Multipath Based Privacy Protection Method for Data Transmission in SDN

Na Dong[1,2], Zhigeng Han[2,3], and Liangmin Wang[1,2(✉)]

[1] School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang 212013, China
wanglm@ujs.edu.cn
[2] Jiangsu Key Laboratory of Security Technology for Industrial Cyberspace, Zhenjiang, China
[3] School of Information Engineering, Nanjing Audit University, Nanjing 211815, China

Abstract. With the development of Software-Defined Networking (SDN), privacy and security issues have become an urgent problem to be solved. Although there are many ways to solve these problems, the existing technology represented by encryption cannot effectively deal with traffic analysis attacks, and there are also key management problems. For this reason, we propose a privacy protection method for SDN data transmission based on multipath, including path searching procedure for searching for all paths between the sender and the receiver, and path filtering procedure for filtering out paths to reduce path correlation, and path selection procedure for randomly selecting one path to disturbed the traffic similarity between multiple transmission. The experiment results show that our method is more effective, less similarity of traffic compared with Multipath-Floyd method and single-path method, respectively. Moreover, it is difficult for attackers to capture the traffic feature and do not need key management, which reduces the cost of the controller.

Keywords: Software-Defined Networking · Multipath filtering · Privacy protection · Traffic analysis attack

1 Introduction

Software-Defined Network (SDN) is defined as a novel network architecture, which consists of application layer, physical layer, and especially control layer. SDN has the ability to decouple data plane and control plane, and it gives the opportunity to solve the control limitations of other infrastructures. In SDN, network resources can be effectively utilized by using the centralized controller for different business requirements [11]. Moreover, it provides an overview of whole underlying network, allowing more flexible and complex management. Recently, SDN architecture has been applied in various scenarios such as data centers and enterprise.

S. Han et al. (Eds.): AICON 2019, LNICST 287, pp. 127–138, 2019.
https://doi.org/10.1007/978-3-030-22971-9_11

With the development of SDN, privacy and security issues have become an urgent problem to be solved. Especially, network programmability and control logical centralization introduce new privacy threats and attack planes. Recently, considerable researchers pay attention to these problems. Kreutz, D. et al. analyzed the potential threat vectors and presented the information about its specificity for SDN [6]. Some of them are specific to SDN as they arise from a new entity introduced subsequently-the centralized controller. And the impact of threats presented in traditional networks may be potentially augmented or expressed differently [13]. Focusing on the security and privacy preservation, the solutions have already been presented in [9,12,13]. Sha et al. [12] designed a method to measure the sensitivity-degree of information and detected the sensitive information covert channel based on the OpenFlow in SDN. There are also some methods [14] using the encryption to solve information disclosure problem. Attackers can still obtain traffic features by traffic analysis attack. At the same time, the security of encryption method highly depends on the secure and reliable key management system.

Multipath has been used to solve privacy and security problems, especially in Mobile Ad-hoc Networks (MANET) [8,10]. In the proposed multipath protocol, messages are split into multiple pieces that sent out via multiple independent paths. Attackers have to collect all pieces of the message. To enhance the communication efficiency, [5,7,15,16] researched the topology-hiding to obscure the traffic features. Multipath transmission is a useful method to improve network service performance, especially in SDN. [1,2,4,11] support routing flows through different paths to overcome the traffic congestion and physical impairment. But there are few studies that use multipath to implement SDN privacy protection.

Compared with the purpose of multipath in the above literature, the method we proposed uses multipath filtering to instead of data partition, which randomly select one trusted paths for data transmission. This avoids traffic analysis attacks that exploit the similarity between multiple transmissions in SDN. Our contributions are as follows.

- We proposed a privacy protection method for SDN data transmission. In the proposed method, the optimal transmission path can be computed and selected by the SDN controller, which resists traffic analysis attack effectively.
- We set constraints and established a novel model to filter multiple paths. Based on this model, we discussed the tradeoff between privacy protection and path correlation degree.
- Experimental results show that numerical results of path correlation degree can be obtained from the tradeoff model. Our method reduces similarity of traffic compared with single-path scenario.

The rest of this paper is organized as follows: In the next section, we introduce the system architecture and attack assumption. Description of method is presented in Sect. 3. In Sect. 4, we discuss the performance and experiment results. Finally, we conclude this paper in the last section.

2 System Description and Assumption

In this section, we introduce the system architecture, network model and attack assumption.

As shown in Fig. 1, the system consists of four entities, i.e., SDN controller, hosts, switches and attacker. Multiple switches are included in underlying network, wherein the switches connected to Host1 and Host2 are sender and receiver, respectively. In SDN controller, the global overview is provided in order to compute the different paths between Host1 and Host2.

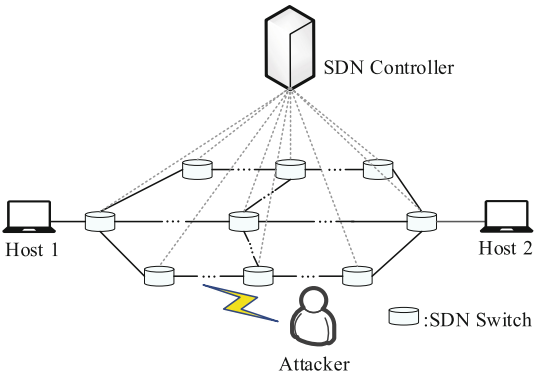


Fig. 1. Architecture, which consist of SDN controller, hosts, switches and attacker.

We formalize our target network problem. Given a graph $G = (V, E)$, where V is set of nodes in the network and E is the set of links. Each $v_i \in V$, and each link $e = (v_i, v_j) \in E$. A path is defined as a list of node $path_i = (v_1, v_2, v_3 ... v_N), \forall i, 1 \leq i \leq N$, where N is the number of nodes. The set of multiple paths is defined as $Path = (path_1, path_2 ... path_M)$, where M is the number of paths. Denoted the source node as v_s, and the destination node as v_d.

In our network model, we assume that:

- Attackers launch traffic analysis attacks by eavesdropping nodes. Attackers can attack on vulnerabilities in controller or switches. But cost of proactive attack is higher than that of reactive attack. Therefore, we assume that attacker is more inclined to reactive attack.
- SDN controller is trusted. If the controller, as the core of the whole network, conspire with the attacker, the whole network will be threatened. So, we assume that SDN controller will not disclose the path information or sensitive data to the attacker.
- The source and destination nodes are both reliable. Others are honest that not expose flow table to attacker.

If sensitive information is transmitted through a single path, attacker can break any switch in the path to cause link failure, or eavesdrop switch to analyze traffic features. While in a multipath scenario, multiple disjoint paths are allowed to be established between source node and destination node and one disjoint path is randomly selected for each transmission. If attackers want to analyze the traffic features by eavesdropping, he must simultaneously listen on all of paths. The question is how to find the multiple disjoint paths in SDN. However, fully disjoint paths are not possible, finding maximally disjoint paths are preferred, wherein it is allowed for paths to share common edges or nodes, as long as the number is minimum [4].

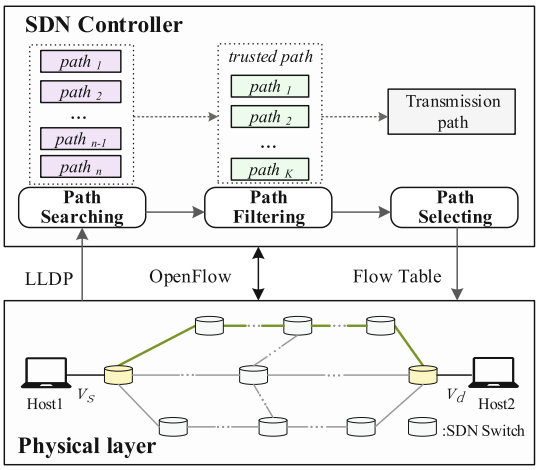


Fig. 2. Proposed method for multipath, where SDN controller focuses on obtaining transmission path and physical layer concentrates on data transmission.

3 Proposed Privacy Protection Method

The idea behind our method is to decrease the correlation between paths then randomly choose one path. Then, we exploit the randomization of multiple transmission paths to resist traffic analysis attack. Even if an attack succeeds to eavesdrop one path, the probability that the next transmission path can be analyzed is low. As shown in Fig. 2, the method is divided into three steps: (1) **Path Searching**, (2) **Path Filtering** and (3) **Path Selection**.

3.1 Path Searching

Based on the idea of Depth-First, SDN controller searches for all paths between the sender v_s and the receiver v_d. Path Searching is implemented in adjacent matrix.

We use a simple topology to describe the searching method as shown in Fig. 3. Let source and destination node be v_1 and v_9, respectively.

We search the first neighbor of v_1, i.e., v_2 as shown in Fig. 4(a), and put v_1 into the stack. Then, we do the same procedure from v_2. Until we come to the destination node, the first path, i.e., $path_1 = (v_1, v_2, v_4, v_6, v_7, v_9)$ can be obtained. Before the second path, we put out the node v_9 at the top of the stack. We start with the node v_7 as shown in Fig. 4(b), repeat the searching method. The whole searching process will be terminated until the stack is empty. Algorithm 1 describes the Path Searching process, which complexity is $O(n^2)$.

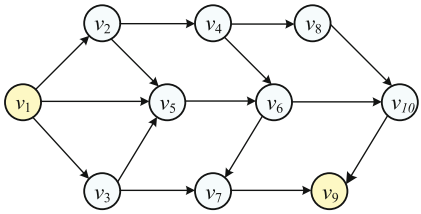
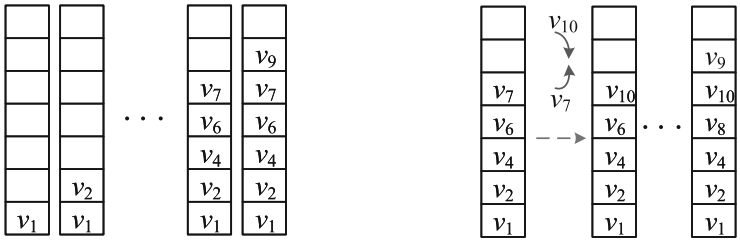


Fig. 3. A network topology with 10 nodes, where v_1 is the source node and v_9 is the destination node.



(a) The node in stack on first searching

(b) The node in stack on second searching

Fig. 4. Nodes in the stack during path searching

3.2 Path Filtering

All paths computed in the previous process need to be filtered to decrease the correlation between paths. And the paths satisfied the conditions are defined as the trusted paths. Detailed filtering conditions as follows:

Algorithm 1. Depth-First Path Searching.

Input: Adjacent matrix $M(n_{dimensionality})$, Start of path V_S, Destination of path V_D.
Output: Set of path list S.

```
while destinationofS_i ≠ V_D do
  set start of S_i: tmp
  for i in [tmp+1: n] do
    if M[tmp][i] ≠ 0 then
      tmp = i
      Add M[tmp][i] to weight of S_i
      Add i to path of S_i
    end if
    return S_i
  end for
end while
```

(1) Path Correlation Degree. We use the number of joint nodes to evaluate path correlation degree ($d(path_i)$).

$$d(path_i) = \sum_{i=1, j \leq n, i \neq n}^{N} \frac{J_{ir}}{N_i} + \frac{J_{ir}}{N_r} \tag{1}$$

where J_{ir} is the number of joint nodes between $path_i$ and $path_r$, whose number of nodes are N_i and N_r. The threshold of path correlation degree is denoted as $maxD$ that $d(path_i)$ of trusted path should less than $maxD$.

(2) Path Cost. To reduce the consumption of network resources, the cost of a single path must be smaller than the average path costs, which is denoted as $c(path_i)$.

$$c(path_i) = \sum_{i=1}^{N} c_i \tag{2}$$

where c_i presents the cost of one node when transmitting a packet. So, the average of link costs is

$$c_{average} = \frac{\sum_{i=1}^{N} c(path_i)}{n} \tag{3}$$

(3) Path Weight. After being filtered by the above conditions, if the number of paths is still large, we select the first K paths in ascending order of weights. The link weight represents delay and is denoted by w, so a path weight is

$$w(path) = \sum_{i=1}^{N} w(v_i, v_{i+1}) \tag{4}$$

3.3 Path Selection

SDN controller randomly selects one path from the set of trusted paths by generating random numbers, whose range is [1, W], where

$$W = \sum_{k=1}^{K} w(path_j) \tag{5}$$

Algorithm 2 describes the Path Selection process, which complexity is O(n^2). This process guarantees that each transmission path is randomized and different. Therefore the similarity between multiple transmission is effectively disturbed.

Algorithm 2. Path Random Selection.

Input: Path weight $T(path_i)$, Sum of path weight W, Account of path K.
Output: The path list L.

```
while i < K do
  Set the random number interval[1,W]
  for i in range (len(T) do
    if W-T[i] < 0 then
      Set the current path i
      Add i to L
      if current path = previous path then
        i++
      end if
      return L
    end if
  end for
end while
```

3.4 Tradeoff Model Between Privacy and *maxD*

In our scenario, the probability of path-attacked represents privacy protection performance, which related to the change of $maxD$. Therefore, we take some considerations: (1) enough number of paths to ensure random selection. (2) less number of joint nodes to reduce the correlation of multiple paths. (3) the more paths, the more joint nodes.

Let the probability of joint nodes and disjoint nodes being attacked be p_a and p_b, p_b<p_a, excluding the source and destination. If one node is compromised, the link is attacked. We assume that there are K trusted path, whose probability is denoted as

$$p(path_i) = 1 - (1 - p_a)^Q (1 - p_b)^{S-Q} \tag{6}$$

then,

$$p(path_i) = 1 - (1 - p_b)^S (\frac{1 - p_a}{1 - p_b})^Q \tag{7}$$

$$\frac{1-p_a}{1-p_b} < 1 \tag{8}$$

where the intermediate nodes number of $path_i$ is S, and the joint number excepted the source and destination is Q.

Under the constraints of paths weight and cost, the total number of nodes S in each path has little difference. So, the number of joint nodes will affect $p(path_i)$ from Eqs. (7) and (8).

The average probability of k paths is

$$p_{average} = \frac{\sum_{k=1}^{K} p(path_k)}{K} \tag{9}$$

We note that the less $p_{average}$ is, the more security paths we use to transmit sensitive information. When we set $maxD$ in Path Selection larger, the K increases, which results the correlation of filtered paths $d(path_i)$ increases. However, if the $maxD$ decreases, correlation of each path decreases, the performance degrades in terms of privacy protection.

In order to minimize $p_{average}$, we propose a tradeoff model as a flowchart illustrated in Fig. 5, which can obtain the optimal threshold $maxD$. First, the p_a and p_b are given, and $maxD$ is set the minimum initially. Then, we filter paths that be searched in Path Searching process and calculate the $p_{average}$. If the $p_{average}$ decreases as threshold $maxD$ increases, the step is repeated. Finally, $maxD$ is determined and the tradeoff is balanced.

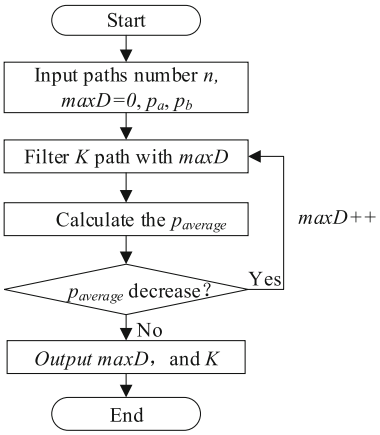


Fig. 5. Tradeoff model between $maxD$ and $p_{average}$

4 Evaluation and Result

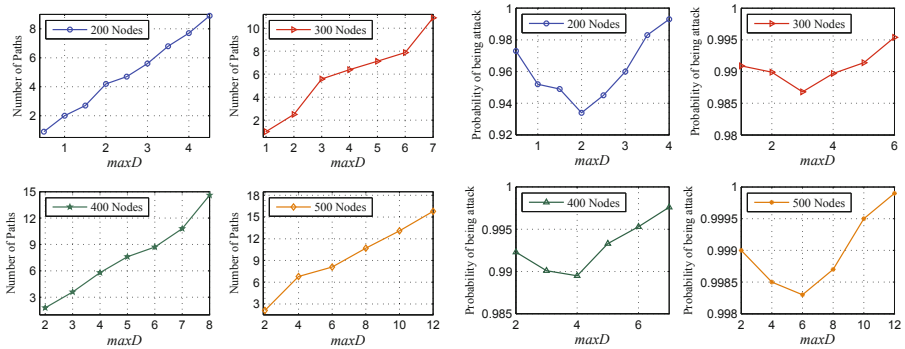
In this section, we show the numerical results to discuss the performance of our method. We tested the implementation on mininet that deployed topologies with

different number of nodes; RYU as the SDN controller. The experiments were conducted using the virtual machine running 64-bit Ubuntu 14.04, with 4 Gb of RAM. And we configured networks at 200, 300, 400, 500 nodes, with a random topologies.

We evaluated performance of the proposed method from two aspects: (1) the probability of path being attacked $p_{averagw}$ when threshold $maxD$ increases. (2) performance of multipath method with the optimal threshold $maxD$.

4.1 Tradeoff Results

Figure 6 indicates the effect of threshold $maxD$ on the number of paths and the probability of path being attacked under different scale network. We assumed p_a and p_b mentioned in Sect. 3.4 is 0.8 and 0.2, respectively. The trend we observed in Fig. 6(a) is, as the $maxD$ is strictly limited, the path number K become large. But the probability $p_{average}$ can be minimized when $maxD$ takes the appropriate value in Fig. 6(b). Combining both figures, we observed the optimal $maxD$ and K that the path are trust. For example, when we set $maxD$ 6, the probability $p_{average}$ is minimum in the network size of 500 nodes and the average number of paths is 8.3.



(a) The number of path (b) Probability of path being attacked

Fig. 6. Tradeoff model of $maxD$ with different number of nodes

4.2 Performance of Multipath

We confirmed the effect of our proposed method by comparing it with the existing method Multipath-Floyd [3]. Figure 7 shows an average of the number of paths. As the size of network increases, the number of paths in our method are greater than Multipath-Floyd. Meanwhile, the correlation between multiple paths decrease relatively. We allow paths to share nodes or edges, while fully disjoint paths are obtained in Multipath-Floyd.

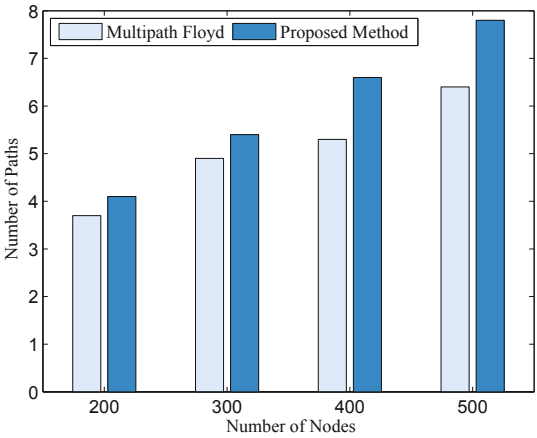


Fig. 7. Comparison between Multipath-Floyd and our method

We also discussed ability to resist traffic analysis attacks. We simulated the a communication between hosts in the different scenarios. Dijkstra algorithm are used to calculate path in single path scenario. From Fig. 8, traffic of single path transmission exhibits high similarity, while the correlation between multiple traffic has been disturbed by our method. Therefore, it is difficult for an attacker to analyze the real path of sensitive data based on traffic similarity.

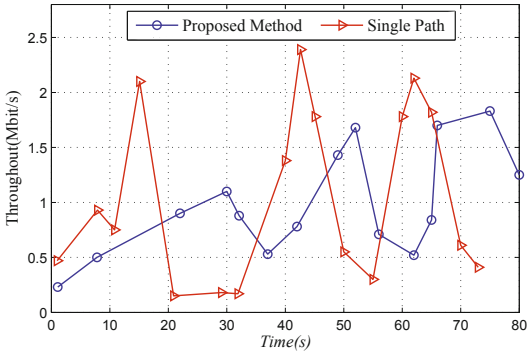


Fig. 8. Comparison between single path and proposed method

5 Conclusion

In this paper, we presented a multipath filtering method to enhance privacy of SDN data transmission, which has the ability to effectively resist traffic attack. First, Depth-First path searching algorithm was adopted to compute each path

from the transmitter to the receiver. Then, we set conditions to filter all paths to obtain the set of trusted paths. Moreover, we established a novel model to discuss privacy protection and $maxD$. The experiments results show that our method obtained optimal $maxD$, which reduces the correlation of paths. Compared with single path scenario, privacy protection can be improved in proposed method. The situation that the probability of nodes being attacked is a variable will be an open topic. We will pay more attention in the future.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (U1736216), the natural science foundation of Jiangsu Province (BK20151460) and the University Natural Science Foundation of Jiangsu Province (16KJB520021).

References

1. Duan, J., Wang, Z., Wu, C.: Responsive multipath TCP in SDN-based datacenters. In: 2015 IEEE International Conference on Communications (ICC), London, UK, pp. 5296–5301, June 2015. https://doi.org/10.1109/ICC.2015.7249165
2. Dulinski, Z., Rzym, G., Cholda, P.: MPLS-based reduction of flow table entries in SDN switches supporting multipath transmission. Networking and Internet Architecture arXiv:1805.07993 (2018)
3. Guan, Y., Lei, W., Zhang, W., Liu, S., Li, H.: Scalable orchestration of software defined service overlay network for multipath transmission. Comput. Netw. **137**, 132–146 (2018)
4. Guillen, L., Izumi, S., Abe, T., Suganuma, T., Muraoka, H.: SDN implementation of multipath discovery to improve network performance in distributed storage systems. In: 2017 13th International Conference on Network and Service Management (CNSM), vol. 1, pp. 1–4. IEEE Computer Society, Tokyo, November 2017. https://doi.org/10.23919/CNSM.2017.8256054
5. Jose, J., Rigi, R.C.: A comparative study of topology enabled and topology hiding multipath routing protocols in MANETs. In: 2015 International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO), Visakhapatnam, India, pp. 1–4, January 2015. https://doi.org/10.1109/EESCO.2015.7254001
6. Kreutz, D., Ramos, F.M., Verissimo, P.: Towards secure and dependable Software-Defined Networks. In: Proceedings of the Second ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking, HotSDN 2013, pp. 55–60. ACM, New York (2013). https://doi.org/10.1145/2491185.2491199
7. Liu, H., Wang, Z., Miao, F.: Concurrent multipath traffic impersonating for enhancing communication privacy. Int. J. Commun. Syst. **27**(11), 2985–2996 (2014)
8. Lou, W., Liu, W., Zhang, Y., Fang, Y.: SPREAD: improving network security by multipath routing in mobile ad hoc networks. Wireless Netw. **15**(3), 279–294 (2009)
9. Nakahara, M., Shinkuma, R., Yamaguchi, K., Yamaguchi, K.: Tradeoff between privacy protection and network resource in community associated network virtualization. In: 2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Hong Kong, China, pp. 2143–2148, August 2015. https://doi.org/10.1109/PIMRC.2015.7343652

10. Othman, J.B., Mokdad, L.: Enhancing data security in ad hoc networks based on multipath routing. J. Parallel Distrib. Comput. **70**(3), 309–316 (2010)
11. Pasca, S.T.V., Kodali, S.S.P., Kataoka, K.: AMPS: application aware multipath flow routing using machine learning in SDN. In: 2017 Twenty-Third National Conference on Communications, Chennai, India (NCC), pp. 1–6, March 2017. https://doi.org/10.1109/NCC.2017.8077095
12. Sha, L., He, L., Fu, J., Sun, J., Li, P.: SDN-based sensitive information SI protection: sensitivity-degree measurement in software and data lifetime supervisor in Software Defined Network. Sec. Commun. Netw. **9**(13), 1944–1957 (2016)
13. Wang, Y., Chau, P., Chen, F.: Towards a secured network virtualization. Comput. Netw. **104**(C), 55–65 (2016)
14. Zeng, T., Meng, S., Wang, M., Zhu, L., Fan, L.: Self-adaptive anonymous communication scheme under SDN architecture. In: 2015 IEEE 34th International Performance Computing and Communications Conference (IPCCC), Nanjing, China, pp. 1–8. December 2015. https://doi.org/10.1109/PCCC.2015.7410337
15. Zhang, Y., Tan, Y., Jie, T., Qi, H., Wang, G., Li, Z.: TOHIP: a topology-hiding multipath routing protocol in mobile ad hoc networks. Ad Hoc Netw. **21**(5), 109–122 (2014)
16. Zhang, Y., Wang, G., Hu, Q., Li, Z., Tian, J.: Design and performance study of a topology-hiding multipath routing protocol for mobile ad hoc networks. In: 2012 Proceedings IEEE INFOCOM, Orlando, FL, USA, pp. 10–18, March 2012. https://doi.org/10.1109/INFCOM.2012.6195468