Architecture design of ML-based spoofing detection solution*

Serhii Bulba[1,†], Danylo Yefimov[1,†], Larysa Kriuchkova[2,†], Roman Kyrychok[2,†], and Sergiy Obushnyi[2,*,†]

[1] State University of Trade and Economics, 19 Kioto str., 02156 Kyiv, Ukraine
[2] Borys Grinchenko Kyiv Metropolitan University, 18/2 Bulvarno-Kudriavska str., 04053 Kyiv, Ukraine

Abstract

The paper explores theoretical and practical approaches to detecting spoofing as a component of modern cyberattacks using machine learning and deep learning methods. The methodological apparatus is based on general scientific methods of analysis, synthesis, induction, and deduction, as well as specialized methods of search and semantic analysis of scientific sources. The aim of the work is to generalize existing approaches to detecting spoofing attacks and to develop an optimal conceptual architecture for a software solution to identify spoofing risks. The empirical basis of the review was formed by systematizing and interpreting 23 relevant scientific papers, which reflect the evolution from single-factor rule-based mechanisms to multi-level architectures with multimodal feature integration. The paper clarifies the essence of spoofing as a hidden cyberthreat that exploits protocol vulnerabilities and the imperfections of authentication mechanisms, while accounting for the human factor. A classification of key types of spoofing (email, IP, DNS, ARP, caller ID, website, GPS) was performed, and their consequences for the confidentiality, integrity, and availability of information resources were determined. Based on the review, requirements for anti-spoofing solutions in corporate environments were formed, in particular for email spoofing and Business Email Compromise scenarios. An Explainable AI module was also integrated to form an evidence package, increase transparency in decision-making, and support operational incident analysis. The features of implementing the architecture in the AWS environment, using managed services for orchestration, streaming processing, storage, and model deployment, were considered. The results obtained confirm that multimodal and ensemble approaches are the most promising for increasing accuracy, resistance to zero-day scenarios, and controllability of anti-spoofing systems.

Keywords

spoofing; phishing; email spoofing; Business Email Compromise; machine learning; deep learning; multimodal detection; ensemble models; Explainable AI; software solution architecture; AWS; near-real-time ML pipeline

1. Introduction

Spoofing is a type of cyberthreat that involves the intentional falsification of biometric data (or other data associated with an identified subject) to imitate the legitimacy of a subject, device, or service and gain access to the corresponding system (for authentication and authorization). Spoofing attacks are usually based on exploiting vulnerabilities in network protocols and the imperfections of authentication approaches, taking into account the influence of the human factor.

The primary danger of spoofing lies in its hidden implementation, which makes timely detection difficult. Spoofing is widely used as a component of complex attacks [1, 2], particularly phishing [3, 4], man-in-the-middle, and distributed denial-of-service attacks. Spoof detection is a basic prerequisite for improving information systems. The development of an optimal architecture for detecting and analyzing spoofing risks enables significantly enhanced protection mechanisms against cyber threats and reduces the likelihood of spoofing as a means to achieve malicious goals. The development of effective approaches to preventing spoofing attacks reduces the risks of unauthorized access, data leakage, and disruption to the stability of information systems.

* CPITS 2026: Cybersecurity Providing in Information and Telecommunication Systems, February 28, 2026, Kyiv, Ukraine
* Corresponding author.
† These authors contributed equally.
serhii.bulba@gmail.com (S. Bulba); d.yefimov@knute.edu.ua (D. Yefimov); l.kriuchkova@kubg.edu.ua (L. Kriuchkova); r.kyrychok@kubg.edu.ua (R. Kyrychok); s.obushnyi@kubg.edu.ua (S. Obushnyi)
0009-0004-5330-2632 (S. Bulba); 0009-0001-2160-7333 (D. Yefimov); 0000-0002-8509-6659 (L. Kriuchkova); 0000-0002-9919-9691 (R. Kyrychok); 0000-0001-6936-955X (S. Obushnyi)

2. Literature review

Modern research in the field of analysis and detection of spoofing and phishing attacks indicates a gradual transition from single-factor methods to complex, multi-level architectures that use machine learning and deep learning models. Abdelnabi, Sahar, Katharina Krombholz, and Mario Fritz [5] emphasize that as attacks become more complex, classical approaches based on static rules become less effective. They proposed the VisualPhishNet system, which analyzes the visual similarity between phishing pages and legitimate sites of well-known brands.

The team of authors consisting of Eleni Kytidou, Theodosia Tsikriki, and Konstantinos Rantos [6] also emphasizes that countering spoofing attacks requires the use of adaptive artificial intelligence models. This opinion is supported by the authors Satwinder Singh and Asadullah Safi [7], who conducted a comprehensive study of existing approaches and architectures for building solutions to analyze phishing threats.

Liu, Dong-Jie, devoted his research to the visual detection of phishing websites [8]. According to the author, visual features are effective in detecting zero-day attacks, even when the characteristics of URLs and text have been intentionally masked. The author uses convolutional neural networks to analyze screenshots of web pages. This allows us to confirm the possibility of using computer vision to detect spoofing. In addition, Kehkashan, T., M. Abdelhaq, and A. S. Al-Shamayleh [9] emphasize the importance of explanatory visual models for practical use in countering phishing attacks.

A significant amount of research has focused on analyzing URLs as the primary source of phishing attack signals. Eman Abdullah Aldakheel, Mohammed Zakariah, Ghada Abdalaziz Gashgari, Fahdah A. Almarshad, and Abdullah I. A. Alzahrani [10] have demonstrated the effectiveness of deep learning for the automatic detection of spoofing features via URL analysis. In particular, Dipankar Kumar Mondal, Bikash Chandra Singh, Haibo Hu, Shivazi Biswas, Zulfikar Alom, and Mohammad Abdul Azim [11] have shown that machine learning methods based on classification models remain effective due to their low computational complexity and high performance. Also, Sunil, D. K., and Ahammad, S. H. emphasize the importance and effectiveness of using machine learning models for detecting spoofing features [12]. In addition, Hayk Ghalechyan, Elina Israyelyan, Avag Arakelyan, Gerasim Hovhannisyan & Arman Davtyan [13] confirmed the ability of classification models to effectively handle spoofing detection and to scale up flexibly using large real-world data sets.

Prasad, Arvind, and Shalini Chandra [14] investigated drifting attacks and proposed a gradual learning approach that enables a response to new phishing patterns. Authors Yang Xiao, Tieshan Li, and Asiri Sultan [15] supplemented the above by using transformer models for URL analysis. By doing so, the authors demonstrated the ability of such architectures to model complex connections in URL character sequences.

An important issue in scientific research on phishing email detection and analysis is the use of CNN models for message content analysis. Fang, Yong, Cheng Zhang, and Cheng Huang [16] proposed an improved RCNN model with an attention mechanism that allows achieving high accuracy in email classification. Further development of this area is presented in the works of Haiwei Li, Jiahai Yang, Yuqi Li, and Kun Li [17], who proposed using the BERT transformer model to analyze phishing attacks that have a spoofing component.

Brij B. Gupta, Akshat Gaurav, Varsha Arya, Razaz Waheeb Attar, Shavi Bansal, Ahmed Alhomoud, and Kwok Tai Chui [18] used CNN models to analyze semantics, context, and emails in their study. Reza Hassanpour, Erdogan Dogdu, Roya Choupani, Onur Goker, and Nazli Nazli [19] also confirm the advantages of deep learning over established approaches in the problem of email phishing.

Studies by Abeer Alhuzali1, Ahad Alloqmani, Manar Aljabri, and Fatemah Alharbi [20] show that the results of analyzing and detecting spoofing risks in email messages depend significantly on the datasets used and the configuration of the attack itself. In particular, Hany F. Atlam and Olayonu Oluwatimilehin [21] highlighted the complexity of business email compromise (BEC)

attacks, which often do not contain malicious URLs or attachments. This requires additional contextual and behavioral analysis. Structural and graph models have been used to improve the obfuscation resistance of phishing pages.

Choon Lin Tan, Kang Leng Chiew, Kelvin S.C. Yong, San Nah Sze, Johari Abdullah, and Yakub Sebastian [22] proposed a graph-based approach to phishing threat detection. According to this approach, a web page is modeled as a graph of connected components. This allows for the detection of anomalous structures. Yan Ding, Nurbol Luktarhan, Keqin Li, and Wushour Slamu [23] demonstrated that even keyword-based combined methods can be effective as an auxiliary component of a multi-layer architecture. Yukun Li, Zhenguo Yang, Xu Chen, Huaping Yuan, and Wenyin Liu [24] used an ensemble model based on URL and HTML features, showing that combining disparate information sources significantly improves detection quality.

Thus, modern scientific research on spoofing and phishing attack detection demonstrates a transition from single-factor rule-based approaches to multi-level architectures based on machine learning and deep learning methods. The works of many authors [5, 8, 9] confirm the effectiveness of visual analysis for detecting zero-day attacks, in particular through the use of convolutional neural networks and explainability mechanisms. A significant part of the research is focused on the analysis of URL addresses, where both classical machine learning models and deep and transformer architectures demonstrate high performance and scalability [6, 10–15, 25]. For the detection of phishing emails, the advantages of CNN-, R-CNN-, and BERT-based models have been proven, capable of effectively analyzing the semantics, context, and hidden intentions of messages [16–19]. The generalization of the results of review and applied works [6, 7, 20, 21, 24] shows that the most promising are multimodal and ensemble architectures that combine URL, content, visual, structural, and behavioral features to improve the quality and stability of spoofing detection systems.

3. Research method

The study's methodological apparatus was based on general scientific and specialized methods. The purpose of the work was to determine the theoretical and practical approaches to spoofing detection using machine learning technologies and to develop an optimal architecture for a software solution that identifies spoofing risks. An important task was to conduct a scientific analysis of theoretical and practical approaches to detecting spoofing attacks. For this purpose, the content of scientific sources was searched and analyzed semantically, as selected as the most relevant to the research topic.

To develop the conceptual architecture for analyzing and detecting spoofing, the methods of scientific analysis, synthesis, induction, and deduction were used. Technical documentation was prepared, and a conceptual decomposition of the requirements for building an MVP version of this software solution was carried out.

The determination of the optimal architecture was based on the use of the following methodological approaches: a thorough analysis of existing requirements for anti-spoofing solutions, taking into account existing business cases; research into implementation options for existing models; study of AWS technical documentation, which allowed building an architectural model of the corresponding software solution.

4. Main material

When discussing the above issues, it is first important to understand the meaning of spoofing and its varieties, which are, to one degree or another, revealed in various scientific works. Table 1 lists out the main types of spoofing, with descriptions of their characteristics and impacts on users of information systems and businesses.

Table1.
Classification for the main types of spoofing attacks

Type of spoofing	Short description	Consequences from the point of view of cyber threats
Email spoofing	Forging the sender address of an email to impersonate a trusted person or organization.	Credential theft, spread of malware, financial losses, compromise of corporate systems.
IP spoofing	The attacker changes the IP address of the source of a network packet to impersonate another device.	Difficulty in detecting attacks, disruption of service availability, possible network penetration.
Website spoofing	Creating a fake website that visually imitates a legitimate resource.	Compromise of accounts, financial losses, theft of personal data.
GPS spoofing	Transmitting fake GPS signals to change the location determined by the device.	Navigation disruption, risks to transport safety, potential accidents and losses.
DNS spoofing (DNS cache poisoning)	DNS record spoofing, which redirects the user to a fake website. The attack is often invisible to the victim.	Theft of logins and passwords, infection of devices, and undermining of trust in web resources.
ARP spoofing	The attacker sends fake ARP messages on the local network, associating his MAC address with the IP address of another device.	Traffic interception, man-in-the-middle attacks (MITM)
Caller ID spoofing	Changing the phone number displayed during a call to deceive the victim.	Financial fraud, phishing of personal data

Email spoofing is a mechanism for substituting services that handle email messages. This type of spoofing is practically an integral part of phishing attacks. IP spoofing is defined as malicious actions associated with the substitution (duplication) of network addresses, which opens up the possibility of hiding traffic sources and initiating a network attack; in particular, substitution of the mobile operator's IP address is used as a technique for concealing the actual location of a subscriber. Website spoofing consists in creating fake web resources that imitate legitimate sites to collect confidential data. DNS spoofing is a specific type of attack on the domain name system. As a result of such attacks, users can be redirected to unauthorized resources. This type of attack is difficult to detect with conventional monitoring tools. ARP spoofing is characterized by a threat to the operation of local networks, which provides the ability to intercept and modify network traffic. Caller ID spoofing involves substituting the subscriber's identification in telecommunications systems in order to implement fraudulent schemes. GPS spoofing is highlighted as a specific type of attack that affects the correctness of determining the location of navigation systems; its practical feasibility has been experimentally demonstrated for aviation surveillance protocols ADS-B and Mode S using software-defined radio [1].

All the described types of spoofing may use different technical approaches, but they are united by a common goal: misleading users or systems. In modern cyberattacks, several types of spoofing can be used simultaneously, which complement each other. Detecting such threats is quite difficult using conventional monitoring tools. To effectively counteract complex spoofing attacks, preventive measures are needed that allow threats to be detected at a stage before they actively manifest.

Identifying the risks of phishing attacks using elements of artificial intelligence and machine learning is achieved through a range of mathematical concepts and models. The first mathematical model we'd like to represent is the Binary model. The basic problem of spoofing detection is formalized as a binary classification problem.

Let us consider a set of objects:

$$X = x_1, x_2, \dots, x_n \quad (1)$$

where every object x_i represents an email, URL, web page, or other signal being analyzed..

The corresponding set of labels could be defined as:

$$y = 0,1 \tag{2}$$

where

$$y_i = \begin{cases} 1, & \text{if } x_i \text{ when spoofing is defined} \\ 0, & \textit{if } \text{x\_i } \textit{when spoofing is not defined} \end{cases} \tag{3}$$

The task is to find the reflection:

$$f : X \to [0,1] \tag{4}$$

which minimizes empirical risk:

$$\beta = \frac{1}{n} \sum_{i=1}^{n} k(y_i, F(x_i)) \tag{5}$$

where β is the loss function (e.g., logistic loss).

Another methematical approach could be defined by multimodal spoofing detection model. Taking into account the multimodal architecture, each object xix_ixi is represented as a tuple:

$$x_i = x_i^{(u)}, x_i^{(t)}, x_i^{(v)}, x_i^{(s)} \tag{6}$$

where $x_i^{(u)}$ is URL related features, $x_i^{(t)}$ textual features, $x_i^{(v)}$ visual features, $x_i^{(s)}$ structural or behavioral features.

A separate detector is defined for each modality:

$$f_k(x_i^{((k))}), \qquad k \in u, t, v, s \tag{7}$$

The final risk score is defined as a weighted aggregation of:

$$R(x_i) = \sum_k \alpha_k \cdot f_k(x_i^{(k)}) \tag{8}$$

where α_k are the weighting factors, $\sum_k \alpha_k = 1$.

The decision is made according to the threshold rule:

$$\widehat{y}_i = \begin{matrix} 1, & R(x_i) \geq \partial \\ 0, & R(x_i) < \partial \end{matrix} \tag{9}$$

where, ∂ is the appropriate threshold value at which a decision is made regarding the presence of spoofing.

Other mathematical models can be used for spoofing validation, but the above are basic and require relatively less effort to implement in real solutions. These mathematical models also determine the key elements of modern spoofing detection systems and are used in many solutions for practical cases of countering spoofing attacks. The combination of binary classification and multimodal analysis provides increased accuracy, stability, and level of explainability of the results obtained in the process of using the models.

The solution architecture for spoofing detection is organized to emphasize increasing the accuracy of results when analyzing data that records the presence of spoofing attack risks. Such architecture is largely classical and aligns with the general principles of building solutions for detection risks. The effectiveness of using the classical spoofing detection model depends on many factors (the quality of the data used to train and test the models, the structure of the data itself and its relevance to the business case being solved, the solution deployment environment, etc.). In general, improving the efficiency of solution operation can be achieved through comprehensive data analysis, the use of modern artificial intelligence methods, and reducing the impact of the limitations of individual mathematical models. A key feature of spoofing detection architecture is the phenomenon's multimodal nature, which allows you to simultaneously consider URL tags, text content, web page visual characteristics, and, if necessary, audio and behavioral data.

The first element of the solution's architecture for analyzing and detecting spoofing risks is the creation of a dataset. It is the accumulation of relevant, consistent data that provides the machine learning model with the necessary information for effective training.

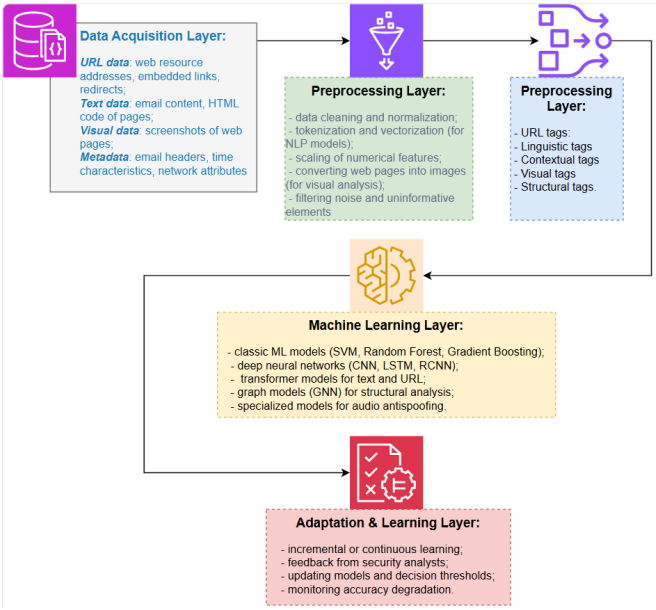

Figure 1: Classical architecture of spoofing detection solution formation using ML

Thanks to data aggregation across channels, the system can capture weak attack signals that may be invisible in a single source. Pre-processing and data normalization increase model stability, eliminating noise and reducing the impact of irrelevant features on the learning process. This directly reduces the number of false positives. The feature extraction level is focused on the formation of informative and discriminative data representations. The use of deep neural networks and transformer models enables automatic detection of complex nonlinear dependencies characteristic of spoofing attacks. The simultaneous use of linguistic, structural, and visual features increases the system's ability to generalize and effectively detect zero-day attacks. A special role is played by the integration of graph and visual features that are resistant to URL and textual content obfuscation.

The detection layer, built on a set of specialized models, provides deep analysis of each data type. The creation of partial risk assessments allows you to store information about each channel's contribution to the overall decision. Further integration of results at the ensemble decision-making level is a key factor in improving the accuracy of spoofing detection. Ensemble mechanisms reduce the impact of errors from individual models and provide more balanced decisions in heterogeneous data.

An important aspect of improving detection quality is the presence of explainability. Explainable models allow you to analyze the reasons for the system's operation and identify potential errors in the decision-making logic. This helps increase trust in the system among cybersecurity experts and facilitates further model improvement. In addition, explainability allows you to quickly adapt the system to new types of spoofing attacks [26].

The level of adaptation and learning ensures the long-term effectiveness of the architecture in the face of constantly evolving threats. Incremental learning and analyst feedback allow models to be updated without significantly reducing performance. Together, all components of the architecture form a holistic system that maximizes the accuracy, robustness, and reliability of spoofing detection, a key requirement for modern cyber defense systems.

5. Practical implementation

Based on theoretical approaches to solution formation in spoofing detection and analysis, a number of functional and non-functional requirements for the designed solution's various use cases were analyzed. From the point of view of simplifying the UOMV for the development of architectural design, several use cases were selected that allow developing an MVP version of the applied

solution. This software component analyzes incoming and outgoing email messages. The data preparation and model training pipeline should be deployed in the AWS environment.

The implementation of the MVP should include only spoofing verification in email messages; however, in the future, the solution can be expanded to include other types of checks. The architectural model of the solution that enables spoofing detection at the email message level is illustrated in Figure 2. The solution aims to improve the efficiency of email spoofing detection by combining machine learning with multimodal analysis and context-oriented decision-making.

Here is a brief description of the architecture. At the top level, the logic of the solution can be represented as the following steps (warfclaw stages): 1) the user receives a message; 2) the solution extracts the relevant metadata from the message and stores it in the repository; 3) the service responsible for pre-processing the stored data, transforms it into a normalized form and stores it in the database; 4) the service responsible for launching the ML pipeline extracts the relevant arrays of normalized data and trains the models; 5) the parameters for detecting spoofing from the trained models are passed to the service responsible for “making decisions” regarding the classification of the message as risky; 6) the service that is integrated with the mail server to provide information to the user that a particular message contains a risk of spoofing.

The solution implementation involves a use case in which users use email via the Microsoft 365 service. The solution involves integrating the Microsoft 365 mail server with the Email Service software component (service). Email Service is a microservice responsible for monitoring email messages arriving in a user's Microsoft 365 mailbox and saving them in the AWS S3 Bucket. Email Service ensures that not all incoming email messages are saved, but only a subset: EML/headers, extracted artifacts (URLs, attachment hashes), and the “evidence packet” (metadata used to analyze spoofing risks). An important element of the architecture is the near-real-time ML pipeline. This model component determines the spoofing risk for each email. The ML pipeline is multimodal, focusing on the following sub-elements: header-detection model, behavioral-detection model, and NLP-detection model.

The header-detection model is used to detect attempts to replace the sender for the corresponding email message. This element uses the results of checks for compliance with email authentication standards used to detect and prevent email spoofing (SPF, DKIM, and DMARC).

The behavioral-detection model analyzes email communication patterns using historical data. This model evaluates interaction patterns between users, including correspondence frequency, as well as deviations from standard behavioral patterns. This approach is especially effective for detecting Business Email Compromise attacks that do not contain malicious attachments or URLs.

The NLP-detection model provides semantic analysis of the letter's text using deep learning methods. It allows you to identify signs of social engineering, hidden intentions, request urgency, and other linguistic patterns characteristic of phishing messages.

The next level is to collect the obtained scoring results from different models and store them in the appropriate data warehouse. It is worth noting that AWS SageMaker is used in combination with the Fusion thresholds component to integrate and aggregate different scoring thresholds. In fact, the Fusion thresholds are a constantly updated JSON.config file whose data is updated based on machine learning results, for which the Scoring ML models component is responsible. This component supports the simultaneous operation of three models: a header-detection model, a behavioral-detection model, and an NLP-detection model. The above approach allows you to reduce the error levels that individual models can produce. That is, integrating and aggregating the results enables you to obtain more objective estimates of the threat of spoofing.

The Explainability (XAI) module is responsible for ensuring the solution's transparency and practical value. From a technical perspective, there are different ways to implement this module. For simplification, within the MVP version of the solution, it is proposed to use a combination of Json.file or DynamoDB and the corresponding API endpoints, which are responsible for performing the following operations: 1) updating the corresponding data repository with information related to Explainability for the implemented solution (containing a characteristic and description of the system components, log and expected results in a structured form); 2) obtaining information about

the system characteristics from the API endpoint. Explanations (information about the model's characteristics) are continually adjusted based on the results of individual models and their integration. Also, the explanatory information is supplemented with additional characteristics regarding individual features, highlighting key fragments of text and forming a generalized evidence package.

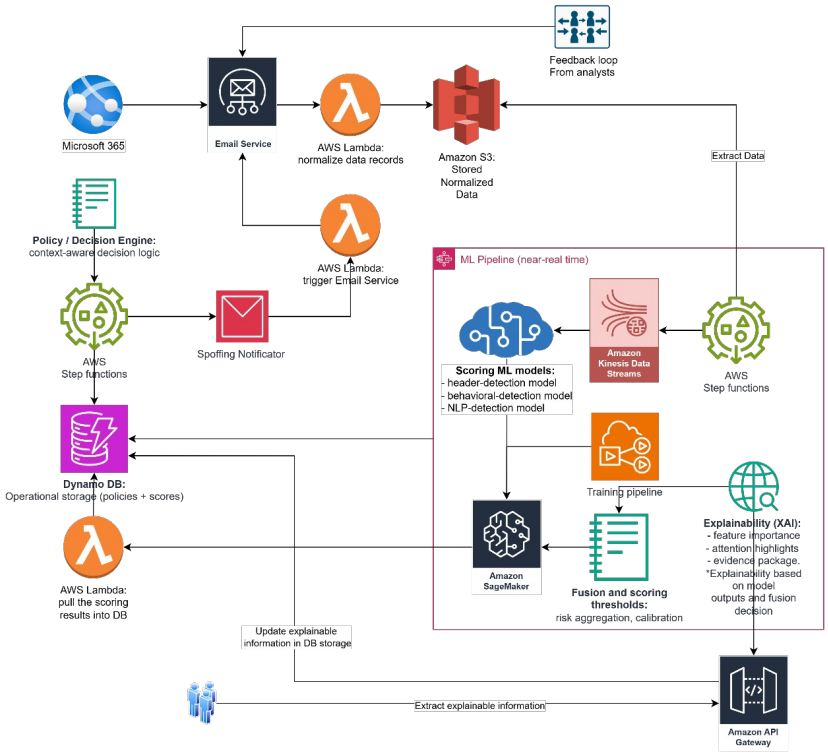

Figure 2: AWS-based microservices architecture of spoofing-detection solution

In our opinion, using Json file is a simpler option and can be used for the MVP version; however, for professional use, it is proposed to store information about Explainable AI in Dynamo DB or Document DB. This simplifies the work of information security analysts and increases confidence in automated decisions.

The last and most important component of the model is the Policy / Decision Engine. This component uses context-oriented logic and performs the following tasks: 1) receives information about the integrated spoofing risk scores for a particular email message; 2) makes a decision about the spoofing risk and generates further notifications for system users. In the event of a suspicious or malicious email, the system initiates appropriate actions, including notifications, message quarantines, or other response measures, implemented through the orchestration mechanism.

Conclusions

The study analyzed modern approaches to detecting spoofing and phishing attacks using machine learning and deep learning methods. It has been proven that traditional single-factor rule-based mechanisms are not effective enough in the face of the increasing complexity and stealth of spoofing attacks. Analysis of existing approaches to architecture design for spoofing detection solutions indicates the feasibility of transitioning to multi-level, multimodal architectures. Such architectural compositions combine protocol, content, behavioral, structural, and visual features. It was determined that the analysis of many criteria by which the spoofing-score is evaluated is an important prerequisite for identifying the risk of sender substitution. It was also shown that spoofing detection effectiveness increases when the solution is supplemented with behavioral and semantic detection methods. The generalization of the results of the review and applied works shows that the most promising are multimodal and ensemble architectures that combine URL,

content, visual, structural, and behavioral features to improve the quality and stability of spoofing detection systems.

Based on the study's results, an improved architecture for an AWS-based solution is proposed to detect spoofing risk. The conceptual architectural scheme is based on independent ML models, with subsequent integration of risk scores across all included models, thereby reducing the error rate of individual detectors. The implementation of a near-real-time ML pipeline enables timely threat detection before they manifest. The integration of Explainable AI mechanisms ensures transparency of decision-making and increases trust in automated cyber defense systems. The use of AWS cloud infrastructure ensures scalability, fault tolerance and manageability of the proposed solution.

Declaration on Generative AI

While preparing this work, the authors used the AI programs Grammarly Pro to correct text grammar and Strike Plagiarism to search for possible plagiarism. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

References

[1] M. TajDini, V. Sokolov, P. Skladannyi, Performing Sniffing and Spoofing Attack Against ADS-B and Mode S using Software Define Radio, in: IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics, 2021, 7–11. doi:10.1109/UkrMiCo52950.2021.9716665

[2] Y. Sadykov, et al., Technology of Location Hiding by Spoofing the Mobile Operator IP Address, in: IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics, 2021, 22–25. doi:10.1109/UkrMiCo52950.2021.9716700

[3] R. Marusenko, V. Sokolov, V. Buriachok, Experimental Evaluation of Phishing Attack on High School Students, Advances in Computer Science for Engineering and Education III, vol. 1247, 2020, 668–680. doi:10.1007/978-3030-55506-1_59

[4] R. Marusenko, V. Sokolov, I. Bogachuk, Method of Obtaining Data from Open Scientific Sources and Social Engineering Attack Simulation, Advances in Artificial Systems for Logistics Engineering, vol. 135, 2022, 583–594. doi:10.1007/978-3-031-04809-8_53

[5] A. Sahar, K. Krombholz, M. Fritz. VisualPhishNet: Zero-Day Phishing Website Detection by Visual Similarity, In: Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security (2020), 1681–1698. doi:10.1145/3372297.3417233

[6] E. Kytidou, Th. Tsikriki, K. Rantos, Machine Learning Techniques for Phishing Detection: A Review of Methods, Challenges, and Future Directions, Intell. Decis. Technol., 2025, doi:10.1177/18724981251366763

[7] S. Singh, A. Safi, A Systematic Literature Review on Phishing Website Detection Techniques, J. King Saud Univ. Comput. Inf. Sci., 2023, doi:0.1016/j.jksuci.2023.01.004

[8] L. Dong-Jie, A CNN-Based SIA Screenshot Method to Visually Identify Phishing Websites, J. Netw. Syst. Manag., 2024, doi:10.1007/s10922-023-09784-7

[9] T. Kehkashan, M. Abdelhaq, A. S. Al-Shamayleh, et al., Explainable Phishing Website Detection for Secure and Sustainable Cyber Infrastructure, Sci. Rep., 15, 2025, doi:10.1038/s41598-025-27984-w

[10] Aldakheel, E. Abdullah, et al., A Deep Learning-Based Innovative Technique for Phishing Detection in Modern Security with Uniform Resource Locators, Sens., 23.9, 2023, doi:10.3390/s23094403

[11] D. K. Mondal, et al., A Novel Learning Approach to Detect Malicious, J. Inf. Secur. Appl., 2021, doi:10.1016/j.jisa.2021.102967

[12] D. Sunil, S. Ahammad, et al., Phishing URL Detection Using Machine Learning Methods, Adv. Eng. Softw., 2022, doi:10.1016/j.advengsoft.2022.103288

[13] H. Ghalechyan, E. Israyelyan, A. Arakelyan, G. Hovhannisyan, A. Davtyan, et al., Detecting Malicious, Sci. Rep., 2024, doi:10.1038/s41598-024-74725-6

[14] P. Arvind, Sh. Chandra, PhiUSIIL: A Diverse Security Profile Empowered Phishing URL Detection Framework based on Similarity Index and Incremental Learning, Comput. Secur., 136, 2024, doi:10.1016/j.cose.2023.103545
[15] Y. Xiao, T. Li, A. Sultan, PhishTransformer: A Novel Approach to Detect Phishing Attacks Using URL Collection and Transformer, Electron., 13, 1, 2024, doi:10.3390/electronics13010030
[16] F. Yong, Ch. Zhang, Ch. Huang, Phishing Email Detection Using Improved RCNN Model With Multilevel Vectors and Attention Mechanism, In: IEEE Access 7, 2019, doi:10.1109/ACCESS.2019.2913705
[17] L. Haiwei, Y. Jiahai, Li Yuqi, Li Kun Li, Email Phishing Attack Detection Based on BERT Transformer Model, In: International Conference on Optics, Electronics, and Communication Engineering, 2024, doi:10.1117/12.3049161
[18] B. Gupta, A. Gaurav, V. Arya, Razaz Waheeb Attar, Shavi Bansal, Ahmed Alhomoud, Kwok Tai Chui, Advanced BERT and CNN-Based Computational Model for Phishing Detection in Enterprise Systems, Comput. Model. Eng. Sci., 2024, 1243–1260. doi:10.32604/cmes.2024.056473
[19] R. Hassanpour, E. Dogdu, R. Choupani, O. Goker, N. Nazli, Phishing E-mail Detection by Using Deep Learning Algorithms, In: Proceedings of the 2018 ACM Southeast Conference, 2018, 1555-1562, doi:10.1145/3190645.3190719
[20] A. Alhuzali1, A. Alloqmani, M. Aljabri, F. Alharbi. In-Depth Analysis of Phishing Email Detection: Evaluating the Performance of Machine Learning and Deep Learning Models Across Multiple Datasets, Appl. Sci., 2025, doi:10.3390/app15063396
[21] H. F. Atlam, O. Oluwatimilehin, Business Email Compromise Phishing Detection Based on Machine Learning: A Systematic Literature Review, Electron., 2023, doi:10.3390/electronics12010042
[22] C. L. Tan, et.al., A Graph-Theoretic Approach for the Detection of Phishing Webpages, Comput. Secur., 2020, doi:10.1016/j.cose.2020.101793
[23] Y. Ding, N. Luktarhan, K. Li, W. Slamu, A Keyword-based Combination Approach for Detecting Phishing Webpages, Comput. Secur., 2019, 256–275. doi:10.1016/j.cose.2019.03.018
[24] Y. Li, Zh. Yang, X. Chen, H. Yuan, W. Liu, Jail-Phish: A Stacking Model Using URL and HTML Features for Phishing Webpage Detection, Future Gener. Comput. Syst., 84, 2018, 27–39, doi:10.1016/j.future.2018.11.004
[25] L. Xin, X. Wang, M. Sahidullah, et al., ASVspoof 2021: Towards Spoofed and Deepfake Speech Detection in the Wild, In: IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2023, 2507–2522, doi:10.1109/TASLP.2023.3285283
[26] Y. Kostiuk, et al., Models and Technologies of Cognitive Agents for Decision-making with Integration of Artificial Intelligence, in: Modern Data Science Technologies Doctoral Consortium (MoDaST), vol. 4005, 2025, 82–96.