American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-11, Issue-04, pp-123-131
www.ajer.org
Research Paper Open Access

Data Stream Event Analysis to Improve Context Awareness

Donghyok Suh, Keunkwang Jung

Abstract
Killer whales have ways of recognizing numerous other marine lives even in murky and dark underwater environments. As it has recently been revealed, killer whales detect and recognize other types of marine life even in water where visibility is not secured. This is because killer whales emit sound waves to other marine creatures and detect their echoes. Currently, what we are paying attention to, which aims to recognize the Internet of Things-based situation, is the standard. We all live underwater together, and there is a central criterion for emitting sound waves and detecting echo waves, and other living groups are explored by that criterion. Therefore, in order to increase the level of situation information by employing a wide variety of sensors in an ideal IoT system, it is necessary to play a central role and a reference role among multiple sensors for situation recognition.

In setting such standards, this study proposes to investigate other sensor data streams based on one data stream among many sensor data flowing in using a network. Through this, it is possible to infer the association of various sensor data and to see different types of sensors perform sensing activities as one situation occurs, which can make an important contribution to situational awareness. In this study, various sensors were used to select one of the various sensor data streams, and another sensor data stream was searched around the reference sensor data stream to select a data section necessary for analysis, and the data stream was used to show better situation recognition.

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

Date of Submission: 04-04-2022 Date of acceptance: 19-04-2022

I. Introduction

Research on real-world situational awareness based on the Internet of Things and various application projects are being actively conducted. With the advent of the Internet of Things, artificial intelligence is being applied to various industries, and intelligence functions can be given to fields that could not be used before. Among the goals of artificial intelligence based on the Internet of Things, the most frequently set goal is 'situation awareness'. A situation recognition system using the Internet of Things protocol has two characteristics. First, various types of sensors are used, and second, the data detected by each sensor is continuously reported as a host over time, so it has a stream form. If so, the data processing issues that must be addressed in the IoT-based situation recognition system can be said to be processing various sensor data and continuous data processing.

Existing studies that analyze and process sensor detection data detected by sensors and reported to hosts through networks have often focused on reducing the amount of data processing. However, since IoT application systems are often aimed at 'situation recognition', it can be said that it is necessary not only to reduce the amount of data processed but also to improve the quality of situation information.

Pic. 1 Killer Whale identity with acoustic wave

Therefore, this study proposes a plan to increase the accuracy of information in recognizing real world situations as well as reducing the amount of data processing in existing IoT-based systems. Since data reported from sensors continues to flow over time, large amounts of data that have not been previously experienced are introduced over time due to the nature of the ideal IoT system, and it is difficult to cope with it by storing and analyzing a certain amount as before. When analyzing data groups that continue to flow over time, it is necessary to select a meaningful data group. In other words, it means setting time intervals over time, throwing away data that is not required for most of the analysis, and appropriately selecting the data that is essential for analysis. If so, the goal of the study will be to devise a method of selecting a data stream interval to determine which data to select from the incoming data. What criteria are set, which time interval data are discarded, and which section of data are selected is an issue.

On the other hand, many researchers who want to solve engineering problems have taken a method of imitating the clues to solving problems from how they operate in the natural world. Radar imitates the eye structure of a dragonfly, and ultrasonic sensors are imitated by bats. In order to solve problems of various artificial systems in the real world and reach functional goals, it is called biomimicry to solve problems by imitating specific mechanisms or algorithms in the natural world. Therefore, this study also intends to derive ideas from how killer whales recognize numerous other marine lives in the underwater environment of the ocean and use them to solve problems. As it has recently been revealed, killer whales use their ability to emit sound waves and detect their echoes in detecting and recognizing their types of marine life in water with insufficient visibility. What we are paying attention to at this time is the standard. We all live underwater together, and there is a central criterion for emitting sound waves and detecting echo waves, and other living groups are explored by that criterion. Therefore, it is necessary to play a central role and a reference role among multiple sensors for situation recognition in order to increase the level of situation information by employing a wide variety of sensors in an ideal IoT system.

In setting such standards, this study proposes a method of investigating other sensor data streams based on one data stream among many sensor data introduced using a network. Through this, it is possible to infer the association between various sensor data and detect the sensing activities of various types of sensors according to one situation, which can contribute significantly to situational awareness. In this study, various sensors can be used to select one of the various sensor data streams and search for another sensor data stream around the reference sensor data stream to show better situation recognition using the obtained data stream.

This study consists of the following. Chapter 2 summarizes related studies, and Chapter 3 introduces and explains the data stream analysis techniques proposed in this study. Chapter 4 conducts experiments and evaluations on the proposed plan and concludes Chapter 5.

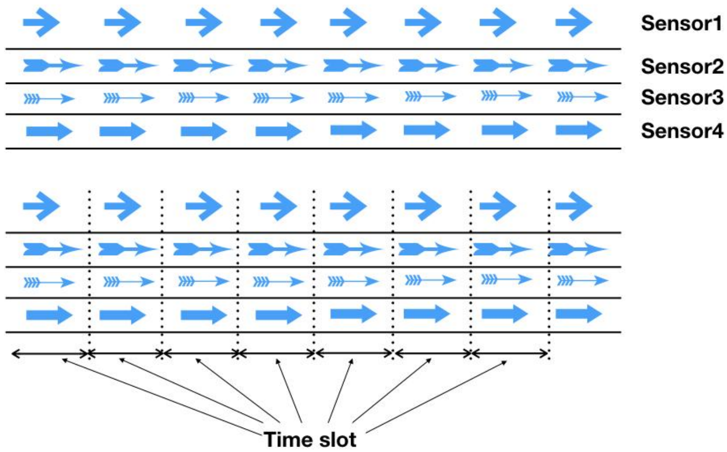
II. Related Works

The Internet of Things (IoT) refers to the information and communication base that connects all objects through a network so that humans and objects can communicate anytime, anywhere. Situation recognition is to provide content suitable for users by recognizing and adapting the surrounding physical environment [1, 2]. Sensor data on the surrounding physical environment should be collected for situational awareness. Abstraction or inference processes follow so that the user can identify the sensor network and process the collected sensor data. Artificial intelligence uses probability-based reasoning, Bayesian networks, and image recognition methods [3,4,12].

Yoo Je-wook and four others studied the situation recognition module, and in consideration of the possibility of an increase in traffic accidents caused by a driver's cognitive decline in the aging era, LEDs and speakers were attached so that users can easily recognize and respond to unexpected situations [5]. Jin So-yeon and four others are demanding a complex situation recognition service amid an increase in information volume due to the development of surveillance and reconnaissance and communication fields. It provides rule-based knowledge that analyzes, and stores collected information and evaluates situations according to teaching norms, and there is a need for a system that can recognize complex situations through impromptu inquiries away from the rule base based on battlefield knowledge. Therefore, a technique for automatically generating related queries and deriving complex inference results from the battlefield knowledge base was presented [6]. Na Sang-hyuk and one other studied whether it was an adaptive real-time situation in which robots perceived, learned, and responded to situations by themselves. The data collection situation was defined as a real-time data stream situation, and a robot's situation was implemented by applying the association rule of data mining. The system structure was proposed and the efficiency of the situational cognitive technique was verified so that it was difficult to apply the association rule exploration technique due to the specificity of the sensor data set collected by the robot.[7-9,12]. Lee Young-dong proposed a system that can monitor the status of wanted boards and monitor them in real time based on the Internet of Things. Control information is transmitted in real time to the Internet, servers, and short-range wireless networks so that the overall status of wanted groups can be quickly managed[10,11].

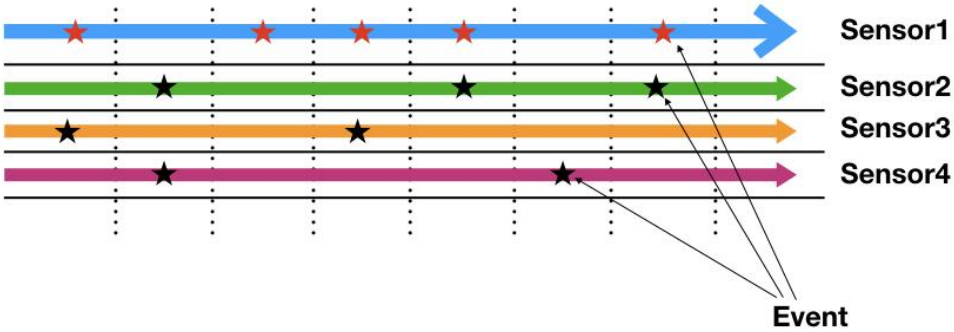
III. Data Stream Event Analysis Plan to Improve Context Awareness

When attempting to recognize an intelligent situation based on the Internet of Things, it is important to note that the sensor detects and reports data. Sensors do not recognize the real-world situation itself, but merely detect events that occur or physical and chemical changes that accompany them. The sensor generates a fine current according to such a change, amplifies it, and transmits a digital value or a numerical value to the host. These simple figures are the basis for situational awareness.



Pic. 2 Data Steam from Sensors

When using the value transmitted by the sensor for context recognition, the focus should be on real-time processing. The real-world situation must be timely. It is less important to belatedly recognize past situations that have passed along with time. It is important to be able to infer the current situation and information for this. It is necessary to recognize and judge past situations that are not too far from the present, even if there is a time difference from recognizing the current time for situational awareness. Because of this need, the method of storing quite a lot of past data and analyzing the data over sufficient time does not meet the current need. In order to reach the situation recognition goal, real-time analysis must be conducted to conclude situation judgment within a short period, and it must be distributed to users quickly.



Pic. 3 Event in the Data Stream from Sensors

When using the value transmitted by the sensor for situational recognition, an important issue is to ensure the accuracy of recognition. Most sensors do not recognize the situation itself. It has a limited function of detecting physical and chemical changes according to situation changes. Individual sensors respond only to specific physicochemical changes, and the situation is very complex and diverse, so special measures are needed to improve situational recognition performance. As a representative method that can be taken to obtain the accuracy of situation recognition, various types of sensors are used. In this case, a difficult problem occurs, and the detected values of sensors different from each other are not uniform.

In addition, the real-time processing, rapid data processing issues mentioned above, and the issues that increase the accuracy of situation recognition have opposite characteristics. If you process it quickly, the accuracy is reduced, and if you increase the accuracy, the processing speed is bound to decrease.

In this study, a data stream analysis method for situation recognition is proposed for data detected and reported by sensors over time in an IoT-based situation recognition system. The need to do this is first to selectively select a processing target to efficiently process a large amount of data. By doing so, the amount of data processed can be reduced. Second, it is necessary to obtain a processing plan to more accurately recognize the situation using different types of data detected and reported by various sensors.

Therefore, the goal of this study is to determine which data to process among sensor data introduced from various sensors over time and to increase the level of situational awareness.

In order to achieve this goal, this study referred to the method of recognizing underwater whales. It is known that killer whales have acoustic recognition boards that investigate sound waves to recognize the goal they want to capture and recognize the sound that is reflected and reflected by the target's beak.

A wide variety of living groups survive and are active in the marine and underwater environment. In this marine and underwater environment, the ability of a pan-whale to quickly recognize the type of each living thing is worth referring to in an IoT situation recognition system that selects and uses data reported and transmitted from numerous sensors.

This study has the following prerequisites.

. An IoT-based situation recognition system was constructed.

. In the present system, sensors are arranged in a terminal, and each sensor is a different type of sensor.

. Each sensor detects and reports to the host the accompanying physicochemical changes when a new situation occurs or daily situation changes.

. Each sensor has a function of detecting physical and chemical changes that are closely related to situational recognition and inference.

. Data reported by each sensor continuously flows into the host over time.

In this study, the goal of the system is to recognize or infer situations that occur in the real world without human intervention. In this case, a method of obtaining only meaningful data from among the data streams continuously introduced over time is proposed, and at the same time, such a method does not reduce situation recognition accuracy. In other words, the amount of data to be processed through data stream analysis can be reduced and the level of situation recognition can be maintained or increased.

To achieve this goal, to identify other creatures in the sea, a reference sensor is selected as a reference point, such as a killer whale emitting sound waves. And when an event to be noted occurs in this reference sensor, it searches for whether there is a sensor in which the event has occurred among other sensors based on this time point. If an event occurs in another sensor, the center and the reference sensor are combined to be used for situational inference and recognition. If two or three sensors continue to increase, the weight will be increased according to the number of each case and used for situational inference and recognition.

The sensor constituting the terminal is selected as capable of contributing to situational inference. Among several sensors, a sensor capable of leading situational inference is selected.

When A sensor was selected. An event occurs in sensor A over time. In this case, it is searched whether there is a sensor in which an event has occurred in addition to sensor A. A sensor data stream that reports events accompanying events of the reference sensor is found by continuing to search for several sensor data flowing into the host to see if any notable events have occurred.

It is a principle not to store almost all the data streams. However, by selecting data to be analyzed for context recognition, all data can be selected as much as necessary without throwing away. By doing so, it is possible to escape the burden of continuously processing all incoming data over time and efficient data processing is possible.

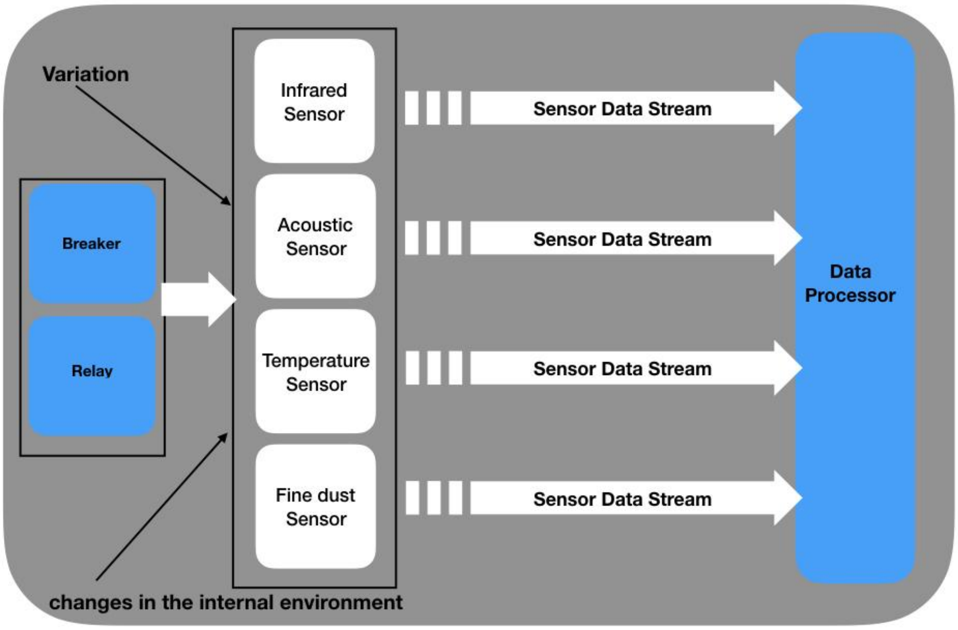
On the other hand, the data stream analysis/processing method proposed in this study cannot stop at finding meaning by reducing the throughput of data. It should also be possible to achieve results in improving the accuracy of situation information through selective data processing. The method proposed in this study has a structure that combines sensing data sent by different types of heterogeneous sensors. In many cases, situation recognition performance may be further improved by using data detected by heterogeneous sensors with different functions compared to recognizing a situation by using events of multiple sensor groups with the same function for the same situation. When intelligent forest fire recognition CCTVs were developed in Korea, a number of CCTVs were used, but there were cases where fog or clouds were recognized as smoke caused by fires. At this time, when the air pressure sensor or humidity sensor was mixed with CCTV, smoke, fog, and clouds caused by the fire were sufficiently identified. In this way, the use of sensing data from multiple sensors that are heterogeneous to each other has an aspect that benefits situational recognition or situational reasoning.

The proposed method in this study is to identify and combine events from data streams detected and transmitted by heterogeneous sensors of different types to recognize situations so that better situation information can be obtained than using sensor data detected and reported by a single sensor.

The following is a formula representing a data stream processing algorithm for situation recognition based on sensor data.

1) Whether an event occurs in any of the K sensors is determined by the step function below.
2) If an event occurs on any sensor, it is reported to the host.
3) If an event does not occur in other sensor data, whether the event continues and changes in any sensor where the event occurs are determined by the step function below.
4) In this case, when the event of the event-detected sensor data shows a decreasing pattern, event data of the corresponding sensor is not received.
5) It is divided at regular time intervals and explores whether each sensor data persists or whether the change in each sensor data event goes to a decreasing pattern.
6) Explore whether any of the events in each sensor data are interrupted.
7) When and, that is, if there is sensor data in which the event is stopped after the reduction pattern, the event occurs again is searched as a step function.
8) If an event occurs again in the sensor and continues, it returns to the process 9) to 11.
9) Interrupt date stream analysis when the sensor stops reporting events after each sensor's event data has entered a decreasing pattern at times and does not resume reporting events.
10) The analysis results are transmitted to the user.

The above algorithm deals with the entire process of processing and storing in response to a data stream introduced from a sensor.



Pic. 4 Data Stream Analysis

The issue covered in this study is the question of how to set the standards. For context recognition, it is to determine a sensor that transmits a reference data stream before checking events of another sensor data stream. In this case, it is preferable to select a sensor having a high correlation with the situation as the reference sensor. In other words, the sensor closest to the characteristics representing the main change in the situation should be the reference. In order to recognize a fire, a heat sensor should be a reference in terms of recognizing a site. In order to recognize fog, a humidity sensor must be the standard. In order to detect the situation of various electrical devices, a current detection sensor is recommended as a reference, and a heat detection sensor is recommended as a reference to detect various electrical abnormalities. In order to detect abnormal situations in a workplace where various industrial devices are installed, it can be said that it is good to use a vibration sensor as a reference. Chapter 4 of the following experiments with the contents proposed in this study.

IV. Experiment and Assessment

The following table shows the experiment with the sensing data reported from the sensor.

Table 1. Sensor Data Stream and Situation Recognition

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

16	84				an increase in humidity
17	86				an increase in humidity
18	87				an increase in humidity
19	86	12	2	2	a humid day
20	87	13	3	4	Rain
21	85	12	4	5	Rain
22	83	13	3	5	Rain
23	82	12	4	6	Rain
24	62				Lack of humidity
25	78				Lack of humidity
26	80				an increase in humidity
27	85				an increase in humidity
28	89				an increase in humidity
29	89				an increase in humidity
30	89	1013			Fog
31	91	1013			Fog
32	92	1758			Fog
33	92	1323			Fog
34	92	906			Fog
35	92	532			Fog
36	92	737			Fog
37	92	601			Fog
38	92	707			Fog
39	92	706			Fog
40	92	554			Fog
41	90	12	2	3	a humid day
42	89	11	3	4	a humid day
43	88	11	4	5	a humid day
44	87	11	On	On	Heavy rain
45	87	11	On	On	Heavy rain
46	81				an increase in humidity
47	88				an increase in humidity
48	89				an increase in humidity
49	89				an increase in humidity
50	89	12	0	0	a humid day
51	89	12	1	1	a humid day
52	88	12	2	2	a humid day
53	88	13	3	3	a humid day
54	88	12	4	4	Rain
55	87	12	On	On	Heavy rain
56	85	13	On	On	Heavy rain
57	86	13	On	On	Heavy rain
58	74				Lack of humidity
59	73				Lack of humidity
60	76				Lack of humidity

61	79				Lack of humidity
62	80				an increase in humidity
63	87				an increase in humidity
64	89				an increase in humidity
65	89				an increase in humidity
66	89	12	3	3	Rain
67	88	12	4	4	Rain
68	88	11	On	On	Heavy rain
69	88	12	On	On	Heavy rain
70	88	12	On	On	Heavy rain
71	92	212	0	4	a humid day

Table 1. Event Data from Sensors

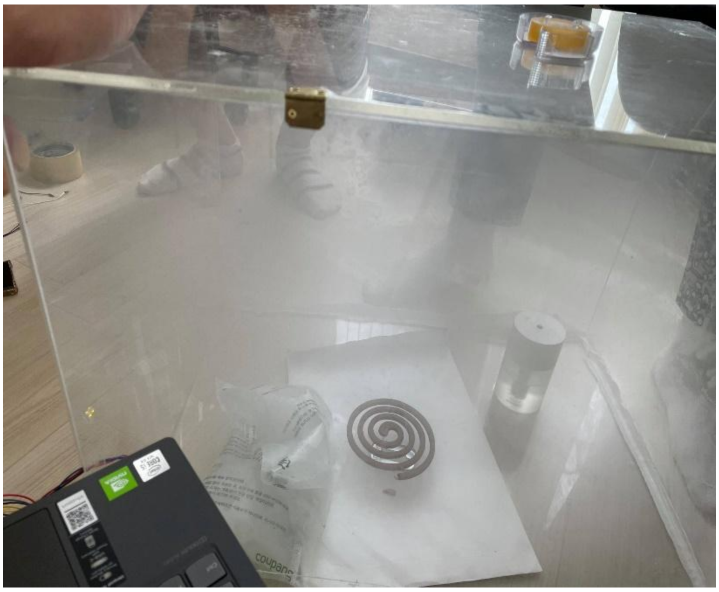
This experiment created an environment in which sensors installed in a closed space, such as a switchboard, transmit detected values, and execute them with data values flowing from each sensor at this time. The sensors used in this experiment were a humidity sensor, a vibration sensor, an acoustic sensor, and a fine dust detection sensor.

1) DHT-11 (humidity sensor)
· Operating voltage : 5 V
· Range : 20-90% / Accuracy : RH ± 5%
· Output format : digital switching output
2) SW-420 (vibration sensor, closed type)
· Operating voltage : DC 3.3V to 5V
· Output format : digital switching output
3) LM393 (acoustic sensor)
· Amplify the sound using a microphone
· Operating voltage: DC 4 to 6V
4) fine dust detection sensor
5) Experimental Platform: Raspberry Pi

Evaluation

Details: The closed narrow space is made of acrylic panels. The four types of sensors were installed inside, and an event was generated to detect and report them.

Evaluation: In this experiment, the humidity sensor was determined as the reference sensor. When the humidity suddenly increased (4), it was confirmed whether events occurred in other sensors, and it was investigated and confirmed that other events occurred in the data table (8). When the detected value of the humidity sensor, which is the reference sensor, decreased, data grouping was not performed, and then events of other sensors were investigated again based on the increase in 15. In (30), it was confirmed that other events occurred, and event cases were viewed and grouped again. In this experiment, the factor that has the greatest influence on electrical/electronic devices was considered humidity, and humidity was determined as a reference sensor according to expert opinions. By examining other events based on significant sensors in an electric device-related environment, it was shown that it is possible to recognize the situation for an internal environment of a specific device and differentiate it according to the situation by referring to refer to the events of other sensors.



Pic. 5 Get Data from Sensors

V. Conclusion

Various sensors are adopted in the IoT-based situation recognition system, and each sensor continuously sends detected data over time to the host. At this time, the necessary method should be to reduce the amount of data to be processed and improve situation recognition performance by selecting data in the section necessary for situation recognition from the continuously flowing sensor data. In this study, a sensor data stream that can be used as a reference among various sensors was selected, and a special change in the data stream from other sensors was detected based on this. As a result, it was possible to check the sensor data stream representing the detection result in connection with each situation, and data analysis was conducted by selecting a section with such connection activities. In this way, it was shown that while reducing the amount of data processed, it was possible to benefit from situational awareness. In the future research direction, follow-up research is needed on how to select sensors that play a role as a reference.

References

[1]. Kapitsaki, G., G. Prezerakos, N. Tselikas, and I. Venieris, "Context-aware service engineering: A survey", The Journal of Systems and Software, Elsevier, Vol.82, No.8, pp.1285-1297, 2009, doi: 10.1016/j.jss. 2009.02.026.

[2]. Knappmeyer, M., N. Baker, S. Liaquat, and R. Tonjes, "Context Provisioning Framework to Support Pervasive and Ubiquitous Applications", In Proceedings of the 4th European Conference on Smart Sensing and Context (EuroSSC 2009), pp.93-106, 2009

[3]. Lee, Y., S. Iyengar, C. Min, Y. Ju, S. Kang, T. Park, J. Lee, Y. Rhess, and J. Song, "MobiCon: A Mobile Context-Monitoring Platform", Communications of the ACM, Vol. 55, No.3, pp.54-65, 2012, doi: 10.1145/2093548.2093567.

[4]. Wang, Y., J. Lin, M. Annavaram, Q. Jacobson, J. Hong, B. Krishnamachari, and N. Sadeh, "A framework of energy efficient mobile sensing for automatic user state recognition", In Proceedings of the 7th international conference on Mobile systems, applications, and services(MobiSys ' 09), pp.172-192, 2009, doi: 10.1145/1555816.1555835.

[5]. Yoo, J. W., Lee, E. H., Kim, M. S., Jang, S. C., and Kim, S. H., “A Development of Perception of Situation Module”, Journal of Korean Institute of Information Technology, Proceedings of KIIT Conference, pp.331–334, June 2019.

[6]. Jin, S.Y., Lee, W. S., Kim, H. J., Jo, S. H., and Kang, Y. R., “A Study on Multiple Reasoning Technology for Intelligent Battlefield Situational Awareness”, JKICS, Vol.45, No.6, pp. 1046-1055, 2020

[7]. Na, S. H., and Lee, W. S., “Adaptive Realtime Context-Aware Technique of Intelligent Robot Based on Association Rule over Data Stream”, JKIIT, Vol.8, No.4, pp.141 – 151, Apr. 2010.

[8]. Kim, T. H., Suh, D. H., Yoon, S. S., and Ryu, K. H., “Noise Reduction in Real-time Context Aware using Wearable Device”, Journal of Digital Contents Society, Vol.19, No.9, pp.1803 – 1810, Sep.2018.

[9]. Kim, T. H., Suh, D. H., Yoon, S. S., and Ryu, K. H., “Error Correction of Real-time Situation Recognition using Smart Device”, Journal of Digital Contents Society, Vol.19, No.9, pp.1779 – 1785, Sep. 2018.

[10]. Lee, Y. D., “Distributing Board Monitoring System based on Internet of Things”, JKIICE, Vol.20, No.1, pp.200–206, Jan. 2016.

[11]. Suh, D. H., and Ryu, J. B., “Unbalanced Data Processing Method in Data Stream Environment of Smart Switchboard”, Journal of the Institute of Electronics and Information Engineers, Vol.58, No.1, pp.77 – 82, Jan. 2021.

[12]. Papageorgiou, I., Kontoyiannis, I., Mertzanis, L., Panotopoulou, A., and Skoularidou, M., “Revisiting Context-Tree Weighting for Bayesian Inference”, 2021 IEEE International Symposium on Information Theory (ISIT), pp.12-20, July 2021