Defending against Distributed Denial-of-Service (DDoS) Attacks Using Routing Assignments and Resource Allocation Strategies under Quality-of-Service (QoS) Constraints

Dr. Frank Yeong-Sung LIN
Department of Information Management, National Taiwan University
Taipei, 10617, Taiwan R.O.C.

Po-Hao TSANG
Department of Information Management, National Taiwan University
Taipei, 10617, Taiwan R.O.C.

Chen-Bin KUO
Department of Information Management, National Taiwan University
Taipei, 10617, Taiwan R.O.C.

ABSTRACT

As networks become more popular, network attack events are occurring more frequently, especially Distributed Denial-of-Service (DDoS) attacks. To counter such attacks and achieve the objective of “continuity of service”, it is essential that networks be well designed with spare resource allocation capacity so that acceptable Quality-of-Service (QoS) levels can be maintained. In this paper, we address the problem of defending against intelligent DDoS attacks by using routing assignments and resource allocation strategies under QoS constraints. The problem is analyzed as a mixed, nonlinear, integer programming optimization problem with a max-min format. The solution approach is based on the Lagrangean Relaxation and subgradient method, which solves this complicated problem effectively. The efficiency and effectiveness of the proposed heuristic are evaluated by computational experiments.

Keyword: Distributed Denial-of-Service (DDoS), Lagrangean Relaxation, Network Attack and Defense, Optimization Problem, Resource Allocation, Survivability

1. INTRODUCTION

The increasing popularity and utility of the Internet raise the issue of defending networks against Distributed Denial of Service (DDoS) attacks [1]. When such attacks occur, a network suffers performance degradation and a waste of resources, or – in a worst case scenario – it may not survive. Thus, how to defend against DDoS attacks and improve the effectiveness of defense mechanisms are critical issues [2, 3]. A number of metrics can be used to evaluate the survivability of a network under attack [4, 5]. To measure survivability under DDoS attacks, we have to evaluate what level of such attacks a network can sustain, since the network operator needs to maintain the Quality of Service (QoS) of the network under DDoS attacks. In this paper, we propose a model that evaluates the survivability of a network under DDoS attack in terms of performance metrics.

A survivable overlay network [4] can resist DDoS attacks by rewiring the architecture and maximizing the end-to-end connectivity between clients and servers. This kind of defense mechanism can be viewed as the final line of defense for the victim network. Therefore, rather than deploy defense mechanisms at the source end and on intermediate routers, we propose a network planning and spare resource allocation method [6] that focuses on the victim end as the final line of defense. Our objective is to design a survivable network that can sustain abnormal traffic when other defense mechanisms cannot work properly. To this end, we construct a mathematical model of potential attack and defense scenarios and then quantitatively analyze the model. Consequently, the budget required to ensure survivability as well as potential losses due to attacks can be estimated more accurately.

To simulate the characteristics of a real network, we adopt the concept of self-similarity [7]. The nature of the self-similarity of network traffic, which is measured by the Hurst parameter (H), has been well studied; however, few works have considered the phenomenon under attack situations. If it is $0.5 < H \leq 1$, we say the traffic is self-similar. In addition, the mixed normal and abnormal traffic is self-similar [8, 9]. In our research, network self-similarity and DDoS attacks are considered jointly, since attacks can be detected based on the nature of DDoS attacks, which influences the network self-similarity of traffic and causes the Hurst parameter value to deviate from normal [3].

We propose a max-min mathematical model to describe the routing assignment and resource allocation strategies of network administrators and the DDoS attack strategies of attackers. After solving the problem optimally, we could provide guidelines for network administrators to block abnormal traffic produced by DDoS attacks. Previous research has shown that attempts to model attack-defense scenarios in an abstract, mathematical way are non-trivial [10].

The remainder of this paper is organized as follows. In Section 2, we formulate the primal AFRB (Attack Flow & Routing assignment and Budgeting allocation strategy) and RB problems. In Section 3, we propose a Lagrangean Relaxation-based solution approach to the problem. The results of computational experiments conducted to evaluate the proposed solution are reported in Section 4. Then, in Section 5, we present our conclusions and indicate possible directions of future research.

2. PROBLEM FORMULATION

Problem description and assumptions

The question we address is: How can a network administrator operating in the Autonomous System (AS) defend against DDoS attacks by using different routing assignment and budget allocation strategies. Proper routing assignment will prevent an excessive traffic load on one communication link, while budget allocation strategies consider the defense needs of network components in order to maintain the communication quality. An attacker outside the target network will try to exhaust the target network's resources through an effective DDoS attack strategy, i.e., by choosing a destination node, the entry nodes for launching the attack, and the volume of attack traffic (i.e., the number of packets).

We model this scenario as a max-min problem. The inner problem, defined as the RB model, represents that, for a given DDoS attack strategy, a network administrator uses routing assignment and budget allocation strategies to minimize the total defense budget under QoS constraints for each Origin-Destination (O-D) pair. The outer problem, defined as the AFRB model, represents that, to maximize the minimized total defense budget, the attacker must determine the volume of the attack flow sent to the designated destination node from specific entry nodes.

For convenience of modeling, we assume that each entry node is connected by two dummy nodes that also belong to the AS. One represents the source of attack traffic, and the other represents the source of normal traffic. Besides considering physical directed links, we use the node splitting technique to generate a virtual link for each node.

Problem formulation for the AFRB model

We now define the notations used in this paper and formulate the problem.

Table 1. Given Parameters

Notation	Description
N	The index set of all nodes in the Autonomous System (AS)
L	The set of directed communication links, $L=L_1\cup L_2$
L_1	The set of directed communication links, each of which links two nodes
L_2	The set of virtual links between two split nodes for all nodes in the AS
W	The set of all Origin-Destination (O-D) pairs
W_{att}	The set of O-D pairs in which all the original nodes are attack source nodes, where $W_{att}\subset W$
P_w	The index set of all candidate paths for an O-D pair w, where $w\in W$
δ_{pl}	The indicator function, which is 1 if link l is on path p; and 0 otherwise (where $l\in L$, $p\in P_w$)
B_l	All possible values of b_l allocated to link l, where $l\in L$
γ_{att}	Total abnormal traffic produced by an attacker
β_w	The traffic requirement (packets/sec) for O-D pair w, where $w\in W$- W_{att}
D_w	The maximum allowable end-to-end delay for O-D pair w, where $w\in W$
H_w	The Hurst parameter used to measure the self-similarity of the traffic for O-D pair w, where $w\in W$
H_{LB}	The Hurst parameter, which is a lower bound, to denote the self-similarity of a link

Table 2. Decision Variables

Notation	Description
γ_w	Abnormal traffic sent from an attack source to a designated destination by the attacker, where $w\in W_{att}$
b_l	The budget allocated to protect link l, where $l\in L$
g_l	The aggregate traffic flow on link l, where $l\in L$
c_l	The capacity (packets/sec) of each link l, where $l\in L$, which is equal to function $\hat{c}_l(b_l)$
H_l	The Hurst parameter used to measure the self-similarity of the aggregate flow on link l, where $l\in L$ (the aggregate flow consists of independent traffic sources)
d_l	The mean traffic delay of each link l, where $l\in L$, which is equal to function $\hat{d}_l(c_l, g_l, H_l)$
x_p	1 if path p is chosen to transmit packets for O-D pair w, and 0 otherwise (where $p\in P_w$, $w\in W$)
t_{wl}	1 if l is used by an O-D pair w, and 0 otherwise (where $l\in L$, $w\in W$)

Objective function:

$$Z_{\mathrm{IP1}} = \max_{\gamma_w} \min_{b_l, x_p} \left[\sum_{l\in L} b_l \right], \qquad \text{(IP 1)}$$

subject to:

$$b_l \in B_l \qquad \forall l\in L \qquad (1\text{-}1)$$

$$\gamma_{att} = \sum_{w\in W_{att}} \gamma_w \qquad (1\text{-}2)$$

$$\gamma_w \ge 0 \qquad \forall w\in W_{att} \qquad (1\text{-}3)$$

$$\sum_{w\in W-W_{att}} \sum_{p\in P_w} x_p \delta_{pl} \beta_w + \sum_{w\in W_{att}} \sum_{p\in P_w} x_p \delta_{pl} \gamma_w = g_l \qquad \forall l\in L \qquad (1\text{-}4)$$

$$0 \le g_l \le c_l = \hat{c}_l(b_l) \qquad \forall l\in L \qquad (1\text{-}5)$$

$$\sum_{p\in P_w} x_p \delta_{pl} H_w \le H_l \qquad \forall w\in W,\ l\in L \qquad (1\text{-}6)$$

$$H_l \in \left\{ H_{LB},\ \sum_{p\in P_w} x_p \delta_{pl} H_w \right\} \qquad \forall w\in W,\ l\in L \qquad (1\text{-}7)$$

$$\sum_{l\in L} d_l \sum_{p\in P_w} x_p \delta_{pl} \le D_w \qquad \forall w\in W \qquad (1\text{-}8)$$

$$\sum_{p\in P_w} x_p \delta_{pl} = t_{wl} \qquad \forall w\in W,\ l\in L \qquad (1\text{-}9)$$

$$\sum_{p\in P_w} x_p = 1 \qquad \forall w\in W \qquad (1\text{-}10)$$

$$x_p = 0 \text{ or } 1 \qquad \forall p\in P_w,\ w\in W \qquad (1\text{-}11)$$

$$t_{wl} = 0 \text{ or } 1 \qquad \forall w\in W,\ l\in L. \qquad (1\text{-}12)$$

The objective function is to maximize the minimized total defense budget. Constraint (1-1) indicates that the budget allocated to a network component belongs to a set containing all possible values. Constraint (1-2) requires that the total amount of abnormal traffic equals a given value. Constraint (1-4) calculates the aggregate traffic flow on link *l*, including normal and abnormal traffic. Constraint (1-5) stipulates that the aggregate traffic flow on link *l* must not exceed the capacity, which is a function of b_l. Constraint (1-6) requires that the Hurst parameter value of the aggregate traffic flow on link *l* is no smaller than the sum of the Hurst parameter values of independent traffic sources. Constraint (1-7) denotes that the Hurst parameter value of the aggregate traffic flow on link *l* belongs to a set, which comprises the sum of the Hurst parameter values of independent traffic sources and the lower bound value of the Hurst parameter. Constraint (1-8) requires that the transmission delay of each O-D pair must not exceed the maximum allowable end-to-end delay QoS requirement. Constraint (1.9) is the relation between t_{wl}, x_p, and δ_{pl}. To simplify the problem, we replace the sum of all $x_p\delta_{pl}$ with the auxiliary set of decision variables, t_{wl}.

Problem formulation for the RB model

To solve the primal problem, we analyze the RB model first. The abnormal traffic, γ_w, becomes a given parameter in the RB model. Furthermore, we can combine the parameters γ_w and β_w to form one parameter α_w, which denotes the traffic of O-D pair *w*.

Objective function:

$$Z_{\mathrm{IP2}} = \min_{b_l, x_p} \left[\sum_{l\in L} b_l \right], \tag{IP 2}$$

subject to:

$$b_l \in B_l \qquad \forall l \in L \tag{2-1}$$

$$\sum_{w\in W}\sum_{p\in P_w} x_p \delta_{pl} \alpha_w = g_l \qquad \forall l \in L \tag{2-2}$$

$$0 \le g_l \le c_l = \hat{c}_l(b_l) \qquad \forall l \in L \tag{2-3}$$

$$\sum_{p\in P_w} x_p \delta_{pl} H_w \le H_l \qquad \forall w \in W,\ l \in L \tag{2-4}$$

$$H_l \in \{H_{LB},\ t_{wl}H_w\} \qquad \forall w \in W,\ l \in L \tag{2-5}$$

$$\sum_{l\in L} d_l t_{wl} \le D_w \qquad \forall w \in W \tag{2-6}$$

$$\sum_{p\in P_w} x_p \delta_{pl} = t_{wl} \qquad \forall w \in W,\ l \in L \tag{2-7}$$

$$\sum_{p\in P_w} x_p = 1 \qquad \forall w \in W \tag{2-8}$$

$$x_p = 0 \text{ or } 1 \qquad \forall p \in P_w, w \in W \tag{2-9}$$

$$t_{wl} = 0 \text{ or } 1 \qquad \forall w \in W,\ l \in L. \tag{2-10}$$

Solution approach for the RB model

We can relax the equality of Constraints (2-2) and (2-7) as $\sum_{w\in W}\sum_{p\in P_w} x_p\delta_{pl}\alpha_w \le g_l$ and $\sum_{p\in P_w} x_p\delta_{pl} \le t_{wl}$ respectively, without affecting the optimality conditions. By applying the Lagrangean relaxation method [11] with a vector of Lagrangean multipliers, we can transform the problem (IP 2) into the following Lagrangean relaxation problem (LR 1), where constraints (2-2), (2-4), (2-6), and (2-7) are relaxed.

$$\begin{aligned} Z_{\mathrm{D}}(\mu^1, \mu^2, \mu^3, \mu^4) = \min_{b_l, x_p} & \left[\sum_{l\in L} b_l \right] \\ &+ \sum_{l\in L} \mu_l^1 \left[\sum_{w\in W}\sum_{p\in P_w} x_p \delta_{pl} \alpha_w - g_l \right] \\ &+ \sum_{w\in W}\sum_{l\in L} \mu_{wl}^2 \left[\sum_{p\in P_w} x_p \delta_{pl} H_w - H_l \right] , \\ &+ \sum_{w\in W} \mu_w^3 \left[\sum_{l\in L} \hat{d}_l(c_l, g_l, H_l) t_{wl} - D_w \right] \\ &+ \sum_{w\in W}\sum_{l\in L} \mu_{wl}^4 \left[\sum_{p\in P_w} x_p \delta_{pl} - t_{wl} \right] \end{aligned} \tag{LR 1}$$

subject to:

$$b_l \in B_l \qquad \forall l \in L \tag{3-1}$$

$$0 \le g_l \le c_l = \hat{c}_l(b_l) \qquad \forall l \in L \tag{3-2}$$

$$H_l \in \{H_{LB},\ t_{wl}H_w\} \qquad \forall w \in W,\ l \in L \tag{3-3}$$

$$\sum_{p\in P_w} x_p = 1 \qquad \forall w \in W \tag{3-4}$$

$$x_p = 0 \text{ or } 1 \qquad \forall p \in P_w, w \in W \tag{3-5}$$

$$t_{wl} = 0 \text{ or } 1 \qquad \forall w \in W,\ l \in L. \tag{3-6}$$

By definition, μ^1, μ^2, μ^3, and μ^4 are the respective vectors of $\{\mu_l^1\}$, $\{\mu_{wl}^2\}$, $\{\mu_w^3\}$, and $\{\mu_{wl}^4\}$, the elements of which are all non-negative. To solve (LR 1) optimally, we decompose it into two independent and easily solvable subproblems as follows.

Subproblem 1 (related to decision variable x_p)

$$\min \left(\sum_{w\in W}\sum_{p\in P_w}\sum_{l\in L} x_p \delta_{pl} \left[\mu_l^1 \alpha_w + \mu_{wl}^2 H_w + \mu_{wl}^4 \right] \right), \tag{SUB 1}$$

subject to (3-4) and (3-5).

(SUB 1) can be further decomposed into a series of |*W*| independent minimum cost path subproblems. In other words, for each OD pair, we can determine the value of x_p individually. Because of the non-negativity constraint of each parameter ($\mu_l^1\alpha_w + \mu_{wl}^2 H_w + \mu_{wl}^4$), which can be treated as the cost of link *l* in OD pair *w* in the minimum cost path subproblems, we can apply Dijkstra's shortest path algorithm to solve the subproblems optimally. The time complexity of (SUB 1) is $O(|W| \times |N|^2)$.

Subproblem 2 (related to decision variables b_l, g_l, H_l, t_{wl})

$$\min \sum_{l\in L} \begin{bmatrix} b_l + \left(-\mu_l^1\right) g_l + \sum_{w\in W} \left(-\mu_{wl}^2\right) H_l \\ + \sum_{w\in W} \left(\hat{d}_l(c_l, g_l, H_l)\mu_w^3 - \mu_{wl}^4 \right) t_{wl} \end{bmatrix}, \tag{SUB 2}$$

subject to (3-1), (3-2), (3-3) and (3-6).

(SUB 2) can further be decomposed into |*L*| independent subproblems. To solve each of these subproblems, Constraint (3-3) is first relaxed to H ∈ $\{H_{LB}, H_w\}$. We assume that the queuing delay function is $\frac{\delta}{c_l(1-\delta)}$, where δ is a function of the Hurst parameter and utilization [12], and then exhaustively assign the values of b_l and H_l to the objective function. For each pair of (b_l, H_l), the objective function of (SUB 2) takes the following format with proper new constraints.

$$\min \left[\left(-\mu_l^1\right) g_l + \sum_{w\in W} \left(\mu_w^3 \frac{\delta}{c_l(1-\delta)} - \mu_{wl}^4 \right) t_{wl} \right]$$

We can now focus on solving g_l and t_{wl}; a similar problem was solved in [13]. The solution algorithm obtains a local minimal objective function value as follows.

Step 1. Solve $\mu_w^3 \frac{\delta}{c_l(1-\delta)} - \mu_{wl}^4 = 0$ for each O-D pair w and call them the break points of g_l.
Step2. Sort the break points and drop infeasible values; feasible regions are defined in Constraint (3-2) and denoted as $g_l^1, g_l^2, \ldots, g_l^n$.
Step3. At each interval $g_l^i \le g_l \le g_l^{i+1}$, t_{wl} is 1 if $\mu_w^3 \frac{\delta}{c_l(1-\delta)} - \mu_{wl}^4 \le 0$, and 0 otherwise.
Step4. Within the interval $g_l^i \le g_l \le g_l^{i+1}$, the local minimal objective function value is either at a boundary point g_l^i or g_l^{i+1}, or at g_l^*, where $\begin{cases} f(g_l^*) \le f(g_l) \\ f(g_l) = \left[-\mu_l^1 g_l + e_l \frac{\delta}{c_l(1-\delta)} \right], \\ e_l = \sum_{w \in W} \mu_w^3 t_{wl}. \end{cases}$ To simplify finding the solution of g_l^*, we assume that the utilization is discrete and search the local optimal solution by increasing the utilization value by 0.001 at each iteration

The global minimum objective function value of (SUB 2) can be found by comparing the local minimum objective function values obtained by the above steps. The time complexity of (SUB 2) is O(|L|×|B_l|×|W|^2×log|W|).

According to the weak Lagrangean duality theorem [11], the optimal value of the problem (LR 1) is, by its nature, the lower bound (for minimization problems) of the objective function's value in the primal problem (IP 2). The tightest Lagrangean lower bound can be derived by tuning the Lagrangean multipliers, i.e., by maximizing the (LR 1) problem. Although there are several ways to solve this problem, the subgradient optimization technique [11] is the most popular.

Getting primal feasible solutions

The solutions to (LR 1) and the multipliers provide some hints about deriving a heuristic to improve the solution quality of (IP 2). Our proposed heuristic is described below.

Step 1. Obtain information from (LR 1) as follows. ● Adopt μ_{wl}^4 as a priority for each O-D pair. ● Assign a candidate routing path x_p for each O-D pair. ● Mark each O-D pair with its priority in a Waiting Queue. ● Construct an empty Candidate Queue in which all the O-D pairs can transmit packets later. ● Define two variables *Max_Searching_Limit* and *Searching_Counter*, whose initial values are =|W|^2/4 and 0 respectively.
Repeat Steps 2 to 4 until either a feasible solution is found or no feasible solution is found after several iterations.
Step 2. Select the first unexamined O-D pair in the Waiting Queue to perform the Path Checking Process described below.
Step 3. Perform the Candidate Queue Checking Process described below.
Step 4. Perform the Searching Limit Checking Process described below.

Path Checking Process

Step 1. Check whether the current candidate routing path of the O-D pair is feasible. If it is, add the O-D pair to the Candidate Queue and terminate this process; otherwise, go to Step 2.
Step 2. Find a minimum end-to-end delay routing path for the O-D pair.
Step 3. Assign a budget to the path and satisfy the capacity constraints. Whether the path is feasible or not, add the O-D pair to the Candidate Queue.

Candidate Queue Checking Process

Step 1. Simulate a scenario where all the O-D pairs in the Candidate Queue are transmitting packets. Reroute each O-D pair to obtain a minimum end-to-end delay routing path.
Step 2. Check the end-to-end delay constraints. If all the O-D pairs in the Candidate Queue are feasible, go to Step 5; otherwise go to Step 3.
Step 3. For each O-D pair with infeasible candidate paths, calculate the value gained by adding one more unit of the budget to each link. The gain function is defined as follows: *gain*=$d_l(b_l)$-$d_l(b_l$+1), for each link l on a candidate path.
Step 4. Find the maximum value gained in Step 3 by adding one more unit of the budget to each link. Repeat Steps 3 and 4 until the candidate path satisfies the end-to-end delay constraints. If all links on the candidate path reach the maximum budget limitation and the candidate path is still infeasible, add the O-D pair to the Waiting Queue and increase the value of the *Searching_Counter* by 1.
Step 5. If the Waiting Queue is empty, terminate the heuristic (because a feasible solution for all O-D pairs has been found).

Searching Limit Checking Process

Step 1. If *Searching_Counter* < *Max_Searching_Limit*, go to the next step; otherwise terminate this process.
Step 2. If all the links in the network reach the maximum budget limitation, terminate the heuristic (because no feasible solution can be found); otherwise, go to next step.
Step 3. Set all the links in the network to their maximum budget. Select the first O-D pair in the Candidate Queue and find a minimum end-to-end delay routing path for it; then put it in the Waiting Queue until the Candidate Queue is empty, and double the value of the *Max_Searching_Limit*.

Solution approach for the AFRB model

The objective of the AFRB model is to maximize the total defense budget by adjusting the abnormal traffic γ_w. Recall that the attacker determines the destination node for an attack, the entry nodes for sending the attack traffic, and the volume of the attack flow. Our proposed heuristic for getting primal feasible solutions to (IP 1) is based on the attack flow adjustment process (described below) for the routing assignment and budget allocation strategy decided by the network administrator.

Attack Flow Adjustment Process

Initialization: Obtain the information about the routing assignment and budget allocation strategy from the RB model.
Step 1. Adopt μ_l^1 as an arc weight to evaluate the importance of each routing path.
Step 2. Extract one unit of attack flow from the unexamined routing path with the lowest weight and add it to the routing path with highest weight.
Step 3. Calculate new total defense budget.
Step 4. Repeat steps 2 and 3 until the total defense budget is maximized.

3. COMPUTATIONAL EXPERIMENTS

Experiment environments

We choose three popular network topologies, i.e., grid, random, and mesh, for our experiments. The capacity of a link is a function of the budget and has a convex form. The maximum allowable end-to-end delay is set to 600ms and 900ms, respectively, in and across the AS. The basic normal traffic requirements in and across the AS are set to 2 and 4 packets per second respectively; and the Hurst parameters (H) of the internal normal traffic, external normal traffic, and attack flow are set to 0.7, 0.75, and 0.85, respectively. We compare the total defense budget of two simple algorithms, SA_1 and SA_2, (described below) with that of our proposed heuristic.

Simple Algorithm 1

Step 1. Find a minimum end-to-end delay routing path for the O-D pair.
Step 2. Allocate a budget to the path and ensure the capacity and end-to-end QoS constraints are satisfied.
Step 3. If any infeasible candidate path exists, go to the next step; otherwise, terminate the algorithm (because a feasible solution for all O-D pairs has been found).
Step 4. For each O-D pair with an infeasible candidate path, repeat Step1. If all the links in the network reach the maximum budget limitation, terminate the algorithm (because no feasible solution can be found).

Simple Algorithm 2

Step 1. Adopt the aggregated flow on links as arc weights and run a shortest path algorithm to find a routing path for each O-D pair.
Step 2. Allocate a budget to the path and ensure the capacity and end-to-end QoS constraints are satisfied.
Step 3. If any infeasible candidate path exists, go to the next step; otherwise, terminate this algorithm (because a feasible solution for all O-D pairs has been found).
Step 4. For each O-D pair with an infeasible candidate path, repeat Step1. If all the links in the network reach the maximum budget limitation, terminate the algorithm (because no feasible solution can be found).

The LR value represents the primal feasible solution derived by our proposed heuristic, while the LB value indicates the lower bound determined by the LR process. The duality gap is calculated by $\frac{LB-LR}{LR}\times 100\%$ [11].

Experiment results for the RB model

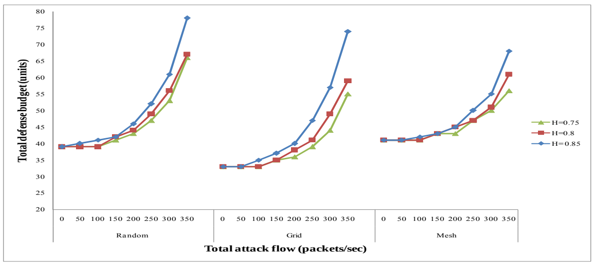

Figure 1. Total defense budget under different H values for different attack flows

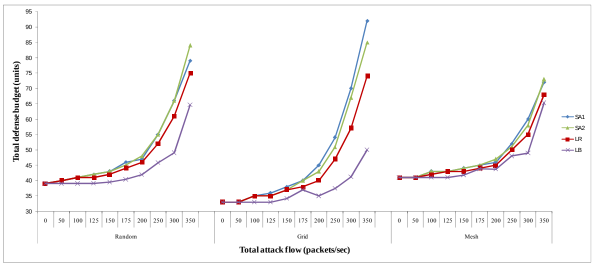

Figure 2. Total defense budget under different attack flows in different network topologies

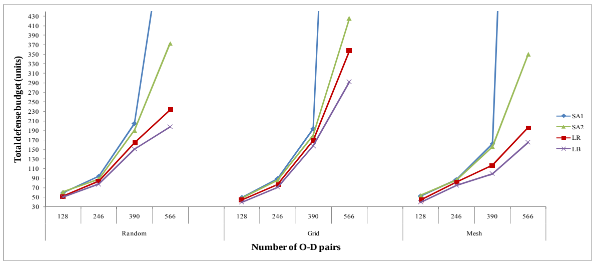

Figure 3. Total defense budget in different network topologies and scales

In Figure 1, we observe that the total defense budget increases continuously as the total attack flow increases. The total defense budget also increases rapidly when the Hurst parameter value of the attack flow is set to 0.85.

The graphs in Figures 2 and 3 show that the proposed heuristic outperforms SA_1 and $SA_{2.}$ Moreover, it improves the ratio of LR to SA_1 and SA_2 as we increase the total attack flows. The efficacy of the LR-based algorithm's solution is clearly demonstrated as the size of the network increases. We also observe that the algorithm performs better in random and grid networks than in a mesh network. The reason might be that, in a mesh network, an O-D pair has more candidate paths for transmitting packets, which would allow SA_1 and SA_2 to find a good routing path for each O-D pair.

Experiment results for the AFRB model

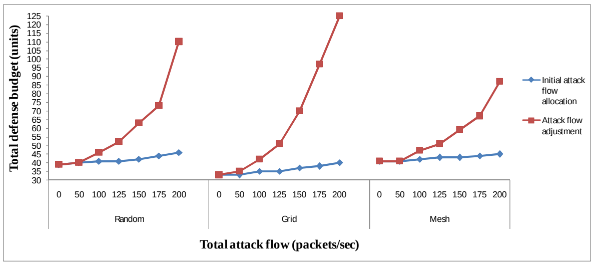

Figure 4. Total defense budget before and after the attack flow adjustment process

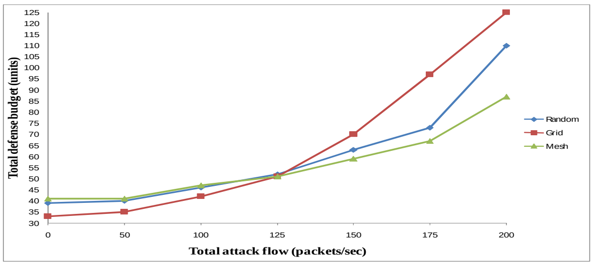

Figure 5. Total defense budget under different network topologies

In Figures 4 and 5, we observe that, after the attack flow adjustment process, the total defense budget increases dramatically when the total attack flow exceeds a threshold. The figures also show that a network's topology strongly influences its total defense budget. Grid networks cannot handle as much attack flow volume as random and mesh networks under the same QoS requirements and maximum budget limitation on each link. The reason is that, compared to random and mesh networks, the network administrator in a grid network has fewer candidate paths and links from which to choose a suitable path for each O-D pair

4. CONCLUSION

Although commercial network security products and mechanisms are being developed constantly, it is still hard to defend against DDoS attacks completely. At the same time, DDoS attacks with higher network self-similarity than normal network traffic consume more network resources, so it is hard to satisfy QoS requirements.

In this paper, we have focused on two issues. First, to improve the security of a network, we have proposed a mathematical model to formulate DDoS attack-defense scenarios and provide defenders with effective routing assignment and resource allocation strategies. Second, we have considered network survivability and evaluated the maximal minimized total defense budget in different scenarios. Furthermore, our mathematical model considers network self-similarity. We first capture the aggregate characteristics of self-similar traffic and then generate DDoS attack flows with higher Hurst parameter values.

From our experiments, we conclude that the total defense budget increases continuously as the total attack flow increases, and it increases rapidly when the Hurst parameter value of the attack flow is set to 0.85. We also note that, after the attack flow adjustment process, the total defense budget increases dramatically when the total attack flow exceeds a threshold. Our experiment results demonstrate that grid networks cannot sustain as much attack flow volume as random and mesh networks under the same QoS requirements and maximum budget limitation on each link.

The mathematical model represents the major contribution of this work. We have carefully researched the security problem's characteristics, identified its objectives and associated constraints, and proposed a well-formulated mathematical model to solve it. To the best of our knowledge, the proposed approach is one of the few that model DDoS attack-defense scenarios as mathematical programming problems in the context of survivability. In addition, we have provided solution approaches to determine the total defense budget. The proposed approach is not only very effective, it is also adaptable to different attack/defense scenarios.

We believe that the proposed model can be extended to different attack-defense scenarios in the context of survivability. In a future work, we will investigate the extent to which our methods can be applied to several ASs to demonstrate the scalability of our model and the concept of collaborative defense. We will also examine the features of the detection and filtering mechanisms for DDoS attacks, and take the end-to-end delay jitter constraint into consideration.

REFERENCE

[1] J. Mirkovic and P. Reiher, "A Taxonomy of DDoS Attack and DDoS Defense Mechanisms", *ACM SIGCOMM Computer Communications Review*, Vol. 34, No. 2, pp.39-53, April 2004.

[2] D. K. Y. Yau, J. C. S. Lui, F. Liang, and Y. Yam, "Defending Against Distributed Denial-of-Service Attacks with Max-Min Fair Server-Centric Router Throttles", *IEEE/ACM Transactions on Networking*, Vol. 13, No. 1, pp. 29-42, February 2005.

[3] Y. Xiang, Y. Lin, W. L. Lei, and S. J. Huang, "Detecting DDoS attack based on network self-similarity", *Proceedings of the IEE Communications*, Vol. 151, No. 3, June 2004.

[4] T. Bu, S. Norden, and T. Woo, "A survivable DoS-Resistant Overlay Network", *Computer Networks*, Vol. 50, pp. 1281-1301, 2006.

[5] D. M. Nicol, W. H. Sanders, and K. S. Trivedi, "Model-Based Evaluation: From Dependability to Security", *IEEE Transactions on Dependable and Secure Computing*, Vol. 1, No. 1, pp. 48-65, January-March 2004.

[6] Z. Zeitlin, "Integer Allocation Problems of Min-Max Type with Quasiconvex Separable Functions", *Operations Research*, Vol. 29, No. 1, pp. 207-211, January-February 1981.

[7] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the Self-Similar Nature of Ethernet Traffic (Extended Version)", *IEEE/ACM Transactions on Networking*, Vol. 2, No. 1, pp.1-15, February 1994.

[8] Q. Yu, Y. Mao, T. Wang, and F. Wu, "Hurst Parameter Estimation and Characteristics Analysis of Aggregate Wireless LAN Traffic", *Proceedings of the IEEE International Conference on Communications, Circuits and System (ICCCAS '05)s*, Vol. 1, pp. 339-345, 2005.

[9] G. Mazzini, R. Rovatti, and G. Setti, "Self-Similarity in Max/Average Aggregated Processes", *Proceedings of the International Symposium on Circuits and System (ISCAS '04).s*, Vol. 5, pp. V-473-V-476, 2004.

[10] F.Y.-S. Lin, P.-H. Tsang, and Y.-L. Lin, "Near Optimal Protection Strategies against Targeted Attacks on the Core Node of a Network", *Proceedings of the 2nd International Conference on Availability, Reliability and Security (ARES'07),* pp. 213-22, April 2007.

[11] M. L. Fisher, "The Lagrangean Relaxation Method for Solving Integer Programming Problems", *Management Science*, Vol. 27, No. 1, pp. 1-18, January 1981.

[12] Y. G. Kim, A. Shiravi, and P. S. Min, "Prediction-Based Routing through Least Cost Delay Constraint", *Proceedings of the 18th International Parallel and Distributed Processing Symposium (IPDPS'04).*

[13] K. T. Cheng and F. Y. S. Lin, "Near-Optimal Delay Constrained Routing in Virtual Circuit Networks", *Proceedings of the 20th IEEE INFOCOM*, Vol. 2, pp. 750-756, 2001.