ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645
DOI: 10.17485/ijst/2015/v8i7/62882

Indian Journal of Science and Technology, *Vol 8(7), 689–694, April 2015*

A Novel Traffic Dividing and Scheduling Mechanism for Enhancing Security and Performance in the Tor Network

K. Sangeetha[1*] and K. Ravikumar[2]

[1]Bharathiyar University, Coimbatore, India; sangeethaphd789@gmail.com
[2]Faculty of Computer Science Department, Tamil University, India; ravikasi2001@yahoo.com

Abstract

Objective: The main intent of this research is to improve the quality-of-service in the Tor anonymous communication network by splitting the traffic and scheduling it according to the traffic load. **Methods:** In this manuscript, a Novel Traffic, Dividing and Scheduling (NTDS) mechanism is introduced for improving performance in the Tor anonymous communication network. The performance of the tor network is described by using various network metrics such as bandwidth, latency, jitter, loss etc. In this mechanism, the traffic load is disseminated across circuits depends on each circuit bandwidth capacity, congestion level, measured latency, throughput. **Results:** The Novel Traffic, Dividing and Scheduling (NTDS) mechanism show higher performance when compared to the existing router selection method. Throughput, congestion level, latency and available bandwidth are measured for the circuit and based on this the traffic is split and scheduled. If the number of nodes is 50, the packet delivery ratio is 70% in NTDS, the throughput in NTDS is 72% and end-to-end delay is 0.69. According to the comparison results, the proposed approach works better than the other existing systems with high quality-of-service. **Conclusion:** The findings demonstrate that the Novel Traffic, Dividing and Scheduling (NTDS) mechanism is presented and suggested that this method has high quality-of-service.

Keywords: Anonymous Communication Network, Bandwidth Estimation, Quality-of-service, Path Selection, Tor Network

1. Introduction

The Tor network is a network consists of effective tunnels that assist users and groups to enhance their privacy and security in the internet. Tor means "The Onion Router" that is a free software for providing online anonymity. The Onion routing[1] is defined as the levels of encryption. Including the original data, the destination also included and this information is encrypted and re-encrypted multiple times and sends via virtual circuit regarding successive, randomly favored Tor relays. The relay node in the network decrypt the level of the encrypted data and given to the successive relay nodes in the circuit. At the end of this process, the data is decrypted in the end level of the encryption process. The original data is send without denoting the destination. By using this method, the possibility of the original data being decreased that is understood in transit and, more predominantly, obscures the routing of it.

The tor network is used for different purposes. Similarly, the increase in the usage of the Tor network has led to create a several attacks on the system. To enhance the performance and anonymity a user-tunable mechanism is suggested for choosing routers based on their bandwidth capabilities. This method effectively integrates the traffic of users with different inclinations and making partitioning attacks difficult[2]. Because of the changing conditions in the tor network, the self-reported bandwidth is repeatedly overrated of the actual node capacity, leading to changeable performance distributed to Tor users.

To solve this problem, an opportunistic bandwidth measurement mechanism is presented instead of using

*Author for correspondence

Tor mechanism. In this method, each and every router has a possibility to interrelate with most other routers and thus scrutinize their performance empirically. This mechanism is a proper alternating method for the self-reported bandwidth that precisely evaluates the performance of the routers and is expansively less susceptible to low-resource attacks. This Tunable Tor shows that the users can accomplish great developments in performance without forfeiting much anonymity, or expansively increase anonymity protection without any loss in performance. But in this method is tor network has a variety of performance problems leads to less quality-of-service.

The Tor network is a widely used low-latency anonymity network, which offers strong privacy guarantees by tunneling a user's Internet traffic through virtual circuits consisting of multiple intermediate overlay routers using a layered encryption scheme based on onion routing. So, a new mechanism is introduced which is called a Novel Traffic, Dividing and Scheduling (NTDS) mechanism for improving performance in the Tor anonymous communication network. In this mechanism, the traffic is split based on the measured metrics like throughput, latency, available capacity, congestion level. The traffic, dividing concept develops load balancing in the tor networks. If routers become over-utilized and high congestion, dividing traffic across semi-disjoint paths can ease the burden on the congested circuit; under this scheme, circuits require only share a common exit router. By dividing data over multiple circuits, the user's throughput can attain up to the aggregate throughput of all circuits rather than a single one. This is predominantly useful when a circuit uses a low-bandwidth router. By using this algorithm, the network metrics like throughput, latency, available capacity, congestion level are measured and based on this measured metric, the traffic is assigned to each circuit. Furthermore, this mechanism extensively enhances the experience of users who watch streaming videos in online.

2. Previous Research

Anonymous communication networks have been developed in recent years. The performance of tor anonymous communication networks depends on various metrics like bandwidth, latency, jitter, loss, etc[3]. The existing relay selection methods are focused on producing paths with various bandwidths. A link-based path generation method is used that relay selection is accomplished by taking high performance links. In this method, the sender selects high performing links[4] to construct the anonymous paths. According to the predicted end-to-end performance, the initiator ranks randomly created paths and computed by aggregating the costs of their component links. By using the set of paths, the initiator selects a path by using the probability distribution weighted by the end to end cost estimates. This method is more flexible routing, and it is able to provide anonymous paths with less delay, less difference in packet arrival time, and loss, in addition to high bandwidth. But the drawback is path selection is less efficient.

To improve security and scalability problems in tor anonymous communication network, a new method which is called Salsa is suggested for managing highly distributed anonymous communications systems[5]. The main idea of salsa is to manage a large anonymous communications network to ensure a random and unbiased node selection for users. The basic design merges a structured overlay with redundant lookups to make sure that arbitrarily selected nodes are not predisposed towards the selection of attackers. The initiator can choose the nodes in the circuit at arbitrary without important bias and without knowledge of the whole set of proxies. The distributed hash table is used based on hashes of the nodes' IP addresses to classify the nodes into groups. By using a virtual tree structure, incomplete knowledge of the other nodes is adequate to route node lookups throughout the system. The redundancy and bounds checking are utilized when performing lookups to avoid malicious nodes from recurring false information without detection. But the main challenge is the Salsa design is the selection of nodes at random from the set of available peers with the only imperfect knowledge of the nodes.

The second generation Onion Router is called as the Tor that supports the forwarding the TCP streams in the internet. The traffic investigation techniques are suggested that make possible the adversaries with only a fractional view of the network to suppose which nodes are utilized to forward the unidentified streams and accordingly diminish the anonymity provided by Tor. There is enormous amount of data in the traffic investigation methods. The conventional wisdom is utilized so that it can be gathered by means of a global passive examiner, which lies exterior the Tor hazard model[6]. The major work is the consciousness but the investigation abilities are not obligatory to complete these attacks. The capability is to forward the packets over the unidentified communication network, that anyone has, can be

utilized to compute the traffic load on accurate Tor nodes precisely adequate to consider the traffic-analysis. But we cannot discover the proficient and best transient signal detectors for traffic-analysis.

An application self-deciding infrastructure called as Onion Routing and it is used for traffic-analysis challenging and unidentified internet connections. The main motivation of the Onion routing is to make a private communications in an efficient manner with highly cost-efficient[7]. Assume the communication in the network is secret so the malicious users are not able to know about the information transmitted between the sender and receiver. The other objective is to accomplish the ambiguity to the sender and receiver. The receiver may able to acquire the messages but not capable to distinguish the sender. Traffic analysis is one of the methods that are used to inform the communication parties to exchange the information in an anonymous manner. As a result, the main goals are considered as anonymity and confidential, communication[8]. An onion routing network contains the onion proxies, Core Onion Routers (CORs), links over which CORs pass fixed length cells, and responder proxies. This can be rescheduled into the application layer data stream. A main challenge is to inspect the traffic on a real onion routing network might examine to take advantage of the topological features, exit procedures and information regarding to the outside communications.

Tor network is one of the most favorable overlay networks for protecting the TCP traffic. Because the part of the superficial strong vagueness properties and its comparatively low latency service, tor network becomes popular in recent years[9,10]. The optimized tor router selection method is used to accomplish less delay in the Tor's capacity. In this method, more routers are supported with huge bandwidth capacity. But an adversary could deploy some nodes which have viewed as the high-bandwidth associations and high-up times. At last, the attacker is said to lie about the resources. There is high possibility to an adversary would be competent to proficiently support the two endpoints the entry node and the exit node of a new Tor client's connection. A low-resource end-to-end traffic analysis is a crucial attack which degrades the Tor performance. The attack stems from Tor's tendency to support routers that assert to be high-resource with high uptime in its routing procedure in an attempt to optimally equilibrium the traffic load. But in this technique, there is no tradeoff between performance and anonymity.

3. Opportunistic Bandwidth Measurement Mechanism

The tor network design is based on the onion routing in which the traffic is forwarded through numerous routers and that traffic is encrypted, with every router eliminating one layer of the encryption. The every router in the tor network reports the IP address, public key, and rules about what traffic it will admit and the value of the bandwidth which is decided by monitoring the peak bandwidth accomplished by the router in a period of time. So, this is an opportunity for attacks in which the adversary nodes can report high bandwidth values instead of the actual bandwidth values so that a larger fraction of tunnels are routed through them[2]. But sometimes the node reporting values are a poor predictor of the available bandwidth because of the varying network circumstances and other aspects. So, to conquer these problems a router selection algorithm is suggested. In this method, firstly computes how the performance of the router is evaluated; and from the list of evaluating routers how the route is preferred.

In this work, the main performance metric is bandwidth accessible to a Tor tunnel, rather than other performance distinctiveness such as latency or jitter. The main motivation is to consider the bandwidth in three folds. In the Tor design, bandwidth is an important factor. Secondly, bandwidth is normally a property of the node rather than the link, whereas the blockage is possible to be close to the node rather than in the intermediate network. So, this motivation makes computations and optimizations much more probable than for link properties, whereas for N nodes there are $O(N^2)$ links. Because of the complete graph topology of the Tor network, every router will have an opportunity to interrelate with most other routers and thus scrutinize their performance empirically. The Tunable Tor shows that users can accomplish great developments in performance without sacrificing much anonymity, or significantly enlarge anonymity protection without degradation in the performance. It also permits for moderate developments in both. This enhanced flexibility should make Tor palatable to a wider range of users, and thus enlarge anonymity for everyone because of a larger community.

Algorithm Description: Opportunistic bandwidth measurement Algorithm

1. Initialize N number of nodes in the network
2. Source S requires to send packets to destination D
3. //Optimal route selection based on bandwidth metric

4. Every router keeps track with current peers in the network
5. // Estimation of available bandwidth in a router
6. $BW = \frac{amount\ of\ bits\ travel\ across\ the\ network}{Seconds}$
7. // Aggregation of values
8. Each router has a large set of bandwidth statistics values
9. Each node aggregates the values A_n to single observation
10. // Calculation of Min-Max Weighted Moving Average
11. $B_{new} = (1-\alpha)\text{Max}(B_{old}, B_{obs}) + \alpha \min(B_{old}, B_{obs})$ //B_{new}= new bandwidth value, B_{old} = old bandwidth value, B_{obs} = observed bandwidth value
12. Nodes share their observed values
13. Computation of the Consensus vector from the observation matrix $N \times N$ to prevent attacks
14. // Router selection
15. Define family of functions
16. $f_s(x) = \frac{1-2^{sx}}{1-2^s}; f_s : [0,1] \to [0,1]$
17. List the routers and rankings based on bandwidth metric
18. List is indexed in 0 to n-1
19. **If** (router is indexed within $n \times f_s(x)$)
20. The router is selected
21. **Else**
22. The router is not selected
23. Based on this route is to be selected
24. Using this path S sends packets to D
25. **End if**

4. Novel Traffic Dividing and Scheduling Mechanism

In this research, a Novel Traffic, Dividing and Scheduling (NTDS) mechanism is introduced for improving the performance in the Tor anonymous communication network. In this method, the number of circuits is built by using OP (Onion Proxy). The number of circuits is interconnected at a common exit Onion router (OR). The Onion Proxy and common exit OR are the end points in the multipath. The responsibility of the Onion proxy is to retrieve and transmits the data to the user's application. The responsibility of the exit Onion router is to transmit and receives data from an external server. Each and every endpoint receives the data and split it into cells and the sequence numbers are added to the cell headers. After that, according to the traffic, dividing method the end point divides the cells across the circuits of the multipath. If the end point retrieves the cells, it gathers and reorders them based on the sequence numbers before delivering their content to their destinations.

This method includes three steps:

(1) Construction of number of circuits
(2) Split the traffic based on Novel Traffic, Dividing and Scheduling (NTDS) mechanism
(3) Gathering and reordering

A novel traffic, dividing and scheduling Algorithm is described as follows. In the traffic splitting scheme measured some of the parameters like,

(1) Throughput
(2) Congestion level
(3) Latency
(4) Available bandwidth capacity

In this algorithm, each and every endpoint divides the cells across the circuits based on the above metrics. Firstly, the throughput is measured for each and every circuit. Throughput is defined as the rate of successful message deliver over a communication channel. It is evaluated in bits per second. Secondly, the congestion level is measured for each circuit. Congestion level is defined as the packet loss per second. If the maximum load of the node exceeds the actual capacity the congestion occurs. Latency is defined as the average time taken by a data packet to reach in the destination. It also contains the delay caused by the route discovery process and the queue in data packet transmission. Finally, the available bandwidth capacity is the amount of bits travel across the network at a particular time interval.

Detailed Description

Algorithm Description: Novel traffic, dividing and scheduling Algorithm

1. Initialize N number of nodes randomly in the network
2. Initialize Onion Proxy (OP), Onion Router (OR)
3. Onion Proxy constructs a number of circuits
4. OP receives and sends data to the client application
5. Exit OR sends and receives from a server
6. Each *EP* receives data and split into cells // EP-end point
7. *EP* divides the cells across the circuits C_i Of the multipath based on the metrics

8. // Traffic splitting Scheme
9. Splitting endpoint measures the following metrics
10. For each circuit C_i Computes these metrics at every time t
11. // Throughput calculation
12. $Th_c = \frac{File\ size}{Transmission\ time}$
13. // Congestion level Calculation
14. $CL_c = \frac{packet\ loss}{time}$
15. // Latency Calculation
16. $L_c = \frac{\sum(arrive\ time - send\ time)}{\sum Number\ of\ connections}$
17. // Available bandwidth capacity
18. $BW_c = \frac{amount\ of\ bits\ travel\ across\ the\ network}{seconds}$
19. Splitting end point periodically updates these parameters
20. $HProb_c$ = Max(Th_c&BW_c)&min(CL_c&L_c) // $HProb_c$ = High Probability circuits
21. Select the circuit with $HProb_c$ as taken for data transmission
22. Data is transmitted by selecting high probability circuits
23. Other endpoint receives the cells
24. Reordering the cells according to the sequence number
25. Deliver the data to the destination.

5. Performance Validation

The performance of the existing and the proposed system is compared in this section. A Router Selection Method (RSM) is presented in the existing system that facilitates the users to control the tradeoff between performance and anonymity. A Novel Traffic, Dividing and Scheduling (NTDS) mechanism is proposed for enhancing the performance in the Tor anonymous communication network. The performance is computed in terms of the packet delivery ratio, throughput and end-to-end delay.

5.1 Packet Delivery Ratio

Packet Delivery Ratio (PDR) is defined as the ratio of number of delivered data packets to the destination.

$$PDR = \frac{\sum Number\ of\ packet\ receive}{\sum Number\ of\ packet\ send}$$

Figure 1 shows that the packet delivery ratio. In the X-axis number of nodes is taken. In the Y-axis packet delivery ratio is taken. In the existing system, a Router Selection Method (RSM) is suggested which balances the performance and anonymity. In the proposed method, a Novel Traffic, Dividing and Scheduling (NTDS) mechanism for improving performance in the Tor anonymous communication network. When compared to the existing method, there is a high packet delivery ratio in the proposed system.

5.2 Throughput

Throughput is defined as the rate of successful message delivery in the communication channel.

Figure 2 shows that the throughput. In the X-axis number of nodes is taken. In the Y-axis throughput is taken. In the existing system, a Router Selection Method

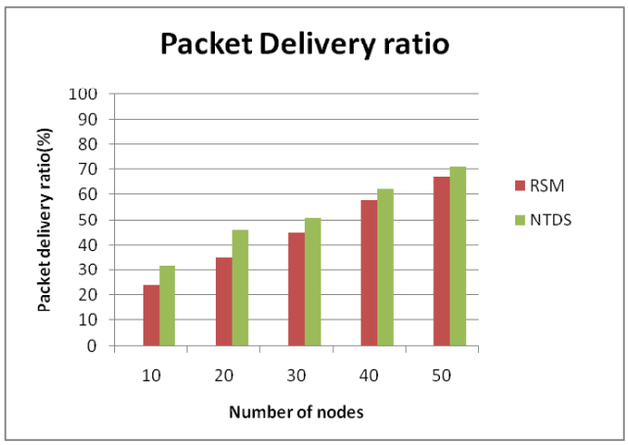

Figure 1. Packet delivery ratio.

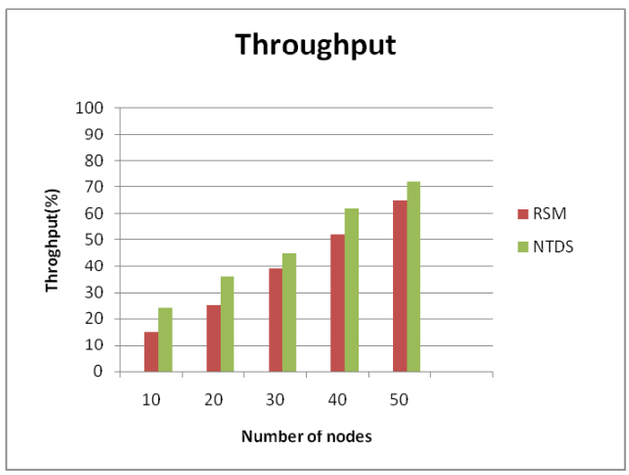

Figure 2. Throughput.

(RSM) is suggested which balances the performance and anonymity. In the proposed method, a Novel Traffic, Dividing and Scheduling (NTDS) mechanism for improving performance in the Tor anonymous communication network. When compared to the existing method, there is a high throughput of the proposed system.

5.3 End-to-end Delay

End-to-end delay is defined as the time taken for a packet to be transmitted in the network from source to destination.

Figure 3 shows that the end-to-end delays. In the X-axis number of nodes is taken. In the Y-axis end-to-end delay is taken. In the existing system, a Router Selection Method (RSM) is suggested which balances the performance and anonymity. In the proposed method, a Novel Traffic, Dividing and Scheduling (NTDS) mechanism for improving performance in the Tor anonymous communication network. When compared to the existing method, there is less end-to-end delay in the proposed system.

6. Conclusion

According to the self-reported bandwidth values, the tor anonymous communication network is preferred to build tunnels. Furthermore, tunnels are owed in proportion to this bandwidth, so there is a possibility for a malicious router operator to magnetize tunnels for cooperation. A router selection method is presented that allows accomplishing the balance between the performance and anonymity. But the problem in this method is there are less quality-of-service and the performance of the network is also less. A Novel Traffic, Dividing and Scheduling (NTDS) mechanism is introduced in the proposed system that improves the performance in the tor anonymous communication network. In the NTDS mechanism, the traffic is splitted according to the some factors like throughput, latency, congestion level, and available bandwidth capacity and assigns the traffic to the circuits. For future work, the interesting uniqueness of the Tor network are observed, which could give insight into the observed behavior of the Tor network, and which we would like to study further.

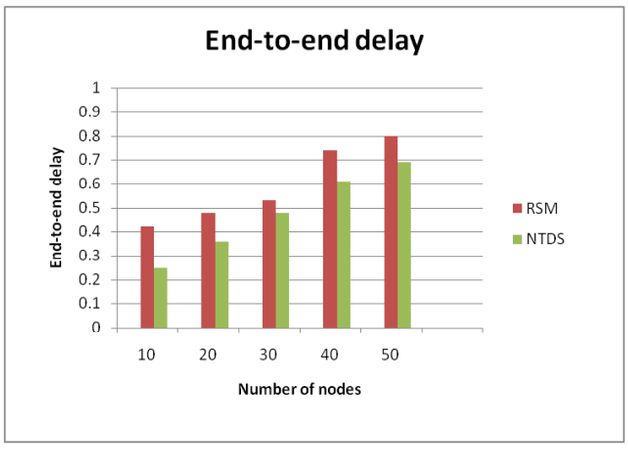

Figure 3. End-to-end delay.

7. References

1. Syverson P, Tsudik G, Reed M, Landwehr C. Towards an analysis of onion routing security. Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability. Springer-Verlag; 2000 Jul. p. 96–114.
2. Dingledine R, Mathewson N. Anonymity loves company: Usability and the network effect. Designing Security Systems That People Can Use. O'Reilly Media; 2005.
3. Dingledine R, Mathewson N, Syverson P. Tor: The second generation onion router. Proceedings of the 13th USENIX Security Symposium (USENIX Security '04); 2004 Aug.
4. Lakshminarayanan K, Padmanabhan VN. Some findings on the network performance of broadband hosts. Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement (IMC03); 2003.
5. Sherr M, Blaze M, Loo BT. Scalable link-based relay selection for anonymous routing. Proceedings of the Ninth Privacy Enhancing Technologies Symposium (PETS'09); 2009 Aug.
6. Bauer K, McCoy D, Grunwald D, Kohno T, Sicker D. Low-resource routing attacks against anonymous systems. Proceedings of the 2007 Workshop on Privacy in the Electronic Society (WPES); 2007 Oct.
7. Durresi A, Paruchuri V, Barolli L, Jain R, Takizawa M. Tokens for anonymous communications in the internet. Proceedings in IEEE Database and Expert Systems Applications; 2006.
8. Snader R, Borisov N. A tune-up for Tor: improving security, performance and anonymity in the Tor network. Proceedings of the Fifteenth Annual Network and Distributed System Security Symposium (NDSS'08); 2008 Feb.
9. Dingledine R, Mathewson N. Anonymity loves company: usability and the network effect. Designing security systems that people can use. O'Reilly Media; 2005.
10. Reardon J, Goldberg I. Improving Tor using a TCP-over-DTLS tunnel. Proceedings of the Eighteenth USENIX Security Symposium; 2009 Aug.