International Journal of Safety and Security Engineering
Vol. 14, No. 6, December, 2024, pp. 1817-1824
Journal homepage: http://iieta.org/journals/ijsse

A Systematic Mapping Study on Security Threats and Solutions in Social Media Environments

Yasir Ali Matnee[1*], Saad Albawi[2]

[1] Basic Education College, University of Diyala, Baqubah 32001, Iraq
[2] College of Engineering, University of Diyala, Baqubah 32001, Iraq

Corresponding Author Email: yaserali@uodiyala.edu.iq

https://doi.org/10.18280/ijsse.140616

Received: 9 October 2024
Revised: 11 November 2024
Accepted: 20 November 2024
Available online: 31 December 2024

Keywords:
social media security, artificial intelligence, machine learning, deep learning, threat detection, data protection, privacy in social networks, security in social media

ABSTRACT

In the digital era, social media has become essential to users' daily lives, providing platforms such as Facebook, Twitter, Instagram, and others with unprecedented opportunities for interaction and information exchange. However, these platforms expose users to many security threats, such as identity theft, phishing, data breaches, and unauthorized access to personal accounts. This study systematically reviews social media security threats and highlights AI-based solutions to mitigate these risks. After conducting extensive research in scientific databases such as (Institute of Electrical and Electronics Engineers) Explore, (Association for Computing Machinery) Digital Library, Science Direct, and others, more than 1000 relevant research papers were found. After applying specific filtering criteria, the number of studies was reduced to 160 and then to 42 papers analyzed in depth. The research uses AI approaches, including machine learning, deep learning, natural language, etc., to detect anomalous behaviors and malicious intent in social media interactions. The findings provide a broad overview of the different research areas and their interrelationships. It also highlights areas that may benefit from increased research attention to enhance the security of this platform. The research found that early detection conserves and adds security benefits, and AI was important for this effort—phishing, data breaches, etc. The current study laid the foundation for future related studies focusing on security models and how they can be improved, emphasizing the need to strike a balance between protecting privacy and ensuring security.

1. INTRODUCTION

Facebook, Twitter, Instagram, and others are examples of these instruments that have become necessary for social and professional communication in the digital era. While these platforms offer a chance for global communication and spreading information on a scale never before seen, they also pose massive security risks. These are sly and malicious cybercriminals that use the weak points of these platforms to have their wicked ways, essentially targeting personal data, including but not limited to identity theft and phishing attacks, amongst other net malpractices [1, 2].

Several studies indicate that social media platforms are fertile environments for security threats like malware, unauthorized account hacking, and data breaches. These threats require a deep understanding and careful analysis, as they are constantly evolving as the platforms themselves evolve. Recent research papers have demonstrated the importance and role of artificial intelligence in detecting and addressing these threats through technologies such as machine learning, deep learning, and natural language processing. These technologies enable the detection of suspicious methods and malicious intentions from user interactions [3, 4].

Artificial intelligence and its technologies provide great potential in detecting security threats in social media programs. However, the integration of these technologies with each other faces many challenges. With the spread and expansion of these platforms (social media platforms), it has become necessary for information security specialists to keep pace with this development and these new threats. On the other hand, monitoring users using artificial intelligence and its technologies raises many concerns, especially about the privacy of user information [5]. Therefore, it has become necessary to balance strengthening and enhancing security measures and protecting users' privacy [6].

Events support the above concerns, such as the Cambridge Analytica scandal, in which information from millions of Facebook users was exploited without the users' knowledge. Due to such scandals, it is necessary to shed light on data protection mechanisms and develop protection methods considering users' privacy [7].

In this research, we seek to analyze the methods used to understand the state of social media security by reviewing and reviewing studies that have addressed the topics of data security for social media, as well as providing solutions based on artificial intelligence to address such threats. Specifically, this study addresses the following research questions:

•What are the key security threats in social media

environments, and how can AI techniques mitigate these threats?

•What mechanisms and methodologies are employed to enhance social media security using AI?

•How effective are current AI-based methods in addressing large-scale security challenges in social media?

•What gaps exist in the application of AI to social media security, and what future directions should research take?

By systematically exploring these questions, the study aims to contribute to developing comprehensive and adaptive security frameworks that leverage AI technologies.

1.1 Search string map

1.1.1 Identifying keywords

It is very important to identify keywords relevant to the topic's core to conduct an effective and comprehensive review. By searching with these keywords, our research strategy will be guided to identify the most relevant sources. Here are some of the keywords we used to support our research:

1.1.2 Search string

It is self-evident that when searching for the titles of studies related to our topic, it is necessary to find a series of keywords to complete the search. Formulating the keywords is very important to obtain good results. Combining these search phrases will help create an effective search series that can provide useful and comprehensive research results [8, 9].

("Social media security" OR "Online platform security" OR "Social networks security") AND ("AI techniques" OR "Deep learning" OR "Machine learning") AND ("Threat analysis" OR "Threat detection" OR "Cyber threat assessment") AND ("Data protection" OR "Privacy in social networks" OR "Information security")

("Cybersecurity in social media" OR "Threat mitigation in online platforms") AND ("Artificial intelligence solutions" OR "Neural networks applications") AND ("Vulnerability management" OR "Big data handling")

("Identity theft prevention" AND "Phishing detection") AND ("Deep learning models" OR "Anomaly detection systems") AND ("User privacy protection" AND "Risk management strategies")

This search string can be applied across various academic databases and search engines to locate studies that align with the research objectives.

Figure 1 highlights the main steps of our systematic mapping process to identify studies closely related to our research area.

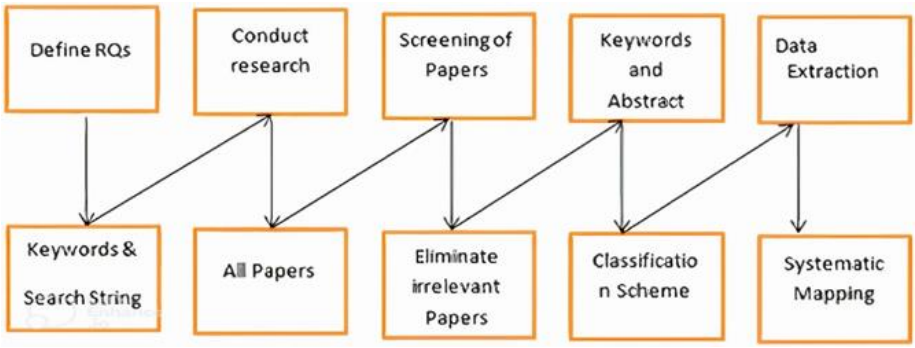

Figure 1. The main steps of the systematic mapping study

2. LITERATURE REVIEW

2.1 Database search methodology

Various databases offer valuable information for analysis in the research domain, and these can be categorized based on the search methods used: automated and manual searches. Researchers typically utilize well-known databases such as the ACM Digital Library, IEEE Xplore Digital Library, ScienceDirect, SCOPUS, and Engineering Village [10]. Additionally, resources like DBLP, academic journals, and conference proceedings also contribute significantly to the research landscape [11, 12].

Scope definition: The scope was defined by identifying key research questions and selecting relevant databases (e.g., IEEE Xplore, ACM Digital Library, ScienceDirect) for literature retrieval [13-15].

Keyword and search string development: Optimized search strings were formulated to ensure comprehensive coverage of relevant literature. These strings incorporated specific terms related to social media security, AI techniques, and threat analysis [16].

Paper collection and filtering: Over 1000 papers were initially retrieved from the databases. A multi-stage filtering process was applied.

Data extraction and classification: Extracted data included publication details, methodologies, AI techniques, and outcomes [17, 18]. Classification schemes were applied to group studies into categories such as research types (e.g., evaluation research, solution proposals) and application areas (e.g., threat detection, data protection) [19-21].

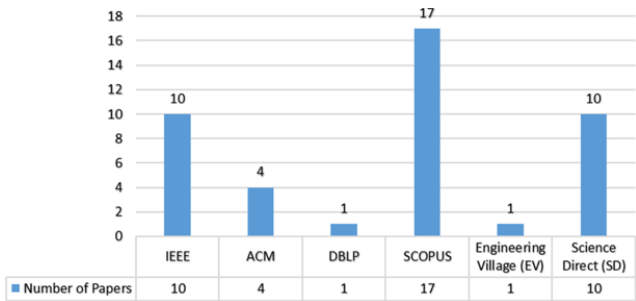

Figure 2. Distribution of papers from sources

Data analysis: Trends and patterns in the literature were analyzed to identify gaps and opportunities. Visual tools like charts and diagrams were used to illustrate relationships

between research domains and methodologies [22-25].

In our study, we conducted a systematic search using these databases. Figure 2 summarizes the distribution of research papers obtained from these sources. This distribution illustrates the prevalence of research on threat detection in social media across different databases, as shown in Table 1.

Table 1. Publication venues

Database	Number of Papers
IEEE	10
ACM	4
DBLP	1
SCOPUS	17
Engineering Village (EV)	1
Science Direct (SD)	10

2.2 Screening of papers

Screening papers is essential to ensure the inclusion of only the most relevant and high-quality studies in this research. This involves applying two key criteria to evaluate and select studies based on the research questions formulated for this study. The criteria for paper selection are outlined.

2.3 Inclusion criteria

Relevance to research questions: Papers must address issues related to security threats, threat detection, or data protection within social media environments [26, 27].

Application of AI techniques: Studies should use advanced AI techniques, such as deep learning, machine learning, or natural language processing, to address and analyze social media security concerns [28-30].

Comprehensive coverage: Selected research papers should provide valuable theoretical or applied insights that clearly and directly advance the study objectives [31].

2.4 Exclusion criteria

Irrelevance: Excluded sources that do not directly address social media or those that do not address the specific research questions in the study [32].

Lack of AI technologies: All studies that do not use or discuss relevant AI technologies within the SSA were excluded [33].

Language and completeness: Documents that will not be published in English, are not yet complete, or are only available in abstract form will be excluded [34].

Quality of insight: Posts that do not reach important insights or new findings relevant to the field will be excluded.

2.5 Analyzing research through various perspectives

A particular topic often requires a lot of multiple aspects. This helps to provide a more diverse view of the available literature. In this study, we will explore the main theme, the purpose of providing meals, and multiple aspects of the research findings [35-40]:

2.5.1 Distribution of primary studies by year

Table 2 shows the history of research over the years through the following number of studies published each year.

Figure 3 distinguishes between research papers using different colors, which helps to identify landmarks and changes in research focus. This graph provides important insights into the field and the current state of research. Its contribution to studies can be followed in detail year by year.

Table 2. Distributed the research by years

Publication Year	Number of Papers
2017	1
2018	6
2019	5
2020	3
2021	4
2023	14
2024	9

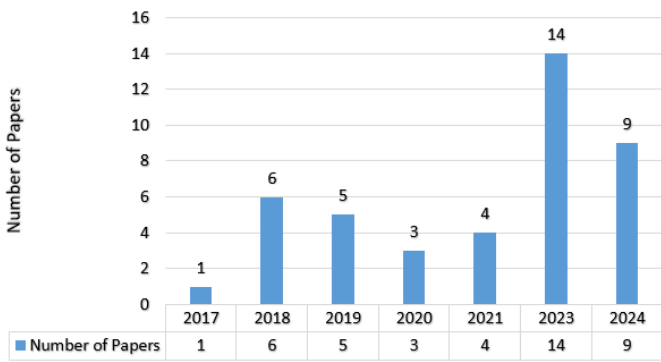

Figure 3. Distribution of research over the years

2.5.2 Venue chart

As shown in Table 3, the venue chart, illustrated in Figure 4, categorizes papers based on where the researches were published, such as conferences, workshops, and journals. It provides a breakdown of publication venues, including the number of papers, publication years, and the length of the papers (short or full) [41-45]. This chart assists in understanding the distribution of research output across different platforms and identifies which venues are most prominent in advancing the field. Figure 4 provides a comprehensive view of the publication venues.

Table 3. Distribution of papers per venue

Category	Number of Papers
Conferences	11
Journals	28
Workshop	3

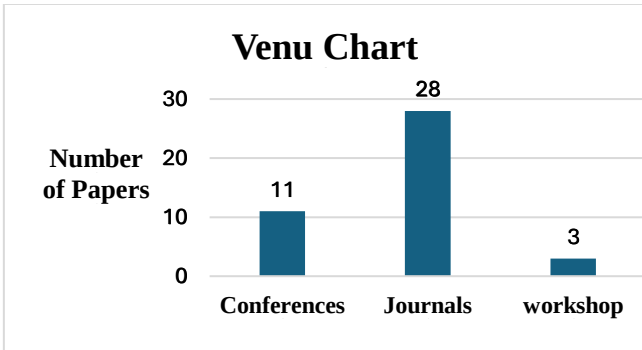

Figure 4. Publication venues

3. CLASSIFICATION SCHEMES FOR DATA EXTRACTION

-Extracted data included publication details, methodologies, AI techniques, and outcomes.

-Classification schemes were applied to group studies into categories: research types (e.g., evaluation research, solution proposals) and application areas (e.g., threat detection, data protection).

Transparency in data extraction:

-A standardized template was used for data extraction, ensuring consistency across all studies.

-Two researchers reviewed Each study independently to minimize bias and ensure accurate categorization.

We utilize three classification schemes to systematically analyze and extract data relevant to security threats in social media environments [46-48]. As shown in Table 4, each scheme provides a distinct perspective on the research data.

Table 4. Research types

Category	Number of Papers
Evaluation Research	12
Validation Research	5
Philosophical Papers	5
Solution Proposals	14
Experience Papers	3
Opinion Papers	3

3.1 Facet 1: Types of research

This facet categorizes the research papers into distinct classes based on their nature and focus:

- Evaluation Research
- Validation Research
- Philosophical Papers
- Solution Proposals
- Experience Papers
- Opinion Papers

Figure 5 provides a detailed visualization of the research types.

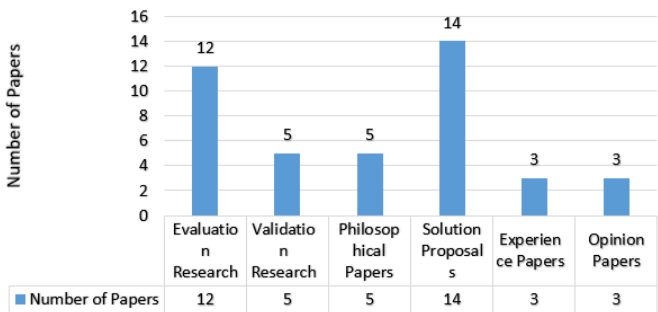

Figure 5. Types of research

3.2 Facet 2: Applications and AI techniques

This classification focuses on different AI techniques used to address security threats in social media, as illustrated in Table 5.

- Deep Learning Models
- Machine Learning Algorithms
- Anomaly Detection Systems
- Natural Language Processing (NLP)
- Risk Management Strategies

Figure 6 provides insights into how these AI techniques are applied to enhance social media security.

Table 5. Techniques of papers

Technique	Number of Papers
Deep Learning Models	7
Machine Learning Algorithms	9
Anomaly Detection Systems	8
Natural Language Processing (NLP)	6
Risk Management Strategies	12

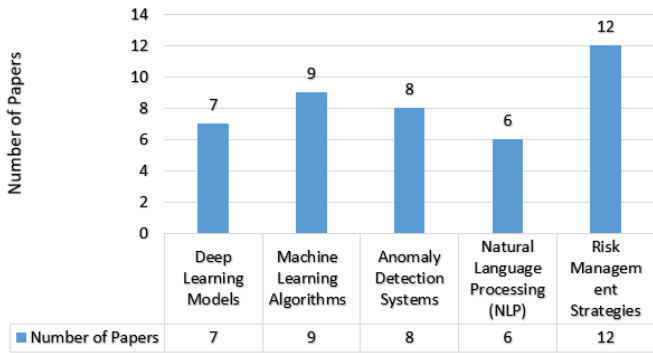

Figure 6. Distribution of papers by technique

3.3 Facet 3: Research domains

Table 6 explains how facet categorizes research based on its specific applications within social media security. Figure 7 illustrates the distribution of papers by domain. The classification highlights AI techniques' focus areas and applications, as detailed in the following categories.

Table 6. Research domains

Application	Number of Papers
Social Media Security	12
Threat Detection	10
Big Data Handling	7
Vulnerability Management	5
Theoretical and Analytical Studies	8

Figure 7. Distribution of papers by domains

Social media security: Research addressing the protection and security of social media platforms, including user data security and privacy measures.

Threat detection: Studies centered on identifying and analyzing threats in social media environments, focusing on

detection methods and strategies.

Big data handling: Papers exploring the application of AI in managing and analyzing large amounts of data from social media platforms for security purposes.

Vulnerability management: Research dedicated to mitigating and identifying vulnerabilities in social media systems.

Theoretical and analytical studies: Papers that provide theoretical insights or analytical frameworks relevant to the application of AI in social media security.

3.4 Facet 4: Datasets used

This facet categorizes research based on the datasets utilized in studies related to social media security. Table 7 explains the types of datasets employed and provides insights into the scope and context of the research. The classification includes:

- Twitter Data
- Facebook Data
- Custom Datasets
- Mixed Social Media

Table 7. Datasets used

Dataset	Number of Studies
Twitter Data	15
Facebook Data	10
Custom Datasets	7
Mixed Social Media	10

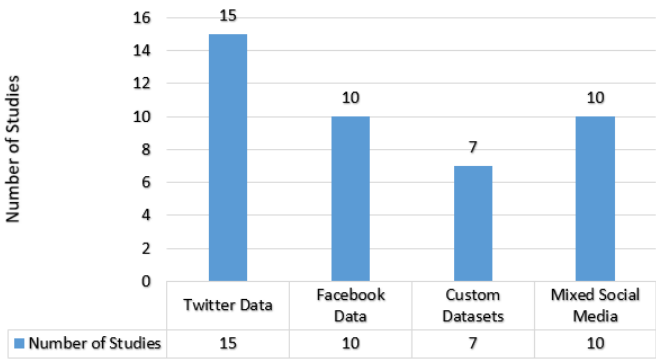

Figure 8. Distribution of papers by dataset

Figure 8 provides a detailed distribution of studies based on the datasets used.

4. ANALYSIS AND DISCUSSIONS

4.1 Merging facets for comprehensive analysis

To provide a holistic view of the research landscape in social media security, we have combined key facets to create a comprehensive analysis. This approach allows us to understand better the interplay between research types, AI techniques, research domains, and datasets used. Below are the two integrated tables that offer insights into these aspects. Figure 9 combines research types with AI techniques to illustrate how various research methodologies are applied across different AI techniques in social media security [49-51]. It provides a detailed look at the AI techniques used in the studies.

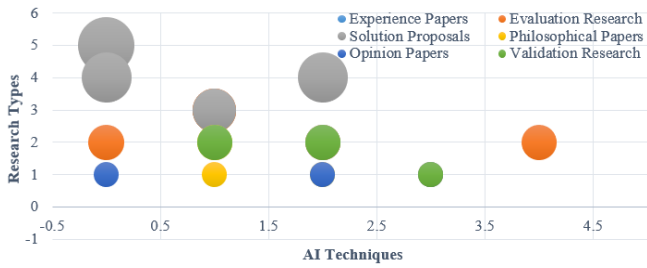

Figure 9. Distribution of research types and AI techniques

4.2 Trends in research types and AI techniques

Distribution of AI techniques: Figure 10 shows a diverse application of AI techniques across various research methodologies in social media security. Validation Research has a strong emphasis, suggesting a focus on verifying the effectiveness of AI solutions [52].

Methodological focus: Experience Papers and Evaluation Research also feature prominently, indicating a balance between theoretical advancements and practical applications in social media security.

Figure 10 combines the research periods with the data sets used to show how different data sets are applied in social media protection. This figure highlights the fields in the festivals used and the popular data sets within this field.

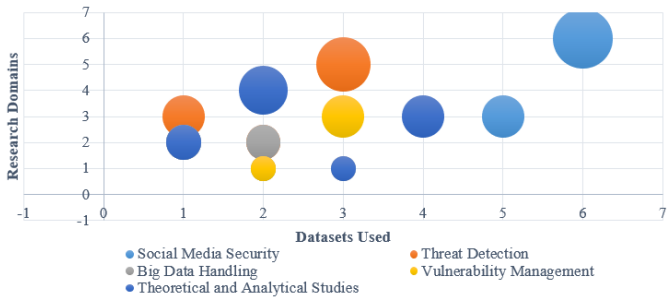

Figure 10. Distribution of research domains and datasets used

4.3 Patterns in research domains and datasets

Thematic concentration: The research is distributed across domains such as threat detection, vulnerability management, and big data handling. This suggests a broad interest in addressing multiple facets of social media security challenges.

Data usage trends: The distribution indicates the popularity of specific datasets, which could highlight benchmark datasets frequently utilized by the research community.

These figures are designed to create a clear and comprehensive view of the research events in the field of social media security, helping to identify trends and directions in the research approach to the use of smart data applications in security and social media.

5. CONCLUSION

In this systematic study, we comprehensively reviewed security threats in social media environments. The research was based on an extensive search process that included more than 1000 scientific research papers collected from different

databases, such as IEEE Xplore, ACM Digital Library, and ScienceDirect, over multiple periods. These inductive papers went through multiple filtering stages, where the number was reduced to 160 papers in the first stage and then to 42 papers that were analyzed in detail in the final stage.

The results showed that AI techniques such as natural language processing and machine learning play an important role in detecting security threats in social media, such as R, phishing, and unauthorized access. The papers were classified based on research types, which include philosophical papers and solution proposals, in addition to classifying research areas such as threat analysis and security risk management. Visualization tools, such as bubble diagrams, have been used to illustrate the relationships between different research. These models highlight areas of research focus, particularly in multi- and multi-solution, while identifying specific gaps in the representation of philosophical and technical research.

The bubble chart also shows that the bubble expansion began with technical solutions, with a heavy focus on personal data security and password security. In contrast, the multi-management and threats categories are less represented than other areas. These gaps indicate a need for further future research towards these less-represented areas.

Insights and Implications

Interdisciplinary approaches: The diverse methodologies and datasets highlight the interdisciplinary nature of social media security research, combining fields like AI, data science, and cybersecurity.

Growth in data utilization: The prominence of big data handling indicates increasing recognition of large-scale data analysis's role in addressing social media threats.

Focus on emerging threats: Threat detection is a dominant research domain, reflecting the critical need to tackle evolving cyber threats on social platforms.

Finally, the study emphasizes the importance of using AI technologies to support social media security. It highlights the need for reliable and adaptable models to the modifications that have been prepared. It calls for more innovative scientific research to be conducted for this purpose.

REFERENCES

[1] Gao, P., Wang, B., Gong, N.Z., Kulkarni, S.R., Thomas, K., Mittal, P. (2018). SYBILFUSE: Combining local attributes with global structure to perform robust Sybil detection. In 2018 IEEE Conference on Communications and Network Security (CNS), Beijing, China, pp. 1-9. https://doi.org/10.1109/CNS.2018.8433147

[2] Hassan, W.U., Hussain, S., Bates, A. (2018). Analysis of privacy protections in fitness tracking social networks-or-you can run, but can you hide? In 27th USENIX Security Symposium (USENIX Security 18), Baltimore, USA, pp. 497-512.

[3] Hogben, G. (2007). Security issues and recommendations for online social networks. ENISA Position Paper.

[4] Liu, M., Zeng, Y., Liu, Y., Liu, Z., Ma, J., Zhu, X. (2019). Collective influence based privacy preservation for social networks. In 2019 International Conference on Networking and Network Applications (NaNA): Daegu, Korea (South), pp. 282-289. https://doi.org/10.1109/NaNA.2019.00056

[5] Kaur, R., Singh, S., Kumar, H. (2018). AuthCom: Authorship verification and compromised account detection in online social networks using AHP-TOPSIS embedded profiling based technique. Expert Systems with Applications, 113: 397-414. https://doi.org/10.1016/j.eswa.2018.07.011

[6] Isaak, J., Hanna, M.J. (2018). User data privacy: Facebook, Cambridge Analytica, and privacy protection. Computer, 51(8): 56-59. https://doi.org/10.1109/MC.2018.3191268

[7] Yoshii, A., Plaut, D.A., McGraw, K.A., Anderson, M.J., Wellik, K.E. (2009). Analysis of the reporting of search strategies in Cochrane systematic reviews. Journal of the Medical Library Association, 97(1): 21-29. https://doi.org/10.3163/1536-5050.97.1.004

[8] Channabasava, U., Raghavendra, B.K. (2022). Ensemble assisted multi-feature learnt social media link prediction model using machine learning techniques. Revue d'Intelligence Artificielle, 36(3): 439-444. https://doi.org/10.18280/ria.360311

[9] İş, H., Tuncer, T. (2021). A profile analysis of user interaction in social media using deep learning. Traitement du Signal, 38(1): 1-11. https://doi.org/10.18280/ts.380101

[10] Abbas, A.K., Fleh, S.Q., Safi, H.H. (2015). Systematic mapping study on managing variability in software product line engineering: Communication. Diyala Journal of Engineering Sciences, pp. 511-520. https://doi.org/10.24237/djes.2008.0110212334

[11] Abbas, E., Qazi, A.A. (2024). Customized AI-powered security and privacy configurations for social media websites. BULLET: Jurnal Multidisiplin Ilmu, 3(1): 108-117.

[12] Fleh, S.Q., Bayat, O., Al-Azawi, S., Uçan, O.N. (2018). A systematic mapping study on touch classification. International Journal of Computer Science and Network Security, 18(3): 7-15.

[13] Abdillah, A., Widianingsih, I., Buchari, R.A., Nurasa, H. (2024). Big data security & individual (psychological) resilience: A review of social media risks and lessons learned from Indonesia. Array, 21: 100336. https://doi.org/10.1016/j.array.2024.100336

[14] Abuali, K.M., Nissirat, L., Al-Samawi, A. (2023). Advancing network security with AI: SVM-based deep learning for intrusion detection. Sensors, 23(21): 8959. https://doi.org/10.3390/s23218959

[15] Abuali, K.M., Nissirat, L., Al-Samawi, A. (2023). Intrusion detection techniques in social media cloud: Review and future directions. Wireless Communications and Mobile Computing, 2023(1): 6687023. https://doi.org/10.1155/2023/6687023

[16] Adekunle, T.S., Alabi, O.O., Lawrence, M.O., Ebong, G.N., Ajiboye, G.O., Bamisaye, T.A. (2024). The use of AI to analyze social media attacks for predictive analytics. Journal of Computing Theories and Applications, 1(4): 386-395. https://doi.org/10.62411/jcta.10120

[17] Momin, K.A., Hasnine, M.S., Sadri, A.M. (2024). Community-based behavioral understanding of crisis activity concerns using social media data: A study on the 2023 Canadian wildfires in New York City. arXiv preprint arXiv:2402.01683. https://doi.org/10.48550/arXiv.2402.01683

[18] AlAjlan, S.A., Saudagar, A.K.J. (2021). Machine learning approach for threat detection on social media

posts containing Arabic text. Evolutionary Intelligence, 14(2): 811-822. https://doi.org/10.1007/s12065-020-00458-w

[19] Alzaabi, F.R., Mehmood, A. (2024). A review of recent advances, challenges, and opportunities in malicious insider threat detection using machine learning methods. IEEE Access, 12: 30907-30927. https://doi.org/10.1109/ACCESS.2024.3369906

[20] Beigi, G., Shu, K., Zhang, Y., Liu, H. (2018). Securing social media user data: An adversarial approach. In Proceedings of the 29th on Hypertext and Social Media, Baltimore, USA, pp. 165-173. https://doi.org/10.1145/3209542.3209552

[21] Bharti, N.S.G., Gulia, P. (2023). Exploring machine learning techniques for fake profile detection in online social networks. International Journal of Electrical and Computer Engineering, 13(3): 2962-2971. https://doi.org/10.11591/ijece.v13i3.pp2962-2971

[22] Chukwuma, C. (2018). Social media and national security: Issues, challenges and prospects. International Journal of Social Sciences and Humanities Reviews, 8(1): 248-255.

[23] Dale, D., McClanahan, K., Li, Q. (2023). AI-based cyber event osint via twitter data. In 2023 International Conference on Computing, Networking and Communications, Honolulu, HI, USA, pp. 436-442. https://doi.org/10.1109/ICNC57223.2023.10074187

[24] Ebrahimi, M., Nunamaker Jr, J.F., Chen, H. (2020). Semi-supervised cyber threat identification in dark net markets: A transductive and deep learning approach. Journal of Management Information Systems, 37(3): 694-722. https://doi.org/10.1080/07421222.2020.1790186

[25] Fraga-Lamas, P., Fernandez-Carames, T.M. (2020). Fake news, disinformation, and deepfakes: Leveraging distributed ledger technologies and blockchain to combat digital deception and counterfeit reality. IT professional, 22(2): 53-59. https://doi.org/10.1109/MITP.2020.2977589

[26] Han, J., Li, Q., Xu, Y., Zhu, Y., Wu, B. (2024). Design of a trusted content authorization security framework for social media. Applied Sciences, 14(4): 1643. https://doi.org/10.3390/app14041643

[27] Han, Z., Li, S., Cui, C., Han, D., Song, H. (2019). Geosocial media as a proxy for security: A review. IEEE Access, 7: 154224-154238. https://doi.org/10.1109/ACCESS.2019.2949115

[28] Huang, S.Y., Ban, T. (2020). Monitoring social media for vulnerability-threat prediction and topic analysis. In 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), Guangzhou, China, pp. 1771-1776. https://doi.org/10.1109/TrustCom50675.2020.00243

[29] Iorga, D., Corlătescu, D., Grigorescu, O., Săndescu, C., Dascălu, M., Rughiniş, R. (2020). Early detection of vulnerabilities from news websites using machine learning models. In 2020 19th RoEduNet Conference: Networking in Education and Research (RoEduNet), Bucharest, Romania, pp. 1-6. https://doi.org/10.1109/RoEduNet51892.2020.9324852

[30] Jain, A., Kumar Gupta, D., Kumar Gupta, K., Authors, C., Kumar Gupta Principal, K. (n.d.). A review on social sentiment based predicting cyber-attacks. In International Journal of Engineering, Management and Humanities (IJEMH), 4(3): 254-260.

[31] Khurana, N., Mittal, S., Piplai, A., Joshi, A. (2019). Preventing poisoning attacks on AI based threat intelligence systems. In 2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP), Pittsburgh, USA, pp. 1-6. https://doi.org/10.1109/MLSP.2019.8918803

[32] Koujalagi, A., Thrupti, N.S., Kurbet, K. (2018). Security threats in Indian cyberspace by social media and cyberhoaxes. International Journal of Trend in Scientific Research and Development (IJTSRD), 2(4): 598-600. https://doi.org/10.31142/ijtsrd13040

[33] Kursuncu, U., Gaur, M., Castillo, C., Alambo, A., et al. (2019). Modeling Islamist extremist communications on social media using contextual dimensions: Religion, ideology, and hate. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW): 151. https://doi.org/10.1145/3359253

[34] Leon, R.D., Marcu, L.M. (2016). Social media platforms as a tool for sharing emotions: A perspective upon the national security agencies. Management Dynamics in the Knowledge Economy, 4(1): 141-152.

[35] Maathuis, C., Chockalingam, S. (2023). Modelling Responsible Digital Security Behaviour for Countering Social Media Manipulation. Indian Institute for Energy Technology.

[36] Maathuis, C., Kerkhof, I., Godschalk, R., Passier, H. (2023). Design lessons from building deep learning disinformation generation and detection solutions. In ECCWS 2023 22nd European Conference on Cyber Warfare and Security, Piraeus, Greece, pp. 285-293. https://doi.org/10.34190/eccws.22.1.1071

[37] Ali, M.A., Alqaraghuli, A. (2023). A survey on the significance of artificial intelligence (AI) in network cybersecurity. Babylonian Journal of Networking, 2023: 21-29. https://doi.org/10.58496/BJN/2023/004

[38] Mustafa, R.U., Japkowicz, N. (2024). Monitoring the evolution of antisemitic discourse on extremist social media using BERT. arXiv preprint arXiv:2403.05548. https://doi.org/10.48550/arXiv.2403.05548

[39] Organiściak, P., Kuraś, P., Kowal, B. (2023). Threats and crisis events detection using machine learning techniques with social media data. Zeszyty Naukowe Wyższej Szkoły Technicznej w Katowicach, 17: 123-141. https://doi.org/10.54264/0079

[40] Fleh, S.Q., Abbas, A.K., Saffer, K.M. (2015). A systematic mapping study on runtime monitoring of services. In the Iraqi Journal for Mechanical and Material Engineering, Special for Babylon First International Engineering Conference, pp. 121-134.

[41] Rasmussen, J. (2021). Share a little of that human touch: The marketable ordinariness of security and emergency agencies' social media efforts. Human Relations, 74(9): 1421-1446. https://doi.org/10.1177/0018726720919506

[42] Rau, M. (2017). The power of (In) security narratives in populist social media: The far-right's attempt of reclaiming conversation. Master thesis, Lund University.

[43] Sarangi, S.S., Singha, G. (2019). A survey on impact of AI and social media for rural development. International Journal of Computer Sciences and Engineering, 7(11): 64-68.

[44] Thapliyal, V., Thapliyal, P. (2024). Machine learning for cybersecurity: Threat detection, prevention, and response. Darpan International Research Analysis

(DIRA Journal), 12(1): 1-7.
[45] Sharma, K.R., Chiu, W.Y., Meng, W. (2023). Security analysis on social media networks via STRIDE model. arXiv preprint arXiv:2303.13075. https://doi.org/10.48550/arXiv.2303.13075
[46] Tosun, N., AltinÃķz, M., Emil, Ã., Turan, Ã., et al. (2020). A SWOT analysis to raise awareness about cyber security and proper use of social media: Istanbul sample. International Journal of Curriculum and Instruction, 12: 271-294.
[47] Vacarelu, M. (2023). Malicious use of artificial intelligence in political campaigns: Challenges for international psychological security for the next decades. In the Palgrave Handbook of Malicious Use of AI and Psychological Security, pp. 203-230. https://doi.org/10.1007/978-3-031-22552-9_8
[48] Van der Walt, E., Eloff, J.H., Grobler, J. (2018). Cyber-security: Identity deception detection on social media platforms. Computers & Security, 78: 76-89. https://doi.org/10.1016/j.cose.2018.05.015
[49] Weissburg, I.X., Arora, M., Pan, L., Wang, W.Y. (2024). Tweets to citations: Unveiling the impact of social media influencers on AI research visibility. arXiv e-prints, arXiv:2401.13782. https://doi.org/10.48550/arXiv.2401.13782
[50] Yuan, S., Wu, X. (2021). Deep learning for insider threat detection: Review, challenges and opportunities. Computers & Security, 104: 102221. https://doi.org/10.1016/j.cose.2021.102221
[51] Zhang, Z., Gupta, B.B. (2018). Social media security and trustworthiness: Overview and new direction. Future Generation Computer Systems, 86: 914-925. https://doi.org/10.1016/j.future.2016.10.007
[52] Herath, T.B., Khanna, P., Ahmed, M. (2022). Cybersecurity practices for social media users: A systematic literature review. Journal of Cybersecurity and Privacy, 2(1): 1-18. https://doi.org/10.3390/jcp2010001.