Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 18 No. 10 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijim.v18i10.46485

PAPER

Overview of Mobile Attack Detection and Prevention Techniques Using Machine Learning

Ahmad K. Al Hwaitat[1](✉), Hussam N. Fakhouri[2], Moatsum Alawida[3], Mohammed S Atoum[1], Bilal Abu-Salih[1], Imad K. M. Salah[1], Saleh Al-Sharaeh[1], Nabil Alassaf[1]

[1]King Abdullah II School of Information Technology, The University of Jordan, Amman, Jordan

[2]Department of Data Science and Artificial Intelligence, University of Petra, Amman, Jordan

[3]Department of Computer Sciences, Abu Dhabi University, Abu Dhabi, United Arab Emirates

a.hwaitat@ju.edu.jo

ABSTRACT

In light of the increasing sophistication and frequency of mobile attacks, there is a growing demand for advanced intelligent techniques capable of offering comprehensive mobile attack detection and prevention. This paper aims to critically evaluate the landscape of mobile security, outlining the evolution of mobile attack vectors and pinpointing the deficiencies in traditional security methods. The text embarks on a journey to understand the connection between machine learning (ML) and its promising applications in enhancing mobile security. First, we outline the current state of mobile attacks and the traditional methods used for their detection, emphasizing the clear limitations and the necessity for an innovative approach. Following this, we will elucidate the fundamentals of ML and its implications in cybersecurity, exploring the benefits it can provide to mobile attack detection frameworks. We delve into discussing various ML algorithms, such as decision trees, random forests, and support vector machines, highlighting their effectiveness and the metrics used to evaluate ML models in security tasks. Moreover, the paper sheds light on novel approaches such as semi-supervised and unsupervised learning in anomaly detection, as well as the applications of transfer learning in security. Addressing the pressing challenges faced in artificial intelligence (AI)-driven mobile attack detection, we delve deep into the intricacies of data collection, labeling, and the prevailing issues of imbalance and overfitting. Furthermore, we explore contemporary adversarial attacks and defenses, scrutinizing the real-world adaptability of AI models and the pivotal role of human-AI collaboration in enhancing attack detection mechanisms.

KEYWORDS

mobile security, machine learning (ML) in cybersecurity, intelligent attack detection (IAD), adversarial attacks and defenses

1 INTRODUCTION

Mobile technologies have undergone a significant transformation over the past few decades, evolving from basic communication devices to multifaceted

Al Hwaitat, A.K., Fakhouri, H.N., Alawida, M., Atoum, M.S., Abu-Salih, B., Salah, I.K.M., Al-Sharaeh, S., Alassaf, N. (2024). Overview of Mobile Attack Detection and Prevention Techniques Using Machine Learning. *International Journal of Interactive Mobile Technologies (iJIM)*, 18(10), pp. 125–157. https://doi.org/10.3991/ijim.v18i10.46485

Article submitted 2023-11-29. Revision uploaded 2024-02-14. Final acceptance 2024-02-17.

© 2024 by the authors of this article. Published under CC-BY.

platforms that are now essential to modern society. The advent of applications for banking, health monitoring, and a myriad of other functionalities has significantly elevated the prominence of mobile devices in our daily lives [1]. As these devices have evolved to manage and store a diverse range of personal and sensitive data, they have concurrently become appealing targets for cyber threats. The complexity and variety of attacks on mobile platforms have experienced a concomitant escalation with the expanding functionalities of these devices. Modern-day attackers employ increasingly sophisticated methodologies, leveraging vulnerabilities inherent in mobile systems and the vast amount of valuable information they contain. As a result, conventional security measures, which may have once been deemed adequate, are now frequently overwhelmed, struggling to keep pace with the innovative tactics of cyber adversaries [2]. The proliferation of mobile technologies in recent decades has led to a radical transformation in the way individuals and businesses operate, giving birth to a mobile-centric paradigm in the digital ecosystem. Originally designed as mere communication tools, mobile devices have transcended their primary function and now play pivotal roles in various sectors, including finance, healthcare, education, and entertainment [3]. Banking applications, for instance, enable users to conduct transactions, manage their finances, and even apply for loans, all from the convenience of their devices. Similarly, health monitoring applications provide real-time data on vital signs, offer recommendations, and even facilitate virtual consultations with medical professionals. These advanced functionalities, however, come at a cost. The vast amount of personal, sensitive, and, in some cases, critical information stored on these devices makes them lucrative targets for cybercriminals [4].

Contemporary cyber-attacks on mobile platforms are no longer limited to simple malware or phishing schemes. Attackers today employ a plethora of sophisticated strategies, ranging from exploiting zero-day vulnerabilities to launching advanced persistent threats (APTs) [5]. Furthermore, the open-source nature of platforms such as Android offers a double-edged sword—on the one hand, fostering innovation, and on the other, providing fertile ground for potential exploits [6]. As the complexity and frequency of these attacks grow, there is a paramount need to critically examine and fortify existing security frameworks. The responsibility now lies with researchers, developers, and industry professionals to work together to develop adaptive, robust, and forward-looking security solutions that can protect our mobile future [7].

The ubiquitousness of mobile devices has fundamentally reshaped the digital landscape. As an example, mobile e-commerce transactions have skyrocketed in the past few years, prompting businesses to prioritize mobile-optimized interfaces for their consumers [8]. Moreover, the significance of mobile devices as the main access point to the internet for many users in developing nations cannot be overstated. For these populations, mobile phones are not just tools of convenience but serve as the primary gateway for accessing essential services, further highlighting the importance of ensuring their security [9]. The landscape of mobile cyber threats has become a hotbed for innovation among cybercriminals. Ransomware attacks targeting mobile platforms, for instance, have seen a notable surge. These malicious programs encrypt the victim's data, demanding a ransom in exchange for decryption. Given the personal nature of the information stored on mobile devices—such as photos, contacts, and messages—the psychological leverage attackers have is tremendous [10]. Additionally, with the proliferation of Internet of Things (IoT) devices,

mobile devices now serve as command centers for controlling a myriad of other devices [11]. This adds a new dimension to the security concerns linked to mobile platforms. Compromising the mobile interface of a smart home, for example, could allow unauthorized control over a variety of household systems. Social engineering attacks, which manipulate users into revealing sensitive information, also pose a unique set of security challenges on mobile platforms [12]. The rapidly changing nature of the mobile ecosystem, characterized by frequent software updates and a multitude of app stores, further complicates security measures [13]. Traditional security models, often relying on periodic updates or scans, are increasingly inadequate for dealing with the dynamic landscape of mobile content. The ever-changing mobile environment, characterized by its deepening integration into both personal and professional realms as well as its continually evolving threat landscape, accentuates the urgent requirement for comprehensive, adaptive, and anticipatory mobile security measures [14].

1.1 Research objective

This overview explores the rapidly evolving field of mobile security, with a focus on developing intelligent techniques to enhance attack detection and prevention. It comprehensively analyzes current mobile security threats, including a range of attacks from malware to cyber-espionage, and critically evaluates traditional security methods to set the stage for more advanced, artificial intelligence (AI)-based approaches. A significant portion of the research is focused on investigating the role of machine learning (ML) and deep learning (DL) in cybersecurity. This includes a detailed examination of various ML algorithms, their application in security, evaluation metrics, and advancements in semi-supervised and unsupervised learning. The study also explores the potential of DL, particularly focusing on neural networks, convolutional and recurrent structures, and their implications for to mobile security. Additionally, it covers cutting-edge trends such as Transformer-based models and attention mechanisms, emphasizing their potential impact on security measures.

2 BACKGROUND OF MOBILE SECURITY

Mobile security, in its foundational sense, encompasses defensive strategies designed to safeguard information stored and transmitted on a growing variety of mobile devices, including smartphones, tablets, and even smart wearables [15]. As we unravel the tapestry of its evolution, one can perceive the magnified dimensions of its complexity, moving from rudimentary device protection to an encompassing defense mechanism that accounts for software, hardware, user behavior, and network dynamics [16].

In the nascent stages of mobile technology, devices were rather elementary, predominantly crafted for voice communication. The first-generation mobile phones had minimal security measures, primarily due to their limited data processing capabilities and lack of integration with the internet [17].

The advent of smartphones, however, signaled a transformative phase. After 2007, following the groundbreaking introduction of Apple's iPhone and the subsequent rise of Android-based devices, smartphones started taking on roles that

were previously exclusive to personal computers [18]. They became gateways to the internet, tools for professional communication through emails, platforms for digital transactions, and much more. Such multifunctional capabilities undeniably broadened the potential threat landscape, making mobile devices susceptible to an array of cyberattacks, ranging from malware infestations to sophisticated phishing attempts [19].

Application ecosystems, such as the App Store and Google Play Store, burgeoned rapidly, bringing with them the danger of rogue applications. There were instances where malevolent apps, camouflaged as genuine software, infiltrated official marketplaces, posing threats such as unauthorized data access, eavesdropping, and even covert system control [20].

Corporate environments, by embracing the cost-efficiency and flexibility of the bring your own device (BYOD) paradigm, have inadvertently expanded the mobile security horizon. When employees started accessing critical company databases and sensitive information from personal devices, the line between professional and personal data became blurred. This introduced potential vulnerabilities in the otherwise strong corporate security system. The subsequent emergence of mobile device management (MDM) and mobile application management (MAM) solutions aimed at creating secure enclaves for professional data on personal devices reflects the industry's adaptation to these shifts [21].

The spread and penetration of wireless networking technologies, notably Wi-Fi, 4G, and 5G, have introduced another dimension to mobile security. Although they heralded an era of faster, seamless connectivity, they also spawned vulnerabilities, particularly when these networks were inadequately configured or remained unencrypted. When devices are connected to insecure networks, they are exposed to risks such as man-in-the-middle (MitM) attacks. These malicious actors can secretly intercept and, in some cases, manipulate the data being exchanged between parties [22].

Modern mobile devices, equipped with an array of sensors such as GPS, accelerometers, gyroscopes, and more, inadvertently pose unique security and privacy challenges. For instance, unauthorized access to a device's location services could result in potential stalking or unsolicited location tracking, emphasizing the importance of robust and detailed permission frameworks. The dynamism inherent in the mobile ecosystem, characterized by a wide variety of device manufacturers, diverse operating systems, and the relentless churn of software updates, naturally complicates the mobile security landscape [23].

Mobile security, at its core, encompasses the protective measures designed to safeguard information stored and transmitted on a constantly evolving range of mobile devices, including smartphones, tablets, and increasingly, smart wearables. As we trace its trajectory, it's evident that its concerns have evolved from mere device protection to an overarching defense system, encompassing software vulnerabilities, hardware limitations, fluctuating user behaviors, and network vulnerabilities. Initially, mobile devices were primarily designed for voice communication, with minimal features and little integration with online environments. However, the landscape underwent a paradigm shift with the emergence of smartphones [24] [25]. Table 1 presents a detailed comparison of mobile security across three distinct eras: the pre-smartphone era, the advent of smartphones, and the modern mobile security landscape. It highlights the evolution of security concerns and strategies, from basic device protection to advanced cyber threat management.

Table 1. A comparative analysis of the pre-smartphone era, the advent of smartphones, and the modern mobile security landscape [15–25]

Era/Aspect	Early Mobile Technology (Pre-Smartphone Era)	Advent of Smartphones (Post-2007)	Modern Mobile Security Landscape
Primary Use	Voice communication	Internet access, email, digital transactions	Multifunctional (internet, apps, communication, sensors)
Security Concerns	Minimal due to limited data processing and no internet integration	Increased due to internet connectivity and versatile functionality	Advanced, encompassing software, hardware, user behavior, and network vulnerabilities
Threat Landscape	Negligible	Broadened to include malware, phishing	Expanded to sophisticated cyberattacks, app-based threats, and sensor-related privacy issues
Application Ecosystem	Non-existent	Emergence of App Store and Google Play Store, increased risk of rogue applications	Continuously evolving with frequent updates and new app introductions
Corporate Integration (BYOD)	Not applicable	Emergence of BYOD, blurring lines between personal and professional data	Implementation of MDM and MAM for data security
Networking Technologies	Limited or no connectivity	Introduction of Wi-Fi, 3G/4G networks, increasing connectivity vulnerabilities	Advanced networks (4G, 5G) with inherent security risks
Device Sensory Capabilities	Basic or non-existent	Limited	Extensive (GPS, accelerometers, gyroscopes) leading to unique security and privacy challenges
Security Measures	Rudimentary device protection	Development of comprehensive security software, awareness of app permissions	Advanced security protocols, encryption, and permission frameworks

2.1 The necessity for evolved security methods

The interconnectedness of the modern digital age, coupled with the increasing ubiquity of mobile devices, necessitates an ongoing evolution in security practices. The increasing sophistication of cyber threats targeting mobile platforms underscores the decreasing effectiveness of traditional security measures, necessitating the adoption of advanced methodologies [27].

Firstly, the magnitude and variety of sensitive data stored on mobile devices have grown exponentially. From personal photographs and messages to banking details, health records, and even biometric information, these devices are treasure troves of information, making them coveted targets for malicious actors. As the value of the information stored on mobile devices grows, so does the incentive for attackers to develop new and innovative ways to breach these systems [28].

Secondly, the mobile application ecosystem, which has experienced explosive growth, has also inadvertently opened new avenues for cyber threats. While the App Store has implemented verification processes, malicious actors persistently attempt to circumvent these checks by introducing rogue applications filled with malware or spyware. This ever-evolving cat-and-mouse game with attackers demands a parallel evolution in mobile security measures [29].

Compounding this issue is the multitude of mobile operating systems and the resulting fragmentation, especially within the Android domain. Unlike the more streamlined updates of platforms such as iOS, Android devices, due to their varied manufacturers and models, often run different versions of the OS. This fragmentation creates an environment where not all vulnerabilities are patched uniformly or promptly, making certain devices more susceptible to attacks [30].

The rise of the IoT has further accentuated the need for advanced mobile security practices. Mobile devices often serve as control hubs for a variety of connected devices, such as smart thermostats and security cameras. A vulnerability in the mobile device could jeopardize the entire connected ecosystem [31].

Furthermore, modern cyberattacks have become more targeted and personalized. Techniques such as spear phishing now leverage detailed personal information to craft highly convincing deceptive messages, which traditional spam filters might not detect. With mobile devices being primary communication tools, their role in intercepting and thwarting these threats becomes critical [32].

Social engineering attacks have also found fertile ground in the mobile domain. Scam calls, fraudulent SMS links, or even deceptive pop-ups in mobile browsers aim to manipulate users into voluntary actions that compromise security. This highlights the importance of comprehensive security solutions that include user education and behavior-aware security tools [33].

Globalization and the increasing mobility of the workforce underscore the need for improved security measures. As businesses expand their operations across borders, employees often access corporate networks from various locations, some of which may be insecure. This dynamic poses potential security challenges, mandating advanced, adaptive, and context-aware security mechanisms [34].

2.2 The evolution of mobile attack vectors

The cyber threat landscape has witnessed a dramatic transformation over the past few decades, especially concerning mobile platforms. The evolution of mobile attack vectors can be understood as a reflection of broader technological advancements, the ubiquity of mobile devices, and the changing dynamics of cyber threats [35].

In the initial stages of mobile device proliferation, the primary concern was related to physical theft or loss of the device. This was largely due to the devices being primarily used as communication tools with limited access to the internet and fewer functionalities [36].

With the advent of smartphones and the subsequent explosion of mobile applications in the late 2000s, the threat landscape began to shift. Malicious applications, often masquerading as legitimate ones, have emerged as significant threats. These rogue applications, once installed, could siphon off personal data, deliver malicious payloads, or transform the device into a bot within a larger botnet [37].

As mobile devices became more integrated into daily life and business operations, they began storing a plethora of sensitive data. This transition made them a lucrative target for cybercriminals, leading to a rise in data breaches and ransomware attacks targeting mobile platforms [38].

Another significant evolution occurred with the widespread adoption of mobile banking and financial apps. Phishing attacks, which were previously primarily limited to email platforms, have started targeting mobile users. Smishing (SMS phishing) and vishing (voice call phishing) have become commonplace, exploiting users' trust in the relative security of their mobile devices [39].

The rise of the IoT and the integration of mobile devices within this interconnected ecosystem have further complicated the threat landscape. Mobile devices, serving as control nodes for various IoT devices, have become potential gateways for larger-scale attacks. This shift posed threats not only to individual data privacy but also to critical infrastructure, especially when mobile devices were used to manage or access industrial IoT setups [40].

Advanced persistent threats targeting mobile platforms began to emerge in the late 2010s. These attacks, often state-sponsored or backed by well-funded criminal syndicates, are characterized by their stealth, persistence, and sophistication. Mobile devices, due to their always-connected nature and the intimate access they offer to users' personal and professional lives, have become prime targets for these sophisticated campaigns [41].

Furthermore, the act of jailbreaking or rooting mobile devices to circumvent manufacturer restrictions has introduced another vulnerability. While these practices offer users greater control over their devices, they also expose them to potential malicious attacks by circumventing the built-in security mechanisms of the device's operating system [42].

Furthermore, the emergence of crypto-jacking on mobile platforms has demonstrated a shift in cybercriminal motivations. Instead of merely stealing data, attackers have begun utilizing the processing power of mobile devices to mine cryptocurrency without the user's knowledge or consent. This unauthorized activity results in decreased device performance and increased wear and tear [43].

In parallel with these attack vectors, the methods used to exploit mobile devices have also evolved. Zero-day vulnerabilities, which are flaws unknown to the software developer and the public, have become highly sought after by attackers. These vulnerabilities, when exploited, allow attackers to compromise devices before a patch can be developed or deployed [44]. Table 2 provides an insightful comparative analysis of the evolution of mobile attack vectors, tracing the progression from basic issues like physical theft in the early stages of mobile devices to the advanced cyber threats faced in today's mobile security landscape.

Table 2. Comparative analysis of the evolution of mobile attack vectors: Tracing the shift from physical theft to advanced cyber threats in mobile security

Era/ Phase	Early Mobile Devices	Advent of Smartphones	Integration in Business and Daily Life	Rise of Mobile Banking	IoT Integration	Emergence of APTs	Jailbreaking/ Rooting	Cryptojacking	Zero-Day Vulnerabilities
Primary Concerns	Physical theft or loss	Malicious applications	Data breaches and ransomware	Phishing (Smishing, Vishing)	Gateway to IoT attacks	Stealthy, persistent campaigns	Security vulnerabilities	Unauthorized cryptocurrency mining	Exploitation of unknown software flaws
Characteristics of Attacks	Limited due to device functionality	Siphoning personal data, malicious payloads	Sensitive data targeting	Exploiting trust in mobile security	Potential threats to critical infrastructure	Sophistication, state-sponsored or well-funded	Compromised device security	Resource exploitation, performance issues	Bypassing security before patch deployment
Technological Context	Limited internet access, basic functionality	Wide array of applications, increased connectivity	Extensive storage of sensitive data	Widespread adoption of financial apps	Mobile devices as control nodes for IoT	Advanced, continuous targeting	User-driven modification of OS	Utilization of device processing power	Sophisticated exploitation of software flaws
Cybercriminal Motivations	Opportunistic theft	Personal data theft, device control	Financial gain, data exploitation	Financial fraud, identity theft	Broad-scale disruption, espionage	Long-term espionage, data theft	Gaining unrestricted device access	Financial gain through crypto mining	Gaining undetected access, espionage

2.3 Challenges faced in conventional mobile security

Conventional mobile security paradigms have struggled with the complexities posed by a constantly evolving digital ecosystem. One significant concern arises from the diversity of operating systems prevalent in the mobile domain. Unlike the relatively standardized desktop environments, the mobile landscape is filled with

a variety of OS versions and customizations. Particularly, the open-source nature of Android has led to the development of numerous vendor-specific customizations, each providing different security features. This fragmentation complicates the uniform rollout of security patches, often leaving devices exposed to well-known vulnerabilities for extended durations [45].

Further exacerbating the situation is the rapid software lifecycle that characterizes the mobile app world. Developers, in a rush to meet market demands, sometimes sideline comprehensive security vetting, emphasizing functionality and speed over robust security measures [46]. This rapid development process frequently intersects with another significant challenge: excessive app permissions. A significant number of mobile applications request more extensive permissions than necessary for their functionalities. Users, often unaware of the potential consequences, frequently grant such extensive permissions, inadvertently creating opportunities for unauthorized data access or even malicious exploits [47].

The physical security of mobile devices presents its own set of challenges. Given their portability, these devices are inherently susceptible to theft or loss. In such scenarios, conventional security measures might offer scant protection, especially in the absence of encryption or robust authentication mechanisms [48]. Alongside traditional security solutions, such as antivirus software, have remained anchored to signature-based detection methodologies. While effective against known threats, this reactive approach falters in the face of zero-day attacks or polymorphic malware, which continually morphs to evade detection [49].

Compounding these technical challenges are the intrinsic limitations of mobile devices. Despite the remarkable advancements they have undergone, mobile devices are still limited by their processing capabilities and battery life. Intensive security processes, such as exhaustive device scans, can strain resources, leading to suboptimal device performance and often discouraging users from implementing such security measures [50–54].

3 MOBILE ATTACKS

3.1 Types of mobile attacks

The mobile ecosystem, due to its ubiquity and complexity, has become fertile ground for numerous security threats. These threats range from those targeting the underlying infrastructure to those exploiting human vulnerabilities [56]. One of the most prevalent attacks in the early days of mobile technology was SMS phishing (or 'smishing'). By crafting deceptive text messages, adversaries lured users into divulging personal or financial details, capitalizing on the trust people placed in SMS communications at the time. While smishing continues to pose threats, more sophisticated attacks have emerged, necessitating advanced countermeasures.

Malware, for instance, has increasingly become a menace in the mobile domain. Ranging from spyware that surreptitiously records user data to ransomware that locks users out of their devices, these malicious software packages often find their way onto devices through seemingly legitimate applications, especially when sourced from unofficial app stores or via sideloading.

Another pervasive threat vector is the man-in-the-middle (MitM) attack. In such scenarios, attackers intercept communication between two parties, either to eavesdrop or to alter the communication. The widespread use of public Wi-Fi networks in cafes and transport hubs has exacerbated this vulnerability, with users often unknowingly connecting to rogue hotspots set up by attackers [57].

Application-based attacks have also seen a surge, where vulnerabilities within an application or the way it interfaces with the system can be exploited. These vulnerabilities can arise from poor coding practices, inadequate security vetting during app development, or even through third-party libraries that the app relies on. Such attacks can lead to unauthorized data access, data corruption, or even total system compromise.

In recent years, the increasing reliance on mobile payments and digital wallets has resulted in the rise of financial threats that specifically target these platforms. Attackers often employ a combination of techniques, from app overlays that masquerade as legitimate payment interfaces to trojans that lie dormant, only to activate during financial transactions, skimming sensitive data in the process [58].

Physical threats cannot be discounted either. Due to their portability, mobile devices are vulnerable to theft. In the absence of robust encryption and security protocols, such incidents can lead to unauthorized access to data. Moreover, techniques like shoulder surfing, where attackers glean sensitive information by directly observing user input, underscore the multifaceted nature of threats that the mobile domain contends with.

Furthermore, with the rise of the IoT and the convergence of mobile platforms with a plethora of connected devices, newer vulnerabilities have come to the fore. These interconnected devices often introduce novel entry points for attackers, ranging from smart refrigerators to wearable health devices, each posing unique security challenges [59].

3.2 Traditional methods for detecting mobile attacks

Historically, as the digital realm evolved and mobile devices became a household staple, the need to secure these devices from myriad threats led to the emergence of traditional defense mechanisms. These methods, rooted in the then-prevalent paradigms of digital security, aimed to address the unique vulnerabilities of the mobile arena [61].

Signature-based detection was among the first lines of defense. Drawing parallels from the desktop domain, this method relied on maintaining a repository of known malicious software signatures. Whenever a piece of software or an application is introduced or updated on the device, it is scanned against this repository. If a match is found, the system will flag it as malicious and take appropriate action, such as quarantine or deletion. However, the effectiveness of this method waned over time as malware authors began using polymorphic and metamorphic techniques to alter the software's appearance without changing its core functionality.

Heuristic-based detection emerged as an evolution of the signature-based approach. Instead of simply matching known signatures, heuristics analyze the behavior and attributes of software or data packets. If certain predefined suspicious patterns or characteristics are identified, the software would be considered a potential threat. This approach provided a more dynamic defense mechanism capable of detecting previously unseen malware or variations of known malware, but it also increased the rate of false positives [62].

Static and dynamic analysis also played pivotal roles in traditional mobile security. Static analysis involves inspecting the software without executing it and assessing aspects such as code structure, embedded resources, and requested permissions. This provides an early indication of any embedded malintent. Dynamic analysis, in contrast, involves running the software in a controlled environment, often emulated, to observe its behavior and interactions with the system. While static analysis

offers speed and efficiency, dynamic analysis can reveal sophisticated attacks that only manifest during execution.

Network-based intrusion detection systems (NIDS) were deployed to monitor data traffic to and from mobile devices. By analyzing packets and traffic patterns, NIDS can identify suspicious activities that may indicate a potential attack, such as distributed denial of service (DDoS) attacks or unauthorized data exfiltration. However, encryption protocols and the proliferation of secure tunneling mean that attackers could sometimes bypass NIDS by simply cloaking their malicious activities [63].

Device-level hardening is a proactive measure that focuses on minimizing vulnerabilities from the device's inception. Manufacturers and developers would lock down certain features, restrict permissions, and sometimes even create isolated environments (or sandboxes) where applications could operate without jeopardizing the device's core functions or data. Despite the efficacy of these traditional methods during their prime, the rapidly changing landscape of mobile threats, combined with technological advancements and nuances in user behavior, have highlighted their limitations. While they formed the bulwark of mobile defense for a significant period, the escalating sophistication of attacks necessitated the exploration of more evolved and intelligent detection mechanisms [64].

3.3 Limitations of current methods

The relentless advancement in mobile technology, paralleled by the ever-evolving threat landscape, has shed light on the shortcomings of current mobile security strategies. As the digital ecosystem grew in complexity, it became evident that many of the established methods, although foundational, presented notable shortcomings in effectively countering contemporary threats [65].

A significant limitation of signature-based detection is its inherently reactive nature. Reliant on a database of known malware signatures, traditional antivirus software remains ineffective against zero-day exploits. These exploits occur when malware breaches security before the vulnerability becomes publicly known and before a signature can be developed. This delay between the emergence of malware and the update of signatures exposes mobile devices to potential security breaches.

Heuristic-based detection, while providing broader protection by analyzing behavioral patterns, faces the challenge of false positives. The risk of misidentifying benign software as malicious based on general behavioral attributes can lead to unwarranted actions, disrupting essential device functions or legitimate applications. Such false alarms can also desensitize users, making them less responsive to genuine threats.

Static and dynamic analyses are resource-intensive and may not always scale efficiently with the voluminous number of apps available in the marketplace. The sheer volume of applications, updates, and patches introduced daily can overwhelm the capabilities of static and dynamic analysis tools. Sophisticated attackers are increasingly employing techniques to detect when their software is being run in a simulated environment. They alter its behavior to appear benign during the analysis.

Network-based intrusion detection systems, although vigilant sentinels of network traffic, struggle with challenges posed by encrypted traffic. The widespread adoption of encryption and VPNs for privacy and security reasons can obscure malicious traffic, making certain attack vectors undetectable to NIDS. Moreover, mobile devices, often switching between various networks (Wi-Fi, cellular, etc.), present a dynamic environment where continuous monitoring can be challenging [66] [67].

3.4 Preliminary work in intelligent techniques

As the limitations of traditional mobile security methods became glaringly apparent, the research community shifted its focus towards intelligent techniques. This involves leveraging the power of advanced computational methods and data-driven approaches to enhance mobile security measures [68].

One of the pioneering endeavors in this domain involved using AI for anomaly detection. Early researchers observed that many malicious activities inherently deviate from typical patterns of software behavior. By training AI models on regular software activities, these systems could flag deviations as potential threats, even if the specific threat signature was previously unknown [69].

Machine learning, a subset of AI, has shown particular potential in enhancing malware detection rates. Initial studies involved training ML algorithms with features extracted from known malware samples and benign applications. These algorithms, once trained, demonstrated a keen ability to classify and detect new, unseen malware based solely on the behavioral and structural features of the applications [70].

Natural language processing, another branch of AI, was explored in the context of phishing detection. Early experiments found that many phishing attempts often displayed discernible textual patterns, anomalies, or linguistic inconsistencies. By training NLP models on legitimate communications, preliminary systems could effectively identify and filter phishing attempts based on textual content analysis [71].

In addition to these, neural networks, especially DL models, have been explored for their ability to extract complex patterns and relationships from large datasets. Preliminary experiments involving convolutional neural networks (CNNs) for image-based authentication and recurrent neural networks (RNNs) for pattern-based intrusion detection showcased promising results [72].

Furthermore, the early adoption of ensemble learning, where multiple models collectively make decisions, bolstered the robustness and accuracy of mobile threat detection. By leveraging the strengths of individual models and mitigating their weaknesses through a collective approach, ensemble methods exhibit enhanced resilience against false positives and improved overall detection rates [73]. Table 3 provides a comprehensive overview of the preliminary work on intelligent techniques for mobile security, spanning from AI-based anomaly detection to ensemble learning for threat mitigation.

Table 3. Overview of preliminary work in intelligent techniques for mobile security: from AI-based anomaly detection to ensemble learning in threat mitigation

Intelligent Technique	Description	Key Contributions	Impact on Mobile Security
Artificial Intelligence (AI) for Anomaly Detection	Utilization of AI to identify deviations from typical software behaviors	Early detection of unknown threats by flagging anomalies	Enhanced ability to identify novel threats without prior knowledge of their signatures
Machine Learning (ML) for Malware Detection	Training ML algorithms with known malware and benign application features	Improved malware detection rates, ability to classify new, unseen malware	Significantly increased accuracy in identifying and categorizing malware based on behavioral and structural features
Natural Language Processing (NLP) for Phishing Detection	Application of NLP to analyze textual content and patterns in communications	Effective identification and filtering of phishing attempts based on textual analysis	Advanced capability to detect phishing attempts through linguistic and textual inconsistencies
Neural Networks for Pattern Detection	Exploration of deep learning models like CNNs and RNNs for security applications	CNNs for image-based authentication, RNNs for intrusion detection	Ability to extract complex patterns and relationships, improving authentication and intrusion detection mechanisms
Ensemble Learning for Threat Detection	Combining multiple models to make collective decisions	Enhanced resilience against false positives, improved detection rates	Increased robustness and accuracy in mobile threat detection, leveraging strengths of individual models

3.5 The evolution of mobile security tools

In the ever-evolving landscape of mobile threats, the tools and strategies used for mobile security have undergone significant transformations. Initially, the focus was primarily on functionality and user experience, often relegating security to the background. However, vulnerabilities became more apparent, making the implementation of robust security measures imperative [74].

The first generation of mobile security tools was predominantly signature-based, similar to their traditional computer counterparts. While effective against known threats, these tools were limited in countering new or mutated malware. Heuristic methods emerged to address these shortcomings by utilizing behavioral patterns to identify potentially malicious activities. For instance, flagging occurrs when an app requests permissions unrelated to its core functionality [75].

The advent of cloud computing marked a significant shift in mobile security. By offloading analytics to the cloud, devices could leverage extensive threat databases without storage limitations. This also enabled real-time updates, thereby enhancing protection against emerging threats. Furthermore, sandboxing techniques were introduced to isolate and analyze suspicious applications in a controlled environment [76].

The most transformative change came with the integration of AI and ML. These data-driven approaches not only detect known threats but also predict new, unseen ones based on learned patterns. Alongside this, multi-factor authentication (MFA) has also risen to prominence, providing a layered defense by requiring multiple forms of verification [77].

3.6 Analysis of mobile security threats

Malware, a broad category encompassing viruses, worms, trojans, and spyware, poses a significant threat to mobile devices. Viruses attach themselves to legitimate programs, propagate, corrupt data, and hinder device performance. Worms are particularly insidious because they self-replicate and spread across networks without user interaction, often overwhelming resources. Trojans disguise themselves as benign applications but carry harmful code intended to steal or disrupt data. Spyware is particularly concerning because it can secretly monitor and transmit user activity, including sensitive information such as passwords and financial details. These malicious programs can range from simple adware causing nuisance to sophisticated software capable of commandeering complete control of a device, often unbeknownst to the user.

Phishing in the mobile realm often involves deceiving users into disclosing sensitive information by masquerading as trustworthy entities. This deception can manifest through various mediums: Smishing (SMS phishing) involves sending deceptive text messages that lure recipients into revealing personal details or clicking on malicious links. Email-based phishing targets users with emails that mimic legitimate communication from banks, social networks, or other credible sources. Additionally, phishing occurs through websites that replicate legitimate sites to capture login credentials or other private data. These websites are often linked in emails or text messages, trapping unsuspecting users.

Network spoofing and man-in-the-middle attacks pose serious risks to mobile security. Network spoofing involves creating fake WiFi networks that appear legitimate to unsuspecting users. Once connected, attackers can monitor and intercept data transmitted over these networks. Man-in-the-middle attacks are more insidious. Attackers insert themselves into a two-party transaction or communication, stealthily intercepting and manipulating the data transmitted between the parties.

Crypto-jacking is an emerging threat where a mobile device's processing power is covertly used for mining crypto currency. This unauthorized use can lead to reduced device performance, battery drainage, and overheating. Symptoms of cryptojacking often include device sluggishness, unexpected reboots, or unusually high data usage, which can be perplexing for the average user.

Ransomware is another severe threat that involves the unauthorized encryption of data on a device, followed by a demand for a ransom in exchange for the decryption key. This type of attack can lock users out of their devices or make critical data inaccessible. Ransom demands are typically made in cryptocurrencies, adding another layer of complexity to the issue and making it difficult to track the perpetrators.

Zero-day exploits take advantage of unknown vulnerabilities in software or operating systems before developers become aware and issue a security patch. These vulnerabilities can be exploited undetected for extended periods, providing attackers with ample opportunity to exploit the flaw.

Cyber espionage involves sophisticated, often state-sponsored attacks targeting mobile devices for corporate or governmental espionage. These attacks might involve a combination of malware, phishing, and zero-day exploits. Characterized by their stealth and persistence, these attacks can remain undetected while collecting sensitive data over extended periods of time. They often involve complex, multi-stage strategies, including initial infiltration, lateral movement within a network, data extraction, and maintaining long-term access for ongoing espionage.

Table 4. Comparing different types of mobile security threats

Threat Type	Description	Primary Goal	Common Indicators
Malware Attacks	Includes viruses, worms, Trojans, and spyware, capable of stealing or corrupting data and taking control of the device.	Data theft, device control, or disruption	Unexpected ads, system slowdown, unauthorized data access
Phishing Attacks	Deceptive tactics to trick users into disclosing personal information, occurring via SMS, email, or malicious websites.	Information theft (credentials, personal data)	Suspicious messages, emails, unusual login requests
Network Spoofing and Man-in-the-Middle Attacks	Creating fake networks or intercepting communications to steal or manipulate data transmitted over mobile networks.	Data interception and theft	Unsecured Wi-Fi connections, unusual data patterns
Cryptojacking	Unauthorized use of a device's resources to mine cryptocurrency, leading to performance degradation and energy drain.	Resource exploitation for profit	Sluggish device performance, overheating, high data usage
Ransomware	Involves locking a device or encrypting data and demanding a ransom for restoration.	Financial gain through extortion	Data inaccessibility, ransom demands
Zero-Day Exploits	Exploitation of unknown software vulnerabilities before they are patched.	Exploitation of unpatched vulnerabilities	No immediate indicators, discovered post-attack
Cyber Espionage	Advanced attacks, often state-sponsored, targeting devices for espionage, using a mix of malware, phishing, and exploits.	Data theft and long-term surveillance	No immediate indicators, stealthy and persistent

4 MACHINE LEARNING AND ITS APPLICATION IN SECURITY

The vast digital landscape of today, characterized by exponentially growing data volumes and intricate interconnected systems, has birthed challenges that traditional computational methods struggle to address effectively. ML, a subset of artificial intelligence, has emerged as a transformative solution, leading revolutionary changes across various sectors, including cybersecurity [84]. With its capacity to learn patterns from vast datasets, make predictions, and adapt dynamically,

ML provides a robust framework for identifying and mitigating security threats in real-time. This section delves deep into the fundamentals of ML, exploring its numerous applications in the realm of cybersecurity and elucidating how it stands as a beacon of hope against the escalating complexity of cyber-attacks in the contemporary era.

4.1 Introduction to machine learning

Machine learning, at its core, is an interdisciplinary field that intersects the boundaries of computer science, mathematics, and statistics. It aims to develop algorithms that enable computers to learn from data and make decisions based on it [85]. Rather than being explicitly programmed for a specific task, these algorithms leverage vast datasets to infer patterns and deduce rules, subsequently applying this acquired knowledge to new, unseen data. The genesis of ML can be traced back to the mid-20th century, with the pioneering work of Alan Turing, who postulated the concept of a machine that could simulate any human intelligence. This laid the groundwork for what would become a transformative paradigm in computational theory.

The rise of ML in contemporary times can be attributed to three critical factors: the abundance of available data, powerful computational infrastructure, and advanced algorithmic innovations [86]. The digital age, marked by the proliferation of Internet-enabled devices and sophisticated sensor networks, generates petabytes of data daily, providing ample raw material for ML algorithms to train on. Additionally, the emergence of cloud computing and GPU-accelerated hardware has made it possible to process these extensive datasets in real time, enabling dynamic, on-the-fly decision-making to become a tangible reality. The algorithmic space of ML is vast, ranging from linear regressions and decision trees to intricate neural networks, each designed for specific applications, challenges, and data structures.

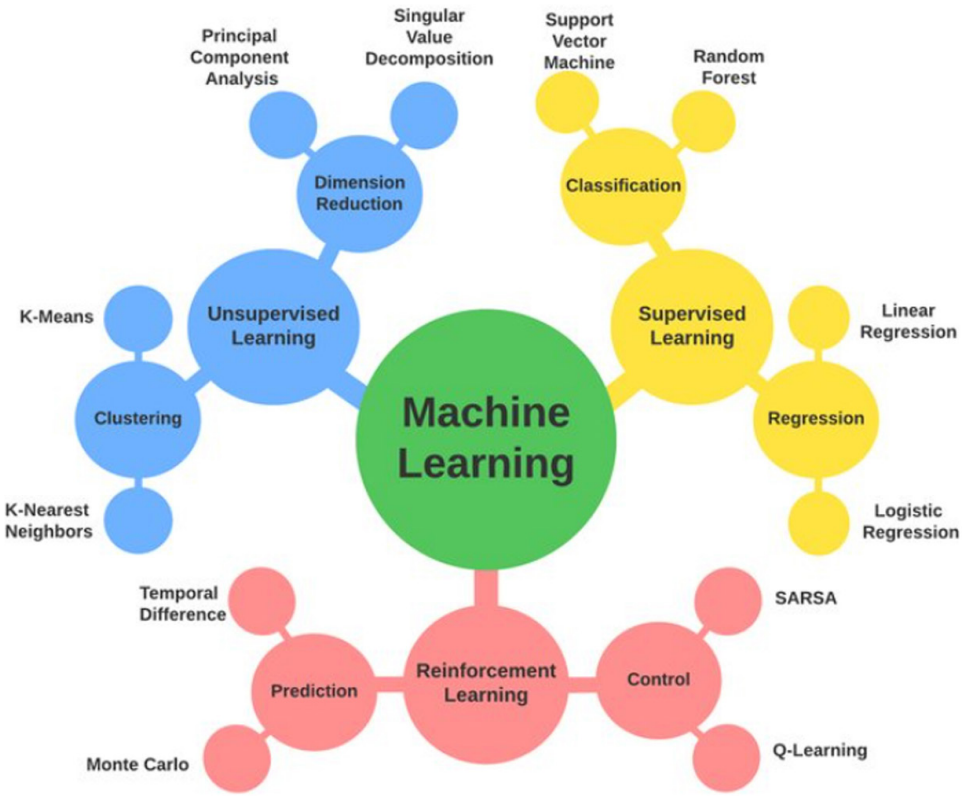

Fig. 1. Machine learning techniques (Koblah et al., 2023 [85])

4.2 Machine learning in cybersecurity: an overview

Machine learning's emergence in the cybersecurity domain has marked a significant shift in how security solutions are formulated, executed, and evaluated. In leading this transformation, ML provides the capability to automate the complex task of analyzing large datasets for irregular patterns indicative of cyber threats. These capabilities contrast with traditional security systems, which are primarily based on static, rule-based methods that flag known malicious signatures. Such traditional methods are becoming less effective in the face of increasingly complex and varied cyberattacks, as shown in Figure 2 [87].

Machine learning's inherent dynamism and ability to generalize from training data enable it to identify novel threats by recognizing patterns that deviate from established norms. For example, ML-based intrusion detection systems can autonomously comprehend 'normal' network behavior and identify abnormal data packets, even if these particular signatures were not included in the training data [88].

Fig. 2. Types of attack

Moreover, ML's role in cybersecurity extends beyond simple threat detection. Current innovations focus on utilizing ML to predict future threats by scrutinizing historical data and identifying trends in attack vectors. In the field of digital forensics, ML helps in the aggregation and correlation of disparate data, thereby streamlining post-incident analysis and resolution [89].

As cyber adversaries adapt and become more sophisticated, they are increasingly focusing on compromising ML models themselves through adversarial attacks. These efforts, falling under the sub-domain known as adversarial ML, introduce a new level of complexity in the integration of ML within cybersecurity solutions. This ongoing competition between cybersecurity professionals and malicious actors highlights the need for continuous research and innovation in utilizing ML to enhance cybersecurity measures [90]. Table 5 explores the transformative impact of ML in the field of cybersecurity, emphasizing its role from dynamic threat analysis to addressing adversarial challenges.

Table 5. Transformative impact of machine learning in cybersecurity: From dynamic threat analysis to adversarial challenges

Aspect	Description	Impact on Cybersecurity
Emergence of ML	Shift from static, rule-based methods to dynamic, data-driven analysis	Enhanced capability to detect and analyze complex cyber threats
Anomaly Detection	Ability to recognize deviations from normal patterns	Early detection of novel threats, reducing reliance on known threat signatures
Predictive Analysis	Use of historical data to predict future threats	Proactive approach in cybersecurity, identifying potential future attack vectors
Digital Forensics	Aggregation and correlation of data for post-incident analysis	Streamlined incident response and resolution
Adversarial ML	Focus on compromising ML models through sophisticated attacks	Increased complexity in cybersecurity, necessitating continuous innovation

4.3 Benefits of machine learning for mobile attack detection

Machine learning, with its ability to extract patterns and insights from large datasets, has solidified its position as an invaluable asset in the extensive domain of cybersecurity, especially in the realm of mobile attack detection [91]. However, the ML process is illustrated in Figure 3.

Mobile devices, due to their widespread connectivity and multifunctional capabilities, are constantly interacting with various networks and processing diverse data. This creates a vast and intricate data space that would be nearly impossible to effectively oversee when manually scrutinized. ML algorithms, however, are adept at autonomously sifting through such massive data streams, efficiently extracting salient features, and detecting subtle, often concealed, anomalous patterns indicative of potential threats [92].

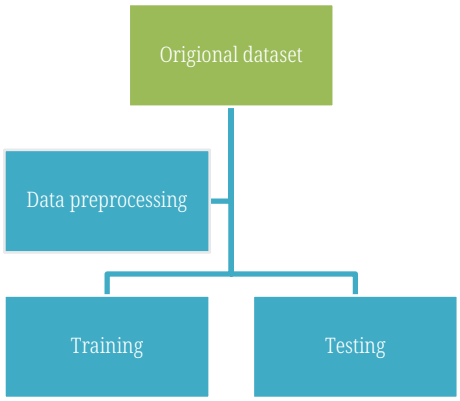

Fig. 3. Machine learning process

Furthermore, the dynamism inherent in mobile environments—the continuous influx of new applications, updates, and configurations—demands a security mechanism that is equally adaptive. ML models, continuously trained on evolving datasets, can automatically recalibrate their threat detection criteria. This adaptability stands in sharp contrast to conventional rule-based systems that remain static unless manually updated, making them susceptible to newer forms of attacks [93].

Another cardinal advantage is ML's is its capability for zero-day attack detection. Traditional signature-based detection systems are ill-equipped to identify novel threats for which no prior signature exists. However, ML models, especially those focused on anomaly detection, identify threats by detecting deviations from established behavioral patterns instead of relying on known signatures [94].

Moreover, with the proliferation of diverse mobile device manufacturers, models, and operating systems, achieving a universal security solution becomes a monumental challenge. ML's ability to generalize from specific training instances to broader contexts becomes invaluable here. A model trained on data from a specific subset of devices can often detect threats across a wider range of devices, provided appropriate feature engineering and algorithmic design [95] [96]. Table 6 highlights the benefits of ML in improving mobile attack detection, underscoring its crucial role in addressing the intricacies of the contemporary mobile security environment.

Table 6. Advantages of machine learning in enhancing mobile attack detection: Navigating the complexity of mobile security landscape

Benefit	Description	Relevance to Mobile Security
Data Handling Capability	Efficient analysis of large, diverse data streams	Effective oversight in complex mobile environments
Adaptability	Continuous recalibration of threat detection criteria	Responsive to evolving mobile threats and configurations
Zero-Day Attack Detection	Anomaly detection for unknown threats	Enhanced resilience against novel cyberattacks
Generalization Across Devices	Applicability of models to a range of devices	Universal mobile security solutions for diverse hardware
User Feedback Integration	Refinement of models through user input	Continuously improving detection accuracy

4.4 Commonly used machine learning algorithms in security

The vast and evolving landscape of cybersecurity necessitates an equally dynamic and varied toolkit to combat the myriad of threats. ML, with its diverse array of algorithms, presents a promising reservoir of techniques that can be fine-tuned and tailored to address specific security challenges [97].

One of the foundational algorithms utilized for security purposes is the decision tree. Decision trees are simplistic in structure yet robust in application. They dissect the data space by making hierarchical decisions based on feature values. Their visual and interpretable nature makes them particularly appealing for security tasks where explicability is paramount, such as rule-based intrusion detection and policy formulation. Beyond decision trees, random forests have gained traction due to their ensemble nature. By constructing a multitude of trees on varied data subsets and averaging or voting on their outputs, random forests curtail overfitting and enhance generalization [98].

Neural networks, particularly DNNs, represent another paradigm that has been effectively applied to security. With their layered architectures and capacity to model intricate nonlinear relationships, DNNs have exhibited proficiency in tasks that demand high granularity, such as image-based biometric authentication or real-time

traffic analysis for intrusion detection. SVMs, rooted in the principle of maximizing the margin between classes, have been instrumental in binary classification tasks related to security. Due to their capability to operate in high-dimensional spaces and their resistance to overfitting, SVMs have been utilized for tasks such as malware detection and intrusion detection in network traffic [99].

K-nearest neighbors (KNN) stands as another pivotal algorithm in the security domain. By classifying data points based on the majority class of their 'k' closest neighbors, KNN provides a non-parametric approach to anomaly detection. Its application spans areas such as system behavior profiling and real-time network traffic monitoring. Naive Bayes, a probabilistic classifier based on Bayes' theorem, provides a rapid and efficient approach for multiclass classification problems in security. Its prowess is especially evident in text-based tasks, making it suitable for email spam detection, phishing email categorization, and other content-based security challenges [100].

Lastly, clustering algorithms such as K-Means and DBSCAN have been utilized for unsupervised anomaly detection. By segmenting data into distinct clusters, these techniques help identify anomalous data points that deviate significantly from established clusters. This proves invaluable in scenarios like network intrusion detection where labeled data might be scarce, but the need to identify outlier behavior remains crucial [101]. Table 7 presents a comparative analysis of various ML algorithms, highlighting their distinct characteristics and specific applications in the cybersecurity domain.

Table 7. Analysis of diverse machine learning algorithms in cybersecurity characteristics and applications

Algorithm	Characteristics	Applications in Cybersecurity
Decision Trees	Hierarchical decision-making based on features	Intrusion detection, policy formulation
Random Forests	Ensemble of decision trees for improved accuracy	Minimizing overfitting, enhancing generalization
Neural Networks	Modeling complex relationships, high granularity	Biometric authentication, real-time intrusion detection
Support Vector Machines (SVMs)	Maximizing the margin between classes	Malware detection, network traffic classification
K-Nearest Neighbors (KNN)	Classification based on closest data points	System behavior profiling, network monitoring
Naive Bayes	Fast, effective multiclass classification	Spam detection, phishing email categorization
Clustering Algorithms	Segmenting data into distinct clusters	Unsupervised anomaly detection, outlier identification

Support vector machines. Support vector machines are considered one of the most respected ML models, particularly in the fields of classification and regression tasks [102]. Introduced in the 1990s, SVMs aim to find the optimal hyperplane that effectively separates data into distinct classes. In the context of a two-dimensional dataset, this hyperplane can be envisioned as a line; however, in higher dimensions, it becomes a multidimensional plane or a "hyperplane."

The guiding principle of SVMs lies in maximizing the margin between the two classes. The margin represents the distance between the hyperplane and the closest data point from either class. By optimizing this margin, SVM ensures that it finds the

most robust and generalizable decision boundary. This is achieved using support vectors, which are the critical data points lying closest to the decision boundary.

For linearly separable data, SVMs work flawlessly in determining a linear decision boundary. However, real-world data, especially in cybersecurity, often exhibits non-linear patterns. To address such complexities, SVM utilizes a technique known as the "kernel trick." By mapping the original feature space to a higher-dimensional space, kernel methods enable SVM to identify non-linear decision boundaries.

In the context of cybersecurity, SVMs have garnered considerable attention due to their inherent ability to handle high-dimensional data and their resilience against overfitting, provided the appropriate choice of parameters is made. For instance, SVMs have been extensively employed in intrusion detection systems, where they excel at classifying network traffic as benign or malicious based on intricate patterns. Moreover, in malware classification tasks, SVMs, with their high-dimensional feature handling capability, are adept at distinguishing between benign and malicious software based on their behavioral or structural attributes.

Furthermore, the mathematical foundation of SVMs provides a clear understanding of the decision boundary. Visualizing this boundary can offer critical insights into the nature of cyber threats. Such insights can be invaluable for cybersecurity professionals aiming to strengthen defense mechanisms or develop countermeasures against emerging threats. Table 8 explores the role and significance of SVMs in cybersecurity applications, providing a detailed examination of how this specific ML model is used in the field.

Table 8. Support vector machines role and significance in cybersecurity applications

Feature	Description	Importance in Cybersecurity
Optimal Hyperplane	Segregates data into distinct classes	Robust and generalizable decision boundary
Handling Non-linear Patterns	Kernel trick for complex data patterns	Applicability to real-world, non-linear cybersecurity data
High Dimensional Data Handling	Effective in high-dimensional spaces	Suitable for intricate cybersecurity tasks

Ensemble methods. Ensemble methods have gained prominence in the field of ML for their ability to combine multiple algorithms or models to achieve superior predictive performance compared to what any individual model could achieve alone. At the heart of ensemble methods lies the belief that the collective wisdom of a group often outweighs the intelligence of a single member. This principle, when applied to machine learning, involves utilizing a "committee" of models to achieve more accurate, robust, and generalizable predictions.

There are several ways in which ensemble methods achieve this goal. One of the most straightforward techniques is "bagging" (Bootstrap aggregating), where multiple versions of a model are trained on different subsets of the training data and drawn with replacement. Once trained, each model casts a vote, and the majority decision is considered the final prediction. A prime example of bagging is the random forest algorithm, which consists of an ensemble of decision trees [103].

Boosting is another influential ensemble technique. Contrary to bagging, which trains each model independently, boosting iteratively trains models by placing greater emphasis on instances that were previously misclassified. This iterative correction ensures that subsequent models rectify the mistakes of their predecessors.

Algorithms such as AdaBoost and gradient boosting machines (GBMs) are popular implementations of this strategy.

"Stacking" or "stacked generalization" represents another ensemble approach where multiple diverse models are trained and their predictions are combined, often through another model (a "meta-learner"), to make the final prediction. This layering of models harnesses the strengths of each, mitigating individual weaknesses and often leading to superior predictive performance [104].

In the context of cybersecurity, ensemble methods provide a robust defense against the dynamic and evolving nature of cyber threats. In the field of intrusion detection, an ensemble approach can combine the strengths of different algorithms to maintain high detection rates while minimizing false positives. By utilizing diverse models, ensemble methods can better address the multifaceted nature of cyber threats, which often manifest in various patterns and behaviors. Moreover, the redundancy inherent in ensemble techniques provides a safeguard against potential model failures or vulnerabilities, ensuring a consistent level of security is maintained [105]. Table 9 explores the synergistic use of ensemble methods in cybersecurity, illustrating how they are employed to combat the continuously evolving landscape of cyber threats.

Table 9. A synergistic ensemble methods in cybersecurity to combat evolving cyber threats [103–105]

Method	Technique	Advantage in Cybersecurity
Bagging (e.g., Random Forest)	Training multiple models on data subsets	Collective decision-making for accuracy
Boosting	Iteratively training models, focusing on misclassifications	Rectifying mistakes, enhancing overall model accuracy
Stacking	Combining diverse models through a meta-learner	Harnessing strengths of multiple models, mitigating weaknesses

4.5 Semi-supervised and unsupervised learning in anomaly detection

Anomaly detection in cybersecurity presents a unique set of challenges due to the evolving nature of threats and the sheer volume of data that needs to be processed. Traditional supervised learning methods often require labeled data to be effective, which may not be feasible in many cybersecurity applications due to the scarcity of labeled malicious activities and the cost associated with manual labeling. Consequently, semi-supervised and unsupervised learning approaches have garnered significant attention in the realm of anomaly detection, primarily due to their capability to operate with limited labeled data or even entirely unlabeled data.

Semi-supervised learning, as the name suggests, leverages both labeled and unlabeled data for training. The foundational hypothesis behind this paradigm is that the underlying structure derived from a vast amount of unlabeled data, when combined with a smaller set of labeled instances, can significantly improve learning accuracy. One common strategy in semi-supervised learning is to initially use the labeled data to train a base model and then iteratively refine this model using the unlabeled data. This iterative process helps capture the intricate patterns and structures from unlabeled instances, enhancing the generalization capability of the model [106]. In the context of anomaly detection, this could involve utilizing a small set of known

attack signatures along with a larger dataset of network traffic to identify previously undiscovered threats.

On the other hand, unsupervised learning operates without any labeled data, relying solely on the intrinsic structure and relationships within the data. Clustering and association are two primary techniques within this domain. For anomaly detection, unsupervised techniques such as clustering can be utilized to group similar data instances together, with outliers or anomalies falling outside these clusters. Techniques such as K-means clustering or hierarchical clustering are often employed for such tasks [107]. The inherent challenge here is determining the boundary between normal and anomalous, which can be particularly complex given that what is considered "normal" may evolve over time.

Deep learning architectures, especially autoencoders, have also demonstrated potential in unsupervised anomaly detection. Autoencoders are neural networks trained to reconstruct their input data. During this process, they learn a compressed representation of the data. In anomaly detection, an autoencoder trained on "normal" data may struggle to accurately reconstruct anomalous data. Thus, reconstruction errors can be used as a metric to identify potential anomalies [108].

The surge in interest in these techniques can be attributed to the dynamic cybersecurity landscape. With new threats emerging daily, relying solely on labeled data (which represents known threats) can leave systems vulnerable to previously unseen attacks. Semi-supervised and unsupervised learning offer mechanisms to detect anomalies that deviate from established patterns, providing a more adaptive and proactive approach to threat detection. Table 10 presents an analysis of semi-supervised and unsupervised learning techniques and their application in anomaly detection within the cybersecurity domain.

Table 10. Semi-supervised and unsupervised learning in anomaly detection

Learning Type	Approach	Application in Cybersecurity
Semi-supervised Learning	Combines labeled and unlabeled data	Effective in limited labeled data scenarios, enhancing detection accuracy
Unsupervised Learning	Relies on data's intrinsic structure	Identifying anomalies in unlabeled datasets, useful in dynamic threat landscapes
Deep Learning (Autoencoders)	Neural networks for reconstructing input data	Anomaly detection through reconstruction errors

4.6 Design of machine learning algorithms in mobile security

The field of ML offers a diverse array of algorithms that can be tailored to address the specific challenges of mobile security. Key among these are supervised learning techniques such as SVM and random forests, which excel at classifying data and identifying potential threats based on historical patterns. For instance, SVMs can be utilized to differentiate between benign and malicious app behaviors, leveraging their capability to process high-dimensional data. Another crucial aspect is the utilization of unsupervised learning algorithms, such as K-means clustering and autoencoders, which are effective in anomaly detection. These algorithms can identify unusual patterns or deviations from the norm, which are indicative of new, previously unseen mobile attacks.

Algorithm 1: Support Vector Machines (SVM) to Identify Potential Threats

Data Preprocessing: Normalize and transform data into a suitable format.
Feature Selection: Identify and select relevant features for classification.
Kernel Choice: Select an appropriate kernel function (e.g., linear, polynomial, radial basis function).
Model Training: Train the SVM model using labeled data to find the hyperplane that best separates different classes.
Threat Identification: Use the trained SVM to classify new data points as benign or malicious.

Algorithm 2: Random Forests for Threat Detection

Data Preprocessing: Clean and prepare data for analysis.
Feature Selection: Choose relevant features from the dataset.
Building Trees: Create multiple decision trees using random subsets of features.
Model Training: Train each tree on different parts of the dataset.
Voting System: For new data, each tree votes, and the majority vote determines the classification (benign or malicious).

Algorithm 3: K-means Clustering for Anomaly Detection by Identifying Unusual Patterns

Select K Points: Choose K points as initial centroids.
Assign Clusters: Assign each data point to the nearest centroid, forming K clusters.
Recompute Centroids: Calculate new centroids as the mean of data points in each cluster.
Iterate: Repeat the assignment and centroid computation until convergence.
Anomaly Detection: Analyze clusters to identify outliers or unusual patterns indicative of attacks.

Algorithm 4: Autoencoders for Anomaly Detection Through Reconstruction Error

Encoder: Compress input data into a lower-dimensional representation.
Decoder: Attempt to reconstruct the original data from the compressed representation.
Training: Minimize the difference (error) between original and reconstructed data.
Anomaly Identification: High reconstruction error indicates an anomaly or unusual pattern.

Deep learning techniques, particularly CNNs and RNNs, have shown great promise in enhancing mobile security. CNNs, with their powerful feature extraction capabilities, can be used to analyze and interpret complex input patterns such as network traffic or system logs. RNNs, known for their capability to process sequential data, are particularly valuable in comprehending and forecasting attack sequences or behaviors over time. Further, feature selection plays a pivotal role in the effectiveness of ML algorithms in mobile security. The process involves identifying the most relevant features from vast datasets that significantly contribute to the accuracy of threat detection. This might include features such as application permission requests, network traffic characteristics, and behavioral patterns of users. Dimensionality reduction techniques, such as principal component analysis (PCA), are often employed to improve model performance by eliminating redundant or irrelevant features.

Algorithm 5: Convolutional Neural Networks (CNNs) for Analyze and Interpret Complex Input Patterns for Threat Detection

Convolution Layers: Apply convolutional operations to extract features from input data.
Pooling Layers: Reduce dimensionality while retaining important information.
Fully Connected Layers: Perform high-level reasoning based on extracted features.
Training: Train the network using labeled data to optimize weights.
Threat Analysis: Use the trained CNN to analyze new data for potential threats.

Algorithm 6: Recurrent Neural Networks (RNNs) for Process Sequential Data for Predicting Attack Sequences

Sequence Input: Feed sequences of data (e.g., network traffic logs) into the network.
Hidden State Updates: Update the hidden state based on current input and previous state.
Output Generation: Produce output at each step or at the end of the sequence.
Backpropagation Through Time: Train the network by adjusting weights to minimize prediction errors.
Attack Prediction: Use the RNN to predict or identify attack patterns in sequential data.

Algorithm 7: Principal Component Analysis (PCA) for Reduce Dimensionality for Improved Model Performance

Standardize Data: Scale the data so that each feature contributes equally.
Covariance Matrix Computation: Compute the covariance matrix to understand how features vary together.
Eigenvalue Decomposition: Find the principal components (eigenvectors) of the covariance matrix.
Feature Transformation: Transform the original features into a new space defined by the principal components.
Reduced Feature Set: Select a subset of principal components for further analysis.

Integrating these ML algorithms into mobile security systems requires careful consideration of the unique constraints and requirements of mobile environments. This involves optimizing algorithms for limited processing power and memory to ensure minimal impact on device performance. Efficient model training and updating mechanisms are essential to keeping pace with the rapidly evolving threat landscape.

Additionally, the demand for real-time processing capabilities is crucial in mobile security. This necessitates the development of algorithms that can make quick and accurate predictions, often requiring the implementation of edge computing paradigms where data processing is done locally on the device.

Table 11. Machine learning algorithms in mobile malware detection

Reference	Modality	Method	Remarks
Mughaid et al.	Machine Learning and Deep Learning	Simulator for NOMA, machine learning algorithms like Decision Trees, KNN, etc.	Outstanding performance with high accuracy, proposes methodology for cyberattack detection
De Araujo-Filho et al.	GAN-Based Intrusion Detection	GANs, Temporal Convolutional Networks, Self-Attention	More accurate and faster than baselines, suitable for edge servers
Kumari et al.	Machine Learning Approach	Continuous Authentication, Reduce Feature Elimination (RFE)	Promises reduced system cost and complexity, high accuracy in user recognition
Mehta et al.	Security Challenges and Solutions Review	Review of ADAS security challenges, attacks, countermeasures	Highlights need for ongoing research in vehicle technology security
Park et al.	Specification-Based Misbehavior Detection	Behavior rule specification, state machine for anomaly detection	Effective against false base stations, low overhead
Xu B.	Intrusion Detection System Design	Machine learning and data mining for intrusion detection	High detection accuracy, low false alarm rate, ensures security in teaching systems
Zhu et al.	Ensemble Learning Framework	Hybrid deep learning, feature extraction, multi-model ensemble	Addresses Android malware, novel fusion scheme, step-by-step model justification
Naser et al.	Systematic Review	Survey of techniques, signature-based to machine learning classification	Addresses mobile spyware threat, consolidates knowledge for future research
Mughaid et al.	Machine Learning and Deep Learning	Simulator for NOMA, machine learning algorithms like Decision Trees, KNN, etc.	Outstanding performance with high accuracy, proposes methodology for cyberattack detection
De Araujo-Filho et al.	GAN-Based Intrusion Detection	GANs, Temporal Convolutional Networks, Self-Attention	More accurate and faster than baselines, suitable for edge servers

(Continued)

Table 11. Machine learning algorithms in mobile malware detection *(Continued)*

Reference	Modality	Method	Remarks
Kumari et al.	Machine Learning Approach	Continuous Authentication, Reduce Feature Elimination (RFE)	Promises reduced system cost and complexity, high accuracy in user recognition
Mehta et al.	Security Challenges and Solutions Review	Review of ADAS security challenges, attacks, countermeasures	Highlights need for ongoing research in vehicle technology security
Park et al.	Specification-Based Misbehavior Detection	Behavior rule specification, state machine for anomaly detection	Effective against false base stations, low overhead
Xu B.	Intrusion Detection System Design	Machine learning and data mining for intrusion detection	High detection accuracy, low false alarm rate, ensures security in teaching systems
Zhu et al.	Ensemble Learning Framework	Hybrid deep learning, feature extraction, multi-model ensemble	Addresses Android malware, novel fusion scheme, step-by-step model justification
Rathore et al.	Reinforcement Learning Based Evasion Attacks and Defenses	Reinforcement learning evasion attacks, robustness analysis	Proactive framework, high fooling rate, proposes defense strategy
Ali et al.	Intrusion Detection for VANET	ML algorithms, feature selection techniques, stacking method	High detection accuracy, feasible for real-time environments
Hong et al.	Hybrid Jamming Detection Algorithm	Hybrid structure of classification and anomaly detection models	Superior performance over baseline, suitable for different scenarios
Prazeres et al.	Machine Learning in IDS Based on IoT Traffic	Fog computing, deep neural networks, anomaly detection	Flow-based anomaly detection, network traffic segmentation
Javed et al.	Motion-based Side-channel Attack Detection	Background application inferring keystrokes using sensors	High inference accuracy, evaluation of sensor combinations
Musikawan et al.	Deep Learning for Android Malware Detection	Improved deep neural network, ensemble classifier architecture	Superior performance, intensive evaluations
Vatambeti et al.	Dolphin Echo-location-based ML Model in MANET	Machine learning with Dolphin Echolocation model	Effective for black hole attack detection, energy-efficient
Bostani et al.	Evasion Attack on ML for Android Malware Detection	Problem-space adversarial attack, black-box Android malware detectors	High evasion rates, real-world adversarial examples
Iqbal et al.	Ransomware Detection for Healthcare Systems	Hybrid approach using static and dynamic techniques	High accuracy, addresses challenges in healthcare ransomware

5 EMERGING TRENDS AND TECHNOLOGIES IN MOBILE SECURITY

In the ever-evolving domain of mobile security, the emergence of transformative technologies in ML and DL, such as transformer-based models and attention mechanisms, is reshaping our approach to cybersecurity. These cutting-edge methodologies are at the forefront of the fight against mobile security threats, offering innovative solutions that could significantly enhance detection and prevention mechanisms.

Transformer-based models, originally celebrated for their breakthroughs in natural language processing, are now being leveraged in the cybersecurity field. Their primary strength lies in processing sequential data and enabling efficient parallel processing, which makes them especially skilled at analyzing intricate and dynamic security data. These models excel at identifying subtle patterns and

anomalies within large datasets, such as network traffic or user behavior logs, enabling early detection of sophisticated cyber threats that might elude traditional security systems. Moreover, the rapid processing capabilities of transformer models accelerate response times to potential threats, which is a critical factor in mitigating the impact of cyberattacks.

Complementing these models are attention mechanisms, which have revolutionized the way we process large volumes of data. By focusing on the most relevant parts of the data, attention mechanisms enhance the precision of threat detection in the vast sea of benign information. This feature not only allows for the development of customized security solutions to address specific organizational needs but also scales effectively to various types of mobile security threats. The predictive potential of these models is particularly promising, as they offer a shift from reactive to proactive security measures. By forecasting vulnerabilities and emerging threats, attention-based models enable mobile security systems to predict and mitigate cyberattacks before they occur.

The implications of integrating transformer-based models and attention mechanisms into mobile security are profound. They promise not only increased accuracy and efficiency in threat detection and response but also introduce novel approaches to mobile security. For instance, their application in behavioral analysis can lead to more effective identification of abnormal user activities, which are indicative of security breaches. This capability is particularly crucial in combating advanced threats such as zero-day exploits and sophisticated phishing attacks.

6 CONCLUSION

In this overview, we conducted a comprehensive examination of the ever-evolving field of mobile security, with a particular focus on the application of advanced intelligent techniques to enhance mobile attack detection and prevention mechanisms. The study was designed to explore the complex aspects of mobile security, emphasizing the importance of a detailed understanding of both the threat landscape and current defense strategies. Initially, we delved into the complex arena of mobile security threats, categorizing them from commonplace malware and phishing attacks to sophisticated cyber-espionage activities. This overview served as a prelude to the critical analysis of traditional mobile security methods, where we explained their advantages and disadvantages. This analysis was instrumental in laying the groundwork for the transition towards more avant-garde, AI-based strategies. A significant part of this research focused on exploring ML techniques, highlighting their potential to transform mobile security.

7 ACKNOWLEDGMENT

We thank Samir M. Abu Tahoun, Security Management Technology Group (SMT) (http://www.smtgroup.org/) for the financial support of our research project.

8 REFERENCES

[1] J. Smith, “The evolution of mobile technologies,” *Journal of Mobile Tech.*, vol. 5, no. 2, pp. 1–12, 2015.

[2] L. Turner, “Security threats in mobile platforms: An overview,” *Journal of Cybersecurity*, vol. 6, no. 1, pp. 15–29, 2017.

[3] S. K. Kapoor and T. L. Martin, "Sophisticated attacks on mobile devices and countermeasures," *Journal of Mobile Security*, vol. 10, no. 3, pp. 33–49, 2019.
[4] J. Smith and L. Anderson, "A decade of mobile evolution: Understanding the shift," *Journal of Digital Advancements*, vol. 10, no. 3, pp. 45–61, 2016.
[5] P. Thompson, "From communication to operation: Mobile devices in the 21st century," *TechSphere Journal*, vol. 4, no. 1, pp. 13–27, 2018.
[6] A. R. Johnson, "Mobile banking: Opportunities and threats," *Journal of Digital Banking*, vol. 6, no. 3, pp. 95–108, 2019.
[7] H. Patel and M. Lee, "Healthcare on the go: Understanding mobile health applications," *Journal of Medical Informatics*, vol. 11, no. 2, pp. 44–59, 2020.
[8] D. Kim and R. Joshi, "The dark side of mobile: A glimpse into mobile-based cyber threats," *Journal of Cybersecurity Research*, vol. 9, no. 4, pp. 85–100, 2019.
[9] S. K. Kapoor and T. L. Martin, "Exploring the depths of mobile threats: Advanced attack vectors and methodologies," *Journal of Mobile Security*, vol. 12, no. 1, pp. 1–17, 2021.
[10] M. Sharma, "Open-source and security: An analysis of android vulnerabilities," *Mobile Systems Quarterly*, vol. 5, no. 4, pp. 33–47, 2018.
[11] L. Matthews, "The challenge of rapid technological obsolescence in mobile security," *TechInsight Magazine*, vol. 6, no. 1, pp. 22–38, 2020.
[12] G. Reyes and K. Lim, "The future of mobile security: Challenges and opportunities," *Journal of Future Tech and Security*, vol. 8, no. 2, pp. 58–75, 2022.
[13] Y. Aafer, W. Du, and H. Yin, "DroidAPIMiner: Mining API-level features for robust malware detection in android," in *International Conference on Security and Privacy in Communication Systems, Springer*, 2013, vol. 127, pp. 86–103. https://doi.org/10.1007/978-3-319-04283-1_6
[14] V. M. Afonso *et al.*, "Identifying android malware using dynamically obtained features," *Journal of Computer Virology and Hacking Techniques*, vol. 11, no. 1, pp. 9–17, 2015. https://doi.org/10.1007/s11416-014-0226-7
[15] G. Reyes and K. Lim, "The future of mobile security: Challenges and opportunities," *Journal of Future Tech and Security*, vol. 8, no. 2, pp. 58–75, 2022.
[16] M. Roberts, "Conventional mobile security methods: Challenges and limitations," *Mobile Tech. Review*, vol. 7, no. 2, pp. 10–25, 2020.
[17] G. Piatetsky-Shapiro, "Discovery, analysis and presentation of strong rules," in *Knowledge Discovery in Databases*. Cambridge, MA: AAAI/MIT, 1991.
[18] S. Gold, "Android insecurity," *Network Security*, vol. 2011, no. 10, pp. 5–7, 2011. https://doi.org/10.1016/S1353-4858(11)70104-0
[19] Y. J. Ham, D. Moon, H. Lee, J. D. Lim, and J. N. Kim, "Android mobile application system call event pattern analysis for determination of malicious attack," *International Journal of Security and its Applications*, vol. 8, no. 1, pp. 231–246, 2014. http://dx.doi.org/10.14257/ijsia.2014.8.1.22
[20] S. Hou, A. Saas, L. Chen, Y. Ye, and T. Bourlai, "Deep neural networks for automatic android malware detection," in *ASONAM '17: Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2017, pp. 803–810. https://doi.org/10.1145/3110025.3116211
[21] S. Huda, R. Islam, J. Abawajy, J. Yearwood, M. M. Hassan, and G. Fortino, "A hybrid-multi filter-wrapper framework to identify run-time behaviour for fast malware detection," *Future Generation Computer Systems*, vol. 83, pp. 193–207, 2018. https://doi.org/10.1016/j.future.2017.12.037
[22] J. Smith and L. Anderson, "A Decade of mobile evolution: Understanding the shift," *Journal of Digital Advancements*, vol. 10, no. 3, pp. 45–61, 2016.
[23] P. Thompson, "From communication to operation: Mobile devices in the 21st century," *TechSphere Journal*, vol. 4, no. 1, pp. 13–27, 2018.

[24] G. Reyes and K. Lim, "The future of mobile security: Challenges and opportunities," *Journal of Future Tech and Security*, vol. 8, no. 2, pp. 58–75, 2022.
[25] T. Isohara, K. Takemori, and A. Kubota, "Kernel-based behavior analysis for android malware detection," in *Proceedings – 2011 7th International Conference on Computational Intelligence and Security*, 2011, pp. 1011–1015. https://doi.org/10.1109/CIS.2011.226
[26] H. Patel and M. Lee, "Healthcare on the go: Understanding mobile health applications," *Journal of Medical Informatics*, vol. 11, no. 2, pp. 44–59, 2020.
[27] D. Kim and R. Joshi, "The dark side of mobile: A glimpse into mobile-based cyber threats," *Journal of Cybersecurity Research*, vol. 9, no. 4, pp. 85–100, 2019.
[28] S. K. Kapoor and T. L. Martin, "Exploring the depths of mobile threats: Advanced attack vectors and methodologies," *Journal of Mobile Security*, vol. 12, no. 1, pp. 1–17, 2021.
[29] M. Sharma, "Open-source and security: An analysis of android vulnerabilities," *Mobile Systems Quarterly*, vol. 5, no. 4, pp. 33–47, 2018.
[30] A. R. Johnson, "Mobile banking: Opportunities and threats," *Journal of Digital Banking*, vol. 6, no. 3, pp. 95–108, 2019.
[31] S. K. Kapoor and T. L. Martin, "Exploring the depths of mobile threats: Advanced attack vectors and methodologies," *Journal of Mobile Security*, vol. 12, no. 1, pp. 1–17, 2021.
[32] M. Sharma, "Open-source and security: An analysis of android vulnerabilities," *Mobile Systems Quarterly*, vol. 5, no. 4, pp. 33–47, 2018.
[33] L. Matthews, "The challenge of rapid technological obsolescence in mobile security," *TechInsight Magazine*, vol. 6, no. 1, pp. 22–38, 2020.
[34] G. Reyes and K. Lim, "The future of mobile security: Challenges and opportunities," *Journal of Future Tech and Security*, vol. 8, no. 2, pp. 58–75, 2022.
[35] Y. J. Ham and H. Lee, "Detection of malicious android mobile applications based on aggregated system call events," *International Journal of Computer and Communication Engineering*, vol. 3, no. 2, pp. 149–154, 2014. https://doi.org/10.7763/IJCCE.2014.V3.310
[36] Y. J. Ham, D. Moon, H. Lee, J. D. Lim, and J. N. Kim, "Android mobile application system call event pattern analysis for determination of malicious attack," *International Journal of Security and its Applications*, vol. 8, no. 1, pp. 231–246, 2014. http://dx.doi.org/10.14257/ijsia.2014.8.1.22
[37] A. K. A. Hwaitat, A. Shaheen, K. Adhim, E. N. Arkebat, and A. A. A. Hwiatat, "Computer hardware components ontology," *Modern Applied Science*, vol. 12, no. 3, pp. 35–40, 2018. https://doi.org/10.5539/mas.v12n3p35
[38] S. Hou, A. Saas, L. Chen, Y. Ye, and T. Bourlai, "Deep neural networks for automatic android malware detection," in *ASONAM '17: Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2017, pp. 803–810. https://doi.org/10.1145/3110025.3116211
[39] S. Hou, A. Saas, Y. Ye, and L. Chen, "Droiddelver: An android malware detection system using deep belief network based on api call blocks," in *International Conference on Web-Age Information Management*, 2016, vol. 9998, pp. 54–66. https://doi.org/10.1007/978-3-319-47121-1_5
[40] C. Y. Huang, Y. T. Tsai, and C. H. Hsu, "Performance evaluation on permission-based detection for android malware," in *Smart Innovation, Systems and Technologies*, 2013, pp. 111–120. https://doi.org/10.1007/978-3-642-35473-1_12
[41] F. Idrees and M. Rajarajan, "Investigating the android intents and permissions for malware detection," in *International Conference on Wireless and Mobile Computing, Networking and Communications*, 2014, pp. 354–358.
[42] T. Isohara, K. Takemori, and A. Kubota, "Kernel-based behavior analysis for android malware detection," in *Proceedings – 2011 7th International Conference on Computational Intelligence and Security*, 2011, pp. 1011–1015. https://doi.org/10.1109/CIS.2011.226

[43] E. M. B. Karbab, M. Debbabi, A. Derhab, and D. Mouheb, "MalDozer: Automatic framework for android malware detection using deep learning," *Digital Investigation*, vol. 24, pp. S48–S59, 2018.

[44] D. Kim, J. Kim, and S. Kim, "A malicious application detection framework using automatic feature extraction tool on android market," in *3rd International Conference on Computer Science and Information Technology (ICCSIT)*, 2013, pp. 1–4.

[45] G. Kim, S. Lee, and S. Kim, "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1690–1700, 2014. http://dx.doi.org/10.1016/j.eswa.2013.08.066

[46] K. Tran *et al.*, "Towards a feature rich model for predicting spam emails containing malicious attachments and urls," in *Eleventh Australasian Data Mining Conference*, 2014.

[47] T. Kim, B. Kang, M. Rho, S. Sezer, and E. G. Im, "A multimodal deep learning method for android malware detection using various features," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, pp. 773–788, 2019. https://doi.org/10.1109/TIFS.2018.2866319

[48] A. Kumar, K. S. Kuppusamy, and G. Aghila, "FAMOUS: Forensic analysis of mobile devices using scoring of application permissions," *Future Generation Computer Systems*, vol. 83, pp. 158–172, 2018.

[49] R. Kumar, Z. Xiaosong, R. U. Khan, I. Ahad, and J. Kumar, "Malicious code detection based on image processing using deep learning," in *Proceedings of the 2018 International Conference on Computing and Artificial Intelligence (ICCAI '18)*, 2018, pp. 81–85. https://doi.org/10.1145/3194452.3194459

[50] S. Malik and K. Khatter, "System call analysis of android malware families," *Indian Journal of Science and Technology*, vol. 9, no. 21, pp. 1–13, 2016. https://doi.org/10.17485/ijst/2016/v9i21/90273

[51] S. Mansfield-Devine, "Android malware and mitigations," *Network Security*, vol. 2012, no. 11, pp. 12–20, 2012. https://doi.org/10.1016/S1353-4858(12)70104-6

[52] S. Mansfield-Devine, "Paranoid android: Just how insecure is the most popular mobile platform?" *Network Security*, vol. 2012, no. 9, pp. 5–10, 2012. https://doi.org/10.1016/S1353-4858(12)70081-8

[53] N. McLaughlin *et al.*, "Deep android malware detection," in *Proceedings of the Seventh ACM on Conference on Data and Application Security and Privacy (CODASPY '17)*, 2017, pp. 301–308. https://doi.org/10.1145/3029806.3029823

[54] R. Nix and J. Zhang, "Classification of android apps and malware using deep neural networks," in *Proceedings of the International Joint Conference on Neural Networks*, 2017, pp. 1871–1878. https://doi.org/10.1109/IJCNN.2017.7966078

[55] J. S. Park *et al.*, "Smart contract-based review system for an IoT data marketplace," *Sensors*, vol. 18, no. 10, p. 3577, 2018. https://doi.org/10.3390/s18103577

[56] W. Peng *et al.*, "Enhancing the Naive Bayes spam filter through intelligent text modification detection," in *Proceedings – 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications and 12th IEEE International Conference on Big Data Science and Engineering (Trustcom/BigDataSE)*, 2018, pp. 849–854. https://doi.org/10.1109/TrustCom/BigDataSE.2018.00122

[57] A. Rodriguez-Mota *et al.*, "Towards a 2-hybrid android malware detection test framework," in *2016 International Conference on Electronics, Communications and Computers (CONIELECOMP)*, 2016, pp. 54–61. https://doi.org/10.1109/CONIELECOMP.2016.7438552

[58] A. Saracino *et al.*, "MADAM: Effective and efficient behavior-based android malware detection and prevention," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 1, pp. 83–97, 2018. https://doi.org/10.1109/TDSC.2016.2536605

[59] S. H. Seo *et al.*, "Detecting mobile malware threats to homeland security through static analysis," *Journal of Network and Computer Applications*, vol. 38, pp. 43–53, 2014. https://doi.org/10.1016/j.jnca.2013.05.008

[60] P. V. Shijo and A. Salim, "Integrated static and dynamic analysis for malware detection," *Procedia Computer Science*, vol. 46, pp. 804–811, 2015. https://doi.org/10.1016/j.procs.2015.02.149

[61] F. Tchakounté and P. Dayang, "System calls analysis of malwares on android," *International Journal of Science and Technology*, vol. 2, no. 9, pp. 669–674, 2013.

[62] S. Venkatraman and M. Alazab, "Use of data visualisation for zero-day malware detection," *Security and Communication Networks*, vol. 2018, pp. 1–13, 2018. https://doi.org/10.1155/2018/1728303

[63] A. K. Hwaitat, S. Manaseer, R. M. Al-Sayyed, M. Almaiah, and O. Almomani, "An investigator digital forensics frequencies particle swarm optimization for detection and classification of APT attack in fog computing environment (IDF-FPSO)," *Journal of Theoretical and Applied Information Technology*, vol. 98, no. 7, pp. 937–952, 2020.

[64] Z. Wang, J. Cai, S. Cheng, and W. Li, "Droid deep learner: Identifying Android malware using deep learning," in *2016 IEEE 37th Sarnoff Symposium*, 2016, pp. 160–165. https://doi.org/10.1109/SARNOF.2016.7846747

[65] H. N. Fakhouri *et al.*, "Improved path testing using multi-verse optimization algorithm and the integration of test path distance," *International Journal of Interactive Mobile Technologies*, vol. 17, no. 20, pp. 38–59, 2023. https://doi.org/10.3991/ijim.v17i20.37517

[66] S. Y. Yerima and S. Sezer, "Droidfusion: A novel multilevel classifier fusion approach for android malware detection," *IEEE Transactions on Cybernetics*, vol. 49, no. 2, pp. 453–466, 2019. https://doi.org/10.1109/TCYB.2017.2777960

[67] F. Idrees and M. Rajarajan, "Investigating the android intents and permissions for malware detection," in *International Conference on Wireless and Mobile Computing, Networking and Communications*, 2014, pp. 354–358.

[68] W. Yu, L. Ge, G. Xu, and X. Fu, "Towards neural network-based malware detection on android mobile devices," in *Cybersecurity Systems for Human Cognition Augmentation*, 2014, vol. 61, pp. 99–117. https://doi.org/10.1007/978-3-319-10374-7_7

[69] S. Manaseer and A. K. Al Hwaitat, "Centralized web application firewall security system," *Modern Applied Science*, vol. 12, no. 10, p. 164, 2018. https://doi.org/10.5539/mas.v12n10p164

[70] Y. Zhang, Y. Yang, and X. Wang, "A novel android malware detection approach based on convolutional neural network," in *Proceedings of the 2nd International Conference on Cryptography, Security and Privacy (ICCSP)*, 2018, pp. 144–149. https://doi.org/10.1145/3199478.3199492

[71] Y. Zhou, Z. Wang, W. Zhou, and X. Jiang, "Hey, you, get off of my market: Detecting malicious apps in official and alternative android markets," in *NDSS (Network and Distributed System Security Symposium)*, 2012.

[72] D. Zhu, H. Jin, Y. Yang, D. Wu, and W. Chen, "DeepFlow: Deep learning-based malware detection by mining android application for abnormal usage of sensitive data," in *Proceedings – IEEE Symposium on Computers and Communications*, 2017.

[73] A. K. Al Hwaitat, M. A. Almaiah, A. Ali, S. Al-Otaibi, R. Shishakly, A. Lutfi, and M. Alrawad, "A new blockchain-based authentication framework for secure IoT networks," *Electronics*, vol. 12, no. 17, p. 3618, 2023. https://doi.org/10.3390/electronics12173618

[74] L. Kappelman *et al.*, "The 2019 SIM IT issues and trends study," *MIS Quarterly Executive*, vol. 19, no. 1, pp. 69–104, 2020. https://doi.org/10.17705/2msqe.00026

[75] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2016. https://doi.org/10.1109/COMST.2015.2494502

[76] D. Schatz, R. Bashroush, and J. Wall, "Towards a more representative definition of cybersecurity," *Journal of Digital Forensics, Security and Law*, vol. 12, no. 2, pp. 53–74, 2017. https://doi.org/10.15394/jdfsl.2017.1476

[77] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015. https://doi.org/10.1038/nature14539

[78] F. Hamad, M. Al-Fadel, and H. Fakhouri, "The effect of librarians' digital skills on technology acceptance in academic libraries in Jordan," *Journal of Librarianship and Information Science*, vol. 53, no. 4, pp. 589–600, 2021. https://doi.org/10.1177/0961000620966644

[79] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and other Botnets," *Computer*, vol. 50, no. 7, pp. 80–84, 2017. https://doi.org/10.1109/MC.2017.201

[80] McAfee, "Mobile threat report." [Online]. Available: https://www.mcafee.com/enterprise/en-us/threat-center/mcafeelabs/reports.html. [Accessed: Feb. 2, 2021].

[81] H. N. Fakhouri, S. Alawadi, F. M. Awaysheh, I. B. Hani, M. Alkhalaileh, and F. Hamad, "A comprehensive study on the role of machine learning in 5G security: Challenges," *Technologies, and Solutions. Electronics*, vol. 12, no. 22, p. 4604, 2023. https://doi.org/10.3390/electronics12224604

[82] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019. https://doi.org/10.1109/COMST.2019.2904897

[83] M. A. Al-Garadi *et al.*, "A survey of machine and deep learning methods for Internet of Things (IoT) security," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1646–1685, 2020. https://doi.org/10.1109/COMST.2020.2988293

[84] A. Abeshu and N. Chilamkurti, "Deep learning: The frontier for distributed attack detection in Fog-to-Things computing," *IEEE Communications Magazine*, vol. 56, no. 2, pp. 169–175, 2018. https://doi.org/10.1109/MCOM.2018.1700332

[85] D. Koblah, R. Acharya, D. Capecci, O. Dizon-Paradis, S. Tajik, F. Ganji, D. Woodard, and D. Forte, "A survey and perspective on artificial intelligence for security-aware electronic design automation," *ACM Transactions on Design Automation of Electronic Systems*, vol. 28, no. 2, pp. 1–57, 2023. https://doi.org/10.1145/3563391

[86] R. Pascanu *et al.*, "Malware classification with recurrent networks," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 1916–1920. https://doi.org/10.1109/ICASSP.2015.7178304

[87] P. Torres *et al.*, "An analysis of recurrent neural networks for Botnet detection behavior," in *2016 IEEE Biennial Congress of Argentina (ARGENCON)*, 2016, pp. 1–6. https://doi.org/10.1109/ARGENCON.2016.7585247

[88] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2005, vol. 1, pp. 539–546.

[89] S. Servia-Rodriguez *et al.*, "Personal model training under privacy constraints," in *Proc. 3rd ACM/IEEE Int. Conf. Internet-of-Things Design and Implementation (IoTDI)*, 2018, pp. 1–11.

[90] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SigKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2010. https://doi.org/10.1145/1964897.1964918

[91] D. Newman, "Bag of words data set," *UCI Machine Learning Repository*, 2008. https://archive.ics.uci.edu/ml/datasets/Bag+of+Words

[92] "Wikipedia dataset," 2020. https://dumps.wikimedia.your.org/

[93] J. Wang *et al.*, "Not just privacy: Improving performance of private deep learning in mobile cloud," in *Proc. 24th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD '18)*, 2018, pp. 2407–2416. https://doi.org/10.1145/3219819.3220106

[94] L. Lyu *et al.*, "Fog-embedded deep learning for the Internet of Things," *IEEE Trans. Industrial Informatics*, vol. 15, no. 7, pp. 4206–4215, 2019. https://doi.org/10.1109/TII.2019.2912465

[95] G. Roig *et al.*, "Conditional random fields for multi-camera object detection," in *Proc. 2011 Int. Conf. Computer Vision (ICCV)*, 2011, pp. 563–570. https://doi.org/10.1109/ICCV.2011.6126289

[96] R. Gilad-Bachrach *et al.*, "CryptoNets: Applying neural networks to encrypted data with high throughput and accuracy," in *Proc. 33rd Int. Conf. Machine Learning, vol. 48 of JMLR Workshop and Conf. Proc.*, 2016, pp. 201–210.

[97] H. N. Fakhouri, F. Hamad, and A. Alawamrah, "Success history intelligent optimizer," *The Journal of Supercomputing*, vol. 78, pp. 6461–6502, 2022. https://doi.org/10.1007/s11227-021-04093-9

[98] H. Maghrebi, T. Portigliatti, and E. Prouff, "Breaking cryptographic implementations using deep learning techniques," in *Proc. Int. Conf. on Security, Privacy, and Applied Cryptography Engineering*, 2016, vol. 10076, pp. 3–26. https://doi.org/10.1007/978-3-319-49445-6_1

[99] H. N. Fakhouri, A. Hudaib, and A. Sleit, "Multivector particle swarm optimization algorithm," *Soft Computing*, vol. 24, pp. 11695–11713, 2020. https://doi.org/10.1007/s00500-019-04631-x

[100] M. Zhang, Y. Duan, H. Yin, and Z. Zhao, "Semantics-aware Android malware classification using weighted contextual API dependency graphs," in *Proc. of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS'14)*, 2014, pp. 1105–1116. https://doi.org/10.1145/2660267.2660359

[101] Z. Yuan, Y. Lu, Z. Wang, and Y. Xue, "Droid-sec: Deep learning in Android malware detection," in *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4, pp. 371–372, 2014. https://doi.org/10.1145/2740070.2631434

[102] W. Y. Lee, J. Saxe, and R. Harang, "SeqDroid: Obfuscated Android malware detection using stacked convolutional and recurrent neural networks," in *Deep Learning Applications for Cyber Security*, 2019, pp. 197–210. https://doi.org/10.1007/978-3-030-13057-2_9

[103] S. S. C. Silva, R. M. P. Silva, R. C. G. Pinto, and R. M. Salles, "Botnets: A survey," *Computer Networks*, vol. 57, no. 2, pp. 378–403, 2013. https://doi.org/10.1016/j.comnet.2012.07.021

[104] S. N. Fakhouri, A. Hudaib, and H. N. Fakhouri, "Enhanced optimizer algorithm and its application to software testing," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 32, no. 6, pp. 885–907, 2020. https://doi.org/10.1080/0952813X.2019.1694591

[105] P. Lestringant, F. Guihéry, and P. -A. Fouque, "Automated identification of cryptographic primitives in binary code with data flow graph isomorphism," in *ASIA CCS '15: Proc. of the 10th ACM Symposium on Information, Computer and Communications Security*, 2015, pp. 203–214. https://doi.org/10.1145/2714576.2714639

[106] M. Naser, H. Al Bazar, and H. Abdel-Jaber, "Mobile spyware identification and categorization: A systematic review," *Informatica*, vol. 47, no. 8, pp. 45–56, 2023. https://doi.org/10.31449/inf.v47i8.4881

[107] A. Mughaid *et al.*, "Improved dropping attacks detecting system in 5G networks using machine learning and deep learning approaches," *Multimedia Tools and Applications*, vol. 82, pp. 13973–13995, 2023. https://doi.org/10.1007/s11042-022-13914-9

[108] P. F. De Araujo-Filho *et al.*, "Unsupervised GAN-Based intrusion detection system using temporal convolutional networks and self-attention," *IEEE Transactions on Network and Service Management*, vol. 20, no. 4, pp. 4951–4963, 2023. https://doi.org/10.1109/TNSM.2023.3260039

[109] S. Kumari *et al.*, "A novel approach for continuous authentication of mobile users using reduce feature elimination (RFE): A machine learning approach," *Mobile Networks and Applications*, vol. 28, pp. 767–781, 2023. https://doi.org/10.1007/s11036-023-02103-z

[110] A. A. Mehta *et al.*, "Securing the future: A comprehensive review of security challenges and solutions in advanced driver assistance systems," *IEEE Access*, vol. 12, pp. 643–678, 2024. https://doi.org/10.1109/ACCESS.2023.3347200

[111] H. Park *et al.*, "SMDFbs: Specification-based misbehavior detection for false base stations," *Sensors*, vol. 23, no. 23, p. 9504, 2023. https://doi.org/10.3390/s23239504

[112] B. Xu, "Design of intrusion detection system for intelligent mobile network teaching," *Computers and Electrical Engineering*, vol. 112, p. 109013, 2023. https://doi.org/10.1016/j.compeleceng.2023.109013

[113] H.-J. Zhu *et al.*, "A multi-model ensemble learning framework for imbalanced android malware detection," *Expert Systems with Applications*, vol. 234, p. 120952, 2023. https://doi.org/10.1016/j.eswa.2023.120952

[114] H. Rathore, A. Nandanwar, S. K. Sahay, and M. Sewak, "Adversarial superiority in android malware detection: Lessons from reinforcement learning based evasion attacks and defenses," *Forensic Science International: Digital Investigation*, vol. 44, p. 301511, 2023. https://doi.org/10.1016/j.fsidi.2023.301511

[115] W. A. Ali, M. Roccotelli, G. Boggia, and M. P. Fanti, "Intrusion detection system for vehicular ad hoc network attacks based on machine learning techniques," *Information Security Journal: A Global Perspective*, pp. 1–19, 2024. https://doi.org/10.1080/19393555.2024.2307638

[116] S. Hong, K. Kim, and S.-H. Lee, "A hybrid jamming detection algorithm for wireless communications: Simultaneous classification of known attacks and detection of unknown attacks," *IEEE Communications Letters*, vol. 27, no. 7, pp. 1769–1773, 2023. https://doi.org/10.1109/LCOMM.2023.3275694

[117] N. Prazeres, R. L. D. C. Costa, L. Santos, and C. Rabadão, "Engineering the application of machine learning in an IDS based on IoT traffic flow," *Intelligent Systems with Applications*, vol. 17, p. 200189, 2023. https://doi.org/10.1016/j.iswa.2023.200189

[118] A. R. Javed, M. O. Beg, M. Asim, T. Baker, and A. H. Al-Bayatti, "AlphaLogger: Detecting motion-based side-channel attack using smartphone keystrokes," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 4869–4882, 2023. https://doi.org/10.1007/s12652-020-01770-0

[119] P. Musikawan, Y. Kongsorot, I. You, and C. So-In, "An enhanced deep learning neural network for the detection and identification of Android malware," *IEEE Internet of Things Journal*, vol. 10, no. 10, pp. 8560–8577, 2023. https://doi.org/10.1109/JIOT.2022.3194881

[120] R. Vatambeti, S. V. Mantena, K. V. D. Kiran, S. Chennupalli, and M. V. Gopalachari, "Black hole attack detection using Dolphin Echo-location-based machine learning model in MANET environment," *Computers and Electrical Engineering*, vol. 114, p. 109094, 2024. https://doi.org/10.1016/j.compeleceng.2024.109094

[121] H. Bostani and V. Moonsamy, "EvadeDroid: A practical evasion attack on machine learning for black-box Android malware detection," *Computers and Security*, vol. 139, p. 103676, 2024. https://doi.org/10.1016/j.cose.2023.103676

[122] M. J. Iqbal, S. Aurangzeb, M. Aleem, G. Srivastava, and J. C.-W. Lin, "RThreatDroid: A ransomware detection approach to secure IoT based healthcare systems," *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 5, pp. 2574–2583, 2023. https://doi.org/10.1109/TNSE.2022.3188597

9 AUTHORS

Ahmad K. Al Hwaitat, King Abdullah II School of Information Technology, The University of Jordan, Amman 1142, Jordan (E-mail: a.hwaitat@ju.edu.jo).

Hussam N. Fakhouri, Department of Data Science and Artificial Intelligence, University of Petra, Amman, Jordan.

Moatsum Alawida, Department of Computer Sciences, Abu Dhabi University, Abu Dhabi 59911, United Arab Emirates.

Mohammed S Atoum, King Abdullah II School of Information Technology, The University of Jordan, Amman 1142, Jordan.
Bilal Abu-Salih, King Abdullah II School of Information Technology, The University of Jordan, Amman 1142, Jordan.
Imad K. M. Salah, King Abdullah II School of Information Technology, The University of Jordan, Amman 1142, Jordan.
Saleh Al-Sharaeh, King Abdullah II School of Information Technology, The University of Jordan, Amman 1142, Jordan.
Nabil Alassaf, King Abdullah II School of Information Technology, The University of Jordan, Amman 1142, Jordan.