ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013

Secure Data Collection using Simple Path Diversity

Geetha V [1], Deepika Singh [2]

Assistant Professor, Information Science and Engineering, RVCE, Bangalore, India [1]
Student – MTECH, Information Science and Engineering, RVCE, Bangalore, India [2]

Abstract: This paper proposes a simple path diversity algorithm (SPDA) for Inter domain routing. The SPDA uses source routing to find an alternate path from a source to a destination, and allows flexible division of traffic over the best and alternate path. The SPDA comprises three steps. First, the BGP protocol is extended to allow the BGP routing table to save the multiple paths for any destination. Second, a given source is able to detect whether the point of congestion along the best path occurs. Third, if congestion occurs, the source will specify an alternate path and direct traffic over the best and alternate path concurrently. Simulation results indicate that the SPDA produces better performance than existing approaches in average end-to-end delay. In this paper, we develop mechanisms that generate randomized multipath routes. Under our designs, the routes taken by the Shares of different packets change over time; so even if the routing algorithm becomes known to the adversary, the adversary still cannot pinpoint the routes traversed by each packet. Besides randomness, the generated routes are also highly dispersive and energy efficient, making them quite capable of circumventing black holes

Keywords: Simple path diversity, Secure Data Collection, Shamirs algorithm, PRP algorithm, NPRP algorithm, MTRP algorithm, SPD algorithm

I. INTRODUCTION

Compromised node and denial of service are two key attacks in wireless sensor networks (WSNs) [1, 2, 3]. In this paper, we study data delivery mechanisms that can with high probability circumvent black holes formed by these attacks. Compromised Node and Denial of Services can both create *Black Holes* which are the areas in which attacker can manipulate or interrupt the packet delivery from source to destination. We argue that classic multipath routing approaches are vulnerable to such attacks, mainly due to their deterministic nature [1]. So once the adversary acquires the routing algorithm, it can compute the same routes known to the source, hence, making all information sent over these routes vulnerable to its attacks. Figure 1 illustrates the conventional way to transmit the packets from source to sink. Packets has been segregated into M secret shares using shamirs algorithm and then propagated towards sink using PRP algorithm.

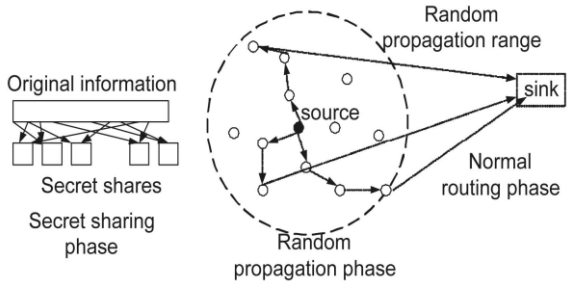

Fig. 1 Conventional way of transmitting data from source to sink

The effect of route dispersiveness on bypassing black holes is illustrated in Fig. 2, where the dotted circles represent the ranges the secret shares can be propagated to in the random propagation phase. A larger dotted circle implies that the resulting routes are geographically more dispersive [7, 8, 9]. Comparing the two cases in Fig. 2, it is clear that the routes of higher dispersiveness are more capable of avoiding the black hole. Clearly, the random propagation phase is the key component that dictates the security and energy performance of the entire mechanism.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013

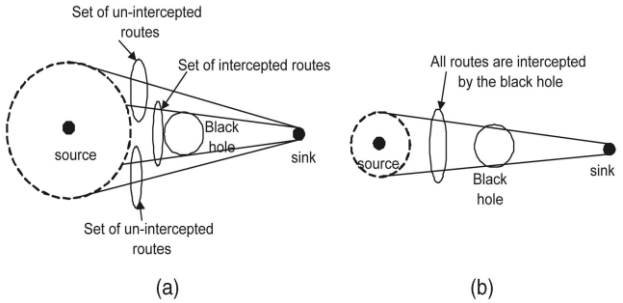

Fig. 2 Implication of route dispersiveness on bypassing the black hole (a) Routes of higher dispersiveness (b) Routes of lower dispersiveness.

This paper, discusses on the problem of constructing. Strong simple path diversity algorithm which along with the conventional algorithms used for transmitting the packets can control the congestion at each node, thus minimizing the denial of services and at the same time it should be dispersive and energy efficient so that it can counter attack the problem of black holes i.e. minimizing the packet loss.

II. EXISTING METHOD

For securing solution current solutions use encryption of data and forwarding the same data in multiple path, but the problem with this solution is compromised nodes can still analyze the data if it know the keys. Also the energy of network will go down rapidly decreasing the life time of the sensor network. Also due to data transferred in multiple path for each data generated, there may be congestion introduced in the network and the packets may be dropped before reaching the base station.

In the classic multipath routing approaches are vulnerable to such attacks, mainly due to their deterministic nature. So once the adversary acquires the routing algorithm, it can compute the same routes known to the source, hence, making all information sent over these routes vulnerable to its attacks. —Compromised node and denial of service are two key attacks in wireless sensor networks.

III. PROPOSED METHOD

To diversify routes, an ideal random propagation algorithm would propagate shares as dispersively as possible [1, 2]. Typically, this means propagating the shares farther from their source. At the same time, it is highly desirable to have an energy-efficient propagation, which calls for limiting the number of randomly propagated hops. The challenge here lies in the random and distributed nature of the propagation: a share may be sent one hop farther from its source in a given step, but may be sent back closer to the source in the next step, wasting both steps from a security standpoint.

A. Shamir's Algorithm

Data to be sent from the source node is split into shares according to shamir's algorithm. The data is split in N total shares in way that only M of the N shares (M < N) is needed to form the original data. This way even if some shares are lost while transmission, if the base station receives minimal shares it can form the original data.

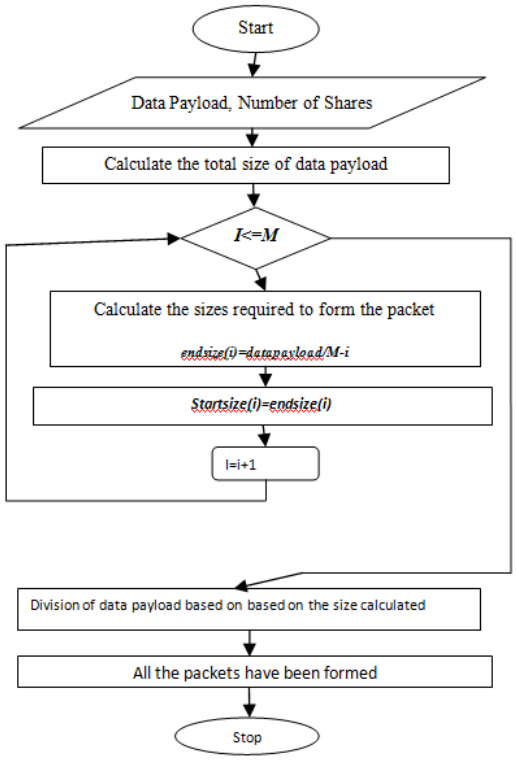

Fig. 3 Shamir's Algorithm

B. Purely Random Propagation Algorithm

In Purely Random Propagation, shares are propagated based on one-hop neighborhood information. More specifically, a sensor node maintains a neighbor list, which contains the ids of all nodes within its transmission range [2, 3]. When a source node wants to send shares to the sink, it includes a TTL of initial value N in each share. It then randomly selects a neighbor for each share, and unicast the share to that neighbor, after receiving the share, the neighbor first decrements the TTL. If the new TTL is greater than 0, the neighbor randomly picks a node from its neighbor list (this node cannot be the source node) and relays the share to it, and so on. When the TTL reaches 0, the final node receiving this share stops the random propagation of this share, and

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013

starts routing it toward the sink using normal min-hop routing.

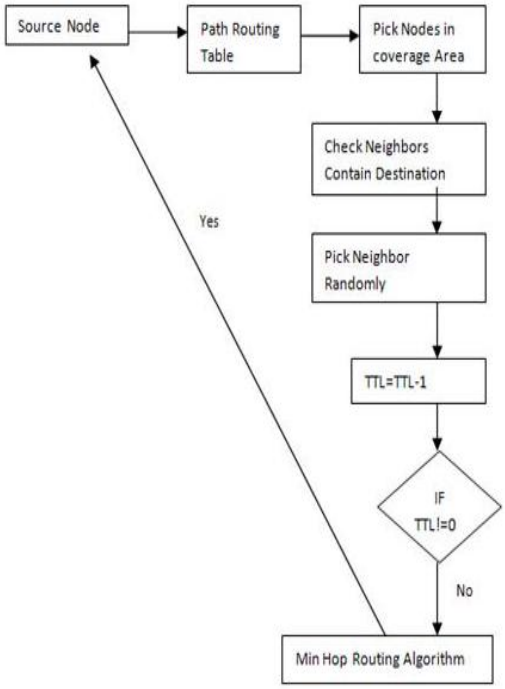

Fig. 4 PRP Algorithm

C. *Non - Repetitive Random Propagation Algorithm*

NRRP is based on PRP, but it improves the propagation efficiency by recording the nodes traversed so far. Specifically, NRRP adds a "node-in-route" (NIR) field to the header of each share. Initially, this field is empty. Starting from the source node, whenever a node propagates the share to the next hop [4, 5], and the id of the upstream node is appended to the NIR field. Nodes included in NIR are excluded from the random pick at the next hop. This non repetitive propagation guarantees that the share will be relayed to a different node in each step of random propagation, leading to better propagation efficiency.

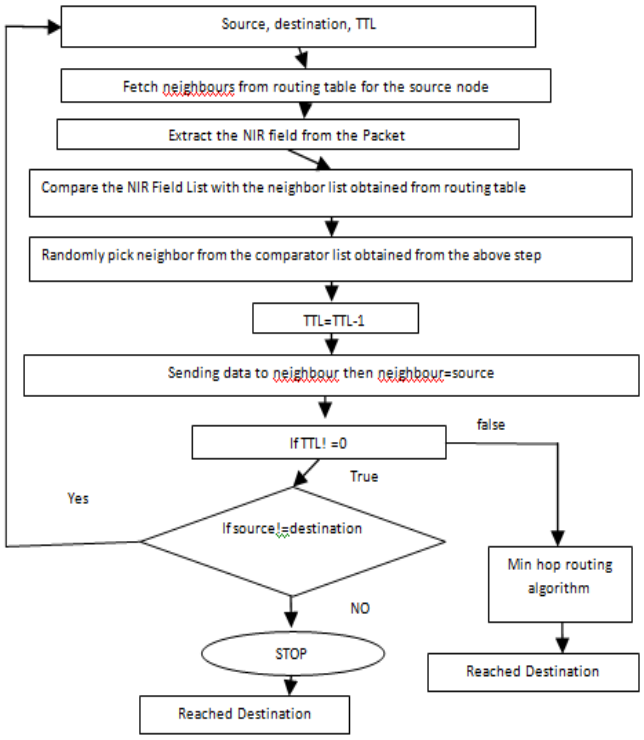

Fig. 5 NRRP Algorithm

D. *Simple Path Diversity Algorithm*

The SPDA uses source routing to find an alternate path from a source to a destination, and allows flexible division of traffic over the best and alternate path. The SPDA comprises three steps. First, the BGP protocol is extended to allow the BGP [6, 7, 8] routing table to save the multiple paths for any destination. Second, a given source is able to detect whether the point of congestion along the best path occurs. Third, if congestion occurs, the source will specify an alternate path and direct traffic over the best and alternate path concurrently. Simulation results indicate that the SPDA produces better performance than existing approaches in average end-to-end delay.

1) Multiple paths in the BGP routing table for any Destination: To save multiple paths in the BGP routing table for any destination, the SPDA extends the functions of BGP. Each BGP routing table saves $pb(s, d)$ and $pa(s, d)$ for any destination d. The $pb(s, d)$ and $pa(s, d)$ in the BGP routing table are indicated by 0 and 1, respectively. Since exchanging the extra path information required for multipath routing consumes extra bandwidth and processing resources , the extended BGP did not allows any node to advertise the saved alternate path along with the best path for a destination. More specifically, a node did not advertise any saved alternate path to its neighboring node. Therefore, the saved alternate path in the BGP routing table is not a second-

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013

best path for a destination. Actually, each node *v* always uses πb(*v*, *d*) to select the next node to forward traffic along the *pb*(*s*, *d*) to destination *d*, and uses πa(*v*, *d*) to select the next node to forward traffic along the *pa*(*s*, *d*) to destination *d*.

2) The point of congestion: Before transmitting IP packets, the source will detect whether the point of congestion along the *pb*(*s*, *d*) occurs. It is possible that there are many points of congestion in the *pb*(*s*, *d*). The key point of SPDA is to detect the first point of congestion in the *pb*(*s*, *d*) and find a *pa*(*s*, *d*) to deflect part of the traffic away from the first point of congestion. The procedures of detection are achieved in three steps.

Step 1: A TCP-based TraceRouteAS is designed and used by the source to measure the round trip time from the source to each intermediate node along the *pb*(*s*, *d*).

Step 2: The source use the information obtained by TraceRouteAS to compute the mean μ and standard deviation σ of round trip time difference between any two neighboring nodes along the *pb*(*s*, *d*). Suppose there are *k* nodes in the *pb*(*s*, *d*) and *di* denotes the round trip time between the source and (*i*+1)*th* node along the *pb*(*s*, *d*).

$$\mu = \frac{\sum_{i=0}^{k}(d_{i+1}-d_i)}{k+1}. \qquad (1)$$

$$\sigma = \sqrt{\frac{\sum_{i=0}^{k}(d_{i+1}-d_i)^2}{k+1} - \mu^2}. \qquad (2)$$

Step 3: For determining the congestion along the path *pb*(*s*, *d*), a threshold value ε is used as defined below:

$$\varepsilon = \mu + \lambda\sigma \qquad (3)$$

Where λ is an adjustment parameter

The first Point of congestion C_{AS} in the route *pb*(*s*, *d*) can be found if and only if the below condition is true.

$$C_{AS} = \min \{ i \mid d_i - d_{i-1} > \varepsilon, i = 1,2, , k+1 \} \quad (4)$$

3) Allocating network traffic over multiple paths: When the first point of congestion C_{AS} occurs along the *pb*(*s*, *d*), the SPDA selects a deflection node prior to the C_{AS} and finds a *pa*(*s*, *d*) to deflect part of the traffic away from the first point of congestion. Since the SPDA still uses the *pb*(*s*, *d*) to forward traffic, the key point is how to reduce the amount of transmitted traffic over the congested *pb*(*s*, *d*), and increase the amount of transmitted traffic over the found *pa*(*s*,*d*).

The SPDA uses two steps to allocate the amount of traffic to individual paths (That is, *pb*(*s*, *d*) and *pa*(*s*, *d*)):

Step 1: The source allocates the amount of traffic to individual paths according to their individual maximum round trip times.

A path with longer duration of round trip means it is carrying more traffic, thus it is advisable to distribute the traffic over each individual path based on their round trip duration.

Consider R_a and Rb as max roundtrip time of path pa(*s*, *d*) and pb(*s*, *d*) respectively. In this case the source will distribute the traffic as shown below:

$$\text{Traffic}_a = Rb / (R_a + R_b) * \text{Total Traffic Mb} \qquad (5)$$

$$\text{Traffic}_b = Ra / (R_a + R_b) * \text{Total Traffic Mb} \qquad (6)$$

Step 2: The source uses the source routing to forward the allocated traffic to each individual path. To use the sourcing routing to forward the allocated amount of IP packets to different paths, some rarely used bits in the existing IP header is used. The options field in the IP header consists of 1-byte code field, a 1-byte length field, and a variable-sized data field. The code field contains three subfields: copy, class, and number. Since two possible values (01 and 11) in the class subfield have not yet been defined, we use 01 and 11 to denote (no congestion) and (congestion) respectively.

If no congestion occurs in the *pb*(*s*, *d*), the value of class subfield carried on the IP packets is set by the source to 01. In this case, each node receiving the IP packets simply forwards the IP packets along the *pb*(*s*, *d*) to the destination. If congestion occurs in the *pb*(*s*, *d*), the value of class subfield carried on the IP packets is set to 11. In this case, the source will set the value of data field carried on the IP packets to a string r = *r0r1r2…rt*, *r i* = 0 or 1, $0 \le i \le t$. The value of *ri* is set to 1 if (*i*+1)*th* node along the *pb*(*s*, *d*) is selected by the SPDA as a deflection node. Otherwise, the value of *ri* is set to 0. Any node will forward the IP packets to the best path if the value of first bit in the string *r* is 0. Similarly, any node will forward the IP packets to the alternate path if the value of first bit in the string *r* is 1.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013

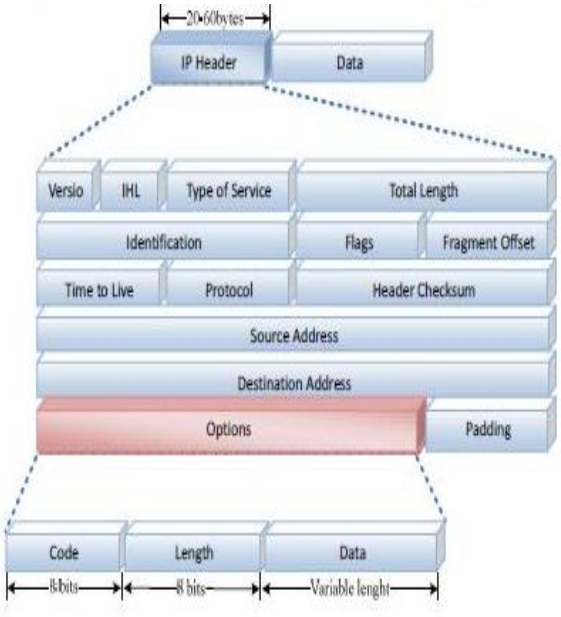

Fig. 6 Options Field in IP Header

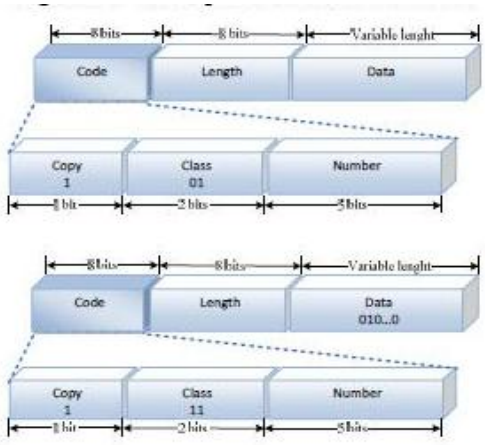

Fig. 7 Class Sub Field Values

IV. CONCLUSION AND FUTURE SCOPE

- **PRP algorithm** propagates the encrypted packets randomly to destination node. Hence the routes are dispersive enough so that the black hole cannot intercept it. The disadvantage of PRP algorithm is that it can propagate the packet back and forth between nodes. Therefore decreasing the propagation efficiency.

- **NRRP propagation** will eliminate the disadvantage of PRP i.e. it eliminates the propagation of packet for the same node from which it arrived. The only disadvantage is as the node propagates towards the destination the NIR field keeps on increasing thereby increasing the overhead.

- In the New Optimized Algorithm is the most performing algorithm as the optimized algorithm does not have any node repetitive in its route and it takes lowest no of hops from source to destination node. The route discovery time is lowest in the new optimized algorithm

- The Congestion based algorithms will find the congestion point and find the alternate path from the point of congestion and we have also network traffic allocation which provides optimized way of routing packets across regular and alternate route.

- The algorithm should also take into the power consideration and energy efficiency in the network and provide an efficient way so that the power consumption is reduced

- The SPDA algorithm proposed in this paper uses source routing to find an alternate path from a source to a destination, and allows flexible division of traffic over the best and alternate path. This has two main advantages. First, the source does not need to maintain some map of the overall network; therefore the proposed algorithm overcomes the scalability problem in source routing. Second, the average end-to-end delay can be reduced by forwarding the packets over the best and alternate path concurrently.

REFERENCES

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Networks," IEEE Comm. Magazine, vol. 40, no. 8, pp. 102-114, Aug. 2002
[2] P.C. Lee, V. Misra, and D. Rubenstein, "Distributed Algorithms for Secure Multipath Routing," Proc. IEEE INFOCOM, pp. 1952-1963, Mar. 2005.
[3] P.C. Lee, V. Misra, and D. Rubenstein, "Distributed Algorithms for Secure Multipath Routing in Attack-Resistant Networks," IEEE/ ACM Trans. Networking, vol. 15, no. 6, pp. 1490-1501, Dec. 2007.
[4] S.J. Lee and M. Gerla, "Split Multipath Routing with Maximally Disjoint Paths in Ad Hoc Networks," Proc. IEEE Int'l Conf. Comm. (ICC), pp. 3201-3205, 2001.
[5] W. Lou, W. Liu, and Y. Fang, "Spread: Enhancing Data Confidentiality in Mobile Ad Hoc Networks," Proc. IEEE INFOCOM, vol. 4, pp. 2404-2413, Mar. 2004.
[6] W. Lou, W. Liu, and Y. Zhang, "Performance Optimization Using Multipath Routing in Mobile Ad Hoc and Wireless Sensor Networks," Proc. Combinatorial Optimization in Comm. Networks, pp. 117-146, 2006.
[7] M. Yannuzzi and et al., "Open Issues in Interdomain Routing: A Survey," IEEE Network, vol. 19, no. 6, pp. 49-56, December 2005
[8] A. Yaar, A. Perrig, and D. Song., "Pi: A path identification mechanism to defend against ddos attacks," In IEEE Symposium on Security and Privacy, 2003.
[9] Feng Wang, Lixin Gao, "Path Diversity Aware Interdomain Routing," IEEE INFOCOM 2009, pp. 307-315.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 5, May 2013

Biography

Deepika Singh is currently pursuing her MTECH in Software Engineering in R V College of Engineering and she is in final semester of her course. Her interests lies in the field of computer networks, network and cyber security and continues her research in these areas.

Geetha V is currently working in RV College of engineering as an assistant professor under the department of information science and engineering. She has published journals on Issues related to QoS routing for Adaptive protocol of MANET" in the month of June 2006 and also have attended various national and international conferences. Her interests lies in the field of Computer Networks.