



REVIEW

Data processing in internet of things networks

Tratamiento de datos en las redes del Internet de las cosas

Taleh Askerov¹, Vugar Abdullayev¹  , Vusala Abuzarova¹, Yitong Niu¹, Khushwant Singh¹

¹National Aviation Academy, Baku, Azerbaijan.

²Azerbaijan State Oil and Industry University, Baku, Azerbaijan.

³School of Industrial Technology, Universiti Sains Malaysia, Penang, Malaysia.

⁴University Institute of Engineering & Technology, Maharshi Dayanand University, Rohtak-124001, India, MDU, Rohtak -124001.

Cite as: Askerov T, Abdullayev V, Abuzarova V, Niu Y, Singh K. Data processing in internet of things networks. *Seminars in Medical Writing and Education*. 2024;3:.595. <https://doi.org/10.56294/mw2024.595>


Submitted: 26-12-2023

Revised: 19-04-2024

Accepted: 11-08-2024

Published: 12-08-2024

Editor: Dr. José Alejandro Rodríguez-Pérez 

Corresponding author: Vugar Abdullayev 

ABSTRACT

As an important component of the IoT ecosystem, data sets are an essential part of the decision-making process. IoT devices generate hundreds of new data sets every second and the problem of managing them appropriately arises. In the process of data management, their processing is a particularly complex and important process.

Various methods and tools are used to process data sets in the IoT ecosystem. Here, data processing allows you to speed up the decision-making process and make it less risky by transforming that data into the required form and making it relatively simple.

The article explores the concept of data, data management and processing in the IoT ecosystem and shows a simple example of data processing.

Keywords: Data Stream; Data Processing; Data Management; Internet of Things; Network.

RESUMEN

Como componente importante del ecosistema IoT, los conjuntos de datos son una parte esencial del proceso de toma de decisiones. Los dispositivos IoT generan cientos de nuevos conjuntos de datos cada segundo y surge el problema de gestionarlos adecuadamente. En el proceso de gestión de datos, su procesamiento es un proceso particularmente complejo e importante.

Se utilizan varios métodos y herramientas para procesar conjuntos de datos en el ecosistema IoT. Aquí, el procesamiento de datos permite acelerar el proceso de toma de decisiones y hacerlo menos riesgoso al transformar esos datos en la forma requerida y hacerlo relativamente simple.

El artículo explora el concepto de datos, la gestión y el procesamiento de datos en el ecosistema IoT y muestra un ejemplo simple de procesamiento de datos.

Palabras clave: Flujo de Datos; Procesamiento de Datos; Gestión de Datos; Internet de las Cosas; Red.

INTRODUCTION

At the stage of ICT development, the connection of all things that surround people (electrical devices, household equipment, vehicles, production facilities, etc.) to the global network has given rise to the term “Internet of things” (IoT). The extended experience of ICT in our daily life plays an important role in the

development of the emerging information society. In developed countries, ICT is used to improve the quality of human life by developing various innovative applications and services to solve society's problems. In modern times, the concept of IoT connects things using network technologies. In this regard, it should be noted that IoT is a network of uniquely identifiable connected things (devices, objects, etc.) that offer intelligent computing services.⁽⁴⁾ IoT-embedded objects are also known as "smart objects" or "smart devices" that enable the rational execution of everyday tasks. The main goal of the article is to examine the components that perform data processing with streams in the IoT network. For this we need to get acquainted with a number of terms.

Data in the Concept of the Internet of Things

A key component of the Internet of Things ecosystem is data. All processes are carried out on this component, and the main goal is to transfer data exchange in the right form at the right time. Which has a key role in the decision-making process.

In general, there are three main concepts related to data. Although they are similar, they have certain differences.⁽⁶⁾

Data: these are raw values obtained from various sources using various methods. In other words, it is a set of values obtained from the original source, but not yet processed (unprocessed). Access to such a set of values in any sector is quite easy, and their number is overwhelming.⁽³⁾

Information: the second concept that comes after data is a set of data obtained by processing, interpreting or structuring data within a specific content or field. In other words, information is a set of processed data.⁽³⁾

Knowledge: the last and most important concept is the collection of information - what we know - derived from information and personalized. Basically, knowledge with a more specific purpose is directed in a specific direction. It is basically a direction that requires a specific act (movement) that has a final goal, such as decision-making, etc.⁽¹²⁾

The relationship among these concepts can be shown as follows: Data- Information- Knowledge.⁽³⁾

What exists in the Internet of Things Ecosystem is data. Their transformation into information or knowledge depends on the subsequent process and degree of necessity.

One of the key approaches here is the management of this data. Data management is a very complex and sensitive process. Because proper data management is important in the decision-making process.

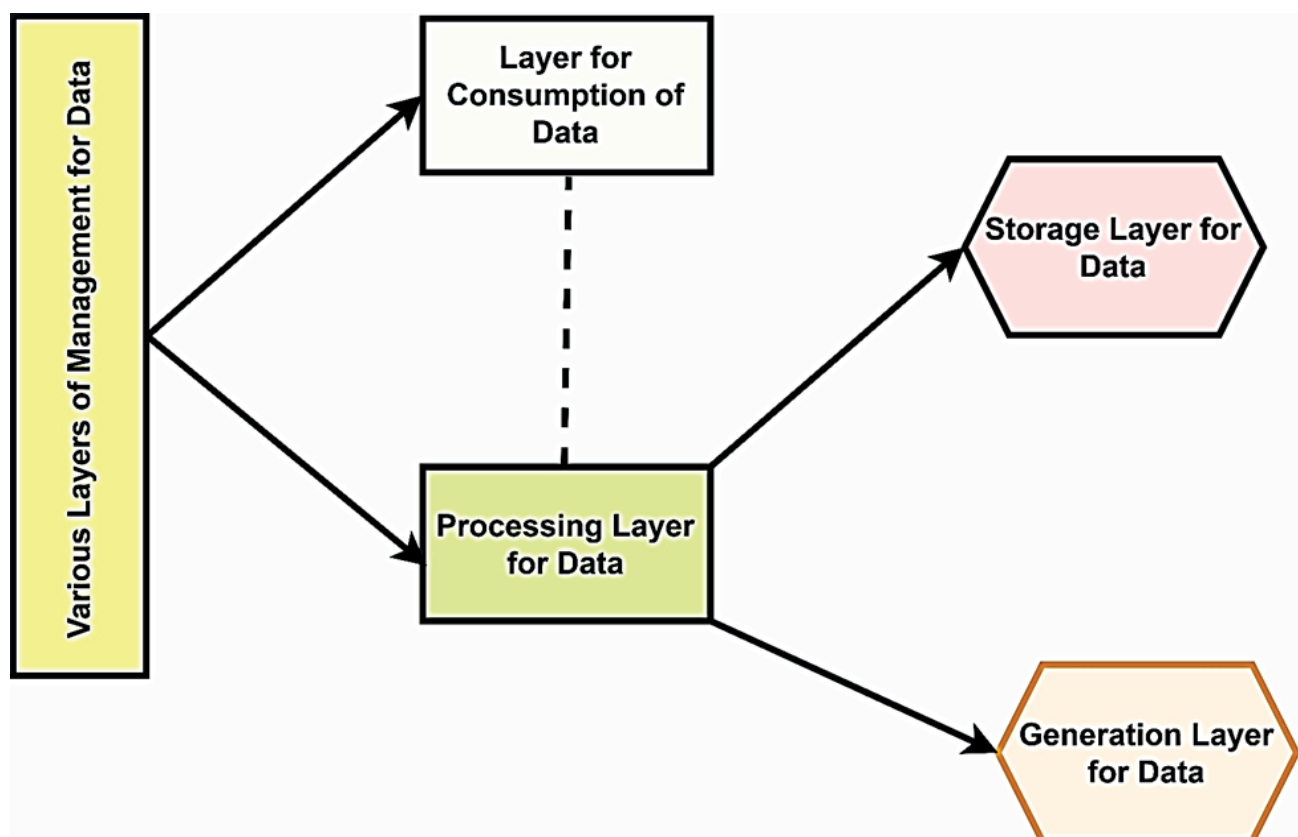


Figure 1. Various layers of Data Management

Although data management is a broader concept, its foundation is the life cycle of data. The life cycle of data consists of the following stages:

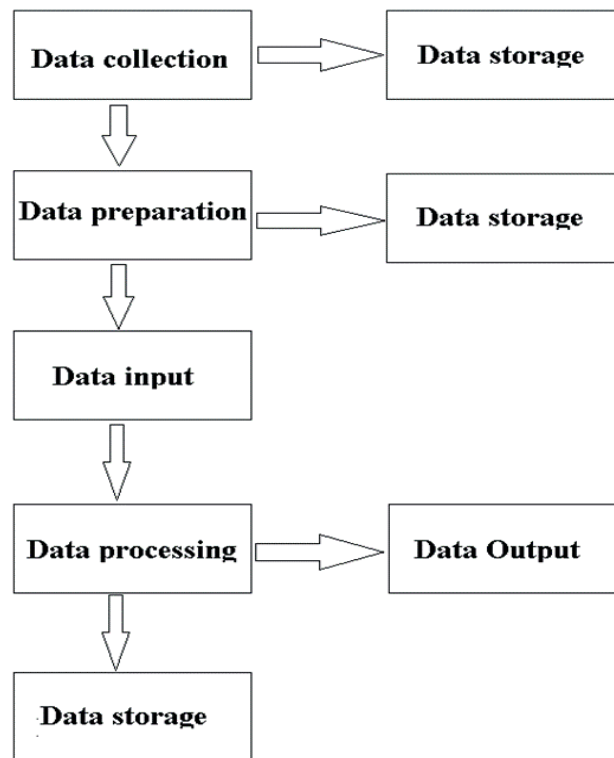


Figure 2. Data

Data Collection - In the IoT ecosystem, smart devices generate new data every second. The process of collecting this data is carried out in various ways depending on demand.

Sensors are the main data collection tool in the IoT ecosystem.

Data preparation and input - there are basically 4 basic processes in data management: data collection, storage, processing and output. Here, data preparation and input is part of the data processing process and is characterized as pre-processing.

This basically involves cleaning, filtering and inputting data for further processing.

Data processing - Data processing refers to the process of extracting the required data by performing various operations on the collected data. And depending on the situation, this data may then be transferred to a third party.

Data processing is the most complex stage in the data management process. It can be performed in different ways depending on the requirements. But the result always ends with the transformation of raw data into the required data.

Data output - after processing, the required data is transmitted to the output. And if necessary, the process of transmitting this data is also performed.

Data storage - the stage that is repeatedly performed throughout the lifecycle is the storage stage. Although it is mentioned as a separate phase, it starts from the data collection process and continues iteratively. In other words, it is repeated.

Data is stored after it is received, data is stored again after the preparation process is completed, and data is stored again after the processing.

In an IoT ecosystem, the communication and data flow between users and devices can be summarized as follows:

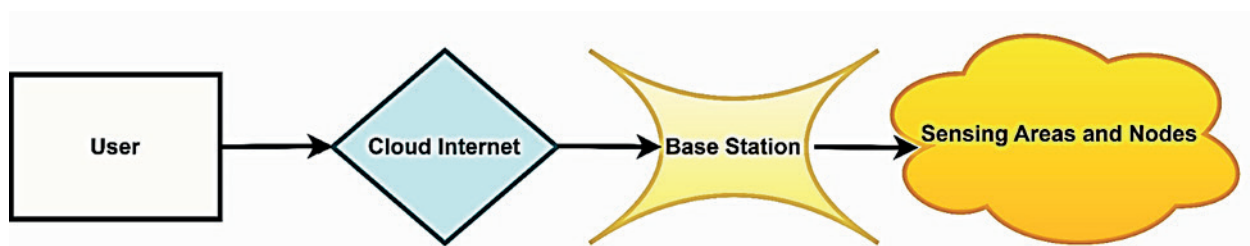


Figure 3. Connection between users and IoT ecosystem devices

All these steps are applied in the same way in the IoT ecosystem. This can be seen in the simplest IoT

architecture. The simplest IoT architecture is as follows:

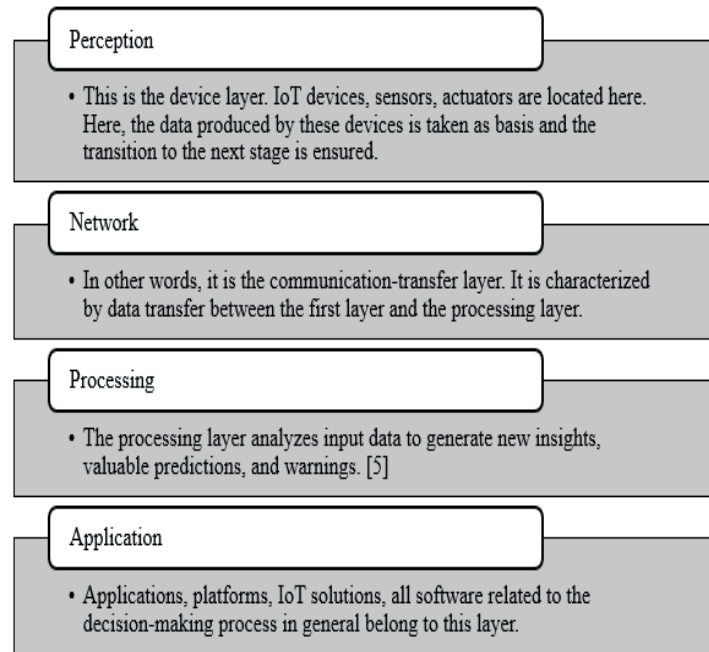


Figure 4. Architecture

The role of Sensors in the IoT ecosystem

As mentioned earlier, the data collection process in the IoT ecosystem is characterized mainly by sensors. Sensors are one of the main parts of the IoT ecosystem, the main agents in the realization of the main goal of the ecosystem.

Sensors are devices that measure an event or change in the environment and convert it into an electronic signal that can be read and calculated. Sensors can detect any aspect of the physical environment and turn it into useful data. Thus, sensors convert stimuli such as heat, light, sound and movement into electrical signals. These signals pass through an interface that converts them into binary code and transmits them to a computer for processing.⁽¹⁾

A sensor network is a group of sensors where each sensor monitors data in a different location and sends that data to a central location for storage, viewing, and analysis. There are two types of sensor networks, wired and wireless. Sensor network components include sensor nodes, sensors, gateways, and control nodes. The four topologies of sensor networks are point-to-point, star, tree, and mesh.^(1,7)

The architecture of the Wireless Sensors is as follows:

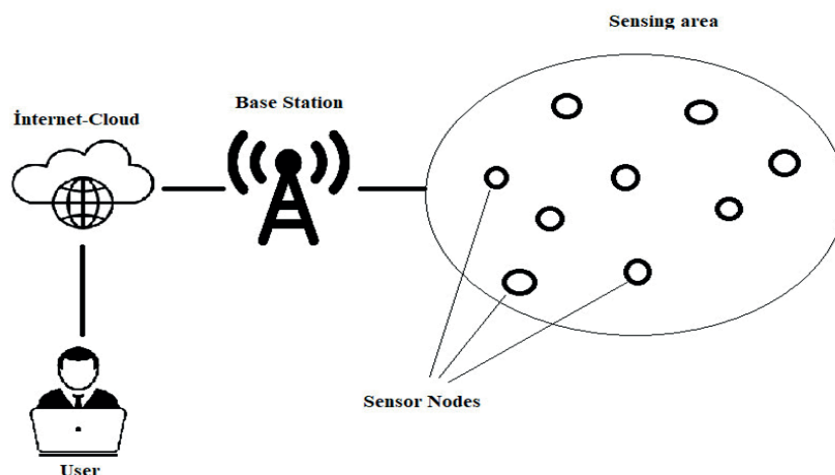


Figure 5. Wireless Sensor Architecture

Looking at IoT architecture, it can be noted that the most basic architecture consists of three layers:

perception, network and application. The first layer contains sensors to collect data about the environment. The second layer connects smart things devices and servers. The third layer provides application services to users.⁽⁸⁾ Here it is possible to understand the degree of use and demand of sensors for the IoT ecosystem.

The main differences that allowed the IoT market to grow faster were:⁽⁹⁾

- Miniaturization of sensors - Technological advances have led to the use of technologies such as microelectromechanical systems (MEMS) to produce sensors on a microscopic scale. This made the sensors small enough to be placed in unique locations such as clothing.
- Improved communication - Wireless connectivity and communication technologies have improved to the point that almost any type of electronic equipment can provide a wireless data connection. These enabled sensors embedded in connected devices to send and receive data over the network.

The data collected by the sensors are then transmitted in real time for analysis. This is basically called data stream.

Data analytics in the IoT ecosystem

One of the main parts of data management in the Internet of Things ecosystem is their analysis.

Data analysis is the process of extracting meaningful data from the various data sets available. This process is mainly done on Big Data. However, the application areas of data analysis are wide. That is, it can be applied to obtain meaningful (necessary) data in many areas. With the help of data analytics, individual or group decision making can be done.⁽¹⁾

There are many types of analytics:⁽¹⁾

- Descriptive analysis: This is the very first type of analysis. Descriptive data is a very large part of all information. Here, meaningful data is obtained from those descriptions. For example, medical descriptive data analysis in Medicine.
- The best thing about descriptive data is that it can describe large amounts of data in a concise and simple form.
- Diagnostic analysis: This is the second type of analysis, which is mainly carried out after descriptive analysis. The data obtained from the descriptive analysis are taken and the analysis is carried out to find the reason for these results.
- Predictive analysis: Predictive analysis is the third type of analysis. As the name suggests, this analysis is used to predict what may happen in the future. In predictive analysis, mainly Machine Learning and Deep Learning algorithms are applied to predict future trends, problems, solutions through the results obtained from diagnostic analysis.
- Instruction analysis: This type of analysis is the most important. At this phase, it is investigated how the results obtained from predictive analysis - predictions - will happen. This analysis allows users to understand future events. In general, data from all previous analyzes is combined in decision making.
- Each of these types is related to each other and they are applied in unity as a part of one general system.

Data Processing in IoT Network

Given the IoT architecture and Data Management mentioned above, we can look at the data processing in the IoT network. Generally, it can take two forms - Edge computing and Cloud computing. AI algorithms, Machine Learning and big data methods are used for data processing.

But security is also important here, which is mainly related to processing such data on the cloud. Using cloud-based systems to process IoT data has a number of limitations, including security risks, latency and missed opportunities to work on powerful real-time concepts.⁽¹¹⁾

Let's look at data processing from a simple IoT network. The data set is taken from <https://www.kaggle.com/datasets/garystafford/environmental-sensor-data-132k?resource=download>.

Table 1. Data					
ts	Device	Humidity	Light	Smoke	Temp
1594512094.38597	b8:27:eb:bf:9d:51	51	FALSE	0,020411	22,7
1594512094.73556	00:0f:00:70:91:0a	76	FALSE	0,013275	19,7
1594512098.07357	b8:27:eb:bf:9d:51	50,9	FALSE	0,020475	22,6
1594512099.58914	1c:bf:ce:15:ec:4d	76,8	TRUE	0,018628	27
1594512101.76123	b8:27:eb:bf:9d:51	50,9	FALSE	0,020448	22,6
1594512104.46841	1c:bf:ce:15:ec:4d	77,9	TRUE	0,018589	27
1594512105.44886	b8:27:eb:bf:9d:51	50,9	FALSE	0,020475	22,6

1594512106.86907	00:0f:00:70:91:0a	76	FALSE	0,013628	19,7
1594512108.27538	1c:bf:ce:15:ec:4d	77,9	TRUE	0,01844	27
1594512109.13668	b8:27:eb:bf:9d:51	50,9	FALSE	0,020457	22,6
1594512112.79851	b8:27:eb:bf:9d:51	50,9	FALSE	0,020425	22,6
1594512115.28854	1c:bf:ce:15:ec:4d	78	TRUE	0,018563	27

The table 1 above shows 12 rows from the dataset as an example. On the other hand, there are more than 40 000 rows in the dataset and they were analyzed. Here we have only used data such as timestamp, device number, humidity, light, smoke and temperature. (There are additional columns in the dataset)

To process data such as humidity, light, smoke and temperature received through sensors, methods such as reading the data, transforming the data, analyzing and visualizing the values over time have been applied. This is the simplest example of data processing.

The images below are the results of data analysis in python:

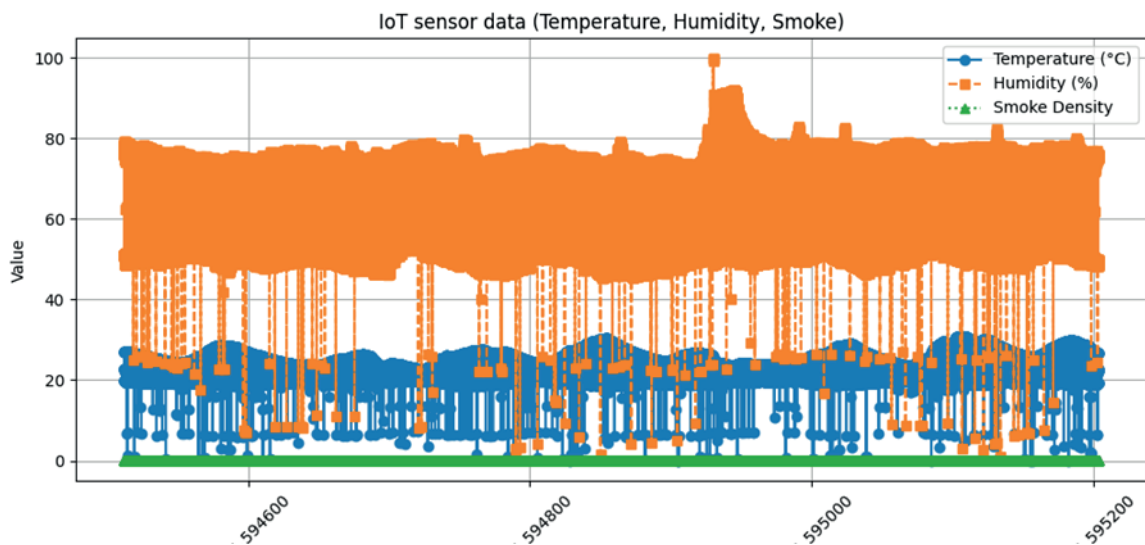


Figure 6. IoT sensor data analysis

Here the proportion of data acquired through IoT sensors is shown. Ambient temperature, humidity and smoke density and their distribution over time are shown.

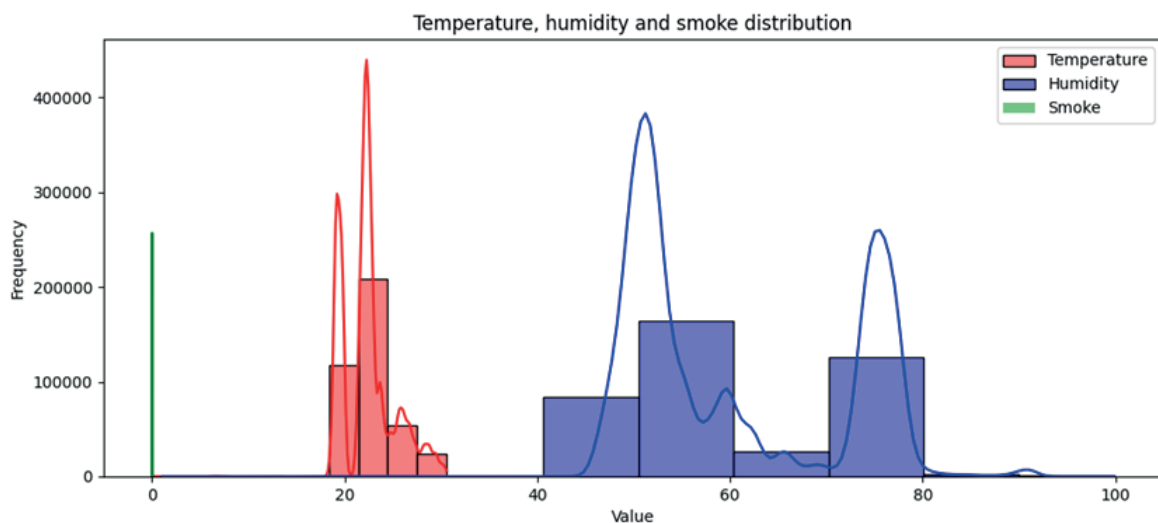


Figure 7. Temperature, humidity and smoke distribution

The same data sets - by frequency and value - are visualized with a histogram.

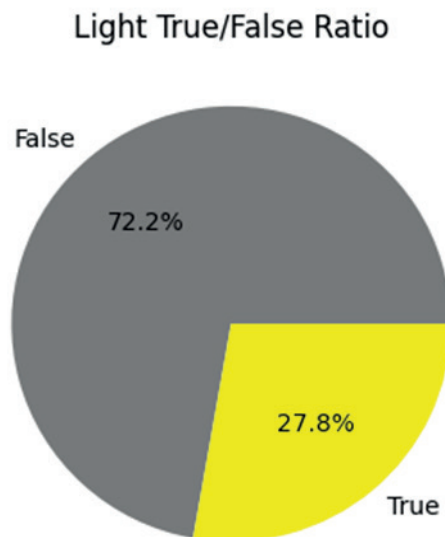


Figure 8. Light Ratio: True of False

By analyzing the data, it was determined whether the environment in which the sensors obtained data was unlit or not. Accordingly, 72,2 % of the data was obtained in a non-illuminated environment and 27,8 % in an illuminated environment.

CONCLUSION

Data is the most fundamental component of the IoT ecosystem. In fact, this ecosystem is built for data exchange. The volume of data in constant motion continues to grow every second, which ultimately poses the problem of data processing. This negatively affects the decision-making process.

In this respect, it is important to manage this data correctly. The main issue here is their processing. In the IoT ecosystem, data collected mainly through sensors passes from pre-processing to the main processing stage. Various methods are applied here depending on the data, its volume and specific area. Artificial intelligence, machine learning algorithms, big data and cloud computing tools are mainly applied.

The images above are the results of a simple data processing process. Thus, an IoT data set of more than 40 000 rows was collected through sensors. Analysis, transformation and visualization were performed on this data.

Considering that there are larger volumes and more complex data in the real-time IoT ecosystem, it would be appropriate to use Big Data analytics methods, artificial intelligence, deep learning, machine learning algorithms to work with such data and to store such data mainly in the cloud.

REFERENCES

1. Alex Khang, Vugar Abdullayev, Babasaheb Jadhav, Shashi Kant Gupta, Gilbert Morris. "AI- Centric Modeling and Analytics - Concepts, Technologies, and Applications", CRC Press, 2023
2. Alex Khang, Sita Rani, Arun Kumar Sivaraman. "AI-Centric Smart City Ecosystems - Technologies, Design, and Implementation", CRC Press, 2022
3. Alex Khang. "AI and IoT Technology and Applications for Smart Healthcare Systems", CRC Press, 2024
4. Muhammad Azhar Iqbal, Sajjad Hussain, Huanlai Xing, Muhammad Imran. "Enabling the Internet of Things", Wiley, 2021
5. Liubomyr Maievskyi, Solution Engineer, "IoT Architecture", 20 Feb. 2023, URL: <https://limestonedigital.com/iot-architecture/>
6. Khang, Alex, et al. "The Key Assistant of Smart City: Sensors and Tools." AI-Centric Smart City Ecosystems. CRC Press, 2022. 271-280.
7. Miranda Junior, Hamilton & Bezerra, Nelson & Bezerra, Marlene & Farias Filho, José. (2017). The internet of things sensors technologies and their applications for complex engineering projects: a digital construction site framework. Brazilian Journal of Operations & Production Management. 14. 567. 10.14488/BJOPM.2017.

v14.n4.a12.

8. Daniel Teachey, “Making sense of streaming data in the Internet of Things”, https://www.sas.com/en_us/insights/articles/big-data/making-sense-of-streaming-data-in-the-internet-of-things.html
9. Alaasam A.B.A. et al. Analytic Study of Containerizing Stateful Stream Processing as Microservice to Support Digital Twins in Fog Computing // *Programming and Computer Software*. 2020. Vol. 46, no. 8. P. 511-525. DOI:10.1134/S0361768820080083
10. “Real-Time Processing of Data for IoT Applications”, August 29, 2021, URL: <https://www.3pillarglobal.com/insights/blog/real-time-processing-of-data-for-iot-applications/>
11. Meehan J., Zdonik S. Data Ingestion for the Connected World // *Cidr*. 2017. <https://builtin.com/data-science/types-of-data-analysis>
12. Triwiyanto T, Vugar Abdullayev, Sabir Mammadov, Shafi Danyalov, Latafat Mikailzade, Ibrahim Abbasov, Taleh Asgarov, Bahar Asgarova, Zarifa Imanova, Vusala Abuzarova, “Significance and Processing of Signals”, 1-12, 2024, <https://itta.cyber.az/2024/papers/42.html>
13. G. Mammadova et al. (Eds.): ITTA 2024, Part 3, pp. 1-12, 2024. <https://doi.org/10.54381/itta2024.42>
14. Margara A., Rabl T. Definition of Data Streams // *Encyclopedia of Big Data Technologies*. Cham: Springer International Publishing, 2019. P. 648- 652. DOI:10.1007/978-3-319-77525-8_188
15. Gavaldà R. Adaptive Windowing // *Encyclopedia of Big Data Technologies*. Cham: Springer International Publishing, 2018. P. 1-6.DOI:10.1007/978-3-319-63962-8_194-1
16. Golab L. Types of Stream Processing Algorithms // *Encyclopedia of Big Data Technologies*. Cham: Springer International Publishing, 2019. P. 1726-1732. DOI:10.1007/978-3-319-77525-8_193.
17. MapReduce Tutorial [Electronic resource]. https://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.html (accessed: 17.10.2021).
18. Shahrivari S. Beyond batch processing: Towards real-time and streaming big data // *Computers*. 2014. Vol. 3, no. 4. P. 117-129. DOI:10.3390/computers3040117
19. Agneeswaran V.S. Big Data Analytics Beyond Hadoop: Real-Time Applications with Storm, Spark, and More Hadoop Alternatives. 2014.
20. Kambatla K. et al. Trends in big data analytics // *Journal of Parallel and Distributed Computing*. 2014. Vol. 74, no. 7. P. 2561-2573. DOI:10.1016/j.jpdc.2014.01.003.
21. Andrade H., Gedik B., Turaga D. Fundamentals of Stream Processing // *Fundamentals of Stream Processing*. Cambridge: Cambridge University Press, 2014. DOI:10.1017/CBO9781139058940.
22. Romero C., Oliveira H.P. Kafka: a Distributed Messaging System for Log Processing // *Proceedings of 6th international workshop on networking meets databases (NetDB)*. Athens, Greece, 2011.
23. Ragimova, N.A., et al. “The Era Of Digital Health And Its Impact On Human Psychology.” 1st INTERNATIONAL CONFERENCE ON THE 4th INDUSTRIAL REVOLUTION AND INFORMATION TECHNOLOGY. Vol. 1. No. 1. Azerbaijan State Oil and Industry University, 2023.
24. Isah H. et al. A Survey of Distributed Data Stream Processing Frameworks // *IEEE Access*. IEEE, 2019. Vol. 7, no. October. P. 154300-154316. DOI:10.1109/ACCESS.2019.2946884. <https://managment.ageditor.uy/index.php/managment>
25. Meehan J., Zdonik S. Data Ingestion for the Connected World // *Cidr*. 2017.
26. Gualtieri M., Yuhanna N. The forrester wave: Big data streaming analytics, Q1 2016 // *Forrester research*. Cambridge, MA, USA, 2016. 15 p. <https://dm.ageditor.ar/index.php/dm>

27. Henning S., Hasselbring W. Theodolite: Scalability Benchmarking of Distributed Stream Processing Engines in Microservice Architectures // Big Data Research. 2021. Vol. 25. DOI:10.1016/j.bdr.2021.100209.
28. Vugar Abdullayev Hajimahmud, Ragimova Nazila Ali, Triwiyanto, Asgarov Taleh Kamran, Mammadov Kanan Hafiz, Abuzarova Vusala Alyar / Application of Industrial Internet of Things Technologies in the Manufacturing Industry/ Book: Machine Vision and Industrial Robotics in Manufacturing, 1st Edition, CRC Press, 2024.
29. Gomes, M. M., Righi, R. da R., da Costa, C. A., & Griebler, D. (2022). STEAM++: An Extensible End-To-End Framework for Developing IoT Data Processing Applications in the Fog. arXiv. <https://arxiv.org/abs/2205.03271>
30. Khan, M. A., Khan, S. U., & Madani, S. A. (2023). Machine Learning Techniques and Big Data Analysis for Internet of Things Applications: A Review Study. ResearchGate. <https://www.researchgate.net/publication/362224734>
31. Mohammadi, M., Al-Fuqaha, A., Sorour, S., & Guizani, M. (2017). Deep Learning for IoT Big Data and Streaming Analytics: A Survey. arXiv. <https://arxiv.org/abs/1712.04301>
32. Sethi, S., Sarangi, S. A., & Panda, S. (2021). IoT Data Preprocessing: A Survey. Webology. <https://www.webology.org/data-cms/articles/20220309031031pmwebology%2018%20%286%29%20-%2021%20pdf.pdf>
33. Usama, M., Jan, M. A., Amanullah, M., & Khan, S. (2023). An Optimized IoT-enabled Big Data Analytics Architecture for Edge-Cloud Computing Using Machine Learning. PMC. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10691823>

FINANCING

This work is supported from Jadara University under grant number [Jadara-SR-Full2023], and Zarqa University.

CONFLICT OF INTEREST

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Data curation: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Formal analysis: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Research: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Methodology: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Project management: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Resources: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Software: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Supervision: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Validation: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Display: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Drafting - original draft: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.

Writing: Taleh Askerov, Vugar Abdullayev, Vusala Abuzarova, Yitong Niu, Khushwant Singh.