Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Categorization and Risk Assessment of Cyber-Attack Chains: Sub-Chain Approach Using LSTM for Predictive Modeling

Parham Rajab nezhad
Shahed University
Shahriar Bijani
Bijani@shahed.ac.ir

Shahed University

Research Article

Keywords: LSTM, Risk Assessment, Cyber-Attack Chains, MITRE ATT&CK, First-order Markov chain

Posted Date: April 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9472866/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Journal of Computer Virology and Hacking Techniques on July 17th, 2026. See the published version at https://doi.org/10.1007/s11416-026-00647-8.

Categorization and Risk Assessment of Cyber-Attack Chains: Sub-Chain Approach Using LSTM for Predictive Modeling

Abstract:

This paper proposes a novel framework for the analysis, classification, and risk prediction of cyber-attack chains, addressing the limitations of existing ATT&CK-based detection methods that overlook sequential dependencies and evolving risks in advanced persistent threat (APT) campaigns. The framework defines conditions for attack chain formation, constructs structured representations, and categorizes them into distinct classes for enhanced interpretability. Probabilistic risk values for sequences of MITRE ATT&CK techniques are computed using a first-order Markov model, integrating each technique's impact score and mitigation effectiveness to produce quantitative risk assessments. These continuous risk scores are further modeled using a Long Short-Term Memory (LSTM) network to predict future risks, which are subsequently transformed into discrete, interpretable categories through K-Means clustering. Experiments on real-world attack datasets demonstrate high predictive accuracy and effective multi-class discrimination, enabling early detection, informed risk prioritization, and proactive defense in complex cyber threat environments. This approach bridges the gap between probabilistic risk assessment and deep sequential modeling, providing actionable insights for Security Operation Centers (SOCs).

Key words: LSTM, Risk Assessment, Cyber-Attack Chains, MITRE ATT&CK, First-order Markov chain.

Introduction

The prevalence and sophistication of advanced persistent threats (APTs) are increasing rapidly, with the global market projected to grow from $10 billion in 2024 to $23 billion by 2028 [1]. This rapid expansion underscores the critical importance of proactive, intelligent cyber defense strategies capable not only of detecting but also diagnosing and interrupting complex attack campaigns before they achieve full penetration. Traditional detection systems often examine individual techniques in isolation and therefore fail to capture sequential dependencies, cause–effect relationships, and the synergistic effects that arise when multiple techniques combine into more destructive attack chains.

To address these limitations, this paper presents a hybrid probabilistic–deep learning framework that systematically defines the conditions for attack-chain formation, constructs structured chain representations, and categorizes them for enhanced interpretability. The framework supports early detection and diagnosis of forming chains by identifying interconnections between individual attack steps and quantifies risk evolution across stages. Probabilistic risk values for sequences of MITRE ATT&CK techniques are estimated using a first-order Markov model that integrates each technique's impact score and mitigation effectiveness to produce stage-wise numerical risk assessments. These continuous scores are then modeled using a Long Short-Term Memory (LSTM) network to capture temporal dependencies and predict future risks, which are discretized into interpretable classes via K-Means clustering to facilitate actionable decision-making.

This integrated approach provides several key benefits. It captures how individual attack steps combine to amplify threats, accounts for both sequential dependencies and mitigation effects, and generates interpretable, operationally useful outputs for Security Operation Centers (SOCs). Stage-wise numerical risk assessments enable effective prioritization of defensive responses, while the combined probabilistic–deep learning pipeline supports early warning, proactive mitigation, and informed decision-making. Overall, the proposed method advances the state-of-the-art in attack-chain prediction and proactive cybersecurity management by providing a unified, data-driven tool for detection, diagnosis, prioritization, and mitigation planning.

Background

One of the most sophisticated methods employed by cybercriminals is the APT which seeks unauthorized access, maintains a foothold, and infiltrates or modifies sensitive data [2]. These attacks are targeted and persistent, often remaining undetected for extended periods [3]. The 2024 IBM Data Breach Report reveals that the average time to identify and contain a breach has decreased to 258 days, a seven-year low [4]. Organizations utilizing AI and automation for security, demonstrated the ability to detect and contain breaches nearly 100 days faster on average, highlighting the effectiveness of these technologies in improving breach response times [4]. APT commonly utilizes multiple attack vectors to exfiltrate or alter data, a process made increasingly easier by the expanding attack surface in the digital ecosystem [5]. The delayed detection and response to these attacks are primarily due to the insufficient support for correlating and predicting attack actions, which are essential for proactive intrusion investigation and mitigation.

MITRE ATT&CK, launched in 2013, was developed to document and classify tactics, techniques, and procedures (TTPs) used in attacks on Microsoft Windows systems, with the objective of enhancing the detection of malicious activities [6]. Over the years, MITRE ATT&CK has expanded to include various platforms and technologies, becoming a comprehensive knowledge base on cyber adversary behavior and a classification system for adversarial activities throughout the attack lifecycle. This knowledge base is made publicly available through various arrays. Studies indicate that more than 80% of organizations use MITRE ATT&CK, with 57% of those surveyed utilizing it to identify security gaps in their infrastructure, 55% applying it to enforce security policies, and 54% using it for threat modeling [7].The groups listed on the MITRE website are formed by combining various techniques, with techniques sharing a common objective being grouped under a specific tactic. Additionally, the MITRE site includes a section dedicated to mitigation measures, which provide countermeasures to address these techniques. The sequence of tactics, techniques and procedures (TTPs) employed during an attack, forming the structured framework of the attack chain.

In a section of our research, we examined the mitigation measures associated with MITRE techniques. The primary motivation for investigating this issue and reviewing prior work was the incorporation of the mitigation parameter into the risk calculation formula. By integrating this parameter, the resulting risk value became more representative of actual conditions. To deepen our understanding of this aspect, we consulted the relevant literature. In [8] provided a comprehensive summary of mitigation strategies for both the ATT&CK Enterprise and ATT&CK ICS matrices, demonstrating that a given mitigation method could apply to one or both matrices. Furthermore, the framework's evaluation results identified underutilized mitigation techniques and overlooked security aspects, which could enable adversaries to exploit specific tactics, techniques, and procedures (TTPs). In essence, this section sought to clarify how adversaries could exploit particular vulnerabilities and weaknesses within information technology (IT) and operational technology (OT) to achieve their malicious objectives. As we shift from risk assessment to cyberattack prediction, several studies have explored different approaches. For example, In [3] introduces IoT-Predictor for predicting IoT-related attacks, In [7] discusses prediction methods for the railway industry, and In [9] examines the use of recommender systems for forecasting cyberattacks. The core idea behind prediction is to leverage historical data and insights from similar events, making it essential to use models that retain information about past occurrences for accurate forecasting.

Limited research has been conducted on predicting or classifying attack chains using the MITRE framework, resulting in a paucity of investigated and implemented methods. Techniques such as K-means clustering, Partition Around Medoids (PAM), and fuzzy clustering were utilized in [10], which ultimately identified hierarchical clustering as the most effective method. For any form of clustering, it is necessary to specify the number of clusters in advance. While numerous methods exist for determining the optimal number of clusters, the authors found the Elbow and Silhouette methods to be suitable for the characteristics of their dataset.

Chain of Cyber Attacks

A chain of cyber-attacks refers to a series of cyber-attacks executed sequentially to compromise a single target. Each attack is assigned to a specific phase within the attack chain, responsible for distinct tasks. This segmentation complicates the identification and diagnosis of an attack chain, as the victim may fail to recognize the connections between the various components.

Conceptually, the formation of an attack chain necessitates the occurrence of several attack phases in succession. For instance, Lockheed Martin's Cyber Kill Chain consists of seven ordered, fixed, and mandatory phases that describe the progression of a cyber-attack [11]. However, when examining real-world attacks, it becomes evident that the phases may not always be executed in a linear fashion. In some instances, not all phases are required, or multiple attacks may be employed within a single phase to achieve the desired outcome. Additionally, the structure of the chain can be adapted based on the strengths and weaknesses of the target system, as the adversary's objectives may evolve during the operation [11].

An additional challenge in identifying attack chains is the use of obfuscation techniques, which obscure critical details and mislead correlation mechanisms, complicating detection. These techniques are generally categorized into three types: expansion attack, attack removal and attack substitution. While they alter the attack's presentation, they maintain its core behavior and impact [12]. By disrupting predictable attack patterns, these methods complicate the detection of attack chains and reduce the effectiveness of security systems in identifying them. Our research is based on MITRE ATT&CK version 11.3 [13], which includes 133 documented attack chains within the MITRE knowledge base. In this framework, each phase of an attack is referred to as a tactic, each individual attack as a technique, and each complete attack chain as a group. For our analysis, we focused exclusively on corporate category attack chains. Consequently, we excluded six attack chains from our collection: ALLANITE

and HEXANE chains, which pertain to the industrial control system sector, and Bouncing Golf, DragonOK, Lotus Blossom, NEODYMIUM chains, which fall under the software category. Thus, our research is conducted with a total of 127 attack chains. Furthermore, the techniques analyzed in this study are limited to those present within these 127 chains, amounting to 372 techniques, rather than encompassing all techniques cataloged in MITRE.

Sub-chain of Cyber Attacks

A significant contribution of this research is the extraction of all possible sub-chains from existing chains. As previously mentioned, each chain comprises a series of techniques. A sub-chain is defined as a subset of these techniques that has a length less than or equal to that of the main chain, with the sequence of techniques adhering to the order established in the main chain. In the initial phase, we extracted all two-member sub-chains. Two-member sub-chains are particularly important for three reasons. First, they allow us to identify the prerequisites for each technique. Second, they provide insight into which technique is typically positioned subsequent to another. Lastly, their third application pertains to the probability calculation, which will be discussed in the subsequent chapter.

In the subsequent step, all sub-chains were extracted from the main chains (comprising 127 chains), irrespective of their length. This task represents a fundamental component of this research. The number of possible sub-chains can be calculated using the formula:

N + (N-1) + (N-2) + … + 1 = number of sub-chains of length N

A chain with 4 elements (A, B, C, D) has 10 sub-chains in total. This is calculated as 1 + 2 + 3 + 4 = 10, representing all possible contiguous subsequences. This can be observed in Table 1.

Table 1: All Sub-chains Derived from a 4-Step Chain

Length1	Length 2	Length 3	Length 4
A	A, B	A, B, C	A, B, C, D
B	B, C	B, C, D	--
C	C, D	--	--
D	--	--	--

As previously indicated, we have identified a total of 372 distinct techniques, each of which has been arranged in various configurations to form a chain of attacks. We are now focused on extracting the specific placement of these techniques within the chains.

A total of 66,446 sub-chains were extracted from 127 main chains, yielding 61,878 unique sub-chains, all of which included the original main chains. The occurrence frequency of each sub-chain was subsequently calculated. Due to the impracticality of presenting the entirety of this data, we instead report the number of sub-chains generated and their respective lengths.

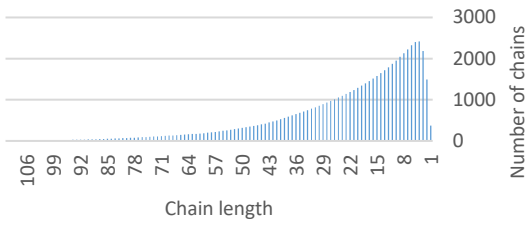


Figure 1: Frequency Distribution of Sub-chain Lengths

As indicated in Figure 1, there are 372 sub-chains of length 1, which corresponds to the total number of available techniques. The largest quantity of sub-chains, 2,414, corresponds to those of length 4. Beyond this point, the number of sub-chains progressively decreases, ultimately reaching a value of 1 at length 109, which represents the longest chain, G0032. Consequently, the lengths of the newly generated sub-chains range from 1 to 109.

Attack Chain Classification

In this section, we categorize the chains of cyber-attacks in various ways. Each classification method, which will be discussed in detail, examines the attack chain from a specific perspective and addresses different questions.

- Which Technique Usually Starts Chains?

A critical aspect of identifying a chain is determining its initial step. Knowing which technique typically initiates a chain can significantly enhance the identification process and increase the vigilance of the security operations center. Figure 2 illustrates the most frequently employed techniques at the onset of an attack chain, specifically those that have been repeated a minimum of three times.

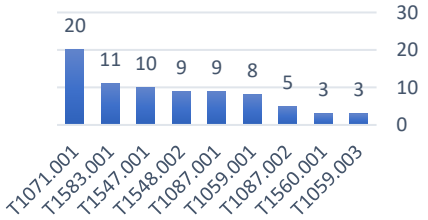


Figure 2: Most Frequent Techniques at the Start of the Chain

- Sharing Techniques

This method classifies chains based on the number of techniques they share. Chains with a greater number of common techniques are grouped together. However, a drawback of this approach is that as the length of the chain increases, the likelihood of that chain being assigned to a group also increases. The chain length thus plays a significant role in its classification. Since the order of techniques is not considered in this method, only the number of shared techniques is taken into

account, which can affect the accuracy of the classification.

- Technique Usage

This component reflects the total number of occurrences of the techniques that form a given attack chain (Procedure Examples) [11]. In this classification system, a shorter chain may see broader application than a longer one. The usage rate was calculated for each attack chain. The usage rates ranged from a minimum of 5 to a maximum of 1,841, and we defined six categories with an equal range of 310. The sixth category, which contained only four chains, G0050, G0016, G0094, and G0032, had usage values of 1,622, 1,679, 1,705, and 1,841, respectively, representing the most efficient chains.

The longest chain, G0032, is classified within the most efficient category, as chain length is recognized as a critical determinant of efficiency. To eliminate the influence of chain length from our classification, we normalized the usage rate by dividing the usage rate of each chain by its respective length. This adjustment ensured that the analysis was independent of the chain's length.

A particularly noteworthy outcome of this normalization was the discovery of an approximately normal distribution of usage, with category occurring every nine units. As illustrated in Figure 3, the average usage per chain falls within the range of 23 to 32 units, which is a critical observation. The most efficient chains, when normalized for length, are those whose usage exceeds 50 units, representing the upper portion of the final category. These include the chains G0013, G0063, G0062, G0103, G0133, and G0084.

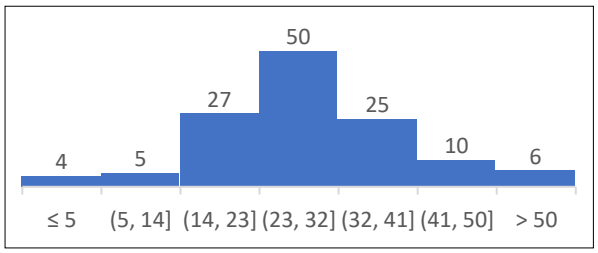


Figure 3: Chain effectiveness is independent of chain length.

- Common Sub-chains

The next objective is to identify sub-chains with a high frequency of occurrence. This task is beneficial from two perspectives. First, understanding the sequence of techniques will facilitate the prevention of attacks from forming chains. For instance, if the frequency of the consecutive techniques A, B, and C is notably high, then the occurrence of attacks A and B suggests a strong likelihood that attack C will follow. By leveraging this information, we can allocate resources more effectively and implement proactive measures to prevent subsequent attacks. Second, we aim to determine which chains a particular sub-chain belongs to. By calculating the number of occurrences and repetitions, we can identify sub-chains that serve as communication links between multiple chains. In the following section, examples are provided for each of the sub-chains with lengths ranging from 2 to 8, which can establish a connection between two or more chains. Table 2 presents these examples to clearly illustrate the concept discussed. For instance, the sub-chain T1082; T1016, with a length of 2, was found in 22 main chains.

Table 2: Sub-chains Common to Multiple Chains

Sub-chain length	Sub-chain	Frequency	Chain of attacks
2	T1082; T1016	22	G0010, G0012, G0018, G0022, G0034, G0038, G0049, G0050, G0059, G0069, G0073, G0081, G0094, G0101, G0102, G0114, G0116, G0121, G0126, G0128, G0129, G0139
3	T1016; T1049; T1033	13	G0004, G0022, G0027, G0032, G0034, G0049, G0050, G0069, G0081, G0093, G0096, G0114, G0116
4	T1082; T1016; T1049; T1033	8	4G0022, G0034, G0049, G0050, G0069, G0081, G0114, G0116
5	T1218.011; T1082; T1016; T1049; T1033	4	G0022, G0034, G0050, G0069
6	T1071.001; T1560.001; T1547.001; T1059.001; T1059.003; T1059.005	3	G0059, G0060, G0069
7	T1082; T1016; T1049; T1033; T1007; T1569.002; T1124	2	G0114, G0116
8	T1052.001; T1203; T1083; T1564.001; T1574.002; T1070.004; T1105; T1036.005	2	G0081, G0129

A total of 903 common sub-chains, ranging in length from 2 to 8, were extracted. Due to limitations in display capacity, only the number of common sub-chains extracted for each specific length is provided in Table 3.

Table 3: Total Number of Mined Sub-chains Common by Length

Sub-chain Length	Total Shared Sub-chains
2	465
3	284
4	102
5	34
6	11
7	6
8	1

Risk Assessment of the Attack Sequence

The majority of existing literature on risk assessment has primarily focused on the banking and financial sectors. However, there exists a significant gap in research that applies risk assessment frameworks to the field of cybersecurity. A review of the current body of work identifies two predominant methodologies: one that relies on data gathered through questionnaires, and another that utilizes data derived from network traffic or pre-existing datasets, such as NSL-KDD and MODBUS. A key distinction between the present study and those employing questionnaires lies in the scope and comprehensiveness of the input data. Specifically, the limited nature of questionnaire-based data raises concerns regarding the varying levels of expertise among respondents, as well as the potential biases introduced by the statistical population, which could affect the validity and generalizability of the findings. Numerous formulas exist for calculating security risks in the field of information security, the most fundamental of which is the formula: Risk = Probability × Impact [14]. Other formulas in this domain that have been adapted and refined for specific applications are categorized as Enhanced Risk Formulas. The formula we propose for calculating risk incorporates three components: the impact of each technique, the probability of that technique occurring, and the methods for mitigating the impact of that technique.

To assess the impact of the techniques, we drew on a comprehensive report [15] from the Center for Threat-Aware Defense, which provided valuable insights and data. The report assigned numerical scores to each technique, which we incorporated into our analysis.

Regarding mitigation measures [16], we systematically extracted relevant data from the MITRE website and analyzed the correlation between these measures and the associated techniques. Some of these measures prove to be more effective, as they mitigate the impact across a greater number of techniques. To highlight the most efficient methods, Table 4 presents the top eight mitigation measures.

Table 4: Most Frequent Mitigation Methods

Mitigation	M1031	M1042	M1038	M1047	M1056	M1018	M1026
Frequency	55	56	61	65	76	79	101

Another table has been compiled to illustrate the inverse relationship between these measures and techniques. Table 5 presents the techniques associated with the greatest number of preventative methods, suggesting that these techniques are comparatively easier to mitigate than others.

Table 5: Techniques with the Greatest Number of Mitigation Methods

Technique	T1059	T1557	T1563.002	T1021.001	T1210	T1003	T1072
Frequency	7	7	7	8	8	9	9

There are also 65 techniques for which no reduction methods are available. These techniques are particularly significant because, if they were present within an attack chain, there would be no means to counter them at that specific stage of the attack. It is important to note that the combination of these 65 techniques forms 176 sub-chains, highlighting their critical nature, as there would be no feasible way to prevent them at any stage of the attack.

In calculating the probability of an attack chain, we employed two distinct methodologies, both of which will be elaborated upon in the following sections.

First Method in Probability: First-Order Markov Property

A crucial aspect of this research is the application of conditional probability to determine the likelihood of each chain occurring. In this section, we will first explain the principles and definitions of conditional probability, followed by its application in our analysis.

Conditional probability measures the likelihood of event A occurring given that event B has occurred, denoted as P(A|B):

$$P(A|B) = \frac{\mathrm{P(A \cap B)}}{\mathrm{P(B)}}$$

where P(A∩B) represents the joint probability of both events A and B occurring. While this concept is initially applied to two events, it can be expanded to more complex scenarios, forming the basis for Bayesian network models (Figure 4) that analyze the relationships between multiple events.

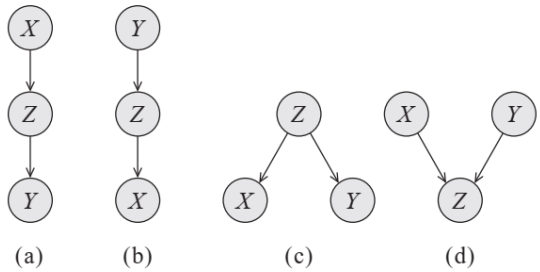


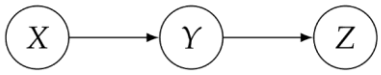
Figure 4: Types of Bayesian Network models

Among the various models used to represent sequences of events, this study focuses specifically on the chain model (also known as the linear model) due to its close alignment with the structure of attack chains in cybersecurity. In this context, the occurrence of each event (or node) in the chain is assumed to depend on its preceding events. Formally, for an N-step chain where steps are

dependent on each other, the overall probability of the chain can be expressed using the chain rule of probability, which states that it is the product of the conditional probabilities of each node given all previous nodes in the sequence:

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_n | X_1, \ldots, X_{i-1})$$

For example, the probability of a chain consisting of three sequential steps can be computed by applying the chain rule as follows:



$$P(x, y, z) = P(X)P(Y|X)P(Z|X, Y)$$

By applying a simplification, it is evident that each step in the chain depends solely on its immediate predecessor. This assumption, known as the first-order Markov property, effectively captures the sequential dependency structure while eliminating the need to condition on the entire history of prior steps [17]. Consequently, we obtain:

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

To estimate the probability of an attack chain using the outlined approach, two main steps are required. First, the marginal probability of occurrence for the initial technique in the chain is calculated. Second, the conditional probabilities for all ordered pairs of techniques (referred to as 2-step sub-chains) are computed.

In the first step, the dataset included 372 distinct techniques with a total of 2,929 occurrences, forming the basis of the sample space for this analysis. The probability of each individual technique was derived from its relative frequency within the dataset. Table 6 summarizes a selection of techniques with the highest estimated marginal probabilities.

Table 6: Most Frequently Repeated Techniques

Technique	Frequency	Probability
T1059.001	62	0/021168
T1027	65	0/022192
T1105	67	0/022875
T1566.001	68	0/023216
T1204.002	71	0/02424

In the second step, the conditional probabilities of all ordered pairs of consecutive techniques—referred to as 2-member sub-chains—within each attack chain are computed. For example, a chain comprising four techniques contains three such 2-step transitions. From a dataset of 127 attack chains, a total of 2,803 ordered 2-member sub-chains were extracted, corresponding to 1,492 unique transitions. Table 7 summarizes the five most frequently observed 2-member sub-chains.

Table 7: Most Frequently Repeated Two-Element Subchains

Technique	Frequency	Probability
T1082, T1016	22	0/007852
T1059.003, T1059.005	27	0/009636
T1566.001, T1566.002	30	0/010707
T1204.001, T1204.002	30	0/010707
T1059.001, T1059.003	40	0/014276

Ultimately, the overall probability of an attack chain is computed by multiplying the marginal probability of the initial technique with the conditional probabilities of each consecutive 2-member sub-chain, reflecting the first-order Markov assumption.

The Second Method of probability: Sample Space

The total number of sub-chains identified was considered the sample space, and the probability of each sub-chain was determined by dividing its frequency of occurrence by this total. It is important to note that this probabilistic analysis would not have been feasible without first generating all sub-chains and accurately recording the frequency of each instance.

Another critical consideration involves the inclusion of mitigation in the denominator of the risk formula. As previously noted, there are 65 techniques for which no mitigation methods are available. In these cases, the denominator of the fraction would become zero. To resolve this issue, we added a unit to all mitigation values. This adjustment ensures that for techniques with a mitigation value of zero, the denominator becomes one, thus preserving the validity of the fraction's result. Since the same unit was added to all values, the overall outcome of the calculations remains unaffected. Additionally, the final result of these calculations is multiplied by 1000 for presentation purposes, yielding the following:

$$\text{Risk} = \frac{\text{Probability} \times \text{Impact}}{\text{Mitigation}+1} \times 1000$$

Risk calculations are conducted not only for the primary chains but also for all 61,878 sub-chains that were generated. This approach allows for the visualization of the risk value at each step across all main chains until the endpoint is reached, enabling a detailed examination of how the risk evolves throughout the chain.

Due to the volume of calculations, Table 8 presents the G0001 chain as a representative example. The results show no direct correlation between chain length and risk value, reflecting the varying effectiveness of mitigation techniques across steps.

The next analysis conducted on the chains and their associated risks allows us to select a specific number of steps and observe the risk levels based on those steps. For example, we can determine which techniques should be employed at specific steps and identify the chains that meet these criteria along with their corresponding risk levels. To illustrate this,

consider the scenario where we are searching for 4-step chains that utilize the T1005 technique in the second step and the T1140 technique in the fourth step. The results of these conditions, including the associated risk levels, are presented in Table 9. As shown, there are six chains that satisfy these conditions.

Table 8: Risk Value at Each Step in the G0001 Chain

Risk	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8	Step 9	Step 10	Step 11	Step 12	Step 13	Step 14	Step 15	Step 16
0/000376245	T1583.002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0/00050166	T1583.002	T1583.003	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0/003396447	T1583.002	T1583.003	T1560	0	0	0	0	0	0	0	0	0	0	0	0	0
0/002867656	T1583.002	T1583.003	T1560	T1584.005	0	0	0	0	0	0	0	0	0	0	0	0
0/003590056	T1583.002	T1583.003	T1560	T1584.005	T1005	0	0	0	0	0	0	0	0	0	0	0
0/003476472	T1583.002	T1583.003	T1560	T1584.005	T1005	T1001.002	0	0	0	0	0	0	0	0	0	0
0/003059804	T1583.002	T1583.003	T1560	T1584.005	T1005	T1001.002	T1189	0	0	0	0	0	0	0	0	0
0/002812627	T1583.002	T1583.003	T1560	T1584.005	T1005	T1001.002	T1189	T1546.008	0	0	0	0	0	0	0	0
0/002891525	T1583.002	T1583.003	T1560	T1584.005	T1005	T1001.002	T1189	T1546.008	T1190	0	0	0	0	0	0	0
0/003069262	T1583.002	T1583.003	T1560	T1584.005	T1005	T1001.002	T1189	T1546.008	T1190	T1203	0	0	0	0	0	0
0/003114951	T1583.002	T1583.003	T1560	T1584.005	T1005	T1001.002	T1189	T1546.008	T1190	T1203	T1003	0	0	0	0	0
0/002917878	T1583.002	T1583.003	T1560	T1584.005	T1005	T1001.002	T1189	T1546.008	T1190	T1203	T1003	T1566	0	0	0	0
0/002640972	T1583.002	T1583.003	T1560	T1584.005	T1005	T1001.002	T1189	T1546.008	T1190	T1203	T1003	T1566	T1563.002	0	0	0
0/00247527	T1583.002	T1583.003	T1560	T1584.005	T1005	T1001.002	T1189	T1546.008	T1190	T1203	T1003	T1566	T1563.002	T1021.001	0	0
0/002441298	T1583.002	T1583.003	T1560	T1584.005	T1005	T1001.002	T1189	T1546.008	T1190	T1203	T1003	T1566	T1563.002	T1021.001	T1553	0
0/002630432	T1583.002	T1583.003	T1560	T1584.005	T1005	T1001.002	T1189	T1546.008	T1190	T1203	T1003	T1566	T1563.002	T1021.001	T1553	T1078

Table 9: Filtered Output of Subchains

Step 1	Step 2	Step 3	Step 4	risk
T1543.003	T1005	T1074.001	T1140	0/0037942
T1213	T1005	T1025	T1140	0/0049478
T1132.001	T1005	T1491.002	T1140	0/0055273
T1555.003	T1005	T1074.001	T1140	0/0066062
T1132.001	T1005	T1039	T1140	0/0071324
T1555	T1005	T1074.001	T1140	0/0095395

ML-based prediction

One of the primary objectives of this research was to predict potential attacks. Thus far, we have observed that the trends in attack chains do not always align with our predictions and may diverge from our expectations. In [18], various prediction methods were evaluated and categorized into three groups: statistical methods, probabilistic methods (including Markov models and Bayesian networks), and machine learning techniques (such as artificial neural networks and genetic algorithms). This analysis uses the CDX2009 dataset, which contains reports generated by the Snort intrusion detection system. Upon comparing these methods, the study concluded that the Long Short-Term Memory (LSTM) algorithm outperformed other techniques in terms of predictive accuracy. To further assess the performance of the LSTM algorithm in prediction tasks, we reviewed additional studies that benchmarked LSTM against other methods. For instance, in [19], techniques such as decision trees, Naive Bayes, random trees, random forests, multilayer perceptron, and support vector machines were compared to LSTM. In this analysis, the NSL-KDD dataset was employed. As illustrated in Figure 5, the results consistently demonstrate the superior performance of the LSTM algorithm across these comparisons.

The primary objective of employing machine learning in this study is to predict the quantitative risk level associated with cyber-attack chains. To enable this, each attack chain was annotated with a corresponding numerical risk value and encoded into a structured format suitable for sequential modeling. This design allows the model to learn both the frequent dependencies among attack techniques and the underlying relationship between each technique and its associated risk. Risk scores for the sub-chains were calculated during earlier stages and served as supervised targets for training a Long Short-Term Memory (LSTM) neural network. As a supervised learning approach, the LSTM requires the dataset to be divided into distinct subsets, with 80% used for training and 20% for testing. Once trained, the model can analyze sequences of techniques—including previously unseen attack chains—and predict a risk score that reflects the severity or potential impact of the sequence. Since the task involves predicting a real-valued score, it is formulated as a regression problem, with performance evaluated using metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). Alternatively, risk scores can be categorized into discrete classes representing different priority levels, thereby converting the problem into a classification task. In such cases, evaluation metrics such as the F1 Score and Receiver Operating Characteristic (ROC) curve are applicable. The model was implemented using the Keras library, with Mean Absolute Error as the loss function and the Adam optimizer for efficient convergence. Training was performed over 10 epochs with a batch size of 32.

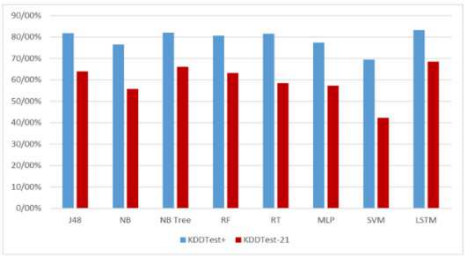


Figure 5: Accuracy Comparison of Machine Learning Algorithms on the KDDTest Dataset

The performance of the final regression model was assessed using several standard evaluation metrics. During training, the model attained a mean squared error (MSE) of 2.2892×10^{-5} and a mean absolute error (MAE) of 0.0013. On the validation dataset, the model yielded a slightly lower MSE of 1.7869×10^{-5} and a corresponding MAE of 0.0034, suggesting effective generalization and no evident overfitting. Evaluation on the independent test set resulted in an MAE of 0.0035 and a root mean squared error (RMSE) of 0.0062, indicating a high level of predictive accuracy with minimal average deviation from true values. Collectively, these results underscore the model's robustness and suitability for the target regression task. For example, when the model was provided with the synthetic sequence "T1213, T1005, T1039, T1140", which does not appear in the original dataset, it predicted a risk level of 0.0034.

To better interpret the model's output, the continuous risk values were categorized into five discrete classes—very low, low, medium, high, and very high risk—using K-Means clustering. This classification strategy enables clearer interpretation of risk levels and supports meaningful comparison with other research in the cybersecurity domain.

The final classification model demonstrated robust performance, achieving a validation accuracy of 96.63% and a training accuracy of 96.11%. The low loss values across both sets indicate stable learning and effective generalization. The model excelled particularly on the dominant risk class, with an F1-score of 0.98, and attained a weighted average F1-score of 0.96, underscoring its reliability across the dataset. Additionally, a macro-ROC AUC score of 0.965 was achieved, reflecting strong overall class separability. These results affirm the effectiveness of the proposed approach in accurately predicting cybersecurity risk levels and support its practical applicability in real-world scenarios.

To contextualize the methodological and performance contributions of the proposed model, a comparative analysis was conducted against recent studies in sequence-based cyberattack modeling. To the best of our knowledge, no prior research has explicitly employed MITRE ATT&CK techniques for the purpose of multi-stage cyber risk regression and subsequent classification. Consequently, the literature review includes works that utilize LSTM-based architectures or operate on MITRE-aligned threat frameworks. Zhou et al. [20] proposed a sequence-to-sequence model based on an LSTM encoder–decoder architecture to detect multi-stage cyberattacks from IDS alert sequences. Their model effectively captured temporal dependencies between attack stages and achieved strong performance on datasets such as DARPA2000 and ISCXIDS2012. However, their approach focused on predicting stage transitions rather than estimating or classifying cyber risk levels, and it did not incorporate structured threat taxonomies like MITRE ATT&CK. Kotenko et al. [21] applied an LSTM classifier to attribute cyberattacks to 66 distinct teams using BASH command sequences from the BAKKA CTF dataset. Their model achieved a training accuracy of 61% and a validation accuracy of 48%, underscoring the difficulty of attacker attribution in the presence of overlapping operational patterns and limited discriminative features. Yuan et al. [22] introduced the DeepOP framework, which shares conceptual alignment through its use of MITRE ATT&CK sequences and a hybrid architecture combining BiLSTM, Transformer encoder layers, and ontology-based reasoning to predict future attack steps. However, DeepOP is designed for sequence continuation rather than risk estimation or classification, and it does not report standard classification metrics such as ROC AUC or F1-score. Collectively, these comparisons highlight the novelty and effectiveness of the proposed approach in integrating structured attack-stage modeling with interpretable, high-accuracy cyber risk prediction. A structured summary of these studies and comparative metrics is presented in Table 10.

Table 10: Comparative Analysis of Cyberattack Modeling Approaches

Study	Task	Dataset	Architecture	Performance	Risk Modeling
Zhou et al.	Multi-stage attack detection	DARPA2000, ISCXIDS2012	LSTM Sequence-to-Sequence	F1-score ≈ 0.99 (binary classification)	✗
Kotenko et al.	Attacker attribution (66 teams)	BAKKA CTF	LSTM classifier	61% training accuracy, 48% validation accuracy	✗
Yuan et al.	Attack chain prediction	MITRE-aligned synthetic sequences	BiLSTM + Transformer + Ontology Reasoning	Hit@3 ≈ 0.72, NDCG@3 ≈ 0.60 (no F1/ROC AUC)	✗
Proposed Model	Risk regression + classification (5 classes)	MITRE-based synthetic sequences	LSTM + Regression + K-Means	96.6% validation accuracy, 0.96 F1, 0.965 ROC AUC	✓

Conclusion

This study advances the understanding and quantification of security risks in cyber-attack chains, addressing a key gap in existing research. While prior work within the MITRE ATT&CK framework has focused on attack chain modeling, comprehensive classification and quantitative risk assessment remain limited. The proposed approach first computes the probability of each sub-sequence using the first-order Markov property, integrating technique impact scores and mitigation values to derive numerical risk scores. These quantified risk values are then modeled and predicted using a Long Short-Term Memory (LSTM)–based temporal framework, enabling accurate forecasting of future risks. The predicted continuous values are subsequently transformed into interpretable risk levels through K-Means clustering, enhancing both operational applicability and interpretability. By capturing sequential dependencies, probabilistic risk factors, and mitigation effects, the method enables precise prioritization of defensive measures. The findings underscore the novelty and effectiveness of the method, demonstrating how the integration of temporal dependency modeling with structured risk stratification yields robust, interpretable, and operationally meaningful outputs for advancing cyber defense in complex threat environments.

Funding

The authors did not receive support from any organization for the submitted work.

References

[1] I. THE RADICATI GROUP, "Advanced Persistent Threat (APT) Protection- Market Quadrant 2024."

[2] A. Ahmad, J. Webb, K. C. Desouza, J. J. C. Boorman, and Security, "Strategically-motivated advanced persistent threat: Definition, process, tactics and a disinformation model of counterattack," vol. 86, pp. 402-418, 2019.

[3] J. Kutscher. (2020). M-Trends 2020: Insights From the Front Lines. Available: https://www.mandiant.com/resources/blog/mtrends-2020-insights-from-the-front-lines

[4] IBM, "Cost of a Data Breach Report," 2024.

[5] Y. Ahmed, A. Taufiq, R. J. C. Md Arafatur, Materials, and Continua, "A Cyber Kill Chain Approach for Detecting Advanced Persistent Threats," vol. 67, no. 2, pp. 2497-2513, 2021.

[6] C. Jaewon Lee, C. CGEIT, and C. CIA, "An Enhanced Risk Formula for Software Security Vulnerabilities," 2014.

[7] A. N. Desk. (2020). MITRE ATT&CK Improves Cloud Security, Yet Many Enterprises Struggle to Implement It. Available: https://aithority.com/it-and-devops/cloud/study-mitre-attck-improves-cloud-security-yet-many-enterprises-struggle-to-implement-it/

[8] A. Georgiadou, S. Mouzakitis, and D. J. S. Askounis, "Assessing mitre att&ck risk using a cyber-security culture framework," vol. 21, no. 9, p. 3267, 2021.

[9] A. Pawlicka, M. Pawlicki, R. Kozik, and R. S. J. S. Choraś, "A systematic review of recommender systems and their applications in cybersecurity," vol. 21, no. 15, p. 5248, 2021.

[10] R. Al-Shaer, J. M. Spring, and E. Christou, "Learning the associations of mitre att & ck adversarial techniques," in 2020 IEEE Conference on Communications and Network Security (CNS), 2020, pp. 1-9: IEEE.

[11] MITRE-ATT&CK®. Frequently Asked Questions. Available: https://attack.mitre.org/resources/faq/

[12] K. Shoushian, A. J. Rashidi, and M. Dehghani, "Modeling of cyber-attacks obfuscation, based on alteration technique of attack %J Journal of Electronic and Cyber Defense," (in Fa), vol. 8, no. 1, pp. 67-77, 2020.

[13] MITRE-ATT&CK®. ATT&CK v11.3. Available: https://attack.mitre.org/versions/v11/

[14] R. Borgovini, S. Pemberton, and M. Rossi, "Failure mode, effects, and criticality analysis (FMECA)," RELIABILITY ANALYSIS CENTER GRIFFISS AFB NY1993.

[15] T. C. f. T.-I. Defense, "Top ATT&CK Techniques."

[16] MITRE-ATT&CK®, "Enterprise Mitigations."

[17] D. Koller and N. Friedman, Probabilistic graphical models: principles and techniques. MIT press, 2009.

[18] P. Soleimani Barijani, A. J. Rashidi, and K. Dadash Tabar Ahmadi, "Presenting a suitable algorithm for predicting cyber attacks based on learning," presented at the Second National Conference on Electrical and Computer Engineering, 2017. Available: https://civilica.com/doc/731172

[19] S. Dami and B. Ghorbani Kondroodi, "Network intrusion detection using long short-term memory (LSTM) neural network," presented at the The Third National Conference on Innovation and Research in Electrical Engineering, Computer Engineering and Mechanical Engineering in Iran, 2017. Available: https://civilica.com/doc/741156

[20] P. Zhou, G. Zhou, D. Wu, M. J. C. Fei, and Security, "Detecting multi-stage attacks using sequence-to-sequence model," vol. 105, p. 102203, 2021.

[21] I. Kotenko, E. Fedorchenko, E. Novikova, and A. J. S. Jha, "Cyber attacker profiling for risk

analysis based on machine learning," vol. 23, no. 4, p. 2028, 2023.
[22] S. Zhang, X. Xue, and X. J. E. Su, "DeepOP: A Hybrid Framework for MITRE ATT&CK Sequence Prediction via Deep Learning and Ontology," vol. 14, no. 2, 2025.