Journal of Computer Science 5 (2): 136-139, 2009
ISSN 1549-3636

Three Dimensional Multidirectional Geographical IP Traceback: Direction Ratio Sampling Algorithm

A. Rajiv Kannan, K. Duraiswamy and K. Sangeetha
Department of Computer Science and Engineering,
KS Rangasamy College of Engineering, Tiruchengode, Namakkal, Tamilnadu, India

Abstract: Problem statement: An important and challenging problem is that of tracing DOS/DDOS attack source. IP traceback is the process of identifying the actual source(s) of attack packets, So that the attackers can be held accountable as also in mitigating them, either by isolating the attack sources or by filtering packets for away from the victim. Several IP traceback schemes have been proposed to solve this problem. Among many IP traceback schemes, a recent development was Directed Geographical Traceback (DGT). Though multidirectional two-dimensional DGT schemes were available, in the real scenario, three dimensional, Multidirectional DGT has potential applications. **Approach:** The Direction Ratio Algorithm (DRA) has the limitation of the impossibility of ensuring sufficient unused space in the packet header for the complete Direction Ratio List (DRL) especially when the length of the path is not known apriori. To overcome this, DRSA was proposed. The methods used in DRSA were random sampling methods, where the sufficient numbers of samples were drawn; one can reconstruct the path of the attack packets and trace the attack source. **Results:** In this study those limitation had been overcome using Direction Ratio Sampling Algorithm (DRSA) which works well for 3-dimensional, multi-directional, geographical IP traceback. This approach enables the attack path reconstruction was easily possible and hence a victim can typically reconstruct the path after receiving 75 packets from the attacker. This same algorithm can efficiently discern multiple attacks. When attackers from different sources produce disjoint edges in the tree structure of reconstruction, the number of packets needed to reconstruct each path is independent of other paths. **Conclusion:** DRSA was found to be a robust scheme of attack path reconstruction in Geographical traceback.

Key words: DOS (Distributed Denial of Service), DGT (Directional Geographical traceback), 3DMDGT (Three dimensional, Multi-Directional Geographical traceback), DRA (Direction Ratio Algorithm), DRSA (Direction Ratio Sampling Algorithm)

INTRODUCTION

DOS attacks[3,4] represent a growing threat to the internet infrastructure, by denying regular internet services from being accessed by legitimate users. IP traceback is the process of identifying the actual source(s) of attack packets, So that the attackers can be held accountable as also in mitigating them, either by isolating the attack sources or by filtering packets for away from the victim, Several IP traceback schemes have been proposed to solve this problem.

DGT (Directed Geographical Traceback) scheme exploits the potential of the geographical topology of the internet for traceback. Gao[1] gave a limited two dimensional, 8 directional DGT scheme. This was generalised by[2,5], to 2^n ($n \geq 4$) directions, though only in 2 dimensions.

Considering the spherical/Ellipsoidal topology of the earth, it is clear that the internet path is three dimensional in nature. In this study, 3 dimensional, Multidirectional, Geographical Traceback, through DRSA (Direction Ratio Sampling Algorithm) is proposed.

Normalized coordinates: Taking the geographical topology of the earth (on which all the routers are) either as the sphere:

$$\xi^2+\eta^2+\Im^2 = a^2 \tag{1}$$

or as the ellipsoid £:

$$\xi^2/a^2 + \eta 2/b^2+\Im^2/c^2 = 1 \tag{2}$$

then the transformation:

Corresponding Author: A. Rajiv Kannan, Department of Computer Science and Engineering,
KS Rangasamy College of Engineering, Tiruchengode, Namakkal, Tamilnadu, India

$$ax = \xi , ay = \eta, az = \Im \quad (3)$$

or

$$ax = \xi, by = \eta, cz = \Im \quad (4)$$

makes (1), (2) into the unit sphere:

$$x^2 + y^2 + z^2 = 1 \quad (5)$$

For all the points on note that (5), except for the points (±1,0,0), (0, ±1,0) and (0,0, ±1), we have:

$$\backslash x\backslash, \backslash y\backslash, \backslash z\backslash < 1 \quad (6)$$

satisfying (5). Thus routers R_i are at points (x_i, y_i, z_i) where:

$$x_i^2+y_i^2+z_i^2 = 1 \quad (7)$$

for all i.

We assume that the routers are numbered serially and that the length of any internet path seldom exceeds 32 hops and hence a 10 bit field in the packet header can accommodate the last 3 digits of the router serial number, throughout its journey. All other assumptions regarding attack packets are the same as in[1,2,5,6].

Direction ratios: In three dimensional space, the direction indicators of a line are the direction cosines (d.c) (Cos α, Cos β, Cos r) where α, β, r are the angles which the line makes with the rectangular coordinate axes ox, oy, oz respectively. It can be shown that:

$$Cos^2\alpha+Cos^2\beta+Cos^2r = 1 \quad (8)$$

for any directional cosines (d.c).

Since Cosθ, in general, is a cumbersome fraction/irrational, we use direction ratios (DR) of a line, which are proportional to directional cosines (d.c); denoted by (a, b, c) where:

$$(a, b, c) \in Z \quad (9)$$

and

$$gcd (a, b, c) = 1 \quad (10)$$

(Z is the set of all integers).

Though DR (a, b, c) do not, in general:
Satisfy:

$$a^2+b^2+c^2 = 1 \quad (11)$$

they can be made into directed cosines(d.c) (a/r, b/r, c/r).
Where:

$$r = \sqrt{a^2 + b^2 + c^2} \quad (12)$$

For any router R_o, we can get a neighborhood direction set of DR (a_i, bi, ci) of neighbor routers R_i by taking:

$$|a_i|, |b_i|, |c_i| \in N \quad (13)$$

Satisfying (10). (where, N is the set of natural numbers).

We can show that DR (n), for n Є N, (the number of neighborhood directions from router R_0) satisfy:

$$(2n-1)^3 < DR (n) < (2n+)^3 \quad (14)$$

In fact DR (1) = 13 and DR (2) = 49 and they are shown in Table 1 and Table 2.

Table 1: Elements of DR (1)

i	Elements of DR (1)	i	Elements of DR (1)
1	(1,0,0)	8	(-1,0,1)
2	(0,1,0)	9	(-1,1,0)
3	(0,0,1)	10	(1,1,1)
4	(0,1,0)	11	(-1,1,1)
5	(0,1,1)	12	(1,-1,1)
6	(1,1,0)	13	(1,1,-1)
7	(0,-1,1)		

Table 2: Elements of DR (2)

i	DR (2)	i	DR (2)
1	(1,0,0)	26	(1,1,2)
2	(0,1,0)	27	(1,2,1)
3	(0,0,1)	28	(2,1,1)
4	(0,1,1)	29	(-1,1,2)
5	(1,0,1)	30	(1,-1,2)
6	(1,1,0)	31	(1,1,-2)
7	(0,-1,1)	32	(-1,2,1)
8	(-1,0,1)	33	(1,-2,1)
9	(-1,1,0)	34	(1,2,-1)
10	(1,1,1)	35	(-2,1,1)
11	(-1,1,1)	36	(2,-1,1)
12	(1,-1,1)	37	(2,1,-1)
13	(1,1,-1)	38	(2,2,1)
14	(0,1,2)	39	(2,1,2)
15	(0,2,1)	40	(1,2,2)
16	(0,-1,2)	41	(-2,2,1)
17	(0,-2,1)	42	(2,-2,1)
18	(1,0,2)	43	(2,2,-1)
19	(2,0,1)	44	(-2,1,2)
20	(-1,0,2)	45	(2,-1,2)
21	(-2,0,1)	46	(2,1,-2)
22	(1,2,0)	47	(-1,2,2)
23	(2,1,0)	48	(1,-2,2)
24	(-1,2,0)	49	(1,2,-2)
25	(-2,1,0)		

One-to-one correspondence between DR at a router R_0 and its neighbor routers:

Theorem: Given router R_0 at (x_0, y_0, z_0),and a set of direction ratios DR(n) for some n Є N then, for each ratio $d_i = (a_i, b_i, c_i)$ Є DR(n),there is a unique neighbor router R_i at (x_i,y_i,z_i) on the unit sphere, given by:

$$x_i = x_0+ra_i,\ y_i = y_0+rb_i,\ z_i= z_0+rc_i \qquad (15)$$

Where:

$$r = -\left[\frac{2(a_i x_0 + b_i y_0 + c_i z_0)}{a_i^2 + b_i^2 + c_i^2}\right] \qquad (16)$$

for i = 1,2,..........

Proof: Any point (x, y, z) on the line through router $R_0(x_0, y_0, z_0)$ in the direction d_i with direction ratios(a_i, b_i, c_i) is:

$$x = x_0+ra_i,\ y = y_0+rb_i,\ z = z_0+rc_i \qquad (17)$$

and it is on:

$$x^2+y^2+z^2 = 1 \qquad (18)$$

at router point $R_i(x_i,y_i,z_i)$ if (18) is satisfied.

∴ We get:

$$r = -2((\sum a_i x_0)/(\sum a_i^2)$$

for i = 1,2,.... and this value of r is unique for each i.

Hence there is one-to-one correspondence between elements of DR(n) at R_0 and its neighbor routers.

MATERIALS AND METHODS

This is a theoretical paper on IP traceback problem using geographical information in three dimensions in a multi-directional environment. The materials are a host of Routers R_i at points (x_i,y_i, z_i) for i = 1 to n, on the earth $x^2+y^2+z^2 = 1$. Also the internet attack packets in flight are materials whose flight path is to be reconstructed for mitigating DOS/DDOS attacks.

The methods used in DRSA are random sampling methods, where, after sufficient number of samples are drawn, one can reconstruct the path of the attack packets and trace the attack source.

Direction Ratio Algorithm (DRA): In this algorithm of traceback, for every packet w arriving from the attacker at router R, we appended the DR $d_j = (a_j,b_j,c_j)$ of the next destination in the packet header of w.

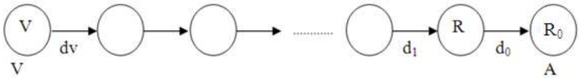
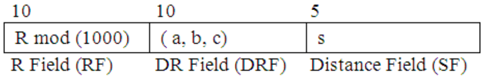


Fig. 1: Flow diagram of DRA



10	10	5
R mod (1000)	(a, b, c)	s
R Field (RF)	DR Field (DRF)	Distance Field (SF)

Fig. 2: IP header format for DRSA

Finally from the suffixes d_0, d_1, d_2...........d_v of w, at the victim router V, we reconstruct the path as in Fig. 1.

This is possible due to the unique (1-1) correspondence between d_j (from any router from R) and its neighbors R_j.

The limitation of this DRA (direction ratio appending algorithm) is the impossibility of ensuring sufficient space in the packet header for appending the DR of every edge of the attack path.

This problem is addressed using DRSA (direction ratio sampling algorithm).

DRSA traceback procedure: We require an address field R, a direction ratio field DR and a distance field S, in the packet header to implement this algorithm.

Assuming that the IP header has (16+8+1) = 25 bits, for DRSA, we can allot 10 bits each. For the address field and DR field and 5 bits for the distance field. This is acceptable since, routers are numbered serially; the 10 digit field can accommodate the last 3 digits of the serial number and is sufficient for R mod (1000). Since a 9 bit field is enough for the 4, 9 direction set of DR (2), 10 bits are sufficient for the DR field. Since any IP path never exceeds 32 hops, a 5 bit distance field is taken at in Fig. 2.

Here R_i is router at (x_i, y_i, z_i) with a given serial number $D_j = (a_j, b_j, c_j)$ = an element of DR (2) indicating the direction ratio of the next router R_j (from R_i). Note that R_i (D_j) = R_j (the router from R_i in the direction D_j is the unique R_j since D_j is in (1-1) correspondence with R_j from a given R_i).

Direction Ratio Sampling Algorithm (DRSA): The marking procedure at a router R_i of every packet w from the attacker is as follows:

Let x be a random number in (0, 1) and p is a chosen probability level[7,8]. If $x<p$, then if the packet is unmarked, then write R_i mod (1000) in RF, D_j in DRF, 0 in SF. Otherwise (if the packet is already marked) or $(x \geq p)$ then only increment the distance field SF.

After sufficient number of samples are drawn, then using the property R_i (D_j) = R_j and the distance field count, the attack path can be reconstructed. The victim uses the DR (along with R) sampled in these packets to create a graph leading back to the source (s) of attack.

RESULTS

After sufficient number of samples are drawn, then using the property R_i $(D_j) = R_j$ and the distance field count, the attack path can be reconstructed. The victim uses the DR (along with R) sampled in these packets to create a graph leading back to the source (s) of attack.

If we constrain p to be identical at each router, then the probability[9] of receiving a marked packet from a router d hops away is p $(1\text{-}p)^{d\text{-}1}$ and this function is monotonic in the distance from the victim. Because the probability of receiving a sample is geometrically smaller, the further away it is from the victim, the time for this algorithm to converge is dominated by the time to receive a sample from the furthest router.

We conservatively assume that samples from all of the d routers (in the path from A toV) appear with the same likelihood as the furthest router. Since these probabilities are disjoint, the probability that a given packet will deliver a sample from some router is at least dp $(1\text{-}p)^{\,d\text{-}1}$ by addition law for disjoint events. As per the well-known Coupon Collector problem[3], the number of trials required to select one of each of d equiprobable items is d (ln(d) + O(1)). Therefore, the number of packets X, required for the victim to reconstruct a path of length d has the bounded expectation:

$$E(x) < \frac{\text{In}(d)}{p(p-1)^{d-1}} \qquad (19)$$

From (19) we can show that E(X) is optimal if p = 1/d ie $dE/dp = 0$, $d^2E/d\,p^2 > 0$ for p = 1/d).

DISCUSSION

For example, if p = 1/d, where d = attack path length, then the victim can typically reconstruct the path after receiving

$E(x) = d^d \ln(d)/(d\text{-}1)^{d\text{-}1}$ packets. For d = 10; $E(x) \le 75$ and hence a victim can typically reconstruct the path after receiving 75 packets from the attacker.

This same algorithm can efficiently discern multiple attacks. When attackers from different sources produce disjoint edges in the tree structure of reconstruction. The number of packets needed to reconstruct each path is independent of other paths.

The limitations imposed by restricting the number of DR to /DR (2)/ = 49 at every stage and using R (mod 1000) instead of the full serial number of router R are marginal in nature. We need more space in the packet header to use elements of DR (3) and the full representation of the R serial number.

CONCLUSION

In conclusion, DRSA is a robust scheme of three dimensional, multi-directional, geographical IP trace back.

REFERENCES

1. Gao, Z. and N. Ansari, 2005. Directed geographical traceback. Proceeding of the 3rd International Conference on Information Technology: Research and Education, June 27-30, IEEE Xplore Press, USA., pp: 221-224. DOI: 10.1109/ITRE.2005.1503108
2. Kannan, A.R. K. Duraiswamy, J. Rajavel, K. Thiyagarajah and V. Suresh, 2008. 16 directional DGT with generalization to 2n(n>4) direction. Int. J. Comput. Sci. Network Secur., 8: 221-225. http://paper.ijcsns.org/07_book/html/200811/200811031.html
3. Savage, S., D. Moore and G.M. Voelker, S. Savage, 2001. Inferring internet denial-of-service activity. Proceeding of the 10th Conference on USENIX Security, Aug. 13-17, USENIX Association Berkeley, CA., USA., pp: 2-2. http://portal.acm.org/citation.cfm?id=1251329
4. Savage, S., D. Wetherall, A. Karlin and T. Anderson, 2000. Practical network support for IP traceback. Comput. Commun. Rev., 30: 295-306. http://portal.acm.org/citation.cfm?id=347560
5. Padmanabhan, V. and L. Subramanian, 2001. Determining the geographic location of internet hosts. Perform. Evaluat. Rev., 29: 324-325. portal.acm.org/citation.cfm?id=384268.378814
6. Padmanabhan, V. and L. Subramanian, 2001. An investigation of geographic mapping techniquesfor internet hosts. Proceeding of the ACMSIGCOMM, Aug. 27-31, San Diego, CA., pp: 173-185. research.microsoft.com/en-us/um/people/padmanab/papers/sigcomm2001.pdf
7. Ferguson, P. and D. Senie, 1998. Network ingress filtering defeating DOS attacks which employ IP source address spoofing. www.faqs.org/ftp/rfc/pdf/rfc2267.txt.pdf
8. Stanford-Chen, S. and L.T. Heberlein, 1995. Holding intruders accountable on the Internet. Proceedings of the Symposium on Security and Privacy, May 8-10, Oakland, CA., pp: 39-49. ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=398921
9. Feller, W., 1966. An Introduction to Probability Theory and its Applications. 3rd Edn., John Wiley and Sons, Icn., pp: 33-122.