RADINKA JOURNAL OF SCIENCE AND SYSTEMATIC LITERATURE REVIEW

ISSN 2988-5825

https://www.rjupublisher.com/ojs/index.php/RJSLR

Detection of Anomalies in Internet of Things (IoT) Devices and Sensors

Solomon Makasda DICKSON*

School of Sciences, Department of Computer Science, National Institute for Nigerian Languages, Aba, Abia State, Nigeria

*Corresponding author: dicksonsolomon1988@gmail.com

Abstract

Though IoT devices and sensors are intended to make our lives easier, they can be vulnerable to malicious activities. As such, a critical part of IoT security is the development of algorithms that can detect anomalies in these devices and sensors. This study aimed to review the various approaches to anomaly detection in the context of IoT devices and sensors. The objectives are to evaluate unsupervised, supervised, reinforcement, and deep learning methods. The research methodology adopted for this study was explanatory. This study focuses on the detection of anomalies in IoT devices and sensors. This involves literature research using sources like web sources, in-print material, journals, textbooks, and articles. In conclusion, this study evaluated the various approaches to anomaly detection in the context of IoT devices and sensors. Methods such as unsupervised, supervised, reinforcement, and deep learning are critically reviewed.

Keywords: anomaly, device, sensor, and detection

Acknowledgments

Thanks to all parties who have supported the implementation of this research. I hope this research can be useful.

For citation:

Dickson, S. M. (2024). Detection of Anomalies in Internet of Things (IoT) Devices and Sensors. *Radinka Journal of Science and Systematic Literature Review*, 2(3), 475-481.

Introduction

The Internet of Things (IoT) is an ever-evolving network of connected devices that generate huge amounts of data, enabling human-driven automation of numerous processes. Ongoing study on anomaly detection for the Internet of Things (IoT) is a rapidly expanding field (M. S. Dickson & Amannah, 2023). This growth necessitates an examination of application trends and current gaps. The vast majority of those publications are in areas such as network and infrastructure security, sensor monitoring, smart home, and smart city applications and are extending into even more sectors (Okonta & Vukovic, 2024). The importance of anomaly

detection is that anomalies in IoT data, which occur sparsely, can yield crucial actionable information in various sectors, including unsupervised, supervised, reinforcement learning, and deep learning methods (Taye, 2023). Anomaly detection in IoT, for example, is employed in the betting and gambling sector to detect insider trading by analyzing trade activity patterns. On the other hand, industrial machines use a detection algorithm to ensure production safety. Anomaly detection according to (Erhan et al., 2021), is a much broader problem, going well beyond the sensors and devices that we scrutinize herein, and dating back many years in the research panorama. It pertains to a vast number of application domains, each one with its peculiarity and constraints. Prominent examples, beyond the general fields of data analysis and artificial intelligence, are cyber security, predictive maintenance, fault prevention, automation, and e-health, to mention but a few (S. M. Dickson & Ijeoma Peace Okechukwu, 2024). Anomaly detection has conventionally been tackled from the statistical viewpoint. The prominence of machine learning (ML) has, however, opened new possibilities for the detection of outliers, particularly thanks to the availability of vast amounts of data to train sophisticated learning models (Li et al., 2020). This is an attractive proposition, particularly in domains such as the IoT, whereby new data patterns make it difficult to use static models (Alexander et al., 2023). The IoT and sensor networks are used for multiple different applications, most of which revolve around monitoring of some kind, whether that be monitoring personal health devices, monitoring an entire city, monitoring the weather, monitoring an industrial system, etc.

theory, an anomaly is simple to comprehend, and a domain expert will spot anomalous data if given enough time (Chatterjee & Ahmed, 2022); (Krzysztoń et al., 2024). However, there are [8]several difficulties in developing an automated model in an IoT environment. Devices and sensors can be vulnerable to malicious activities. As such, there is a need to develop modalities for device and sensor anomaly detection to make our lives easier. This study reviewed the various approaches to anomaly detection in the context of IoT devices and sensors. Methods such as unsupervised learning, supervised learning, reinforcement learning deep learning, and the concept of anomaly detection are evaluated. The study aimed to evaluate anomalies in IoT devices and sensors (AIDS). The specific objectives included to review: approaches to anomaly detection in the context of IoT devices and sensors; detection of anomaly methods in IoT devices and sensors; concept of anomaly detection in IoT devices and sensors; The study was restricted to the approaches to anomaly detection and the handshake of IoT, shared with devices and sensors. It also unearths the methods and concept of an anomaly as is related to IoT devices and sensors value chain**.**

Methodology

The research methodology adopted for this study was explanatory. This study focuses on the detection of anomalies in IoT devices and sensors. This involves literature research using sources like web sources, in-print material, journals, textbooks, and articles.

Literature review

This section highlights with critical evaluation scientifically connected to the research problem. (Kraft-Todd & Rand, 2021), that science is characterized by reduction, repeatability, and retaliation. The successful application of the method of science in the physical systems is due to their closeness. The IoT systems theory holds that an organization is a social system made up of integrated parts. It also observed that the Internet of Things system refers to the integration of different components and relationships among them and their attributes such that they form a functionally related whole. The hierarchical and lateral structure in any system and their associated integrations are geared towards achieving the goal. It is known that any dislocations of the part of

a system will jeopardize the efficient working of the system. The study is anchored on system theory. The IoT is a system where many individuals play different roles in their levels to achieve one common aim which is the interconnection of sensors for information sharing.

IoT-Enabled Smart Agriculture: a Pathway to Nigeria's Economic Recovery and Sustainability. The rapid growth of the Internet of Things (IoT) and the introduction of Smart Agriculture present an opportunity for Nigeria to create meaningful economic growth and improved livelihoods. This study aimed to examine the potential of IoT-enabled Smart Agriculture to provide a pathway to economic recovery and sustainable development in Nigeria. The specific objectives are to review an overview of Nigeria's current agricultural sector and the challenges it faces, including security, privacy, cost, data management, energy, lack of access to knowledgeable farming techniques, and oversight of land use. The method adopted in this study was the incremental descriptive model (IDM). It then provides the potential solutions of IoT-enabled Smart Agriculture to address these challenges and provide Nigeria's agricultural sector with the resources it needs to both increase yields and improve environmental sustainability. In conclusion with the right policies and measures put in place, the system can provide a pathway for economic growth and environmental sustainability. Moreover, the potential of the system to reduce poverty and provide job opportunities for rural populations offers a crucial opportunity for Nigeria to remain competitive in the global economy while simultaneously supporting its environmental goals.

Results

Internet of Things (IoT) Devices

The desktop, tablet, and cellphone remain integral parts of IoT as the command centre and remotes. The desktop: Provides the user with the highest level of control over the system and its settings. The tablet: Provides access to the key features of the system in a way resembling the desktop, and also acts as a remote. The cellphone: Allows some essential settings modification and also provides remote functionality. Other key connected devices include standard network devices like routers and switches. Enabling devices for the Internet of Things is a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving information and communication technologies. With the Internet of Things, communication is extended via the Internet to all the things that surround us. The Internet of Things is much more than machine-to-machine communication, wireless sensor networks, sensor networks, 2G/3G/4G/5G, GSM, GPRS, RFID, WI-FI, GPS, microcontroller, microprocessor, etc.

These are considered to be the enabling technologies that make “Internet of Things” applications possible. Enabling technologies for the Internet of Things are considered into the following groups (Arefin et al., 2024): (1) technologies that enable “things” to acquire contextual information, (2) technologies that enable “things” to process contextual information, and (3) technologies to improve security and privacy. The first two categories can be jointly understood as functional building blocks required to build “intelligence” into “things”, which are indeed the features that differentiate the IoT from the usual Internet. The Internet of Things could be seen as the coming together of all the computing mechanisms and sensors. As information devices, it can store, retrieve, and process data for communication between individuals and organizations or between one device to another device. Certain things make up the Internet of Things as a system. Such things include sensors, biochip transponders of farm animals, and heart monitoring in plants. Others are electronic coastal waters, agriculture crops monitoring, and devices for environmental food monitoring among other sensors. These sensors collect information on already existing devices and transmit it to them for more productive and reactionary purposes.

The concept of anomaly

Anomalies are also known as outliers, abnormalities, and deviants. (Min et al., 2021) define an outlier as "an observation, which deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism." An important thing to consider in anomaly detection is the type of the anomaly. Anomalies could be classified into the following categories (Toivonen et al., 2019). Point anomalies: when an individual data point is different from the rest of the data; Contextual anomalies: when a data instance is anomalous in a specific context only, meaning that in all other situations, it would be perceived as normal. Collective anomalies: when a group of related data points is anomalous compared to the dataset; the individual data points could represent normality, while it is their actual sequence that represents an anomaly. Another important aspect to consider is the type of input data, as this requires certain techniques and poses several challenges. Some details to consider are dimensionality (one-dimensional, or multidimensional); number of attributes (unilabiate, or multivariate); and any existing relationship between the data instances. It is common for the data points to be related to each other, as is the case for sequence data, spatial data, and graph data (Shekhar et al., 2015). For sequence data, also known as temporal data, the data points are linearly ordered, representing an ordered series of events. Common examples include time series (both discrete and continuous), genome sequences, and so on. Spatial data consists of points that are related in space. Data can be both spatial and temporal, as is the case for images and video, and it is referred to as spatiotemporal. Other types of data which pose difficulties for anomaly detection include:

Evolving data: the environment which one monitors to detect anomalies can be non-stationary, meaning the data changes in time, as the characteristics of the system dynamically change. As a result, the models and algorithms need to be adaptive and to account for the changes that occur. Classic stationary algorithms do not work as the underlying data distribution changes constantly. Streaming data: it is a sequence of data points that is mainly generated by real-time data sources. It is characterized by a continuous flow and a high number of data instances over a short period. It poses multiple technical challenges such as storing and processing the data on the go or online. Furthermore, the data can also evolve and change throughout time. Correlated data: in certain situations, when multiple data sources are used to monitor one particular system, it is often likely that the generated data streams are correlated. In these cases, for detecting anomalies, it is worth analyzing also the combination of the data streams as certain data instances may not be anomalous by themselves, but they may indicate an anomaly when looked at altogether. The data instances that can be related to each other can also be part of the same stream, as is the case with sequence data. In the latter case, to detect anomalies, one needs to analyze a sequence of data instances of a certain length. Heterogeneous data: in IoT systems, it is often the case that the collected data is heterogeneous. Generally speaking, analyzing such data poses challenges in how to combine and interpret the amount of information provided.

Contaminated data: refers to situations in which the data source is affected by noise from the environment, or in which there are missing values in the data, for instance as a result of different hardware or software malfunctions. The challenge arises from the fact that it is difficult to distinguish between true anomalies and different degrees of contamination. Furthermore, the contamination could strongly affect the anomaly detection. Big data: the challenge when working with big data comes from the overwhelming size of the data, as captured by the five Vs: volume, velocity, variety, veracity, and value.

Approaches to Anomaly Detection in the Context of IoT Devices and Sensors

An anomaly is a data point that is not associated with the predicted behaviour in a modelled system (Blázquez-García et al., 2022). Anomalies are rare events or observations that deviate significantly from conventional behaviour or patterns observed in a single data point, a specific

context, or slice of time (for example, a season or quarter), or the entire dataset. In principle, anomalies result from external factors, such as sensor failure or external attack, and the purpose of a detection algorithm is to identify where an anomaly has occurred and classify/infer the cause. In the binary classification of an anomaly, the approximation model that best fits the expected data behaviour is crucial. Furthermore, the complexities of many situations require a distinct detection strategy for each application (Lin et al., 2020).

Unsupervised Learning Method

Unsupervised learning refers to learning from data where the desired output is not available or identifying anomalies with no prior knowledge of the data. A major challenge to deploying an anomaly detection system in sensor networks is the need for labelling data for use by learning algorithms. With unsupervised learning algorithms, it builds detection models without the need for manual labelling, thus reducing the deployment cost. Unsupervised learning algorithms are based on the assumption that the anomalies are rare and significantly different from the normal instances (Xiao et al., 2021).

Supervised Learning method

Supervised learning refers to machine learning techniques that train a model using a set of examples with a target output (labels). Supervised learning for anomaly detection has peculiar challenges when compared to other supervised learning applications. In practical cases, the rare samples are usually fewer in the training instances. There are modifications to algorithms for such class-balanced scenarios, to increase the impact of the rare instances onto the models. One popular approach is to make use of cost-sensitive learning (Yin et al., 2022).

Reinforcement learning

(Dang et al., 2021), explore a computation approach to learning from interaction, namely reinforcement learning (RL). This approach is the closest to how humans learn and focuses on mapping situations to actions to maximize a goal through a numerical reward signal. The main actors are a software agent, which can take actions, and the environment, whose state is affected by the actions that the agent takes. What differentiates reinforcement learning from the other types of learning is its independence from supervision and focus on decision-making in pursuit of a defined goal. Reinforcement learning is often used in IoT scenarios with multi-agent settings. For reinforcement learning to be employed in the field of anomaly detection, more attention needs to be paid to how the problem can be formulated. (Poltronieri et al., 2023) the system needs to interpret the signals and the states from a team of agents to give an alarm in the case of an intrusion (intrusion detection).

Deep Learning Method

Dang et al. (2021) explained deep learning as emerged from the traditional Artificial Neural Network (ANN). In terms of architecture, what makes deep learning different is that the hidden layers are more in-depth (deep) than the traditional Artificial Neural Network (ANN), which has few hidden layers (shallow network). The additional layers enable deep learning to learn from massive data. Another difference is that, in the more conventional machine learning, the features are manually extracted from the input, before being fed into the network for the learning process. Deep learning can be used in either a supervised or unsupervised manner. Recent works have emerged that deployed deep learning for anomaly detection in sensor systems.

Conclusion

In conclusion, the study explains the importance of having an efficient, accurate, and robust anomaly detection system, and the need for further research in this area. This study reviewed the various approaches to anomaly detection in the context of IoT devices and sensors. Methods such as unsupervised, supervised, reinforcement, and deep learning are critically evaluated.

Conflicts of Interest

The authors declare no conflict of interest

References:

Alexander, C., Han, Y., & Meng, X. (2023). Static and dynamic models for multivariate distribution forecasts: Proper scoring rule tests of factor-quantile versus multivariate GARCH models. *International Journal of Forecasting*, *39*(3), 1078–1096. https://doi.org/10.1016/j.ijforecast.2022.04.004

Arefin, Md. S., Rahman, M. M., Hasan, Md. T., & Mahmud, M. (2024). A Topical Review on Enabling Technologies for the Internet of Medical Things: Sensors, Devices, Platforms, and Applications. *Micromachines*, *15*(4), 479. https://doi.org/10.3390/mi15040479

Blázquez-García, A., Conde, A., Mori, U., & Lozano, J. A. (2022). A Review on Outlier/Anomaly Detection in Time Series Data. *ACM Computing Surveys*, *54*(3), 1–33. https://doi.org/10.1145/3444690

Chatterjee, A., & Ahmed, B. S. (2022). IoT anomaly detection methods and applications: A survey. *Internet of Things*, *19*, 100568. https://doi.org/10.1016/j.iot.2022.100568

Dang, T.-B., Le, D.-T., Nguyen, T.-D., Kim, M., & Choo, H. (2021). Monotone Split and Conquer for Anomaly Detection in IoT Sensory Data. *IEEE Internet of Things Journal*, *8*(20), 15468–15485. https://doi.org/10.1109/JIOT.2021.3073705

Dickson, M. S., & Amannah, C. I. (2023). Augmented IoT Model for Smart Agriculture and Farm Irrigation Water Conservation. *International Journal of Intelligence Science*, *13*(04), 131–163. https://doi.org/10.4236/ijis.2023.134007

Dickson, S. M. & Ijeoma Peace Okechukwu. (2024). Internet of Things (IoT) Data Communication Challenges, and Solutions. *JOURNAL OF DIGITAL LEARNING AND DISTANCE EDUCATION*, *2*(11), 803–808. https://doi.org/10.56778/jdlde.v2i11.226

Erhan, L., Ndubuaku, M., Di Mauro, M., Song, W., Chen, M., Fortino, G., Bagdasar, O., & Liotta, A. (2021). Smart anomaly detection in sensor systems: A multi-perspective review. *Information Fusion*, *67*, 64–79. https://doi.org/10.1016/j.inffus.2020.10.001

Kraft-Todd, G. T., & Rand, D. G. (2021). Practice what you preach: Credibility-enhancing displays and the growth of open science. *Organizational Behavior and Human Decision Processes*, *164*, 1–10. https://doi.org/10.1016/j.obhdp.2020.10.009

Krzysztoń, E., Rojek, I., & Mikołajewski, D. (2024). A Comparative Analysis of Anomaly Detection Methods in IoT Networks: An Experimental Study. *Applied Sciences*, *14*(24), 11545. https://doi.org/10.3390/app142411545

Li, C., Mo, L., Tang, H., & Yan, R. (2020). Lifelong Condition Monitoring Based on NB-IoT for Anomaly Detection of Machinery Equipment. *Procedia Manufacturing*, *49*, 144–149. https://doi.org/10.1016/j.promfg.2020.07.010

Lin, X.-X., Yeh, E.-H., & Lin, P. (2020). Anomaly Detection for IoT Systems. In X. (Sherman) Shen, X. Lin, & K. Zhang (Eds.), *Encyclopedia of Wireless Networks* (pp. 18–20). Springer International Publishing. https://doi.org/10.1007/978-3-319-78262-1_183

Min, M., Lee, J. J., Park, H., & Lee, K. (2021). Detecting Anomalous Transactions via an IoT Based Application: A Machine Learning Approach for Horse Racing Betting. *Sensors*, *21*(6), 2039. https://doi.org/10.3390/s21062039

Okonta, D. E., & Vukovic, V. (2024). Smart cities software applications for sustainability and resilience. *Heliyon*, *10*(12), e32654. https://doi.org/10.1016/j.heliyon.2024.e32654

Poltronieri, F., Stefanelli, C., Tortonesi, M., & Zaccarini, M. (2023). Reinforcement Learning vs. Computational Intelligence: Comparing Service Management Approaches for the Cloud Continuum. *Future Internet*, *15*(11), 359. https://doi.org/10.3390/fi15110359

Shekhar, S., Jiang, Z., Ali, R., Eftelioglu, E., Tang, X., Gunturi, V., & Zhou, X. (2015). Spatiotemporal Data Mining: A Computational Perspective. *ISPRS International Journal of Geo-Information*, *4*(4), 2306–2338. https://doi.org/10.3390/ijgi4042306

Taye, M. M. (2023). Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. *Computers*, *12*(5), 91. https://doi.org/10.3390/computers12050091

Toivonen, T., Heikinheimo, V., Fink, C., Hausmann, A., Hiippala, T., Järv, O., Tenkanen, H., & Di Minin, E. (2019). Social media data for conservation science: A methodological overview. *Biological Conservation*, *233*, 298–315. https://doi.org/10.1016/j.biocon.2019.01.023

Xiao, Q., Wang, J., Lin, Y., Gongsa, W., Hu, G., Li, M., & Wang, F. (2021). Unsupervised Anomaly Detection with Distillated Teacher-Student Network Ensemble. *Entropy*, *23*(2), 201. https://doi.org/10.3390/e23020201

Yin, C., Zhang, S., Wang, J., & Xiong, N. N. (2022). Anomaly Detection Based on Convolutional Recurrent Autoencoder for IoT Time Series. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *52*(1), 112–122. https://doi.org/10.1109/TSMC.2020.2968516