MIS Quarterly

RESEARCH ARTICLE

GROWTH AND SUSTAINABILITY OF MANAGED SECURITY SERVICES NETWORKS: AN ECONOMIC PERSPECTIVE

Alok Gupta
Department of Information and Decision Sciences, Carlson School of Management, University of Minnesota,
Minneapolis, MN 55455 U.S.A. {alok@umn.edu}

Dmitry Zhdanov
Department of Operations & Information Management, School of Business, University of Connecticut,
Storrs, CT 06269 U.S.A. {dmitry.zhdanov@business.uconn.edu}

Appendix A

MSSP Simulation Design Based on KDD Cup Data Set

The original dataset consists of data on over four million connections each described by 42 attributes (e.g., duration, protocol, etc.) and identified either as normal traffic or one of 24 attack types. Our simulation is designed as follows:

1. A total of 20,000 connections were randomly selected from the original dataset to form the simulation training set; 16 attack types were represented in the simulation set.
2. The simulation training set was duplicated to represent the simulation test set with the same distribution of attacks as in the training set.
3. The 20,000 connections in the training set were randomly split into 20 groups of 1,000. These groups represent firms. It is assumed that each firm can independently observe 1,000 connections.
4. One firm was chosen to start the network (network size = 1, pool of connections = 1,000).
5. The proportion of attack connections in the connection pool was computed and provided the probability of attack, P_a.
6. Based on the pool of connections, the decision tree was built to classify the attacks using the C4.5 algorithm (Quinlan 1993).
7. The output of C4.5 algorithm was tested against a random subset of attacks from the testing set. The testing subset is half the size of the training subset. The proportion of misclassified attacks in the testing subset was computed and provided the probability of attack success, P_s (on the assumption that if the attack was not identified correctly, then no appropriate defense would be activated).
8. Network size was incremented by 1 (until it reached 20; for example, after first iteration, network size = 2, pool of connections = 2,000). Return to step 5.

The data was averaged over 10 simulation runs.

Reference

Quinlan, J. R. 1993. *C4.5: Programs for Machine Learning,* San Mateo, CA: Morgan Kauffman.

Appendix B

Comparative Dynamics of Hiding and Knowledge Effects

Let us define the hiding effect in the network of size N as $H(N)$ and the knowledge effect as $K(N)$.

Recall from equation (1) that damage on the network of size N is defined as

$$D(N) = P_a(N) \times P_s(N) \times N$$

where N is network size, and P_a and P_s are probability of attack taking place and attack success, respectively.

By definition, the hiding effect captures the difference in exposure between being a part of an MSSP network and providing security alone for any particular firm.

$$H(N) = D(N) / N - D(1) / 1$$

Similarly, the knowledge effect captures the improvement in security state due to the addition of a new node to the network.

$$K(N) = P_a(N) \times P_s(N) - P_a(N-1) \times P_s(N-1)$$

To explore the relative magnitude of these two effects, consider a network of size $N + 1$.

$$\begin{aligned} &H(N+1) - K(N+1) = \\ &[P_a(N+1) \times P_s(N+1) - P_a(1) \times P_s(1)] - [P_a(N+1) \times P_s(N+1) - (P_a(N) \times P_s(N)] = \\ &P_a(N) \times P_s(N) - P_a(1) \times P_s(1) = H(N) \end{aligned}$$

We see that for a network of any size, addition of a new node produces an incremental change in the magnitude of hiding effect that is equal to the magnitude of the current knowledge effect. This result may be formulated as a lemma.

Lemma 1. Given the definitions of the hiding and knowledge effects, for any network size N, the following relationship holds:

$$H(N+1) - K(N+1) = H(N)$$

The implications of this result are twofold. First, it means that, in general, the hiding effect dominates the knowledge effect. Second, the extent of this dominance grows with the size of the network, if both effects are present and monotonic (these conditions hold in our formulation).

Appendix C

Correspondence of MSSP Growth Model with the Constructs of the Cooperative Game Theory

1. General properties of cooperative games

1.1 In general, cooperative game theory models assume that there is a finite number of players (N). The MSSP model allows for an unlimited number of potential clients. However, the solution area for the MSSP problem includes a finite number of clients N_{max}. Growing the network beyond this size is going to lead to the loss of efficiency since the gross costs of this approach outweigh gross benefits. Thus, we can consider a game which has N_{max} players.

1.2 A coalition game with transferrable utility is a pair $G = (N, v)$ where N is a coalition and v is a function that associates a real number $v(S)$ with each subset S of N. In our case, the set of players is N_{max}, and the coalition function $v(S)$ is defined based on the consortium rules.

1.3 A game (N, v) is superadditive if $(S, T \subseteq N \text{ and } S \cap T = \varnothing) \Rightarrow v(S \cup T) \geq v(S) + v(T)$. Clearly, this condition holds in our case, since any two subsets of firms cannot achieve a better outcome than a larger consortium.

1.4 A game is weakly superadditive if $v(S \cup \{i\} \geq v(S) + v(\{i\})$ for all $S \subseteq N$ and $i \notin S$. This condition holds in our formulation, since addition of each new member for a consortium increases total consortium benefits by the amount which exceeds an individual value of being alone.

1.5 A game is monotonic if $S \subseteq T \subseteq N \Rightarrow v(S) \leq v(T)$. In our setting, smaller consortia have a smaller amount of total benefit, thus the monotonicity property holds.

2. Solution concepts of cooperative games

2.1 For a game (N, v), a feasible payoff vector is defined as $X^*(N, v) = \{x \in \mathbf{R}^N \mid x(N) \leq v(N)\}$. Basically, feasible payoff should not exceed the total worth of the game. Let Γ be a set of games. Then, a solution on Γ is defined as a function σ that associates with each game $(N, v) \in \Gamma$ a subset $\sigma(N, v)$ of $X^*(N, v)$. Intuitively, a solution is a system of reasonable restrictions on X^*.

2.2 One possible solution for a cooperative game is known as the core. The core $C(N,v)$ is defined as $C(N,v) = \{x \in X^*(N, v) \mid x(S) \geq v(S) \forall S \subseteq N\}$. In the case of the MSSP game, the winning coalition (members of the consortium) get a payoff equal to the worth of the consortium, so the concept of the core is valid in our case.

2.3 The core as a solution to the cooperative game is anonymous (independent of the names of players) and Pareto optimal. Clearly, in our case, any firm can be a member of an MSSP and their particular identities are not relevant. Our solution is also Pareto optimal, since in the formed consortium, no member may be made better off without making another member worse off.

2.4 A solution must be reasonable from above and below. Let b_i represent the $\boldsymbol{i}^{th}$ member incremental contribution to the coalition. Then, the solution is reasonable from above if $((N, v) \in \Gamma, x \in \sigma(N, v)) \Rightarrow x^i \leq b^i_{max}(N, v) \forall i \in N_i$.

It is reasonable from below if $((N, v) \in \Gamma, x \in \sigma(N, v)) \Rightarrow x^i \geq b^i_{max}(N, v) \forall i \in N_i$.

This means that each member of the coalition must be paid an amount that does not exceed the maximum individual contribution to the coalition, while also providing individual incentives to join the coalition. The core is reasonable from above and below. In the MSSP case, there is individual rationality for each member to join (reasonable from below), and none of the members are paid more than the maximum possible individual contribution (reasonable from above). Thus, our solution achieves the same results as the core with respect to the reasonableness requirement.

2.5 Another solution concept is known as the Shapley value. It is based on the *a priori* evaluation of the coalition game by each of its players. Besides Pareto optimality, it also satisfies the null player property and the equal treatment property. The null player property states that each player without an impact on the solution should get a payoff of zero. The equal treatment property states that players with equal contributions to the coalition should get the same payoffs. In our case, the players are identical and they get equal payoffs. There are no dummy players in our formulation.

The Shapley value is defined as $\left\{ \phi^i(v) = \sum_{S \subseteq N} \frac{|S|!(n-|S|-1)!}{n!} \left(v(S \cup \{i\} - v(S) \right) \right\}$

According to Shapley (1953, p. 316),

> The players in N agree to play the game v in a grand coalition, formed in the following way: 1. Starting with a single member, the coalition adds one player at a time until everybody has been admitted. 2. The order in which the players are to join is determined by chance, with all arrangements equally probable. 3. Each player, on his admission, demands and is promised the amount which his adherence contributes to the value of the coalition (as determined by the function v). The grand coalition then plays the game "efficiently" so as to obtain $v(N)$—exactly enough to meet all the promises.

It is clear from this description that our formulation implements the Shapley value mechanism. In addition, we provide a description of the revenue sharing mechanism that implements the equal treatment property, and we prove the optimality of that mechanism.

Therefore, the formulation of the MSSP dynamic growth process has all of the important properties of cooperative games, such as monotonicity and superadditivity. It also corresponds in properties to the common solution concepts such as the core and Shapley value. However, our approach makes fewer assumptions and provides additional results such as the implementation of an optimal, equal treatment-based consortium value distribution mechanism.

References

Peleg, B., and, Sudhölter, P. 2003. *Introduction to the Theory of Cooperative Games*, Boston: Springer.

Shapley, L. S. 1953. "A Value for n-Person Games," in *Contributions to the Theory of Games* (Volume 2), H. Kuhn and A. W. Tucker (eds.), Princeton, NJ: Princeton University Press, pp. 307-317.

Appendix D

Proofs of Propositions

Proposition 1 (The Optimal Size of Consortium without Investment). The optimal size of a consortium-based MSSP with no initial investment, N^*_{cn}, will be less than or equal to the welfare maximizing MSSP network size N^*_s (i.e., $N^*_{cn} \leq N^*_s$).

Proof: Suppose that value function $V(N)$ is concave and the damage function $R(N)$ is convex. Then, the difference between the partial derivates of a concave, $V'(j)$, and a convex function $R'(j)$ of a variable, j, is non-increasing as j increases. Therefore, since the R.H.S. in equation $V'(j) - R'(j) = [V(j) - R(j)] / j$ is a positive number, $j \leq k$ where $V'(k) - R'(k) = 0$—the optimality condition for a welfare maximizing solution. Q.E.D.

Proposition 2 (Equal Sharing and MSSP Network Viability). Let there exist a network size n that allows investment recovery and network viability with the equal sharing rule. Then, it is a minimum viable network size and the equal sharing rule is optimal.

Proof (By Contradiction): We will show that there is no other sharing rule that results in a smaller network size than the equal sharing rule.

Let the investment be recovered at a minimum network size of n using the equal sharing rule

$\Rightarrow$ $[V(n) - R(n)] / n \geq [R(i) - V(i)] / n \; \forall \; n$ members (D4)

Now let there exist a rule such that the initial investment is not equally shared and the investment is recovered at size $m < n$.

$$\Rightarrow [V(m) - R(M)] / m \geq L_j \; \forall \; j = 1, \ldots, m \qquad \text{(D5)}$$

where L_j is the share of investment shared by member j.

However, note that since $m < n$, and the investment is not equally shared, $L_j > [R(i) - V(i)] / m$ for at least some member j.

$$\Rightarrow [V(m) - R(m)] / m \geq [R(i)] / m \qquad \text{(D6)}$$

However, equation (D6) implies that investment should have been recovered using the equal sharing rule at size $m < n$—a contradiction since by assumption n was the minimum network size to recover the investment using the equal sharing rule. Q.E.D.

Proposition 3 (Optimal Size of MSSP Consortium with Investment). The optimal size of an MSSP consortium that requires the initial investment to the overcome critical mass problem, N_c^*, is equal to or greater than the optimal MSSP network size without investment (i.e., $N_c^* \geq N_{cn}^*$).

Proof: Suppose that value function $V(N)$ is concave and the requirement resource function $R(N)$ is convex. Then, the optimal consortium size without investment is a solution to equation (D7):

$$N_{cn}^* = j\text{: } V'(j) - R'(j) = [V(j) - R(j)] / j \qquad \text{(D7)}$$

Further, optimal consortium size with investment is a solution to equation (D8):

$$N_c^* = j\text{: } V'(j) - R'(j) = [V(j) - R(j) - C] / j \qquad \text{(D8)}$$

Since C is a positive number, the R.H.S. of equation (D8) is smaller than the R.H.S. of equation (D7).

Since the difference $V'(j) - R'(D)(j))$ is decreasing in j, it follows that the solution to equation (14), N_{cn}^*, is smaller than the solution to equation (20) (i.e., $N_c^* \geq N_{cn}^*$). Q.E.D.

Lemma 4 (Minimum Viable Initial MSSP Consortium). The minimum starting network size is given by $I^* = \min\{i\text{: } V(N_s^*) - R(N_s^*) \geq R(i) - V(i)\}$.

Proof: Since $R(i) - V(i)$ is decreasing in $i < N_0$ and maximum recoverable investment is $V(N_s^*) - R(N_s^*)$, the smallest viable initial network size is given by $I^* = \min\{i\text{: } V(N_s^*) - R(N_s^*) \geq R(i) - V(i)\}$. Q.E.D.

Proposition 5 (Monopolist MSSP Versus Social Net Benefit Size): The monopolist MSSP may have a larger network size than the social net benefit maximizing size if it can provide sufficient compensation for all current members of the consortia who lose value due to the addition of another member beyond the social benefit optimal (i.e., $P_{N_s^*}^m > \sum_{j \in \Omega} P_j^m \big/ (x+1)$, where x is the number of firms whose benefits are reduced below the price charged to them due to the introduction of the new customer and Ω is the set of individual firms so affected).

Proof: Suppose that the net benefit maximizing network size is N_s^* and the monopolist MSSP is viable.

(1) Note that the profits of a monopolist MSSP cannot be maximized on a network size that is smaller than the social net benefit maximizing size; that is, N_m^* cannot be less than N_s^*.

Assume the contrary, that profits are maximized at $N_m < N_s^*$. Then, there are one or more potential customers in interval $(N_m ; N_s^*]$ who will get positive benefit from joining the network, since N_s^* is the socially optimal size. Charging these customers any positive price up to their willingness to pay and letting them join the network will increase the monopolist's profit. But, N_m was a profit-maximizing point for the monopolist—a contradiction. Thus, N_m^* is at least equal to N_s^*.

(2) We now just need to prove that under certain circumstances $N_m^* > N_s^*$. Consider a case when a monopolist provider attracts one more customer than at the optimal net benefit maximizing network size N_s^*. Then, by definition,

$$V(N_s^*) - R(N_s^*) > V(N_s^* + 1) - R(N_s^* + 1) \tag{D9}$$

However, there may be other firms $k \leq N_s^*$ such that

$$V(k) - R(k) > V(N_s^* + 1) - R(N_s^* + 1) \tag{D10}$$

Let the set of these customers be defined as $\Omega = \{k : V(k) - R(k) > V(N_s^* + 1) - R(N_s^* + 1)\}$.

Each of these customers will require compensation defined by

$$Comp_k = [V(k) - R(k)] / k - [V(N_s^* + 1) - R(N_s^* + 1)] / (N_s^* + 1) \tag{D11}$$

Since $[V(k) - R(D(k))] / k = P_k^m$ and $[V(N_s^* + 1) - R(N_s^* + 1)] / (N_s^* + 1) = P_{N_s^*+1}^m$ and , we can rewrite (D11) as

$$Comp_k = P_k^m - P_{N_s^*+1} \tag{D12}$$

The total compensation then is

$$\sum_{k\in\Omega} Comp_k = \sum_{k\in\Omega} P_k^m - xP_{N_s^*+1}^m \text{ where } x = |\Omega|\text{, cardinality set of } \Omega \tag{D13}$$

Since the price charged to this $(N_s^* + 1)^{st}$ customer should be enough to cover the total compensation, we have

$$P_{N_s^*+1}^m \geq \sum_{k\in\Omega} P_k^m - xP_{N_s^*+1}^m \Rightarrow P_{N_s^*+1}^m \geq \sum_{k\in\Omega} P_k^m \Big/ (x+1) \qquad \text{Q.E.D.} \tag{D14}$$

Corollary **5.1 (Monopolist MSSP Versus Consortium MSSP Size)**: The monopolist MSSP, if viable, will have a network not smaller than a consortium MSSP.

Proof: From Proposition 5, the monopolist MSSP size is greater than the social benefit, $N_m^* \geq N_s^*$. However, the consortium provider will not grow its network beyond N_s^*, as it decreases total and average benefits to its members: $\forall N > N_s^*, W(N) < W(N_s^*) \Rightarrow W(N) / (N) < W(N_s^*) / N_s^*$. Thus, $N_m^* \geq N_s^* \geq N_c$. Q.E.D.

Corollary **5.2 (Monopolist MSSP Versus Consortium MSSP Viability)**: There may be instances when a consortium MSSP is viable, while a monopolist MSSP is not viable.

Proof: Recall that the viability condition for a MSSP network is given by the need to recover the initial investment: $I^* = \min\{i : V(N_s^*) - R(N_s^*) \geq R(i) - V(i)\}$.

In the worst case scenario, the monopolist will start with a network of size 1, while a consortium may have more founding members. Since $R(i) - V(i)$ is decreasing in $i < N_0$, there is a chance that the monopolist will have to make a larger investment. Even if this investment is equal to that of the consortium, the monopolist is following zero-price strategy for the initial few clients, while the consortium begins the cost recovery immediately via its pricing and reallocation scheme. Finally, in some cases the monopolist needs to collect the compensatory payments (if any) in the amount of $\sum_{k\in\Omega} Comp_k = \sum_{k\in\Omega} P_k^m - xP_{N_s^*+1}^m$.

Since the consortium is not facing any of these costs, it may survive on a network with a smaller total potential value than the monopolist. Therefore, there may be cases when, all things being equal, the consortium is profitable and will start, while the monopolist is not profitable and will not start. The reverse is not the case. Q.E.D.

Appendix E

Pseudocode for Profit Maximizing Provider's Pricing and Allocation

```
If (k < N_0)
    Set P_k = 0
End If
Else If (k >= N_0 and k <= N_s^*)
    Set P_k = [V(k) – R(k)] / k
End Else If
Else If (k >= N_s^*)
    Set P_k = [V(k) – R(k)] / k
    For (n = N_0 to N_s^*) do
        If (P_n > [V(k) – R(k)] / k)
            Total refund = 0
            For (m = n to k – 1) do
                Refund firm m amount (R_m) = P_m – [V(k) – R(k)] / k
                Total refund = total refund + P_m – [V(k) – R(k)] / k
            End For
            If (Total refund > [V(k) – R(k)] / k)
                Reject entry to firm k
            Else
                For (m = n to k – 1) do
                    Commit R_m
                    P_m = [V(k) – R(k)] / k
                End For
            End Else
        End If
    End For
End Else If
```