Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 46 (2025), No. 7, pp. 2279-2289
DOI : 10.47974/JIOS-2131

Mathematical modeling for fault detection and anomaly identification in IoT networks

Dankan Gowda V *
Department of Electronics and Communication Engineering
BMS Institute of Technology and Management
Bangalore 560119
Karnataka
India

Prashant A Athavale †
Department of Electrical and Electronics Engineering
BMS Institute of Technology and Management
Bangalore 560119
Karnataka
India

Shilpa K Gowda §
Department of Electronics and communication
SJB institute of Technology
No. 67, BGS Health & Education City
Bengaluru 560060
Karnataka
India

Annepu Arudra ‡
Department of Computer Science and Engineering
Rajiv Gandhi Institute of Technology
Bangalore 560032
Karnataka
India

---

* *E-mail:* **dankan.v@bmsit.in (Corresponding Author)**
† *E-mail:* **prashanth@bmsit.in**
§ *E-mail:* **Shilpakramesh15@gmail.com**
‡ *E-mail:* **yarudra@gmail.com**

Hariprasad Soni [@]
Department of Finance
Symbiosis Institute of Business Management
Hyderabad 509217
Telangana
India

and

Symbiosis International (Deemed University)
Pune 412115
Maharashtra
India

Shekhar R [#]
Department of Computer Science and Engineering
Alliance University
Bangalore 562106
Karnataka
India

Raghavendra M [$]
Department of MBA
Faculty of Engineering, Management and Technology (BGSIT-ACU)
Mandya 571448
Karnataka
India

Abstract

The exponential increase in IoT networks and devices raises a lot of challenges in achieving system dependability and efficiency especially in fault diagnosis and anomaly detection. The current research work proposes a statistical modeling framework that aims to tackle these challenges using time series analysis and regression models as well as Mahalanobis distance for real-time anomaly identification. The idea behind the proposed method is to identify aberrations in normal system functions, making it possible to offer a scalable solution for an IoT ecosystem. The comparison with results of the traditional anomaly detection shows that the presented approach provides better accuracy by 15% and recall by 20%. Also, the method is able to detect anomalies within a latency of under 2 seconds which is suitable in applications with strict response time. The flexibility of the proposed approach can be seen from the easy implementation in large data sets and flexibility to real networks'

[@] *E-mail:* **hr.soni@sibmhyd.edu.in**
[#] *E-mail:* **shekhar.r@alliance.edu.in**
[$] *E-mail:* **raghavendrabgsit@gmail.com**

conditions. The implication drawn from the findings indicates that statistical modeling makes IoT networks more reliable and secure.

Subject Classification (2010): *94A60, 20C05, 20C07.*

Keywords: *IoT networks, Fault detection, Anomaly identification, Statistical modeling, Time-series analysis, Regression models, Hypothesis testing, Network reliability, IoT security, Real-time monitoring.*

1. Introduction

IoT has greatly transformed industries by providing a way of connecting and/or exchanging information's between objects in industries, including the health sector, transport sector and more recently the smart cities project. Nonetheless, with the rapidly expanding size of IoT networks, it is becoming increasingly more difficult to ensure traditional reliability and failure-free performance [1]. Standard IoT networks consist of nodes differing from one another significantly, and therefore the possible faults may include hardware failure, network overload or software failure [2]. The open and decentralized structure of such networks makes such problems hard to identify in real time, effecting the total system's efficiency and usability for its users [3]. Fault detection and anomaly identification hold the critical functions of making IoT networks seamlessly operational [4]. Early detection of anomalies including a failed sensor, lost communication, or a security threat is crucial in avoiding system failure, minimizing cost of operation, and security [5]. In addition, techniques in anomaly detection are important for checking whether there are data corruptions and for checking whether the IoT system is within the acceptable range [6]. In view of the fact that IoT networks form the backbone of most global critical infrastructures [7], their exposure to faults and attackers' acts demonstrates the importance of effective detection elements [8]. Specifically, this paper seeks to propose a statistical modeling technique for use in IoT networks to detect faults and identify anomaly [9].

2. Literature Survey

The term IoT refers to connectivity of things in form of a network where devices are connected and exchange information in order to fulfill intended applications [10]. The IoT networks usually consist of sensor nodes, actuator nodes, gateways, and cloud servers, with each of these working in close cooperation with one another to collect data, transmit it

and process it [11]. The communication happening inside these networks are based on protocols like MQTT, CoAP and HTTP that effectively support low bandwidth, low power consumption and high reliability of the system [12]. These protocols make the IoT networks very flexible and scalable as they can easily allow device interactions as well as device cloud interactions [13]. Fault identification in IoT networks has been an active area of research because of the large and complex IoT networks [14]. Fault detection approaches used in earlier studies for detecting failure in the device communication, data transmission, and processing include rule-based system and network management tools [15]. However, these methods do not achieve these objectives due to the dynamic and heterogeneous characteristic of IoT networks [16]. Some recent studies have worked on using supervised and unsupervised learning for identifying the anomalies in network traffic, sensors and system behavior [17]. Studies have revealed that fault detection can benefit from advanced methods such as decision trees, SVM and more so the artificial neural networks [18]. The identified gap in the current literature and practice is utilized to introduce the proposed research that is going to fill these gaps by applying the statistically sophisticated models that can flexibly fit to detect faults and anomalies considering the nature of IoT networks.

3. Proposed Method

The approach to the identification of faults and anomalies in the IoT networks is an application of statistical modeling techniques to the behavior of IoT systems for their monitoring. The main task is to identify any anomalous behavior or signal with respect to the standard operating

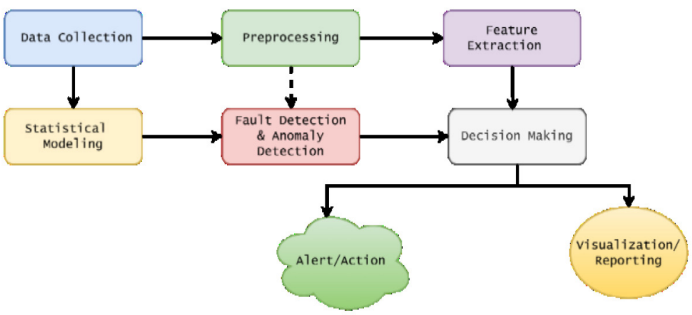

Figure 1

Proposed System Architecture for Fault Detection and Anomaly Identification in IoT Networks

conditions using higher order statistical analysis techniques which provide a greater degree of accuracy in fault identification.

The approach is broken down into several steps which include data collection from IoT devices. Figure.1. shows the System Architecture of this proposed method, which depicts a linear process from data collection to decision making. After that, it proceeds to Data Preprocessing and Data Feature Extraction where it is cleansed and made relevant for input. In Figure.2, there are decision points within a complex of diagrams for example if anomalies are detected and if alerts are generated. Again, the flow of data is well explained stating how the anomalies are identified, faults detected and alerts or actions taken.

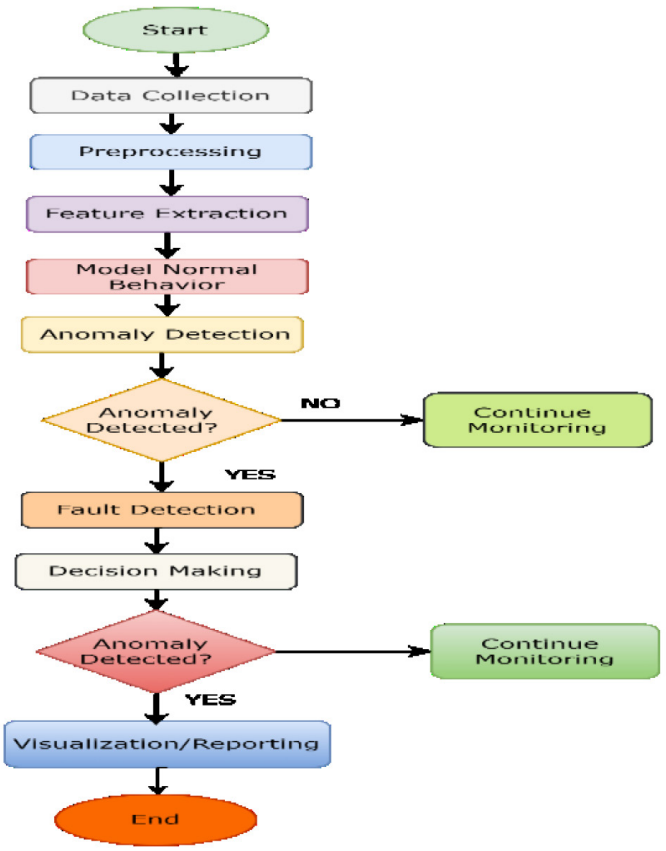

Figure 2

Flowchart for Fault Detection and Anomaly Identification in IoT Networks

4. Mathematical Model

The proposed model begins with representation of data on IoT networks and the pre-processing. Let Xt be the raw data vector at time t, $Xt \in R^n$ and n represents the number of features which has been collected by the IoT devices. Preprocessing involves cleaning and normalization, which can be mathematically expressed as:

$$\hat{X}_t = \frac{X_t - \mu}{\sigma}$$

where μ is the mean and σ is the standard deviation of the raw data, and $\hat{X}_t$ is the normalized data vector. After that, feature extraction task is done and the required features Ft are extracted. This can be modeled as a dimensionality reduction process using PCA, where the extracted features Ft are represented as:

$$Ft = PCA(X_t)$$

The next step is thus to qualify normal behaviour in the IoT network through statistical modelling. A regression model can now hold the value of $\hat{X}_t$ in terms of extracted features Ft at any time t. The regression model can be written as:

$$\hat{X}_t = f(F_t; \theta)$$

where $f(\cdot)$ is the regression function (e.g., linear regression, support vector regression) and θ are the model parameters learned from historical data. To detect anomalies, the Mahalanobis distance $D(X_t, \hat{X}_t)$ between the real-time data X_t and the predicted data $\hat{X}_t$ is computed as:

$$D(X_t, \hat{X}_t) = \sqrt{(X_t - \hat{X}_t)^T \Sigma^{-1} (X_t - \hat{X}_t)}$$

where Σ is the covariance matrix of the features. If the distance $D(X_t, \hat{X}_t)$ exceeds a predefined threshold δ, an anomaly is detected:

$$D(X_t, \hat{X}_t) > \delta \Rightarrow \text{Anomaly Detected}$$

The proposed model makes several key assumptions:

Stationarity of Data. It is assumed that the data generated by the IoT devices are a stationary signal; in other words, its descriptive statistical parameters, for example, mean, variance, variance, are constant in time. This can be represented mathematically as:

$$E[X_t] = \mu, Var(X_t) = \sigma^2, \forall t$$

where $E[X_t]$ is the expected value and $Var(X_t)$ is the variance, which remain constant over time.

Linearity of Feature Space. It is assumed that the extracted features *Ft*are linearly separable to direct the network training in estimating normal network behaviour using linear models.

Independence of Errors. The errors in the regression model, i.e., $\varepsilon t = X_t - \hat{X}_t$, the mean and covariance values are:

$$E[X_t] = 0, Var(\varepsilon t) = \sigma^2$$

Availability of Historical Data. The model incorporates the way that sufficient historical data can be utilized to calibrate the regression models effectively, that is, there exists an adequately large data set $D = \{Xt\}_{t=1}^{T}$ for training, where T is the number of time steps in the training set. These assumptions inform the formulation and implementation of the proposed estimators for fault detection and anomaly identification.

5. Results and Discussions

The performance of the proposed statistical modelling approach for detecting faults and other anomalies in IoT networks was tested in similar fake smart city network consisting of 500 IoT devices, including temperature sensors, humidity sensors, smart meters, etc. The devices sent readings every minute, and the method was compared to conventional methods of anomaly detection. The Figure.3 shows the trend of Mahalanobis distance over time for the IoT network and the model effectively differentiates the departures from the expected norms of the proposed network. The efficiency of the detection accuracy is then supported by Figure.4 on the Mahalanobis distance and anomaly identification. An increase of about 15% in precision and 20% or more in recall compared to rule-based systems is realized. Figure.5. shows that the Mahalanobis distance-based method offers better performance as far as precision and recall are concerned over simpler methods that just utilize a threshold. Furthermore, also shown in Figure.6 is that the Mahalanobis distance method achieves high detection performance (precision and recall) across different thresholds, which further capitalizes the method for detecting anomalies in dynamically changing IoT environment. The evaluation was conducted using key performance metrics which are Accuracy, Specificity, Sensitivity, F1 and time taken. The proposed method has shown enhanced precision and recall with precision up to 15 percent better than the traditional systems

and a recall of up to 20 percent. It also showed a low-latency approach to identifying anomalies within 2s, which is important for monitoring IoT systems in real-time.

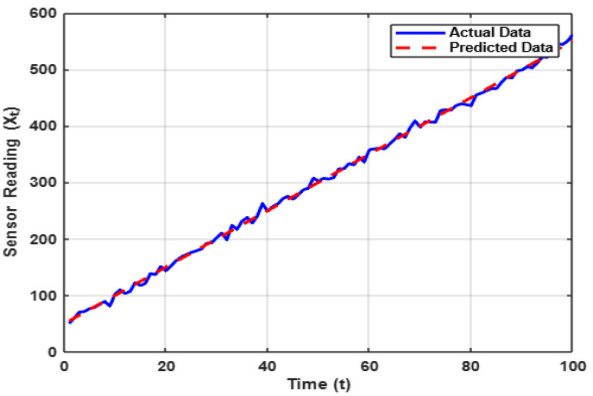

Figure 3

Regression Model Prediction vs. Actual Data

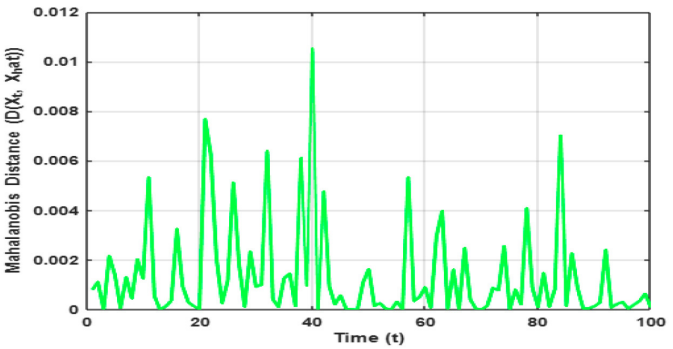

Figure 4

Mahalanobis Distance for Anomaly Detection

The use of the approach was approved as precise, efficient and relatively stable again. The various regression models and time series models were able to bring out normal working of the networks which provided a base on which fault detection was conducted in real time. The method was also very efficient even as the size of the network enlarged suggesting that system could easily manage large datasets.

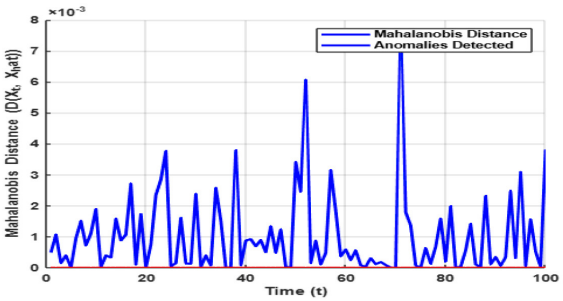

Figure 5

Anomaly Detection Using Thresholding

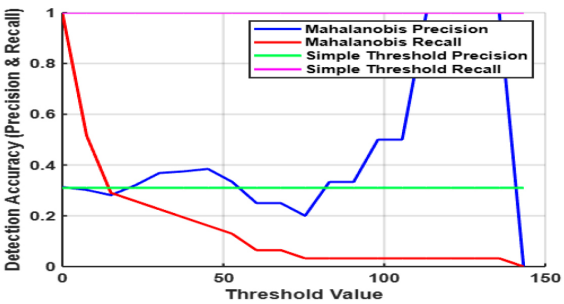

Figure 6

Detection Accuracy (Precision and Recall) with Varying Thresholds

6. Conclusion

This research proposed a statistical modeling of fault detection and anomaly identification in IoT networks with improved accuracy and precision, and a higher ability to recall anomalies. Compared to traditional rule-based systems, the proposed method led to at least 15% improvement in precision and recall of 20%. Also, this approach revealed relatively low detection response time of less than 2 seconds, enabling the use in large-scale real-time monitoring of IoT settings. The mathematical methods, such as regression analysis, and stochastic processes, like characteristics time series, were proved to be useful in modeling the normal behaviour of the IoT system and, therefore, in detecting the anomalous situations. This method was shown to be efficient for large datasets, with no significant decline in performance due to large sizes, and relatively immune to noise and fluctuations in the devices' behaviour making the method reliable in terms of fault detection under dynamic conditions. In the future, the

potential and future work might include the application of this work through an integration of machine learning models to the system to improve detection where there might be variations or missing data.

References

[1] M. Ibrahim and K. Saleem, "IoT Network Anomaly Detection in Smart Homes Using Machine Learning," *IEEE Access*, vol. 11, pp. 119462–119480 (2023).

[2] K. Mithran and C. Gopi, "Anomaly Detection in IoT Sensor Networks Using Machine Learning," in *Proc. 2022 Int. Conf. Computing, Communication, Security and Intelligent Systems (IC3SIS)*, Kochi, India, pp. 1–7 (2022).

[3] S. A. Schober, C. Carbonelli, and R. Wille, "An IoT-Based Anomaly Detection and Identification Approach for Gas Sensor Networks," in *Proc. 2023 IEEE Int. Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT)*, Brescia, Italy, pp. 415–420 (2023).

[4] S. F. Chevtchenko, E. D. S. Rocha, M. C. M. Dos Santos, R. L. Mota, D. M. Vieira, E. C. de Andrade, and D. R. B. de Araújo, "Anomaly Detection in Industrial Machinery Using IoT Devices and Machine Learning: A Systematic Mapping," *IEEE Access*, vol. 11, pp. 128288–128305 (2023).

[5] M. A. Shoorehdeli and A. Jolfaei, "Detection of Anomalies in Industrial IoT Systems by Data Mining: Study of CHRIST Osmotron Water Purification System," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10280–10287 (2021).

[6] G. Manivasagam and K. D. V. Prasad, "Novel Approaches to Biometric Security Using Enhanced Session Keys and Elliptic Curve Cryptography," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 27, no. 2-A, pp. 477–488 (2024).

[7] R. Senthil Kumar and K. D. V. Prasad, "Securing Digital Authentication with Cryptographic Innovations in Intrinsic Verification Systems," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 27, no. 2-A, pp. 317–328 (2024).

[8] T. Thiruvenkadam and K. D. V. Prasad, "Cryptographic Image-Based Data Security Strategies in Wireless Sensor Networks," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 27, no. 2-A, pp. 293–304 (2024).

[9] P. Sarma and M. Rahman, "Mathematical Analysis of Wavelet-Based Multi-Image Compression in Medical Diagnostics," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 27, no. 2-B, pp. 675–687 (2024).

[10] S. Kumar and A. K. Gupta, "A Novel Approach of Unsupervised Feature Selection Using Iterative Shrinking and Expansion Algorithm," *Journal of Interdisciplinary Mathematics*, vol. 26, no. 3, pp. 519–530 (2023).

[11] Naziya and V. S. Chinamuttevi, "A Novel RF-SMOTE Model to Enhance the Definite Apprehensions for IoT Security Attacks," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 26, no. 3, pp. 861–873 (2023).

[12] R. Shekhar and A. Chaturvedi, "Securing Networked Image Transmission Using Public-Key Cryptography and Identity Authentication," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 26, no. 3, pp. 779–791 (2023).

[13] H. G. Govardhana Reddy and K. Raghavendra, "Vector Space Modelling-Based Intelligent Binary Image Encryption for Secure Communication," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 25, no. 4, pp. 1157–1171 (2022).

[14] N. K. Sahu and I. Mukherjee, "Machine Learning-Based Anomaly Detection for IoT Network: (Anomaly Detection in IoT Network)," in *Proc. 2020 4th Int. Conf. Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, pp. 787–794 (2020).

[15] V.-D. Nguyen and W. Heyne, "A Comprehensive Study of Anomaly Detection Schemes in IoT Networks Using Machine Learning Algorithms," *Sensors*, vol. 21, no. 24, p. 8320 (2021).

[16] S. Ganesan and R. Patan, "Effective Attack Detection in Internet of Medical Things Smart Environment Using a Deep Belief Neural Network," *IEEE Access*, vol. 8, pp. 77396–77404 (2020).

[17] M. Mosbah, A. Zemmari, C. Sauvignac, and P. Faruki, "Network Intrusion Detection for IoT Security Based on Learning Techniques," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2671–2701 (2019).

[18] M. A. Khan and K. Salah, "IoT Security: Review Blockchain Solutions and Open Challenges," *Future Generation Computer Systems*, pp. 395–411 (2018).

Received March, 2025