MODIFIED ANT COLONY SYSTEM ALGORITHM FOR IP-TRACEBACK PROBLEM

Prathamesh Vijay Tarare[1], AnurajMalav[2], G. Usha Devi[3]

[1]*School of Information Technology and Engineering, VIT University Vellore*
[2]*School of Information Technology and Engineering, VIT University Vellore*
[3]*School of Information Technology and Engineering, VIT University Vellore*

Abstract

Determining the origin of packet over the internet is called IP traceback. When the network is formed, it also contains the attacker in the form of viruses, malwares, Trojans etc. It is very difficult to identify the exact location of these attackers which is termed as IP tracebackproblem. To deal with this problem, ant colony algorithm can be effectively used. Hence the ant colony optimization algorithm (ACO) which is a probabilistic strategy for rectifying the computational issues is used to find optimal attack path. This algorithm is inspired by behaviour of real ants.Ants initially move randomly, and when they find food on any path they fetch it and lay down Pheromone for other ants to follow the same. On the off chance that different ants find such a way, they are likely not to continue going aimlessly, but instead follow the trail. But the main issue with these algorithms is that they converge to local suboptimal solutions and the feasible attack path is not identified properly. To overcome this problem an algorithm,called modified-ACS scheme is proposed. A global heuristic mechanism is used by this algorithm to find a feasible attack path. The algorithm is implemented in NS2.To obtain sufficient routing information of the network, we are using Hamiltonian Routing Strategy. Hence, we conclude that modified-ACS algorithm gives slightly more convergence time than conventional algorithms but yields more globally optimised solution.

Keywords- *Ant Colony System Algorithm, IP Traceback, Hamiltonian Routing Strategy, Attack path, Byzantine attack simulation, NS2*

----------***----------

1. INTRODUCTION

From the past few years, a special attention is given to secure the internet infrastructure as it has become an important media for communication, transaction, transmission and storage.It is continuously being threatened by malwares, viruses, IP-spoofing attacks etc. Today, DDoS attacks evolved in strategy and tactics. These days the problem known as “Smokescreening” is acquiring more attention. The aggressors make utilization of DDoS to exasperate IT staff while embeddings malware to rupture organization database and other essential data. More than half of assaulted organizations reported burglary of assets, information or licensed innovation. Such cyber-attacks lives for short duration of time but are very dangerous, more harmful than sustained strikes whose goal is extended downtime. While DDoS[12-14] attacks were successful in distracting IT and security teams.Criminals snatch and clone private information to siphon off protected innovation, assets and that's only the tip of the iceberg. In one case, hackers used DDoS[6] for stealing Bankcustomer’saccreditations and channel $9 million from ATMs in only 48 hours. Such incident is known as smokescreening. According to Neustar report in 2013, those attacks that requiremore than 6 individuals to reduce almost multiplied to 56 percent contrasted with 33 percent in 2012. Besides, DDoS[15,16] assaults that requires 10 individuals and more have put out the fire dramatically increased, expanding from 20 percent In 2012 to almost 41 percent in 2013. Such problems come under combinatorial optimization problem.To deal with this we make use of Ant Colony Optimisation algorithm, which is a heuristic artificial intelligence algorithm.This algorithm is based on the food gathering behaviour of ants. It makes use of routing information to find the attack path.However in practice it may not be possible to obtain this information by service provider. Thereby it uses partial information to construct attack path. Moreover by using the advance searching technique time complexity and convergence speed is minimised comparedto conventional algorithms. The whole algorithm is implemented in NS2 platform.

2. RELATED WORK

This section provides a various existing methods for solving IP traceback problem caused by DDoS attacks.

Existing IP Tracebackframeworks for DDOS Attacks

Distributed Denial of Service Attack (DDoS) attack [3-5] is a kind of attack that has an objective to compile several systems around the internet having infected by zombies/agents and form botnet networks. Targeted network is being attack by zombies using different types of packets. These attacker remotely control the targeted system and launch the packet flood.One of the methods chosen by Alan Saied, Richard E.Over ill, Tomasz Radzik is Artificial

Neural Network (ANN) algorithm.This algorithm is based on identifying specific patterns that are responsible for separating DDoS attack traffic from other authentic traffic.They expanded the chances of detecting known as well as unknown DDoS attacks [(Alan Saied, 2015)]by training their algorithm with up-to-date patterns and avoiding repetitive patterns. This is on account of the ANN calculation gains from situations and distinguishes zero-day designs that are like what it was prepared with. The reason behind it is that ANN algorithm adapts learning from scenarios and also detects zero-day patterns having similarity with what it was trained with.Similarly, to detect DDoS[9-11] attack Jie-Hao and Ming [2] make use of comparison of detection outcome with entropy, ANN, decision tree and Bayesian. There are numerous proposals available concentrating on various methodologies. Probabilistic packet marking (PPM) was proposed by Savage et al. [3]. In this approach the computational research load was reduced. During the forwarding process each packet was probabilistically marked using partial path information. The IP traceback information was stored in IP fragment identification field. This PPM approach has limitations too, as creation of spoofed IP addresses is still an easy task and it doesn't trace multiple attack paths. The modification to Savage method was proposed by Song and Perrig in 2001. By storing a hash of each IP addresses they reduced the storage requirements [4]. After reassembling the edge fragments, a comparison between hashes of router IP addresses andhashes of resulting IP addresses. This comparison helps for reconstructing the attack path.

Hash-based IP traceback scheme named as source path isolation engine (SPIE) [5] was proposed by Soneren et al. This mechanism uses a 32-bit hash function for created packet digest log rather than IP data packets. The results yielded by this mechanism give compelling decrease in packet storage.Colorni et al. first proposed ACO algorithm called Ant System [17]. Later new search strategies known as: ant-density, ant-quantity and ant-cycle were emerged. Ant Colony system was developed by Dorigo and Gambardella to search most feasible path under complex time and cost constraint [7].To prevent premature convergence to local suboptimal solution,Stuetzle and Hoos proposed [8] a max-min Ant system involving pheromone intensity update range.

ANT Algorithm

ANT Algorithm was motivated by the conduct of characteristic ants and after that connected to a wide range of discrete improvement issues, for example, vehicle directing and resource scheduling. In an ANT algorithmcalculation, different operators, spoken to by ants, collaborate with one another utilizing backhanded correspondence interceded by pheromone. This algorithm was initially acquainted to tackle the travelling salesman problem.A moving ant lays some pheromone on the ground, hence denoting the way it takes after by a trail of this substance. While a separated ant moves irregularly, an ant experiencing a formerly laid trail can identify it and choose with a high likelihood to tail it, fortifying the trail with its own pheromone. A noteworthy normal for an ant algorithm is the reasonability of autocatalytic procedures. A "single ant" autocatalytic procedure generally unites rapidly to an awful problematic arrangement. Fortunately, the association of numerous autocatalytic procedures can prompt fast convergence to a subspace of the arrangement space that contains numerous great arrangements, creating the hunt movement to discover rapidly a decent arrangement, without getting stuck in it.At the end , every one of the ants convergenot to a solitary arrangement, but rather to a subspace of arrangements,from that point they continue hunting down upgrades of the best discovered arrangement. Thus, we trust this component will be useful for discovering DoS path. The pheromone can support, rather than short edge, suspicious DoS path and after that shape a positive input prompting the DoS beginning.

3. PROPOSED-WORK

3.1 Hamiltonian Routing Strategy

In general, in an m × n 2D mesh network with an aggregate number of N = m×nrouters, every routers area compares to the Cartesian coordinate framework. Hence, every router i,0≤i≤N is determined with its direction ($u_{i,}v_i$) such that it demonstrate the router's position along the U and V headings, separately. In the Hamiltonian Routing strategy, every router is allocated a mark between 0 and N-1. Beginning from the first router with (0,0) directions, the routers in even rows are given names from left to right, and the routers situated in odd rows are marked from right to left.

The naming function could be summed up as

$$l(u,v) = \begin{cases} v \times n + u & if\ v\ is\ even \\ v \times n + n - u - 1 & if\ v\ is\ odd \end{cases}$$

Any path in a graph that is visited by each node precisely once is called Hamiltonian path [16]. In light of the checking, two disjoint subnetworks could be recognized in the System: a low-channel (HL) subnetwork.And a high-channel (HH) subnetwork . Beginning from (0,0) , nodes are visited in an ascending order in $H_H(H_L)$. Routing in the subnetwork takes place when the destination label is more prominent then source. This routing methodology ensures live lock and deadlock free routing. This is because of the way that the packets get near to the destinations by going by the nodes in an entirely rising request along the paths in $H_H(H_L)$. In this manner, the routing has been confined such that no cyclic reliance is framed between the nodes.

3.2 Modified-ACS method

(i) Network Topology Generation

Topology in the network is created through Waxman model. Through this model adjacent nodes x and y are connected via probability of

$$P(x,y)=\mu \exp\frac{-\delta(x,y)}{L\gamma}$$

Where δ(x,y) is the Euclidean distance and L is the largest probable distance between two nodes. μ And γ are parameters in interval of [0, 1].

(ii) Attack Path Construction

(a)A module known as Byzantine attack simulation which is a protocol independent is used to construct attack paths in NS2.

(b)The initial path chosen by ant and further for going from one node to other is constructed through

(iii) State Transition Rule

$$k = \begin{cases} argmax\left\{\left[\tau_{xy}(t)^{\alpha}\right]\left[\mu_{xy}\ (t)^{\beta}\right]\right\} & if\ m \le m^0 \\ B & otherwise \end{cases}$$

$$G=\begin{cases} \frac{[\tau_{xy}(t)]^{\propto}[\mu_{xy}(t)]^{\beta}}{\sum[\tau_{xy}(t)]^{\propto}[\mu_{xy}(t)]^{\beta}} & k\ \in\ N_i \\ 0 & otherwise \end{cases}$$

Wherem^0distribution ratio and m israndom number having value between 0≤m≤1. $\mu_{xy}(t)$ is figured as number of routing packets in routers x and y in time t-1 and t

(c) Local Updating Rule

On a visit, every ant builds the measure of pheromone on the navigated trail by applying the local updating rule. A localsuboptimalsolution is obtained by preventing ants from selecting same trail.

$$\tau_{xy}(t+1) = (1-\sigma) \times \tau_{xy}(t) + \sigma\Delta\tau_{xy}(t)$$

Where σ is local pheromone decay rate having value in interval [0,1].

(d) Global updating Rule

When every ant have finished their visits in the present cycle, the force of the pheromone appointed to every circular segment of the most probable path is recalculated

$$\tau_{xy}(t+1) = (1-\sigma) \times \tau_{xy}(t) + \sigma\Delta\tau_{xy}(t)$$

$$\Delta\tau_{xy}(t) = \begin{cases} \frac{C}{Lp} & if\ route(i,j)\ is\ the\ optimal\ path \\ 0 & otherwise \end{cases}$$

Where C is constant and Lp is the quantity of nodes along the most likely path.

Table1 Abbrevations

Notations	Meaning
τ	Pheromone intensity
σ	Pheromone decay rate
T	Time interval [0,1]
K	State transition rule
S	
m	Random number
Lp	Number of nodes along most probable path
δ	Euclidean distance
ω, μ	Parameters
m^0	Distribution ratio

4. RESULTS AND DISCUSSION

This segment describes about three vital viewpoints connected with utilizing the ACO group of plans (i.e., AS, ACS, and modified-ACS) to solve the IP traceback:
(a) appropriate selection of the ACO parameters utilized for attack path recreation
(b) factors affecting the convergence speed
(c) relationship between the quantity of packets required to reproduce the attack path for different routing distances
(d) topology scale analysis.

4.1 Parameter Selection

Three control variables must be allocated for executing the ACO models: weighting element of pheromone intensityαweighting element of visibility, β and pheromone decay rateσ. The qualities allotted to these three parameters specifically impact the rate at which the directing parcels join to the genuine attack path of the three parameters, the pheromone power weighting parameter (α) and the Perceivability weighting parameter β .in this manner, an affectability Investigation was performed to distinguish the ideal extents of α and β. A case-control trial was performed in all the ACO models (i.e., AS, ACS, or modified-ACS) by utilizing a system topology involving p=400 nodes and different qualities of α and β in the extent 0.8 – 2.0 and 0.4 – 1.6, respectively. After 500 test runs, the normal number of packets passing through each node in the attack path was evaluated to identify the convergence performance of the solution procedure; both the parameters significantly influenced the convergence execution of the ACO plans. The ideal convergence performance was obtained for α in the interval [0.8, 2.0] and β in the interval [1.5, 1.8]

Table2 Experimental values of α and β

P=400	$\boldsymbol{\alpha}$**=0.8**	$\boldsymbol{\alpha}$**=1.1**	$\boldsymbol{\alpha}$**=1.4**	$\boldsymbol{\alpha = 1.7}$	$\boldsymbol{\alpha = 2.0}$
$\boldsymbol{\beta}$**=0.4**	15.1	15.9	13.1	19.4	24.9
$\boldsymbol{\beta}$**=0.7**	23.2	16.6	15.2	19.7	25.1
$\boldsymbol{\beta = 1.0}$	37.4	23.5	16.9	18.1	24.4
$\boldsymbol{\beta = 1.3}$	90.1	41.9	24.8	20.2	24.1
$\boldsymbol{\beta = 1.6}$	191.3	97.7	47.3	27.4	22.9

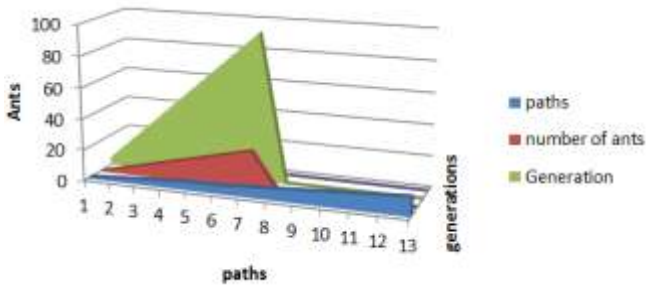

Fig1 Evolution of feasible attack paths in network topology

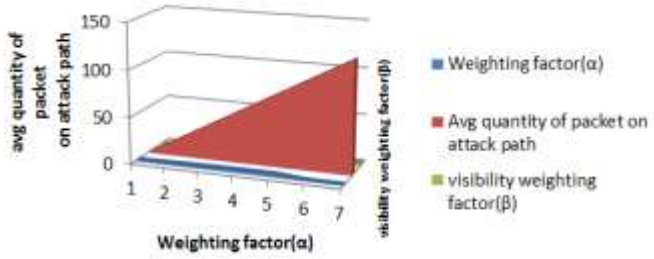

Fig 2. Performance evaluation results for α and β

4.2 Convergence Time Analysis

This segment assesses the convergence performance of the AS, ACS, and modified-ACS schemes for attack paths of different lengths. Records of the joining time of the ACS and modified-ACS calculations in a system comprising p=200 nodesbridges the convergence time of the distinctive plans as a system's element size. All in all, the outcomes exhibited that the proposed modified-ACS plan has a slower convergence rate than does the conventional ACS plan. In addition the convergence time of the modified-ACS plan increments with the quantity of subgroups, that is, with a diminishing number of ants in every subgroup

4.3 Quantity of Packets Required for Attack Path Construction

As a rule, a vast system topology requires more network packets to effectively recreate the attack path. The variation of the coverage percentage of the modified-ACS plan, with the quantity of network packets, as a component of δ. At consistent δ, the convergence percentage expanded with the quantity ofnetwork packets. Besides, for a consistent number of network packets, the convergence percentage diminished with expanding δ. The outcomes ramifications exhibited with respect to the quantity of packets required for attack path reconstruction. Similarly, the quantity of packets required for attack path remaking in unmistakable topology scales was assesse

4.4 Topology Scale Analysis

As the network size topology expands, so does the quest space for the attack path. Hence, the AS's adequacy, ACS, and modified-ACS plans in investigating every plausible path inside of the search space is a key concern. That the convergence percentage for all the three plans decreases as the quantity of nodes in the network topology increments. The AS and ACS plans yield a particularly poor execution as the network size increments. For instance, the AS plan neglects to identify the right attack path in topologies with more than 300 nodes, though the conventional ACS plan comes up short in network containing 500 nodes.

It is seen that, the attack source and victim were picked self-assertively inside of the network topology; that is, they were not obliged to the end nodes. In this way, numerous possible attack paths existed. Thus, in the AS and ACS plots, the ants may overextend (i.e., miss the victim) and meet towards a mistaken attack path. In addition, the worldwide overhauling guideline, May trap the arrangement technique in neighbourhood suboptimal in light of the fact that the short lengths of the ways in a topology with more network nodes results in an excessively fast gathering of pheromones. By complexity, the execution of the modified-ACSplan improves with increasing network density as a result of its utilization of conveyed swarm insight in discovering the genuine attack path

5. CONCLUSION

The results exhibited an ACS-based plan, modified-ACS, for taking care of the IP traceback issue. In the proposed plan, the ant's proficiency to inquire the entire optimum path inside of the arrangement space was upgraded by apportioning the ant colony into subgroups, where every sub gathering applies an alternate pheromone overhauling tene. The utilization of a subgroup strategy lessened the rate at which the pheromone force on the most plausible attack way was upgraded, hence enhancing the worldwide optimality of the last arrangement. The simulation results have affirmed the capacity of the proposed modified-ACS plan in finding the genuine attack way even without the whole directing data or when the assailant's character is camouflaged utilizing a parodied IP address. An affectability examination was performed to research the fundamental's impacts ACS model parameters (pheromone force and perceivability weightings) on the meeting execution of the proposed plan. Trial results of the Hamiltonian routing strategy accept the proficiency of the proposed approach by selecting a proper routing path, nullifying congested region and provides better distribution of network traffic. Moreover, the system's impact size on the quantity of parcels required to develop the attack path and the convergence time of the proposed plan, were examined and contrasted with those of the routine AS and ACS plans. In short, the outcomes showed that in spite of the fact that the convergence time of the modified-ACS plan is marginally slower than that of the conventional ACS scheme, a more solid convergence execution is obtained, especially in systems including various nodes. However, gathering and grouping attack data continuously is a nontrivial assignment. Thus, directing data from different attack occasions must be gathered and solidified in a database ahead of time. This encourages web safeguards in using so as to recognize the attack path using a Modified–ACSin a consequent DDoS attack.

REFERENCES

[1]. Alan Saied, Richard E. Overill, Tomasz Radzik, "*Detection of known and unknown DDoS attacks using Artificial Neural Networks*", Neurocomputing (2015)
[2]. C. Jie-Hao; C. Feng-Jiao, Zhang, "*DDoSdefense system with test and neural network*", in: Proceedings of the IEEE International Conference on Granular Computing (GrC), Hangzhou, China, 11–13 Aug. 2012
[3]. S. Savage, D. Wetherall, A. Karlin, et al., "*Network support for IP traceback"*, IEEE/ACM Trans. Netw. 9 (3) (2001) 226–237
[4]. D.X. Song, A. Perrig, "*Advanced and authenticated marking schemes for IP traceback"*, in: Proceedings of the 20th Conference on Computer Communications (INFOCOM'01), 2, 2001, pp. 878–886
[5]. A.C. Snoeren, C. Partridge, L.A. Sanchez, C.E. Jones, "*Hash-based IP traceback*", in: Proceedings of the Special Interest Group on Data Communication (SIGCOMM), 2001, pp. 27–31.
[6]. I. Corona, G. Giacinto, F. Roli, "*Adversarial attacks against intrusion detection systems*", Taxon. solutions Open Issue Inf. Sci. 239 (2013) 201–225.
[7]. M. Dorigo, L.M. Gambardella, "*Ant colony system*: *A cooperative learning approach to the traveling salesman problem*", IEEE Trans. Evol. Comput. 1 (1) (1997) 53–66
[8]. T. Stuetzle, H. Hoos, "*Improvements on the ant system: Introducing max-min ant system"*, in: Proceedings of the International Conference on Artificial Neural Networks and Genetic Algorithms(ICANNGA'97), SpringerVerlag,Wien, 1997
[9]. M.H. Yang, M.C. Yang, RIHT, "*A novel hybrid IP traceback scheme"*, IEEE Trans. Inf. Forensics Secur. 7 (4) (2012) 789–797.
[10]. M. Muthuprasanna, G. Manimaran, "*Space-time encoding scheme for DDoS attack traceback*", in: Proceedings of the IEEE Global Communications Conference,2005
[11]. D. Basheer, G. Manimaran, " *novel packet marking scheme for IP traceback"*, in: Proceedings of the 10th IEEE International Computer Performance and Dependability Symposium, 2004
[12]. A. Belenky, N. Ansari, "*On IP traceback"* ,IEEECommun. Mag.41 (7) (2003) 142–153.
[13]. B.M. Waxman, "*Routing of multipoint connections"*, IEEE J. Sel. Area Commun. 6 (9) (1988) 1617–1622.
[14]. G. Di Caro, M. Dorigo, "*AntNet: Distributed stigmergetic control for communications networks"*, J. Artif. Intell. Res. 9(1998)317–365.
[15]. P. Wang, H.T. Lin, T.C. Wang, "*A revised ant colony optimization scheme for discovering Attack paths of botnet"*, in: 2011 IEEE International Workshop on Network and System Security, 2011, pp. 918–923.8) 317–365.
[16]. Poona Bahrebar , Dirk Stroobandt , "*The Hamiltonian-based odd–even turn model for maximally adaptive routing in 2D mesh networks-on-chip"*, in: 2015 www.elsevier.com/locate/compeleceng
[17]. A.Colorni, M.Dorigo, V.Maniezzo, "*Distributed optimization by ant colonies"*, in: Proceedings of the European Conference on Artificial Life, Paris, France, 1991, pp.134-142.