UDC 004.7.056.5

DETECTING CYBER-ATTACKS BASED ON NETWORK STEGANOGRAPHY

[1]Duong Ngoc Anh, [2]Dao Thanh Huong
Educational Institution "Belarusian State University of Informatics and Radioelectronics", Minsk, Republic of Belarus

Annotation: With the rapid development of information technology and the growing number of network attacks, ensuring the security of computer networks is particularly important. One of the covert methods of data transmission is network steganography, which allows information to be embedded in legitimate network traffic. Such approaches are used by attackers to control malware, establish covert communication channels, and leak confidential data. This paper examines the main attack methods based on network steganography and analyzes modern approaches to their detection.

Keywords: network steganography; covert channel detection; steganalysis; information hiding; machine learning; deep learning; multilayer hybrid detection; Zero Trust.

Introduction

Network steganography is a type of steganography that uses network protocols of the Open Systems Interconnection (OSI) reference model as carriers of secret data. Generally, network steganography is a family of methods for modifying data in network protocol headers and packet payload fields, altering the structure of packet transmission, and using hybrid methods within a particular network protocol (sometimes several at once) [1].

The fundamental principle of network steganography is to utilize inherent features of network communication in order to transmit information without raising suspicion. These techniques can involve modifying protocol fields, manipulating packet payloads, or altering temporal characteristics such as inter-packet delays. As a result, network steganography enables attackers to bypass traditional security mechanisms, making it a critical challenge in the field of cybersecurity.

Main part

Network steganography methods can be classified according to how they modify protocol data units (PDUs), and are generally divided into two main categories: intra-protocol techniques and inter-protocol techniques. Intra-protocol methods operate within a single protocol and can be grouped into three main approaches: modifying protocol PDU fields, hybrid methods, and modifying PDU timing relationships. These approaches include techniques such as payload embedding (e.g., DNS tunneling), protocol header manipulation (e.g., IP, TCP, SCTP fields), and timing-based strategies such as packet reordering, intentional packet loss, or delay alteration. In contrast, inter-protocol methods exploit the interaction between multiple protocols, with PadSteg being a prime example.

To clarify further, here are some specific examples:

1. Payload-based methods. They embed hidden data directly into packet payloads. A common example is DNS tunneling, where data is encoded within domain names in DNS queries, enabling covert data transmission through legitimate traffic. Another approach involves modifying the least significant bits of application-layer data (e.g., HTTP or VoIP), making the changes difficult to detect.

2. Protocol-based methods. Protocol-based techniques modify fields within network protocol headers, often using unused or optional fields in IP, TCP, or UDP. For example, HICCUPS introduces intentional checksum errors in wireless networks to transmit hidden data, which can only be decoded by specific receivers. Similarly, PadSteg uses Ethernet frame padding to carry covert information, as these fields are typically ignored.

3. Timing-based methods. These methods encode information by altering packet timing rather than content. The LACK method delays selected VoIP packets so they appear lost to normal receivers but are used by covert ones. In TCP packet dropping signaling, packet loss patterns represent data. Another technique modifies inter-PDU delays, where different time intervals correspond to encoded information.

Given the diversity of network steganography techniques, no single detection method can address all threats. Each approach – whether signature-based, anomaly-based, machine learning-driven, or active defense – offers distinct advantages while suffering from specific limitations. Signature-based methods provide reliable detection of known attacks but cannot identify novel variants. Anomaly-based techniques can uncover unknown threats yet generate high false positive rates. Machine learning and deep learning achieve impressive accuracy but demand substantial data and computational resources. Multilayer approaches balance these trade-offs at the cost of architectural complexity, while active defense mechanisms prevent attacks proactively but may impact network performance.

A comparative understanding of these approaches is therefore essential for selecting appropriate countermeasures. As steganographic attacks continue to evolve in sophistication, understanding the trade-offs between different detection methods becomes increasingly crucial. Table provides a structured comparison of the approaches discussed in this paper, summarizing fundamental principles of network steganography detection methods.

Comparison of Network Steganography Detection Approaches

Approach	Main Methods	Fundamental Principle	Strengths	Limitations
Signature-based Detection	Pattern matching.Detection of traces from known steganography tools	Relies on a database of signatures extracted from known hiding techniques	Simple implementation. High accuracy for known attacks	Unable to detect novel/zero-day attacks. Requires frequent signature updates
Anomaly-based Detection	Statistical analysis (entropy, frequency). Inter-packet delay analysis. Detection of abnormal header values	Establishes baseline of normal protocol behavior and identifies deviations	Capable of detecting unknown attacks. Independent of signature databases	High false positive rate. Difficult to determine optimal thresholds

Completing the table

Approach	Main Methods	Fundamental Principle	Strengths	Limitations
Machine Learning Detection	Supervised learning (Random Forest, SVM). Unsupervised learning (Clustering)	Trains classification models on labeled datasets to distinguish malicious traffic	Automates detection process. Adaptable to various attack types	Requires large, high-quality training data. Limited interpretability ("black box" issue)
Deep Learning Detection	Convolutional Neural Networks (CNN). Recurrent Neural Networks (RNN/LSTM)	Automatically extracts hierarchical features from raw data to detect complex patterns	Highest detection accuracy. Eliminates manual feature engineering	High computational resource demands. Requires massive amounts of training data
Multilayer/Hybrid Detection	Combination of multiple techniques. Selective analysis based on suspicion level	Performs lightweight analysis at upper layers, triggers deep inspection only when anomalies are detected	Optimizes system performance. Balances accuracy and processing speed	Complex architectural design. Challenging parameter optimization
Active Defense (Wardens)	Packet normalization. Overwriting unused/reserved header fields	Proactively sanitizes packets to destroy hidden data before reaching destination	Complete prevention capability. No detection required	May degrade network performance. Cannot distinguish sophisticated attacks

Conclusion

This paper has presented a comprehensive overview of detection methods for network steganography-based cyber attacks. The comparative analysis reveals that each approach possesses distinct strengths and limitations. Signature-based detection offers simplicity and high accuracy for known attacks but fails to identify novel threats. Anomaly-based detection can uncover unknown variants yet suffers from high false positive rates and threshold configuration challenges. Machine learning and deep learning techniques achieve superior accuracy through automated feature extraction; however, they

demand substantial computational resources and large, high-quality training datasets.

For practical deployment in real-world environments, multilayer/hybrid detection emerges as the most promising direction. By combining multiple detection layers with increasing complexity, this approach effectively balances detection accuracy, processing speed, and resource utilization – critical factors for operational network infrastructures. Furthermore, the continued advancement of artificial intelligence and deep learning will remain central to enhancing detection capabilities against increasingly sophisticated steganographic techniques.

Notably, as traditional network perimeters continue to dissolve, the Zero Trust model – with its core principle of "never trust, always verify" – becomes increasingly relevant. Zero Trust architecture mandates inspection of every packet and data stream regardless of origin, enabling the discovery of even the most concealed covert channels. The integration of advanced detection techniques within a Zero Trust framework creates a robust defense mechanism, protecting systems against the growing threat landscape of steganography-based attacks.

References

1. Peskova O.Yu., Khalaburda G.Yu. Application of network steganography to conceal data transmitted over communication channels // Bulletin of the Southern Federal University. Technical Sciences, 2012. – T. 137. – № 12 (137). – C. 167-176.

2. N. Singh, J. Bhardwaj, and G. Raghav. Network Steganography and its Techniques: A Survey, International Journal of Computer Applications, vol. 174, no. 2, pp. 30-35, Sep. 2017. [Electronic resourse].

3. Badar, L. T., Carminati, B., & Ferrari, E. (2025). A comprehensive survey on stegomalware detection in digital media, research challenges and future directions. Signal Processing, 231, 109888.

Information about the authors

Duong N.A., student of the Department of Information Security, Educational Institution "Belarusian State University of Informatics and Radioelectronics", dngocanh2412@gmail.com.

Dao T.H., student of the Department of Information Security, Educational Institution "Belarusian State University of Informatics and Radioelectronics", daothanhhuong2k3ad@gmail.com.