Acta Cybernetica **22** (2016) 735–769.

The Structure of Rooted Weighted Trees Modeling Layered Cyber-security Systems

Geir Agnarsson*, Raymond Greenlaw†, and Sanpawat Kantabutra‡

Abstract

In this paper we consider the structure and topology of a layered-security model in which the containers and their nestings are given in the form of a rooted tree T. A *cyber-security model* is an ordered three-tuple $M = (T, C, P)$ where C and P are multisets of *penetration costs* for the containers and *target-acquisition values* for the prizes that are located within the containers, respectively, both of the same cardinality as the set of the non-root vertices of T. The problem that we study is to assign the penetration costs to the edges and the target-acquisition values to the vertices of the tree T in such a way that minimizes the total prize that an attacker can acquire given a limited *budget*. The attacker breaks into containers starting at the root of T and once a vertex has been broken into, its children can be broken into by paying the associated penetration costs. The attacker must deduct the corresponding penetration cost from the budget, as each new container is broken into. For a given assignment of costs and target values we obtain a *security system*. We show that in general it is not possible to develop an optimal security system for a given cyber-security model M. We define P- and C-models where the penetration costs and prizes, respectively, all have unit value. We show that if T is a rooted tree such that any P- or C-model $M = (T, C, P)$ has an optimal security system, then T is one of the following types: (i) a rooted path, (ii) a rooted star, (iii) a rooted 3-caterpillar, or (iv) a rooted 4-spider. Conversely, if T is one of these four types of trees, then we show that any P- or C-model $M = (T, C, P)$ does have an optimal security system. Finally, we study a duality between P- and C-models that allows us to translate results for P-models into corresponding results for C-models and vice versa. The results obtained give us some mathematical insights into how layered-security defenses should be organized.

Keywords: cyber-security model, duality, graph minors, rooted tree, security system, system attack, tree types, weighted rooted tree

*Department of Mathematical Sciences, George Mason University, Fairfax, VA 22030, E-mail: geir@math.gmu.edu

†Department of Cyber Sciences, United States Naval Academy, Annapolis, Maryland 21402, E-mail: greenlaw@usna.edu

‡The Theory of Computation Group, Computer Engineering Department, Chiang Mai University, Chiang Mai, 50200, Thailand, E-mail: sanpawat@alumni.tufts.edu

DOI: 10.14232/actacyb.22.4.2016.2

1 Introduction

According to [6], the global cyber-security market cost in 2017 is expected to top 120 billion US dollars. This site also reports that there are 18 victims of a cyber crime every single second! Other sources report similarly alarming and worsening statistics. There is agreement that the number of cyber attacks is increasing rapidly, and the consequences of such attacks are greater than ever on economics, national security, and personal data. Threats come from nation states with advanced cyber warfare commands, nation states having less technical capabilities but intent on doing harm, ideologically motivated groups of hackers or extremists, profit-seeking criminals, and others. As a result, quite a bit of work has been done where cyber-security systems, or more generally layered computer systems, are modeled as a fixed weighted trees. For example, in [1, 3, 4, 8, 10, 12] the authors consider finding *weight-constrained, maximum-density subtrees* and similar structures given a fixed weighting of a tree as part of the input. In these cases weights are specified on both vertices and edges. There has also been some research on *network fortification* and problems related to that topic. For example, in [13] stochastic linear programming games are studied and it is demonstrated how these can, among other things, model certain network fortifications. In [14] the problem of network interdiction is studied – how to minimize the maximum amount of flow an adversary/enemy can push through a given network from a source s to a sink t. There each edge/arc is provided with a fixed integer capacity and an integer resource (required to delete the edge/arc). This is a variation of the classical Max-Flow-Min-Cut Theorem. Although interesting in their own way, neither of these papers or related papers that we have found in the literature address directly what we study in this paper. To build secure systems requires first principles of security. "In other words, we need a *science of cyber-security* that puts the construction of secure systems onto a firm foundation by giving developers a body of laws for predicting the consequences of design and implementation choices" [11]. To this end, Schneider called for more models and abstractions to study cyber security [11]. This paper is a step in that direction. We hope that others will build on this work to develop even better and more realistic models, overcome the shortcomings of our model, as well as develop additional foundational results.

Building on the work done in [3], in this paper we study a layered-security model and strategies for assigning *penetration costs* and *target-acquisition values* so as to minimize the amount of damage an attacker can do to a system. That is, we examine *security systems*. The approach we take here is to assign weights to the vertices and edges of a tree in order to build a cyber defense that minimizes the amount of prize an attacker can accumulate given a limited budget. To the best of our knowledge this approach is new in that the usual approach is to consider a particular weighted tree as input. In [3] the following question was posed: Can one mathematically prove that the intuition of storing high-value targets deeper in the system and having higher penetration costs on the outer-most layers of the system results in the best or at least good security? In this paper we answer this question and obtain more general and specific results. We define three types of security

systems: *improved*, *good*, and *optimal*. We show that not all *cyber-security models* admit optimal security systems, but prove that paths and stars do. We define and study *P-* and *C-models* where all penetration costs, or all prizes, are set to one, respectively. We classify the types of trees that have optimal security systems for both P- and C-models. We then discuss a duality between P- and C-models, which provides a dictionary to translate results for P-models into corresponding results for C-models, and vice versa.

The outline of this article is as follows. In Section 2 we present the rationale for our layered-security model. In Section 3 we define the framework for security systems and present the definitions of improved, good, and optimal security systems, and state some related observations and examples. In Section 4 we explore optimal security systems and prove that they do not always exist, but they exist if and only if the underlying tree T of the given security system is either a path rooted at a leaf, or a star rooted at its center vertex. In Section 5 we define P- and C-models and show that any cyber-security model $M = (T, C, P)$ is equivalent to both a P-model M' and a C-model M''. We further show that if T is a rooted tree such that any P- or C-model M has an optimal security system, then T is one of the following four types: (i) a rooted path, (ii) a rooted star, (iii) a rooted 3-caterpillar, or (iv) a rooted 4-spider. In Section 6 we prove that if T is one of the four types of rooted trees mentioned above, then any P-model does indeed have an optimal security system. In Section 7 we define a duality between equivalence classes of P-models and equivalence classes of C-models that serves as a dictionary allowing us to obtain equivalent results for C-models from those of the P-models that were obtained in Section 6. In particular, we obtain Theorem 7.2 that completely classifies which P- and which C-models have optimal security systems. Conclusions and open problems are discussed in Section 8.

2 Rationale for Our Layered-Security Model

In defining our layered-security model to study defensive cyber security, we need to strike a balance between simplicity and utility. If the model is too simple, it will not be useful to provide insight into real situations; if the model is too complex, it will be too cumbersome to apply, and we may get bogged down in too many details. The model described in this paper is a step toward gaining a better understanding of a broad range of security systems in a graph-theoretical setting for a layered-security model.

Many systems contain layered security or what is commonly referred to as *defense-in-depth*, where valuable assets are hidden behind many different layers or secured in numerous ways. For example, a *host-based defense* might layer security by using tools such as signature-based vendor anti-virus software, host-based systems security, host-based intrusion-prevention systems, host-based firewalls, encryption, and restriction policies, whereas a *network-based defense* might provide defense-in-depth by using items such as web proxies, intrusion-prevention systems, firewalls, router-access control lists, encryption, and filters [9]. To break into such

a system and steal a valuable asset requires that several levels of security be penetrated, and, of course, there is an associated cost to break into each level, for example, money spent, time used, or the punishment served for getting caught.

Our model focuses on the layered aspect of security and is intended to capture the notion that there is a cost associated with penetrating each additional level of a system and that attackers have finite resources to utilize in a cyber attack. Defenders have the ability to secure targets using defense mechanisms of various strengths and to secure targets in desired locations and levels. We assume that the structure where targets will be stored, that is, the container nestings; is given as part of the input in the form of a rooted tree. In this way we can study *all* possible structures at a single time, as they can be captured in the definition of our problems. This methodology is as opposed to having the defender actually construct a separate defense structure for each input.

For any specific instance of a problem, a defender of a system will obviously consider the exact details of that system and design a layered-security approach to fit one's actual system. Similarly, a traveling salesman will be concerned about constructing a tour of *his* particular cities, not a tour of any arbitrary set of cities with any arbitrary set of costs between pairs of cities. Nevertheless, researchers have found it extremely helpful to consider a general framework in which to study the Traveling Salesman Problem. And, in studying the general problem, insights have been gained into *all* instances of the problem. Thus, we believe it is worthwhile to consider having a fixed structure as part of our input, and this approach is not significantly different from that used in complexity theory to study problems [5, 7].

In this paper we focus on a static defense. We pose as an open problem the question of how to create a defense and an attack strategy if the defender is allowed to move targets around dynamically or redistribute a portion of a prize. We also consider the total prize as the sum of the individual values of the targets collected although one could imagine using other or more complex functions of the target values to quantify the damage done by an attacker. Our defensive posture is formed by assigning to the edges and vertices of the rooted tree in question the input-provided penetration costs and target-acquisition values, respectively. We formalize the model, the notion of a security system, and the concept of a system attack in the next section.

3 Cyber-Security Model and Security Systems

Let $\mathbb{N} = \{1, 2, 3, \ldots\}$, $\mathbb{Q}$ be the rational numbers, and $\mathbb{Q}_+$ be the non-negative rational numbers.

Definition 3.1. *A* cyber-security model (CSM) *M is given by a three-tuple $M = (T, C, P)$, where T is a directed tree rooted at r having $n \in \mathbb{N}$ non-root vertices, C is a multiset of* penetration costs *$c_1, \ldots, c_n \in \mathbb{Q}_+$, and P is a multiset of* target-acquisition-values *(or* prizes *for short) $p_1, \ldots, p_n \in \mathbb{Q}_+$.*

Remark. As mentioned right after Observation5.1, strictly speaking, we could have stated the above definition using the set $\mathbb{N}$ of natural numbers instead of non-negative rationals $\mathbb{Q}_+$ for possible penetrations costs and prizes. We do, however, prefer the most general definition we can discuss.

Throughout $V(T) = \{r, u_1, \ldots, u_n\}$, where r is the designated root that indicates the start of a *system attack*, and $E(T) = \{e_1, \ldots, e_n\}$ denotes the set of edges of T, where our labeling is such that u_i is always the head of the edge e_i. The prize at the root is set to 0. The penetration costs model the expense for breaking through a layer of security, and the target-acquisition-values model the amount of prize one acquires for breaking through a given layer and exposing a target. The penetration costs will be weights that are assigned to edges in the tree, and the target-acquisition-values, or the prizes, are weights that will be assigned to vertices in the tree.

Sometimes we do not distinguish a target from its acquisition value or prize, nor a container, which is a layer of security, from its penetration cost. Note that one can think of each edge in the rooted tree as another container, and as one goes down a path in the tree, as penetrating additional layers of security. We can assume that the number of containers and targets is the same. Since if we have a container housing another container (and nothing else), we can just look at this "double" container as a single container of penetration cost equal to the sum of the two nested ones. Also, if a container includes many prizes, we can just lump them all into a single prize, which is the sum of them all.

Recall that in a rooted tree T, each non-root vertex $u \in V(T)$ has exactly one parent, and that we assume the edges of T are directed naturally away from the root r in such a way that each non-root vertex has an in-degree of one. The root is located at *level* 0 of the tree. *Level* 1 of the tree consists of the children of the root, and, in general, *level* i of the tree consists of the children of those vertices at level $i-1$ for $i \geq 1$. We next present some key definitions about a CSM that will allow us to study questions about *security systems*.

Definition 3.2. *A* security system (SS) *with respect to a cyber-security model $M = (T, C, P)$ is given by two bijections $c : E(T) \to C$ and $p : V(T) \setminus \{r\} \to P$. We denote the security system by (T, c, p).*

A system attack (SA) *in a security system (T, c, p) is given by a subtree τ of T that contains the root r of T.*

- *The* cost *of a system attack τ with respect to a security system (T, c, p) is defined by*
$$\mathrm{cst}(\tau, c, p) = \sum_{e \in E(\tau)} c(e).$$
- *The* prize *of a system attack τ with respect to a security system (T, c, p) is defined by*
$$\mathrm{pr}(\tau, c, p) = \sum_{u \in V(\tau)} p(u).$$

- *For a given* budget $B \in \mathbb{Q}_+$ *the maximum prize* $\mathrm{pr}^\star(B,c,p)$ *with respect to* B *is defined by*

$$\mathrm{pr}^\star(B,c,p) := \max\{\mathrm{pr}(\tau,c,p) : \textit{for all system attacks } \tau \subseteq T, \textit{ where } \mathrm{cst}(\tau,c,p) \le B\}.$$

A system attack τ *whose prize is a maximum with respect to a given budget is called an* optimal attack.

The bijection c in Definition 3.2 specifies how difficult it is to break into the various containers, and the bijection p specifies the prize associated with a given container. Note that for any SS (T,c,p) we have $\mathrm{cst}(r,c,p) = 0 \le B \in \mathbb{Q}_+$. When $T = (\{r\},\varnothing)$, then $\mathrm{pr}^\star(B,c,p) = 0$ for any $B \in \mathbb{Q}_+$. When two bijections are given specifying a SS, we call the resulting weighted tree a *configuration of the CSM*. Any configuration represents a defensive posture and hence the name security system. Note that the CSM can be used to model any general security system and not just cyber-security systems. We are interested in configurations that make it difficult for an attacker to accumulate a large prize. It is natural to ask if a given defensive stance can be improved. Next we introduce the notion of an *improved security system* that will help us to address this question.

Definition 3.3. *Given a CSM* $M = (T,C,P)$ *and a SS* (T,c,p)*, an* improved security system (improved SS) with respect to (T,c,p) *is a SS* (T,c',p') *such that for any budget* $B \in \mathbb{Q}_+$ *we have* $\mathrm{pr}^\star(B,c',p') \le \mathrm{pr}^\star(B,c,p)$*, and there exists some budget* $B' \in \mathbb{Q}_+$ *such that* $\mathrm{pr}^\star(B',c',p') < \mathrm{pr}^\star(B',c,p)$.

Definition 3.3 captures the idea of a better placement of prizes and/or penetration costs so that an attacker cannot do as much damage. That is, in an improved SS one can never acquire a larger overall maximum prize with respect to any budget B; and furthermore, there must be at least one particular budget where the attacker actually does worse. Notice that there can be an improved SS (T,c',p'), where for some budget $B \in \mathbb{Q}_+$, there is a SA τ whose cost is less than or equal to B for both SSs such that $\mathrm{pr}(\tau,c',p') > \mathrm{pr}(\tau,c,p)$. In this case an attacker obtains a larger prize in the improved SS; and, of course, this situation is undesirable and means a weaker defense against this specific attack. We, however, are interested in improved SSs with respect to a given budget rather than a particular SA. Since we have exactly n penetration costs and n prizes to assign, it is difficult to imagine an improved SS for all but the most-restricted trees in which all SAs would be improved in the sense just described. Next, we formalize the notion of an *optimal security system*.

Definition 3.4. *Let* $M = (T,C,P)$ *be a given CSM. (i) For a budget* $B \in \mathbb{Q}_+$*, a SS* (T,c,p) *is* optimal w.r.t. B *if there is no other SS* (T,c',p') *for* M *such that* $\mathrm{pr}^\star(B,c',p') < \mathrm{pr}^\star(B,c,p)$*. (ii)* (T,c,p) *is* optimal *if it is optimal w.r.t. any budget* $B \in \mathbb{Q}_+$.

Notice that an optimal SS is not necessarily the best possible. We could define a *critically optimal security system* to be one where for every single SA the SS was

at least as good as all others and for at least one better. And, in a different context, these SSs might be interesting. However, in light of Theorem 4.1 in the following section, which shows that even an optimal SS may not exist for a given CSM, we do not pursue critically optimal SSs further in this paper. By Definitions 3.3 and 3.4 we clearly have the following.

Observation 3.1. *A SS (T, c, p) for a CSM $M = (T, C, P)$ is optimal if and only if no improved SS for (T, c, p) exists.*

We next introduce the concept of two closely-related configurations of a CSM, and this notion will give us a way to relate SSs.

Definition 3.5. *Given a CSM $M = (T, C, P)$, the two configurations (T, c, p), and (T, c', p') are said to be* neighbors *if*

1. *there exists an edge $(u, v) \in E(T)$ such that*

$$\begin{aligned} p'(v) &= p(u) \\ p'(u) &= p(v) \\ p'(w) &= p(w), \text{ otherwise, or} \end{aligned}$$

2. *there exist two edges $(u, v), (v, w) \in E(T)$ such that*

$$\begin{aligned} c'((u, v)) &= c((v, w)) \\ c'((v, w)) &= c((u, v)) \\ c'((x, y)) &= c((x, y)), \text{ otherwise.} \end{aligned}$$

The notion of neighboring configurations will be useful in developing algorithms for finding *good security systems*, which we define next.

Definition 3.6. *A* good security system (good SS) *is a SS (T, c, p) such that no neighboring configuration results in an improved security system.*

Given a SS (T, c, p) for a CSM M, a natural question to pose is whether a local change to the SS can be made in order to strengthen the SS, that is, make the resulting SS improved. In a practical setting one may not be able to redo the security of an entire system, but instead may be able to make local changes.

Suppose $(u, v) \in E(T)$ where $p(u) \geq p(v)$, and let p' be the prize assignment obtained from p by swapping the prizes on u and v, that is $p'(u) = p(v)$, $p'(v) = p(u)$, and $p'(w) = p(w)$ otherwise. If now τ is any SA, then $\mathrm{pr}(\tau, c, p') = \mathrm{pr}(\tau, c, p)$ if either both $u, v \in V(\tau)$ or neither u nor v are vertices of τ, or $\mathrm{pr}(\tau, c, p') \leq \mathrm{pr}(\tau, c, p)$ if $u \in V(\tau)$ and $v \notin V(\tau)$. In either case $\mathrm{pr}(\tau, c, p') \leq \mathrm{pr}(\tau, c, p)$ and therefore we have for any budget B that

$$\mathrm{pr}^\star(B, c, p') \leq \mathrm{pr}^\star(B, c, p). \tag{1}$$

Similarly, if $(u, v), (v, w) \in E(T)$ where $c((u, v)) \leq c((v, w))$, let c' be the cost assignment obtained from c by swapping the costs on the incident edges (u, v) and

(v,w) and leave all the other edge-costs unchanged, that is $c'((u,v)) = c((v,w))$, $c'((v,w)) = c((u,v))$ and $c'(e) = c(e)$ otherwise. If τ is a SA, then clearly we always have $\mathrm{pr}(\tau, c', p) = \mathrm{pr}(\tau, c, p)$. Also, if either both $(u,v),(v,w) \in E(\tau)$ or neither (u,v) nor (v,w) are edges in τ, then $\mathrm{cst}(\tau, c', p) = \mathrm{cst}(\tau, c, p)$, and if $(u,v) \in E(\tau)$ and $(v,w) \notin E(\tau)$, then $\mathrm{cst}(\tau, c', p) \geq \mathrm{cst}(\tau, c, p)$. In either case we have $\mathrm{cst}(\tau, c', p) \geq \mathrm{cst}(\tau, c, p)$. Hence, if B is any budget, then by mere definition we have that

$$\mathrm{pr}^\star(B, c', p) \leq \mathrm{pr}^\star(B, c, p). \tag{2}$$

By (1) and (2) we have the following proposition.

Proposition 3.1. *Let $M = (T, C, P)$ be a CSM. A SS given by (T, c, p) is a good SS if for all $(u,v),(v,w) \in E$ we have $c((u,v)) \geq c((v,w))$ and for all non-root vertices $u, v \in V(T)$ with $(u,v) \in E(T)$ we have $p(u) \leq p(v)$.*

Note that Proposition 3.1 says that on any root to leaf path in T the penetration costs occur in decreasing order and the prizes occur in increasing order.

From any configuration resulting from a SS (T, c, p) for a CSM, Proposition 3.1 gives a natural $O(n^2)$ algorithm for computing a good SS by repeatedly moving to improved neighboring configurations until no more such neighboring configurations exist. We can do better than this method by first sorting the values in C and P using $O(n \log n)$ time, and then conducting a breath-first search of T in $O(n)$ time. We can then use the breath-first search level numbers to define bijections c and p that meet the conditions of a good SS. We summarize in the following.

Observation 3.2. *Given a CSM $M = (T, C, P)$, there is an $O(n \log n)$ algorithm for computing a good SS for M.*

If we could eliminate the sorting step, we would have a more efficient algorithm for obtaining a good SS, or if we restricted ourselves to inputs that could be sorted in $O(n)$ time. Also, notice that a good SS has the heap property, if we ignore the root. However, in our case we cannot "choose" the shape of the heap, but we must use the structure that is given to us as part of our input.

Suppose that our SS (T, c, p) for M satisfies a *strict* inequality $p(u) > p(v)$ for some $(u,v) \in E(T)$, or that $c((u,v)) < c((v,w))$ for some incident edges $(u,v),(v,w) \in E(T)$. A natural question is whether the prize and cost assignments p' and c' as in (1) and (2) will result in an improved SS as in Definition 3.3. In Example 3.1 we will see that that is not the case.

Convention: Let $T_p(\ell)$ denote the rooted tree whose underlying graph is a path on $2\ell + 1$ vertices $V(T_p(\ell)) = \{r, u_1, \ldots, u_{2\ell}\}$ and directed edges

$$E(T_p(\ell)) = \{(r,u_1),(r,u_2),(u_1,u_3),(u_2,u_4),\ldots,(u_{2\ell-3},u_{2\ell-1}),(u_{2\ell-2},u_{2\ell})\}$$

rooted at its center vertex. We label the edges by the same index as their heads: $e_1 = (r,u_1)$, $e_2 = (r,u_2)$,..., $e_{2\ell-1} = (u_{2\ell-3},u_{2\ell-1})$, and $e_{2\ell} = (u_{2\ell-2},u_{2\ell})$, see Figure 1.

Example 3.1.

Let $(T_p(3), c, p)$ be a SS for a CSM M where

$$\begin{aligned} c(e_1, e_2, e_3, e_4, e_5, e_6) &:= (1, 1, 1, 1, 1, 2), \\ p(u_1, u_2, u_3, u_4, u_5, u_6) &:= (10, 2, 10, 3, 10, 40), \end{aligned}$$

where the penetration costs and the prizes have been simultaneously assigned in the obvious way. We see that for any budget $B \in \mathbb{Q}_+$ we have

$$\mathrm{pr}^\star(B, c, p) = \begin{cases} 10\lfloor B \rfloor & \text{for } 0 \leq B < 4, \\ 10\lfloor B \rfloor + 5 & \text{for } 4 \leq B \leq 7, \\ 75 & \text{for } 7 < B. \end{cases}$$

If now $p'(u_1, u_2, u_3, u_4, u_5, u_6) = (10, 3, 10, 2, 10, 40)$ is the prize assignment obtained from p by swapping the prizes on the neighboring vertices u_2 and u_4, and $c'(e_1, e_2, e_3, e_4, e_5, e_6) = (1, 1, 1, 2, 1, 1)$ be the edge-cost assignment obtained from c be swapping the costs of the incident edges e_4 and e_6, then

$$\mathrm{pr}^\star(B, c, p') = \mathrm{pr}^\star(B, c', p) = \mathrm{pr}^\star(B, c, p),$$

for any non-negative budget $B \in \mathbb{Q}_+$, showing that locally swapping either prize assignments on adjacent vertices, or edge-costs on incident edges, does not necessarily improve the SS.

In Theorem 4.1 in Section 4, we show that there are CSMs for which no optimal SS exists. In such cases obtaining a locally optimal SS, as defined in Definition 3.6, may provide us with a reasonable defensive posture.

4 Optimal Security Systems

One of the most natural and important questions to consider for a given CSM M is whether an optimal SS exists and if it does, what it would look like. Unfortunately, Theorem 4.1 shows that there are small and simple CSMs for which no optimal SS exists. Still we would like to know for what CSMs optimal SSs do exist, and, if possible, have a way to find these optimal SSs efficiently. Corollary 4.1 and Theorem 4.2 show that optimal SSs exist for CSMs $M = (T, C, P)$ when T is a path or a star, respectively. These theorems also yield $O(n \log n)$ algorithms for producing optimal SSs in these cases. But, these results are not satisfying, as they are limited. In Sections 5, 6, and 7 we study *P-* and *C-models* and completely characterize the types of trees that have optimal SSs.

We begin with a lemma showing that all optimal SSs must have the highest penetration costs assigned to the edges involving the root and level-one vertices.

Lemma 4.1. *Let $M = (T, C, P)$ be a CSM, where T rooted at r contains at least one non-root vertex. Let $V_1 \subseteq T(V)$ denote the level-one vertices of T, and let C_L be the multiset of the largest $|V_1|$ values in C. If an optimal (T, c, p) SS for M, exists, then $c(e) \in C_L$ for $e \in \{(r, v) \mid v \in V_1\}$.*

Proof. Suppose we have an optimal SS (T, c, p) that does not meet the conditions of the lemma. Let $c_s \notin C_L$ be the smallest penetration cost assigned by c to an edge between the root r and a vertex $v_s \in V_1$, that is, $c((r, v_s)) = c_s \leq c((r, v))$ for all $v \in V_1 - \{v_s\}$. Let $e_s = (r, v_s)$ and let e_l be an edge not between the root and a level-one vertex where $c(e_l) \in C_L$. We know that such an edge exists because (T, c, p) does not meet the conditions of the lemma. To show that (T, c, p) cannot be an optimal SS, we define a SS (T, c', p) by letting $c'(e_s) = c(e_l)$, $c'(e_l) = c(e_s)$, and $c'(e) = c(e)$ otherwise. Notice that for the budget $B = c_s$, we have $\mathrm{pr}^\star(B, c, p) = p(v_s) > 0 = \mathrm{pr}^\star(B, c', p)$. This fact contradicts that (T, c, p) is an optimal SS. □

If an optimal SS exists, Lemma 4.1 tells us something about its form. In the next theorem we show that there are CSMs for which no optimal SS exists.

Theorem 4.1. *There is a CSM $M = (T, C, P)$ for which no optimal security system exists.*

Proof. Consider $M = (T, \{1, 2, 3\}, \{1, 2, 3\})$, Where T is the tree given by $V(T) = \{r, u_1, u_2, u_3\}$ and $E(T) = \{e_1, e_2, e_3\}$ where $e_1 = (r, u_1)$, $e_2 = (r, u_2)$, and $e_3 = (u_1, u_3)$. By Lemma 4.1 we know that an optimal SS (T, c, p) has $c(e_3) = 1$, and we can further assume that $p(u_3) = 3$. By considering the budget of $B = 2$, we can also assume the prize of the head of the edge of cost 2 to by 1. Therefore, we have only two possible optimal SSs for M: (T, c, p) with $c(e_1, e_2, e_3) = (3, 2, 1)$ and $p(u_1, u_2, u_3) = (2, 1, 3)$, or (T, c', p') with $c'(e_1, e_2, e_3) = (2, 3, 1)$ and $p'(u_1, u_2, u_3) = (1, 2, 3)$, see Figure 2. Since $\mathrm{pr}^\star(3, c, p) = 2$ and $\mathrm{pr}^\star(3, c', p') = 4$, we see that (T, c', p') is not optimal, and since $\mathrm{pr}^\star(4, c, p) = 5$ and $\mathrm{pr}^\star(4, c', p') = 4$, we see that (T, c, p) is not optimal either. Hence, no optimal SS for M exists. □

Although Theorem 4.1 showed that there are CSMs for which no optimal SS exists, we are interested in finding out for which trees T optimal SSs do exist. We should point out that the values of the weights in C and P also play an important role in whether or not an optimal SS exists for a given tree. In the next theorem we show that an optimal SS exists for CSMs in which the tree in the model is a path, and this result is independent of the values of the weights in C and P.

Consider a CSM $M = (T, C, M)$ where T is a path rooted at a leaf, so

$$V(T) = \{u_0, u_1, \ldots, u_n\}, \quad E(T) = \{e_1, \ldots, e_n\}, \tag{3}$$

where $u_0 = r$ and $e_i = (u_{i-1}, u_i)$, for each $i \in \{1, \ldots, n\}$. For a SS (T, c, p) for M, then for convenience let $p_i = p(u_i)$ and $c_i = c(e_i)$ for each i. If we have $p_i \leq p_{i+1}$ and $c_i \geq c_{i+1}$ for each $i \in \{1, \ldots, n-1\}$ (so the prizes are ordered increasingly and the edge-costs decreasingly as we go down the path from the root), then by Proposition 3.1 the SS (T, c, p) is a good SS as in Definition 3.6. But, we can say slightly more here when T is a path, in terms of obtaining an improved SS as in Definition 3.3.

Lemma 4.2. *Let $M = (T, C, M)$ be a CSM where T is a path with its vertices and edges labeled as in (3).*

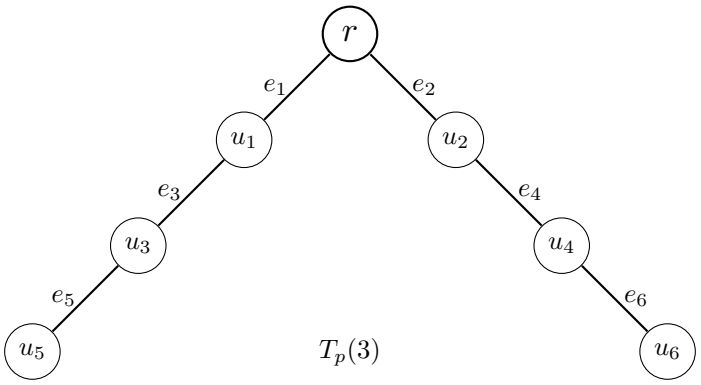


Figure 1: $T_p(3)$ is a path on seven vertices rooted at its center.

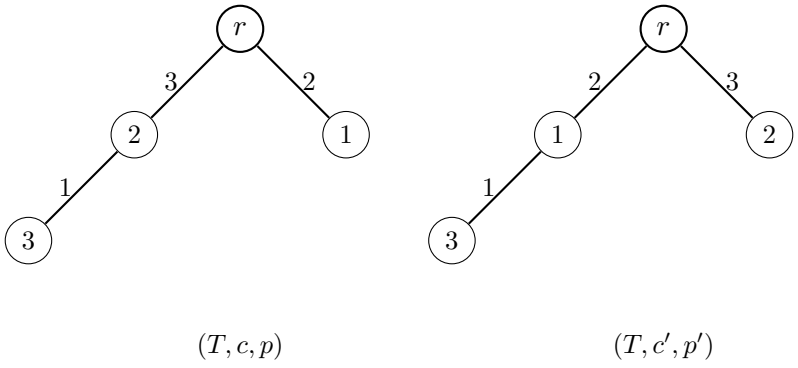


Figure 2: Only two possible SSs for $M = (T, \{1, 2, 3\}, \{1, 2, 3\})$.

(i) If (T, c, p) is a SS for M and there is an i with $p_i > p_{i+1}$ and $c_{i+1} > 0$, then the SS (T, c, p') where p' is obtained by swapping the prizes on u_i and u_{i+1} is an improved SS.

(ii) If (T, c, p) is a SS for M and there is an i with $c_i < c_{i+1}$, then the SS (T, c', p) where c' is obtained by swapping the edges costs on e_i and e_{i+1} is an improved SS.

Proof. By Proposition 3.1 we only need to show (i) there is a budget B' such that $\mathrm{pr}^\star(B', c, p') < \mathrm{pr}^\star(B', c, p)$ and (ii) a budget B'' such that $\mathrm{pr}^\star(B'', c', p) < \mathrm{pr}^\star(B'', c, p)$. For each j let $\tau_j = T[e_1, \ldots, e_j]$ be the rooted sub-path of T that contains the first j edges of T.

For $B' = c_1 + \cdots + c_i$ we clearly have

$$\begin{aligned} \mathrm{pr}^\star(B', c, p') &= \mathrm{pr}(\tau_i, c, p') \\ &= p_1 + \cdots + p_{i-1} + p_{i+1} \\ &< p_1 + \cdots + p_i \\ &= \mathrm{pr}(\tau_i, c, p) \\ &= \mathrm{pr}^\star(B', c, p), \end{aligned}$$

showing that (T, c, p') is an improved SS for M.

Likewise, we have

$$\begin{aligned} \mathrm{pr}^\star(B', c', p) &= \mathrm{pr}(\tau_{i-1}, c', p) \\ &= p_1 + \cdots + p_{i-1} \\ &< p_1 + \cdots + p_i \\ &= \mathrm{pr}(\tau_i, c, p) \\ &= \mathrm{pr}^\star(B', c, p), \end{aligned}$$

showing that (T, c', p) is also an improved SS for M. □

Given any SS (T, c, p) for M as in Lemma 4.2 when T is a rooted path, by bubble sorting the prizes and the edge costs increasingly and decreasingly respectively, as we go down the path T from the root, we obtain by Lemma 4.2 a SS (T, c', p') such that for any budget B we have $\mathrm{pr}^\star(B, c', p') \leq \mathrm{pr}^\star(B, c, p)$. We therefore have the following corollary.

Corollary 4.1. *If $M = (T, C, M)$ is a CSM where T is a rooted path with its vertices and edges labeled as in (3), then there is an optimal SS for M, and it is given by assigning the penetration costs to the edges and the prizes to the vertices in a decreasing order and increasing order respectively from the root.*

We now show that an optimal SS exists for $M = (T, C, P)$ when T is a star. Let T be a star with root r and non-root vertices $u_1, \ldots, n_n$ and edges $e_i = (r, u_i)$ for $i = 1, \ldots, n$. Suppose the costs and prizes are given by $C = \{c_1, \ldots, c_n\}$ and $P = \{p_1, \ldots, p_n\}$. When considering an arbitrary security system (T, c, p) where $c(u_i) = c_i$ and $p(e_i) = p_i$ for each i, we can without loss of generality assume the edge-costs to be in an increasing order $c_1 \leq \cdots \leq c_n$.

Lemma 4.3. *Suppose T is a star and (T, c, p) is a SS as above. If p' is another prize assignment obtained from p by swapping the prizes p_i and p_j where $i < j$ and $p_i \leq p_j$, then for any budget B we have* $\mathrm{pr}^\star(B, c, p) \leq \mathrm{pr}^\star(B, c, p')$.

Proof. Let B be a given budget and $\tau \subseteq T$ an optimal attack with respect to p, so $\mathrm{pr}(\tau, c, p) = \mathrm{pr}^\star(B, c, p)$. We consider the following cases.

Case one: If both of u_i and u_j are in τ, or neither of them are, then we haef $\mathrm{pr}^\star(B, c, p) = \mathrm{pr}(\tau, c, p) = \mathrm{pr}(\tau, c, p') \leq \mathrm{pr}^\star(B, c, p')$.

Case two: If $u_i \in V(\tau)$ and $u_j \notin V(\tau)$, then $\mathrm{pr}^\star(B, c, p) = \mathrm{pr}(\tau, c, p) \leq \mathrm{pr}(\tau, c, p) - p_i + p_j = \mathrm{pr}(\tau, c, p') \leq \mathrm{pr}^\star(B, c, p')$.

Case three: If $u_i \notin V(\tau)$ and $u_j \in V(\tau)$, then $\tau' = (\tau - u_j) \cup u_i$ is a rooted subtree of T with $c(\tau') = c(\tau) - c_j + c_i \leq B$ and is therefore within the budget B. Hence, $\mathrm{pr}^\star(B, c, p) = \mathrm{pr}(\tau, c, p) = \mathrm{pr}(\tau', c, p') \leq \mathrm{pr}^\star(B, c, p')$.

Therefore, in all cases we have $\mathrm{pr}^\star(p, c, B) \leq \mathrm{pr}^\star(p', c, B)$. □

Since any permutation is a composition of transpositions, we have the following theorem as a corollary.

Theorem 4.2. *Let $M = (T, C, P)$ be a CSM where T is a star rooted at its center vertex. Then there is an optimal SS for M, and it is given by assigning the prizes to the vertices in the same increasing order as the costs are assigned increasingly to the corresponding edges.*

For rooted trees on n non-root vertices, Corollary 4.1 and Theorem 4.2 give rise to natural sorting-based $O(n \log n)$ algorithms for computing optimal SSs. Notice that in an optimal SS in a general tree, the smallest prize overall must be assigned to a level-one vertex u which has the largest penetration cost assigned to its corresponding edge, (r, u), to the root. And, furthermore, we cannot say more than this statement for arbitrary trees as the next assignment of a prize will depend on the relative values of the penetration costs, prizes, and structure of the tree. In view of the fact that optimal SSs do not exist, except for paths and stars as we will see shortly in Observation 5.1, we turn our attention to restricted CSMs and classify them with respect to optimal SSs.

5 Specific Security Systems, P-Models, and C-Models

In this section we extend CSMs to include penetration costs and prizes of value zero. For a CSM $M = (T, C, P)$ with no optimal SS and a rooted super-tree $T^\dagger$ of which T is a rooted subtree, we can always assign the prize of zero to the nodes in $V(T^\dagger) \setminus V(T)$ and likewise the penetration cost of zero to the edges in $E(T^\dagger) \setminus E(T)$, thereby obtaining a CSM $M^\dagger = (T^\dagger, C^\dagger, E^\dagger)$ that also has no optimal SS. Note that if T is the rooted tree in the proof of Theorem 4.1, then the only rooted trees that do not have T as a rooted subtrees are paths rooted at one of their leaves or stars rooted at their center vertices. Hence, by the example provided in the proof of Theorem 4.1, we have the following observation.

Observation 5.1. *If T is a rooted tree, such that for any multisets C and P of penetration costs and prizes, respectively, the CSM $M = (T, C, P)$ has an optimal SS, then T is either a path rooted at one of its leaves, or a star rooted at its center vertex.*

In light of Observation 5.1, we seek some natural restrictions on our CSM M that will guarantee it having an optimal SS. Since both the penetration costs and the prizes of $M = (T, C, P)$ take values in $\mathbb{Q}_+$ we can, by an appropriate scaling, obtain an equivalent CSM where both the costs and prizes take values in $\mathbb{N} \cup \{0\}$, that is, we may assume $c(e) \in \mathbb{N} \cup \{0\}$ and $p(u) \in \mathbb{N} \cup \{0\}$ for every $e \in E(T)$ and $u \in V(T)$, respectively.

First, we consider the restriction on a CSM $M = (T, C, P)$ where C consists of a single penetration-cost value, that is, $C = \{1, 1, \ldots, 1\}$ consists of n copies of the unit penetration cost one. From a realistic point of view, this assumption seems to be reasonable; many computer networks consist of computers with similar password/encryption security systems on each computer (that is, the penetration cost is the same for all of the computers), whereas the computers might store data of vastly distinct values (that is, the prizes are distinct).

Convention: In what follows, it will be convenient to denote the multiset containing n (or an arbitrary number of) copies of 1 by I. In a similar way, we will denote by $\mathbf{1}$ the map that maps each element of the appropriate domain to 1. As the domain of $\mathbf{1}$ should be self-evident each time, there should be no ambiguity about it each time.

Definition 5.1. *A* P-model *is a CSM $M = (T, I, P)$ where T has n non-root vertices and where I is constant, consisting of n copies of the unit penetration cost.*

Consider a SS (T, c, p) of a CSM $M = (T, C, P)$. We can obtain an *equivalent* SS $(T', \mathbf{1}, p')$ of a P-model $M' = (T', I, P')$ in the following way: for each edge $e = (u, v) \in E(T)$ with penetration cost $c(e) = k \in \mathbb{N}$ and prizes $p(u), p(v) \in \mathbb{N}$ of its head and tail, respectively, replace the 1-path (u, e, v) with a directed path of new vertices and edges $(u, e_1, u_1, e_2, u_2, \ldots, u_{k-1}, e_k, v)$ of length k. We extend the penetration cost and prize functions by adding zero-prize vertices where needed, that is, $\mathbf{1}(f) = 1$ for each $f \in E(T')$, and we let

$$p'(u) = p(u), \quad p'(v) = p(v), \text{ and } p'(u_1) = p'(u_2) = \cdots = p'(u_{k-1}) = 0.$$

In this way we obtain a SS (T', c', p') of a P-model $M' = (T', I, P')$. We view the vertices $V(T)$ of positive prize as a subset of $V(T')$ (namely, those vertices of T' with positive prize).[1]

Recall that T is a *rooted contraction* of T' if T is obtained from T' by a sequence of simple contractions of edges, and where any vertex contracted into the root remains the root. This means precisely that T is a rooted *minor* of T' [2, p. 54].

[1] Note that there are some redundant definitions on the prizes of the vertices when considering incident edges, but the assignments do agree, as they have the same prize values as in T.

Proposition 5.1. *Any SS (T, c, p) of a CSM $M = (T, C, P)$ is equivalent to a SS $(T', \mathbf{1}, p')$ of a P-model $M' = (T', I, P')$ where (i) T is rooted minor of T', and (ii) $p'(u) = p(u)$ for each $u \in V(T) \subseteq V(T')$, and $p'(u) = 0$, otherwise.*

Proof. (Sketch) Given a budget $B \in \mathbb{Q}_+$, clearly any optimal attack τ on a SS (T, c, p) with $\mathrm{pr}(\tau, c, p) = \mathrm{pr}^\star(B, c, p)$ has an equivalent attack τ' on a SS $(T', \mathbf{1}, p')$ of the same cost $\mathrm{cst}(\tau', \mathbf{1}, p') = \mathrm{cst}(\tau, c, p)$ and hence within the budget B, where τ' is the smallest subtree of T' that contains all of the vertices of τ. By construction, we also have that $\mathrm{pr}(\tau', \mathbf{1}, p') = \mathrm{pr}(\tau, c, p) = \mathrm{pr}^\star(B, c, p)$ since all of the vertices from τ are in τ' and have the same prize there, and the other vertices in τ' have prize zero. This shows that $\mathrm{pr}^\star(B, c, p) \leq \mathrm{pr}^\star(B, \mathbf{1}, p')$.

Conversely, an optimal attack τ' on $(T', \mathbf{1}, p')$ with $\mathrm{pr}(\tau', \mathbf{1}, p') = \mathrm{pr}^\star(B, \mathbf{1}, p')$ yields an attack τ on (T, c, p) by letting τ be the subtree of T induced by the vertices $V(\tau') \cap V(T)$. In this way $\mathrm{pr}(\tau, c, p) = \mathrm{pr}(\tau', \mathbf{1}, p')$ and $\mathrm{cst}(\tau, c, p) \leq \mathrm{cst}(\tau', \mathbf{1}, p')$, since some of the vertices of τ' might have zero prize, as they are not in τ. By definition of $\mathrm{pr}^\star(\cdot)$ we have that $\mathrm{pr}^\star(B, \mathbf{1}, p') \leq \mathrm{pr}^\star(B, c, p)$. Hence, the SS (T, c, p) and $(T', \mathbf{1}, p')$ are equivalent. □

Secondly, and dually, we can restrict our attention to the case where the multiset of prizes P consists of a single unit prize value, so $P = I = \{1, 1, \ldots, 1\}$ consists of n copies of the unit prize.

Definition 5.2. *A* C-model *is a CSM $M = (T, C, I)$, where T has n non-root vertices and where I is constant, consisting of n copies of the unit prize.*

As before, consider a SS (T, c, p) of a CSM $M = (T, C, P)$. We can obtain an equivalent SS $(T'', c'', \mathbf{1})$ of a C-model $M'' = (T'', C'', I)$ in the following way: for each edge $e = (u, v) \in E(T)$ with penetration cost $c(e) = k \in \mathbb{N}$ and prizes $p(u), p(v) \in \mathbb{N}$ of its head and tail, respectively, replace the 1-path (u, e, v) with a directed path of new vertices and edges $(u, e, u_1, e_1, u_2, \ldots, u_{k-1}, e_{k-1}, v)$ of length k. We extend the penetration cost and prize functions by adding zero-cost edges where needed, that is, $\mathbf{1}(w) = 1$ for every $w \in V(T'')$, and we let

$$c''(e) = c(e) \quad \text{and } c''(e_1) = c''(e_2) = \cdots = c''(e_{k-1}) = 0.$$

In this way we obtain a SS $(T'', c'', \mathbf{1})$ of a C-model $M'' = (T'', C'', I)$, where the multiset of prizes consists of a single unit prize value ($\sum_{u \in V(T) \setminus \{r\}} p(u)$ copies of it). We also view the edges $E(T)$ of positive penetration cost as a subset of $E(T'')$ (namely, those edges of T'' with positive penetration cost). We also have the following proposition that is dual to Proposition 5.1.

Proposition 5.2. *Any SS (T, c, p) of a CSM $M = (T, C, P)$ is equivalent to a SS $(T'', c'', \mathbf{1})$ of a C-model $M'' = (T'', C'', I)$, where (i) T is rooted minor of T'', and (ii) $c''(e) = c(e)$ for each $e \in E(T) \subseteq E(T'')$, and $c''(e) = 0$, otherwise.*

Proof. (Sketch) Suppose we are given a budget $B \in \mathbb{Q}_+$ and an optimal attack τ on a SS (T, c, p) with $\mathrm{pr}(\tau, c, p) = \mathrm{pr}^\star(B, c, p)$. Here $(T'', c'', \mathbf{1})$ has an equivalent attack

τ'', where τ'' is the largest subtree of T'' that contains all of the edges of τ and no other edges of T. Note that $\mathrm{cst}(\tau'', c'', \mathbf{1}) = \mathrm{cst}(\tau, c, p)$ since all of the additional edges of τ'' that are not in $V(\tau)$ have zero penetration cost, and so τ'' is within the budget B. Also, by construction we have $\mathrm{pr}(\tau'', c'', \mathbf{1}) = \mathrm{pr}(\tau, c, p) = \mathrm{pr}^\star(B, c, p)$. This result shows that $\mathrm{pr}^\star(B, c, p) \leq \mathrm{pr}^\star(B, c'', \mathbf{1})$.

Conversely, consider an optimal attack τ'' on $(T'', c'', \mathbf{1})$ with $\mathrm{pr}(\tau'', c'', \mathbf{1}) = \mathrm{pr}^\star(B, c'', \mathbf{1})$. By the optimality of τ'', every leaf of τ'' is a tail of an edge of T, since otherwise we can append that edge (of zero penetration cost), and thereby obtain an attack with a prize strictly more than $\mathrm{pr}(\tau'', c'', \mathbf{1})$, a contradiction. The edges $E(\tau'') \cap E(T)$ induce a subtree τ of T of the same cost $\mathrm{cst}(\tau, c, p) = \mathrm{cst}(\tau'', c'', \mathbf{1})$; and moreover, τ'' is, by its optimality, the largest subtree of T'' that contains exactly all of the edges of τ, and so $\mathrm{pr}(\tau, c, p) = \mathrm{pr}(\tau'', c'', \mathbf{1}) = \mathrm{pr}^\star(B, c'', \mathbf{1})$. This result shows that $\mathrm{pr}^\star(B, c'', \mathbf{1}) \leq \mathrm{pr}^\star(B, c, p)$. This proves that the SS (T, c, p) and $(T'', c'', \mathbf{1})$ are equivalent. □

We now present some examples of both P- and C-models that will play a pivotal role in our discussion to come.

Definition 5.3. *Let $T(2)$ denote the rooted tree given as follows:*

$$\begin{aligned} V(T(2)) &= \{r, u_1, u_2, u_3, u_4, u_5\}, \\ E(T(2)) &= \{(r, u_1), (r, u_2), (u_1, u_3), (u_2, u_4), (u_2, u_5)\}. \end{aligned}$$

Note that $T(2)$ has all of its non-root vertices on two non-zero levels. Similarly, let $T(3)$ denote the rooted tree given as follows:

$$\begin{aligned} V(T(3)) &= \{r, u_1, u_2, u_3, u_4\}, \\ E(T(3)) &= \{(r, u_1), (r, u_2), (u_2, u_3), (u_3, u_4)\}. \end{aligned}$$

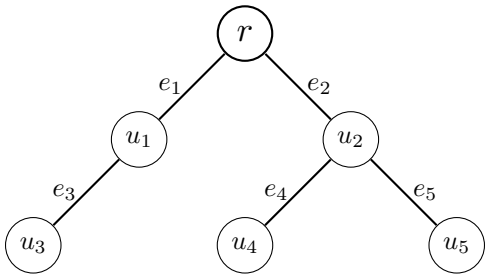
Note that $T(3)$ has all of its vertices on three non-zero levels.

Convention: For convenience we label the edges of both $T(2)$ and $T(3)$ with the same index as their heads (see Figures 3 and 4):

$$\begin{aligned} T(2) &: e_1 = (r, u_1),\ e_2 = (r, u_2),\ e_3 = (u_1, u_3),\ e_4 = (u_2, u_4),\ e_5 = (u_2, u_5). \\ T(3) &: e_1 = (r, u_1),\ e_2 = (r, u_2),\ e_3 = (u_1, u_3),\ e_4 = (u_3, u_4). \end{aligned}$$

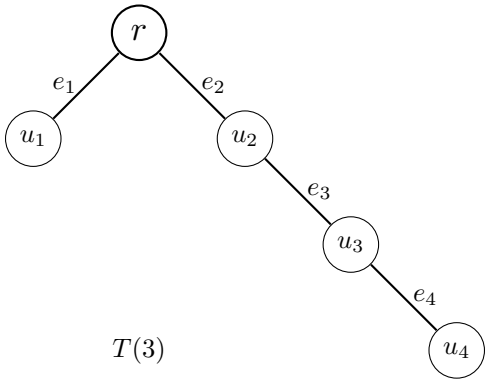
Example 5.1.
Consider a P-model (with $c = \mathbf{1}$) on the rooted tree $T(2)$, where the prize values are given by $P = \{0, 1, 2, 2, 3\}$.

Prize Assignment (A): Consider the case where the prizes have been simultaneously assigned to the non-root vertices of $T(2)$ by $p(u_1, u_2, u_3, u_4, u_5) := (0, 1, 3, 2, 2)$ in the obvious way. We will use a similar shorthand notation later for the bijection c. In this case we see that for budgets of $B = 2, 3$, we have $\mathrm{pr}^\star(2, \mathbf{1}, p) = 3$ and $\mathrm{pr}^\star(3, \mathbf{1}, p) = 5$, respectively.



$T(2)$

Figure 3: $T(2)$ has all of its non-root vertices on two non-zero levels.



$T(3)$

Figure 4: $T(3)$ has all of its non-root vertices on three non-zero levels.

Prize Assignment (B): Consider now the case where the prizes have been simultaneously assigned to the non-root vertices of $T(2)$ by $p'(u_1, u_2, u_3, u_4, u_5) := (1, 0, 3, 2, 2)$. In this case we see that for the same budgets of $B = 2, 3$ as in (A), we have $\mathrm{pr}^\star(2, \mathbf{1}, p') = 4$ and $\mathrm{pr}^\star(3, \mathbf{1}, p') = 4$, respectively.

From these assignments we see that for budget $B = 2$, the SS in (A) is better than the one in (B), and for $B = 3$, the SS in (B) is better than the one in (A).

Example 5.2.

Consider a P-model on the rooted tree $T(3)$, where the prize values are given by $P = \{0, 0, 1, 1\}$.

Prize Assignment (A): Consider the case where the prizes have been simultaneously assigned to the non-root vertices of $T(3)$ by $p(u_1, u_2, u_3, u_4) := (0, 0, 1, 1)$. In this case we see that for budgets of $B = 1, 3$, we have $\mathrm{pr}^\star(1, \mathbf{1}, p) = 0$ and $\mathrm{pr}^\star(3, \mathbf{1}, p) = 2$, respectively.

Prize Assignment (B): Consider now the case where the prizes have been simultaneously assigned to the non-root vertices of $T(3)$ by $p'(u_1, u_2, u_3, u_4) := (1, 0, 0, 1)$. In this case we see that for the same budgets of $B = 1, 3$ as in (A), we have $\mathrm{pr}^\star(1, \mathbf{1}, p') = 1$ and $\mathrm{pr}^\star(3, \mathbf{1}, p') = 1$, respectively.

From these assignments we see that for budget $B = 1$, the SS in (A) is better than the one in (B), and for $B = 3$, the SS in (B) is better than the one in (A).

Considering the budget $B = 1$ for the P-model in Example 5.1, we see that in order for a prize assignment to be optimal we must have the prizes of u_1 and u_2 to be 0 and 1. Considering further $B = 2$ we see that an optimal prize assignment in this case must be p or p' as in Example 5.1, or p'' where $p''(u_1, u_2, u_3, u_4, u_5) := (1, 0, 2, 3, 2)$. Since $\mathrm{pr}^\star(B, \mathbf{1}, p'') = \mathrm{pr}^\star(B, \mathbf{1}, p)$ for any B, we see that the P-model in Example 5.1 has no optimal SS. As the P-model in Example 5.2 can be analysed in the same way, we have the following observation.

Observation 5.2. *For general prize values P, neither of the P-models $M = (T(2), I, P)$ nor $M = (T(3), I, P)$ have optimal SSs.*

We will now consider the dual cases of the C-models.

Example 5.3.

Consider a C-model (with $p = \mathbf{1}$) on the rooted tree $T(2)$, where the penetration costs are given by $C = \{0, 1, 1, 2, 3\}$.

Cost Assignment (A): Consider the case where the penetration costs have been simultaneously assigned to the edges of $T(2)$ by $c(e_1, e_2, e_3, e_4, e_5) := (3, 2, 0, 1, 1)$. In this case we see that for budgets of $B = 2, 4$, we have $\mathrm{pr}^\star(2, c, \mathbf{1}) = 1$ and $\mathrm{pr}^\star(4, c, \mathbf{1}) = 3$, respectively.

Cost Assignment (B): Consider now the case where the penetration costs have been assigned to the edges of $T(2)$ by $c'(e_1, e_2, e_3, e_4, e_5) := (2, 3, 0, 1, 1)$. In this case we see that for the same budgets of $B = 2, 4$ as in (A), we have $\mathrm{pr}^\star(2, c', \mathbf{1}) = 2$ and $\mathrm{pr}^\star(4, c', \mathbf{1}) = 2$, respectively.

From these assignments we see that for budget $B = 2$, the SS in (A) is better than the one in (B), and for $B = 4$, the SS in (B) is better than the one in (A).

Example 5.4.

Consider now a C-model on the rooted tree $T(3)$, where the penetration costs are given by $C = \{0, 0, 1, 1\}$.

Cost Assignment (A): Consider the case where the penetration costs have been simultaneously assigned to the edges of $T(3)$ by $c(e_1, e_2, e_3, e_4) := (1, 1, 0, 0)$. In this case we see that for budgets of $B = 0, 1$, we have $\mathrm{pr}^\star(0, c, \mathbf{1}) = 0$ and $\mathrm{pr}^\star(1, c, \mathbf{1}) = 3$, respectively.

Cost Assignment (B): Consider now the case where the penetration costs have been assigned to the edges of $T(3)$ by $c'(e_1, e_2, e_3, e_4) := (0, 1, 1, 0)$. In this case we see that for the same budgets of $B = 0, 1$ as in (A), we have $\mathrm{pr}^\star(0, c', \mathbf{1}) = 1$ and $\mathrm{pr}^\star(1, c', \mathbf{1}) = 2$, respectively.

From these assignments we see that for budget $B = 0$, the SS in (A) is better than the one in (B), and for $B = 1$, the SS in (B) is better than the one in (A).

In a similar way as we obtained Observation 5.2, we see from the previous two examples the following.

Observation 5.3. *For general penetration costs C, neither of the C-models $M = (T(2), C, I)$ nor $M = (T(3), C, I)$ have optimal SSs.*

Remark 5.1. (i) Note that in Examples 5.1 and 5.3 involving the rooted tree $T(2)$, we have that the prize assignments to the non-root vertices and cost assignments to the corresponding edges sum up to a constant vector for both assignments *(A)* and *(B)*:

$$\begin{aligned}
(A) \quad &: \quad p(u_1, u_2, u_3, u_4, u_5) + c(e_1, e_2, e_3, e_4, e_5) \\
&\quad\; = (0, 1, 3, 2, 2) + (3, 2, 0, 1, 1) = (3, 3, 3, 3, 3), \\
(B) \quad &: \quad p'(u_1, u_2, u_3, u_4, u_5) + c'(e_1, e_2, e_3, e_4, e_5) \\
&\quad\; = (1, 0, 3, 2, 2) + (2, 3, 0, 1, 1) = (3, 3, 3, 3, 3),
\end{aligned}$$

and similarly for the rooted tree $T(3)$:

$$\begin{aligned}
(A) \quad &: \quad p(u_1, u_2, u_3, u_4) + c(e_1, e_2, e_3, e_4) = (0, 0, 1, 1) + (1, 1, 0, 0) = (1, 1, 1, 1), \\
(B) \quad &: \quad p'(u_1, u_2, u_3, u_4) + c'(e_1, e_2, e_3, e_4) = (1, 0, 0, 1) + (0, 1, 1, 0) = (1, 1, 1, 1).
\end{aligned}$$

This duality is not a coincidence and will discussed in more detail in Section 7. (ii) Although special cases of Theorems 6.1, 6.2, 7.3 and 7.4, it is an easy combinatorial exercise to see that both a C- or P-model $M = (T, C, P)$, where T is a proper rooted subtree of either $T(2)$ or $T(3)$ does indeed have an optimal SS, and so $T(2)$ and $T(3)$ are the smallest rooted trees, in either model, with no optimal SS. This point will also be discussed and stated explicitly in Sections 6 and 7.

Consider now a given rooted tree T and another rooted tree $T^\dagger$ containing T as a rooted subtree, so $T \subseteq T^\dagger$. Assume that the P-model $M = (T, I, P)$ has no optimal SS. Extend M to a P-model on $T^\dagger$ by adding a zero prize for each vertex in $V(T^\dagger) \setminus V(T)$, so $P^\dagger = P \cup Z$, where Z is the multiset consisting of $|V(T^\dagger)| - |V(T)|$ copies of 0. In this case we have the following.

Observation 5.4. *If $M = (T, I, P)$ is a P-model with no optimal SS, and $T^\dagger$ contains T as a rooted subtree, then the P-model $M^\dagger = (T^\dagger, I, P^\dagger)$ has no optimal SS.*

Proof. (Sketch) For any budget consisting of $B = m$ edges and a SS $(T, \mathbf{1}, p)$, there is a rooted subtree τ of T with m edges such that $\mathrm{pr}(\tau, \mathbf{1}, p) = \mathrm{pr}^\star(m, \mathbf{1}, p)$. Let $\mathbf{1}$ and $p^\dagger$ be the obvious extensions of $\mathbf{1}$ and p to $T^\dagger$, by letting $\mathbf{1}(e) = 1$ for all $e \in E(T^\dagger)$ and $p^\dagger(u) = 0$ for any $u \in V(T^\dagger) \setminus V(T)$. If τ' is a rooted subtree of $T^\dagger$ with m edges, then $\tau' \cap T$ is a rooted subtree of both T and $T^\dagger$ on m or fewer edges. Since any vertex of $V(\tau') \setminus V(T)$ has zero prize, we have

$$\mathrm{pr}(\tau', \mathbf{1}, p^\dagger) = \mathrm{pr}(\tau' \cap T, \mathbf{1}, p^\dagger) = \mathrm{pr}(\tau' \cap T, \mathbf{1}, p) \leq \mathrm{pr}^\star(m, \mathbf{1}, p),$$

with equality for $\tau' = \tau$ since $\tau \subseteq T \subseteq T^\dagger$. Hence, $\mathrm{pr}^\star(m, \mathbf{1}, p^\dagger) = \mathrm{pr}^\star(m, \mathbf{1}, p)$, and we conclude that if $M = (T, I, P)$ has no optimal SS, then neither does $M^\dagger = (T^\dagger, I, P^\dagger)$. □

Dually, assume that we have a C-model $M = (T, C, I)$ that has no optimal SS, and similarly, let $T^\dagger$ be a rooted subtree containing T as a rooted subtree. Extend M to a C-model on $T^\dagger$ by adding penetration costs of ∞[2] for each edge of $T^\dagger$ that is not in T, so $C^\dagger = C \cup Y$, where Y is the multiset consisting of $|E(T^\dagger)| - |E(T)|$ copies of ∞.

Observation 5.5. *If $M = (T, C, I)$ is a C-model with no optimal SS, and $T^\dagger$ contains T as a rooted subtree, then the C-model $M^\dagger = (T^\dagger, C^\dagger, I)$ has no optimal SS.*

Proof. (Sketch) The proof is similar to the one for Observation 5.4. For any budget $B \in \mathbb{Q}_+$ and a SS $(T, c, \mathbf{1})$ of M, there is a rooted subtree τ of T with m edges such that $\mathrm{pr}(\tau, c, \mathbf{1}) = \mathrm{pr}^\star(B, c, \mathbf{1})$. Let $c^\dagger$ and $\mathbf{1}$ be the obvious extensions of c and $\mathbf{1}$ to $T^\dagger$, by letting $c^\dagger(e) = \infty$ for all $e \in E(T^\dagger) \setminus E(T)$. If τ' is a rooted subtree of $T^\dagger$ within the attacker's budget of $B < \infty$, then every edge of τ' must be in T, and so $\tau' \subseteq T \subseteq T^\dagger$. Since $c^\dagger$ agrees with c on the edges of T we have

$$\mathrm{pr}(\tau', c^\dagger, \mathbf{1}) = \mathrm{pr}(\tau', c, \mathbf{1}) \leq \mathrm{pr}^\star(B, c, \mathbf{1}),$$

with equality for $\tau' = \tau$. Hence, $\mathrm{pr}^\star(B, c^\dagger, \mathbf{1}) = \mathrm{pr}^\star(B, c, \mathbf{1})$, and we conclude that if $M = (T, C, I)$ has no SS, then neither does $M^\dagger = (T^\dagger, C^\dagger, I)$. □

By Observations 5.2, 5.3, 5.4, and 5.5 we have the following corollary.

[2]Where here we can choose ∞ to be the number of edges of T plus one, that is, a large number exceeding any sensible attack budget.

Corollary 5.1. *If T is a rooted tree such that any P- or C-model $M = (T, C, P)$ has an optimal SS, then T contains neither $T(2)$ nor $T(3)$ as rooted subtrees.*

Let T be a rooted tree such that any CSM $M = (T, C, P)$ has an optimal SS. Assume further that T is not a path rooted at one of its two leaves. If T has at least three non-zero levels (we consider the root r to be the unique level-0 vertex), then T must contain $T(3)$ as a rooted subtree and hence, by Corollary 5.1, there is a CSM $M = (T, C, P)$ with no optimal SS, contradicting our assumption on T. Consequently, T has at most two non-zero levels.

If T has at most two non-zero levels, and it has two leaves of distance four apart (with the root r being midways between them), then neither parent of the leaves is of degree three or more, because then T has $T(2)$ as a rooted subtree. And, so again, by Corollary 5.1, there is a CSM $M = (T, C, P)$ with no optimal SS. This observation again contradicts our assumption on T. As a result, either (i) T has a diameter of three and is obtained by attaching an arbitrary number of leaves to the end vertices of a single edge and then rooting it at one of the end-vertices of the edge, or (ii) T has diameter of four and each level-one vertex has degree at most two.

Recall that a *caterpillar tree* is a tree where each vertex is within distance one of a central path, and that a *spider tree* is a tree with one vertex of degree at least three and all other vertices of degree at most two.

Definition 5.4. *A* rooted path *is a path rooted at one of its two leaves.*

A rooted star *is a star rooted at its unique center vertex.*

A 3-caterpillar *is a caterpillar tree of diameter three.*

A rooted 3-caterpillar *is a 3-caterpillar rooted at one of its two center vertices.*

A 4-spider *is a spider tree of diameter four with its unique center vertex of degree at least three.*

A rooted 4-spider *is a 4-spider rooted at its unique center vertex.*

By Corollary 5.1 and the discussion just before Definition 5.4, we therefore have the following main theorem of this section.

Theorem 5.1. *If T is a rooted tree such that any P- or C-model $M = (T, C, P)$ has an optimal SS, then T is one of the following types: (i) a rooted path, (ii) a rooted star, (iii) a rooted 3-caterpillar, or (iv) a rooted 4-spider.*

It remains to be seen whether or not a rooted 3-caterpillar or a rooted 4-spider T is such that any P- or C-model $M = (T, C, P)$ has an optimal SS. This item will be the main topic of the next two sections.

6 P-models with Optimal Security Systems

In this section we prove that if T is one of the four types of rooted trees mentioned in Theorem 5.1, then any P-model $M = (T, I, P)$ indeed has an optimal SS. The

C-models will be discussed in Section 7. We already have that any P-model $M = (T, I, P)$ (in fact, any CSM $M = (T, C, P)$), where T is a rooted path or a rooted star, does have an optimal SS, so it suffices to consider rooted 3-caterpillars and rooted 4-spiders.

Let T be a rooted 3-caterpillar on vertices $\{r, u_1, \ldots, u_n\}$ with edges given by

$$E(T) = \{(r, u_1), \ldots, (r, u_k), (u_1, u_{k+1}), \ldots, (u_1, u_n)\}, \tag{4}$$

where $2 \leq k \leq n-1$. As before, we label the edges by the index of their heads, so $e_i = (r, u_i)$ for $i \in \{1, \ldots, k\}$ and $e_i = (u_1, u_i)$ for $i \in \{k+1, \ldots, n\}$. Our first result is the following.

Theorem 6.1. *Let $M = (T, I, P)$ be a P-model where T is a rooted 3-caterpillar and $P = \{p_1, \ldots, p_n\}$ is a multiset of possible prizes indexed increasingly $p_1 \leq p_2 \leq \cdots \leq p_n$. Then the SS $(T, \mathbf{1}, p)$, where $p(u_i) = p_i$ for each $i \in \{1, \ldots, n\}$ is an optimal SS for M.*

Proof. Let $B = m \in \{0, 1, \ldots, n\}$ be the attacker's budget, that is the number of edges an adversary can afford to penetrate. We want to show that $\mathrm{pr}^\star(m, \mathbf{1}, p) \leq \mathrm{pr}^\star(m, \mathbf{1}, p')$ for any prize assignment p' to the vertices of the rooted 3-caterpillar T.

Let $\tau \subseteq T$ be a rooted subtree of T on m edges with $\mathrm{pr}(\tau, \mathbf{1}, p) = \mathrm{pr}^\star(m, \mathbf{1}, p)$. There are two cases we need to consider.

First case: $e_1 \in E(\tau)$. Since all the leaves are connected to one of the end-vertices of $e_1 = (r, u_1)$, the remaining $m-1$ edges of τ must be incident to the $m-1$ maximum prize vertices, and so $\mathrm{pr}^\star(m, \mathbf{1}, p) = \mathrm{pr}(\tau, \mathbf{1}, p) = p_n + p_{n-1} + \cdots + p_{n-m+2} + p_1$. If p' is another prize assignment to the vertices of T, then $p'(u_1) = p_c$, where $c \in \{1, \ldots, n\}$. Therefore, $\mathrm{pr}^\star(m, \mathbf{1}, p') \geq \mathrm{pr}(\tau', \mathbf{1}, p')$, where τ' is a rooted subtree of T that contains e_1 and contains all the remaining $m-1$ maximum prizes, and so

$$\mathrm{pr}(\tau', \mathbf{1}, p') = \begin{cases} p_n + p_{n-1} + \cdots + p_{n-m+1} & \text{if } c \in \{n-m+1, \ldots, n\}, \\ p_n + p_{n-1} + \cdots + p_{n-m+2} + p_c & \text{if } c \notin \{n-m+1, \ldots, n\}. \end{cases}$$

In either case we have $\mathrm{pr}(\tau', \mathbf{1}, p') \geq p_n + p_{n-1} + \cdots p_{n-m+2} + p_1 = \mathrm{pr}^\star(m, \mathbf{1}, p)$, and so $\mathrm{pr}^\star(m, \mathbf{1}, p') \geq \mathrm{pr}^\star(m, \mathbf{1}, p)$ in this case.

Second case: $e_1 \notin E(\tau)$. For this case to be possible we must have $m \leq k-1$, since otherwise e_1 must be in τ. Secondly, we must have that τ contains all the maximum prize vertices on level one and so $\mathrm{pr}^\star(m, \mathbf{1}, p) = \mathrm{pr}(\tau, \mathbf{1}, p) = p_k + p_{k-1} + \cdots + p_{k-m+1}$. In particular, we must have

$$p_k + p_{k-1} + \cdots + p_{k-m+1} \geq p_n + p_{n-1} + \cdots + p_{n-m+2} + p_1,$$

since a tree containing e_1 does not have a greater total prize than τ. If p' is another prize assignment to the vertices of T, then let $\{\ell_1, \ldots, \ell_k\}$ be the indices of the prizes assigned to vertices on level one by p', that is, $\{p_{\ell_1}, \ldots, p_{\ell_k}\} = \{p'(u_1), \ldots, p'(u_k)\}$

as multisets. If now τ' is the rooted subtree of T with m edges containing the m vertices with the largest prizes, then, since $p_{\ell_i} \geq p_i$ for each $i \in \{1, \ldots, k\}$, we have

$$\begin{aligned} \mathrm{pr}^\star(m, \mathbf{1}, p') &\geq \mathrm{pr}(\tau', \mathbf{1}, p') \\ &= p_{\ell_k} + p_{\ell_{k-1}} + \cdots + p_{\ell_{k-m+1}} \\ &\geq p_k + p_{k-1} + \cdots + p_{k-m+1} \\ &= \mathrm{pr}^\star(m, \mathbf{1}, p), \end{aligned}$$

in this case as well. This completes the proof that the SS (T, p) is optimal. □

Now, let T be a rooted 4-spider on vertices $\{r, u_1, \ldots, u_n\}$ with edges given by

$$E(T) = \{(r, u_1), \ldots, (r, u_k), (u_1, u_{k+1}), (u_2, u_{k+2}), \ldots, (u_{n-k}, u_n)\}, \tag{5}$$

where $n/2 \leq k \leq n-2$. As before, the edges are labeled by the index of their heads: $e_i = (r, u_i)$ for $i \in \{1, \ldots, k\}$ and $e_i = (u_{i-k}, u_i)$ for $i \in \{k+1, \ldots, n\}$. Our second result is the following.

Theorem 6.2. *Let $M = (T, I, P)$ be a P-model, where T is a rooted 4-spider and $P = \{p_1, \ldots, p_n\}$ is a multiset of possible prizes indexed increasingly $p_1 \leq p_2 \leq \cdots \leq p_n$. Then the SS $(T, \mathbf{1}, p)$, where $p(u_i) = p_i$ for $i \in \{1, \ldots, k\}$ and $p(u_i) = p_{n+k+1-i}$ for $i \in \{k+1, \ldots, n\}$ is an optimal SS for M.*

Before we prove Theorem 6.2, we need a few lemmas that will come in handy for the proof.

Lemma 6.1. *Let T be a 4-spider presented as in (5) and $m \in \mathbb{N}$. Let p be a prize assignment on $V(T)$ such that $p_i = p(u_i) \leq p(u_j) = p_j$, where u_i is on level one and u_j is a leaf of T. If p' is the prize assignment obtained from p by swapping the prizes of u_i and u_j, then $\mathrm{pr}^\star(m, \mathbf{1}, p) \leq \mathrm{pr}^\star(m, \mathbf{1}, p')$.*

Proof. If $j = k + i$, so u_j is the unique child of u_i, then the lemma holds by (1). Hence, we can assume that u_j is not a child of u_i. Let $\tau \subseteq T$ be a max-prize rooted subtree on m edges, so $\mathrm{pr}(\tau, \mathbf{1}, p) = \mathrm{pr}^\star(m, \mathbf{1}, p)$. We now consider the following cases.

If either both u_i and u_j are vertices of τ, or neither of them are, then clearly $\mathrm{pr}^\star(m, \mathbf{1}, p) = \mathrm{pr}(\tau, \mathbf{1}, p) = \mathrm{pr}(\tau, \mathbf{1}, p') \leq \mathrm{pr}^\star(m, \mathbf{1}, p')$.

If $u_i \in V(\tau)$ and $u_j \notin V(\tau)$, then

$$\mathrm{pr}^\star(m, \mathbf{1}, p) = \mathrm{pr}(\tau, \mathbf{1}, p) \leq \mathrm{pr}(\tau, \mathbf{1}, p) - p_i + p_j = \mathrm{pr}(\tau, \mathbf{1}, p') \leq \mathrm{pr}^\star(m, \mathbf{1}, p).$$

If $u_i \notin V(\tau)$ and $u_j \in V(\tau)$, then, since u_i is on level one and u_j is a leaf of τ, we have that $\tau' = (\tau - u_j) \cup u_i$ is also a rooted subtree of T on m vertices and $\mathrm{pr}^\star(m, \mathbf{1}, p) = \mathrm{pr}(\tau, \mathbf{1}, p) = \mathrm{pr}(\tau', \mathbf{1}, p') \leq \mathrm{pr}^\star(m, \mathbf{1}, p')$, which completes our proof. □

Let $M = (T, I, P)$ be a P-model where T is a rooted 4-spider, $P = \{p_1, \ldots, p_n\}$, and p' be an arbitrary prize assignment on $V(T)$. Since every vertex of T on level two is automatically a leaf, we can, by repeated use of Lemma 6.1, obtain a prize assignment with smaller max-prize with respect to any m that has its $n-k$ largest prizes on its level-two vertices, and hence has its k smallest prizes on the level-one vertices $u_1, \ldots, u_k$ of T. By further use of the same Lemma 6.1 when considering these level-one vertices of T, we can obtain a prize assignment p that has its smallest prizes on the non-leaf vertices on level one and yet with smaller max-prize, so $\mathrm{pr}^\star(m, \mathbf{1}, p) \leq \mathrm{pr}^\star(m, \mathbf{1}, p')$ for any m. Note that our p satisfies

$$p(\{u_1, \ldots, u_{n-k}\}) = \{p_1, \ldots, p_{n-k}\}, \;\; p(\{u_{k+1}, \ldots, u_n\}) = \{p_{k+1}, \ldots, p_n\}.$$

As the level-one vertices of T can be assumed to be ordered by their prizes, we summarize in the following.

Corollary 6.1. *From any prize assignment p' we can by repeated use of Lemma 6.1 obtain a prize assignment p on our 4-spider T, presented as in (5), such that*

$$p(u_i) = p_i \text{ for all } i \in \{1, \ldots, k\}, \text{ and } p(u_i) = p_{\pi(i)} \text{ for all } i \in \{k+1, \ldots, n\},$$

where π is a permutation of $\{k+1, \ldots, n\}$, and with $\mathrm{pr}^\star(m, \mathbf{1}, p) \leq \mathrm{pr}^\star(m, \mathbf{1}, p')$ for any $m \in \mathbb{N}$.

Our next lemma provides our final tool in proving Theorem 6.2.

Lemma 6.2. *Let T be a 4-spider presented as in (5) and $m \in \mathbb{N}$. Let p be a prize assignment on $V(T)$ such that for some $i, j \in \{1, \ldots, n-k\}$ with $i < j$, we have $p(u_i) \leq p(u_j)$ and $p(u_{i+k}) \geq p(u_{j+k})$. If p' is a prize assignment where the prizes on u_{i+k} and u_{j+k} have been swapped, then $\mathrm{pr}^\star(m, \mathbf{1}, p) \leq \mathrm{pr}^\star(m, \mathbf{1}, p')$.*

Proof. Let $\tau \subseteq T$ be a max-prize rooted subtree on m edges with respect to p, so $\mathrm{pr}(\tau, \mathbf{1}, p) = \mathrm{pr}^\star(m, \mathbf{1}, p)$. We now consider the following cases.

If either both u_{i+k} and u_{j+k} are vertices of τ, or neither of them are, then clearly $\mathrm{pr}^\star(m, \mathbf{1}, p) = \mathrm{pr}(\tau, \mathbf{1}, p) = \mathrm{pr}(\tau, \mathbf{1}, p') \leq \mathrm{pr}^\star(m, \mathbf{1}, p')$.

If $u_{i+k} \notin V(\tau)$ and $u_{j+k} \in V(\tau)$, then

$$\begin{aligned} \mathrm{pr}^\star(m, \mathbf{1}, p) &= \mathrm{pr}(\tau, \mathbf{1}, p) \\ &\leq \mathrm{pr}(\tau, \mathbf{1}, p) - p(u_{j+k}) + p(u_{i+k}) \\ &= \mathrm{pr}(\tau, \mathbf{1}, p') \\ &\leq \mathrm{pr}^\star(m, \mathbf{1}, p'). \end{aligned}$$

If $u_{i+k} \in V(\tau)$ and $u_{j+k} \notin V(\tau)$, then we consider two (sub-)cases. If $u_j \in V(\tau)$, then since u_j is a leaf in τ, we have that $\tau' = (\tau - u_{i+k}) \cup u_{j+k}$ is also a rooted subtree of T on m vertices and $\mathrm{pr}^\star(m, \mathbf{1}, p) = \mathrm{pr}(\tau, \mathbf{1}, p) = \mathrm{pr}(\tau', \mathbf{1}, p') \leq \mathrm{pr}^\star(m, \mathbf{1}, p')$. If $u_j \notin V(\tau)$, then $\tau'' = (\tau - \{u_i, u_{i+k}\}) \cup \{u_j, u_{j+k}\}$ is also a rooted subtree of T on

m vertices, and

$$\begin{aligned} \mathrm{pr}^\star(m, \mathbf{1}, p) &= \mathrm{pr}(\tau, \mathbf{1}, p) \\ &\leq \mathrm{pr}(\tau, \mathbf{1}, p) - p(u_i) - p(u_{j+k}) + p(u_j) + p(u_{i+k}) \\ &= \mathrm{pr}(\tau'', \mathbf{1}, p') \\ &\leq \mathrm{pr}^\star(m, \mathbf{1}, p'), \end{aligned}$$

which completes the proof. □

Proof of Theorem 6.2. Let T be a 4-spider, p a prize assignment as given in Theorem 6.2, and $m \in \mathbb{N}$. Let p' be an arbitrary prize assignment of the vertices of T. By Corollary 6.1 we can obtain a prize assignment p'' such that

$$p''(u_i) = p_i \text{ for all } i \in \{1, \ldots, k\}, \text{ and } p''(u_i) = p_{\pi(i)} \text{ for all } i \in \{k+1, \ldots, n\},$$

where π is a permutation of $\{k+1, \ldots, n\}$, and with $\mathrm{pr}^\star(m, \mathbf{1}, p'') \leq \mathrm{pr}^\star(m, \mathbf{1}, p')$ for any $m \in \mathbb{N}$. By Lemma 6.2 we can obtain a prize assignment p on $V(T)$ from p'' simply by ordering the prizes on the level-two leaves in a decreasing order, thereby obtaining the very prize assignment p from Theorem 6.2 that satisfies $\mathrm{pr}^\star(m, \mathbf{1}, p) \leq \mathrm{pr}^\star(m, \mathbf{1}, p'')$ for any $m \in \mathbb{N}$. This proves that for any $m \in \mathbb{N}$ we have $\mathrm{pr}^\star(m, \mathbf{1}, p) \leq \mathrm{pr}^\star(m, \mathbf{1}, p'') \leq \mathrm{pr}^\star(m, \mathbf{1}, p')$, and since p' was an arbitrary prize assignment, the proof is complete. □

As a further observation, we can describe the optimal SAs on the P-model $M = (T, I, P)$, where T is a rooted 4-spider with the vertices and edges labeled as in (5), as follows.

Observation 6.1. *Let T be a 4-spider, p a prize assignment as in Theorem 6.2, and $m \in \mathbb{N}$. Then there is a max-prize rooted subtree $\tau \subseteq T$ on m edges with respect to p, so $\mathrm{pr}(\tau, \mathbf{1}, p) = \mathrm{pr}^\star(m, \mathbf{1}, p)$, with the following property:*

1. *If $n \leq 2k-1$, then all the leaves of τ are leaves in T, and hence in the set $\{u_{n-k+1}, \ldots, u_n\}$.*
2. *If $n = 2k$, then τ has at most one leaf on level one, in which case it can assumed to be u_k.*

Proof. Suppose τ has two leaves $u_i, u_j \in \{u_1, \ldots, u_{n-k}\}$. In this case $\tau' = (\tau - u_j) \cup u_{k+i}$ is also a rooted subtree of T on m edges and has $\mathrm{pr}(\tau', \mathbf{1}, p) \geq \mathrm{pr}(\tau, \mathbf{1}, p)$. Hence, we can assume τ to have at most one leaf from $\{u_1, \ldots, u_{n-k}\}$.

Suppose τ has one leaf $u_i \in \{u_1, \ldots, u_{n-k}\}$. We now consider the two cases; $k > n-k$ and $k = n-k$.

First case: $k > n-k$ or $n \leq 2k-1$. If τ has another additional leaf $u_j \in \{u_{n-k+1}, \ldots, n_k\}$, then, as above, $\tau' = (\tau - u_j) \cup u_{k+i}$ has $\mathrm{pr}(\tau', \mathbf{1}, p) \geq \mathrm{pr}(\tau, \mathbf{1}, p)$. Otherwise, τ has no leaves from $\{u_{n-k+1}, \ldots, n_k\} \neq \emptyset$. In this case $\tau'' = (\tau - u_i) \cup u_k$ is a rooted subtree of T on m edges with $\mathrm{pr}(\tau'', \mathbf{1}, p) \geq \mathrm{pr}(\tau, \mathbf{1}, p)$. Hence, we can assume that τ has no leaves from $\{u_1, \ldots, u_{n-k}\}$, which proves or claim in this case.

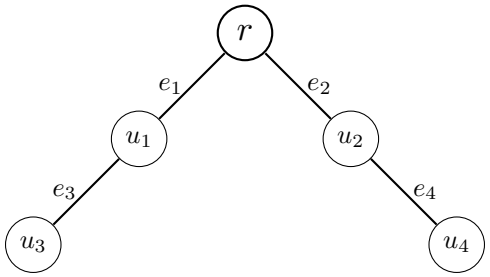
SECOND CASE: $k = n-k$ or $n = 2k$. In this case τ has the unique level-one leaf u_i. If $i < k$, then u_k has a unique child u_{2k} in τ, and so $\tau' = (\tau - u_{2k}) \cup u_{k+i}$ has the unique level-one leaf u_k and $\mathrm{pr}(\tau', \mathbf{1}, p) \geq \mathrm{pr}(\tau, \mathbf{1}, p)$. Hence, we can assume that τ has its unique level-one leaf u_k. □

Remark. Note that in the case $n \leq 2k-1$ in the proof of Observation 6.1, all the level-one leaves of τ can be assumed to be from $\{u_{n-k+1}, \ldots, u_k\}$. If we have ℓ of them, then they can further be assumed to be $u_{k-\ell+1}, \ldots, u_k$.

7 Duality between P- and C-Models

In this section we state and use a duality between the P- and C-models, which then can be used to obtain similar results for C-models that we obtained for P-models in the previous section. In particular, we will demonstrate that if T is one of the four types of rooted trees mentioned in Theorem 5.1, then any C-model $M = (T, C, I)$ indeed has an optimal SS, as we proved was the case for the P-model. As with the P-model, we already have that any C-model $M = (T, C, I)$ (in fact, any CSM $M = (T, C, P)$), where T is a rooted path or a rooted star, does have an optimal SS.

As mentioned in Remark 5.1 right after Observation 5.3, we now explicitly examine an example of a rooted proper subtree $T_p(2)$ of $T(2)$, for which any P- or C-model $M = (T_p(2), C, P)$ has an optimal security system. For the next two examples, and just as in the convention right before Example 3.1, let $T_p(2)$ denote the rooted tree, whose underlying graph is a path, on five vertices $V(T_p(2)) = \{r, u_1, u_2, u_3, u_4\}$ and edges $E(T_p(2)) = \{(r, u_1), (r, u_2), (u_1, u_3), (u_2, u_4)\}$ rooted at its center vertex. We continue the convention of labeling the edges by the same index as their heads: $e_1 = (r, u_1)$, $e_2 = (r, u_2)$, $e_3 = (u_1, u_3)$, and $e_4 = (u_2, u_4)$, see Figure 5.



$T_p(2)$

Figure 5: The underlying graph of $T_p(2)$ is a path on five vertices.

Example 7.1.
Consider a P-model (with $c = \mathbf{1}$) on the rooted tree $T_p(2)$ where the prize values $P = \{p_1, p_2, p_3, p_4\}$ are general real positive values ordered increasingly $p_1 \leq p_2 \leq p_3 \leq p_4$. By Theorem 6.2 an optimal SS for our CSM $M = (T_p(2), I, P)$ is obtained by assigning the prizes as $p(u_1, u_2, u_3, u_4) := (p_1, p_2, p_4, p_3)$. We can explicitly obtain the max-prize subtree for each given budgets $B \in \mathbb{R}$ that yields the following:

$$\mathrm{pr}^\star(B, \mathbf{1}, p) = \begin{cases} 0 & \text{for } B < 1, \\ p_2 & \text{for } 1 \leq B < 2, \\ \max(p_1 + p_4, p_2 + p_3) & \text{for } 2 \leq B < 3, \\ p_1 + p_2 + p_4 & \text{for } 3 \leq B < 4, \\ p_1 + p_2 + p_3 + p_4 & \text{for } 4 \leq B. \end{cases}$$

Example 7.2.
Consider a C-model (with $p = \mathbf{1}$) on the rooted tree $T_p(2)$ where the penetration cost values $C = \{c_1, c_2, c_3, c_4\}$ are general real positive values ordered decreasingly $c_1 \geq c_2 \geq c_3 \geq c_4$. It is now an easy combinatorial exercise to verify directly that an optimal SS for our CSM $M = (T_p(2), C, I)$ can be obtained by assigning penetration costs as $c(u_1, u_2, u_3, u_4) := (c_1, c_2, c_4, c_3)$, in the same (index-)order as for the P-model in Example 7.1. We explicitly obtain the max-prize subtree for each given budget $B \in \mathbb{R}$ that yields the following:

$$\mathrm{pr}^\star(B, c, \mathbf{1}) = \begin{cases} 0 & \text{for } B < c_2, \\ 1 & \text{for } c_2 \leq B < \min(c_1 + c_4, c_2 + c_3), \\ 2 & \text{for } \min(c_1 + c_4, c_2 + c_3) \leq B < c_1 + c_2 + c_4, \\ 3 & \text{for } c_1 + c_2 + c_4 \leq B < c_1 + c_2 + c_3 + c_4, \\ 4 & \text{for } c_1 + c_2 + c_3 + c_4 \leq B. \end{cases}$$

Let K be a sufficiently large cost number (any real number $\geq \max(c_1, \ldots, c_4) + 1$ will do), and write each edge-cost of the form $c_i = K - c_i'$. In this way $\mathrm{pr}^\star(B, c, \mathbf{1})$ will take the following form

$$\mathrm{pr}^\star(B, c, \mathbf{1}) = \begin{cases} 0 & \text{for } B < K - c_2', \\ 1 & \text{for } K - c_2' \leq B < 2K - \max(c_1' + c_4', c_2' + c_3'), \\ 2 & \text{for } 2K - \max(c_1' + c_4', c_2' + c_3') \leq B < 3K - (c_1' + c_2' + c_4'), \\ 3 & \text{for } 3K - (c_1' + c_2' + c_4') \leq B < 4K - (c_1' + c_2' + c_3' + c_4'), \\ 4 & \text{for } 4K - (c_1' + c_2' + c_3' + c_4') \leq B. \end{cases}$$

From the above we see the evident resemblance to the expression for $\mathrm{pr}^\star(B, \mathbf{1}, p)$ of the P-model in Example 7.1. This is a glimpse of a duality between the P-models and the C-models that we will now describe.

Convention: In what follows, it will be convenient to view the cost and prize assignments c and p not as functions as in Definition 3.2, but rather as vectors $\tilde{c} = (c_1, \ldots, c_n)$ and $\tilde{p} = (p_1, \ldots, p_n)$ in the n-dimensional Euclidean space $\mathbb{R}^n$, which can be obtained by a fixed labeling of the n non-root vertices $u_1, \ldots, u_n$ and a corresponding labeling of the edges $e_1, \ldots, e_n$, with our usual convention that for each i the vertex u_i is the head of e_i, and by letting $c_i := c(e_i)$ and $p_i := p(u_i)$.

For a given $n \in \mathbb{N}$, let $\mathcal{B}(\mathbb{R}^n)$ denote the group of all bijections $\mathbb{R}^n \to \mathbb{R}^n$ with respect to compositions of maps. For $a \in \mathbb{Q}_+$ and $b \in \mathbb{Q}$ the affine map $\alpha : \mathbb{R}^n \to \mathbb{R}^n$ given by $\alpha(\tilde{x}) = a\tilde{x} + b\tilde{1}$, where $\tilde{1} = (1, \ldots, 1) \in \mathbb{R}^n$, is bijective with an inverse $\alpha^{-1}(\tilde{x}) = \frac{1}{a}\tilde{x} - \frac{b}{a}\tilde{1}$ of the same type. Further, if $\alpha'(\tilde{x}) = a'\tilde{x} + b'\tilde{1}$ is another such map, then the composition $(\alpha' \circ \alpha)(\tilde{x}) = a'a\tilde{x} + (a'b + b')\tilde{1}$ is also a bijection of this very type. Since the identity map of $\mathbb{R}^n$ has $a = 1 \in \mathbb{Q}_+$ and $b = 0 \in \mathbb{Q}$, we have the following.

Observation 7.1. *If $n \in \mathbb{N}$ then $G_n = \{\alpha \in \mathcal{B}(\mathbb{R}^n) : \alpha(\tilde{x}) = a\tilde{x} + b\tilde{1},$ for some $a \in \mathbb{Q}_+$ and $b \in \mathbb{Q}\}$ is a subgroup of $\mathcal{B}(\mathbb{R}^n)$.*

By letting G_n act on the set $\mathbb{R}^n$ in the natural way, $(\alpha, \tilde{x}) \mapsto \alpha(\tilde{x})$, then the group orbits $G_n(\tilde{x}) = \{\alpha(\tilde{x}) : \alpha \in G_n\}$ yield a partition of $\mathbb{R}^n$ into corresponding equivalence classes $\mathbb{R}^n = \bigcup_{\tilde{x} \in \mathbb{R}^n} G_n(\tilde{x})$. By intersecting with $\mathbb{Q}_+^n$ we obtain the following equivalence classes that we seek.

Definition 7.1. *For each $\tilde{x} \in \mathbb{Q}_+^n$ let $[\tilde{x}]$ denote the equivalence class of $\tilde{x}$ with respect to the partition of $\mathbb{R}^n$ into the G_n orbits: $[\tilde{x}] = G_n(\tilde{x}) \cap \mathbb{Q}_+^n$.*

We now justify the above equivalence of vectors of $\mathbb{Q}_+^n$. The following observation is obtained directly from Definition 3.2.

Observation 7.2. *Let T be a rooted tree on n labeled non-root vertices and edges, τ a rooted subtree of T, and $\alpha \in G_n$ given by $\alpha(\tilde{x}) = a\tilde{x} + b\tilde{1}$. If $\tilde{c}, \tilde{p} \in \mathbb{Q}_+^n$ are a cost and prize vector, respectively, then we have*

$$\begin{aligned} \mathrm{pr}(\tau, \tilde{c}, \alpha(\tilde{p})) &= a\mathrm{pr}(\tau, \tilde{c}, \tilde{p}) + |E(\tau)|b, \\ \mathrm{cst}(\tau, \alpha(\tilde{c}), \tilde{p}) &= a\mathrm{cst}(\tau, \tilde{c}, \tilde{p}) + |E(\tau)|b. \end{aligned}$$

If $J \subseteq \{1, \ldots, n\}$ and $\Sigma_J : \mathbb{R}^n \to \mathbb{R}$ is given by $\tilde{x} \mapsto \sum_{i \in J} x_i$, then we clearly have

$$\Sigma_J(\alpha(\tilde{x})) \leq \Sigma_J(\alpha(\tilde{y})) \Leftrightarrow \Sigma_J(\tilde{x}) \leq \Sigma_J(\tilde{y}), \tag{6}$$

and hence the following corollary.

Corollary 7.1. *Let T be a rooted tree on n labeled non-root vertices and edges, $B \in \mathbb{Q}_+$ a budget, and $\alpha \in G_n$ given by $\alpha(\tilde{x}) = a\tilde{x} + b\tilde{1}$.*

(i) If $\tilde{p} \in \mathbb{Q}_+^n$ is a prize vector, then we have

$$\mathrm{pr}^\star(B, \tilde{1}, \alpha(\tilde{p})) = a\mathrm{pr}^\star(B, \tilde{1}, \tilde{p}) + b\lfloor B \rfloor. \tag{7}$$

Further, both max prizes in (7) are attained at the same rooted subtree τ of T where $|E(\tau)| = \lfloor B \rfloor$.

(ii) If $\tilde{c} \in \mathbb{Q}_+^n$ is a cost vector, then we have

$$\mathrm{pr}^\star(aB + bm, \alpha(\tilde{c}), \tilde{1}) = m \Leftrightarrow \mathrm{pr}^\star(B, \tilde{c}, \tilde{1}) = m,$$

and further, both max prizes are attained at the same rooted subtree τ of T within the budget; that is, $|E(\tau)| = m$ and $\mathrm{cst}(\tau, \tilde{c}, \tilde{1}) \leq B$.

Remark. (i) That both max prizes are attained at the same rooted subtree τ in (i) in Corollary 7.1 simply means that

$$\mathrm{pr}(\tau, \tilde{1}, \alpha(\tilde{p})) = \mathrm{pr}^{\star}(B, \tilde{1}, \alpha(\tilde{p})) \Leftrightarrow \mathrm{pr}(\tau, \tilde{1}, \tilde{p}) = \mathrm{pr}^{\star}(B, \tilde{1}, \tilde{p}),$$

which is a direct consequence of Observation 7.2 and (7). (ii) Also, for a rooted subtree τ with $|E(\tau)| = m$ and $\mathrm{cst}(\tau, \tilde{c}, \tilde{1}) \leq B$, then by Observation 7.2 we also have $\mathrm{cst}(\tau, \alpha(\tilde{c}), \tilde{1}) \leq aB + bm$, and

$$\mathrm{pr}(\tau, \tilde{c}, \tilde{1}) = m = \mathrm{pr}^{\star}(B, \tilde{c}, \tilde{1}) \Leftrightarrow \mathrm{pr}(\tau, \alpha(\tilde{c}), \tilde{1}) = m = \mathrm{pr}^{\star}(aB + bm, \alpha(\tilde{c}), \tilde{1}).$$

We can, in fact, say a tad more than Corollary 7.1 for C-models $M = (T, C, I)$.

Definition 7.2. *Let $M = (T, C, I)$ be a C-model. For a given cost vector $\tilde{c} \in \mathbb{Q}_+^n$ let $B_m(\tilde{c})$ denote the smallest cost $B \in \mathbb{Q}_+$ with $\mathrm{pr}^{\star}(B, \tilde{c}, \tilde{1}) = m$.*

Note that

$$\mathrm{pr}^{\star}(B, \tilde{c}, \tilde{1}) = m \Leftrightarrow B_m(\tilde{c}) \leq B < B_{m+1}(\tilde{c}).$$

We also have the following useful lemma.

Lemma 7.1. *If $\alpha \in G_n$ is given by $\alpha(\tilde{x}) = a\tilde{x} + b\tilde{1}$, then $B_m(\alpha(\tilde{c})) = aB_m(\tilde{c}) + bm$.*

Proof. By definition of $B_m(\tilde{c})$ we have $\mathrm{pr}^{\star}(B_m(\tilde{c}), \tilde{c}, \tilde{1}) = m$, and hence by Corollary 7.1 $\mathrm{pr}^{\star}(aB_m(\tilde{c}) + bm, \alpha(\tilde{c}), \tilde{1}) = m$ as well. Suppose that $\mathrm{pr}^{\star}(B', \alpha(\tilde{c}), \tilde{1}) = m$, where $B' < aB_m(\tilde{c}) + bm$. If now $B' = aB'' + bm$, then $B'' < B_m(\tilde{c})$ and we have again by Corollary 7.1 that $\mathrm{pr}^{\star}(B'', \tilde{c}, \tilde{1}) = m$. This contradicts the definition of $B_m(\tilde{c})$. Hence, $B_m(\alpha(\tilde{c}) = aB_m(\tilde{c}) + bm$. □

Proposition 7.1. *For $m \in \{0, 1, \ldots, n\}$ and a cost vectors $\tilde{c}$ and $\tilde{c}'$ we have $B_m(\tilde{c}) \geq B_m(\tilde{c}')$ if and only if for every budget B with $\mathrm{pr}^{\star}(B, \tilde{c}, \tilde{1}) = m$ we have $\mathrm{pr}^{\star}(B, \tilde{c}, \tilde{1}) \leq \mathrm{pr}^{\star}(B, \tilde{c}', \tilde{1})$.*

Proof. Suppose $B_m(\tilde{c}) \geq B_m(\tilde{c}')$, and let B be a budget with $\mathrm{pr}^{\star}(B, \tilde{c}, \tilde{1}) = m$. By definition we then have $B \geq B_m(\tilde{c})$ and hence $B \geq B_m(\tilde{c}')$ and therefore $\mathrm{pr}^{\star}(B, \tilde{c}', \tilde{1}) \geq m = \mathrm{pr}^{\star}(B, \tilde{c}, \tilde{1})$.

Conversely, if for every budget B with $\mathrm{pr}^{\star}(B, \tilde{c}, \tilde{1}) = m$ we have $\mathrm{pr}^{\star}(B, \tilde{c}, \tilde{1}) \leq \mathrm{pr}^{\star}(B, \tilde{c}', \tilde{1})$, then, in particular for $B = B_m(\tilde{c})$ we have $m = \mathrm{pr}^{\star}(B_m(\tilde{c}), \tilde{c}, \tilde{1}) \leq \mathrm{pr}^{\star}(B_m(\tilde{c}), \tilde{c}', \tilde{1})$, and hence, by definition, $B_m(\tilde{c}') \leq B_m(\tilde{c})$. □

Convention: For a vector $\tilde{x} = (x_1, \ldots, x_n) \in \mathbb{Q}_+^n$ let $\{\tilde{x}\}$ denote its underlying multiset. So if $(T, \tilde{c}, \tilde{p})$ is an SS for a CSM $M = (T, C, P)$, then we necessarily have $C = \{\tilde{c}\}$ and $P = \{\tilde{p}\}$ as multisets. Also, we have $\{\tilde{1}\} = I$ as the multiset containing n copies of 1.

Suppose $\mathrm{pr}^{\star}(B, \tilde{1}, \tilde{p}) \leq \mathrm{pr}^{\star}(B, \tilde{1}, \tilde{p}')$ for all $\tilde{p}'$ with $\{\tilde{p}'\} = \{\tilde{p}\}$. Then by Corollary 7.1 we get for any $\alpha \in G_n$ with $\alpha(\tilde{x}) = a\tilde{x} + b\tilde{1}$, that

$$\mathrm{pr}^{\star}(B, \tilde{1}, \alpha(\tilde{p})) = a\mathrm{pr}^{\star}(B, \tilde{1}, \tilde{p}) + b\lfloor B \rfloor \leq a\mathrm{pr}^{\star}(B, \tilde{1}, \tilde{p}') + b\lfloor B \rfloor = \mathrm{pr}^{\star}(B, \tilde{1}, \alpha(\tilde{p}')),$$

and so we have the following.

Proposition 7.2. *The SS $(T, \tilde{1}, \tilde{p})$ is optimal for the P-model $M = (T, I, \{\tilde{p}\})$ with respect to the budget $B \in \mathbb{Q}_+$ if and only if the SS $(T, \tilde{1}, \alpha(\tilde{p}))$ is optimal for the P-model $M = (T, I, \{\alpha(\tilde{p})\})$ with respect to B.*

In a similar way, we have by Proposition 7.1 that $\mathrm{pr}^\star(B, \tilde{c}, \tilde{1}) = m \leq \mathrm{pr}^\star(B, \tilde{c}', \tilde{1})$ whenever $B_m(\tilde{c}) \leq B < B_{m+1}(\tilde{c})$ and $\{\tilde{c}'\} = \{\tilde{c}\}$ if and only if $B_m(\tilde{c}) \geq B_m(\tilde{c}')$, which by Lemma 7.1 holds if and only if

$$B_m(\alpha(\tilde{c})) = aB_m(\tilde{c}) + bm \geq aB_m(\tilde{c}') + bm = B_m(\alpha(\tilde{c}')).$$

In other words, $\mathrm{pr}^\star(B, \tilde{c}, \tilde{1}) \leq \mathrm{pr}^\star(B, \tilde{c}', \tilde{1})$ when $B_m(\tilde{c}) \leq B < B_{m+1}(\tilde{c})$ holds if and only if $\mathrm{pr}^\star(B', \alpha(\tilde{c}), \tilde{1}) \leq \mathrm{pr}^\star(B', \alpha(\tilde{c}'), \tilde{1})$ when $B_m(\alpha(\tilde{c})) \leq B' < B_{m+1}(\alpha(\tilde{c}))$. Since this holds for every $\alpha \in G_n$, which is a group with each element having an inverse, then we have the following.

Proposition 7.3. *The SS $(T, \tilde{c}, \tilde{1})$ is optimal for the C-model $M = (T, \{\tilde{c}\}, I)$ with respect to $B \in [B_m(\tilde{c}), B_{m+1}(\tilde{c})[\cap\mathbb{Q}_+$ if and only if the SS $(T, \alpha(\tilde{c}), \tilde{1})$ is optimal for the C-model $M' = (T, \{\alpha(\tilde{c})\}, I)$ with respect to $B' \in [B_m(\alpha(\tilde{c})), B_{m+1}(\alpha(\tilde{c}))[\cap\mathbb{Q}_+$.*

Combining Propositions 7.2 and 7.3, we have the following summarizing corollary.

Corollary 7.2. *Let $\alpha \in G_n$.*

The SS $(T, \tilde{1}, \tilde{p})$ is optimal for the P-model $M = (T, I, \{\tilde{p}\})$ if and only if the SS $(T, \tilde{1}, \alpha(\tilde{p}))$ is optimal for the P-model $M' = (T, I, \{\alpha(\tilde{p})\})$.

The SS $(T, \tilde{c}, \tilde{1})$ is optimal for the C-model $M = (T, \{\tilde{c}\}, I)$ if and only if the SS $(T, \alpha(\tilde{c}), \tilde{1})$ is optimal for the C-model $M' = (T, \{\alpha(\tilde{p})\}, I)$.

Corollary 7.2 shows that optimality of security systems of both C- and P-models is G_n-invariant when applied to the prize and cost vector, respectively.

Recall the equivalence class $[\tilde{x}] = G_n(\tilde{x}) \cap \mathbb{Q}_+^n$ from Definition 7.1. We can now define induced equivalence classes of SS of both C- and P-models. By Corollary 7.2 the following definition is valid (that is, the terms are all well defined).

Definition 7.3. *For a C-model $M = (T, C, I)$ and a SS $(T, \tilde{c}, \tilde{1})$ of M, we let*

$$[(T, \tilde{c}, \tilde{1})] := \{(T, \tilde{x}, \tilde{1}) : \tilde{x} \in [\tilde{c}]\}.$$

We say that $[(T, \tilde{c}, \tilde{1})]$ is optimal *if one $(T, \tilde{x}, \tilde{1}) \in [(T, \tilde{c}, \tilde{1})]$ is optimal for its corresponding $M = (T, \{\tilde{x}\}, I)$, since then each element in $[(T, \tilde{c}, \tilde{1})]$ is also optimal.*

Likewise, for a P-model $M = (T, I, P)$ and a SS $(T, \tilde{1}, \tilde{p})$ of M, we let

$$[(T, \tilde{1}, \tilde{p})] := \{(T, \tilde{1}, \tilde{y}) : \tilde{y} \in [\tilde{p}]\}.$$

We say that $[(T, \tilde{1}, \tilde{p})]$ is optimal *if one $(T, \tilde{1}, \tilde{y}) \in [(T, \tilde{1}, \tilde{p})]$ is optimal for its corresponding $M = (T, I, \{\tilde{y}\})$, since then each element in $[(T, \tilde{1}, \tilde{p})]$ is also optimal.*

With the setup just presented we now can define the dual of both vector classes and SS classes for C- and P-models in the following.

Definition 7.4. *For a vector* $\tilde{x}$ *and* $[\tilde{x}] = G_n(\tilde{x}) \cap \mathbb{Q}_+^n$ *as in Definition 7.1, then* $[\tilde{x}]^* := [-\tilde{x}]$ *is the* dual vector class *of* $[\tilde{x}]$.

For a C-model $M = (T, C, I)$ *and a SS* $(T, \tilde{c}, \tilde{1})$ *of* M, *then* $[(T, \tilde{c}, \tilde{1})]^* := [(T, \tilde{1}, -\tilde{c})]$ *is the corresponding* dual P-model security system class (dual P-model SS class) *of the C-model class* $[(T, \tilde{c}, \tilde{1})]$.

Likewise, for a P-model $M = (T, I, P)$ *and a SS* $(T, \tilde{1}, \tilde{p})$ *of* M, *then the class* $[(T, \tilde{1}, \tilde{p})]^* := [(T, -\tilde{p}, \tilde{1})]$ *is the corresponding* dual C-model security system class (dual C-model SS class) *of the P-model class* $[(T, \tilde{1}, \tilde{p})]$.

Note that the double-dual yields the original class in each case: $[\tilde{x}]^{**} = [-\tilde{x}]^* = [\tilde{x}]$, and

$$[(T, \tilde{c}, \tilde{1})]^{**} = [(T, \tilde{1}, -\tilde{c})]^* = [(T, \tilde{c}, \tilde{1})], \quad [(T, \tilde{1}, \tilde{p})]^{**} = [(T, -\tilde{p}, \tilde{1})]^* = [(T, \tilde{1}, \tilde{p})].$$

For a P-model $M = (T, I, P)$ and a SS P-model class $[(T, \tilde{1}, \tilde{p})]$ we can always assume the prize vector $\tilde{p}$ is such $p_i \in [0, 1] \cap \mathbb{Q}_+$ for each i, since $\alpha(\tilde{x}) = a\tilde{x}$ is indeed an element of G_n for any $a > 0$. In this way $\tilde{c} = \tilde{1} - \tilde{p} \in ([0, 1] \cap \mathbb{Q}_+)^n$ is a legitimate cost vector, and we have $[\tilde{p}]^* = [\tilde{1} - \tilde{p}]$ and $[(T, \tilde{1}, \tilde{p})]^* = [(T, \tilde{1} - \tilde{p}, \tilde{1})]$. In what follows, we will call such a prize vector *scaled*. The following is easy to show.

Claim 7.1. *For a scaled prize vector* $\tilde{p}$ *with* $p_i \in [0, 1] \cap \mathbb{Q}_+$ *for each* i, *and a rooted subtree* τ *of* T *with* $|E(\tau)| = m$, *then* $\mathrm{pr}(\tau, \tilde{1}, \tilde{p}) + \mathrm{cst}(\tau, \tilde{1} - \tilde{p}, \tilde{1}) = m$.

Let $\tilde{p}$ be a scaled prize vector and assume B is a budget with $\mathrm{pr}^\star(B, \tilde{1} - \tilde{p}, \tilde{1}) = m$. Then there is a rooted subtree τ of T on m edges such that $\mathrm{cst}(\tau, \tilde{1} - \tilde{p}, \tilde{1}) \leq B$, and hence there is such a τ of smallest cost. Hence, we may assume τ is indeed such a rooted subtree of smallest cost. By Claim 7.1 applied to $\tilde{1} - \tilde{p}$, which is also scaled, we then have $\mathrm{pr}(\tau, \tilde{1}, \tilde{p}) = m - \mathrm{cst}(\tau, \tilde{1} - \tilde{p}, \tilde{1})$ with the smallest $\mathrm{cst}(\tau, \tilde{1} - \tilde{p}, \tilde{1})$ among rooted subtrees τ on m edges, and hence $\mathrm{pr}(\tau, \tilde{1}, \tilde{p})$ is maximum among all rooted subtrees τ on m edges, and so $\mathrm{pr}(\tau, \tilde{1}, \tilde{p}) = \mathrm{pr}^\star(m, \tilde{1}, \tilde{p})$. Hence,

$$B \geq \mathrm{cst}(\tau, \tilde{1} - \tilde{p}, \tilde{1}) = m - \mathrm{pr}(\tau, \tilde{1}, \tilde{p}) = m - \mathrm{pr}^\star(m, \tilde{1}, \tilde{p}).$$

Since $\mathrm{cst}(\tau, \tilde{1} - \tilde{p}, \tilde{1})$ is the smallest cost among all rooted subtrees on m edges, then

$$B' = \mathrm{cst}(\tau, \tilde{1} - \tilde{p}, \tilde{1}) = m - \mathrm{pr}^\star(m, \tilde{1}, \tilde{p})$$

is indeed the smallest cost with $\mathrm{pr}^\star(B', \tilde{1} - \tilde{p}, \tilde{1}) = m$. By Definition 7.2 we then have the following.

Lemma 7.2. *For* $m \in \{0, 1, \ldots, n\}$ *and a scaled (prize) vector* $\tilde{p}$, *we have*

$$B_m(\tilde{1} - \tilde{p}) = m - \mathrm{pr}^\star(m, \tilde{1}, \tilde{p}).$$

As a direct consequence of Lemma 7.2, we then have

Corollary 7.3. *For any* $m \in \{0, 1, \ldots, n\}$ *and scaled vectors* $\tilde{p}$ *and* $\tilde{p}'$, *we have*

$$B_m(\tilde{1} - \tilde{p}) \geq B_m(\tilde{1} - \tilde{p}') \Leftrightarrow \mathrm{pr}^\star(m, \tilde{1}, \tilde{p}) \leq \mathrm{pr}^\star(m, \tilde{1}, \tilde{p}').$$

We can now prove one of the main results in this section.

Theorem 7.1. *Let $M = (T, I, P)$ be a P-model, $(T, \tilde{1}, \tilde{p})$ a SS for M where $\tilde{p}$ is scaled, and $m \in \{0, 1, \ldots, n\}$. Then $\mathrm{pr}^\star(m, \tilde{1}, \tilde{p}) \leq \mathrm{pr}^\star(m, \tilde{1}, \tilde{p}')$ for any $\tilde{p}'$ with $\{\tilde{p}'\} = P$ if and only if $\mathrm{pr}^\star(B, \tilde{1} - \tilde{p}, \tilde{1}) \leq \mathrm{pr}^\star(B, \tilde{1} - \tilde{p}', \tilde{1})$ for any budget B with $\mathrm{pr}^\star(B, \tilde{1} - \tilde{p}, \tilde{1}) = m$ and for any $\tilde{p}'$ with $\{\tilde{p}'\} = P$.*

Proof. By Corollary 7.3 we have that $\mathrm{pr}^\star(m, \tilde{1}, \tilde{p}) \leq \mathrm{pr}^\star(m, \tilde{1}, \tilde{p}')$ for any $\tilde{p}'$ with $\{\tilde{p}'\} = P$ if and only if $B_m(\tilde{1} - \tilde{p}) \geq B_m(\tilde{1} - \tilde{p}')$ for any $\tilde{p}'$ with $\{\tilde{p}'\} = P$ which, by Proposition 7.1, holds if and only if $\mathrm{pr}^\star(B, \tilde{1} - \tilde{p}, \tilde{1}) \leq \mathrm{pr}^\star(B, \tilde{1} - \tilde{p}', \tilde{1})$ for all budgets B with $\mathrm{pr}^\star(B, \tilde{1} - \tilde{p}, \tilde{1}) = m$ and for all $\tilde{p}'$ with $\{\tilde{p}'\} = P$. □

Note that by Theorem 7.1 we have that $\mathrm{pr}^\star(B, \tilde{1}, \tilde{p}) \leq \mathrm{pr}^\star(B, \tilde{1}, \tilde{p}')$ for any budget B and any $\tilde{p}'$ with $\{\tilde{p}'\} = \{\tilde{p}\}$, if and only if $\mathrm{pr}^\star(B, \tilde{1} - \tilde{p}, \tilde{1}) \leq \mathrm{pr}^\star(B, \tilde{1} - \tilde{p}', \tilde{1})$ for any budget B and any $\tilde{p}'$ with $\{\tilde{p}'\} = \{\tilde{p}\}$. Hence, by Corollary 7.2 and Theorem 7.1 we therefore have the main conclusion of this section in light of Definition 7.3.

Corollary 7.4. *For a rooted tree T and a prize vector $\tilde{p} \in \mathbb{Q}_+^n$, then $[(T, \tilde{1}, \tilde{p})]$ is an optimal P-model SS class if and only if the dual C-model SS class $[(T, \tilde{1}, \tilde{p})]^* = [(T, -\tilde{p}, \tilde{1})]$ is optimal.*

In particular, if $\tilde{p}$ is scaled, then the SS $(T, \tilde{1}, \tilde{p})$ is optimal for the P-model $M = (T, I, \{\tilde{p}\})$ if and only if the SS $(T, \tilde{1} - \tilde{p}, \tilde{1})$ is optimal for the C-model $M = (T, \{\tilde{1} - \tilde{p}\}, I)$.

Consequently, by Corollary 4.1, Theorems 4.2, 5.1, 6.1 and 6.2 and Corollary 7.4, we have the following summarizing result.

Theorem 7.2. *For a rooted tree T on n non-root vertices the following are equivalent:*

1. *Any P-model $M = (T, I, P)$ has an optimal SS.*
2. *Any C-model $M = (T, C, I)$ has an optimal SS.*
3. *T is one of the following types: (i) a rooted path, (ii) a rooted star, (iii) a rooted 3-caterpillar, or (iv) a rooted 4-spider.*

Note that by (6) we have, in particular, that each $\alpha \in G_n$ preserves the order of the entries of each $\tilde{x} \in \mathbb{Q}_+^n$, so each $\tilde{x} \in [\tilde{p}]$ has the same order of its entries as $\tilde{p}$ does. But clearly, the dual operation on $[\tilde{x}]^* = [-\tilde{x}]$ is order reversing, that is, we have that $x_i \leq x_j$ for any $\tilde{x} \in [\tilde{p}]$ if and only if $y_i \geq y_j$ for any $\tilde{y} \in [-\tilde{p}] = [\tilde{p}]^*$. Since the optimal assignments of prizes from a given multiset P are given in Theorems 6.1 and 6.2, we then have by Corollary 7.4 the following theorems for C-models as well.

Theorem 7.3. *Let $M = (T, C, I)$ be a C-model where T is a rooted 3-caterpillar as in (4) and $C = \{c_1, \ldots, c_n\}$ is a multiset of possible edge-costs indexed decreasingly $c_1 \geq c_2 \geq \cdots \geq c_n$. Then the SS $(T, c, \mathbf{1})$, where $c(e_i) = c_i$ for each $i \in \{1, \ldots, n\}$ is an optimal SS for M.*

Theorem 7.4. *Let $M = (T, C, I)$ be a P-model, where T is a rooted 4-spider as in (5) and $C = \{c_1, \ldots, c_n\}$ is a multiset of possible edge-costs indexed decreasingly $c_1 \geq c_2 \geq \cdots \geq c_n$. Then the SS $(T, c, \mathbf{1})$, where $c(e_i) = c_i$ for $i \in \{1, \ldots, k\}$ and $c(e_i) = c_{n+k+1-i}$ for $i \in \{k+1, \ldots, n\}$ is an optimal SS for M.*

8 Summary and Conclusions

This paper defined a cyber-security model to explore defensive security systems. The results obtained mathematically support the intuition that it is best to place stronger defenses in the outer layers and more-valuable prizes in the deeper layers. We defined three types of SSs: improved, good, and optimal. We showed that it is not always possible to find an optimal SS for a given CSM, but demonstrated for rooted paths and stars that optimal SSs do exist. The results mathematically show that a path produces the best cyber-security, however, burying something n levels deep for large n may prevent the friendly side from accessing the "information" effectively. The results show, in general, that trees having greater depth provide more security in this setting.

We showed that any CSM is equivalent to a CSM where either all the edge penetration costs are unit priced (a P-model) or where all the vertices have a unit prize (C-model), by allowing larger underlying rooted trees. We then characterised for which trees a P-model has an optimal SSs, and we also did that for the C-models. We noted that the P- and C-models have optimal SSs for exactly the same types of rooted trees. This was then explained by obtaining a duality between the P- and C-models in the penultimate section of the paper.

We gave an $O(n \log n)$ algorithm for producing a good SS that was based on sorting. It is not clear how strong such a good SS is, as there may be many such good SSs, and some may be better than others. It would be interesting to come up with a comparison metric to rank various good SSs. We must continue to explore models of cyber-security systems to develop the foundations needed to combat the ongoing and increasing number of cyber attacks. This work is but one step in that direction.

We conclude the paper with a number of questions.

1. Can we find an efficient algorithm to develop optimal SSs in the cases where all penetration costs or all targets are from a finite set of possible values? Say, if we have two possible penetrations costs or three? Similarly for prizes?

2. In a two-player version of the model, what would be the best strategy for a defender who is allowed to reposition a prize or a portion of a prize after each move by an attacker? And, what would the complexity of this problem be?

3. Are there on-line variants of the model that are interesting to study? For example, a version where the topology of the tree changes dynamically or where only a partial description is known to the attacker.

4. Could a dynamic programming approach be used to obtain a SS that is somehow quantifiably better than a good SS or allow us to pick the "best" good SS?

5. Is there a more useful definition of neighboring configuration that could lead to an efficient algorithm for producing better SSs, for example, perhaps a definition where sibling vertices or edges can have their prizes or penetration costs swapped, respectively?

Acknowledgments

This work was supported by the Office of Naval Research. The work was also supported by Thailand Research Fund grant No. RSA5480006. – Finally, sincere thanks to the two anonymous referees for all their helpful comments and for spotting some well-hidden (and embarrassing!) typos. This greatly improved the presentation of the paper.

References

[1] El Houssaine Aghezzaf, Thomas L. Magnanti, and Laurence A. Wolsey. Optimizing Constrained Subtrees of Trees. *Mathematical Programming*, **71(2)**:113–126, Series A, (1995).

[2] Geir Agnarsson and Raymond Greenlaw. Graph Theory, Modeling, Applications, and Algorithms. *Prentice Hall*, (2007).

[3] Geir Agnarsson, Raymond Greenlaw, and Sanpawat Kantabutra. On Cyber Attacks and the Maximum-Weight Rooted-Subtree Problem *Acta Cybernetica*, **22**:591–612, (2016).

[4] Sofie Coene, Carlo Filippi, Frits Spieksma, and Elisa Stevanato. Balancing Profits and Costs on Trees. *Networks*, **61(3)**:200–11, (2013).

[5] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, New York, (1979).

[6] Go-Gulf. *Cyber Crime: Statistics and Trends*, `www.go-gulf.com/blog/cyber-crime`, retrieved March 21, (2015).

[7] Raymond Greenlaw, H. James Hoover, and Walter Larry Ruzzo. *Limits to Parallel Computation: P-Completeness Theory*, *Oxford University Press*, (1995).

[8] Sun-Yuan Hsieh and Ting-Yu Chou. Finding a Weight-constrained Maximum-density Subtree in a Tree. *Algorithms and Computation, Lecture Notes in Computer Science*, **3827**:944–953, Springer, Berlin, (2005).

[9] Robert Johnston and Clint LaFever. `Hacker.mil`, Marine Corps Red Team (PowerPoint Presentation). (2012).

[10] Hoong Chuin Lau, Trung Hieu Ngo, and Bao Nguyen Nguyen. Finding a Length-constrained Maximum-sum or Maximum-density Subtree and Its Application to Logistics. *Discrete Optimization*, **3(4)**:385–391, (2006).

[11] Fred B. Schneider. Blueprint for a Science of Cybersecurity, *The Next Wave*, **19(2)**:47–57, (2012).

[12] Hsin-Hao Su, Chin Lung Lu, and Chuan Yi Tang. An Improved Algorithm for Finding a Length-constrained Maximum-density Subtree in a Tree. *Information Processing Letters*, **109(2)**:161–164, (2008).

[13] Nelson A. Uhan. Stochastic linear programming games with concave preferences. *European Journal of Operations Research*, **243(2)**:637–646, (2015).

[14] R. Kevin Wood. Deterministic network interdiction. *Mathematical and Computer Modelling*, **17(2)**:1–18, (1993).

Received 10th June 2016