Distributing the Cost of Securing a Transportation Infrastructure*

Sudarshan S. Chawathe

Computer Science Department
University of Maine
Orono, Maine 04469, USA

chaw@cs.umaine.edu
http://www.cs.umaine.edu/ chaw/

Abstract. We address the problem of fairly distributing the cost of system-wide improvements to the security of a transportation infrastructure over the beneficiaries. We present a framework that models transportation links and the emergence (magnitude and frequency) and propagation of threats. The cost-distribution is based on a weighted sum that characterizes the expected reduction in the vulnerability of a site as a result of the security improvements.

1 Introduction

Securing the transportation infrastructure to protect it from hostile agents is an increasingly important task that is the subject of much recent work. No matter what strategy one uses for improving the security of the infrastructure, there are substantial and varied costs related to personnel, equipment, impediments to traffic, loss of revenue due to slow or rerouted traffic, etc. Once such costs have been determined, an important question is how they are borne by the typically numerous parties involved in the infrastructure. Indeed, lack of agreement on such division of costs has been the topic of much political controversy and threatens to derail initiatives for securing the transportation infrastructure.

For example, consider a proposal to implement additional checkpoints on some highways of a regional network and to disallow hazardous-material carriers on certain routes. Such actions incur the obvious direct costs associated with setting up checkpoints and enforcing new regulations. However, there are also indirect costs such as noise, pollution, and danger of rerouted hazardous-material carriers. Further, additional checkpoints may lead to congestion which may result in loss of business in the affected areas. It is not surprising, then, that even modest proposals that affect the functioning of the transportation infrastructure often elicit strong protests.

Given the increased awareness of security, it is likely that major disagreement is not about whether additional security is necessary, but rather about who

* This work was supported in part by the U.S. National Science Foundation with grants IIS-9984296, IIS-0081860, and CNS-0426683.

should shoulder what portion of its cost. In this paper, we present a model of the costs and benefits of improvements to transportation-infrastructure security. Using this model, we can determine a cost distribution that has a sound basis and is thus likely to be considered fair by the concerned parties.

2 Model

We model a network of transportation links using a graph whose vertices represent locations of interest, or simply intersections, and whose edges represent links. A link (i, j) between locations i and j permits, in general, travel both from i to j as well as from j to i. However, travel in these two directions is modeled using separate parameters, as described below. Figure 1 on page 3 suggests a small transportation network modeled in this manner.

One-way links do not pose a problem to this model. The disallowed direction may simply be assigned a very low probability of traversal. A nonzero probability of traversal in the disallowed direction of a one-way link, such as a one-way street, may often model the link more accurately than a zero probability because, for example, it is quite likely that agents perpetrating an attack do not hold traffic regulations in high regard.

Threats may originate at any vertex of the graph representing the transportation network. Threats that originate at locations on a link between two vertices, such as a rail link between two stations, are modeled by inserting an additional vertex between those vertices. In other words, locations at which threats originate are, by definition, locations of interest and are therefore modeled using vertices in the graph.

We use two parameters to describe threats originating at a vertex i: a *magnitude* m_i and a *frequency* f_i. Intuitively, the magnitude represents the seriousness of a threat, modeling quantities such as the amount of damage and the affected area. The frequency indicates how often a threat is likely to materialize at i. Our methods do not depend on any particular interpretation of these parameters. Further, our work uses these parameters only in conjunction, as the product $m_i f_i$, which represents the expected magnitude, per unit time, of a threat originating at vertex i. Determining appropriate values for m_i, f_i, and the other parameters of our model is an important problem, but not one that is the focus of this paper. Our focus is on how such data, once obtained, may be used to allocate system-wide costs of securing the transportation network.

When a threat appears at a location in the network, it may either be executed at that location or be transported to another location using one of the links. A threat may appear at a location either because it originates at that location, as described earlier, or because it traversed a link from another location to that location. We use e_i to denote the probability that a threat appearing at a location is executed at that location. More precisely, the *probability of execution* e_i is the conditional probability of a threat executing at i given that it has appeared at i. Similarly, we use t_{ij} to denote the conditional *probability of traversal* from i to j, given a threat appearing at i. In general, t_{ij} and t_{ji} are not the same. Let

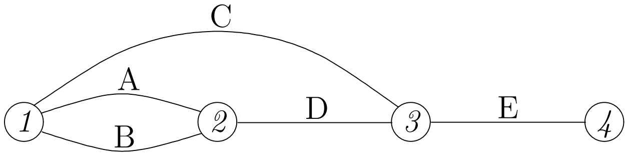

Fig. 1. A network of transportation links used by the running example.

us use nbd(i) to denote the *neighborhood* of i, i.e., the set of vertices that are the targets of links from i. Since e_i and t_{ij} represent probabilities we must have the following for every vertex i in the graph: $e_i + \sum_{j \in \text{nbd}(i)} t_{ij} \leq 1$. However, the terms on the left-hand side of this inequality need not sum to one because the threat may disappear (e.g., a planned attack may be abandoned).

We denote the cost of improving the security on link (i, j) by c_{ij}. The resulting (lower) link traversal probability is denoted by t'_{ij}. Finally, we define $s_{ij} = 1 - t_{ij}$ and $s'_{ij} = 1 - t'_{ij}$ for notational convenience.

3 Vulnerabilities

The model of Section 2 allows us to quantify the *vulnerability* of each location of interest. Intuitively, the vulnerability of a location of interest is the expected magnitude of a threat executed at that location. In order to keep the calculations simple, we henceforth assume that a threat is executed at its intended target (location of interest) as soon as it arrives at that target. That is, we may restrict our attention to traversals that do not visit any vertex more than once (acyclic paths). It is conceptually easy to do away with this assumption by using the steady-state distribution obtained by interpreting the graph of Section 2 as a Markov process [1].

Figure 1 suggests a small transportation network that we shall use as a running example. The four locations of interest are identified by the numbers within the circles: $V = \{1, 2, 3, 4\}$. The five links are identified by the letters above each link: $L = \{A, B, C, D, E\}$. The magnitude and frequency of a threat originating at vertex 1 are 1024 and 4, respectively, so that $f_1 m_1 = 4096$. (The numbers are chosen to minimize fractions in the calculations that follow but nothing in our model depends on such carefully chosen values.) For all other vertices in this example, $f_i m_i = 0$. That is, a nontrivial threat originates only at vertex 1. We use an execution probability of 1/4 at each vertex. That is, $e_i = 1/4$ for all $i \in [1, 4]$. Traversal probabilities for all links (in either direction) are uniformly 1/4. That is, for all $i, j \in [1, 4]$, $i \neq j$, we have $t_{ij} = t_{ji} = 1/4$ (and thus $s_{ij} = s_{ji} = 3/4$). We may verify that these values satisfy $e_i + \sum_{j \in \text{nbd}(i)} t_{ij} \leq 1$ for all vertices i in our example.

By a slight abuse of notation, we shall use the link identifiers, such as A and B, to denote both the links themselves and traversals of those links. More precisely, given a link $X = (i, j)$ with $i < j$, a traversal from i to j is denoted by X while a traversal from j to i is denoted by X'. Paths are denoted by concatenating these labels for the traversals, in sequence. Thus, given Figure 1, AD denotes a path from vertex 1 to vertex 3 via vertex 2, while $A'CE$ denotes a path from vertex 2 to vertex 4 via vertices 1 and 3. Further, we use $\overline{X}$ to denote a non-traversal from i to j, and similarly $\overline{X'}$ to denote a non-traversal from j to i.

Let us now calculate the vulnerability of vertex 3. Since there is only one origin of threats (vertex 1) in our example, the vulnerability of any other vertex in our example depends only on the paths leading to that vertex from vertex 1. A threat from vertex 1 may arrive at vertex 3 either directly, using link C, or via vertex 2, by using either of links A and B followed by link D. Therefore, we may calculate the probability of a threat from vertex 1 arriving at vertex 3 as follows.

$$\begin{aligned} P(((A \text{ or } B) \text{ and } D) \text{ or } C) &= 1 - P(\overline{(((A \vee B) \wedge D) \vee C)}) \\ &= 1 - P(\overline{C})(1 - (1 - P(\overline{A})P(\overline{B}))P(\overline{D})) \end{aligned}$$

Using our notation for the traversal probabilities from Section 2, we have the following expression for the vulnerability of vertex 3:

$$v_3 = m_1 f_1 (1 - s_C(1 - (1 - s_A s_B) t_D)) \qquad (1)$$

Substituting the parameter values from our running example yields

$$v_3 = 1024 \cdot 4 \cdot (1 - (1/4)(1 - (1 - (3/4)(3/4))(1/4)) = 1360$$

The interpretation of this number depends on the interpretation used in assigning values to the parameters m_1 and f_1. For instance, if m_1 represents the number of persons affected by a bomb and if f_1 represents the number of times a year such a bomb is expected to originate at site 1, then 1360 is the expected number of people affected yearly by the bomb, given our model. However, our work is equally applicable to any other interpretation that fits our model described in Section 2.

The above calculations are based on the state of the transportation network before any security improvements are made, i.e., the base state. In general, the vulnerability v_i of vertex i depends on the set of links on which security improvements have been made. Therefore, we use $v_i(S)$ to denote the vulnerability of i given a set S of improved links. The left-hand side of Equation 1 is expressed as $v_3(\emptyset)$ in this notation, which we shall henceforth use.

Continuing our running example (Figure 1), suppose that improving the security of a link halves the probability of traversal. Recall that we have traversal probabilities $t_{ij} = 1/4$ for all $i, j \in [1, 4]$, $i \neq j$. Using the notation of Section 2, we have, for all $i, j \in [1, 4]$, $i \neq j$, $t'_{ij} = 1/8$ and $s'_{ij} = 7/8$.

We may calculate $v_3(S)$ for all $S \subseteq L$ using Equation 1 by substituting s_A, s_B, s_C, and t_D with, respectively, s'_A, s'_B, s'_C, and t'_D depending on whether A, B, C, and D (respectively) belong to S. For $S = \{A, C\}$, substituting s'_A for s_A

Improved set S	Vulnerability $v_3(S)$
$\emptyset$, $\{E\}$	1360
$\{A\}$, $\{A, E\}$	1288
$\{B\}$, $\{B, E\}$	1288
$\{C\}$, $\{C, E\}$	904
$\{D\}$, $\{D, E\}$	1192
$\{A, B\}$, $\{A, B, E\}$	1204
$\{A, C\}$, $\{A, C, E\}$	820
$\{A, D\}$, $\{A, D, E\}$	1156
$\{B, C\}$, $\{B, C, E\}$	820
$\{B, D\}$, $\{B, D, E\}$	1156
$\{C, D\}$, $\{C, D, E\}$	708
$\{A, B, C\}$, $\{A, B, C, E\}$	722
$\{A, B, D\}$, $\{A, B, D, E\}$	1114
$\{A, C, D\}$, $\{A, C, D, E\}$	666
$\{B, C, D\}$, $\{B, C, D, E\}$	666
$\{A, B, C, D\}$, $\{A, B, C, D, E\}$	617

Table 1. Vulnerability of vertex 3 of Figure 1 (page 3) for different sets of improved links, based on the discussion in Section 3.

and s'_C for s_C yields the following:

$$\begin{aligned} v_3(\{A, C\}) &= m_1 f_1(1 - s'_C(1 - (1 - s'_A s_B)t_D)) \\ &= 1024 \cdot 4 \cdot (1 - (7/8)(1 - (7/8)(3/4))(1/4)) = 820 \end{aligned}$$

The result of such calculations for all subsets S in our running example summarized in Table 1. There is no origin of a threat at vertex 4 in our example. Therefore, as indicated by Equation 1, link E is immaterial for calculating the vulnerability of vertex 3. This fact explains the two sets in the first column of each row of Table 1.

The benefit of improving the security of a link is, in general, different for each the vertex. A fair scheme for distributing the cost of improving links over the vertices in the network reflects these differing benefits using the above framework. We defer the details of the distribution scheme to a forthcoming paper.

4 Related Work

While our work abstracts away some of the details of how various model parameters are determined, such determination is nevertheless very important and forms the basis of our model by providing the important parameters. For example, Shao presents a method for allocating redundant resources for disaster-recovery planning [2]. Similarly, recent work by Park et al. may be used to determine the vulnerabilities of nodes in a computer-network infrastructure [3]. Sinai discusses how work in the social and behavioral sciences may be applied to model and assess threats of terrorism [4]. Such work is key to determining the parameters, such as threat magnitude and frequency, used by our model in this paper. Information resources such as the MIPT system described by Ellis provide a means for efficiently accessing a variety of information necessary for threat assessment [5]. A similar effort in the context of the spread of infectious diseases is described by Zeng et al. [6]. Park and Ho describe a method for addressing insider threats [7], which are an important category of threats in any environment, including the one in this paper. Lin et al.'s user-acceptance study based on the COPLINK system [8] highlights the importance of solutions that make a compelling case for acceptance, which is also one of the motivations of our work in this paper. Xu et al. present a method to analyze and visualize criminal networks, focusing on dynamics [9]. Introducing the dynamic element into our model in this paper is an interesting avenue for further work.

References

1. Breiman, L.: Probability. Society for Industrial and Applied Mathematics (1992)
2. Shao, B.B.M.: Optimal redundancy allocation for disaster recovery planning in the network economy. In: Proceedings of the Symposium on Intelligence and Security Informatics (ISI). Volume 3073 of Lecture Notes in Computer Science (LNCS)., Tucson, Arizona (2004) 484–491
3. Park, E., Seo, J.T., Im, E.G., Lee, C.W.: Vulnerability analysis and evaluation within an intranet. In: Proceedings of the Symposium on Intelligence and Security Informatics (ISI). Volume 3073 of Lecture Notes in Computer Science (LNCS)., Tucson, Arizona (2004) 514–515
4. Sinai, J.: Utilizing the social and behavioral sciences to assess, model, forecast and preemptively respond to terrorism. In: Proceedings of the Symposium on Intelligence and Security Informatics (ISI). Volume 3073 of Lecture Notes in Computer Science (LNCS)., Tucson, Arizona (2004) 531–533
5. Ellis III, J.O.: MIPT: Sharing terrorism information resources. In: Proceedings of the Symposium on Intelligence and Security Informatics (ISI). Volume 3073 of Lecture Notes in Computer Science (LNCS)., Tucson, Arizona (2004) 520–525
6. Zeng, D., Chen, H., Tseng, C., Larson, C., Eidson, M., Gotham, I., Lynch, C., Ascher, M.: West nile virus and botulism portal: A case study in infectious disease informatics. In: Proceedings of the Symposium on Intelligence and Security Informatics (ISI). Volume 3073 of Lecture Notes in Computer Science (LNCS)., Tucson, Arizona (2004) 28–41
7. Park, J.S., Ho, S.M.: Composite role-based monitoring (CRBM) for countering insider threats. In: Proceedings of the Symposium on Intelligence and Security Informatics (ISI). Volume 3073 of Lecture Notes in Computer Science (LNCS)., Tucson, Arizona (2004) 201–213
8. Lin, C., Hu, P.J., Chen, H., Schroeder, J.: Technology implementation management in law enforcement: COPLINK system usability and user acceptance evaluations. In: Proceedings of the National Conference on Digital Government Research, Boston, Massachusetts (2003)
9. Xu, J., Marshall, B., Kaza, S., Chen, H.: Analyzing and visualizing criminal network dynamics: A case study. In: Proceedings of the Symposium on Intelligence and Security Informatics (ISI). Volume 3073 of Lecture Notes in Computer Science (LNCS)., Tucson, Arizona (2004) 359–377