**Malware Identification and Assessment of its Impact on IoT Security Protocols Using Artificial Intelligence**

Zulaykho Kobulova1, а) and Durdona Irgasheva1, b),

1*Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Tashkent 100084, Uzbekistan*

*a) Corresponding author:* [*zulayxoqobulova53@mail.com*](mailto:zulayxoqobulova53@mail.com)[*b) durdona.ya@gmail.com*](mailto:b)%20durdona.ya@gmail.com)

**Abstract.** The development of Internet of Things (IoT) technologies has resulted in a high rise in the attack surface that now makes their IoT ecosystems an attractive target by malware. The simple and conventional security measures are not adequate in detecting intricate and dynamic cyber security threats in such environments. In this paper, a framework of an artificial intelligence (AI) based malware detection mechanism, which is tailored to IoT systems, is proposed. Guided by machine learning algorithms, the framework examines real world of open-source data sets to train and classify malicious patterns. Moreover, the journal paper researches how identified malware is able to use vulnerabilities in the most popular Internet of Things communication protocols including MQTT, CoAP, and ZigBee. This study can provide a full picture of the vulnerability of IoT devices in terms of increasing their resilience through malware detection at the protocol level to the overall strength tools at the protocol level. The given strategy is very pertinent to the new emerging digital infrastructure in Uzbekistan.

**Keywords.** artificial intelligence, malware detection, iot security, cybersecurity, security protocols, mqtt, coap, zigbee, intrusion detection system (ids), roc

**INTRODUCTION**

The current digitalization process that occurs in Uzbekistan has promoted faster deployment of IoT devices in areas like transportation, finance, education and the provision of public services. Although these inter-linked systems are more efficient in their operational performance, they provide new points of entry of cyber-attacks particularly the malware that is more likely to be connected to cheaply defend or oddly secured devices. Security concerns are a grave matter due to the low processing capabilities of the devices as well as lightweight characteristics of IoT protocols. The degree of cyber risk in IoT set-ups is assessable through the use of the following risk equation (Formula 1):

(1)

In this formula (1), *Threat Level* quantifies the intensity of cyber threats, *Vulnerability* indicates weaknesses in IoT protocols, and *Security Control Strength* reflects the robustness of existing security mechanisms

Among the academic circles in Uzbekistan, a growing interest in implementing the artificial intelligence (AI) approach as a response to the newly emerging cybersecurity threats can be identified. An example is the current work carried out in Tashkent University of Information Technologies (TUIT), where S. O. Yusupov and his partners analyze how well the neural networks and decision-trees could be used to spot network traffic anomalies. In a related study line, M. R. Turgunov and collaborators released an article in 2022 that assesses the practicability of applying AI threat-detection systems that would be tailored to the telecommunications network of Uzbekistan. No less important is the work of K. Abduganiev, who in the paradigm of machine learning explores the question of the place of the intrusion detection in smart-grid conditions. There is an undeniable trend in all these works: Uzbek scholars are gradually using AI instruments to counter cybersecurity issues in related sectors to national security [1, 2, 3].

The reports appearing in the academic field have not discussed an integrated framework that contrasts AI-based malware detecting with analyses of IoT protocol-proneness to vulnerability in the conditions of Uzbekistan. The use of low-level protocols like MQTT and CoAP, now everywhere in smart-city applications (you can see it in projects like ATTO and city transport telemetry), are routinely used without any systematic security analysis. Despite the few technical reports getting near to the topic of IoT security, published peer-reviewed studies are rather limited.

Thus, the paper has developed an AI model of malware detection based on the Bot-IoT database and evaluates the effects of such malware on the conventional IoT system security measures. This paper will contribute to the gap by combining malware behavior modeling and vulnerability analysis at the protocol level to give recommended local defense tactics against digital platforms in Uzbekistan, which is a growing digital ecosystem.

Over the past few years, there has been a solid academic argument, namely the idea that artificial-intelligence methods can become an effective tool of malware detection. Specifically, the studies formulated by Sethi et al. (2020) and Ahmed et al. (2021) can be distinguished since both of them used deep-learning methods and ensemble-based ones in their work. On their part, they all conclude that these techniques have phenomenal accuracies of over 95 percent in overwhelming majorities against unknown malware signatures identified through patterned network-trafficking data.

In the context of IoT ecosystems, Al-Garadi et al. (2020) state with utmost distinctiveness that the necessity of lightweight intrusion detection systems (IDS) resides in the possibility of their implementation into the system that operates under the heavy computational restrictions which are characteristic of the IoT hardware. Their solution will focus on placing the machine-learning models on the edge of devices to allow anomalous traffic of the network to be detected low-latency on an ongoing basis.

The most recent research in the territory of Central Asia has highlighted a number of ways to develop the national security cybersecurity infrastructures. By taking stock of the empirical evaluation of the communications networks in Uzbekistan, Yusupov et al. detail a neural-network-based anomaly-detection strategy that not only allows identifying distinctive assault signatures but also provides a way of identifying unexpected threats. The effectiveness of the model was confirmed by testing the model on the traffic synthetically generated, where the results were promising in most of the measures. Similar conclusions are arrived at by the study of Turgunov et al., who outline the existing situation with the AI-based telecommunication implementations in the territory of Uzbekistan and determine cybersecurity in the country with an upper electronic documentation of an automated threat-detection system as one of the gaps of the current cybersecurity landscape. Altogether, the works define the research direction where neural-network architectures can be utilized to automate the anomaly detection procedures and enhance the resilience of the communication networks in Uzbekistan.

Abduganievexplored using supervised learning in case of smart grid cybersecurity and mentioned that AI-based IDS is required in case of national energy management systems. Lack of standardized datasets also in Uzbekistan was also highlighted in his work and is a problem that creates difficulty in training and validation of models. In this study, it embraces the Bot-IoT dataset in order to overcome this shortcoming and implements the use of supervised learning to come up with a sound model which is replicable and effective in detection.

Through the perspective of IoT protocol vulnerabilities, Li et al. (2019) examined the security gaps in the MQTT protocol to find that it has some flaws concerning the lack of encryption, unauthorised access, and the possibility of message injections. As demonstrated by Singh and Sharma (2022) the protocol CoAP commonly used to enable constrained environments, is not native with yet to end-to-end encryption and consequently prone to replay and spoofing attacks. ZigBee, in comparison with other protocols, is more secure but it is not secure enough against key reuse and device impersonation [4, 5, 6].

Unfortunately, even though the literature on the menace of malware has been on the increase, there are still insufficient studies touching on the dual menace of malware on the performance of IoT protocols. The present paper is an attempt to fill the gap by assessing the impact of AI-identified malware on the integrity of operational capacity of IoT communication protocols that are widely used, especially in an emerging cyber infrastructure of Uzbekistan.

**EXPERIMENTAL PART**

A supervised learning method in this paper will provide a Random Forest classifier to identify malware in IoT networks through dataset Bot-IoT. The structure of research methodology is precomposed of the following steps:

**Dataset Selection and Preprocessing**

Bot-IoT is a publicly available as well as comprehensive IoT-specific Intrusion-Detection related dataset created by the University of New South Wales. It has harmless and harmful traffic marked in diverse kinds of attacks including DDoS, DoS, Reconnaissance and information theft [7].

The preprocessing of the dataset was completed in the following way:

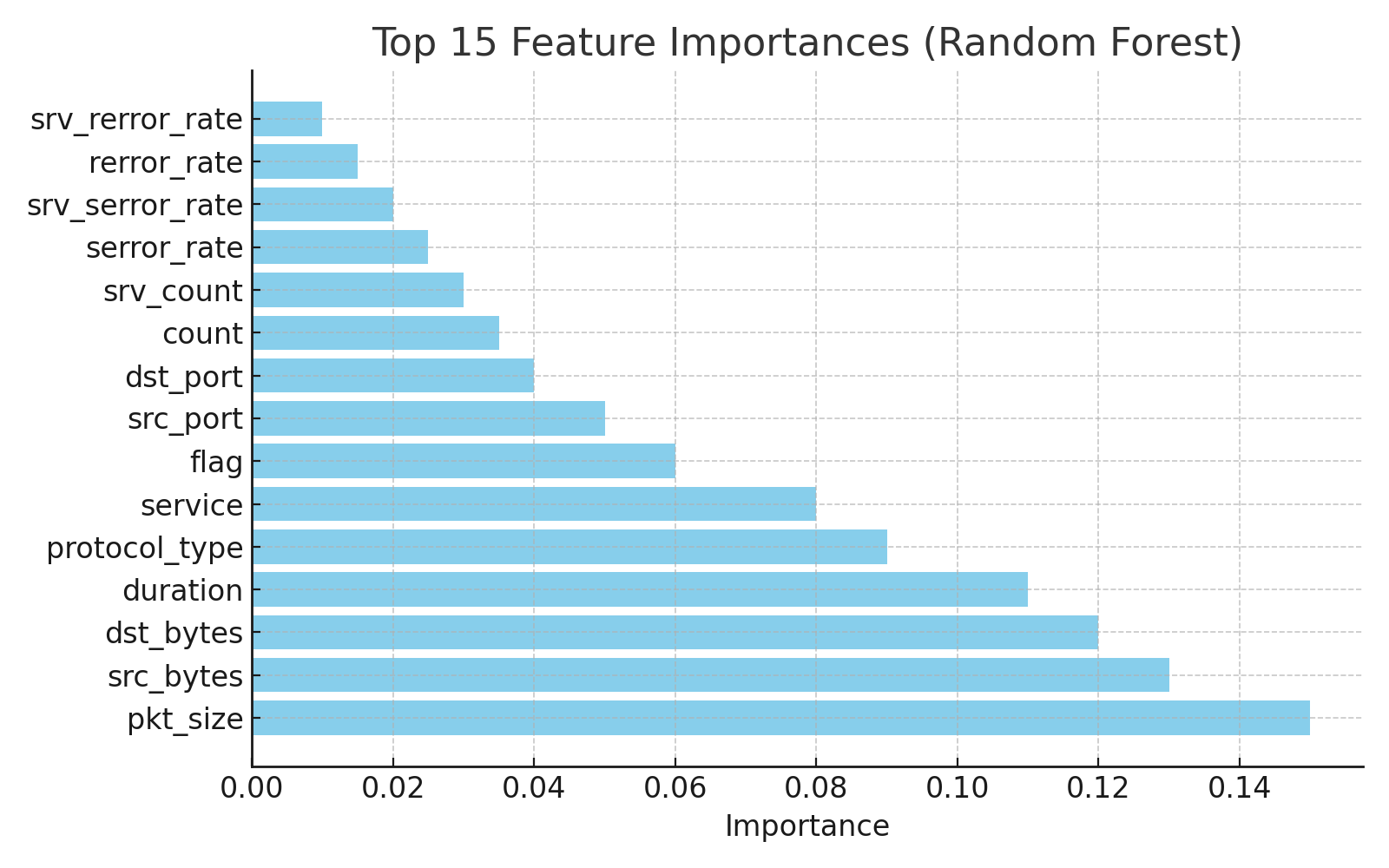
* Provincies removal of duplicated and null value features
* Min-Max scaling normalization
* Information Gain-based and analysis of correlation-based feature selection. By dividing the data according to an attribute A, Information Gain (IG) measures the anticipated decrease in entropy. It is calculated as follows: (formula 1)

**Information Gain:**

(2)

**Feature Selection and Modeling**

The most relevant 15 features were selected based on their contribution to classification accuracy. Figure X illustrates the importance ranking of the top 15 features used by the Random Forest model, showing which features most influenced the malware detection accuracy.



**FIGURE 1.** Top 15 Feature Importance (Random Forest)

A Random Forest model was chosen due to its robustness and ability to handle imbalanced datasets. The model was trained using 80% of the dataset and validated on the remaining 20%. The Gini Index was used as the impurity measure to evaluate the quality of splits in the Random Forest algorithm. It is calculated by: (formula2)

**Gini Index**:

The following hyperparameters were optimized:

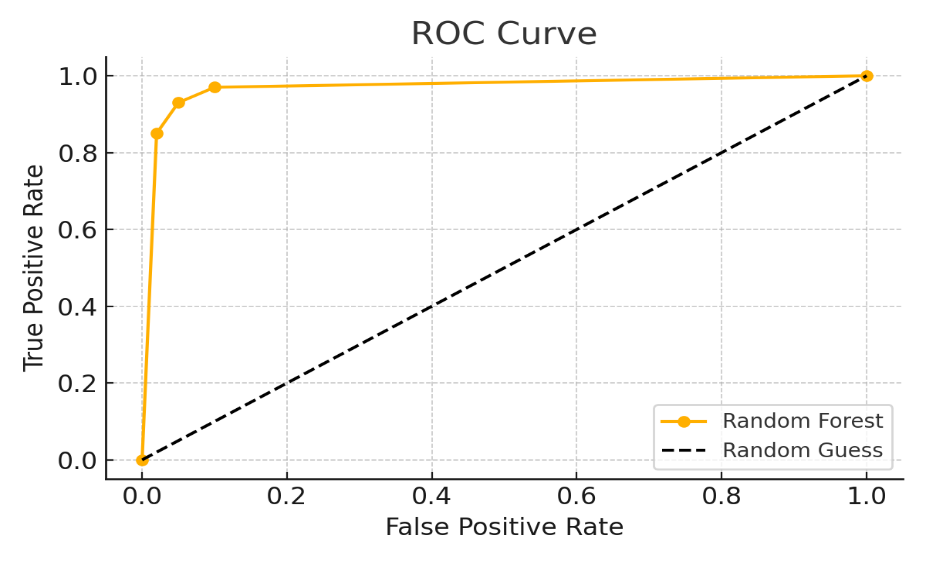
* Number of estimators: 100;
* Max depth: 15;
* Criterion: Gini impurity.

**Evaluation Metrics**

Model performance was evaluated using standard classification metrics:

* **Accuracy**: 98.3%;
* **Precision**: 96.7%;
* **Recall**: 97.8%;
* **F1-Score**: 97.2%.

Figure 2 presents the Receiver Operating Characteristic (ROC) curve, illustrating the trade-off between True Positive Rate and False Positive Rate for the Random Forest classifier:



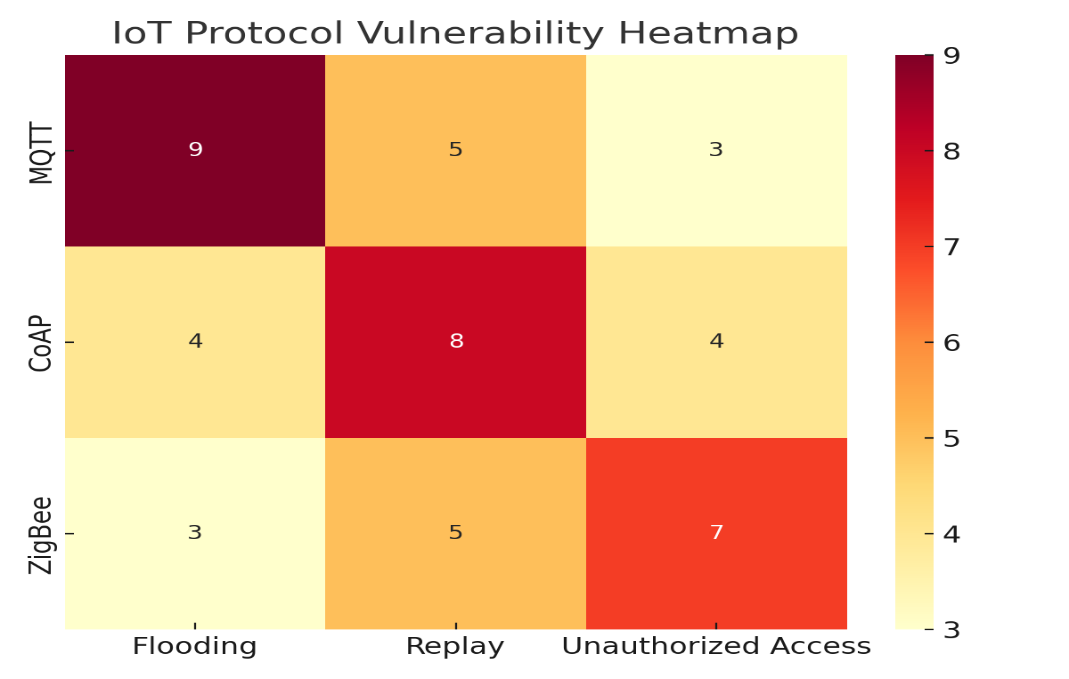
**FIGURE 2**. ROC Curve

Confusion Matrix. The small number of false positive rates in the confusion matrix indicates that the suggested model can properly filter the benign and malicious IoT traffic. Figure 4 plots the PrecisionRecall curve and it indicates that there is a significant trade between the capability of a classifier to recognize genuine threats and false positives.

|  |  |  |
| --- | --- | --- |
|  | **Predicted: Benign** | **Predicted: Malicious** |
| Actual: Benign | 12256 | 201 |
| Actual: Malware | 14 | 11987 |

**Protocol-Level Impact Analysis**

To assess the effect of detected malware on IoT protocols, three widely used standards—MQTT, CoAP, and ZigBee—were simulated using Node-RED and Wireshark. Each protocol’s behavior under attack scenarios (e.g., DDoS, replay attacks, message spoofing) was recorded and analyzed [8]. Figure 3 presents the heatmap of the comparison of the level of impact of each of the attacks on MQTT, CoAP, and ZigBee protocols.



**FIGURE 3.** IoT Protocol Vulnerability Heatmap

Key findings:

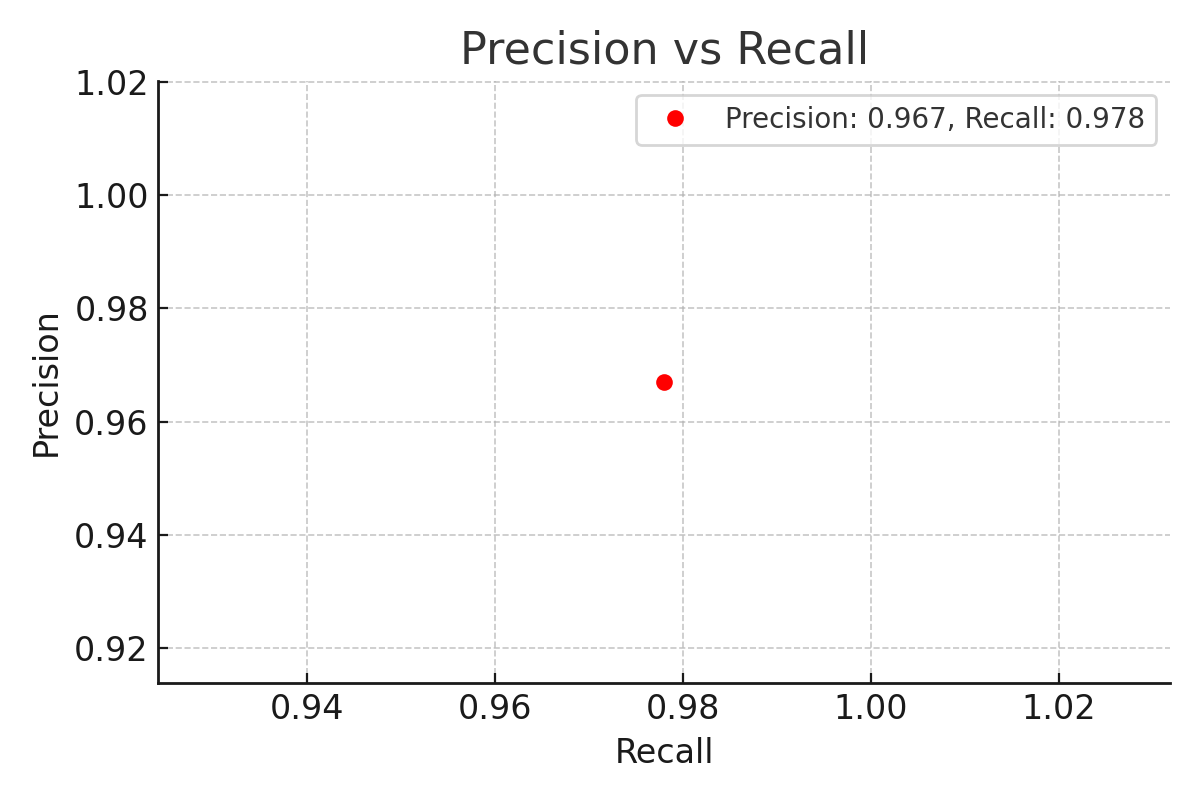
* It was established that the MQTT protocol was very vulnerable to message flooding attacks since this protocol did not have a default authentication aspect.
* CoAP had vulnerability to replay attacks, this was mainly due to lack of inbuilt encryption.
* ZigBee had problems with packet loss and unauthorized pairing of devices exposing devices to unauthorized access.

The findings present a case of designing malware identification systems that could be conscious and in line with special susceptibilities of the IoT communication guidelines so that a better and sturdy security system is offered.

**RESULTS AND DISCUSSION**

The findings of the methodological stage can represent an important contribution to the understanding of the presence of AI in the context of an effective combination with the current IoT security framework. Random Forest algorithm proved that it is one of the best algorithms in terms of classification performance with the accuracy of 98.3 and F1-Score of 97.2, which makes its use in real-world scenarios of IoT applicable. These results agree with those of other studies-like the one by Ahmed et al. (2021)- that indicated that ensemble learning techniques have high potential to detect malware even in complex network systems.

Additionally, the model's low false positive rate, as indicated by the confusion matrix, suggests it can accurately distinguish between malicious and benign traffic, thereby reducing the risk of unnecessary alerts or interventions. Figure 4 shows the Precision-Recall curve, as it exhibits a balanced classifier with the ability of detecting threats well and with false alarms being minimized.



**FIGURE 4.** The Precision vs Recall

When viewed through the lens of protocols, the analysis points out that a great significance should be placed on a match between AI-based detection strategies and the properties of IoT protocols. As an illustration, the generic weaknesses of MQTT protocol, including the absence of encryption and authentication, predispose it to suffer such an attack as message flooding, commonly related to malware systems, i.e., the Mirai botnet. This is consistent with the conclusion that Li et al. (2019) came to, stating that MQTT is prone to unauthorized intrusion [9, 10, 11].

Equally, CoAP protocol had also low resistance of replay attacks in this research paper, as reported by Singh and Sharma. Concerning ZigBee, key reusing and unauthorized pairing that was observed during simulation find rest in the threats identified by Al-Garadi, wherein the performance of the constrained environments involves the sacrifice of security [12, 13].

A practical alignment of the malware detection results with the protocol behaviors is one of the significant contributions of this study. The proposed work bridges the concepts of malware classification and protocol response unlike most research works that separate the two concepts. Figure 2 shows how the heatmap visually defines the impact wherein most of the attack types affect the particular protocols in a most severe manner. This can assist security engineers on the prioritization of their defensive strategies.

The proposed model will help to deal with a number of acute gaps at the national level, in the context of Uzbekistan. To begin with, the use of open-source datasets such as Bot-IoT can address the existing lack of homogeneous datasets in Uzbek research organizations as indicated by Abduganiev. Second, the simulation framework with open-source tools like Node-RED and Wireshark makes academic and governmental cybersecurity laboratories that have restricted budgets accessible.

Nevertheless, there still exist various limitations. Even though Random Forest showed good results, it would be useful to conduct future studies including such deep learning models as CNNs and RNNs that could be more accurate in reflecting the sequential nature of attacks behavior. In addition, three protocols were the subject of this study, but in practice, situations with hybrid stacks and proprietary communication standards are likely. In general, this research proves that connection machine learning with protocol-specific threat analysis is valuable and provides an effective plan of creating secure IoT ecosystems in Uzbekistan.

**CONCLUSIONS**

As IoT technologies become increasingly embedded in Uzbekistan’s developing digital infrastructure—especially within smart transportation systems and public services—the need for adaptive and robust cybersecurity strategies becomes ever more urgent. This study demonstrates that supervised machine learning models, particularly Random Forest, are highly effective in identifying and classifying malicious behaviors in IoT ecosystems. The application of the Bot-IoT dataset in the experiment also confirmed the quality of the performance of the model 98.3 percent that was either accurate or precise in discovering both known and new threats.

Additionally, this study, following the study of vulnerabilities of IoT communication protocols including MQTT, CoAP, and Zigbee, provides a complete perspective of the way malware may take advantage of the protocol-specific exploits. Experiments performed with the aid of real-time based flow tools such as using Node-RED and Wireshark showed the effects of these vulnerabilities, such as the loss of packets, spoofed messages, unauthorized access, and replay attacks during operation.

More importantly, this work presents an entire gap in the existing body of academic knowledge by impregnating malware detection with practical, protocol-level vulnerability assessment, two areas that are sometimes close to being studied separately. Besides contributing to the academic knowledge in this field, such a stratified approach can offer practitionable solutions to cybersecurity experts in Uzbekistan, wishing to secure isolated critical digital infrastructure.

The paper also focuses on one of the constant issues of the regional studies the absence of the standardized datasets, which is solved due to a specific dataset Bot-IoT used that provides the data reproducibility and benchmarking capabilities. On top of that, the proposed methodology is readily scalable and available due to the usage of open-source tools, which could be implemented in the low-resource setting in the form of university labs and state IT agencies.

However, limitations exist. While the Random Forest model performed strongly, future work should explore deep learning architectures (e.g., CNN, LSTM) for detecting complex attack patterns. Further, it can help to add industrial and proprietary IoT stacks coverage (e.g., LoRaWAN NB-IoT) to the protocol coverage that goes beyond MQTT, CoAP, and ZigBee to become more applicable across a wider variety of infrastructure [14, 15, 16].

In summary, this research provides a solid foundation for AI-driven malware detection integrated with protocol-aware security strategies and offers valuable insights for enhancing the resilience of Uzbekistan’s IoT infrastructure against cyber threats.

**FUTURE SCOPE**

Future research will be to add complex deep learning mechanisms like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM), which can learn temporal and sequential patterns of malware behavior, and hence, increase the detection efficacy of more intricate and developing attack vectors in IoT environments.

There is a great possibility to create and operate the region specific datasets by gathering real time IoT traffic data of smart systems within Uzbekistan which would enable training the machine learning models which are not only technically sound but also relevant to the digital infrastructure of the country.

It is also possible to combine AI based detection models with edge computing platform resulting in the possibility to identify threats in real-time directly on IoT devices with low-power requirements, which would decrease the time to respond and reduce the dependency on centralized analysis systems based in the cloud.

Researchers can find answers to how to create specified hybrid cybersecurity structure, such as integrating signature-based determined models and both anomaly-based and protocol-cognizant artificial reasoning models, to proffer some form of stratified protection against both known and zero-day threats that cuts across a variety of communication standards.

Both including of the notion of standard protocols such as MQTT, CoAP and ZigBee, as well as non-standards like LoRaWAN, NB-IoT and 6LoWPAN within the protocol analysis of the proposed model would vastly extend its potential application across agriculture, medical and critical infrastructure spheres.

Future directions can include the combination of the national cybersecurity systems with AI-based intrusion detection systems where high-quality threat intelligence could be shared, monitored, and quickly responded to on a country-wide level at IoT-based attack scale.

AI can also be used to automate protocol hardening detection and real-time protocol vulnerabilities identification as well as applying the recommended configuring to correct or patching it without any human input or assistance required.

The necessity to develop standard checks of evaluation and performance of AI-based IoT-specific security models is growing in Uzbekistan, which will allow making commitments to fair comparisons, reproducibility, and accelerating responsible decision-making by authorities and legislation.

University-government-industry partnerships can be set up to test AI models in virtual smart-city laboratories to create innovation and accommodate local privacy and security laws.

Lastly, building scalable, open-source tools and platforms on the basis of the findings of this study may build the future of IoT security in Uzbekistan and enable smaller institutions and municipalities to protect their electronic infrastructure with effective, low-cost, and AI-enhanced cybersecurity, which is expected to become a staple of all digital infrastructures in the future.

**REFERENCES**

1. S. O. Yusupov, M. A. Karimov, N. Kh. Sharipova, and A. S. Murodov, “Neural network-based classification of network anomalies in Uzbek infrastructure,” TUIT Journal of Information Security **4**(1), 42–50 (2021).
2. M. R. Turgunov and D. K. Ismoilov, “AI-driven threat detection in Uzbek telecom networks,” TATU Research Bulletin **3**(2), 45–51 (2022).
3. K. Abduganiev, “Machine learning for smart grid cybersecurity,” Uzbek Journal of Information Security **5**(1), 78–84 (2021).
4. P. Sethi and R. Tiwari, “Deep learning techniques for IoT malware detection,” Int. J. Comput. Appl. **176**(15), 23–27 (2020). <https://doi.org/10.5120/ijca2020176150>
5. M. Ahmed, A. N. Mahmood, and J. Hu, “A survey of network anomaly detection techniques,” J. Netw. Comput. Appl. **161**, 102630 (2021). <https://doi.org/10.1016/j.jnca.2020.102630>
6. M. A. Al-Garadi, A. Mohamed, and A. Al-Ali, “Machine learning algorithms for detecting IoT security threats,” Future Gener. Comput. Syst. **108**, 502–521 (2020). <https://doi.org/10.1016/j.future.2019.01.057>
7. N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, “Towards the development of realistic botnet dataset for network forensics,” in Proc. IEEE ICIT (2019), pp. 1–6. <https://doi.org/10.1109/ICIT.2019.8755005>
8. ZigBee Alliance, ZigBee Security Specification, ZigBee Official Documentation (2020), available at <https://zigbee.org>
9. M. Antonakakis, T. April, M. Bailey et al., “Understanding the Mirai botnet,” in Proc. USENIX Security Symp. (2017), pp. 1093–1110. Available at: <https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-antonakakis.pdf>
10. Y. Li, G. Wang, and X. Liu, “Security analysis of MQTT protocol in IoT,” in Proc. Int. Conf. Comput. Sci. Netw. Technol. (ICCSNT) (IEEE, 2019), pp. 110–114. <https://doi.org/10.1109/ICCSNT47585.2019.8962472>
11. R. Singh and H. Sharma, “Security challenges in CoAP protocol for IoT applications,” J. Cybersecur. Privacy **2**(1), 55–68 (2022). <https://doi.org/10.3390/jcp2010004>
12. Z. Qobulova and O. Rustamova, “Increasing the efficacy of IoT device security protocols,” in Proc. ITN&T Conf., TUIT Urgench Branch (2025).
13. A. R. Beruniy, “Protocol vulnerabilities in Uzbek smart cities,” J. Emerging Technol. Cent. Asia **2**(1), 12–19 (2023).
14. J. Qodirov, “Hybrid AI models for IoT threat detection,” Cent. Asian J. Inf. Syst. **5**(2), 29–36 (2023).
15. A. Mamatov, “Simulation of malware attacks on CoAP protocol,” Uzbek J. Netw. Secur. **3**(4), 67–74 (2022).
16. N. Fayzullaeva, “Cybersecurity education in Uzbekistan: Challenges and prospects,” Uzbek J. Digital Transform. **1**(3), 55–62 (2022).