Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 19, SEPTEMBER 2024

DOI: 10.37943/19UQOC7381

Aigul Adamova
PhD, Assistant Professor, Department of Computer Engineering
aigul.adamova@astanait.edu.kz, orcid.org/0000-0001-7773-9522
Astana IT University, Kazakhstan

Tamara Zhukabayeva
PhD, Professor, Department of Information Systems
tamara_kokenovna@mail.ru, orcid.org/0000-0001-6345-5211
L.N. Gumilyov Eurasian National University, Kazakhstan
Astana IT University, Kazakhstan

DEVELOPMENT OF A METHODOLOGY FOR DATA NORMALISATION AND AGGREGATION TO ENHANCE SECURITY LEVELS IN INTERNET OF THINGS INTERACTIONS

Abstract: The number of interacting devices is increasing every day, and with this constant innovation, serious security challenges arise. The concept of the Internet of Things is being actively applied in both domestic and industrial settings. Researchers are increasingly highlighting the challenges and importance of network security. Data preprocessing plays an important role in security by transforming the input data corresponding to algorithmic criteria and thereby contributing to the prediction accuracy. The data preprocessing process is determined by many factors, including the processing algorithm, the data, and the application. Moreover, in Internet of Things interactions, data normalisation and aggregation can significantly improve security and reduce the amount of data used further decision making. This paper discusses the challenges of data normalisation and aggregation in the IoT to handle large amounts of data generated by multiple connected IoT devices. A secure data normalisation and aggregation method promotes successful minimised data transfer over the network and provides scalability to meet the increasing demands of IoT deployment. The proposed work presents approaches used in data aggregation protocols that address interference, fault tolerance, security and mobility issues. A local aggregation approach using the run-length encoding algorithm is presented. The proposed technique consists of data acquisition, data preprocessing, data normalisation and data aggregation steps. Data normalisation was performed via the Z-score algorithm, and the LEACH algorithm was used for data aggregation. In the experimental study, the percentage of faulty nodes reached 35%. The performance of the proposed solution was 0.82. The results demonstrate a reduction in resource consumption while maintaining the value and integrity of the data.

Keywords: Internet of Things; security; data normalisation; data aggregation; z-score; LEACH

Introduction

In today's world, digitalisation is a key driver of economic growth. Kazakhstan is developing e-commerce, digital financial services and digital platforms to improve business processes. One of the areas of digitalisation is the Internet of Things (IoT). An increasing number of devices are becoming 'smart', with the IoT network penetrating homes, cities, industry, medicine and

Received: 29.07.2024 Accepted: 24.09.2024 Published: 30.09.2024

agriculture. This opens new opportunities to improve efficiency, comfort and safety. Kazakhstan aims to be at the forefront of change, opening new horizons of opportunity for its citizens. According to early estimates by research firm J'son & Partners Consulting, the IoT market size in Kazakhstan will grow at a CAGR of 21.2% and will triple the market size in 2019 by the end of 2024 [1]. The benefit of the IoT is reflected in the global numbers as well. According to Statista, the number of IoT-connected devices worldwide is expected to grow from 8-6 billion in 2019 to 29.42 billion by 2030 [2].

As IoT devices become more pervasive, security and privacy become more important. Many IoT devices are vulnerable to various cyber threats that can jeopardise the security and privacy of sensitive data [3]. IoT devices collect large amounts of personal data, raising concerns about privacy and data protection [4]. This study explored data normalisation and aggregation methods to improve data security when interacting with the Internet of Things.

S. Abbasian Dehkordi, et al. presented research results of data aggregation methods in wireless sensor network (WSN) with respect to terrestrial, underground, underwater and body functional domains. The applications, advantages and disadvantages of using each method have been described [5]. The authors also note that no comprehensive and consistent approach to data aggregation has been presented that is able to combine security, communication overhead, energy consumption and the data compression ratio.

In IoT interactions, a key issue is how important information can be processed in a more energy-efficient way. Kamal Gulati, et al. highlighted the importance of energy-efficient data aggregation techniques. As a result of their research, they presented various approaches and algorithms for energy-efficient data aggregation in IoT-WSN systems [6].

Liu, X. et al. proposed a data aggregation method to guarantee IoT security, eliminate data collection redundancy and improve the energy efficiency of smart nodes. The proposed approach results in higher security, longer lifetimes and better accuracy. Although the security assurance is demonstrated from the experimental results, but at the same time, it is important to note that further research is needed to investigate the impact of different packet loss rates on aggregation accuracy [7].

To improve the reliability level and uptime of the network supporting IoT services, Haseeb, K. et al. proposed a structure-based data aggregation method with optimal data forwarding processing [8]. In [9], the authors proposed a method for edge computing that supports the IoT. The proposed method covers three mechanisms, namely, block header construction, sensitive task decomposition and task receiver separation to prevent privacy disclosure. By exploiting the distributed nature and other advanced features of fog computing, the authors of [10] proposed an approach that reduces communication overheads and energy consumption while maintaining secure and reliable aggregation of healthcare data between medical sensors and cloud servers. Diène, B., et al. presented a classification of data aggregation techniques based on a network model, topology, key cryptography technique, encryption method, application, authentication mechanism and data recovery capability to ensure security [11].

In general, data normalisation and aggregation for secure IoT interactions is an important concept, as evidenced by the existence of various studies in this direction. However, the covered studies, within the scope of our study, do not present an approach that combines elements of both normalisation and aggregation of data received from IoT devices. The aim of study is to develop and evaluate the effectiveness of a data normalisation and aggregation approach to improve the security of data transmission in IoT networks. To achieve this goal, the researchers studied multiple methods for collecting data in IoT interactions, developed a new approach based on the Z-score and LEACH algorithms, and then conducted a real-time experimental study of the proposed approach to determine how well it works.

Methods and Materials

This section discusses the architecture of the IoT, presents possible threats with respect to each layer of the architecture, various data normalisation methods, and data aggregation techniques, and describes the local aggregation algorithm.

Architecture. Smart cities, smart manufacturing, healthcare, and agriculture are some of the types of IoT applications where data normalisation and aggregation are used. For example, in the smart city sector, data normalisation and aggregation can optimise the management of city infrastructure operations. In the healthcare sector, it can help improve patient care [12]. In addition to these benefits, there are also threats. The traditional IoT architecture consists of three layers: the physical device layer, the network layer and the application layer [13], [14]. Figure 1 summarises the levels of the IoT architecture and the possible threats at each level.

IoT devices generate large volumes of data that need to be collected and transmitted to servers for processing. The IoT architecture must be designed to ensure reliable and efficient data transmission, considering bandwidth and network constraints. The collected data need to be processed and stored for further analysis. The IoT architecture should support various data processing techniques such as filtering, aggregation and analysis. The data storage must be scalable and reliable to accommodate large amounts of data.

In summary, IoT architecture, data normalisation and aggregation are complementary elements that work together to ensure that IoT systems function effectively.

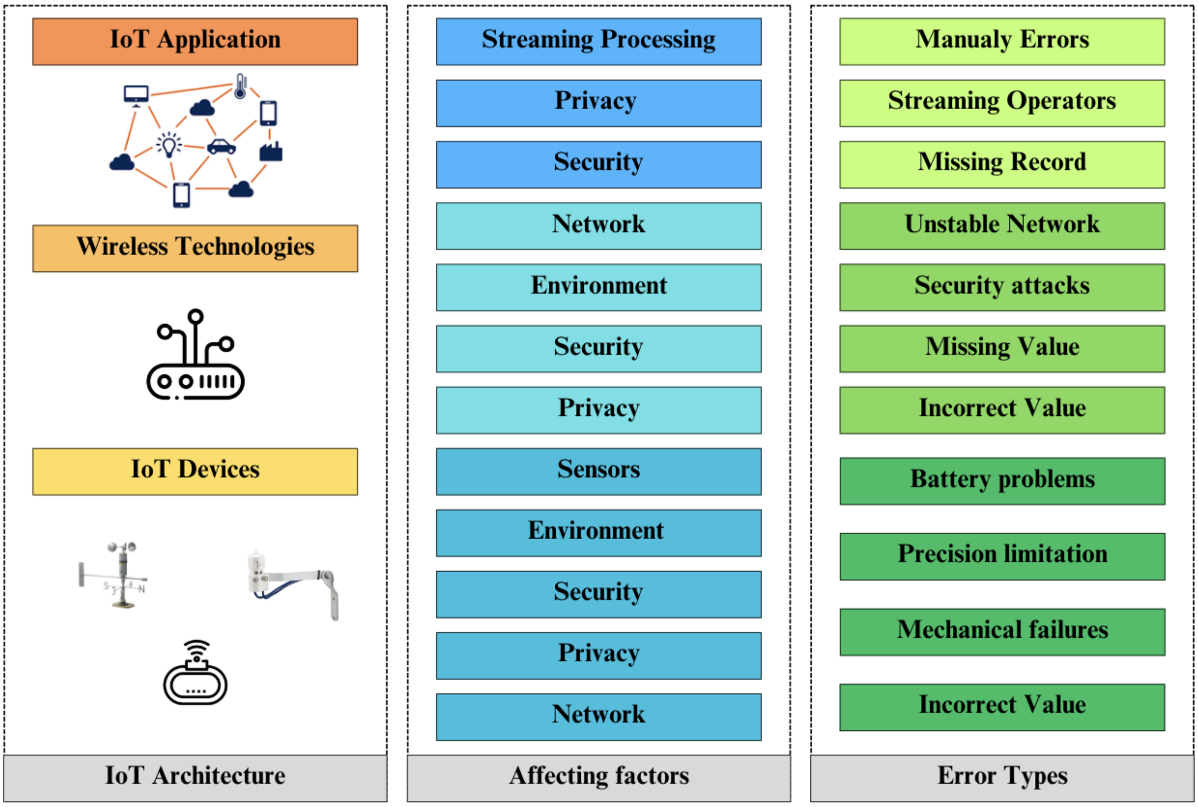


Figure 1. Threats in relation to each layer of the architecture

Data normalisation methods. The process of data normalisation, which aims to bring data values to a common scale [15]. Data normalisation is the process of preparing clean data [16], [17]. Through data normalisation, data are organised so that they become similar in all records and fields. This standardisation increases the cohesion of different records, allowing data cleaning and improving data quality. The process ensures that data are stored logically, eliminating unstructured data and data redundancy. If data normalisation is carried out according to the steps shown in Figure 2, the result is a standardised input of information.

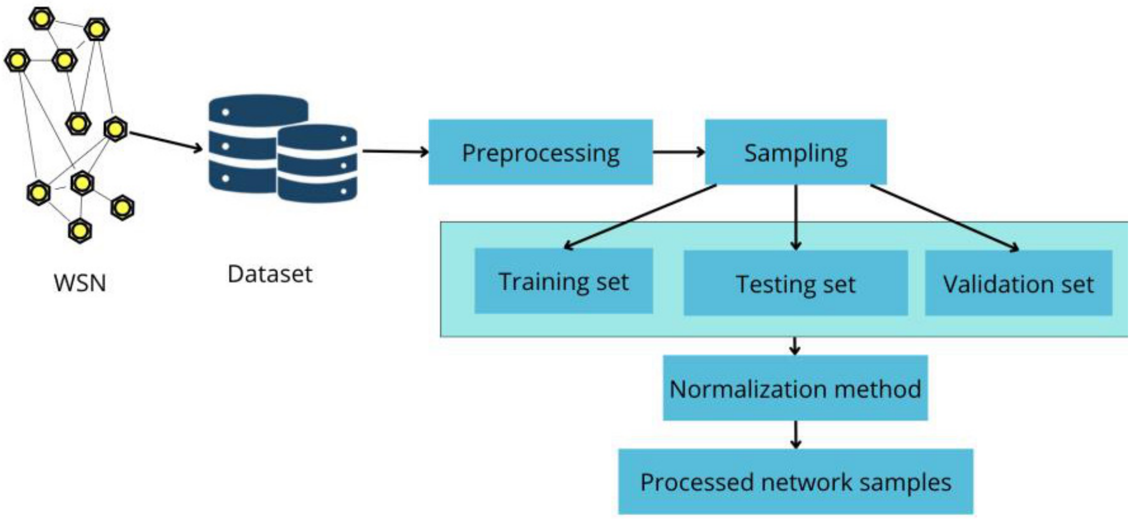


Figure 2. Stages of data normalisation

When selecting the IoT data normalisation method, methods such as Min-Max Scaling, Z-score normalisation, Decimal scaling, Unit conversion, and Encoding categorical data were investigated. The advantages and disadvantages of the data normalisation methods have been identified and are presented in Table 1.

Table 1. IoT data normalisation methods

Method	Work	Advantages	Disadvantages
Min-Max Scaling	[18, 19]	Simple and efficient for numerical data with a well-defined value range	Sensitive to outliers that can skew the scaling
Z-score normalisation	[19, 20]	Less sensitive to outliers compared to Min-Max scaling	Requires knowledge of the entire dataset for calculating mean and standard deviation
Decimal scaling	[21]	Preserves the relative order of data points and is efficient for data with similar scales	Doesn't address inherent differences in data ranges or units
Unit conversion	[22]	Enables comparison of data from devices using different sensors or measurement systems	Requires knowledge of the original units and potential loss of information during conversion
Encoding categorical data	[23]	Enables algorithms to understand relationships between categorical variables	Can increase data dimensionality depending on the number of categories

The choice of the appropriate normalisation method depends on the type of data, the desired result and the purpose of using the normalised data. Importantly, normalisation methods are more suitable for numeric data, whereas categorical data require coding methods.

Data aggregation techniques. The process of data aggregation combines numerical values into a single representative value, a process that is performed via aggregation functions [24]. These functions are used when data reduction, data quality enhancement, information extraction, etc., are needed. The different data aggregation techniques are shown in Figure 3.

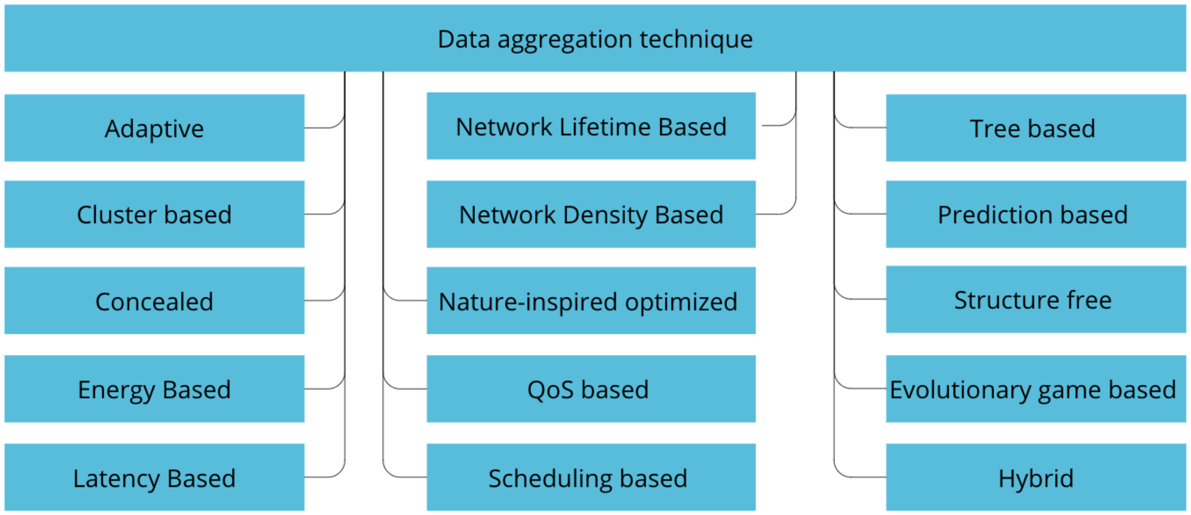


Figure 3. Data aggregation techniques

The IoT architecture must support different levels of data aggregation, from simple summaries to complex analytical models. Data aggregation needs to guarantee security and privacy. Table 2 summarises the aggregation approaches used in the research.

Table 2. Analysing research on IoT data aggregation

Paper	Year	Main goal	Open issue
[25]	2024	Investigation of data aggregation by adding noise, random permutation and parameter estimation	Trade-offs between utility and privacy are not covered
[26]	2024	The authors propose a scheme for data aggregation and exchange using homomorphic encryption with multiple keys	Technological errors are not highlighted
[27]	2024	A fog-based data aggregation method using Douglas-Peucker algorithm and random response mechanism is investigated	Integration of homomorphic encryption techniques for data integrity assurance
[28]	2024	An approach for securely collecting and transmitting aggregated data is investigated	Aggregation and redundancy-aware data transfer based on sensor node
[29]	2024	A data aggregation scheme with multifunctionality is proposed	A privacy-focused scheme using common cryptographic techniques is investigated
[30]	2024	The authors propose a parameter aggregation technique using a mechanism to protect client privacy in federated learning process	An optimisation technique for the FL system from a local learning perspective
[31]	2024	Data aggregation protocols are investigated to address network topology, interference, fault tolerance, mobility and security issues	Optimal utilisation of network resources, workflow management and routing protocols

The above research uses various aggregation mechanisms to ensure privacy, security and efficiency in general. However, the direction of data aggregation is not exhausted and has several open questions concerning trade-offs between utility and privacy, robustness to technological failures, integration with encryption techniques and optimisation of network resources.

Localised aggregation. The Run-Length Encoding algorithm provides a simple and efficient way to compress data with long sequences of repetitive symbols [32]. With a linear time complexity of O(n) and a linear spatial complexity of O(n), it is suitable for various applications where data transmission over networks is required (Algorithm 1).

The multitude of data generated by IoT devices has increased the complexity of data management. A previous study [33] presented the need to efficiently reduce the amount of IoT data efficiently while maintaining data integrity via the Run-Length Encoding algorithm. The algorithm is used locally and is applied for secure and energy-efficient data processing in IoT edge networks [34-36].

Algorithm 1. Algorithm Run-Length Encoding

```
public static class RLE
{
    public static string Encode(string input)
    {
        if (string.IsNullOrEmpty(input))
        {
            return string.Empty;
        }
        StringBuilder sb = new StringBuilder();
        int count = 1;
        char current = input[0];
        for (int i = 1; i < input.Length; i++)
        {
            if (input[i] == current)
            {
                count++;
            }
            else
            {
                sb.Append(count.ToString() + current);
                count = 1;
                current = input[i];
            }
        }
        sb.Append(count.ToString() + current);
        return sb.ToString();  }
```

Proposed System

In this section, a data normalisation and aggregation technique, which consists of data collection, data preprocessing, data normalisation and data aggregation steps, is proposed.

Implementing data integration in the IoT is a complex process, as shown in Figure 4.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031   ISSN (E): 2707-904X
VOLUME 19, SEPTEMBER 2024

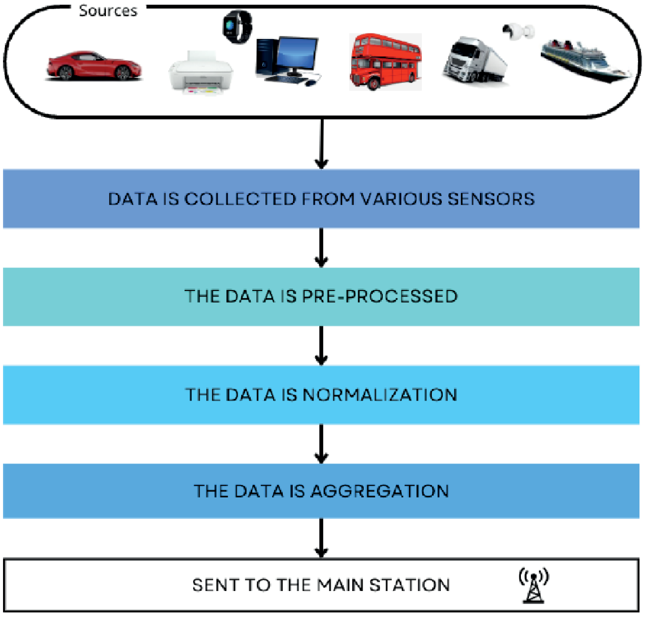

Sources
DATA IS COLLECTED FROM VARIOUS SENSORS
THE DATA IS PRE-PROCESSED
THE DATA IS NORMALIZATION
THE DATA IS AGGREGATION
SENT TO THE MAIN STATION

Figure 4. Data aggregation and normalisation process

Initially, data is collected from multiple IoT devices. The data may vary in type and source. However, its raw and unstructured nature makes it difficult to directly select the most important information, so it is then moved to a preprocessing stage. In the next stage, the data are cleaned and standardised. The main purpose of this stage is to prepare the data for statistical analysis, i.e., to eliminate inconsistencies and recover missing values. Next, the preprocessed data are compiled into a form that is easy to read and integrate. This process involves transforming the data through calculations, sorting or grouping to produce an output that can provide useful information. Finally, the collected data are analysed. Here, the data are scrutinised to identify trends, patterns and correlations. Using data analysis tools, valuable information is extracted from the collected data, which can then be used to make decisions.

During the data aggregation stage, fault tolerance is an important factor. The aggregation of network traffic when IoT devices communicate with each other can help solve the fault tolerance problem and summarise the capacity of the data links involved in the aggregation. In fact, most IoT network traffic flows are aggregated and come from various sensors and transducers.

If the target process is represented as a sequence $S = (S_1, S_2, S_3, \dots S_k)$, N is the number of elements of the sequence, and φ^2 is the variance, where $k = 1, 2, 3, \dots$. The autocorrelation function $F(l)$ is calculated according to formula (1).

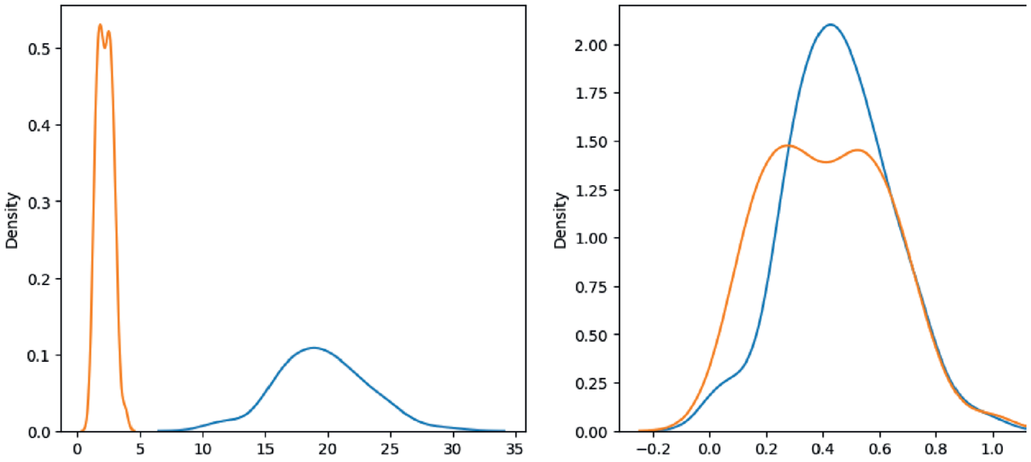
$$F(l) = \frac{\sum_{i=1}^{N-l} 0(S_i - \underline{S})(S_{i+l} - \underline{S})}{(N-l)\phi^2} \tag{1}$$

The aggregated flow is specified via sequential blocks whose elements are obtained from the original flow by averaging.

Results

The experimental study resulted in normalised and aggregated data, which were obtained via Z-score normalisation and the LEACH algorithm. The initial data were generated by sensors and nodes that were randomly located in a 500 by 500 metre network area. The number of nodes in the network ranges from 20-100, with 5 clusters and the packet sizes are 1024 bytes.

Z-score normalisation transforms each column of the dataset so that its mean is zero and its standard deviation is one. The results of the data normalisation process are shown in Figure 5. In the first graph, the mean value for the selected column is 0.25 and the standard deviation is 0.15. The second graph shows that the values of the column after normalisation range from -0.2 to 1.2.



a) before normalisation

b) after normalisation

Figure 5. Data normalisation process

Z-score normalisation helps to improve convergence by preparing the data for further processing using machine learning algorithms. At the same time, data normalisation provides an opportunity to compare different features per model. However, in addition to these advantages, there are also disadvantages such as loss of information and possible distortion of the mean and standard deviation. The application depends on the task at hand and the original data.

The LEACH algorithm was used for data aggregation to maintain balanced load balancing among all nodes in the IoT network. The protocol has a hierarchical structure with three types of participants: receivers, cluster head nodes and cluster nodes. The cluster nodes have two main functions: data collection and data transmission. After the cluster head node receives information from all the nodes in its cluster, it goes into sleep mode to save energy. Upon waking, the cluster head node aggregates the received information and sends the aggregated data to the sink node. The percentage of data aggregation decreases in the presence of faulty nodes. In the experimental study, the percentage of faulty nodes varies from 0 to 35%. Figure 6 shows the percentage of data aggregation where the abscissa axis shows the number of sensors. The investigated LEACH algorithm has an average value of 0.82.

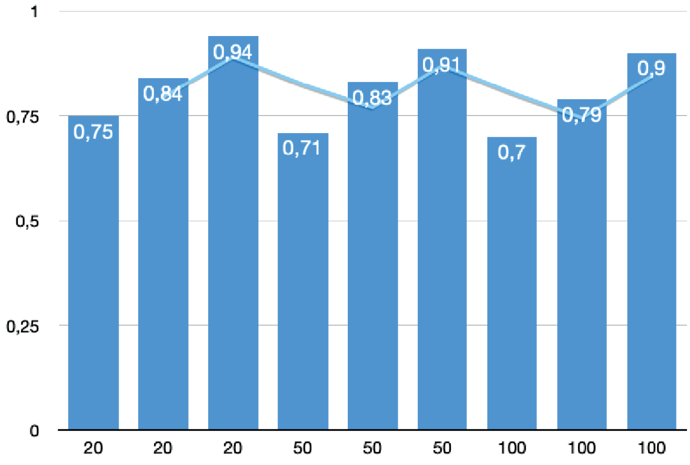


Figure 6. Data normalisation process

The proposed data normalisation and aggregation technique enables efficient load balancing among the nodes of the IoT network while providing a hierarchical control structure. Table 3 shows the comparison results of the proposed methodology with other research results.

Table 3. Comparison of the obtained results with those of other studies

Paper	Methods	Result
[37]	F-LEACH	improvement by 5-20%
[38]	Data aggregation and data driven discretisation method	improvement by 16,5%
[39]	Data gathering and aggregation with selective transmission, prefix frequency filtering	the solutions performance indicator 0,73: 0,69
Proposed System	Z-Score, LEACH algorithm	the solutions performance indicator 0,82

The table presents the results of the performance comparisons of different IoT data normalisation and aggregation methods. The results show that using F-LEACH can improve the performance by 5-20% [37], whereas the methods used in [38] can improve the performance by 16.5%. The solution obtained in [39] demonstrates the data collection and aggregation performance using the selective transmission method by 0.73 and the prefix frequency filtering by 0.69. The proposed model, which combines Z-Score normalisation and the LEACH algorithm, achieves the best performance, with a solution performance score of 0.82.

Discussion

Data normalisation and aggregation enable real-time tracking of IoT devices, which in turn facilitates rapid responses to attacks, anomalies, problems or emerging opportunities. One of the greatest challenges in IoT data normalisation and aggregation is ensuring data integrity. The sheer volume and variety of data from multiple IoT devices make it susceptible to inconsistencies, inaccuracies, and duplication. Protecting the accuracy and reliability of data is a significant hurdle. Another major challenge is the privacy concerns associated with data normalisation and aggregation. In turn, data aggregation can provide anonymisation; it must balance the collection of sufficient data while ensuring the privacy of individuals.

The experimental results showed that the proposed method based on Z-score and LEACH algorithms can effectively reduce the amount of transmitted data while maintaining high accuracy. The Z-score algorithm successfully detected and removed outliers, which improved the data quality for subsequent aggregation. The LEACH algorithm ensured uniform load balancing among the network nodes and minimised the energy cost. Comparison with traditional aggregation methods showed that our approach provided higher accuracy and scalability. However, it should be noted that the effectiveness of the method may decrease when the network is highly dynamic or in the presence of complex data patterns. Adaptive aggregation methods that can automatically adjust to changing network conditions can be considered as a future research direction.

Conclusion

The presented research is aimed at solving the actual problem of data security in IoT networks. The authors proposed and verified a methodology of data normalisation and aggregation in the interaction of the IoT, which contributes to reducing the volume of transmitted data, increasing the security and efficiency of the system as a whole. The proposed data processing methodology for IoT systems includes the stages of data collection, preprocessing, normalisation and aggregation. For normalisation, a Z-score algorithm was used which successfully detected and handled outliers in the data. Data aggregation was performed using the LEACH algorithm, which reduced the load on the network and optimised the power consumption. To validate the proposed methodology, sensors and nodes of 20, 50 and 100 sensors and nodes arranged in random order have been considered. The data normalisation and aggregation technique can successfully enable the network to minimise data transmission and energy consumption when the data aggregation process is performed in a secure manner. Therefore, it is very important to ensure that the network is secure when implementing the data normalisation and aggregation process so that it is possible to retrieve the original information from the data owner within a short period of time. On this basis, this paper presents the research results of scholars who have studied and proposed various methods to secure the data aggregation process in their own way. The presented methodology for aggregation and normalisation of data received from IoT devices is necessary to highlight their operational mechanism, advantages and limitations. A similar comparative study of the approaches used in the process of data normalisation and aggregation was presented, which addresses the issues of interference, fault tolerance, security and mobility. While discussing the methods of data normalisation and aggregation, a study of their applicability to IoT networks was presented. Aggregating data from different IoT devices allows us to identify patterns, trends and correlations. The results of the study can be useful for IoT system developers and cybersecurity researchers. Future work will explore different scenarios via methods that consider limited client resources, heterogeneous client data, server capacity and high communication costs.

Acknowledgement

This research was funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan, Grant No. AP14973006 “Investigation of methods of maintaining confidentiality in the process of interaction of the Internet of Things”.

References

[1] Report by J’son & Partners Consulting. (2019). website: https://json.tv

[2] Statista. (2022). website: https://statista.com

[3] A. Schlemitz and V. Mezhuyev, “Approaches for data collection and process standardization in smart manufacturing: Systematic literature review,” Journal of Industrial Information Integration, vol. 38, p. 100578, Mar. (2024), https://doi.org/10.1016/j.jii.2024.100578.

[4] A. Adamova, T. Zhukabayeva, and Y. Mardenov, “Machine Learning in Action: An Analysis of its Application for Fault Detection in Wireless Sensor Networks,” 2023 IEEE International Conference on Smart Information Systems and Technologies (SIST), May 2023, https://doi.org/10.1109/sist58284.2023.10223548.

[5] S. Abbasian Dehkordi, K. Farajzadeh, J. Rezazadeh, R. Farahbakhsh, K. Sandrasegaran, and M. Abbasian Dehkordi, “A survey on data aggregation techniques in IoT sensor networks,” Wireless Networks, vol. 26, no. 2, pp. 1243–1263, Sep. 2019, https://doi.org/10.1007/s11276-019-02142-z.

[6] K. Gulati, R. S. Kumar Boddu, D. Kapila, S. L. Bangare, N. Chandnani, and G. Saravanan, “A review paper on wireless sensor network techniques in Internet of Things (IoT),” Materials Today: Proceedings, vol. 51, pp. 161–165, 2022, https://doi.org/10.1016/j.matpr.2021.05.067.

[7] X. Liu et al., “Secure Data Aggregation Aided by Privacy Preserving in Internet of Things,” Wireless Communications and Mobile Computing, vol. 2022, pp. 1–14, Mar. 2022, https://doi.org/10.1155/2022/4858722.

[8] K. Haseeb, N. Islam, T. Saba, A. Rehman, and Z. Mehmood, “LSDAR: A light-weight structure based data aggregation routing protocol with secure internet of things integrated next-generation sensor networks,” Sustainable Cities and Society, vol. 54, p. 101995, Mar. 2020, https://doi.org/10.1016/j.scs.2019.101995.

[9] X. Wang, S. Garg, H. Lin, G. Kaddoum, J. Hu, and M.S. Hossain, "A Secure Data Aggregation Strategy in Edge Computing and Blockchain-Empowered Internet of Things," IEEE Internet of Things Journal, vol. 9, no. 16, pp. 14237–14246, Aug. 2022, https://doi.org/10.1109/jiot.2020.3023588.

[10] C. Chakraborty, S. B. Othman, F. A. Almalki, and H. Sakli, "FC-SEEDA: fog computing-based secure and energy efficient data aggregation scheme for Internet of healthcare Things," Neural Computing and Applications, vol. 36, no. 1, pp. 241–257, Jan. 2023, https://doi.org/10.1007/s00521-023-08270-0.

[11] L. Bolognini, S. Ziegler, P. Annicchino, F. Capparelli, and A. Audino, "9. Data Protection Compliance Requirements for the Internet of Things," Security Risk Management for the Internet of Things: Technologies and Techniques for IoT Security, Privacy and Data Protection, 2020, https://doi.org/10.1561/9781680836837.ch9.

[12] M. Narimani Zaman Abadi, A. Jalaly Bidgoly, Y. Farjami, and E. Hossein Khani, "A comprehensive soft security model for Cognitive Internet of Things," Internet of Things, vol. 23, p. 100858, Oct. 2023, https://doi.org/10.1016/j.iot.2023.100858.

[13] S.N.G. Aryavalli and H. Kumar, "Top 12 layer-wise security challenges and a secure architectural solution for Internet of Things," Computers and Electrical Engineering, vol. 105, p. 108487, Jan. 2023, https://doi.org/10.1016/j.compeleceng.2022.108487.

[14] A.A. Abba Ari et al., "Enabling privacy and security in Cloud of Things: Architecture, applications, security &amp; privacy challenges," Applied Computing and Informatics, vol. 20, no. 1/2, pp. 119–141, Jul. 2020, https://doi.org/10.1016/j.aci.2019.11.005.

[15] X. Larriva-Novo, V.A. Villagrá, M. Vega-Barbas, D. Rivera, and M. Sanz Rodrigo, "An IoT-Focused Intrusion Detection System Approach Based on Preprocessing Characterization for Cybersecurity Datasets," Sensors, vol. 21, no. 2, p. 656, Jan. 2021, https://doi.org/10.3390/s21020656.

[16] F.T. Lima and V.M.A. Souza, "A Large Comparison of Normalization Methods on Time Series," Big Data Research, vol. 34, p. 100407, Nov. 2023, https://doi.org/10.1016/j.bdr.2023.100407.

[17] A. Alabrah, "An Improved CCF Detector to Handle the Problem of Class Imbalance with Outlier Normalization Using IQR Method," Sensors, vol. 23, no. 9, p. 4406, Apr. 2023, https://doi.org/10.3390/s23094406.

[18] M. Shantal, Z. Othman, and A.A. Bakar, "A Novel Approach for Data Feature Weighting Using Correlation Coefficients and Min–Max Normalization," Symmetry, vol. 15, no. 12, p. 2185, Dec. 2023, https://doi.org/10.3390/sym15122185.

[19] M. Pagan, M. Zarlis, and A. Candra, "Investigating the impact of data scaling on the k-nearest neighbour algorithm," Computer Science and Information Technologies, vol. 4, no. 2, pp. 135–142, Jul. 2023, https://doi.org/10.11591/csit.v4i2.p135-142.

[20] K. Cabello-Solorzano, I. Ortigosa de Araujo, M. Peña, L. Correia, and A. J. Tallón-Ballesteros, "The Impact of Data Normalization on the Accuracy of Machine Learning Algorithms: A Comparative Analysis," 18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023), pp. 344–353, 2023, https://doi.org/10.1007/978-3-031-42536-3_33.

[21] S. Aparna and V. P Raghu, "Improving Anomaly Classification using Combined Data Transformation and Machine Learning Methods," International Journal of Performability Engineering, vol. 20, no. 2, p. 68, 2024, https://doi.org/10.23940/ijpe.24.02.p2.6880.

[22] Y. Wang, K. Yang, W. Wan, Y. Zhang, and Q. Liu, "Energy-Efficient Data and Energy Integrated Management Strategy for IoT Devices Based on RF Energy Harvesting," IEEE Internet of Things Journal, vol. 8, no. 17, pp. 13640–13651, Sep. 2021, https://doi.org/10.1109/jiot.2021.3068040.

[23] M. K. Dahouda and I. Joe, "A Deep-Learned Embedding Technique for Categorical Features Encoding," IEEE Access, vol. 9, pp. 114381–114391, 2021, https://doi.org/10.1109/access.2021.3104357.

[24] S. Pramanik, "An Effective Secured Privacy-Protecting Data Aggregation Method in IoT," Achieving Full Realization and Mitigating the Challenges of the Internet of Things, pp. 186–217, Mar. 2022, https://doi.org/10.4018/978-1-7998-9312-7.ch008.

[25] Z. Wang, J. Tao, and D. Zou, "Privacy-Preserving Data Aggregation in IoTs: A Randomize-then-Shuffle Paradigm," 2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring), Jun. 2023, https://doi.org/10.1109/vtc2023-spring57618.2023.10199427.

[26] K. Jastaniah, N. Zhang, and M.A. Mustafa, "Efficient User-Centric Privacy-Friendly and Flexible Wearable Data Aggregation and Sharing," IEEE Transactions on Cloud Computing, pp. 1–18, 2024, https://doi.org/10.1109/tcc.2024.3375801.

[27] Q. Cao, F. Xu, H. Xu, and Y. Jin, "A Fog Based Privacy Preserving Data Aggregation Method for Vehicular Internet of Things," 2024 7th International Conference on Communication Engineering and Technology (ICCET), Feb. 2024, https://doi.org/10.1109/iccet62255.2024.00010.

[28] M.A. Mughal, A. Ullah, X. Yu, W. He, N. Z. Jhanjhi, and S. K. Ray, "A secure and privacy preserved data aggregation scheme in IoMT," Heliyon, vol. 10, no. 7, p. e27177, Apr. 2024, https://doi.org/10.1016/j.heliyon.2024.e27177.

[29] J. Zhang and J. Wei, "PFDAM: Privacy-Preserving Fine-Grained Data Aggregation Scheme Supporting Multifunctionality in Smart Grid," IEEE Internet of Things Journal, vol. 11, no. 15, pp. 25520–25533, Aug. 2024, https://doi.org/10.1109/jiot.2024.3356593.

[30] M.A.P. Putra, R.N. Alief, S.M. Rachmawati, G.A. Sampedro, D.-S. Kim, and J.-M. Lee, "Proof-of-authority-based secure and efficient aggregation with differential privacy for federated learning in industrial IoT," Internet of Things, vol. 25, p. 101107, Apr. 2024, https://doi.org/10.1016/j.iot.2024.101107.

[31] B.A. Begum and S.V. Nandury, "Data aggregation protocols for WSN and IoT applications – A comprehensive survey," Journal of King Saud University - Computer and Information Sciences, vol. 35, no. 2, pp. 651–681, Feb. 2023, https://doi.org/10.1016/j.jksuci.2023.01.008.

[32] N.A. Khairi and A. Bahari Jambek, "Study on data compression algorithm and its implementation in portable electronic device for Internet of Things applications," EPJ Web of Conferences, vol. 162, p. 01073, 2017, https://doi.org/10.1051/epjconf/201716201073.

[33] Y. Idir, I. Moumen, J. Abouchabaka, and N. Rafalia, "Enhancing IoT Data Integrity and Effectiveness through hybrid Compression Method: A Step Towards Energy Efficiency," E3S Web of Conferences, vol. 477, p. 00042, 2024, https://doi.org/10.1051/e3sconf/202447700042.

[34] S. Patidar, R. Jindal, and N. Kumar, "A secure and energy-efficient edge computing improved SZ 2.1 hybrid algorithm for handling iot data stream," Multimedia Tools and Applications, Mar. 2024, https://doi.org/10.1007/s11042-024-18765-0.

[35] E. al. Madhu M Nashipudmath, "Smart Data Management in IoT: Leveraging Wireless Sensor Networks for Efficient Information Processing," Journal of Electrical Systems, vol. 19, no. 2, pp. 01–08, Jan. 2024, https://doi.org/10.52783/jes.669.

[36] A. K. Idrees and L. W. Jawad, "Energy-efficient Data Processing Protocol in edge-based IoT networks," Annals of Telecommunications, vol. 78, no. 5–6, pp. 347–362, Mar. 2023, https://doi.org/10.1007/s12243-023-00957-8.

[37] S. Sajedi, M. Maadani, M. Nesari Moghadam, "F-LEACH: a fuzzy-based data aggregation scheme for healthcare IoT systems," The Journal of Supercomputing, vol. 78(1), pp. 1030-1047, 2022, https://doi.org/10.1007/s11227-021-03890-6.

[38] Z. Guo, A.R. Coffman, J. Munk, P. Im, T. Kuruganti, and P Barooah, "Aggregation and data driven identification of building thermal dynamic model and unmeasured disturbance," Energy and Buildings, vol. 231, p.110500, 2021, https://doi.org/10.1016/j.enbuild.2020.110500.

[39] A.K.M. Al-Qurabat, and A. Kadhum Idrees, "Data gathering and aggregation with selective transmission technique to optimize the lifetime of Internet of Things networks," International Journal of Communication Systems, vol. 33(11), p.e4408, 2020, https://doi.org/10.1002/dac.4408.