Proof of Theorem 1

Proof. Consider a threat screening game. As stated in the main paper, the TSG naturally separates into time windows w with the given resources type capacities and screenee numbers for each window. Assume we have $|W|$ such time windows and as a result we can construct $|W|$ sub-games, one for each time window.

In each sub-game G^w the adversary chooses a screenee category to pose as during screening and chooses an attack method. The Strong Stackelberg equilibrium of sub-game G^w can be represented as a probability distribution $\mathbf{q}^w$ over the valid pure strategies for time window w, $\hat{P}^w$. Also, there is an adversary action space defined as $A^w \subseteq C^w \times M$, from which the adversary finds one screenee category / attack method pair $a \in A^w$ the most attractive. Let U_s^{w*} and U_θ^{w*} be the utility of the screener and adversary respectively in the Strong Stackelberg equilibrium of each sub-game in time window w, and because of the zero-sum-like assumption $U_\theta^{w*} = -U_s^{w*}/c$.

We wish to show that $\mathbf{q}$ and the set of a such that $a \in \cup_{w\in\hat{w}} A^w$, where $\hat{w} = \text{argmax}_w\, U_\theta^{w*}$, is a Strong Stackelberg equilibrium of the original game with the screener utility being $\max_{w\in\hat{w}} U_s^{w*}$.

To show that it is a Strong Stackelberg equilibrium, we first show that, given the screener strategy, the set of $a \in \cup_{w\in\hat{w}} A^w$ that the adversary will choose from is indeed the best response. Next, we show the contradiction that if there is any other probability distribution achieving higher utility for the screener, then we can improve the utility for at least one of the sub-games.

First, note that the utility of the adversary and screener when any a is chosen in time window w depends only on the allocation of the resources available in that time window to the screenee categories in the same time window, which is given by the probability distribution over pure allocations. If the adversary chooses any $a \in \cup_{w\in\hat{w}} A^w$, where $\hat{w} = \text{argmax}_w\, U_\theta^{w*}$, then he obtains payoff $\max_w U_\theta^{w*}$. This is because for any $a \in A^w$, given distribution q^w on the allocation, we know that the adversary would obtain payoff U_θ^{w*}, since $\mathbf{q}^w$ is the screener's Stackelberg strategy for this sub-game. If the adversary chooses any other a in time window w he obtains payoff $< U_\theta^{w*}$, thus, overall for any other choice of a (other than $\cup_{w\in\hat{w}} A^w$) the adversary obtains payoff $< U_\theta^{w*}$. This means that the choices in $\cup_{w\in\hat{w}} A^w$ form the best response set for the adversary against the overall allocation $\mathbf{q}$.

Next, since $\hat{w} = \text{argmax}_w\, A^w$, we know (from the zero-sum-like assumption) that for any $w \in \hat{w}$, U_s^{w*} is the lowest equilibrium screener utility over all w. Thus, the screener's utility in equilibrium is $\min_w U_s^{w*}$. Suppose for contradiction, there is a probability distribution $\mathbf{q}_M$ that provides utility $U_s > \min_w U_s^{w*}$. Let this utility be achieved when the adversary chooses screenee category / attack method pair given by a. As the game is zero-sum-like the adversary gets his max utility from choosing a, which is $U_\theta = -U_s/c$. Thus, using $U_s > \min_w U_s^{w*}$, we have $-U_\theta > \min_w U_s^{w*}/c$ or rearranging $U_\theta < \max_w U_\theta^{w*}$. Suppose $w_0 \in \text{argmax}_w\, U_\theta^{w*}$, hence $U_\theta^{w_0*} = \max_w U_\theta^{w*}$ and $U_s^{w_0*} = \min_w U_s^{w*}$. Then, since the adversary's best response is to choose a, his utility in choosing any other a' in time window w_0 is $\leq U_\theta$. Consider the marginal of $\mathbf{q}_M$ for time window w_0, given by $q_M^{w_0}$. For the distribution $q_M^{w_0}$ the adversary's utility by playing his best response is some U'_θ where $U'_\theta \leq U_\theta < \max_w U_\theta^{w*}$. Thus, the screener obtains utility $U'_s = -cU'_\theta > -\max_w cU_\theta^{w*} = \min_w U_s^{w*} = U_s^{w_0*}$. That is, the screener can obtain utility better than the Stackelberg utility $U_s^{w_0*}$ with marginal distribution $q_M^{w_0}$ in the sub-game for time window w_0, which is a contradiction. □

Suboptimal Decomposition

Consider a TSG with two windows w_1 and w_2, four screenee categories c_1, c_2, c_3, c_4 each consisting of a single screenee with c_1 and c_2 arriving during w_1 and c_3 and c_4 arriving during w_2, one attack method m, one screening resource r with $C_r^w = 1$, and two screening teams $t_1 = \{r\}$ and $t_2 = \{\emptyset\}$ with $E_m^{t_1} = 1$ and $E_m^{t_2} = 0$. We decompose that TSG into two sub-games G^{w_1}, G^{w_2}, with the screener and adversary utilities shown below.

G^{w_1}	c_1		c_2	
	d	u	d	u
U_s	0	-20	0	-22
U_θ	3	5	3	5

G^{w_2}	c_3		c_4	
	d	u	d	u
U_s	0	-2	0	-3
U_θ	2	5	2	5

The optimal strategy for the screener in G^{w_1} is $\{P^{w_1}_{c_1,t_1} = 1$ with probability 0.5 and $P^{w_1}_{c_2,t_1} = 1$ with probability $0.5\}$. The adversary chooses c_1 and gets a payoff of 4 while the screener gets a payoff of -10. Similarly, the optimal strategy for screener in G^{w_2} is $\{P^{w_2}_{c_3,t_1} = 1$ with probability 0.5 and $P^{w_2}_{c_4,t_1} = 1$ with probability $0.5\}$. The adversary chooses c_3 and gets a payoff of 3.5 while the screener gets a payoff of -1.

Now, observe that simply combining the strategies of the two sub-games would cause the adversary to choose c_1 and get a payoff of 4, while the screener gets a payoff of -10. However, if in G^{w_2}, the strategy is changed to $\{P^{w_2}_{c_4,t_1} = 1$ with probability $1\}$, then the adversary chooses c_3 and gets a payoff of 5, while the screener gets a payoff -2. Thus, simply combining the equilibrium of the two sub-games is not the most optimal solution.

Proof of Theorem 2

Proof. We do a reduction from the independent set problem. First, the independent set problem can be given as an integer program: have one variable $P_t \in \{0, 1\}$ for each vertex, for each edge t, t' add a constraint $P_t + P_{t'} \leq 1$, which specifies choosing one vertex (by setting vertex variable to 1) linked by this edge. Then, the objective $\sum_t P_t$ finds the maximum independent set.

The core hardness in $MixedStrategyLP$ is due to team formation. Thus, we work with the special case with one screenee category c and one attack method m. The adversary's choice is obviously to make the single choice available of screenee category c and attack method m, and con-

straint (5) in $MixedStrategyLP$ becomes

$$x_{c,m}N_c = \sum_{P\in\hat{P}} q_P \sum_{t\in T} E_m^t P_{c,t}$$

Also, as there is just one θ, thus, the objective is just

$$\max_{x_{c,m},q_P} x_{c,m}(U_{s,c}^d - U_{s,c}^u)$$

Then the optimization problem can be stated as follows (by eliminating x)

$$\begin{array}{ll} \max_{q_P} & \frac{U_{s,c}^d - U_{s,c}^u}{N_c} \sum_{P\in\hat{P}} q_P \sum_{t\in T} E_m^t P_{c,t} \\ \text{subject to} & \sum_{P\in\hat{P}} q_P = 1, q_P \geq 0 \end{array}$$

As is well known, the extreme points of feasible space (probability simplex) of this LP is one of the integral points in which $q_P = 1$ for some P. This is equivalent to choosing a pure strategy P. Thus, the maximum value of this LP is same as that of the integer LP with the objective $\max_P \frac{U_{s,c}^d - U_{s,c}^u}{N_c} \sum_{t\in T} E_m^t P_{c,t}$ and constraints given by Equations 1 and 2: $\sum_{t\in T\setminus T^*} I_r^t P_{c,t} \leq L_r$, $\sum_{t\in T} P_{c,t} = N_c$, $P_{c,t}$ is a non-negative integer. Thus, in the reduction in the next paragraph we reduce to this integer LP formulation.

Given an independent set problem with V vertices, we construct a TSG with $T = \{1,\ldots,V+1\}$ team types (with variable $P_{c,t}$), where each team type in $1,\ldots,V$ corresponds to a vertex. The $V+1^{th}$ team is special in the sense that it does not correspond to any vertex and it is made up of just one resource with a very large capacity. Choose $N_c = |T|$. Choose $U_{s,c}^d - U_{s,c}^u(1) = |T|$ and efficiencies $E_m^t = 1$ for all teams, except $E_m^t = 0$ for $t = V+1$. Then, the objective of the integer LP is $\sum_{t\in T\setminus V+1} P_{c,t}$. Also, we can treat the number of screenees screened by team type $V+1$ as slack variables, and instead of the equality constraint $\sum_{t\in T} P_{c,t} = N_c$ consider the constraint $\sum_{t\in T\setminus V+1} P_{c,t} \leq N_c$ in the integer LP.

For each vertex, assign a resource type r_t associated with it. For each edge, assign a resource type $r_{t,t'}$, which is included in team t and t'. Take capacity L_r for each r to be 1. The resource types r_t yields constraint $P_{c,t} \leq 1$ and resource types $r_{t,t'}$ for each edge yields constraint $P_{c,t} + P_{c,t'} \leq 1$ for the integer LP. The variables $P_{c,t}$ can take values in $0,1$, the constraint $\sum_{t\in T\setminus V+1} P_{c,t} \leq N_c$ is redundant and any integral setting of $P_{c,t}$ for $t = 1,\ldots,V$ specify an independent set. As the objective maximizes $\sum_{t\in T\setminus V+1} P_{c,t}$, the maximum of the integer LP is the size of the maximum independent set. □

Non-Implementable Marginal Strategy

Consider 10 resource types that come together to form 5 teams types with each resource type being present in a unique set of 3 teams types. For each resource type $r \in R$, we set $L_r = 8$. Also, there is only one screenee category c with $N_c = 12$. Valid pure strategies must satisfy Equations 1 and 2. Equation 1 yields the following constraint:

$$\forall \text{ distinct } i,j,k.\ P_{c,t_i} + P_{c,t_j} + P_{c,t_k} \leq 8$$

where t_i, t_j, and t_k are the three teams types that contain a given resource type r. Since we have only one screenee category with $N_c = 12$, Equation 2 yields the following constraint:

$$\sum_{i=1}^{5} P_{c,t_i} = 12$$

Similarly, for marginal strategies we obtain the following constraints from Equations 8 and 9:

$$\forall \text{ distinct } i,j,k.\ n_{c,t_i} + n_{c,t_j} + n_{c,t_k} \leq 8$$

$$\sum_{i=1}^{5} n_{c,t_i} = 12$$

A feasible marginal strategy is then $\mathbf{n} = (\frac{8}{3},\frac{8}{3},\frac{8}{3},\frac{8}{3},\frac{4}{3})$. However, $\mathbf{n}$ has no equivalent mixed strategy $\mathbf{q}$, and thus cannot be formed as a convex combination of valid pure strategies. To see that $\mathbf{n}$ is not in the convex hull, note that the maximum value of $n_{c,t_1} + n_{c,t_2} + n_{c,t_3} + n_{c,t_4}$ in the convex hull must lie at a vertex, and hence must be integral. This maximum is $3+3+2+2 = 10$. But, for $\mathbf{n}$ that sum is $n_{c,t_1} + n_{c,t_2} + n_{c,t_3} + n_{c,t_4} = 32/3$, hence $\mathbf{n}$ cannot belong to convex hull.

Proof of Theorem 3

First, we define a canonical two-sided constraint structure (CTS). We say H is a canonical two-sided constraint structure if H contains all rows (i.e., sets of the form $\{c\} \times T$ for each $c \in C$) and all columns (i.e., sets of the form $C \times \{t\}$ for each $t \in T$).

Proof. We start with a constraint structure H for a given TSG. By definition, H contains all rows as Equation 9 creates a screenee category assignment constraint of the form $\{c\} \times T$ for each $c \in C$. However, H does not contain all columns by definition as the resource type capacity constraints from Equation 8 cover multiple columns, i.e., the set of columns corresponding to the team types that contain a given resource type r. Therefore, we include additional constraints such that H forms a CTS. By definition, the following must hold for each team type t:

$$\sum_{c\in C} n_{c,t} \leq \min_{\{r|I_r^t=1\}} L_r$$

That is, team type t can never be used more than the resource type r contained in t, i.e., $I_r^t = 1$, with the lowest capacity L_r. Thus, such a constraint can be added to H for each column, i.e., team type t, without altering the initial feasbile marginal strategy space. Adding these constraints on the columns results in H forming a CTS.

□

Proof of Lemma 1

Proof. First, any point $\mathbf{n}_i \in \mathbf{n}(H_i)$ can be written as the convex combination of integral allocations $P_1^i,\ldots,P_k^i$ for some k, where $P_j^i \in \mathbf{n}(H_i)$. This is true since a bi-hierarchy implies universal implementability (Budish et al. 2013). And since $\mathbf{n}(H_i) \subset \mathbf{n}(H)$, all of $P_1^i,\ldots,P_k^i$ lie

within $\mathbf{n}(H)$. Further, any point $\mathbf{n} \in conv(\mathbf{n}(H_1), \ldots)$ can be written a convex combination of one point, say $\mathbf{n}_i$, in each $\mathbf{n}(H_i)$. Now expanding $\mathbf{n}_i$ as a convex combination of integral allocations as stated above reveals that $\mathbf{n}$ can be written as a convex combination of integral allocations all lying in $\mathbf{n}(H)$. Thus, $\mathbf{n}$ is implementable. □

Proof of Theorem 4

Proof. Since the space of implementable strategies is a subset of the space of marginal strategies, if the implementable solution $\mathbf{n}'$ is same as the optimal marginal $\mathbf{n}^*$, then $\mathbf{n}'$ is also optimal. Also, if the tight resolution is never used, then in the single leaf we have a bihierarchy that contains $\mathbf{n}^*$, since the first two resolutions preserve $\mathbf{n}^*$. Thus, the solution $\mathbf{n}'$ in this case must provide same utility as $\mathbf{n}^*$ (or else we can choose the feasible point $\mathbf{n}^*$ to improve the solution quality).

The sampling technique is used after the MGA outputs $\mathbf{n}', \lambda_i$'s, $\mathbf{n}_i$'s are obtained. As stated in the convex hull constraints, we have $\mathbf{n}' = \sum_i \mathbf{n}_i$ or in other words, $\mathbf{n}' = \sum_i \lambda_i(\mathbf{n}_i/\lambda_i)$. That is, $\mathbf{n}'$ is the convex combination of $\mathbf{n}^i/\lambda_i$ for different i's. Further, since $A_i\mathbf{n}_i \leq \lambda_i b$, we know that $\mathbf{n}_i/\lambda_i \in \mathbf{n}(H_i)$. Thus, since H_i is a bihierarchy, by (Budish et al. 2013) $\mathbf{n}_i/\lambda_i$ is a convex combination of pure strategies $P_1^i, \ldots, P_k^i$ for some k with coefficients μ_k, where $P_j^i \in \mathbf{n}(H_i)$. As each $\mathbf{n}(H_i) \subset \mathbf{n}(H)$, all $P_k^i \in \mathbf{n}(H)$. Thus, sampling from a convex combination of pure strategies with coefficients $\lambda_i\mu_k$ for all i, k is equivalent to first sampling the set of pure strategies corresponding to λ_i and then from those sample a single pure strategy using coefficients μ_k. □

Proof of Theorem 5

Proof. First, we show that in each application of full resolution, we do not lose any integral (pure) strategies. Take any pure (integral) strategy P that satisfies $\mathbf{n}[S_r] \leq L_r$. Since, each $P_{c,t}$ is integral $P[S_{r,r'}]$ is also an integer. Thus, $P[S_{r,r'}]$ and $P[S_r \backslash S_{r,r'}]$ must satisfy $\mathbf{n}[S_{r,r'}] \leq L_r - i$, $\mathbf{n}[S_r \backslash S_{r,r'}] \leq i$ for some $i \in \{0, 1, \ldots, L_r\}$.

Suppose all the feasible pure strategies are $P_1, \ldots, P_k$, and assume at the end of FRA we obtain bihierachies $H_1, \ldots, H_l$. Then, we show that $conv(\mathbf{n}(H_1), \ldots, \mathbf{n}(H_k)) = conv(P_1, \ldots, P_k)$, and since $conv(P_1, \ldots, P_k)$ is precisely the set of implementable strategies, we will obtain the optimal implementable solution by using $conv(\mathbf{n}(H_1), \ldots, \mathbf{n}(H_k))$.

As argued earlier each P_i belongs to at least one of the bihierarchies, since we do not lose P_i in the full resolution step. Thus, all P_i belong to $conv(\mathbf{n}(H_1), \ldots, \mathbf{n}(H_k))$, and since the convex hull of some points is the smallest convex set containing those points, we have $conv(P_1, \ldots, P_k) \subseteq conv(\mathbf{n}(H_1), \ldots, \mathbf{n}(H_k))$. Also, by Lemma 1, we get that $conv(\mathbf{n}(H_1), \ldots, \mathbf{n}(H_k))$ is subset of implementable strategies, i.e $conv(\mathbf{n}(H_1), \ldots, \mathbf{n}(H_k)) = conv(P_1, \ldots, P_k)$. Thus, we have proved $conv(\mathbf{n}(H_1), \ldots, \mathbf{n}(H_k)) = conv(P_1, \ldots, P_k)$.

□