circumstances on their own, according to research by Na Sang-hyuk and another person. Real-time data streams were used to describe the data gathering scenario, and the association rule of data mining was used to implement the situation for a robot. Due to the complexity of the sensor data set the robot acquired and the suggested system structure, it was challenging to use the situational cognitive approach and the association rule exploration technique. (Al-Shargabi & Siewe, 2020) On the basis of the Internet of Things, Lee Young-dong designed a system that can track the status of wanted posters and monitor them in real time. To swiftly manage the general status of wanted groups, control information is broadcast in real time to the Internet, servers, and short-range wireless networks(Hong et al., 2009).

3. Analysis of Data Stream Events a strategy to raise context awareness

It's crucial to remember that the sensor detects and reports information while trying to identify an intelligent scenario based on the Internet of Things. Sensors just detect events that take place or the accompanying physical and chemical changes rather than the actual real-world situation itself.(Sarker, 2019) In response to such a change, the sensor creates a fine current, amplifies it, and sends a digital value or a numerical value to the host. Situational awareness is based on these basic numbers. (Kim et al., 2018)

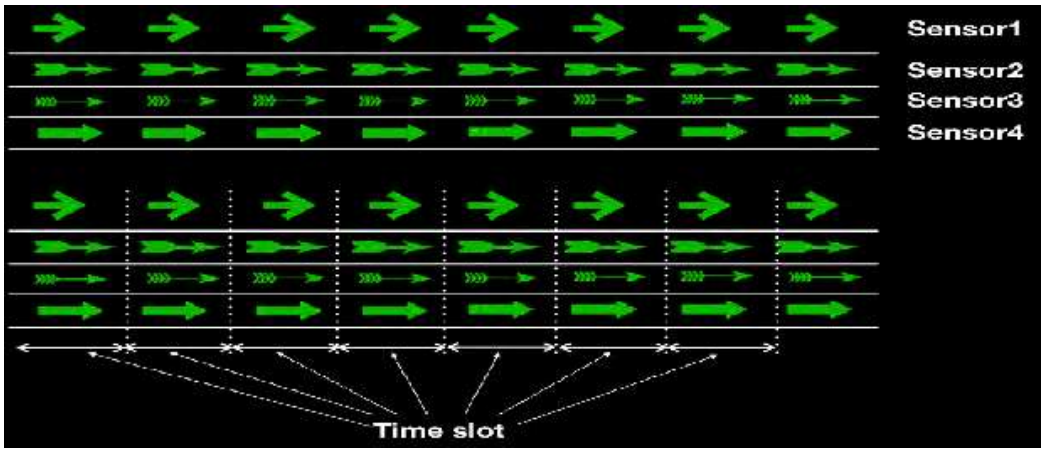

Figure 2: Information from sensors

Real-time processing should be the main consideration when employing the sensor's value for context recognition. The situation in the actual world must be timely.(Lee, 2016) Recognizing retroactively occurrences that have passed with time is less significant. For this, the ability to infer the present circumstances and information is crucial. Even though there is a temporal gap from identifying the present moment, it is required for situational awareness to notice and evaluate former circumstances that are not too far from the present. Due to this requirement, the approach of keeping a sizable amount of historical data and evaluating it over an adequate period is insufficient. Real-time analysis must be done to swiftly arrive at a scenario judgement and deliver it to users in order to fulfil the situation identification aim.(Suh & Lee, n.d.)

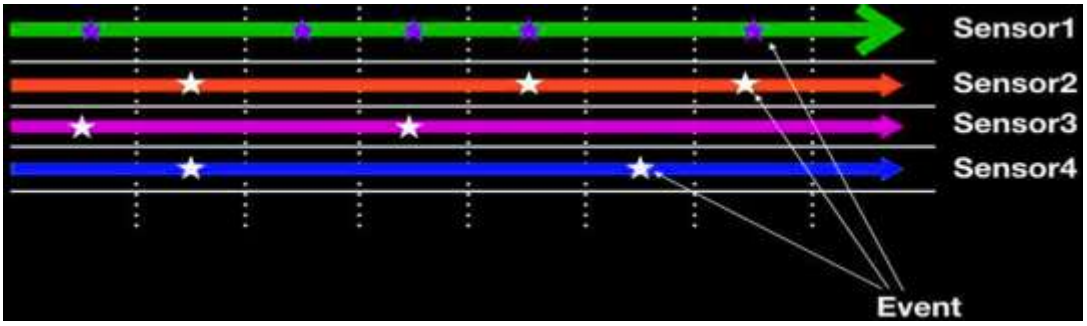

Figure 3: Event in the Sensor Data Stream

A crucial concern is making sure the recognition is accurate when utilising the sensor's value for situational recognition. Most sensors cannot detect the actual situation. Its inadequate ability to detect physical and chemical changes in response to environmental changes. The scenario is exceedingly complicated and diversified, and since individual sensors only react to physicochemical changes, extra methods are required to enhance situational recognition performance.(Gope et al., 2021; Mahdavinejad et al., 2018) Several types of sensors are given as an example of a method that may be utilised to determine the accuracy of scenario detection. In this instance, a challenging issue arises because the observed values from sensors that are not identical to one another are not uniform.

However, the problems with real-time processing and quick data processing outlined above as well as the problems that improve scenario detection accuracy are opposing in nature. Accuracy is diminished by processing it fast, and processing speed must inevitably fall as accuracy rises.
For data collected and reported over time by sensors in an IoT-based scenario identification system, a data stream analysis approach for situation recognition is presented in this paper. To effectively handle a vast volume of data, it is necessary to first choose a processing objective with care. The amount of data handled can be lowered by doing this. Second, a processing strategy is required to better understand the situation utilising various forms of data that are collected and reported by various sensors.

The purpose of this study is to improve situational awareness by selecting which data to analyse from the sensor data that has been introduced from numerous sensors throughout time.
This study made reference to the technique for identifying underwater whales in order to accomplish this purpose. Killer whales are reported to have acoustic identification systems that listen for sound waves that match the target they're aiming for as well as the sound that the target's beak is reflecting.
The marine and undersea environments support a vast diversity of living groups that are active and able to live. A pan-capacity whale's to swiftly identify the species of each living item in this marine and undersea environment is something that should be taken into consideration by an IoT scenario identification system that chooses and employs data recorded and sent from various sensors.

The following are the requirements for this study.

- An Internet of Things-based scenario recognition system was built.
- Each sensor in the current system is a separate sort of sensor, and they are placed in a terminal. As a new scenario arises or the everyday condition changes, each sensor detects the accompanying physicochemical changes and reports them to the host.
- Physical and chemical changes that are directly connected to situational detection and inference can be detected by each sensor.
- Over time, data recorded by each sensor flows continually into the host.

The system's objective in this study is to recognise or infer events that naturally occur without human involvement. In this instance, a solution is suggested that limits the data streams continually added over time to those that include only significant information while maintaining scenario recognition accuracy. In other words, it is possible to preserve or improve the quality of scenario recognition while reducing the amount of data that has to be processed using data stream analysis.
A reference sensor, such as a killer whale generating sound waves, is used as a reference point to help achieve this aim of identifying other marine species. Moreover, depending on this time point, it checks to see whether any additional sensors contain a sensor in which an event that has to be documented has already happened. The centre and the reference sensor are combined to be utilised for situational inference and identification if an event happens in another sensor. The weight will be adjusted for each case's number and utilised for situational inference and identification if the number of sensors stays at two or three.
The sensor that makes up the terminal is chosen because it can help with situational inference. A sensor that can provide situational inference is chosen from a group of sensors.
When sensor A was chosen. In sensor A, an event gradually takes place. In this instance, in addition to sensor A, it is looked to see whether any other sensors have had an event. Continue searching for several sensor data streams coming into the host to check if any remarkable events have occurred. This will reveal a sensor data stream that reports events that accompany occurrences of the reference sensor.
It is a general rule to avoid storing practically all data streams. Yet, all data may be picked as much as necessary without being wasted by choosing data to be studied for context recognition. By doing this, the strain of processing all incoming data constantly throughout time may be avoided, and effective data processing is made feasible.

Nevertheless, the data stream analysis/processing approach suggested in this study cannot just end at determining meaning by lowering the data throughput. With the use of selected data processing, it should also be able to improve the accuracy of situational information. The strategy put out in this work is structured to aggregate the sensing information transmitted by many kinds of heterogeneous sensors. When compared to detecting a scenario by using events from numerous sensor groups with the same purpose, combining data detected by heterogeneous sensors with distinct functions often results in better situation identification performance. Several CCTVs were employed when intelligent forest fire detection systems were established in Korea, however occasionally fog or clouds were mistaken for smoke from fires. At this point, the smoke, fog, and clouds brought on by the fire were appropriately detected when the air pressure sensor or humidity sensor was combined with CCTV. Hence, using sensing data from various senses that are heterogeneous to one another is an advantage for scenario reasoning or situational identification. In order to recognise circumstances and gain better situational information than utilising sensor data detected and reported by a single sensor, the suggested approach in this work is to identify and integrate events from data streams detected and sent by heterogeneous sensors of different types.

A data stream processing algorithm for scenario identification based on sensor data is represented by the following formula.

1) The step function below determines if an event occurred in any one of the K sensors.
2) The host is informed if any sensor experiences an event.
3) The step function below determines if an event persists and changes in any sensor where it happens if it does not occur in other sensor data.
4) In this situation, event data from the relevant sensor is not received when the event of the event-detected sensor data displays a declining pattern.
5) It examines if each sensor data event changes to a decreasing pattern or remains over time and is broken into sections at regular intervals.
6) Examine each sensor's data to see if any events are missing.
7) The event is searched as a step function when and, that is, if there is sensor data in which the event is ended after the reduction pattern.
8) If an event repeats in the sensor and goes on, the procedure is resumed. 9) to 11.
9) Stop date stream analysis when a sensor ceases reporting events when the event data from each sensor occasionally enters a diminishing trend and does not start reporting events again.
10) The user receives the analysis findings.

The aforementioned algorithm takes care of the complete procedure of processing and storing in response to a data stream presented from a sensor.

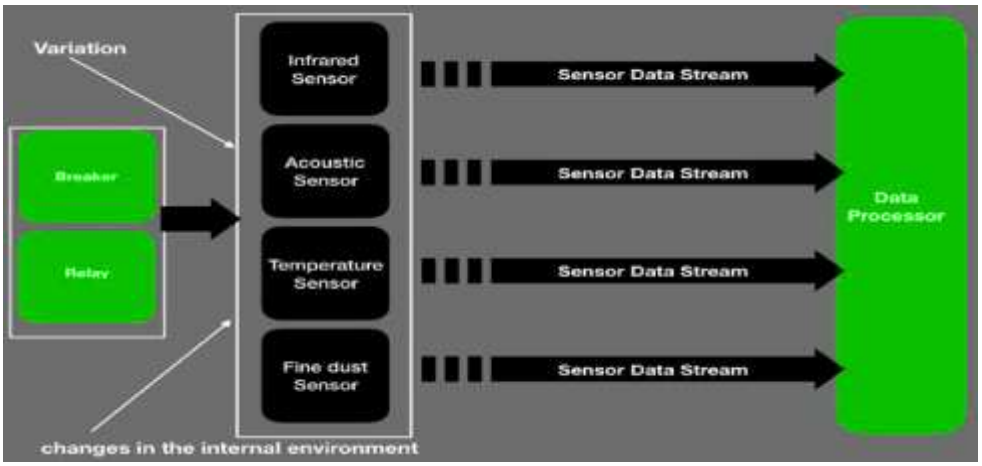

Figure 4: Stream Analysis of Data

How to create the standards is the topic that this study addresses. Before examining the events of another sensor data stream for context recognition, it is necessary to identify a sensor that provides a reference data stream. In this circumstance, using a sensor with a good connection to the situation as the reference sensor is preferred. In other words, the reference should be the sensor that is closest to the features that signify the biggest change in the circumstance. A heat sensor should be used as a guide for identifying a location in order to recognise a fire. A humidity sensor must be the norm in order to detect fog. A current detection sensor is suggested as a reference to detect the state of various electrical devices, and a heat detection sensor is suggested as a reference to detect various electrical problems. It can be claimed that it is a good idea to utilise a vibration sensor as a reference in order to detect abnormal conditions in a workspace where numerous industrial equipment are installed. Experiments with the topics suggested in this research are included in Chapter 4 of the next book.

4. Test and Evaluation

The experiment is presented in the table below along with the sensor's reported sensing data.

Table 1: Situation Recognition and Sensor Data Stream

No	Humidity Sensor (%)	Fine dust Sensor (times)	Vibration Sensor (times)	Acoustic Sensor(times)	Context
1	59				Lack of humidity
2	59				Lack of humidity
3	60				Lack of humidity
4	90				Lack of humidity
5	90				an increase in humidity
6	90				an increase in humidity
7	91				an increase in humidity
8	89	11	1	0	a humid day
9	89	12	2	0	a humid day
10	88	12	3	0	Rain
11	88	14	4	0	Rain
12	88	12	5	0	Rain
13	57				Lack of humidity
14	70				Lack of humidity
15	82				an increase in humidity

In this experiment, an environment was constructed where sensors located in a confined area, such as a switchboard, could broadcast detected values and carry out operations while data values were simultaneously flowing from each sensor. A humidity sensor, a vibration sensor, an acoustic sensor, and a tiny dust detection sensor were the sensors employed in this experiment.

1) DHT-11 (humidity sensor) (humidity sensor): • Accuracy: RH 5%; Operational voltage: 5 V; Range: 20-90% • Digital switching output as the output format
2) SW-420 (vibration sensor, closed type) (vibration sensor, closed type): • Operating voltage: DC 3.3V to 5V • Digital switching output as the output format
3) LM393 (acoustic sensor): • Use a microphone to amplify the sound • DC 4 to 6V operating voltage
4) A sensor for tiny dust detection
5) Raspberry Pi Experimental Platform

4.1. Process evaluation details

Acrylic panels are used to cover the small, enclosed space. Inside, the four different kinds of sensors were installed, and an event was created to recognise and report them.

Evaluation: The humidity sensor was used as the reference sensor in this experiment. It was determined whether events happened in other sensors when the humidity suddenly rose (4), and it was looked into and determined that further events occurred in the data table (8). Data grouping was not carried out when the humidity sensor, the reference sensor, had a drop in measured value; instead, events involving other sensors were re-examined based on the rise in 15. When it was established that more events had taken place in (30), the event cases were once more evaluated and sorted. Humidity was chosen as the reference sensor in this experiment because, in the view of experts, it has the greatest impact on electrical and electronic systems. It was demonstrated that it is possible to recognise the situation for an internal environment of a specific device and differentiate it according to the situation by referring to the events of other sensors by looking at other events based on significant sensors in an environment related to electric devices.

Figure 5: Get Sensor Data

5. Conclusion

The IoT-based scenario identification system uses a variety of sensors, and each sensor continually transmits discovered data to the host over time. Nowadays, picking data from the constantly streaming sensor data in the section required for scenario identification should be the essential way to decrease the quantity of data to be processed and increase situation recognition performance. In this study, a particular change in the data stream from other sensors was identified using a sensor data stream that may be used as a reference among multiple sensors. This made it possible to examine the sensor data stream that represented the detection outcome in relation to each circumstance, and data analysis was carried out by choosing a segment with such connection activities. It was demonstrated in this way that situational awareness might still be advantageous despite processing less data. Follow-up study is required to determine how to choose sensors that serve as a reference in the future research path.

6. References

Al-Shargabi, A. A., & Siewe, F. (2020). A Lightweight Association Rules Based Prediction Algorithm (LWRCCAR) for Context-Aware Systems in IoT Ubiquitous, Fog, and Edge Computing Environment. Proceedings of the Future Technologies Conference, 22–30.

Gope, P., Sikdar, B., & Millwood, O. (2021). A Scalable Protocol Level Approach to Prevent Machine Learning Attacks on PUF-based Authentication Mechanisms for Internet-of-Medical-Things. IEEE Transactions on Industrial Informatics.

Hong, J., Suh, E.-H., Kim, J., & Kim, S. (2009). Context-aware system for proactive personalized service based on context history. Expert Systems with Applications, 36(4), 7448–7457.

Kim, T. H., Suh, D. H., Yoon, S. S., & Ryu, K. (2018). Error Correction of Real-time Situation Recognition using Smart Device. Journal of Digital Contents Society, 19(9), 1779–1785.

Kiris, I., Kapan, S., Kılbas, A., Yılmaz, N., Altuntaş, I., Karahan, N., & Okutan, H. (2008). The protective effect of erythropoietin on renal injury induced by abdominal aortic-ischemia-reperfusion in rats. Journal of Surgical Research, 149(2), 206–213.

Knappmeyer, M., Baker, N., Liaquat, S., & Tönjes, R. (2009). A context provisioning framework to support pervasive and ubiquitous applications. European Conference on Smart Sensing and Context, 93–106.

Lee, Y.-D. (2016). Distributing board monitoring system based on internet of things. Journal of the Korea Institute of Information and Communication Engineering, 20(1), 200–206.

Mahdavinejad, M. S., Rezvan, M., Barekatain, M., Adibi, P., Barnaghi, P., & Sheth, A. P. (2018). Machine learning for Internet of Things data analysis: A survey. Digital Communications and Networks, 4(3), 161–175.

Sarker, I. H. (2019). Context-aware rule learning from smartphone data: survey, challenges and future directions. Journal of Big Data, 6(1), 1–25.

Suh, D., & Lee, M. (n.d.). The Imbalanced Data Processing for the Data Stream Processing in the Smart Distribution Panel.

Vavoulas, G., Pediaditis, M., Chatzaki, C., Spanakis, E. G., & Tsiknakis, M. (2014). The mobifall dataset: Fall detection and classification with a smartphone. International Journal of Monitoring and Surveillance Technologies Research (IJMSTR), 2(1), 44–56.

Wang, Y., Lin, J., Annavaram, M., Jacobson, Q. A., Hong, J., Krishnamachari, B., & Sadeh, N. (2009). A framework of energy efficient mobile sensing for automatic user state recognition. Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services, 179–192