Postprint: Anonymous Communication Mechanism Based on Improved Network Coding

Authors: Yang Kang, Zhai Jiangtao, Dai Yuewei

Date: 2018-11-29T00:00:00+00:00

Abstract

Existing network-coding-based anonymous systems suffer from low data transmission success rates due to the instability of forwarding nodes and multi-path forwarding of anonymous messages. This paper proposes ACSNC (anonymous communication system based on network coding), a novel anonymous communication mechanism based on network coding and data redundancy methods. First, a data redundancy mechanism is employed to perform redundant fragmentation on the data to be sent and transmit the fragments along multiple paths. Then, intermediate forwarding nodes randomly encode the information before forwarding, causing the statistical characteristics of the information to change as it passes through the nodes. Finally, the destination node recovers the anonymous information based on the received data fragments and encoding coefficients. Simulation results demonstrate that this mechanism significantly improves data transmission success rates while effectively ensuring communication anonymity and security.

Full Text

Preamble

Title: Research on an Anonymous Communication Mechanism Based on Improved Network Coding

Authors: Yang Kang, Zhai Jiangtao†, Dai Yuewei
(School of Electronics & Information, Jiangsu University of Science & Technology, Zhenjiang, Jiangsu 212003, China)

Abstract: Existing anonymous systems based on network coding suffer from low data transmission success rates due to the instability of forwarding nodes and multipath forwarding of anonymous messages. To address this issue, this paper proposes a new anonymous communication mechanism called ACSNC

(Anonymous Communication System based on Network Coding). The mechanism first employs a data redundancy mechanism to redundantly fragment the data to be sent and transmit it along multiple paths. Intermediate forwarding nodes then randomly encode the information before forwarding, causing the statistical characteristics of the information to change as it passes through the nodes. Finally, the destination node recovers the anonymous information based on the received data fragments and coding coefficients. Simulation results demonstrate that this mechanism can effectively guarantee communication anonymity and security while significantly improving data transmission success rates.

Keywords: anonymous communication; network coding; data redundancy; multipath; network security

Introduction

Since David Chaum first proposed anonymous communication technology in 1981 [3], scholars have continuously discussed and developed this field for over three decades, yielding numerous classical research achievements. Ubiquitous network surveillance and traffic censorship pose severe threats to covert communication on the Internet. When transmitting sensitive data, people desire not only transparent data transfer but also anonymity for both communicating parties. However, network communication itself does not provide identity protection, making anonymous communication mechanisms a research hotspot for identity information protection [1].

Anonymous communication refers to the use of technical means to protect the relationship between communicating parties during transmission, preventing attackers from directly obtaining or indirectly inferring the communication relationship or any party' s identity information [2]. Mix-Net technology originated from this concept. Traditional anonymous communication systems generally rely on trusted key infrastructure and require third-party participation in key negotiation, making overall system security heavily dependent on key security. Moreover, traditional anonymous systems struggle to adapt to network environments without a key trust center. Therefore, research on anonymous communication without a key center holds significant importance.

Reference [9] implements anonymous path construction without a key center through source information segmentation strategies. Analysis shows that this anonymous system significantly improves attack resistance during the path construction phase, but it does not employ coding techniques to process anonymous information during the information transmission phase. Building upon this, reference [10] further proposes transmitting anonymous communication information through information segmentation and network coding. However, the encoding and forwarding of anonymous messages by forwarding nodes in this approach increases system delay, while multipath transmission of anonymous messages reduces the success rate of anonymous message delivery. To address these problems, the anonymous communication mechanism proposed in this paper combines the advantages of current network coding, eliminating the need for a key mechanism and reducing system deployment costs. Simultaneously, the introduced data redundancy mechanism improves the success rate of anonymous information transmission and enhances system efficiency and stability.

1.1 Architecture

Network coding [11,12] is an information exchange technology that integrates routing and coding. Its core idea is to perform linear or nonlinear processing on information received from various channels at each node in the network before forwarding it to downstream nodes, with intermediate nodes acting as encoders or signal processors. Ahlswede et al., using the butterfly network [13] as a case study, demonstrated that network coding can achieve the maximum flow bound of multicast routing transmission, thereby improving information transmission efficiency. Network coding technology can enhance existing network throughput while also improving network reliability and attack resistance. In anonymous communication, processing transmitted information through encoding at intermediate nodes changes the form and content of data flows as they enter and exit nodes, preventing attackers from tracking information flows. However, existing network coding-based multicast scenarios struggle to guarantee system data transmission success rates. To address this, this paper improves the anonymous communication mechanism based on network coding by introducing a data redundancy mechanism.

The data redundancy fragmentation mechanism is illustrated in [Figure 1: see original paper]. The file is read and fragmented into n fragments. XOR operations are then performed on each row of fragments, with results saved to auxiliary fragments. XOR operations are also performed on each column, with results saved to auxiliary fragments. Ultimately, $n + m + 1$ fragments are obtained. Randomly deleting k fragments may delete original fragments or auxiliary fragments, yet the original data can still be recovered using the remaining fragments [14].

As shown in [Figure 2: see original paper], the source sending node is S, the destination node is O, and forwarding nodes are $N_1, N_2, \ldots, N_{m-1}$. To construct an anonymous path $S \to N_1 \to N_2 \to \cdots \to N_{m-1} \to O$, when source node S sends path construction information to N_1, it randomly selects n nodes (excluding path construction nodes) to build n coding forwarding nodes through interaction with the directory server. Subsequently, source node S encodes and forwards the path construction information to destination node O. Repeating the above steps for other nodes' routing information construction, each forwarding node in the entire coding forwarding network eventually obtains the address of its successor node.

1.2 Design Philosophy

After establishing the anonymous path, the source sending node uses a data redundancy mechanism to fragment messages and send them through multiple paths. Upon receiving anonymous data, forwarding nodes encode and forward data packets using random coding coefficients, changing the statistical characteristics of data packets as they flow into and out of the node. The random coding coefficients and encoded data are then forwarded together to the next node until reaching the destination node. Finally, the receiver obtains all message fragments and random coding coefficients to recover the anonymous message through corresponding decoding rules. The system architecture is shown in [Figure 2: see original paper].

1.3 Improved Method

This section details the communication process of the ACSNC mechanism. lists the definitions of commonly used symbols in the system.

Table 1 Symbol Definitions in the System

1) **Anonymous Forwarding Network Establishment**: The routing construction of forwarding nodes in the anonymous network primarily draws on the anonymous path construction mechanism based on network coding from reference [9]. However, reference [9]'s multipath forwarding of coding coefficients increases system complexity and delay. This paper improves the forwarding mechanism by having forwarding nodes forward through the current path, reducing delay while improving system fault tolerance. Ultimately, each forwarding node only knows the address of its next node.

2) **Anonymous Message Forwarding Phase**:

    - The source sending node first performs redundancy processing on anonymous data. The redundant coding model is shown in [Figure 3: see original paper]. The anonymous data M is split into n data packets. Then, through XOR algorithms, 1 or multiple redundant data packets are generated. The n data packets and generated redundant data are sent through m different links.
    - Forwarding nodes N_i in the anonymous link, after receiving anonymous data from the previous node, select random coding coefficients $p_i(k)$ to encode the received anonymous data $M_i(k)$. The encoded data $M_i(k+1)$ and random coding coefficients $p_i(k)$ are then sent together to the next node N_{i+1}.
    - The destination node receives all or partial fragment data $M_m(k)$ and random coding coefficients $p_m(k)$. It first decodes each message fragment using all obtained coding coefficients, then decodes the original anonymous data message through the redundancy algorithm.

The anonymous communication process is shown in [Figure 4: see original paper].

2.1 Anonymity Analysis

1) Identity Anonymity: Each node in the system only knows its predecessor and successor nodes. Even if a node is malicious, observing its data packets cannot reveal the real addresses of the source and destination nodes. Furthermore, through encoding and forwarding by intermediate nodes, this system' s links become untraceable, further protecting the real addresses of source and destination nodes.

2) Communication Anonymity: The source node fragments communication messages through the redundancy mechanism and transmits them via different paths. Attackers cannot easily control nodes on all paths to reconstruct the communication message. The probability of attackers obtaining communication messages through collusion attacks is defined as anonymity degree D. A smaller D value indicates stronger system anonymity [17].

In the coding forwarding network, to obtain communication messages, attackers must control one node on each of the m links and have these m nodes share data to decode and recover the source communication message. The probability of attackers obtaining communication messages through collusion attacks is $D = C_{mn}^{m}/C_{mn}^{n}$, where m and n represent the number of links in the anonymous forwarding network and the number of nodes per link, respectively.

As shown in [Figure 5: see original paper] and [Figure 6: see original paper], the anonymity degree of ACSNC decreases with increasing m and n, but is more significantly affected by the number of transmission links n. [Figure 6: see original paper] shows that when $n \geq 6$, the system' s anonymity degree D approaches 0, approaching absolute anonymity, while also indicating that further increasing the number of links and forwarding nodes has limited effect on improving system anonymity. Information leakage primarily results from attackers controlling forwarding nodes. The more fragments anonymous messages are divided into and the more paths they traverse, the more difficult it becomes for attackers to control one node on each forwarding path, resulting in stronger system anonymity. Simultaneously, increasing the number of forwarding nodes per path lengthens anonymous paths, further enhancing system anonymity.

2.2 Performance Analysis

When the source node sends information, ACSNC performs redundant fragmentation processing on the information through the data redundancy mechanism. This redundant fragmentation mechanism can improve the system' s data transmission success rate. By adding one or more redundant packets to the original data, the receiver can still recover the original data even after losing one or more data packets on m links. Assuming the transmission success rate on each link is identical at η, the successful transmission probability for m links without redundant packets is given by formula (1): $P_1 = \eta^m$. The successful transmission probability after adding a single redundant packet is given by formula (2): $P_2 = \sum_{i=0}^{1} C_m^i (1-\eta)^i \eta^{m-i}$.

The simulation results for generating a single redundant packet are shown in [Figure 7: see original paper]. ACSNC' s redundancy mechanism also supports generating multiple redundant packets, with simulation results shown in [Figure 8: see original paper].

As shown in [Figure 7: see original paper] and [Figure 8: see original paper], transmitting data packets through the redundancy mechanism significantly improves data transmission success rates. [Figure 7: see original paper] demonstrates that generating a single redundant packet improves system data transmission success rate by 20%. [Figure 8: see original paper] shows that generating multiple redundant packets yields even more significant improvement, increasing success rate by over 30%. Therefore, introducing the redundant fragmentation mechanism into anonymous communication greatly enhances system stability and fault tolerance.

2.3 Security Analysis

1) Collusion Attacks: The source node first fragments messages into multiple pieces and transmits them through multiple paths via the redundancy mechanism. Each forwarding node in the system only knows its predecessor and successor node addresses, making it difficult for attackers to probe the entire system' s anonymous communication relationships through collusion attacks. As analyzed in Section 2.1, as long as the number of links and forwarding nodes in the system is properly controlled, the probability of successful attack can basically be ignored.

2) Traffic Analysis Attacks: In ACSNC' s forwarding network, each intermediate forwarding node encodes and forwards data, changing the form of transmitted information as it flows into and out of nodes. This makes it extremely difficult for attackers to trace data packet transmission paths through traffic analysis to discover the addresses of source and destination nodes.

3 Conclusion

This paper proposes an improved anonymous communication mechanism based on network coding (ACSNC) by introducing a data redundancy mechanism. ACSNC inherits the advantages of existing network coding-based anonymous systems while the introduced data redundancy mechanism significantly improves the success rate of anonymous data transmission. Theoretical analysis and simulation results demonstrate that ACSNC offers good stability, anonymity, and security. This study assumes identical forwarding probabilities for all forwarding nodes; future work will consider system efficiency in more complex network environments.

References

[1] Haraty R A, Assi M, Rahal I. A systematic review of anonymous communication systems [C]// Proc of International Conference on Enterprise Information Systems. [S. l.] : SciTePress, 2017: 211-220.

[2] Zhu Tingwei, Feng Dan, Wang Fang, et al. Efficient anonymous communication in SDN-Based data center networks [J]. IEEE//ACM Trans on Networking, 2017, PP (99): 1-14.

[3] Chaum D. Untraceble electronic mail, return addresses, and digital pseudonyms [J]. Communications of the ACM, 1981, 24 (2): 84-90.

[4] Fanti G, Viswanath P. Algorithmic advances in anonymous communication over network [C]// Proc ofConference on Information Science and Systems. Piscataway, NJ: IEEE Press 2016: 12-21.

[5] Justin B. The anonymizer: protecting user privacy on the Web [J]. Computer-Mediated Communication Magazine, 1997, 4 (9): 34-38.

[6] Reed M G, Syverson P F, Goldschlag D M. Anonymous connections and onion routing [C]// Selected Areas in Communications, Piscataway, NJ: IEEE Press, 1998, 16 (4): 482-494.

[7] Dingledine R, Mathewson N, Syverson P. Tor: the second-generation onion router [J]. Journal of the Franklin Institute, 2004, 239 (2): 135-139.

[8] Reiter M K, Rubin A D. Anonymous Web transcations with crowds [J]. Communications of the ACM, 1999, 42 (2): 32-48.

[9] 段桂华, 王伟平, 王建新, 等. 一种基于多路径网络编码的匿名通信机制 [J]. 软件学报, 2010, 21 (9): 2338-2351. (Duan Guihua, Wang Weiping, Wang Jianxin, Yang Luming. Anonymous communication mechanism with Multi-Paths network coding [J]. Journal of Software, 2010, 21 (9): 2338-2351.)

[10] 周彦伟, 杨波, 吴振强, 等. 基于网络编码的匿名通信模型 [J]. 中国科学: 信息科学, 2014, 44 (12): 1560-1579. (Zhou Yanwei, Yang Bo, Wu Zhenqiang, et al. Anonymous communication model based on network coding [J]. Chinese Science: Information Science, 2014, 44 (12): 1560-1579.)

[11] Chen Fei, Xiang Tao, Yang Yuanyuan, et al. Securecloud storage meets with secure network coding [J]. IEEE Trans on Computers, 2016, 65 (6): 1936-1948.

[12] Yu Mingchao, Sadeghi P, Aboutorab N. Performance characterization and transmission schemes for instantly decodable network coding in wireless broadcast [J]. Eurasip Journal on Advances in Signal Processing, 2015, 2015 (1): 1987-1996.

[13] Ahlswede R, Cai N, Yeung R W, et al. Network information flow [J]. IEEE Trans on Information Theory, 2002, 46 (4): 1204-1216.

[14] 王少辉, 蒋季宏, 肖甫. 基于重路由匿名通信系统的设计 [J]. 计算机科学, 2016, 43 (10): 154-159. (Wang Shaohui, Jiang Jihong, Xiao Wei. New design of rerouting-based anonymous communication system [J]. Computer Science, 2016, 43 (10): 154-159.)

[15] 谭庆丰, 时金桥, 方滨兴, 等. 匿名通信系统不可观测性度量方 [J]. 计算机研究与发展, 2015, 52 (10): 2373-2381. (Tan Qingfeng, Shi Jinqiao, Fang Binxing, et al. Towards measuring unobservability in anonymous communication systems [J]. Computer Research and Development, 2015, 52 (10): 2373-2381.)

[16] 周彦伟, 杨波, 张文政. 新型的多路径匿名通信系统 [J]. 电子学报, 2017, 45 (5): 1234-1239. (Zhou Yanwei, Yang Bo, Zhang Wenzheng. A new anonymous communication system with multi-path [J]. Chinese Journal of Electronics, 2017, 45 (5): 1234-1239.)

[17] Bagai R, Malik N, Jadliwala M. Measuring anonymity of pseudonymized data after probabilistic background attacks [J]. IEEE Trans on Information Forensics & Security, 2017, PP (99): 1.

Note: Figure translations are in progress. See original paper for figures.