# A novel approach based on ant colony system for IP traceback

Menglin Liu [a], Zhengping Jin [b]

State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China.

[a]lml_bupt@163.com, [b]zhpjin@bupt.edu.cn

**Abstract.** Ant colony algorithm is often used to solve the IP traceback problem without the entire network routing information. However, such algorithm is easy to converge to a local suboptimal solution. Especially with the increase of network topology size, the real attacker is more difficult to find. In this paper, a novel approach based on colony system is proposed to identify the source of the DDoS attack. This approach brings in a new concept of the network topology division before tracing the attack by the ant colony algorithm. The division based on flow information can enhance the ability of the ants to search a more globally optimal solution for the attack path, even if the network topology is large-scale. The performance of the novel approach in reconstruction the attack path is evaluated through a series of ns2 simulations. The simulations results show that the proposed scheme has better performance than the conventional ACS algorithm, e.g. the performance of the novel ACO approach has been about 10% higher than the improvement ACO scheme, when the network topology is p=800.

**Keywords:** DDoS; IP traceback; Ant colony optimization (ACO); Louvain algorithm.

## 1. Introduction

DDoS attack is causing great harm to the entire Internet and cause serious economic losses, which is one of the attacks hardest to defend. IP traceback is crucial to solve the network security issues, and this kind of technology is used to reverse tracking data packets to the source of the attack. The main method of the current IP traceback, such as link-testing, packet-marking, messaging, and logging. For more about the methods, you can refer to Karanpreet Singh et al. [6].

The methods that mentioned above either requiring to modify the network infrastructure or requiring all routing information along the DDoS attacking path between attack and victim. In view of the drawback of the above methods, Gu Hsin Lai et al. [1] proposed an Ant-based IP traceback scheme. This scheme utilized the existing traffic flow information as the trace for ants finding the attack path, even being able to identify the origin as the source address could be spoofed. Ping Wang et al. [11] improved the scheme and proposed a modified ACS (ant colony system) optimization scheme with a good effect. The latest one is ACS-TPTBK scheme which is verified that have a significant increase in the effect of convergence, but when the network topology becomes greater, it’s difficult for ants to explore enough areas, which may leads to a local suboptimal solution, and the real attacker could be missed.

In view of the above problems, in this paper, we propose a novel method based on ACO for IP traceback problem. We took the network topology partition method to divide the larger network topology into some small communities. Then we use ACO algorithm to search for the optimal solution in the community. Finally, we do the overall global optimization, and last step community optimization results provides global optimization for initialization of pheromone intensity. Due to the higher pheromone intensity attracts more ants, ants can expand a broader area according to pheromone concentration, and thus the most properly attack path was obtained through repeated iteration. So that ants can search for the most probable attack path with limited routing information and computational time, even the network topology scale is large. The effectiveness of the proposed approach was demonstrated using a series of ns2 simulations.

The remainder of this paper is organized as follows: Section 2 introduces in detail the novel approach in this study. Section3 presents the effects of the novel approach for IP traceback compared with previous ACS method. Section 4 provides a brief summary and prospects for the future in this field.

## 2. The novel approach for IP traceback

The novel IP traceback scheme mainly consists of three key steps: network topology division, optimizing in the community, and global optimization.

We use the Louvain Method to carry out the network topology division. The Louvain Method is a simple, efficient and easy-to-implement method for identifying communities in large networks created by Vincent Blondel et al. [12].

The attack path reconstruction in response to a DDoS attack is a type of graph optimization. Network topology can be viewed as a directed graph, G= (V, E), where V represents a set of nodes, V$\{v_1, v_2, \dots, v_n\}$: $V_s$ is a set of attack source nodes, $V_d$ is a set of victims nodes, and E denotes the graph edges. The IP traceback is mainly in order to determine the most probable attack path between $V_s$ and $V_d$ . The distance between each node is regarded as the weight of the directed graph. Then the network topology can be divided into some community, as shown in Fig. 1

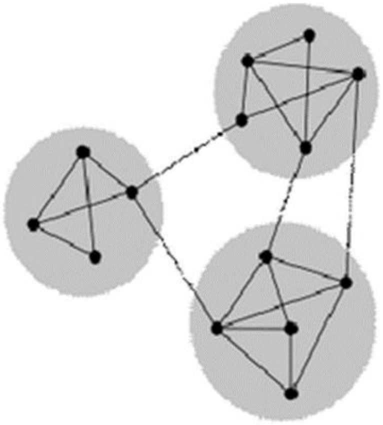


Fig. 1. network topology division

After the large network topology has been divided into $C_n$ communities, we use the improved ACO scheme to search the most probable attack path in each community. The improved ACS schemes shows specifically as follows:

If an ant wants to move from node i to node j, the choice probability of node j ($p_{ij}(t)$) is calculated by the state transition rule. The state transition rule is as follows:

$$j = \begin{cases} \arg \quad \max_{j \notin tabu_k} \left\{ \left[ \tau_{ij}(t)^{\alpha} \right] \left[ \eta_{ij}(t)^{\beta} \right] \right\}, \text{if } q \le q^0 \\ \qquad\qquad S, \end{cases} \tag{1}$$

$$S = p_{ij}(t) = \begin{cases} \dfrac{\left[\tau_{ij}(t)\right]^{\alpha} \left[\eta_{ij}(t)\right]^{\beta}}{\sum \left[\tau_{ij}(t)\right]^{\alpha} \left[\eta_{ij}(t)\right]^{\beta}}, j \in N_i \\ 0, \qquad\qquad \text{otherwise} \end{cases} \tag{2}$$

Where $\tau_{ij}(t)$ denotes the intensity of pheromone on path ij at time t, and $\eta_{ij}(t)$ is calculated as the number of routing packets between routers i and j in duration of time t-1 to t. Furthermore, $\alpha$ and $\beta$ are the weighting factors. Finally, $N_i$ is a set of nodes in the neighborhood of node i, which the ant has not yet visited. Where $q^0$ is a user-defined parameter which specifies the distribution ratio of the two policies, and is $q$ random number in the interval [0, 1].

On a tour, each ant updates the pheromone intensity on the traversed trail by applying the local updating rule. Once all the ants have completed their tours, the pheromone intensity on a path is updated once again through the global updating rule.

### 2.1 Local updating rule

The pheromone intensity on each arc of the path is updated as follows when an ant traverses a path.

$$T_{ij}(t+1) = (1-\omega)\tau_{ij}(t) + \omega \Delta \tau_{ij} \tag{3}$$

Where w represents the decay rate of the local pheromone and has a value in the interval [0, 1]. Moreover, $\Delta\tau_{ij}(t)$ represents the additional pheromone deposited on path ij during the current period, and it is calculated by the formulation (4) and (5), expressed below.

$$\Delta\tau_{ij}(t) = \sum_{k=1}^{m} \Delta\tau_{ij}^{k}(t) \tag{4}$$

Where $\Delta\tau_{ij}^{k}(t)$ is the amount of pheromone laid on path $ij$ by the kth ant between time $(t-1)$ and $t$ , and m is the total number of ants in the colony. The value of $\Delta\tau_{ij}^{k}(t)$ is calculated by the strategy of Ant-cycle system, that is:

$$\Delta\tau_{ij}^{k}(t) = \begin{cases} \dfrac{Q_k}{L_k}, \text{if node } j \text{ in path visited by } k\text{th ant} \\ 0, \qquad\qquad \text{otherwise} \end{cases} \tag{5}$$

Where $Q_k$ the total amount of the octets belonging to the DDoS is attack on the kth ant's path, and $L_k$ is the number of nodes along the path traversed by ant $k$

### 2.2 Global updating rule

Once all the ants have completed their tours in the current iteration, the intensity of the pheromone assigned to each arc the most probable path is recalculated as follows:

$$\tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t) + \rho\Delta\tau_{ij}(t) \tag{6}$$

$$\Delta\tau_{ij}(t) = \begin{cases} \dfrac{C}{L_G}, \text{ if route } (i,j) \text{ is on the optimal path} \\ 0, \qquad\qquad \text{othewise} \end{cases} \tag{7}$$

Where C is a constant and $L_G$ is the number of nodes along the most probable path. Eqs.(1)-(7) were applied repeated iteration until the counter pre-set was reached.

We assume that the number of ants in each community is $m_1, m_2, \cdots, m_{C_n}$ , respectively. Because of the flow of traffic on the DDoS attack path is obviously higher than the normal, we put the ants on the largest traffic node in the community at the beginning of search. The select probabilities of the next node and the pheromone intensity of the node is calculated by Eqs.(1)-(7), and the initial pheromone intensity of each node in each community is set to a same value $\tau_0$. In this optimization procedure, we assume that the pre-set iteration cycle is Max_loop1 for all communities.

When all communities' optimization were completed, the pheromone intensity of each node is recorded in the moment, and we begin to search the most probable attack path with ACO scheme in the whole network topology. We assumed that the ant colony is composed of m ants, and the pre-set iteration cycle is Max_loop2. The select probabilities of the next node and the pheromone intensity of the node is calculated by Eqs. (1)-(7), and the initial pheromone intensity of each node is set to the value obtained by last step that has been recorded. Optimization process is repeated iterative until the counter Max_loop2 is reached, and the most probable attack path can be obtained at this time.

## 3. Testing and validation

In order to investigate the effectiveness of the conventional ACS, improved ACS, and the novel scheme proposed in this paper, a series of ns2 simulations was conducted. The simulation were performed using a PC with an Inter Dual core CPU 3.2G, 4GB RAM and Windows 7 operating system. The simulations followed a three-step procedure:

### 3.1 Creating the network topology

Various experimental network topologies were constructed using a random graph generator according to the Waxman model [13]. The generator randomly placed 200,300,500,600,800 nodes (p=200,300,500,600,800) at integer coordinates distributed over a rectangular area of size 50 × 50 or 100 × 100. Adjacent nodes, $v_i$ and $v_j$, were connected to form edges with a probability of

$$P(i,j) = \eta \exp(\frac{-d(i,j)}{L_y}) \tag{8}$$

Where $\mathrm{d}(i,j)$ is the Euclidean distance between nodes $v_i$ and $v_j$ , and L is the maximum possible distance between any two nodes in the topology. Furthermore, η and γ are parameters with values in the interval [0, 1] and used to vary the graph characteristics. For this five experiments, various values of η in the range 0.3-0.7and γ in the range 0.2-0.6 were set.

### 3.2 Reconstructing the attack paths

The attack paths were constructed using the following three-step procedure:

The first step is attack simulation. To simulate 600 random attacks in the network topology that has been generated, the Monte Carlo method was adopted. Sixty ants were placed randomly in the network nodes to generate routing information for each network node in each loop. A total of 2000 loops (MAX_LOOP) were performed, and the average number of packets collected by each node was computed and used as the basis for updating the pheromone intensity $\tau_{ij}(t)$ at the corresponding node for assisting ants to search the routes.

The second step is network topology division which used the Louvain Method.

The last step is to trace back of attack paths. The routing information generated in the first step was used as the input dataset for the test. To implement the novel scheme, the ant colony was assumed to have 30 members in each community that has been divided in the second step. Furthermore, Max_loop1 and Max_loop2 were set as 60 and 100, respectively. Finally, the routing information was averaged to determine the mean number of routing packets collected at each node. In addition, α (pheromone intensity weighting factor) and β (visibility weighting factor) were assigned values of 1.5 and 1.2, respectively. The decay rate of the pheromone in (3) and (6) were set as 0.4 and 0.6, respectively. The constant C in (6) was set as 300.

A total of 17 feasible attack paths were identified in the network topology with p=500. Fig. 2 depicts the number of ants which traversed each path in each iteration of the optimization procedure. We can draw a conclusion through the chart that path 9 is the most probable attack path

### 3.3 Performance comparison of the scheme for IP traceback

In order to compare the effectiveness of these scheme for IP traceback, we conducted a series of comparative experiments. We assumed that the ant colony had 32 members in conventional ACO scheme and improved scheme, and the ant colony (m=32) was divided into 8 subgroups (each subgroup having four ants) for ACS-IPTBK scheme.

We adopt the evaluation criteria proposed by [11], that is:

$$\text{coverage percentage(\%)}=\frac{\text{Average number of packets / attack path}}{\text{Total number of routing p}} \tag{9}$$

Where the average number of packets on the attack path is computed as the total number of packets on the path divided by the routing distance(in hops). If the converged solution is not the true attack node, the average number of packets on the path is reset to zero and the search for the true path is resumed. Fig. 3(a)-(c) present the performance evaluation results for the above schemes for IP traceback with p=300, 500, 800 nodes.

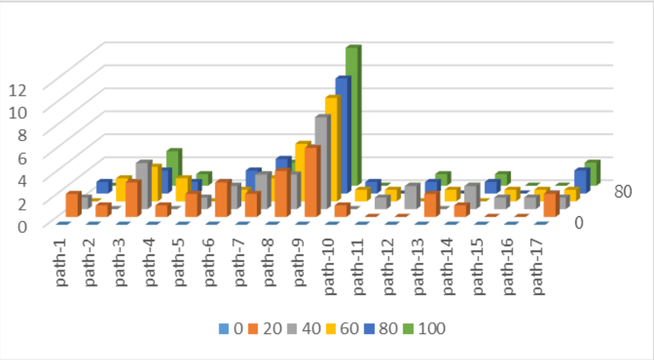
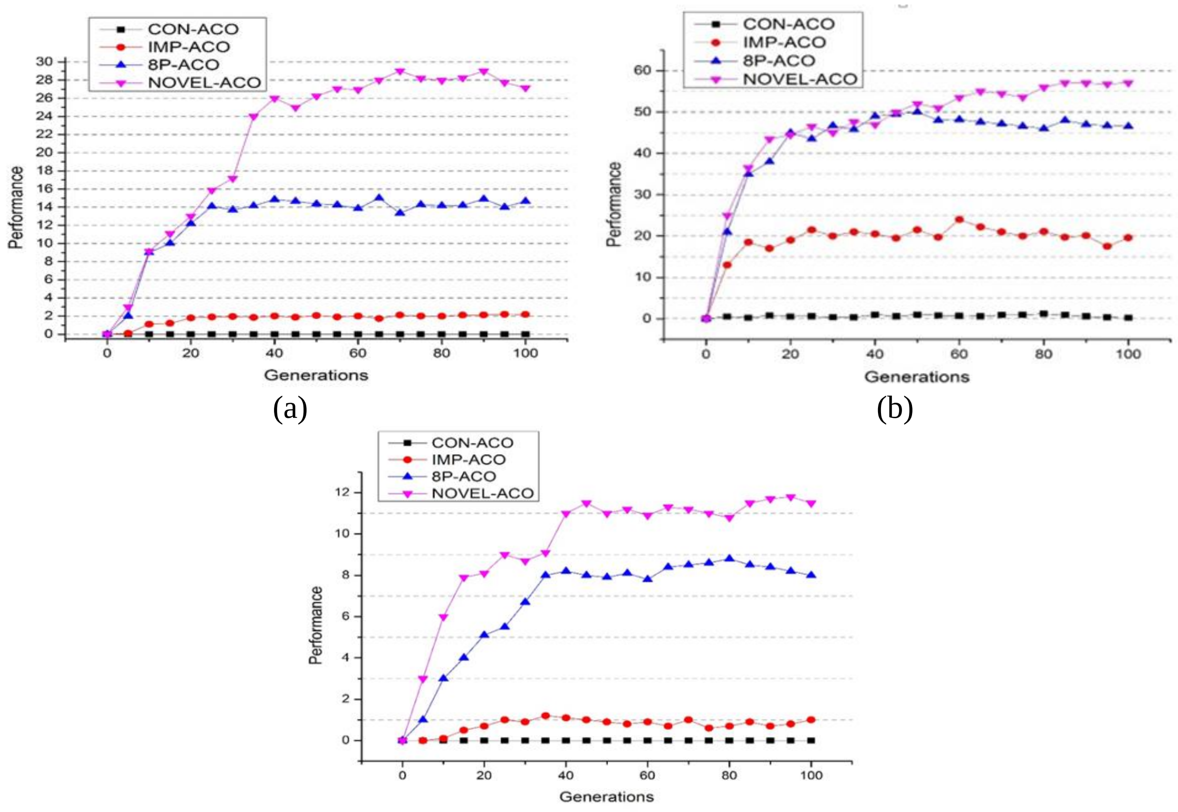


Fig. 2. X-the attack path ants searching, Y- generation of the the search process, Z-the number of ants on each path.



(a) (b)

(c)

Fig. 3. (a) Percentage of ants on most probable attack route (p=300), (b) Percentage of ants on most probable attack route (p=500), (c) Percentage of ants on most probable attack route (p=800).

The results presented in Fig. 3(a) show that for a small network topology (i.e.,p=300), the novel scheme performance is almost the same as ACS-IPTBK scheme(8-p), and significantly than that of the conventional ACO scheme. For the large network topologies, the novel scheme performance preceded over others too much.

## 4. Conclusion

This study presented a novel scheme that combining ACO algorithm and network division, so that the ability of ants for searching the attack path was enhanced. The simulation results have confirmed that the proposed novel scheme has better search effect than previous schemes, especially in the case that the network topology scale is large. The proposed scheme get rid of the limitations of the network topology size for IP traceback.

However, in this paper, specific parameter analysis and spoofed IP attack of the proposed scheme for IP traceback were not been considered too much. For these aspects and other related issues about IP traceback problem, we will leave after to detailed study. Furthermore, collecting and classifying attack information in real time is still a nontrivial task. Thus, routing information from various attack events must be collected and consolidated in a database in advance.

## Acknowledgments

This work is supported by NSFC (Grant Nos. 61300181, 61502044), the Fundamental Research Funds for the Central Universities (Grant No. 2015RC23).

## References

[1] Lai G. H., Chen C. M., Jeng B. C., Chao W., Ant-based IP tracebak, Expert Syst.Appti, 34(2008) 3071-3080.

[2] Hamedi-Hamzehkolaie M, Shamani MJ, Ghaznavi-Ghoushchi MB. Low rate DOS traceback based on sum of flows. In: Proceedings of the International Symposium on Telecommunications; 2012, 1142–6.

[3] Savage S, Wetherall D, Karlin A, Anderson T. Practical network support for IP traceback. ACM SIGCOMM Comput Commun Rev 2000; 30(4):295–306.

[4] Hsu F, Chiueh TA. Path information caching and aggregation approach to traffic source identification. In: Proceedings of the International Conference on Distributed Computing Systems; 2003, 332–9.

[5] Snoeren AC, Partridge C, Sanchez LA, Jones CE, Tchakountio F, Kent ST, et al. Hash-based IP traceback. ACM SIGCOMM Comput Commun Rev 2001;31(4):3–14.

[6] Karanpreet S, Paramvir S, Krishan K. A systematic review of IP traceback schemes for denial of service attacks. computers & security 56 (2016)111–139.

[7] Dorigo M., Maniezzo V., & Colorni A. The ant system: An autocatalytic optimizing process, Technical Report No. 91-016(1991) Rev, Politecnico di Milano, Italy.

[8] Bell J. E., McMullen P. R., Ant colony optimization techniques for the vehicle routing problem, Adv. Eng. Inf.1 (8) (2004) 41–48.

[9] R. Schoonderwoerd, O. Holland, J. Bruten, L. Rothkrantz, Ant-based load balancing in telecommunications networks, Adapt. Behav. 5(1997)169–207.

[10] R. Hadji, M. Rahoual, E. Talbi, V. Bachelet, Ant colonies for the set covering problem, in: Proceedings of ANTS 2000, Second International Work shop on Ant Algorithms: From Ant Colonies to Artificial Ants, 2000:63–66.

[11] Ping W., Hui-Tang L., Tzy-Shiah W. An improved ant colony system algorithm for solving the IP traceback problem. Information Sciences. 326(2016)172-187.

[12] Vincent D Blonde, Jean-Loup Guillaume, Renaud Lambiotte and Etienne Lefebvre, Fast unfolding of communities in large networks, J. Stat. Mech. (2008) P10008.

[13] B. M. Waxman, Routing of multi point connections, IEEEJ. Sel. AreaCommun. 6 (9) (1988) 1617–1622.