JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 11 Issue 01
2026

Research Article
https://doi.org/10.56979/1101/2026/1348

A Hybrid Scenario-Driven Risk Analysis of Security and Privacy in Extended Reality Ecosystems

Mamoon Obiedat[1], Ahmad Alkhatib[2], Qais Al-Na'amneh[3], Mahmoud Aljawarneh[3], Fadi Bata[3], and Ayoub Alsarhan[1&4*]

[1]Department of Information Technology, Faculty of Prince Al-Hussein Bin Abdallah II for Information Technology, The Hashemite University, Zarqa, Jordan.
[2]Cyber security department, Alzaytoonah university of Jordan.
[3]Department of Cybersecurity and Cloud Computing, Applied Science Private University, Amman, Jordan.
[4]Department of Data Science and Artificial intelligence, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan.
Corresponding Author: Ayoub Alsarhan. Email: a.alsarhan@ammanu.edu.jo

Received: February 01, 2026 **Accepted:** May 07, 2026

Abstract: Extended Reality (XR) is a recent innovation in human-computer interaction that joins both the real and digital worlds. XR systems are actively used in such critical areas as the healthcare, manufacturing, and education sphere, and the security and privacy of the XR systems should be investigated strictly. The immersive and data-heavy quality of XR presents a more distinct and broadened threat variety, which involves technical, humanistic, and perceptual vulnerable regions that lack proper approach measures on conventional approaches to securing cyberspace. The article introduces an XR ecosystem-specific hybrid and scenario-based risk assessment methodology. We blend formal measures of impact of Common Vulnerability Scoring System (CVSS) with a multi-factorial likelihood model of our own to offer a quantitative examination of practical vectors of attack. We perform and monitor scenarios involving malicious application sideloading to steal data and man-in-the-middle attacks to steal credentials using an experimental testbed involving a consumer-grade VR headset. The results of our study indicate that the social engineering vulnerability to the user and the insecure defaults of developer-specific features constitute the most prominent factors of the high-risk vulnerabilities. Credential theft via network interception and data exfiltration using unauthorized permissions prove to be the immediate threats that are most critical according to the results. We summarize by suggesting a series of practical action plan recommendations to be taken by platform vendors, application developers, and organizations in the effort to reduce the risk of these, including the use of context-sensitive security devices and developing effective user education.

Keywords: Extended Reality (XR); Cybersecurity; Virtual Reality (VR); Risk Assessment; CVSS; Social Engineering; Man-in-the-Middle Attacks; Data Privacy

1. Introduction

Extended Reality (XR) which is the general title behind the Virtual (VR), Augmented Reality (AR), and Mixed Reality (MR) technology is quickly developing beyond a new technology to become a fundamental platform of the new generation of computing. XR will transform not only the process of remote work and collaboration but also industries such as industrial manufacturing, medical training, and K-12 education through the development of immersive, interactive, and persistent digital products and services to change how work is done and information is accessible across various industries [1]–[4]. The XR market in the

world is expected to rise exponentially, indicating the deep-seated nature of these systems into the system of our life at both the personal and professional level. [5].

Nevertheless, such transformative potential is coupled with an unexplored increase in the digital attack surface. Contrary to traditional computers, XR devices are advanced sensory data platforms of data collection [6], [7]. They constantly record intimate biometric information (eye tracking, facial expression), detailed behavioral patterns (gait, gestures), and detailed 3D maps of physical environment of their users [8]. The same aspects that allow such powerful immersive experiences also provide new channels of security breaches and privacy violations with a potentially far-reaching impact. The resulting threat terrain has the form of a multi-layered ecosystem, which includes physical user vulnerabilities to cloud services they are connected to, as illustrated in Fig. 1.

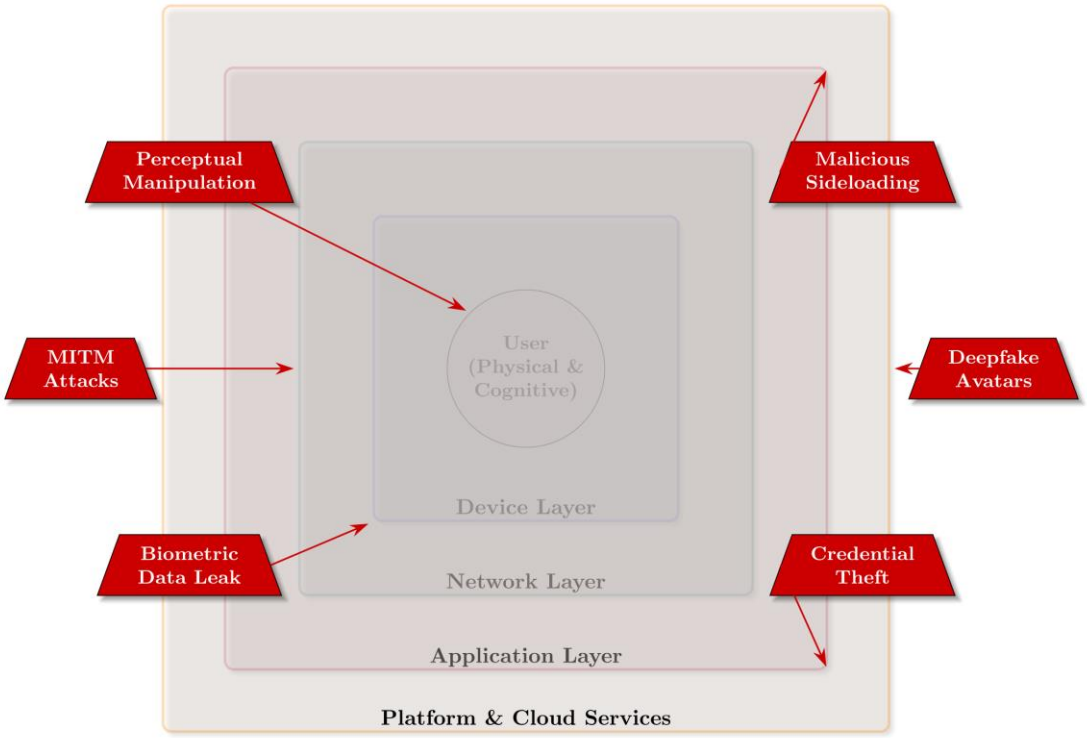

Figure 1. The Layered XR Ecosystem Threat Landscape, showing multi-domain attack surface from the user's cognitive state to cloud-based platform services.

A potential adversary can delete a virtual safety fence in an industrial AR overlay, or can create a photorealistic deepfake avatar of a workmate to cause social engineering in a VR meeting. Those threats obscure the concept of cybersecurity and physical security, and they are a challenge that the current security paradigm is not prepared to address at all, on the one hand, because of its ill-equipped nature, on the other hand, due to its inability to consider the traditional privacy aspects of the interconnection between the virtual realm and the real one [9]–[12].

The human factors increase a large part of these security challenges. Users, typically fascinated by the novelty of XR, can have a high level of implicit trust in the systems in which they interact, which reduces their security posture to a minimum of possible levels of security positioning in systems of interaction with users [13]. This is made worse by the fact that insecure practices like the continuous activation of the so-called Developer Mode on consumer products usually go around basic security protocols like application verification during installations [14]. The presence of a new and non-user-friendly technology is combined with the lack of mental models that enable users to perceive the usage of the new security method. It is a fertile ground where exploitative actions can be performed.

Conventional risk assessment frameworks [15] are necessary but do not usually work when used in XR. They find it difficult to measure risks which greatly depend on user context, cognitive load, and psychological vulnerability. It underscores an urgent requirement of a novel type of risk analysis approaches, which will combine formal technical severity measurements with subtle and context-sensitive likelihood models. This is the main reason why this work should address this gap.

The primary contributions of this paper are therefore:

1. A thorough examination of the various and distinctive threat environment of contemporary XR ecosystems, technical, human-focused, and perceptual attack Vectors.
2. The development and application of a hybrid risk assessment system that integrates the CVSS in quantifying the impact with a customized and weighted likelihood model to generate a holistic risk score.
3. An experimental analysis of this framework based on a sequence of scenario-based experiments on a commercial, off-the-shelf VR platform, which delivers tangible information about vulnerabilities of high risk.
4. A series of concrete and practical recommendations to key actors such as vendors of the platform, developers, and enterprise adopters to build more secure XR systems.

This paper is organized as follows. Part 2 looks at related literature in XR security. Section 3 outlines our approach and experimental design to hybrid risk assessment. The analysis of our scenario analysis is presented in section 4. Section 5 is the discourse on the broader implications of what we found and Section 6 is a conclusion of the paper and directions to research in the future.

2. Related Work

The XR security and privacy research is a fast-growing field. The initial research was based mainly on threat detection and classification whereas the latter research has started to investigate the particular defense mechanisms and human factor influence.

2.1. Threat Taxonomies and Modeling in XR

There is a considerable amount of literature devoted to listing the various threats to XR ecosystems [16], [17]. These dangers include traditional attacks like malware and denial-of-service to those specific to XR. The article at the source has given a detailed review of these issues [18] suggesting a multifaceted approach to them, which involves integrating technical solutions with privacy-focused policies. On the same note, analysis discusses the dangers of biometric data gathering, the development of persuasive deepfakes, and network-related attacks, as well as others, [19]–[21]. Lake et al., applied the STRIDE framework to list information disclosure and tampering as the most serious threats, particularly the risk of so-called immersive manipulation, which might lead to patient outcomes being affected by the approach of this kind of manipulation to the patient-focused issue of risk management in healthcare systems and facilities [22]. Such works offer a good background to the extent of the possible attacks but are usually qualitative in terms of evaluation.

2.2. Human-Centric Vulnerabilities and Usable Security

The user has become a highly popular topic of research due to the critical role of the user in the XR security chain. This high mental load of XR can greatly reduce the capacity to make effective security decisions, as it is shown by Cayir et al. that a user is unable to make the right choices when facing a risky permission request, even in an immersive setting, because of this mental load [23]. This mental weakness predisposes users to social engineering to a great extent. The so-called virtual trust the implicit trust that a user is putting into the virtual world and the members of the virtual world can be abused by the evil forces in the form of a misguided avatar or a spoofed interface. This has triggered a study of usable security of XR, which seeks to come up with security mechanisms that are effective and not disruptive. Proposed solutions include context-aware permission systems and novel methods for visualizing security information within a 3D space [14].

2.3. Proposed Defense Mechanisms

To address these threats, researchers have come up with a number of mitigation measures. Technically, there is a lot of concern regarding strong authentication protocols that could be applied in XR including continuous authentication relying on behavioral biometrics, including gait or head movement patterns, etc. [24]. The Privacy-Enhancing Technologies (PETs) are currently being considered to deal with the privacy challenge posed by data collection of immense magnitude. Chen et al. came up with a federated learning architecture to train user behavior models without centralizing sensitive raw information, which showed a way to characterize privacy-guaranteed analytics. [25]. At the system level, on-device sandboxing and runtime monitoring are common proposals to contain potentially malicious applications.

2.4. Risk Assessment Frameworks

While threat catalogs are extensive, the formal, quantitative assessment of XR risks is less developed. Some studies have adapted existing cybersecurity frameworks like STRIDE [22], [26] or have applied the CVSS to XR vulnerabilities [27], [29]. hese frameworks, however, tend to be unable to provide the entire context of XR threats, especially when it comes to the interaction between technical exploits and human factors. It is not just that the probability of a perceptual attack, e.g., depends on the technical exploitability of the attack, but it also varies depending on user context and cognitive state. This underscores this research gap to which our work is connected. In order to put our contribution into perspective, Table 1 assembles the major existing studies and provides an overview of the research gap that our hybrid methodology fills.

Table 1. Comparison of Existing Research in XR Security Risk Assessment.

Ref.	Focus Area	Risk Assessment Method	Key Limitation Addressed by Our Work
[18]	Threat Taxonomy	Qualitative, Categorical	Lacks quantitative risk scoring for prioritization.
[22], [28]	Healthcare Apps, STRIDE	Semi-Quantitative (STRIDE Framework)	Does not systematically integrate human factors into the likelihood calculation.
[23]	Usable Security, Human Factors	User Studies, Qualitative Analysis	Focuses on user perception rather than developing a holistic, quantifiable risk score.
This Work	Hybrid Risk Assessment	Quantitative (CVSS + Custom Weighted Likelihood Model)	Integrates technical impact with context-dependent, human-centric likelihood factors for a more realistic risk score.

3. Materials and Methods

Our methodology will measure and estimate security risks in XRs in a quantitative and organized manner. Our modelling involves a scenario-based approach to modeling realistic attack vectors, with the basis of the analysis being a hybridization of the CVSS framework with a bespoke likelihood model.

3.1. Experimental Testbed Architecture

In order to make our findings practically relevant, we have built an experimental testbed to represent a realistic consumer XR usage setting as shown in Fig. 2.

Hardware and OS: he main device was a Meta Quest 2 headset, which was selected due to its unavoidability on the market and Android-based operating system. The "Developer Mode" was turned on to allow the application sideloading and diagnostic observation with the help of the Android debug bridge (adb).

Network Environment: The headset was linked to a separated Wi-Fi access point. An adversarial platform was an operating machine using Kali Linux and connected to the same network. This arrangement provided the possibility to intercept and manipulate network traffic without impacting external networks.

Analysis Tools: A combination of generic penetration testing tools was used: Wireshark to capture network packets passively; mitmproxy to launch active Man-in-the-Middle (MITM) attacks; adb to interact with the device and APKTool to reverse-engineer application packages.

Malicious Application: We have created a proof-of-concept malicious VR application that was created with the Unity engine. It was a masqueraded application that was to seek permission to access the microphones and steal the audio recordings to a server controlled by an attacker.

Ethics and Safety Statement: All experiments involving the malicious application and phishing scenarios were conducted in a strictly controlled, isolated laboratory environment using test devices and dummy user accounts. No human participants were involved in the testing, and no real-world user data was exposed or collected. Consequently, Institutional Review Board (IRB) approval was not required for this study.

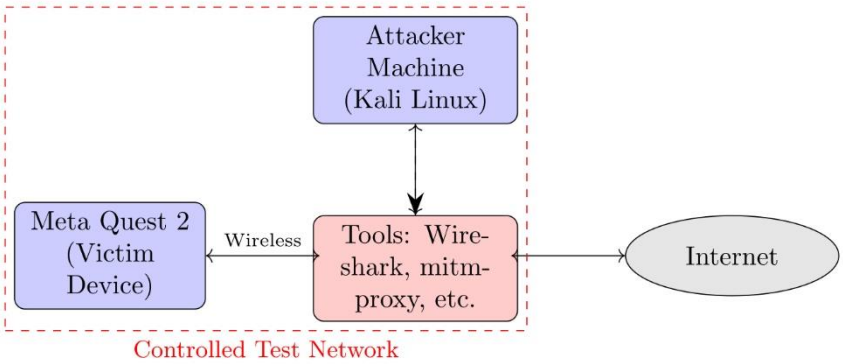

Figure 2. System architecture of the experimental testbed, illustrating the relationship between the victim device, the attacker machine, and the isolated network.

3.2. Scenario-Driven Analysis

We designed three attack scenarios to evaluate a range of vulnerabilities.

Scenario 1: Data Exfiltration via Malicious Sideloaded Application. This scenario tests the risks associated with insecure application installation practices. The attacker's objective is to steal ambient audio from the user's environment by convincing the user to install a malicious application from an untrusted source.

Scenario 2: Credential Theft via Network Interception. This scenario evaluates the threat of MITM attacks on a compromised network. The attacker's objective is to intercept the user's network traffic and inject a phishing page to capture credentials for a third-party service.

Scenario 3: Perceptual Manipulation via Chaperone Hijacking (Conceptual). This scenario explores a high-impact, XR-specific attack. Unlike Scenarios 1 and 2, which were executed experimentally, this scenario is presented purely as a theoretical model to assess its potential severity. Assuming an attacker has already achieved Remote Code Execution (RCE), their objective is to maliciously alter the virtual safety boundaries (the "Chaperone" or "Guardian") to cause physical harm or disorientation.

3.3. Hybrid Risk Quantification Model

Our risk model defines risk as the product of Impact and Likelihood.

Table 2. Scoring Rubric for the Custom Likelihood Assessment Model.

Factor (Weight)	Score = 1 (Low)	Score = 3 (Moderate)	Score = 5 (High)
User Susceptibility (0.4)	Requires expert user to make multiple, unlikely errors; clear warnings provided.	Non-technical user might be tricked by a moderately convincing lure or overlook a subtle warning.	Attack exploits default user behavior, cognitive biases, or trust; warning is absent, ambiguous, or normalized.
System Exposure (0.3)	Requires a rare, non-default system configuration or physical access.	Preconditions are common on public networks or among specific user groups (e.g., enthusiasts using Dev Mode).	Affects default "out-of-the-box" user configurations on common networks.
Exploit Availability (0.2)	Requires custom-built, non-public tools and deep, specialized expertise.	Requires common penetration testing tools but some configuration and knowledge.	Can be executed with readily available, easy-to-use, and well-documented software (e.g., SDKs, popular tools).

Attack Detectability (0.1)	Attack is immediately obvious to the user (e.g., system crash, major visual glitch, clear alert).	Anomalies might be noticed in system logs or by an observant, technical user, but are not user-facing.	Attack is stealthy, consuming negligible resources and leaving no user-facing trace of compromise.

3.3.1. Impact Assessment (CVSS)

We used the CVSS v3.1 Base Score to provide a standardized, objective measure of a vulnerability's intrinsic severity. The Base Score (0.0-10.0) is derived from metrics evaluating exploitability (e.g., Attack Vector, User Interaction) and impact (Confidentiality, Integrity, Availability) [14]. This allows for a consistent comparison of disparate vulnerability types.

3.3.2. Likelihood Assessment Model

To quantify the probability of a successful attack, we developed a custom likelihood model. The Likelihood Score (0.1-1.0) is a weighted average of four factors, designed to capture the context-dependent aspects of XR threats:

$$Likelihood = \frac{1}{5}\sum_{i=1}^{n} w_i \cdot F_i \quad (1)$$

where F_i is the score [1-5] for each factor and w_i is its assigned weight. The factors are:

User Susceptibility (US) [Weight: 0.4]: The propensity of a non-technical user to perform the actions required by the attacker (e.g., accepting a prompt, ignoring a warning).

System Exposure (SE) [Weight: 0.3]: The frequency of the necessary preconditions for the attack (e.g., use of public Wi-Fi, enabling Developer Mode).

Exploit Availability (EA) [Weight: 0.2]: The public availability and use of tools required for the attack.

Attack Detectability (AD) [Weight: 0.1]: The likelihood that the attack will be detected by the user or by defense systems (an inverse metric: 1=Easily Detected, 5=Stealthy).

The weights were established through expert elicitation, involving a panel of three cybersecurity researchers specializing in XR human-computer interaction, reflecting the consensus that human-centric factors are currently the most prominent enablers of XR attacks. The impact of these weights on the final risk ranking is evaluated through a sensitivity analysis detailed in Section 4.4. To ensure consistent application of these factors, we developed a detailed scoring rubric, presented in Table 2.

Risk Score Thresholds: To translate the final risk score into actionable guidance for enterprise prioritization, we define three risk bands: Low (0.1--3.9) indicates threats requiring routine monitoring; Medium (4.0--6.9) identifies vulnerabilities needing scheduled remediation; and High (7.0--10.0) represents critical risks demanding immediate intervention. The final Risk Score is calculated as:

$$\text{Risk Score} = \text{CVSS Base Score} \times \text{Likelihood Score}$$

The complete process, from scenario definition to final risk score calculation, is visualized in the flowchart in Figure 3. This hybrid approach, outlined in Algorithm 1, allows for a nuanced evaluation that balances intrinsic technical severity with the practical likelihood of exploitation in a real-world setting.

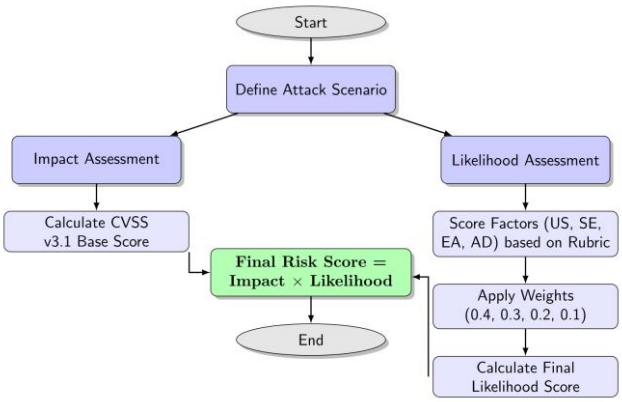

Figure 3. Flowchart of the Hybrid Risk Quantification Process, showing the parallel assessment of technical impact and context-aware likelihood.

Algorithm 1 Hybrid Risk Assessment Procedure

```
procedure ASSESSRISK(scenarios)
    Input: A set of XR security scenarios
    Output: A prioritized list of risks
    for all scenario in scenarios do
        vuln ← IdentifyPrimaryVulnerability(scenario)
        CVSS_Score ← CalculateCVSSv3.1(vuln)
        US ← AssessUserSusceptibility(scenario)          ▷ [1-5]
        SE ← AssessSystemExposure(scenario)              ▷ [1-5]
        EA ← AssessExploitAvailability(scenario)         ▷ [1-5]
        AD ← AssessAttackDetectability(scenario)         ▷ [1-5]
        L_Score ← (0.4*US + 0.3*SE + 0.2*EA + 0.1*AD) / 5
        RiskScore ← CVSS_Score * L_Score
        Store scenario, RiskScore
    end for
    return RankByRiskScore(results)
end procedure
```

4. Results

In this section, we show quantitative outcomes of the use of our hybrid risk assessment framework to the specified scenarios. The ultimate risk scores, presented in Table 3, allow the prioritized ranking of threats and evaluated.

4.1. Scenario 1: Malicious Sideloaded Application

This situation, which aimed at the exfiltration of data through an application that was not installed securely, was identified as a risky one.

4.1.1 Execution and Observation

The malicious Unity application was installed on the Quest 2 headset through adb with success. When initially launched, the new app gave the regular Android OS microphone access permission warning. The prompt was generic, in which there was no particular context on why it was necessary. Upon receiving the permission, audio recording was started by a background service. Wireshark packet capture verified that exfiltration of the audio data carried in UDP packets was periodic and to our server under the control of the attacker. The malicious code remained completely transparent to the user who was presented with a harmless loading screen and the malicious code used minimal system resources so that it remained undetected.

4.1.2 Risk Quantification

Impact (CVSS): 6.5 (Medium). The vulnerability was rated as AV:L/AC:L/PR:N/UI: R/S:U/C:H/I:N/A:N. The key driver is the High impact on Confidentiality (C:H), as all ambient audio is compromised. User Interaction (UI:R) is required.

Likelihood: 0.84. The factors contributing to the attack were high: User Susceptibility was 4/5, because user is trained to take permissions; System Exposure was 4/5, because developer mode is a normal SDK; and Exploit Availability was 5/5, because the tools are standard SDKs.

Final Risk Score: 5.46 (High). Calculated as 6.5×0.84.

4.2. Scenario 2: Credential Theft via Network Interception

The situation of MITM attack was evaluated as the most practical risk scenario out of the tested ones mainly because of a high level of user vulnerability to phishing.

4.2.1. Execution and Observation

Using 'arpspoof' and 'mitmproxy', we successfully established a MITM position on the network. We found that intercepting traffic to modern services was heavily constrained by Transport Layer Security (TLS) and HTTP Strict Transport Security (HSTS). To circumvent these limitations, we targeted the headset's captive portal detection phase—a process that typically occurs over unencrypted HTTP upon connecting to a new Wi-Fi network. By spoofing the captive portal, we presented a convincing "Meta

Account Login" phishing page as the first screen the user saw before they could access the broader internet. Because this interaction leverages an expected operating system behavior (captive portal sign-in) rather than a browser warning bypass, it proved highly reproducible and reliable across multiple test connections. Credentials entered on this page were successfully captured by 'mitmproxy'. This vector was highly effective as it appeared to be a legitimate part of the network connection process. Visual evidence of this attack, from the user's perspective to the attacker's successful data capture, is provided in Figure 4.

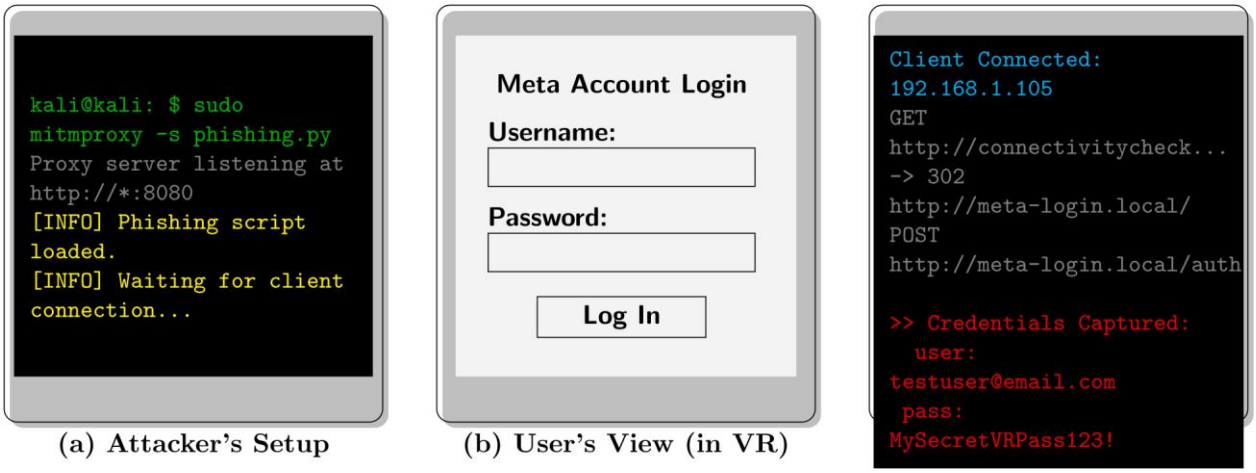

Figure 4. Visual Evidence from Scenario 2 (Credential Theft via MITM), showing the attacker's setup, the user-facing phishing page, and the captured credentials.

4.2.2. Risk Quantification

Impact (CVSS): 8.8 (High).

Scored as AV:A/AC:L/PR:N/UI:R/S:C/C:H/I:H/A:N. The Attack Vector is Adjacent (AV:A). The high score is driven by the High impact on both Confidentiality (C:H) and Integrity (I:H) of the targeted user account, and a change in Scope (S:C) from the device to the online service.

Likelihood: 0.74. User Susceptibility was maximal at 5/5, as phishing is notoriously effective. System Exposure was 3/5, reflecting the frequent use of untrusted public Wi-Fi.

Final Risk Score: 6.51 (High). Calculated as 8.8×0.74.

4.3. Scenario 3: Perceptual Manipulation (Conceptual)

This scenario was assessed based on its theoretical potential, assuming a prerequisite RCE vulnerability. It represents the most severe long-term threat.

4.3.1. Risk Quantification

Impact (CVSS): 9.9 (Critical).

Scored as AV:L/AC:L/PR:H/UI:N/S:C/C:N/I:H/A:H. This near-maximum score reflects the potential for direct physical harm. The attack has a High impact on the Integrity of the system's safety features and a High impact on their Availability.

Likelihood: 0.60. While User Susceptibility and Attack Detectability are maximal (5/5), the Likelihood is constrained by the low probability of the prerequisite RCE, with System Exposure and Exploit Availability rated at 1/5.

Final Risk Score: 5.94 (High). Calculated as 9.9×0.60.

4.4. Sensitivity Analysis and Uncertainty

To assess the robustness of our ranking against the subjective weighting of likelihood factors, we performed a sensitivity analysis. By applying equal weights (0.25) to all four likelihood factors (User Susceptibility, System Exposure, Exploit Availability, Attack Detectability) instead of our expert-elicited weights, the Likelihood Scores become: S1 (0.85), S2 (0.65), and S3 (0.60). The resulting Risk Scores shift to: S1 (5.53), S2 (5.72), and S3 (5.94). Under equal weighting, the ranking changes to prioritize the conceptual S3 attack over the empirically validated MITM attack. This highlights the value of our human-centric weighting scheme, which appropriately prioritizes highly probable user-interaction threats (S2) over purely technical worst-case scenarios (S3).

Detailed CVSS v3.1 Vector Analysis:

S1 (Sideloading): AV:L/AC:L/PR:N/UI:R/S:U/C:H/I:N/A:N (Base Score: 6.5)

S2 (MITM Theft): AV:A/AC:L/PR:N/UI:R/S:C/C:H/I:H/A:N (Base Score: 8.8)

S3 (Perceptual Manipulation): AV:L/AC:L/PR:H/UI:N/S:C/C:N/I:H/A:H (Base Score: 9.9)

Table 3. Summary of Hybrid Risk Assessment Results for All Scenarios.

Scenario	CVSS Score (Impact)	User Susceptibility	System Exposure	Exploit Availability	Attack Detectability	Final Risk Score
S1: Malicious Sideloading	6.5 (Medium)	4/5	4/5	5/5	4/5	5.46 (High)
S2: Credential Theft (MITM)	8.8 (High)	5/5	3/5	3/5	2/5	6.51 (High)
S3: Perceptual Manipulation	9.9 (Critical)	5/5	1/5	1/5	5/5	5.94 (High)

5. Discussion

The analysis reveals several critical insights. The highest risk score (6.51) belongs to the credential theft scenario, closely followed by perceptual manipulation (5.94) and malicious sideloading (5.46), as shown in Table [tab:risk_summary]. All three scenarios fall into the "High" risk category, but for different reasons.

The risk matrix in Figure 6 indicate that the most urgent threat in the practical risk is the MITM attack (S2) that integrates both high impact and high probability. The least impact but the most likely attack is the sideloading attack (S1), and therefore, the most possible way that the average user can be compromised. On the other hand, the critical impact of the perceptual manipulation attack (S3) has a low current likelihood. Furthermore, relying exclusively on a standard CVSS assessment would rank S3 (9.9) as the absolute highest priority, followed by S2 (8.8) and S1 (6.5). However, our hybrid model corrects this bias by emphasizing the immediate, real-world probability of S2 over the purely theoretical S3. It is also important to note that traditional CVSS metrics (Confidentiality, Integrity, Availability) are not perfectly calibrated to capture the physical or psychological harm induced by perceptual manipulation (S3). While we map physical harm to a loss of system Integrity and Availability, future XR security standards will require dedicated metrics for bodily safety. This implies a prioritization plan: in the short term, there should be a strategy of mitigation of network attacks and better application vetting, and long-term research and design should be aimed at assuring the architectural integrity of the perceptual pipeline.

One of the fundamental results is the overweight of human factors. User susceptibility was mainly the cause of the high likelihood scores in S1 and S2. This ascertains that the user is nowadays the weakest component in the XR security chain. The excessive role of the user-centric determinants in the total probability of an attack is quantitatively presented in figure 5, as the probability score of each scenario is broken down.

This observation seriously suggests that technical solutions will alone not be sufficient. It must be a socio-technical approach, which is to provide solid system design with efficient user education and user-friendly security interfaces. Moreover, the Developer Mode risk is not insecure in the traditional meaning of the term, but rather insecure default settings. This brings out the need of the vendors of the platform to build features that are secure by default even by the power users. It is up to all stakeholders to come up with a concerted effort to translate these findings into practice. Table 3 outlines a set of specific, actionable recommendations designed to address the most critical vulnerabilities identified in our analysis.

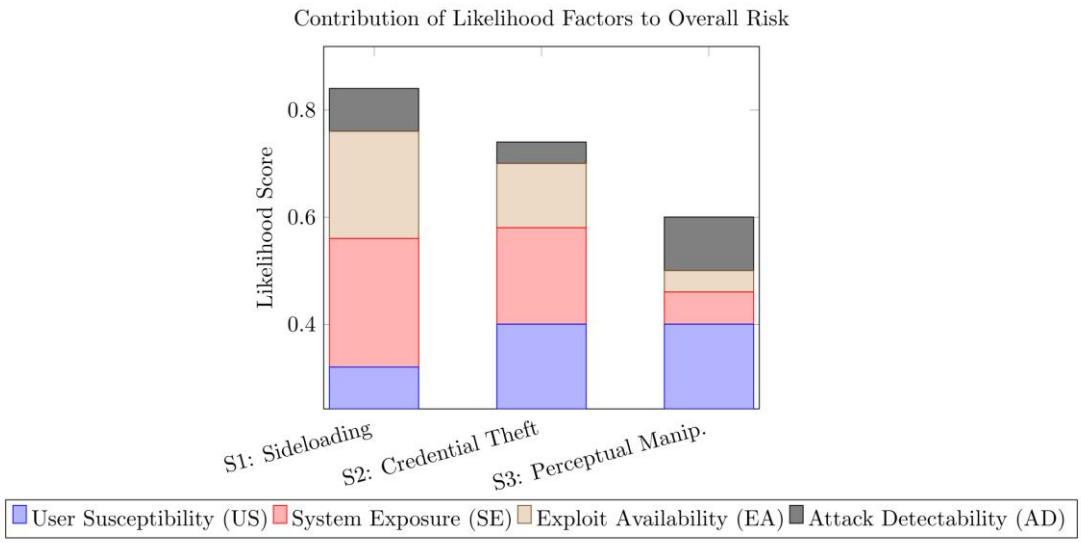

Figure 5. Contribution of weighted likelihood factors to the final Likelihood Score for each scenario, highlighting the dominant role of User Susceptibility (US).

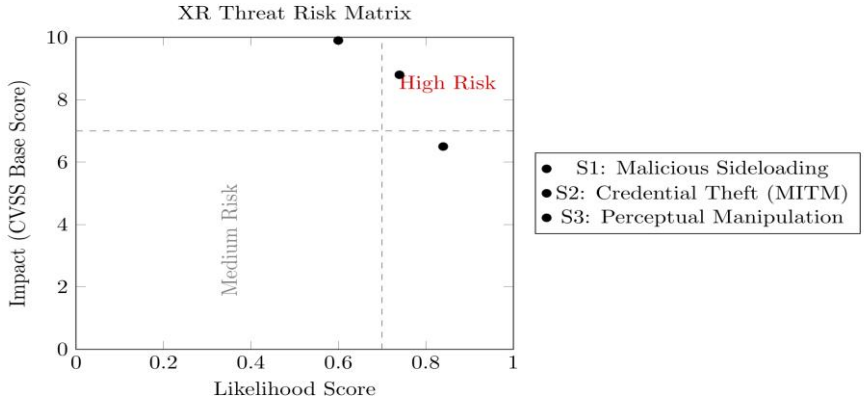

Figure 6. Risk Matrix visualizing the Impact vs. Likelihood scores for the evaluated scenarios. The top-right quadrant represents the most critical threats.

Table 4. Recommended Mitigation Strategies for Key Stakeholders.

Stakeholder	Identified Threat / Vulnerability	Recommended Mitigation Strategy
Platform Vendors	Malicious Sideloading via Developer Mode	Implement an auto-disable timer for Developer Mode. Enhance permission prompts with context-specific justifications and visual risk indicators.
	Network-based Phishing (Captive Portals)	Improve captive portal detection and display prominent, unambiguous warnings for unencrypted portals before loading any web content.
App Developers	Data Exfiltration via Excessive Permissions	Strictly adhere to the principle of least privilege. Provide clear, user-friendly privacy policies detailing exactly what data is collected and why.
Organizations & Enterprise Adopters	Social Engineering and Insecure User Practices	Mandate user security training focused on XR-specific threats (phishing, avatar impersonation).

	Establish and enforce clear policies against using XR devices on untrusted public Wi-Fi.

6. Conclusions

In this paper, we introduced and applied a hybrid, scenario-driven risk assessment framework to quantify security and privacy threats in contemporary XR ecosystems. Our methodology, which integrates the formal CVSS impact metrics with a custom likelihood model, provides a more holistic view of risk than traditional approaches. The experimental results identified credential theft via MITM network attacks and data exfiltration from maliciously sideloaded applications as the most critical immediate threats. A conceptual analysis of perceptual manipulation attacks confirmed their potential for causing physical harm, marking them as a severe long-term concern.

A central conclusion of our work is that human factors are currently the primary enabler of high-risk exploits in consumer XR. The combination of user trust in immersive environments, a lack of established security protocols, and insecure system defaults creates significant vulnerabilities. Consequently, mitigating these risks requires a multi-pronged strategy as outlined in our recommendations. Platform vendors must redesign security-critical interactions to be more context-aware and secure by default. Application developers must adhere to the principle of least privilege. Finally, a significant investment in targeted user education is necessary to equip users with the mental models needed to navigate the unique security challenges of XR.

This study has several limitations. Crucially, our experimental testbed relied exclusively on the Meta Quest 2 operating in Developer Mode; our findings may not fully generalize to other AR/MR platforms, restrictive enterprise configurations, or different operating systems. Our analysis was confined to a single hardware platform, and our likelihood model relied on expert judgment rather than large-scale empirical user data. Specifically, our conclusions regarding human factors are drawn from the predictive risk model rather than direct behavioral measurements of user participants under cognitive load or phishing stimuli. Future work should expand this analysis across a wider range of XR devices and involve rigorous, human-subject empirical testing to identify platform-specific vulnerabilities. A critical area for future research is the development and empirical testing of "usable security" interfaces for XR that can convey security information effectively without disrupting immersion. Finally, long-term research must focus on building a resilient and verifiable rendering pipeline to provably defend against the threat of perceptual manipulation attacks.

References

1. S. K. Jagatheesaperumal, K. Ahmad, A. Al-Fuqaha, and J. Qadir, "Advancing education through extended reality and internet of everything enabled metaverses: Applications, challenges, and open issues," *IEEE Transactions on Learning Technologies*, vol. 17, pp. 1120–1139, 2024.
2. S. Pahi and C. Schroeder, "Extended privacy for extended reality: XR technology has 99 problems and privacy is several of them," Notre Dame J. on Emerging Tech., vol. 4, p. 1, 2023.
3. A. Almaini, A. Al-Dubai, I. Romdhani, M. Schramm, and A. Alsarhan, "Lightweight edge authentication for software defined networks," Computing, vol. 103, no. 2, pp. 291–311, 2021.
4. M. Aljaidi, G. Samara, M. K. Singla, A. Alsarhan, M. Hassan, M. Safaraliev, P. Matrenin, and A. Tavlintsev, "A particle swarm optimizer-based optimization approach for locating electric vehicles charging stations in smart cities," International Journal of Hydrogen Energy, vol. 87, pp. 1047–1055, 2024.
5. D. Jones, S. Ghasemi, D. Gračanin, and M. Azab, "Privacy, safety, and security in extended reality: User experience challenges for neurodiverse users," in International conference on human-computer interaction, 2023, pp. 511–528.
6. L. Hallal, J. Rhinelander, and R. Venkat, "Recent trends of authentication methods in extended reality: A survey," Applied System Innovation, vol. 7, no. 3, p. 45, 2024.
7. J. Tooker, "Privacy in the era of constant reality capture: Informed consent in extended reality (xr)," PhD thesis, 2021.
8. M. Z. Iqbal, X. Xu, V. Nallur, M. Scanlon, and A. G. Campbell, "Security, ethics and privacy issues in the remote extended reality for education," in Mixed reality for education, Springer, 2023, pp. 355–380.
9. K. Lebeck, K. Ruth, T. Kohno, and F. Roesner, "Towards security and privacy for multi-user augmented reality: Foundations with end users," in 2018 IEEE symposium on security and privacy (SP), 2018, pp. 392–408.
10. C. Pereira, A. Marto, R. Ribeiro, A. Gonçalves, N. Rodrigues, C. Rabadão, R. L. de C. Costa, and L. Santos, "Security and privacy in physical–digital environments: Trends and opportunities," Future Internet, vol. 17, no. 2, p. 83, 2025.
11. M. El-Hajj, "Cybersecurity and privacy challenges in extended reality: Threats, solutions, and risk mitigation strategies," in Virtual worlds, 2024, vol. 4, p. 1.
12. J. Happa, A. Steed, and M. Glencross, "Privacy-certification standards for extended-reality devices and services," in 2021 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW), 2021, pp. 397–398.
13. T. Li, Y. Zheng, W. Ma, G. Wang, Z. Li, and L. Wang, "P-4.33: Trustworthy metaverse: A comprehensive investigation into security risks and privacy issues in artificial intelligence-extended reality systems," in SID symposium digest of technical papers, 2024, vol. 55, pp. 872–877.
14. E. Hine, I. N. Rezende, H. Roberts, D. Wong, M. Taddeo, and L. Floridi, "Safety and privacy in immersive extended reality: An analysis and policy recommendations," Digital Society, vol. 3, no. 2, p. 33, 2024.
15. J. A. de Guzman, A. Seneviratne, and K. Thilakarathna, "Unravelling spatial privacy risks of mobile mixed reality data," Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 5, no. 1, pp. 1–26, 2021.
16. 16. S. Oh and T. Shon, "Digital forensics for analyzing cyber threats in the XR technology ecosystem within digital twins," Electronics, vol. 13, no. 13, p. 2653, 2024.
17. L. Schmidt and E. Yigitbas, "Taxonomy and analysis of security vulnerabilities, privacy violations and potential mitigation strategies to XR systems," in Proceedings of the 18th ACM international conference on PErvasive technologies related to assistive environments, 2025, pp. 368–375.
18. S. Kaur, S. Rajvanshi, and G. Kaur, "Privacy and security concerns with augmented reality/virtual reality: A systematic review," Augmented Reality and Virtual Reality in Special Education, pp. 209–231, 2024.
19. R. Acheampong, D.-M. Popovici, T. C. Balan, A. Rekeraho, and I.-A. Oprea, "A cybersecurity risk assessment for enhanced security in virtual reality," Information, vol. 16, no. 6, p. 430, 2025.
20. McGill and Mark, "White paper-the IEEE global initiative on ethics of extended reality (XR) report–extended reality (XR) and the erosion of anonymity and privacy," Extended Reality (XR) and the Erosion of Anonymity and Privacy-White Paper, pp. 1–24, 2021.
21. Q. Al-Na'amneh, M. Aljaidi, A. Nasayreh, H. Gharaibeh, R. E. Al Mamlook, A. S. Jaradat, A. Alsarhan, and G. Samara, "Enhancing IoT device security: CNN-SVM hybrid approach for real-time detection of DoS and DDoS attacks," Journal of Intelligent Systems, vol. 33, no. 1, p. 20230150, 2024.

22. K. Lake, A. Mc Kittrick, M. Desselle, A. P. L. Bo, R. A. M. Abayasiri, J. Fleming, N. Baghaei, D. D. Kim, and others, "Cybersecurity and privacy issues in extended reality health care applications: Scoping review," JMIR XR and Spatial Computing (JMXR), vol. 1, no. 1, p. e59409, 2024.
23. D. Cayir, A. Acar, R. Lazzeretti, M. Angelini, M. Conti, and S. Uluagac, "Augmenting security and privacy in the virtual realm: An analysis of extended reality devices," IEEE Security & Privacy, vol. 22, no. 1, pp. 10–23, 2024.
24. A. S. Jaradat, A. Nasayreh, Q. Al-Na'amneh, H. Gharaibeh, and R. E. Al Mamlook, "Genetic optimization techniques for enhancing web attacks classification in machine learning," in 2023 IEEE intl conf on dependable, autonomic and secure computing, intl conf on pervasive intelligence and computing, intl conf on cloud and big data computing, intl conf on cyber science and technology congress (DASC/PiCom/CBDCom/CyberSciTech), 2023, pp. 0130–0136.
25. C. Chen, J. Liu, H. Tan, X. Li, K. I.-K. Wang, P. Li, K. Sakurai, and D. Dou, "Trustworthy federated learning: Privacy, security, and beyond," Knowledge and Information Systems, vol. 67, no. 3, pp. 2321–2356, 2025.
26. M. Aljaidi, A. Alsarhan, D. Al-Fraihat, A. Al-Arjan, B. Igried, S. M. El-Salhi, M. Khalid, and Q. Al-Na'amneh, "Cybersecurity threats in the era of ai: Detection of phishing domains through classification rules," in 2023 2nd international engineering conference on electrical, energy, and artificial intelligence (EICEEAI), 2023, pp. 1–6.
27. A. Abu-Zaid, M. Aljaidi, Q. Al-Na'amneh, G. Samara, A. Alsarhan, and B. Qadoumi, "Advancements and challenges in the internet of drones security issues: A comprehensive review," Machine Intelligence Applications in Cyber-Risk Management, pp. 1–24, 2025.
28. D. Jones, S. Fealy, D. Evans, and R. Galvez, "The use of extended realities providing better patient outcomes in healthcare," Frontiers in medicine, vol. 11. Frontiers Media SA, p. 1380046, 2024.
29. Hussain, T., Faiz, R. B., Aljaidi, M., Khattak, A., Samara, G., Alsarhan, A., & Alazaidah, R. (2023). Maximizing Test Coverage for Security Threats Using Optimal Test Data Generation. Applied Sciences, 13(14), 8252. https://doi.org/10.3390/app13148252.