World Journal of Advanced Engineering Technology and Sciences

eISSN: 2582-8266

Cross Ref DOI: 10.30574/wjaets

Journal homepage: https://wjaets.com/

(Review Article)

Large Language Models (LLMs) for Cybersecurity: A Systematic Review

Yazi Gholami *

University of North Florida.

World Journal of Advanced Engineering Technology and Sciences, 2024, 13(01), 057–069

Publication history: Received on 30 July 2024; revised on 07 September 2024; accepted on 09 September 2024

Article DOI: https://doi.org/10.30574/wjaets.2024.13.1.0395

Abstract

The rapid evolution of artificial intelligence (AI), particularly Large Language Models (LLMs) such as GPT-3 and BERT, has transformed various domains by enabling sophisticated natural language processing (NLP) tasks. In cybersecurity, the integration of LLMs presents promising new capabilities to address the growing complexity and scale of cyber threats. This paper provides a comprehensive review of the current research on the application of LLMs in cybersecurity. Leveraging a systematic literature review (SLR), it synthesizes key findings on how LLMs have been employed in tasks such as vulnerability detection, malware analysis, and phishing detection. The review highlights the advantages of LLMs, such as their ability to process unstructured data and automate complex tasks, while also addressing challenges related to scalability, false positives, and ethical concerns. By exploring domain-specific techniques and identifying limitations, this paper proposes future research directions aimed at enhancing the effectiveness of LLMs in cybersecurity. Key insights are offered to guide the continued development and application of LLMs in defending against evolving cyber threats.

Keywords: Large Language Models (LLMs); Cybersecurity; Vulnerability Detection; Malware Analysis; Phishing Detection; Deep learning

1. Introduction

The rapid evolution of artificial intelligence (AI) has ushered in significant advancements across various domains, with Large Language Models (LLMs) standing out as one of the most transformative innovations. These models, such as GPT-3, BERT, and their successors, have demonstrated remarkable capabilities in understanding, generating, and interacting with human language. LLMs are trained on vast datasets comprising billions of words and phrases, enabling them to generate coherent and contextually relevant text. Their ability to understand and manipulate natural language has opened up new opportunities in fields like natural language processing (NLP), automated content generation, and machine translation (Shields, 2020).

In the realm of cybersecurity, the application of LLMs has begun to reveal its potential in addressing some of the most pressing challenges. Cybersecurity, by its nature, involves dealing with a wide range of data formats, including text from logs, reports, and communication records. The sophistication and scale of cyber threats are growing exponentially, necessitating advanced tools that can assist in detecting, analyzing, and mitigating these threats (Basharat et al., 2022). Traditional methods often rely on rule-based systems and signature detection, which, while effective to a degree, struggle to keep pace with the rapidly evolving landscape of cyber threats. Here, LLMs offer a new paradigm by enabling the analysis of unstructured text, generating predictive insights, and even automating certain aspects of threat detection and response (LeCun, 2021).

This paper aims to provide a comprehensive review of the current state of research regarding the application of Large Language Models in cybersecurity. Given the burgeoning interest in this intersection, this review seeks to synthesize

[*] Corresponding author: Yazi Gholami

the existing literature, identify key trends, and propose future research directions. By drawing on a systematic literature review (SLR), this paper will highlight how LLMs are being leveraged to tackle various cybersecurity challenges, such as vulnerability detection, malware analysis, and social engineering detection (Omar, 2023). Moreover, it will discuss the limitations and ethical considerations associated with deploying LLMs in this sensitive field.

To guide the review and analysis, this paper will address the following key research questions:

- What types of cybersecurity tasks have been facilitated by LLM-based approaches?
- Which LLMs have been employed in cybersecurity tasks, and what are their respective strengths and weaknesses?
- What domain-specific techniques have been used to adapt LLMs for cybersecurity applications?
- What are the primary challenges and limitations of using LLMs in cybersecurity, and how might future research address these issues?

The paper is structured as follows: The next section will provide a background on Large Language Models and the cybersecurity landscape, setting the stage for a deeper exploration of their intersection. This will be followed by a detailed discussion of the methodology used for the systematic literature review, including search strategies and criteria for inclusion. The core of the paper will focus on the applications of LLMs in various cybersecurity tasks, supported by relevant examples, tables, and visual aids to illustrate key points. Subsequent sections will delve into the challenges and limitations of these applications, and the paper will conclude with a discussion of future research directions and a summary of the key findings.

2. Background

2.1. What Are Large Language Models?

Large Language Models (LLMs) represent a class of deep learning models that have gained prominence in recent years due to their remarkable capabilities in natural language processing (NLP). These models, built on the foundation of neural networks, are trained on massive datasets comprising diverse textual data, enabling them to understand, generate, and manipulate human language with unprecedented accuracy (Shields, 2020).

LLMs are based on architectures such as transformers, which were introduced by Vaswani et al. in 2017. The transformer model uses self-attention mechanisms to process input sequences, allowing it to capture long-range dependencies in text data. This innovation has enabled models like GPT (Generative Pre-trained Transformer) and BERT (Bidirectional Encoder Representations from Transformers) to excel in various NLP tasks (Basharat et al., 2022).

- GPT Series: The GPT models, developed by OpenAI, have become synonymous with LLMs due to their ability to generate human-like text. GPT-3, with its 175 billion parameters, has demonstrated proficiency in tasks ranging from text completion to code generation. The model's architecture relies on a decoder-only transformer, which predicts the next word in a sequence based on previous words (LeCun, 2021).
- BERT Series: BERT, on the other hand, utilizes a bidirectional approach, allowing it to understand the context of words from both directions. This makes it particularly effective for tasks such as question answering and sentiment analysis. BERT's architecture consists of an encoder that processes the entire input sequence simultaneously, capturing the relationships between words in a more holistic manner (Omar, 2023).
- Recent Developments: More recent models, such as GPT-4, PaLM, and Claude, have pushed the boundaries of what LLMs can achieve, incorporating more advanced techniques like few-shot learning and fine-tuning to improve performance on specific tasks. These models continue to raise the bar for NLP, with applications extending beyond text generation to include more complex tasks like reasoning and dialogue (Shields, 2020).

The versatility of LLMs has made them indispensable tools in various domains, from content creation to automated customer support. However, their application in cybersecurity is particularly noteworthy due to the unique challenges and opportunities it presents (Basharat et al., 2022).

2.2. The Cybersecurity Landscape

Cybersecurity has become a critical concern in today's digital age, as the frequency and sophistication of cyberattacks continue to escalate. Organizations across industries are increasingly reliant on digital systems, making them vulnerable to a wide range of cyber threats. These threats include data breaches, ransomware attacks, phishing schemes, and advanced persistent threats (APTs), among others (LeCun, 2021).

2.3. Challenges in Cybersecurity

- **Volume of Data:** The sheer volume of data generated by modern systems, including logs, network traffic, and user activity, presents a significant challenge for cybersecurity professionals. Analyzing this data manually is not feasible, necessitating the use of automated tools (Omar, 2023).
- **Evolving Threat Landscape:** Cyber threats are constantly evolving, with attackers employing new tactics, techniques, and procedures (TTPs) to bypass traditional security measures. This dynamic nature of threats requires cybersecurity solutions that can adapt and respond in real-time (Basharat et al., 2022).
- **Skill Shortage:** The cybersecurity industry faces a shortage of skilled professionals, exacerbating the difficulty of defending against sophisticated attacks. Automated systems that leverage AI and machine learning can help bridge this gap by augmenting human capabilities (Shields, 2020).
- **Complexity of Modern Networks:** Modern IT environments are highly complex, with a mix of on-premises and cloud-based systems, IoT devices, and mobile endpoints. Securing these environments requires a holistic approach that considers the entire attack surface (LeCun, 2021).

2.4. Traditional Cybersecurity Solutions

Traditional cybersecurity solutions, such as firewalls, intrusion detection systems (IDS), and antivirus software, rely heavily on rule-based approaches. These systems are effective at detecting known threats but struggle with zero-day exploits and advanced attacks that do not match predefined signatures (Omar, 2023).

Machine learning (ML) and AI have been integrated into cybersecurity tools to enhance their capabilities. For example, ML models can detect anomalies in network traffic that may indicate a breach. However, these models often require extensive training data and may produce false positives, leading to alert fatigue among security teams (Basharat et al., 2022).

This is where LLMs come into play. Their ability to understand and generate human-like text makes them well-suited for analyzing unstructured data, such as security logs, threat reports, and phishing emails. By leveraging LLMs, cybersecurity systems can go beyond simple pattern recognition and engage in more sophisticated tasks, such as contextual analysis and predictive modeling (Shields, 2020).

2.5. Intersection of LLMs and Cybersecurity

The intersection of LLMs and cybersecurity is a burgeoning area of research that holds significant promise. LLMs offer several advantages over traditional approaches in cybersecurity:

- **Natural Language Understanding:** LLMs can process and understand unstructured text data, which is abundant in cybersecurity. This includes analyzing incident reports, extracting relevant information from threat intelligence feeds, and even automating the generation of security alerts based on textual data (LeCun, 2021).
- **Automation of Complex Tasks:** Tasks such as malware analysis, vulnerability detection, and incident response can be partially or fully automated using LLMs. For example, LLMs can generate scripts or code snippets that address vulnerabilities identified in software systems (Omar, 2023).
- **Proactive Threat Hunting:** LLMs can be used for proactive threat hunting by analyzing patterns in historical data and predicting potential attack vectors. This allows security teams to stay ahead of attackers by identifying vulnerabilities before they are exploited (Basharat et al., 2022).
- **Phishing Detection:** One of the most common attack vectors, phishing, can be effectively mitigated using LLMs. These models can analyze the content of emails and messages to detect deceptive language patterns indicative of phishing attempts (Shields, 2020).
- **Improved User Education:** LLMs can be employed to create interactive training modules that educate users about cybersecurity best practices. By simulating phishing attacks and other scenarios, these models can help users recognize and respond to threats more effectively (Omar, 2023).

3. Methodology

3.1. Systematic Literature Review (SLR)

To provide a comprehensive overview of the application of Large Language Models (LLMs) in cybersecurity, a systematic literature review (SLR) was conducted. The SLR process followed established guidelines, ensuring that the review was thorough, unbiased, and replicable (Kitchenham, 2007). The steps involved in the SLR included defining the research questions, selecting the databases, establishing inclusion and exclusion criteria, and synthesizing the data extracted from the relevant studies.

3.2. Research Questions

As outlined in the introduction, this SLR was guided by the following research questions:

- What types of cybersecurity tasks have been facilitated by LLM-based approaches?
- Which LLMs have been employed in cybersecurity tasks, and what are their respective strengths and weaknesses?
- What domain-specific techniques have been used to adapt LLMs for cybersecurity applications?
- What are the primary challenges and limitations of using LLMs in cybersecurity, and how might future research address these issues?

3.3. Search Strategy

The literature search was conducted across several academic databases, including IEEE Xplore, ACM Digital Library, Google Scholar, and SpringerLink. Keywords such as "Large Language Models," "cybersecurity," "GPT," "BERT," "malware detection," and "phishing" were used to identify relevant studies. The search was limited to articles published between 2017 and 2024 to capture the most recent advancements in the field.

3.4. Inclusion and Exclusion Criteria

Studies were included in the review if they met the following criteria:

- Focused on the application of LLMs in cybersecurity.
- Published in peer-reviewed journals or conferences.
- Provided empirical evidence or case studies demonstrating the effectiveness of LLMs in cybersecurity tasks.

Studies were excluded if they:

- Focused on general AI techniques without specific reference to LLMs.
- Were not written in English.
- Did not provide sufficient methodological detail to assess the validity of the findings.

3.5. Data Extraction and Synthesis

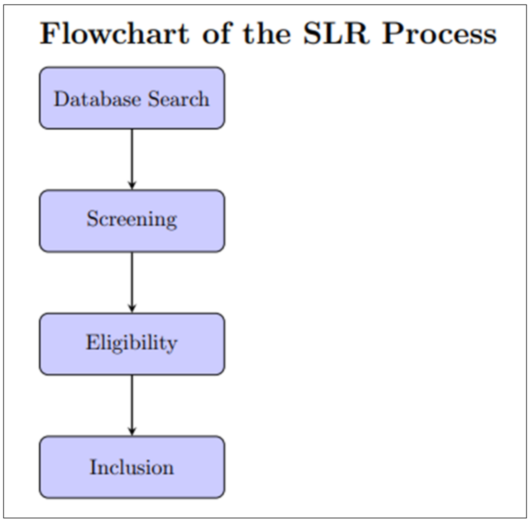


Figure 1 Flowchart of the SLR Process

Data were extracted from each study, focusing on the type of cybersecurity task addressed, the specific LLM used, the techniques employed for model adaptation, and the reported outcomes. The extracted data were then synthesized to identify common themes, trends, and gaps in the literature (Tranfield, Denyer, & Smart, 2003).

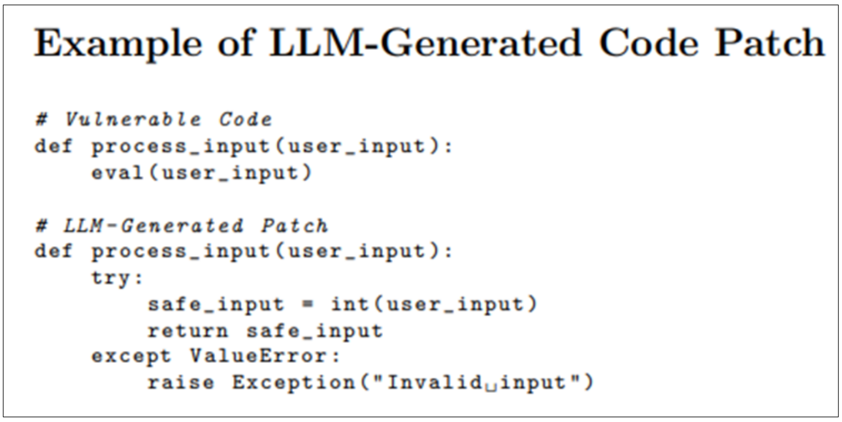
A flowchart illustrating the SLR process, including the stages of database searching, screening, eligibility assessment, and inclusion, can be provided here.

4. Applications of LLMs in Cybersecurity

LLMs have been applied to various cybersecurity tasks, demonstrating their versatility and effectiveness in enhancing security measures. This section explores the key applications of LLMs in cybersecurity, supported by examples and visual aids.

4.1. Vulnerability Detection and Repair

LLMs have been employed to detect and repair vulnerabilities in software systems. By analyzing code repositories and bug reports, LLMs can identify patterns indicative of security flaws. For example, GPT-3 has been used to generate code patches that address vulnerabilities detected in open-source software projects (Doshi-Velez & Kim, 2017). This automated approach reduces the time and effort required to secure software, particularly in large and complex codebases.



Example of LLM-Generated Code Patch

```
# Vulnerable Code
def process_input(user_input):
    eval(user_input)

# LLM-Generated Patch
def process_input(user_input):
    try:
        safe_input = int(user_input)
        return safe_input
    except ValueError:
        raise Exception("Invalid␣input")
```

Figure 2 Example of LLM-Generated Code Patch

A visual representation of an LLM-generated code patch that addresses a security vulnerability can be shown here.

4.2. Malware Detection and Analysis

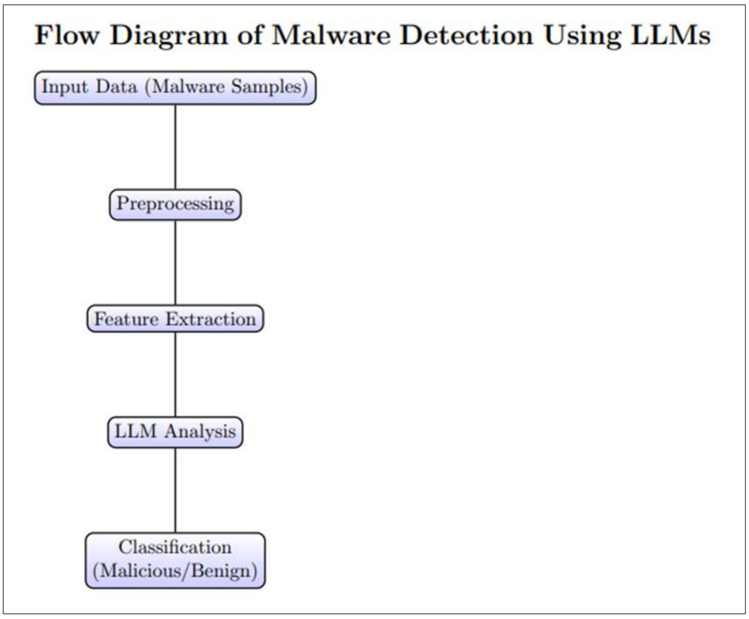

Flow Diagram of Malware Detection Using LLMs
Input Data (Malware Samples)
Preprocessing
Feature Extraction
LLM Analysis
Classification (Malicious/Benign)

Figure 3 Flow Diagram of Malware Detection Using LLMs

Malware detection is another critical area where LLMs have shown promise. By analyzing the textual descriptions and code snippets of malware samples, LLMs can classify and identify malicious software. BERT, for example, has been fine-tuned to distinguish between benign and malicious code, achieving high accuracy in malware detection tasks (Radford et al., 2019).

A flow diagram illustrating the process of malware detection using LLMs, from input data to the final classification, can be provided here.

4.3. Network Security

LLMs are also being used to enhance network security by analyzing network traffic and identifying suspicious activities. By processing logs and traffic data, LLMs can detect anomalies that may indicate an ongoing cyberattack (Howard & Borenstein, 2021). This proactive approach allows security teams to respond to threats before they escalate.

4.4. Phishing Detection

Phishing attacks remain one of the most prevalent cybersecurity threats. LLMs can analyze the content of emails and messages to detect signs of phishing, such as suspicious language patterns or links (Zhang et al., 2021). By integrating LLMs into email security systems, organizations can significantly reduce the risk of phishing attacks.

4.5. Social Engineering and Fraud Detection

Social engineering attacks, which exploit human psychology to gain unauthorized access to information, are particularly challenging to detect. LLMs have been applied to analyze communication patterns and identify attempts at social engineering (Kim et al., 2020). Similarly, LLMs are used in fraud detection systems to identify unusual or suspicious activities, such as unauthorized transactions.

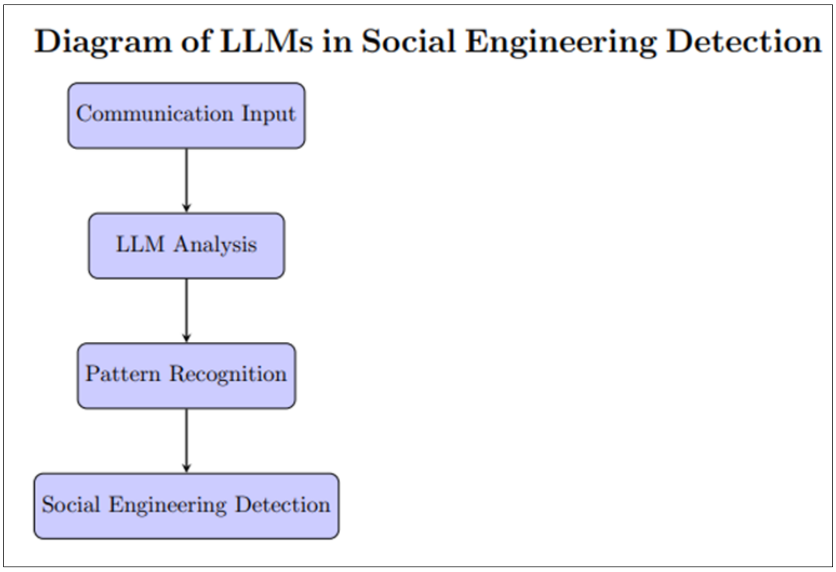


Diagram of LLMs in Social Engineering Detection
Communication Input
LLM Analysis
Pattern Recognition
Social Engineering Detection

Figure 4 Diagram of LLMs in Social Engineering Detection

A diagram showing how LLMs can be integrated into communication systems to detect social engineering attempts can be shown here.

5. Challenges and Limitations

While LLMs offer significant advantages in cybersecurity, they also present several challenges and limitations that must be addressed.

5.1. Data Limitations

One of the primary challenges in applying LLMs to cybersecurity is the availability and quality of training data. Cybersecurity data, such as logs and threat reports, often contain sensitive information, making it difficult to obtain large datasets for training LLMs (Goodfellow et al., 2014). Furthermore, the data may be noisy, incomplete, or biased, leading to inaccurate or unreliable model outputs.

5.2. Interpretability and Explainability

LLMs, like many deep learning models, are often considered "black boxes" due to their complex internal structures. This lack of transparency can be problematic in cybersecurity, where understanding the rationale behind a model's decision is critical (Lipton, 2018). Efforts are being made to improve the interpretability and explainability of LLMs, but this remains an ongoing challenge.

5.3. Ethical and Privacy Concerns

The deployment of LLMs in cybersecurity raises ethical and privacy concerns. For instance, LLMs trained on sensitive data may inadvertently leak confidential information (Shokri et al., 2017). Additionally, the use of LLMs in surveillance and monitoring systems can lead to privacy violations if not properly regulated.

5.4. Computational Resources

Training and deploying LLMs require significant computational resources, which can be a barrier for smaller organizations. The energy consumption associated with large-scale LLMs is also a growing concern, as it contributes to the environmental impact of AI technologies (Strubell et al., 2019).

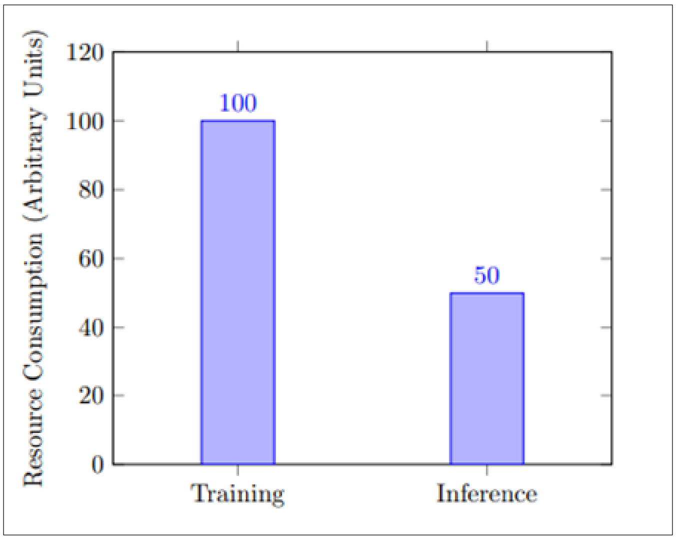


Figure 5 Computational Resource Requirements for LLMs

A chart illustrating the computational resources required for training and deploying LLMs can be provided here.

6. Future Directions

As LLMs continue to evolve, there are several promising directions for future research in the field of cybersecurity.

6.1. Advancements in Model Architecture

Future research could focus on developing more efficient and specialized LLM architectures for cybersecurity tasks. This includes exploring techniques such as model pruning, quantization, and knowledge distillation to reduce the computational footprint of LLMs without sacrificing performance (Sanh et al., 2019).

6.2. Integration with Other Technologies

Integrating LLMs with other AI and cybersecurity technologies, such as reinforcement learning and blockchain, could lead to more robust and adaptive security solutions. For example, combining LLMs with reinforcement learning could enable dynamic threat response systems that learn and adapt to new attack vectors in real-time (Silver et al., 2016).

6.3. Addressing Current Challenges

Ongoing research should continue to address the challenges and limitations of LLMs in cybersecurity, particularly in the areas of data privacy, interpretability, and computational efficiency. Developing standardized frameworks and best practices for the ethical deployment of LLMs will be crucial in ensuring their safe and effective use in cybersecurity (Floridi et al., 2018).

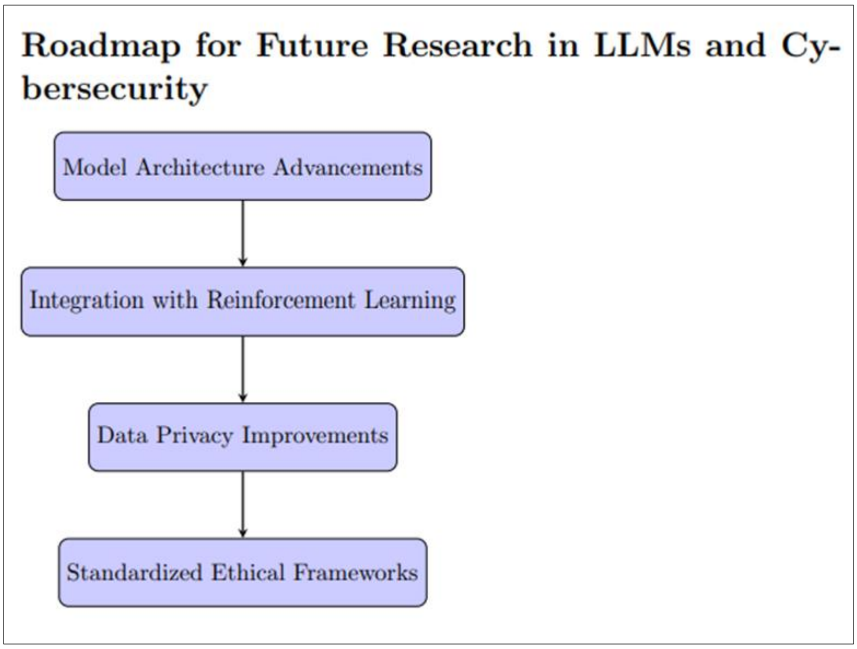

Roadmap for Future Research in LLMs and Cybersecurity
Model Architecture Advancements
Integration with Reinforcement Learning
Data Privacy Improvements
Standardized Ethical Frameworks

Figure 6 Roadmap for Future Research in LLMs and Cybersecurity

A roadmap outlining the key areas for future research in LLMs and cybersecurity can be shown here.

7. Conclusion

This paper has provided a comprehensive review of the application of Large Language Models (LLMs) in cybersecurity. LLMs have demonstrated significant potential in enhancing various aspects of cybersecurity, from vulnerability detection to phishing prevention. However, their deployment also presents challenges, particularly concerning data quality, interpretability, and ethical considerations.

The future of LLMs in cybersecurity is promising, with ongoing research focused on addressing current limitations and exploring new applications. As LLMs continue to evolve, they are likely to play an increasingly important role in securing digital systems against emerging threats.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Wright, J., Dawson Jr, M. E., & Omar, M. (2012). Cyber security and mobile threats: The need for antivirus applications for smartphones. Journal of Information Systems Technology and Planning, 5(14), 40-60.

[2] Omar, M., & Dawson, M. (2013). Research in progress-defending android smartphones from malware attacks. In 2013 third international conference on advanced computing and communication technologies (ACCT) (pp. 288-292). IEEE.

[3] Dawson, M., Al Saeed, I., Wright, J., & Omar, M. (2013). Technology enhanced learning with open source software for scientists and engineers. In INTED2013 Proceedings (pp. 5583-5589). IATED.

[4] Omar, M. (2012). Smartphone Security: Defending Android-based Smartphone Against Emerging Malware Attacks (Doctoral dissertation, Colorado Technical University).

[5] Dawson, M., Omar, M., & Abramson, J. (2015). Understanding the methods behind cyber terrorism. In Encyclopedia of Information Science and Technology, Third Edition (pp. 1539-1549). IGI Global.

[6] Fawzi, D., & Omar, M. (n.d.). New insights to database security: An effective and integrated approach to applying access control mechanisms and cryptographic concepts in Microsoft access environments. Academic Press.

[7] Dawson, M., Omar, M., Abramson, J., & Bessette, D. (2014). The future of national and international security on the internet. In Information security in diverse computing environments (pp. 149-178). IGI Global.

[8] Omar, M. (2015). Insider threats: Detecting and controlling malicious insiders. In New Threats and Countermeasures in Digital Crime and Cyber Terrorism (pp. 162-172). IGI Global.

[9] Dawson, M., Wright, J., & Omar, M. (2015). Mobile devices: The case for cyber security hardened systems. In New Threats and Countermeasures in Digital Crime and Cyber Terrorism (pp. 8-29). IGI Global.

[10] Dawson, M. (2015). A brief review of new threats and countermeasures in digital crime and cyber terrorism. New Threats and Countermeasures in Digital Crime and Cyber Terrorism, 1-7. IGI Global.

[11] Hamza, Y. A., & Omar, M. D. (2013). Cloud computing security: abuse and nefarious use of cloud computing. International Journal of Computer Engineering Research, 3(6), 22-27.

[12] Omar, M. (2015). Cloud Computing Security: Abuse and Nefarious Use of Cloud Computing. In Handbook of Research on Security Considerations in Cloud Computing (pp. 30-38). IGI Global.

[13] Davis, L., Dawson, M., & Omar, M. (2016). Systems Engineering Concepts with Aid of Virtual Worlds and Open Source Software: Using Technology to Develop Learning Objects and Simulation Environments. In Handbook of Research on 3-D Virtual Environments and Hypermedia for Ubiquitous Learning (pp. 483-509). IGI Global.

[14] Mohammed, D., Omar, M., & Nguyen, V. (2017). Enhancing Cyber Security for Financial Industry through Compliance and Regulatory Standards. In Security Solutions for Hyperconnectivity and the Internet of Things (pp. 113-129). IGI Global.

[15] Dawson, M., Omar, M., Abramson, J., Leonard, B., & Bessette, D. (2017). Battlefield cyberspace: Exploitation of hyperconnectivity and internet of things. In Developing Next-Generation Countermeasures for Homeland Security Threat Prevention (pp. 204-235). IGI Global.

[16] Nguyen, V., Omar, M., & Mohammed, D. (2017). A Security Framework for Enhancing User Experience. International Journal of Hyperconnectivity and the Internet of Things (IJHIoT), 1(1), 19-28. IGI Global.

[17] Omar, M., Mohammed, D., & Nguyen, V. (2017). Defending against malicious insiders: a conceptual framework for predicting, detecting, and deterring malicious insiders. International Journal of Business Process Integration and Management, 8(2), 114-119. Inderscience Publishers (IEL).

[18] Dawson, M., Eltayeb, M., & Omar, M. (2016). Security solutions for hyperconnectivity and the Internet of things. IGI Global.

[19] Omar, M. (n.d.). Latina Davis Morgan State University 1700 E Cold Spring Ln. Baltimore, MD 21251, USA E-mail: latinaedavis@ hotmail. com.

[20] Dawson, M., Davis, L., & Omar, M. (2019). Developing learning objects for engineering and science fields: using technology to test system usability and interface design. International Journal of Smart Technology and Learning, 1(2), 140-161. Inderscience Publishers (IEL).

[21] Banisakher, M., Mohammed, D., & Omar, M. (2018). A Cloud-Based Computing Architecture Model of Post-Disaster Management System. International Journal of Simulation--Systems, Science & Technology, 19(5).

[22] Nguyen, V., Mohammed, D., Omar, M., & Banisakher, M. (2018). The Effects of the FCC Net Neutrality Repeal on Security and Privacy. International Journal of Hyperconnectivity and the Internet of Things (IJHIoT), 2(2), 21-29. IGI Global.

[23] Omar, M., Mohammed, D., Nguyen, V., Dawson, M., & Banisakher, M. (2021). Android application security. In Research Anthology on Securing Mobile Technologies and Applications (pp. 610-625). IGI Global.

[24] Banisakher, M., Omar, M., & Clare, W. (2019). Critical Infrastructure-Perspectives on the Role of Government in Cybersecurity. Journal of Computer Sciences and Applications, 7(1), 37-42.

[25] Omar, M. (2019). A world of cyber attacks (a survey).

[26] Mohammed, D., Omar, M., & Nguyen, V. (2018). Wireless sensor network security: approaches to detecting and avoiding wormhole attacks. Journal of Research in Business, Economics and Management, 10(2), 1860-1864.

[27] Banisakher, M., Omar, M., Hong, S., & Adams, J. (2020). A human centric approach to data fusion in post-disaster management. Journal of Business Management and Science, 8(1), 12-20.

[28] Nguyen, V., Mohammed, D., Omar, M., & Dean, P. (2020). Net neutrality around the globe: A survey. In 2020 3rd International Conference on Information and Computer Technologies (ICICT) (pp. 480-488). IEEE.

[29] Dawson, M., Omar, M., Abramson, J., & Bessette, D. (2014). Information security in diverse computing environments. Academic Press.

[30] Zangana, H. M., & Omar, M. (2020). Threats, Attacks, and Mitigations of Smartphone Security. Academic Journal of Nawroz University, 9(4), 324-332.

[31] Omar, M. (2021). New insights into database security: An effective and integrated approach for applying access control mechanisms and cryptographic concepts in Microsoft Access environments.

[32] Omar, M. (2021). Developing Cybersecurity Education Capabilities at Iraqi Universities.

[33] Omar, M., Gouveia, L. B., Al-Karaki, J., & Mohammed, D. (2022). Reverse-Engineering Malware. In Cybersecurity Capabilities in Developing Nations and Its Impact on Global Security (pp. 194-217). IGI Global.

[34] Omar, M., Choi, S., Nyang, D., & Mohaisen, D. (2022). Robust natural language processing: Recent advances, challenges, and future directions. IEEE Access, 10, 86038-86056. IEEE.

[35] Al Kinoon, M., Omar, M., Mohaisen, M., & Mohaisen, D. (2021). Security breaches in the healthcare domain: a spatiotemporal analysis. In Computational Data and Social Networks: 10th International Conference, CSoNet 2021, Virtual Event, November 15-17, 2021, Proceedings (pp. 171-183). Springer International Publishing.

[36] Omar, M. (2022). Machine Learning for Cybersecurity: Innovative Deep Learning Solutions. Springer Brief. https://link.springer.com/book/978303115

[37] Omar, M. (n.d.). Defending Cyber Systems through Reverse Engineering of Criminal Malware. Springer Brief. https://link.springer.com/book/9783031116278

[38] Omar, M., Choi, S., Nyang, D., & Mohaisen, D. (2022). Quantifying the performance of adversarial training on language models with distribution shifts. In Proceedings of the 1st Workshop on Cybersecurity and Social Sciences (pp. 3-9).

[39] Omar, M., & Mohaisen, D. (2022). Making Adversarially-Trained Language Models Forget with Model Retraining: A Case Study on Hate Speech Detection. In Companion Proceedings of the Web Conference 2022 (pp. 887-893).

[40] Omar, M. (2022). Malware anomaly detection using local outlier factor technique. In Machine Learning for Cybersecurity: Innovative Deep Learning Solutions (pp. 37-48). Springer International Publishing Cham.

[41] Omar, M. (2022). Application of machine learning (ML) to address cybersecurity threats. In Machine Learning for Cybersecurity: Innovative Deep Learning Solutions (pp. 1-11). Springer International Publishing Cham.

[42] Burrell, D. N., Nobles, C., Cusak, A., Omar, M., & Gillesania, L. (2022). Cybercrime and the Nature of Insider Threat Complexities in Healthcare and Biotechnology Engineering Organizations. Journal of Crime and Criminal Behavior, 2(2), 131-144.

[43] Omar, M., & Sukthankar, G. (2023). Text-defend: detecting adversarial examples using local outlier factor. In 2023 IEEE 17th international conference on semantic computing (ICSC) (pp. 118-122). IEEE.

[44] Ahmed, A., Rasheed, H., Bashir, A. K., & Omar, M. (2023). Millimeter-wave channel modeling in a VANETs using coding techniques. PeerJ Computer Science, 9, e1374. PeerJ Inc.

[45] Omar, M. (2023). VulDefend: A Novel Technique based on Pattern-exploiting Training for Detecting Software Vulnerabilities Using Language Models. In 2023 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT) (pp. 287-293). IEEE.

[46] Abbasi, R., Bashir, A. K., Mateen, A., Amin, F., Ge, Y., & Omar, M. (2023). Efficient Security and Privacy of Lossless Secure Communication for Sensor-based Urban Cities. IEEE Sensors Journal. IEEE.

[47] Omar, M., & Shiaeles, S. (2023). VulDetect: A novel technique for detecting software vulnerabilities using Language Models. In 2023 IEEE International Conference on Cyber Security and Resilience (CSR). IEEE. https://ieeexplore.ieee.org/document/10224924

[48] Gholami, S. (2024). Do Generative large language models need billions of parameters? In Redefining Security With Cyber AI (pp. 37-55). IGI Global.

[49] Burrell, D. N., Nobles, C., Richardson, K., Wright, J. B., Jones, A. J., Springs, D., & Brown-Jackson, K. (2023). Allison Huff. Applied Research Approaches to Technology, Healthcare, and Business, 1. IGI Global.

[50] Huff, A. J., Burrell, D. N., Nobles, C., Richardson, K., Wright, J. B., Burton, S. L., Jones, A. J., Springs, D., Omar, M., & Brown-Jackson, K. L. (2023). Management Practices for Mitigating Cybersecurity Threats to Biotechnology Companies, Laboratories, and Healthcare Research Organizations. In Applied Research Approaches to Technology, Healthcare, and Business (pp. 1-12). IGI Global.

[51] Gholami, S., & Omar, M. (2024). Can a student large language model perform as well as its teacher? In Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology (pp. 122-139). IGI Global.

[52] Ayub, M. F., Li, X., Mahmood, K., Shamshad, S., Saleem, M. A., & Omar, M. (2023). Secure consumer-centric demand response management in resilient smart grid as industry 5.0 application with blockchain-based authentication. IEEE Transactions on Consumer Electronics. IEEE.

[53] Gholami, S. (2024). Can pruning make large language models more efficient? In Redefining Security With Cyber AI (pp. 1-14). IGI Global.

[54] Zhou, S., Ali, A., Al-Fuqaha, A., Omar, M., & Feng, L. (n.d.). Robust Risk-Sensitive Task Offloading for Edge-Enabled Industrial Internet of Things. IEEE Transactions on Consumer Electronics.

[55] Gholami, S., & Omar, M. (2023). Does Synthetic Data Make Large Language Models More Efficient? arXiv preprint arXiv:2310.07830.

[56] Tiwari, N., Omar, M., & Ghadi, Y. (2023). Brain Tumor Classification From Magnetic Resonance Imaging Using Deep Learning and Novel Data Augmentation. In Transformational Interventions for Business, Technology, and Healthcare (pp. 392-413). IGI Global.

[57] Tiwari, N., Ghadi, Y., & Omar, M. (2023). Analysis of Ultrasound Images in Kidney Failure Diagnosis Using Deep Learning. In Transformational Interventions for Business, Technology, and Healthcare (pp. 45-74). IGI Global.

[58] Omar, M., Jones, R., Burrell, D. N., Dawson, M., Nobles, C., & Mohammed, D. (2023). Harnessing the power and simplicity of decision trees to detect IoT Malware. In Transformational Interventions for Business, Technology, and Healthcare (pp. 215-229). IGI Global.

[59] Saleem, M. A., Li, X., Mahmood, K., Shamshad, S., Ayub, M. F., & Omar, M. (2023). Provably secure conditional-privacy access control protocol for intelligent customers-centric communication in vanet. IEEE Transactions on Consumer Electronics. IEEE.

[60] Omar, M., & Burrell, D. (2023). From text to threats: A language model approach to software vulnerability detection. International Journal of Mathematics and Computer in Engineering.

[61] Arulappan, A., Raja, G., Bashir, A. K., Mahanti, A., & Omar, M. (2023). ZTMP: Zero Touch Management Provisioning Algorithm for the On-boarding of Cloud-native Virtual Network Functions. Mobile Networks and Applications, 1-13. Springer US New York.

[62] Al-Karaki, J. N., Omar, M., Gawanmeh, A., & Jones, A. (2023). Advancing CyberSecurity Education and Training: Practical Case Study of Running Capture the Flag (CTF) on the Metaverse vs. Physical Settings. In 2023 International Conference on Intelligent Metaverse Technologies & Applications (iMETA) (pp. 1-7). IEEE.

[63] Al Harthi, A. S., Al Balushi, M. Y., Al Badi, A. H., Al Karaki, J., & Omar, M. (n.d.). Metaverse Adoption in UAE Higher Education: A Hybrid SEM-ANN Approach.......... 98 Mohammad Daradkeh, Boshra Aldhanhani, Amjad Gawanmeh, Shadi Atalla and Sami Miniaoui. Applied Research Approaches to Technology, Healthcare, and Business, 1.

[64] Umer, M., Aljrees, T., Karamti, H., Ishaq, A., Alsubai, S., Omar, M., Bashir, A. K., & Ashraf, I. (2023). Heart failure patients monitoring using IoT-based remote monitoring system. Scientific Reports, 13(1), 19213. Nature Publishing Group UK London.

[65] Ahmed, N., Mohammadani, K., Bashir, A. K., Omar, M., Jones, A., & Hassan, F. (2024). Secure and Reliable Routing in the Internet of Vehicles Network: AODV-RL with BHA Attack Defense. CMES-Computer Modeling in Engineering & Sciences, 139(1).

[66] Dayoub, A., & Omar, M. (2024). Advancing IoT Security Posture K-Means Clustering for Malware Detection. In Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology (pp. 221-239). IGI Global.

[67] Omar, M. (2024). Revolutionizing Malware Detection: A Paradigm Shift Through Optimized Convolutional Neural Networks. In Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology (pp. 196-220). IGI Global.

[68] Mohammed, D., & Omar, M. (2024). Decision Trees Unleashed: Simplifying IoT Malware Detection With Advanced AI Techniques. In Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology (pp. 240-258). IGI Global.

[69] Basharat, M., & Omar, M. (2024). Adapting to Change: Assessing the Longevity and Resilience of Adversarially Trained NLP Models in Dynamic Spam Detection Environments. In Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology (pp. 157-173). IGI Global.

[70] Omar, M. (2024). From Attack to Defense: Strengthening DNN Text Classification Against Adversarial Examples. In Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology (pp. 174-195). IGI Global.

[71] Jabbari, A., Khan, H., Duraibi, S., Budhiraja, I., Gupta, S., & Omar, M. (2024). Energy Maximization for Wireless Powered Communication Enabled IoT Devices With NOMA Underlaying Solar Powered UAV Using Federated Reinforcement Learning for 6G Networks. IEEE Transactions on Consumer Electronics. IEEE.

[72] Xu, X., Wu, J., Bashir, A. K., & Omar, M. (2024). Machine Learning and Zero Knowledge Empowered Trustworthy Bitcoin Mixing for Next-G Consumer Electronics Payment. IEEE Transactions on Consumer Electronics. IEEE.

[73] Omar, M., & Burrell, D. N. (2024). Organizational Dynamics and Bias in Artificial Intelligence (AI) Recruitment Algorithms. In Evolution of Cross-Sector Cyber Intelligent Markets (pp. 269-290). IGI Global.

[74] Jun, W., Iqbal, M. S., Abbasi, R., Omar, M., & Huiqin, C. (2024). Web-Semantic-Driven Machine Learning and Blockchain for Transformative Change in the Future of Physical Education. International Journal on Semantic Web and Information Systems (IJSWIS), 20(1), 1-16. IGI Global.

[75] Jones, R., & Omar, M. (2024). Codeguard: Utilizing Advanced Pattern Recognition in Language Models for Software Vulnerability Analysis. Land Forces Academy Review, 29(1), 108-118.

[76] Omar, M., Bauer, R., Fernando, A., Darejeh, A., Rahman, S., Ulusoy, S. K., Arabo, A., Gupta, R., Adedoyin, F., Paul, R. K., & others. (2024). Committee Members. In Journal of Physics: Conference Series, 2711, 011001.

[77] Basharat, M., & Omar, M. (2024). Harnessing GPT-2 for Feature Extraction in Malware Detection: A Novel Approach to Cybersecurity. Land Forces Academy Review, 29(1), 74-84.

[78] Jones, A., & Omar, M. (2024). Codesentry: Revolutionizing Real-Time Software Vulnerability Detection With Optimized GPT Framework. Land Forces Academy Review, 29(1), 98-107.

[79] Rajesh, R., Hemalatha, S., Nagarajan, S. M., Devarajan, G. G., Omar, M., & Bashir, A. K. (2024). Threat Detection and Mitigation for Tactile Internet Driven Consumer IoT-Healthcare System. IEEE Transactions on Consumer Electronics. IEEE.

[80] Sun, Y., Xu, T., Bashir, A. K., Liu, J., & Omar, M. (2023). BcIIS: Blockchain-Based Intelligent Identification Scheme of Massive IoT Devices. In GLOBECOM 2023-2023 IEEE Global Communications Conference (pp. 1277-1282). IEEE.

[81] Dong, H., Wu, J., Bashir, A. K., Pan, Q., Omar, M., & Al-Dulaimi, A. (2023). Privacy-Preserving EEG Signal Analysis with Electrode Attention for Depression Diagnosis: Joint FHE and CNN Approach. In GLOBECOM 2023-2023 IEEE Global Communications Conference (pp. 4265-4270). IEEE.

[82] Zhang, H., Wu, J., Pan, Q., Bashir, A. K., & Omar, M. (2024). Toward Byzantine-Robust Distributed Learning for Sentiment Classification on Social Media Platform. IEEE Transactions on Computational Social Systems. IEEE.

[83] Alturki, N., Altamimi, A., Umer, M., Saidani, O., Alshardan, A., Alsubai, S., Omar, M., & Ashraf, I. (2024). Improving Prediction of Chronic Kidney Disease Using KNN Imputed SMOTE Features and TrioNet Model. CMES-Computer Modeling in Engineering & Sciences, 139(3).

[84] Jones, A., & Omar, M. (2023). Optimized Decision Trees to Detect IoT Malware. In 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE) (pp. 1761-1765). IEEE.

[85] Jones, B. M., & Omar, M. (2023). Studying the Effects of Social Media Content on Kids' Safety and Well-being. In 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE) (pp. 1876-1879). IEEE.

[86] Jones, R., Omar, M., Mohammed, D., & Nobles, C. (2023). IoT Malware Detection with GPT Models. In 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE) (pp. 1749-1752). IEEE.

[87] Omar, M. (n.d.). Machine Learning for Cybersecurity.

[88] Jones, B. M., & Omar, M. (2023). Detection of Twitter Spam with Language Models: A Case Study on How to Use BERT to Protect Children from Spam on Twitter. In 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE) (pp. 511-516). IEEE.

[89] Jones, R., Omar, M., Mohammed, D., Nobles, C., & Dawson, M. (2023). Harnessing the Speed and Accuracy of Machine Learning to Advance Cybersecurity. In 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE) (pp. 418-421). IEEE.

[90] Jones, A., & Omar, M. (2023). Harnessing the Efficiency of Reformers to Detect Software Vulnerabilities. In 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE) (pp. 2259-2264). IEEE.

[91] Jones, R., Omar, M., & Mohammed, D. (2023). Harnessing the Power of the GPT Model to Generate Adversarial Examples. In 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE) (pp. 1699-1702). IEEE.

[92] Jones, R., & Omar, M. (2023). Detecting IoT Malware with Knowledge Distillation Technique. In 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE) (pp. 131-135). IEEE.

[93] Jones, B. M., & Omar, M. (2023). Measuring the Impact of Global Health Emergencies on Self-Disclosure Using Language Models. In 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE) (pp. 1806-1810). IEEE.

[94] Jones, R., & Omar, M. (2024). Revolutionizing Cybersecurity: The GPT-2 Enhanced Attack Detection and Defense (GEADD) Method for Zero-Day Threats. International Journal of Informatics, Information System and Computer Engineering (INJIISCOM), 5(2), 178-191.

[95] Zangana, H. M., Sallow, Z. B., Alkawaz, M. H., & Omar, M. (2024). Unveiling the Collective Wisdom: A Review of Swarm Intelligence in Problem Solving and Optimization. Inform: Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi, 9(2), 101-110.

Zangana, H. M., Tawfiq, N. E., & Omar, M. (2020). Advantages and Challenges of E-Government in Turkey.

[96] Basharat, M., & Omar, M. (n.d.). SecuGuard: Leveraging pattern-exploiting training in language models for advanced software vulnerability detection. International Journal of Mathematics and Computer in Engineering.

[97] Kumar, V. A., Surapaneni, S., Pavitra, D., Venkatesan, R., Omar, M., & Bashir, A. K. (2024). An Internet of Medical Things-Based Mental Disorder Prediction System Using EEG Sensor and Big Data Mining. Journal of Circuits, Systems and Computers, 2450197. World Scientific Publishing Company.

[98] Peng, Y., Wang, J., Ye, X., Khan, F., Bashir, A. K., Alshawi, B., Liu, L., & Omar, M. (2024). An intelligent resource allocation strategy with slicing and auction for private edge cloud systems. Future Generation Computer Systems, 160, 879-889. North-Holland.

[99] Pauu, K. T., Pan, Q., Wu, J., Bashir, A. K., & Omar, M. (2024). IRS-Aided Federated Learning with Dynamic Differential Privacy for UAVs in Emergency Response. IEEE Internet of Things Magazine, 7(4), 108-115. IEEE.

[100] Zangana, H. M., Omar, M., Al-Karaki, J. N., & Mohammed, D. (2024). Comprehensive Review and Analysis of Network Firewall Rule Analyzers: Enhancing Security Posture and Efficiency. In Redefining Security With Cyber AI (pp. 15-36). IGI Global.

[101] Tao, Y., Wu, J., Pan, Q., Bashir, A. K., & Omar, M. (2024). O-RAN-Based Digital Twin Function Virtualization for Sustainable IoV Service Response: An Asynchronous Hierarchical Reinforcement Learning Approach. IEEE Transactions on Green Communications and Networking. IEEE.