symmetry

MDPI

Article

MFBS: Multiple Factor Bandwidth Strategy Scheme for Anonymity Assessment

Tianbo Lu [1,2,*], Meng Luo [1,2], Ru Yan [1,2] and Zhimin Lin [1]

[1] School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China; haha@bupt.edu.cn (M.L.); yanru@bupt.edu.cn (R.Y.) zhiminlin1155@163.com (Z.L.)

[2] Key Laboratory of Trustworthy Distributed Computing and Service, Beijing University of Posts and Telecommunications, Ministry of Education, Beijing 100876, China

* Correspondence: lutb@bupt.edu.cn

Received: 30 November 2018; Accepted: 31 January 2019; Published: 11 February 2019

Abstract: This paper proposes a multiple factor bandwidth strategy (MFBS), an anonymity assessment scheme based on bandwidth strategy. We first analyzed the path selection algorithm mechanism based on bandwidth strategy and anonymity constraint conditions and then elaborated the overall architecture and the key module design of the MFBS scheme. A detailed design was carried out so that it can be applied for the evaluation of Tor's anonymous communication system. Finally, according to the running data in the node resource file in the anonymous network, we analyzed anonymity from different dimensions. By evaluating the bandwidth consumed by Tor in the actual network, the anonymity of the Tor could be calculated, and a more comprehensive anonymity assessment could be performed.

Keywords: multiple factor bandwidth strategy; path selection algorithm; anonymity constraint conditions; anonymity assessment

1. Introduction

Anonymous communication is mainly aimed at protecting user information and safeguarding personal privacy of users in network communication. Its main purpose is to hide the identity and communication of users in the network so as to hide sensitive user information. Through anonymous communication technology, the IP addresses used in the communication between users can be made unlinkable, and users can hide their real identity and online activity information in the network.

In anonymous communication technology, the mix anonymous network was originally proposed by Chaum [1]. The main idea is to use multiple mix agents to obfuscate and encrypt messages and then redirect them. This can hide the identity between the sender and the recipient so that the two persons can communicate with each other anonymously. In order to ensure the anonymity of the system [2], some new anonymous communication systems have gradually emerged based on the mix anonymity network, such as Mixmaster [3] and Mixminion [4] anonymous email systems, web mix [5] anonymous browsing tools, Tor [6] anonymous communication system of onion routing agent, etc. These anonymous communication systems can provide users with anonymous services, enabling them to communicate, send, and receive mails and publish information securely.

In the mix anonymous network, it is crucial to evaluate the anonymity of the network. In general, in the absence of security defenses, attackers often prioritize factors that have a significant impact on anonymity, such as path selection strategy, path length, user configuration, geographic location, etc. Once the information is obtained and utilized by attackers, it may cause the anonymity in the mix anonymous network to be damaged to varying degrees.

Anonymity evaluation in the mix anonymous network has prompted widespread interest and discussion among academics. Many scholars have proposed methods to evaluate the anonymity of the mix anonymous network. In terms of qualitative assessment, in 1998, Reiter and Rubin [2] proposed the concept of degree of anonymity, which can be divided into six levels of anonymity. In the same year, the anonymity of mix anonymous networks was first evaluated using the set theory [7]. Since 2002, some scholars have proposed entropy-based metrics [8,9], which can evaluate anonymity from a global perspective. In 2007, Edman et al. [10] proposed a system-wide anonymity evaluation scheme from the perspective of the attacker by constructing a complete bipartite graph framework between input and output relationships. The scheme can obtain correspondence between the input and output of each message in the mix anonymous network, and two evaluation indicators can be used. In 2010, Venkitasubramaniam et al. [11] proposed the idea of zero-sum game evaluation. Here, the authors studied the game between attackers and network designers and used the degree of anonymity to solve the problem of maximizing anonymity in the mix anonymous network from the perspective of game theory.

With the widespread use of mix anonymous communication systems, more and more scholars are proposing new anonymity assessment schemes from different perspectives. It has now become common to study anonymity in mix anonymous networks through available models and frameworks. Keeping with this trend, the present study also uses some typical anonymity assessment programs for further analysis and research. In an actual anonymous communication system, the anonymity evaluation scheme needs to be simple and effective. At the same time, it needs to reflect the actual usage of the anonymous communication system in real-life situations. Therefore, in the process of evaluating anonymity, the following aspects need to be considerez [12]:

(1) The anonymity assessment should take into account the topology of the anonymous network or any additional content defined in the anonymous communication system. The attacker can influence the anonymity of the system by collecting the topology information in the network.

(2) Anonymity assessment can be used as a method to evaluate the effectiveness of an anonymous communication system and should be independent of the number of users of the system.

(3) The anonymity assessment must be independent of the attacker's threat model because the attacker may use multiple attack techniques to undermine anonymity.

2. Related Works

This section focuses on anonymity constraints and bandwidth-based path selection algorithms that are commonly considered in mix anonymous networks.

2.1. Anonymity Constraints

1. Node type constraint

According to the location of the relay node, the node classification of Tor is mainly divided into the following types [13]:

Entrance node

The first hop relay node located in the communication link, called the guard node, is usually represented by "*G*".

Intermediate node

A second hop relay node located in the communication link, called the middle node, is usually represented by "*M*".

Exit node

A third hop relay node located in the communication link, called the exit node, is usually represented by an "*E*".

Mixed node

A relay node that can be located at the first hop entry location in the communication link or at the third hop exit location can be regarded as a compatible node and is usually represented by "$G + E$" or "D".

2. Path construction constraint

According to the attributes of the relay nodes, there are two main types in the path construction of Tor [14]:

Stable path

A stable path consists of stable nodes, which refers to the relay node whose running time exceeds half of the Tor network or 30 days. A stable path will be used for connections that have long-lived links and will communicate via a fixed port number.

Fast path

A fast path consists of fast nodes. A fast node means that the bandwidth declared by the node is within the range of the first seven-eighths of the Tor network, or the bandwidth is at least 100 kB/s.

2.2. Analysis of Path Selection Algorithm Based on Bandwidth Strategy

In the Tor anonymous network, the Tor selects the path based on two points [14]. One is the Tor node attribute provided by the Tor directory authority, and the other is the specific configuration requirement according to the user's request. The main path selection algorithm types are as follows:

Random bandwidth strategy

In the initial path selection algorithm design of the anonymous network, the path node selection is performed by a random uniform distribution method in order to ensure the anonymity of Tor. In this case, the probability that each node is selected is equal.

Weighted bandwidth strategy

In order to achieve load balancing, the node selection probability in the path is proportional to the bandwidth value issued by the node, and the weight can be assigned according to the node type, thereby selecting the relay node corresponding to the location. According to the description of the path algorithm, the weight ratio of each type of node (entrance node, intermediate node, and exit node) can be set to 1/3.

Tunable bandwidth strategy

Tunable bandwidth strategy can be seen as an improvement to the weighted bandwidth strategy, which optimizes performance in the anonymous network. In the algorithm, an exit node is first selected, and the egress node is then restricted in the selection process. After the exit node is selected, different node types are determined under different conditions according to the changeable weight pol. Lastly, the intermediate node and the entrance node corresponding to the path are selected.

2.3. Path Compromise Rate under Bandwidth Strategy

In the anonymity evaluation for the Tor anonymous communication system, a more common method is to evaluate the anonymity [15] effect in the current network by calculating the path compromise rate in the anonymous network. Therefore, based on the ability of the attacker to obtain the bandwidth budget, this paper makes a reasonable assessment of the anonymity of the MIX anonymous network under different attack situations. In the anonymity evaluation scheme, the path compromise rate is mainly limited by some constraints in the path, such as the node bandwidth value, the number of nodes, the port service type, the geographic location corresponding to the node IP address, and the user configuration [16,17].

According to the communication protocol of the Tor path selection algorithm, the node selection scheme of the first hop (entrance node) and the last hop (exit node) is selected according to the bandwidth budget in the protocol. By analyzing the total bandwidth budget [18] corresponding to the two types of nodes and the bandwidth budget obtained by the attacker, the calculation formula of the path compromise rate, which indicates the ratio of malicious nodes on anonymous networks, can be obtained.

Random bandwidth strategy mode

The algorithm considers the overall bandwidth distribution of the system, and the number of nodes of each type can be counted. Defined as follow:

(1) The optional node collection in the path can be given as follows:

$$S = (G + D) \cdot (E + D) - D \tag{1}$$

G indicates the number of entrance node types, E indicates the number of exit node types, and D indicates the number of nodes of the hybrid node type.

(2) The collection of malicious nodes in the path can be given as follows:

$$S_c = (G_c + D_c) \cdot (E_c + D_c) - D_c \tag{2}$$

where G_c represents the number of malicious node types, E_c indicates the number of malicious node types of the exit node, and D_c indicates the number of the hybrid nodes.

(3) Node compromise rate

$$P_c = \frac{S_c}{S} \tag{3}$$

where S_c represents a set of malicious nodes controlled by the attacker. S represents the total set of entrance and exit nodes in the path.

Weighted bandwidth strategy mode

According to the path selection algorithm of the weighted bandwidth strategy, the bandwidth budget of each node type needs to be divided so that the path compromise rate can be calculated. Defined as follows: where B_i indicates the bandwidth value of the node X_i, S_g represents a collection of entrance node types, S_e represents a collection of exit nodes, S_d represents a collection of mixed nodes, and S_m represents a collection of intermediate nodes.

$$P_c = \frac{(\sum_{i \in G_c} B_i + \sum_{i \in D_c} B_i) \cdot (\sum_{i \in E_c} B_i + \sum_{i \in D_c} B_i)}{S} - \frac{\sum_{i \in D_c} B_i^2}{S} \tag{4}$$

$$S = (\sum_{i \in S_g} B_i + \sum_{i \in S_d} B_i) \cdot (\sum_{i \in S_e} B_i + \sum_{i \in S_d} B_i) - \sum_{i \in S_d} B_i^2 \tag{5}$$

Adjustable bandwidth strategy mode

According to the design goal of the path selection algorithm under the adjustable parameter bandwidth strategy, it is possible to continue constructing the resource path after adding the node parameter of the adjustable weight factor when the number of nodes of a certain type is scarce. Thus, it may do some damage to the anonymity of the system.

The calculation formula of the node compromise rate is as follows:

$$P_c = \frac{G_c \cdot E_c - D_c}{G \cdot E - D} \tag{6}$$

3. Design of MFBS

This section mainly describes the design and overall architecture of the multiple factor bandwidth strategy (MFBS). The main purpose of MFBS is to evaluate the anonymity of the anonymous communication system based on the mix. It can be applied to low-latency anonymous communication systems, such as Tor.

3.1. Design Goals

The anonymity evaluation design for MFBS should have the following objectives:

1. Feasibility

By analyzing the functional requirements of the MFBS in the evaluation process, the detailed design of each module can be investigated. The evaluation scheme is technically feasible and is feasible for the system's users.

2. Accuracy

The data acquisition in the design of the MFBS is derived from the real network file information in the Tor network communication. We can calculate the bandwidth budget situation that can be controlled by the first node and the last node in the attack mode, and according to the consumption of the bandwidth in the actual anonymous network, we can then calculate the anonymity of the Tor network, which can truly and effectively reflect the anonymity.

3. Extensibility

In the actual anonymous network, the Tor proxy node may be in an ever-changing state. The specific implementation scheme can acquire the network data in real time and can expand the data information in the system at any time.

3.2. The Architecture Design

The design of the MFBS is based on the operational data of the Tor, and the network information between the entrance node and the exit node of Tor needs to be processed. According to the functional requirements of the system, the functional modules of the MFBS can be divided into the following: the node resource acquisition module, the node resource processing module, and the anonymity evaluation module. The overall architecture of the design is shown in Figure 1.

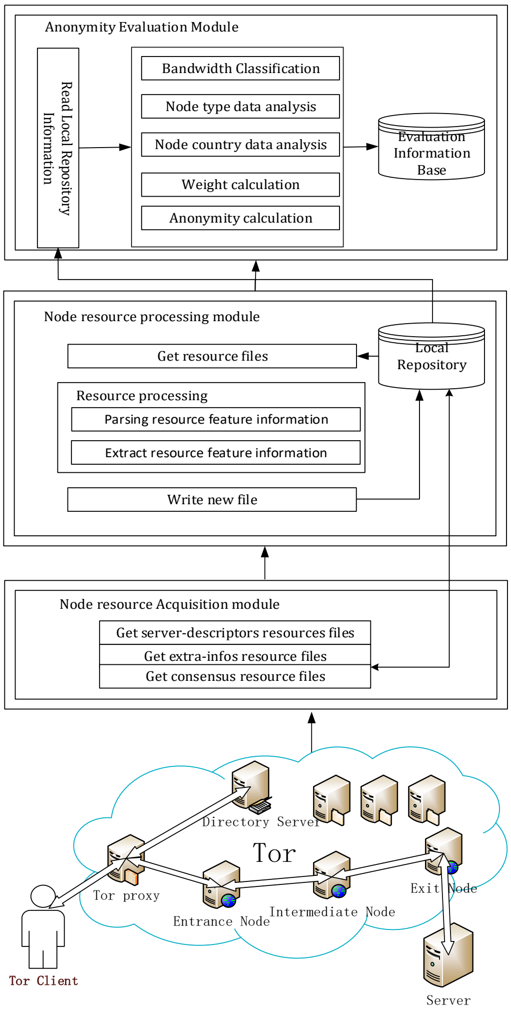
Anonymity Evaluation Module
Read Local Repository Information
Bandwidth Classification
Node type data analysis
Node country data analysis
Weight calculation
Anonymity calculation
Evaluation Information Base
Node resource processing module
Get resource files
Local Repository
Resource processing
Parsing resource feature information
Extract resource feature information
Write new file
Node resource Acquisition module
Get server-descriptors resources files
Get extra-infos resource files
Get consensus resource files
Directory Server
Tor
Tor proxy
Entrance Node
Intermediate Node
Exit Node
Tor Client
Server

Figure 1. Overall architecture of multiple factor bandwidth strategy (MFBS).

1. Node resource acquisition module

In order to evaluate the global anonymity of the Tor, it is necessary to first obtain the data flow information of the Tor relay node and collect the node information in the anonymous network. In order to ensure the reliability of the node information, we can obtain the actual data of the nodes in the anonymous network of Tor from the CollecTor [19,20] official website.

2. Node resource processing module

According to the relevant information of the Tor network node obtained by the node resource acquisition module, this module uses the Python language to design a parsing script to parse and process user data. By processing the node-related information, key information can be extracted, such as node IP address, node bandwidth value, and node port. After analyzing the log files of node information, it will be stored in the database to support follow-up anonymity assessments.

3. Anonymity evaluation module

For the anonymity evaluation module, we propose an improved anonymity calculation scheme based on multifactor bandwidth strategy combined with the path selection algorithm of the tunable parameter bandwidth strategy in the Tor path protocol. After processing the node data, the data can be statistically analyzed and quantified. According to the current node data to be evaluated, the bandwidth budgets of different node types and country distributions can be counted. According to the objective weighting algorithm, the corresponding scheme of bandwidth weight calculation can be designed. Finally, the path compromise rate in the attack mode can be obtained so that the global anonymity can be evaluated.

4. Modular design of MFBS

4.1. Design of Node Resource Acquisition Module

The main execution flow of the node resource acquisition modules is shown in Figure 2. The module is a preliminary collection of Tor network data. In the design of the module, the actual operating data of the node in the Tor network was obtained from the official website of CollecTor, a service provided by the Tor network that regularly collects data services from Tor relay nodes, bridge agents, etc. The entire Tor network information can be provided after aggregation, and it will guarantee the objectivity of this module.

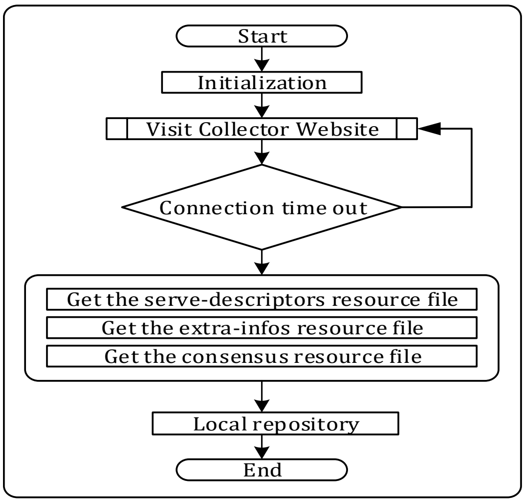
Start
Initialization
Visit Collector Website
Connection time out
Get the serve-descriptors resource file
Get the extra-infos resource file
Get the consensus resource file
Local repository
End

Figure 2. Node resource acquisition module.

4.2. Design of Node Resource Processing Module

The main execution flow of the node resource processing module is shown in Figure 3. The purpose of this module is to extract the information of the network nodes and then extract the characteristic information in the nodes. After processing, the related information needs to be integrated into the new file for data analysis to evaluate anonymity. In this module, it is first necessary to read and parse the above resource files of node and then extract the node feature information required for anonymous evaluation of Tor. This includes the IP address, bandwidth value, geographic location, port service, running time, release time, and so on.

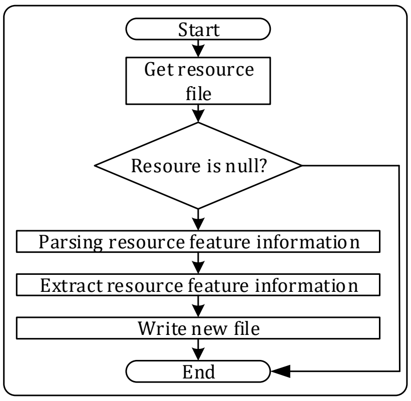
Start
Get resource file
Resoure is null?
Parsing resource feature information
Extract resource feature information
Write new file
End

Figure 3. Node resource processing module.

4.3. Design of Anonymity Evaluation Module

The main execution flow of the anonymity evaluation module is shown in Figure 4. In actual anonymous networks, the bandwidth distribution of the node type and the node IP address will have an impact on the anonymity. Some scholars have analyzed the statistical results of node data in anonymous networks through qualitative analysis, indicating that the distribution of node countries will affect the anonymity of anonymous networks; however, the subjective judgment factors here are large. Therefore, the MFBS scheme based on bandwidth strategy is proposed in this section. Considering the influence of geographical distribution in an anonymous network, we can quantitatively analyze the node specifically and objectively influence the bandwidth distribution of anonymous networks by adding the constraints of node distribution. According to the bandwidth distribution of each node type corresponding to the country type, we can more comprehensively evaluate the anonymity of the anonymous network in which the attacker gets different bandwidth budgets. According to the specific distribution of current network node data, the purpose of this module is to statistically analyze the bandwidth budget of different node types and design the corresponding weight calculation scheme using the objective weighting algorithm. We can calculate the path compromise rate in attack mode so that the global anonymity can be evaluated.

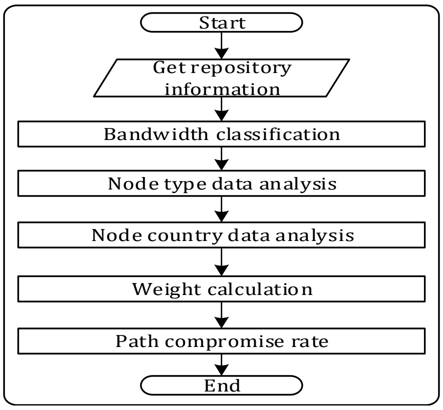

Figure 4. Anonymity evaluation module.

The anonymity assessment algorithm proposed in this section is based on the path selection algorithm of the tunable bandwidth strategy. Based on the anonymity calculation method described in Section 2, we can obtain the calculation method of the path compromise rate and evaluate the anonymous effect of the anonymous system. To design a specific evaluation algorithm, bandwidth grading is first performed by the bandwidth distribution interval, and the relay nodes are then classified by the node type. Based on this classification, the bandwidth distribution effect corresponding to the Top 10 country type is increased. C_i represents the i-th country. By obtaining the weights of the node country types, we can calculate the weighting effect of each node type in the path construction. Therefore, the final bandwidth budget in the anonymous network can be obtained under the factors of geographical distribution, and we can evaluate the anonymous network by calculating the path compromise rate. The plan for the specific anonymity evaluation algorithm is shown in Figure 5.

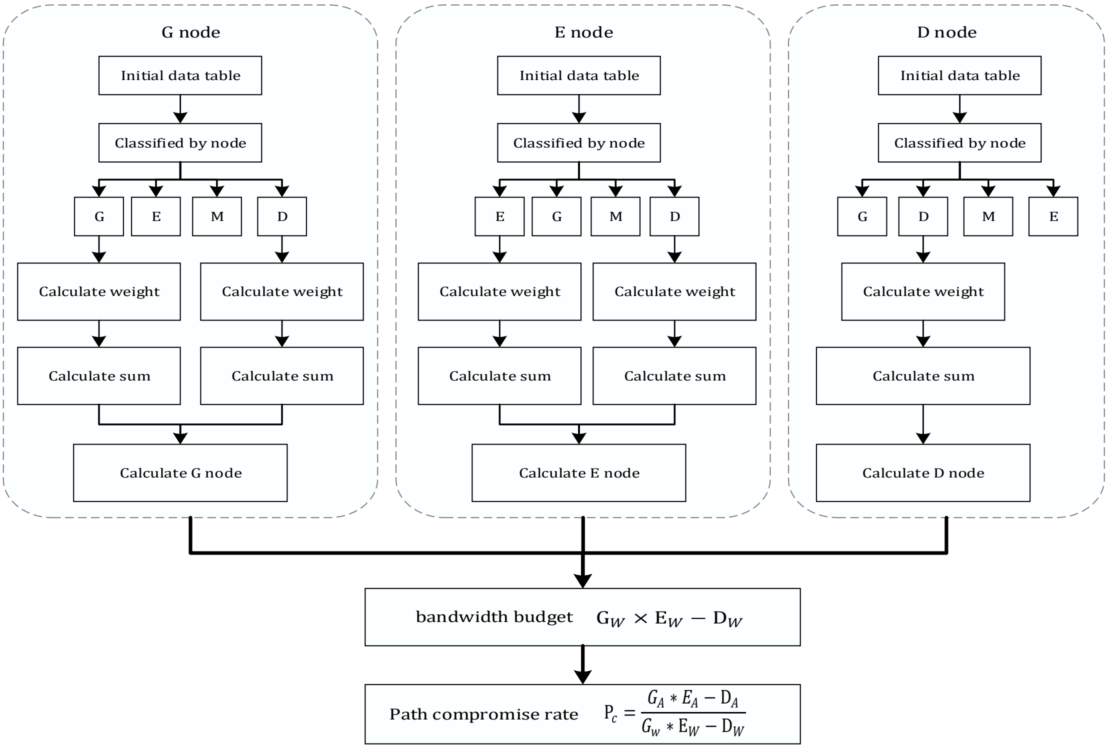

Figure 5. Anonymity evaluation algorithm.

According to the formula in [21,22], the following path compromise rate calculation formula can be obtained:

$$p_c = \frac{G_A \cdot E_A - D_A}{G_W \cdot E_W - D_W} \tag{7}$$

$$G_A = W_{gg} \cdot \sum\nolimits_{i \in S_G} (W_{C_i} \cdot \sum\nolimits_{j \in A_{C_i}} b_{ij}) + W_{gd} \cdot \sum\nolimits_{i \in S_D} (W_{C_i} \cdot \sum\nolimits_{j \in A_{C_i}} b_{ij}) \tag{8}$$

$$E_A = W_{ee} \cdot \sum\nolimits_{i \in S_E} (W_{C_i} \cdot \sum\nolimits_{j \in A_{C_i}} b_{ij}) + W_{ed} \cdot \sum\nolimits_{i \in S_D} (W_{C_i} \cdot \sum\nolimits_{j \in A_{C_i}} b_{ij}) \tag{9}$$

$$D_A = W_{gd} \cdot W_{ed} \cdot \sum\nolimits_{i \in S_D} (W_{C_i} \cdot \sum\nolimits_{j \in A_{C_i}} {b_{ij}}^2) \tag{10}$$

$$G_W = W_{gg} \cdot \sum\nolimits_{i \in S_G} (W_{C_i} \cdot \sum\nolimits_{j \in C_i} b_{ij}) + W_{gd} \cdot \sum\nolimits_{i \in S_D} (W_{C_i} \cdot \sum\nolimits_{j \in C_i} b_{ij}) \tag{11}$$

$$E_W = W_{ee} \cdot \sum\nolimits_{i \in S_E} (W_{C_i} \cdot \sum\nolimits_{j \in C_i} b_{ij}) + W_{ed} \cdot \sum\nolimits_{i \in S_D} (W_{C_i} \cdot \sum\nolimits_{j \in A_{C_i}} {b_{ij}}^2) \tag{12}$$

$$D_W = W_{gd} \cdot W_{ed} \cdot \sum\nolimits_{i \in S_D} (W_{C_i} \cdot \sum\nolimits_{j \in A_{C_i}} {b_{ij}}^2) \tag{13}$$

5. Analysis of MFBS

For the analysis of anonymity, we used the evaluation method from [23]. Assuming that the attacker set has a fixed bandwidth budget, they can control the percentage of the total bandwidth value in the entire path. Then, if the attacker can control the ideal bandwidth budget ratio, the attacker controls the entrance node and the exit node in a path. As a result, the path will be compromised, and this will reduce the overall anonymity of the system. Therefore, in this experiment, it was assumed that the attacker is under optimal attack capabilities. Based on the results of the anonymity evaluation module, it was possible to evaluate the global anonymity of the anonymous network.

In this paper, all odd month data sets from January to November were selected. The anonymity assessment result of the attacker under different bandwidth budgets is shown in Figure 6. The abscissa indicates the bandwidth budget ratio that the attacker can control, and the ordinate indicates the percentage corresponding to the path compromise rate.

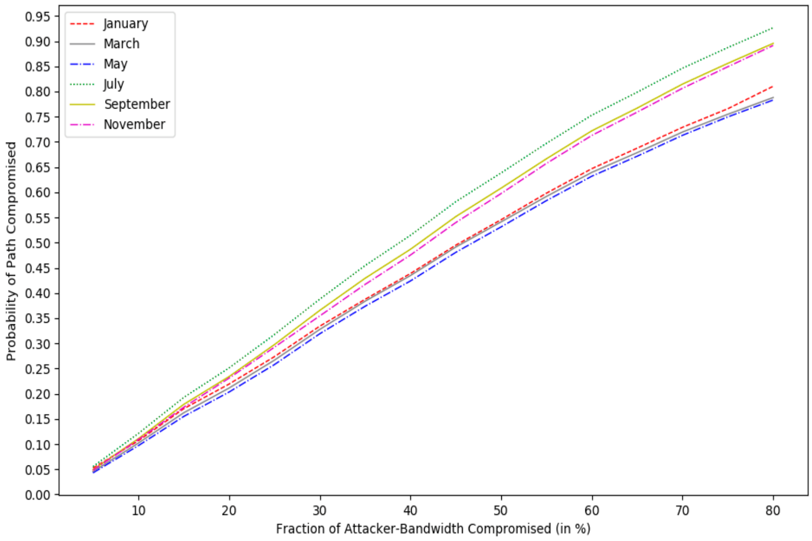

Figure 6. Anonymity assessment results.

By analyzing the results of anonymity of the anonymous system, shown in Figure 6, we drew the following conclusions:

From the overall trend of change, the path compromise rate increased as the attacker's controllable bandwidth budget increased.

When the attacker's bandwidth budget was between 10% and 30%, the path compromise rate increased rapidly. When the bandwidth budget exceeded 30%, the growth rate of the path compromise rate gradually began to slow down.

From the trend of changes in the six odd months, the trend of path compromise rates in January, March, and May were very close, and the trend of path compromise rates in July, September, and November were relatively close.

When the attacker-controlled path bandwidth ratio was the same, the path compromise rate in July was larger, and the path compromise rate in March was relatively small. When the attacker controlled the path bandwidth ratio to 80%, the path compromise rate in July exceeded 0.9, and the path compromise rate in March was below 0.8.

6. Conclusions

We developed a new multifactor anonymity calculation scheme based on bandwidth strategy for anonymous assessment. For this, we used the real node information in the Tor network to guarantee the objectivity of the results. We studied the detailed calculation process relating to the calculation of the path compromise rate. Finally, we assumed that the attacker could control the bandwidth budget ratio and obtained the anonymous result of the anonymous module. The implementation of MFBS will be beneficial to users and designers of anonymous systems. It can help users in the network to intuitively understand the anonymity of their network so that users can better protect personal, sensitive information in the network. At the same time, it can help developers calculate the anonymity of the network so that they can better optimize and improve it.

Author Contributions: Conceptualization, T.L.; Methodology, T.L., M.L., M.L., C.L; Software, Z.L., M.L.; Validation, M.L., Z.L., R.Y.; Formal Analysis, M.L., M.L., C.L.; Investigation, Z.L.; Resources, T.L., M.L., Z.L; Data Curation, M.L., Z.L.; Writing-Original Draft Preparation, M.L.; Writing-Review & Editing, M.L., R.Y.; Visualization, M.L., R.Y.; Supervision, M.L., R.Y.; Project Administration, T.L.; Funding Acquisition, T.L.

Funding: This research was funded by the National Natural Science Foundation of China (Grant No. 61170273) and the China Scholarship Council (Grant No. [2013]3050).

Acknowledgments: This work is supported by the following grants: The National Natural Science Foundation of China under Grant No. 61170273 and the China Scholarship Council under Grant No. [2013]3050. In addition, we express our sincere gratitude to Lingling Gong, Meng Luo, Zhimin Lin, Peiyuan Li and the anonymous reviewers for their valuable comments and suggestions.

Conflicts of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. Chaum, D.L. Untraceable Electronic mail, return addresses, and digital pseudonyms. *Commun. ACM* **1981**, *24*, 84–90. [CrossRef]
2. Reiter, M.K.; Rubin, A.D. Crowds: Anonymity for Web transactions. *ACM Trans. Inf. Syst. Secur.* **1998**, *1*, 66–92. [CrossRef]
3. Moeller, U.; Cottrell, L.; Palfrader, P.; Sassaman, L. *Mixmaster Protocol—Version 2*; Network Working Group Internet-Draft; The Internet Society: Reston, VA, USA, 2004.
4. Danezis, G.; Dingledine, R.; Mathewson, N. Mixminion: Design of a Type III Anonymous Remailer Protocol. In Proceedings of the IEEE Symposium on Security and Privacy, Berkeley, CA, USA, 11–14 May 2003; p. 2.
5. Berthold, O.; Federrath, H.; Kopsell, S. Web MIXes: A system for anonymous and unobservable Internet access. In Proceedings of the International Workshop on Designing Privacy Enhancing Techologies: Design Issues in Anonymity and Unobservability, Berkeley, CA, USA, 25–26 July 2000; pp. 115–129.
6. Dingledine, R.; Mathewson, N.; Syverson, P. Tor: The Second-Generation Onion Router. In Proceedings of the 13th USENIX Security Symposium, San Diego, CA, USA, 9–13 August 2004; p. 21.
7. Chaum, D. The dining cryptographers problem: Unconditional sender and recipient untraceability. *J. Cryptol. Arch.* **1998**, *1*, 65–75. [CrossRef]
8. Diaz, C.; Seys, S.; Claessens, J.; Preneel, B. Towards measuring anonymity. In Proceedings of the Second International Conference on Privacy Enhancing Technologies, San Francisco, CA, USA, 14–15 April 2002; pp. 54–68.
9. Serjantov, A.; Danezis, G. Towards an information theoretic metric for anonymity. In Proceedings of the Second International Conference on Privacy Enhancing Technologies, San Francisco, CA, USA, 14–15 April 2002; pp. 41–53.
10. Edman, M.; Sivrikaya, F.; Yener, B. A Combinatorial Approach to Measuring Anonymity. In Proceedings of the 2007 IEEE Intelligence and Security Informatics, New Brunswick, NJ, USA, 23–24 May 2007; pp. 356–363.
11. Venkitasubramaniam, P.; Tong, L. A game-theoretic approach to anonymous networking. *IEEE/ACM Trans. Netw. Arch.* **2012**, *20*, 892–905. [CrossRef]
12. Zhu, Y.; Bettati, R. Anonymity vs. information leakage in anonymity systems. In Proceedings of the 25th IEEE International Conference on Distributed Computing Systems, Columbus, OH, USA, 6–10 June 2005; pp. 514–524.
13. Tor Path Specification Index: Torspec. Available online: https://gitweb.torproject.org/torspec.git/tree/tor-spec.txt (accessed on 7 September 2018).
14. Tor Path Specification Index: Torspec. Available online: https://gitweb.torproject.org/torspec.git/tree/path-spec.txt (accessed on 7 September 2018).
15. Sebastian, C.; Schiffner, S. Structuring anonymity metrics. In Proceedings of the Second ACM Workshop on Digital Identity Management, Alexandria, VA, USA, 3 November 2006; pp. 55–62.

16. Loesing, K.; Murdoch, S.J.; Dingledine, R. A Case Study on Measuring Statistical Data in the Tor Anonymity Network. In Proceedings of the International Conference on Financial Cryptograpy and Data Security, Tenerife, Spain, 25–28 January 2010; pp. 203–215.
17. Perry, M. Torflow: Tor network analysis. In Proceedings of the HotPETs—The 2nd Hot Topics in Privacy Enhancing Technologies (Co-Located with PETS), Seattle, WA, USA, 5–7 August 2009; pp. 1–14.
18. Hamel, A.M.; Gregoire, J.; Goldberg, L. The Mis-Entropists: New Approaches to Measures in Tor. Available online: http://cacr.uwaterloo.ca/techreports/2011/cacr2011-18.pdf (accessed on 29 November 2018).
19. Index of Torproject. Available online: https://collector.torproject.org/ (accessed on 7 September 2018).
20. Alsabah, M.; Goldberg, I. Performance and Security Improvements for Tor: A Survey. *ACM Comput. Surv.* **2016**, *49*, 32.
21. Kondakci, S. Network Security Risk Assessment Using Bayesian Belief Networks. In Proceedings of the 2010 IEEE Second International Conference on Social Computing, Minneapolis, MN, USA, 20–22 August 2010; pp. 952–960.
22. Wacek, C.; Tan, H.; Bauer, K.S.; Sherr, M. An Empirical Evaluation of Relay Selection in Tor. In Proceedings of the NDSS Symposium 2013, San Diego, CA, USA, 24–27 February 2013; Volume 13, pp. 24–27.
23. Azzalini, A.; Bowman, A.W.; Härdle, W. A candidate's formula: A curious result in Bayesian prediction. *Biometrika* **1989**, *76*, 183.