Balancing Congestion and Security in the Presence of Strategic Applicants with Private Information

Xiaofang Wang  
Renmin University of China

Jun Zhuang  
University of Buffalo, The State University of New York, jzhuang@buffalo.edu

Follow this and additional works at: http://research.create.usc.edu/published_papers

Recommended Citation
http://research.create.usc.edu/published_papers/102

This Article is brought to you for free and open access by CREATE Research Archive. It has been accepted for inclusion in Published Articles & Papers by an authorized administrator of CREATE Research Archive. For more information, please contact gribben@usc.edu.
Balancing congestion and security in the presence of strategic applicants with private information

Xiaofang Wang, Jun Zhuang

School of Business, Renmin University of China, China
Department of Industrial and Systems Engineering, University at Buffalo, State University of New York, USA

Abstract

Concerns on security and congestion appear in security screening which is used to identify and deter potential threats (e.g., attackers, terrorists, smugglers, spies) among normal applicants wishing to enter an organization, location, or facility. Generally, in-depth screening reduces the risk of being attacked, but creates delays that may deter normal applicants and thus, decrease the welfare of the approver (authority, manager, screener). In this paper, we develop a model to determine the optimal screening policy to maximize the reward from admitting normal applicants net of the penalty from admitting bad applicants. We use an M/M/1 queueing system to capture the impact of security screening policies on system congestion and use game theory to model strategic behavior, in which potential applicants with private information can decide whether to apply based on the observed approver’s screening policy and the submission behavior of other potential applicants. We provide analytical solutions for the optimal non-discriminatory screening policy and numerical illustrations for both the discriminatory and non-discriminatory policies. In addition, we discuss more complex scenarios including imperfect screening, abandonment behavior of normal applicants, and non-zero waiting costs of attackers.

1. Introduction

Security screenings play an important role in many fields, including visa issuance, cargo inspection, and airport security screening. In-depth examination of applicants reduces security risk, but can entail high congestion which may deter normal applications and in turn may conflict with the approver’s interests. Since September 11, 2001, the security levels for the United States visa issuance have been greatly tightened, significantly affecting regular academic exchanges, and the interactions that go along with these exchanges such as work obligations, collaboration networking, and family reunions (Seife and Ding, 2003; Bhattacharjee, 2004a,b; Bagla, 2006; Stone, 2008). For example, the US Government Accountability Office (2004) states that the average waiting period for a visa security clearance is 67 days, indicating that “students and scholars with science backgrounds might decide not to come to the US, and technological advancements that serve the US and global interests could be jeopardized”.

Concerns on security and congestion also appear in the setting of container security at ports of entry. In airport security, as early as the 1970’s, researchers have applied queueing models to study airport congestion resulting from security screenings (Gilliam, 1979). Since September 11, 2001, many researchers have studied airport congestion due to the enhanced screening. Blalock et al. (2007) find that the tightened post-9/11 baggage-screening policy has reduced passenger volume by eight percent. Following the attempted terrorist attack to Northwest Airlines Flight 253 on December 25, 2009, the US started a new screening policy requiring citizens of 14 nations (including Pakistan, Saudi Arabia and Nigeria), who are flying to the US, to be subject to special screening at airports worldwide (Lipton, 2010). This nationality-based discriminatory screening policy has been widely criticized (Fisher, 2010), while its effectiveness remains unclear.

Motivated by the above examples, we apply game theory and queueing theory in this paper to study the approver’s optimal
screening policy, considering both system congestion and the strategic behavior of potential applicants with private information. Each potential applicant could be either a terrorist (bad type) or not (good type), which is his private information and not directly observable by the approver. The approver’s screening policy can be non-discriminatory, or discriminatory based on the applicants’ observable attributes (such as age, nationality, gender, and security history). There might exist some correlation between the unobservable information and the observable attributes.

The novelty of this research, compared to previous studies, is that it allows strategic decisions by all types of potential applicants. In particular, potential applicants could adapt their behavior according to a disclosed security policy (Zhuang and Bier, 2007). For example, smugglers may choose the weakest port to enter; leisure travelers may choose not to travel because of congestion; and foreign students may no longer apply to the US schools because of the long waiting period for visas.

To our knowledge, no research to date has simultaneously considered both the applicant’s strategic behavior and congestion in determining the optimal screening policy (with an exception of a recent paper by Bakshi and Gans (2010), which considers congestion and the terrorist’s strategic choice but does not consider the normal people’s choice); our research aims to fill this important theoretical gap. In particular, we first analytically derive the approver’s optimal non-discriminatory screening policy, and then use numerical examples to illustrate the benefit of discriminatory policies based on observable attributes.

We recognize that there exists close work on security screening using game theory and economic modeling. For example, in the setting of container security at ports of entry, Bier and Haphuriwat (in press) study the optimal proportion of containers to inspect using game theory to capture the fact that attackers are simultaneously trying to maximize their expected rewards. In critical infrastructure protection, optimal defense investment in the face of adaptive terrorists has been studied (Golany et al., 2009; Zhuang and Bier, 2007; Zhuang et al., 2010; Golalikhani and Zhuang, forthcoming). In the economics literature, Basuchoudhary and Razzolini (2006) and Yetman (2004) have studied the idea of discriminatory screening. In contrast to our paper, all of these papers fail to consider congestion or normal applicants’ decisions.

In recent years, selective screening based on passenger profiling has been proposed to reduce the inconvenience caused to normal passengers and improve the effectiveness of screening devices. McLay et al. (2006) show that using different screening procedures on different passenger types (passenger risk information) can lead to more effective security screening strategies. Babu et al. (2006) study the effects of classifying passengers into different groups and applying different levels of screening to them. Nie et al. (2009) extend the work of Babu et al. (2006) by allowing passengers to have heterogeneous threat levels. Recently, Cavusoglu et al. (2010) study the impacts of profiling on the design of the screening systems and other performance measures, such as the reliability of screening device signals and the inconvenience caused to normal passengers. However, these papers do not consider strategic passengers’ behavior. Along the same line, in other application domains which are different from transportation security, Viane et al. (2007) study the optimal screening policy for detecting fraudulent automobile insurance claims, and Claeyts et al. (2010) study group screening policies in a queueing model with dynamic arrivals.

In the service operations management literature, there exists some work using a game-theoretic formulation of a queueing problem to study the interaction between the service provider and the customers who consider congestion when deciding whether or not to queue (for an excellent overview, see Hassin and Haviv, 2003). This formulation proves to be effective in modeling service systems with customer behaviors and thus, fits for our settings. An example using this formulation is Anand et al. (2011) which studies the similar tradeoff between quality (analogous to security in the above literature) and congestion trade-off in a setting where customers decide whether to use the service based on their evaluations of quality and speed. As in Anand et al. (2011), we endogenize the trade-off between quality and congestion into individual decisions. But in comparison, Anand et al. (2011) assume the service provider directly decides service speed and price to maximize her revenue, while our work assumes that the service provider determines screening probabilities and service speed is a parameter, not a decision. In addition, the quality/time interaction that Anand et al. (2011) outline is modeled explicitly as a monotonic function: as the service speed increases, the quality of the service decreases which depresses demand. In contrast, the quality (security)/time interaction in our paper lies implicitly in the fact that the more applicants are screened, the more bad applicants can be caught while more congestion is created at the approver. Due to the nature of the security context, we address differentiated screening strategies in the presence of heterogeneous customers (potential applicants or travelers) with observable information which is not discussed by Anand et al. (2011).

The remainder of the paper is organized as follows. A description of the model is presented in Section 2. Section 3 provides an analysis of the best responses of potential applicants. Section 4 provides an analytical solution for the approver’s optimal non-discriminatory policy. Section 5 then provides numerical analysis of both discriminatory and non-discriminatory policies and indicates under which conditions discriminatory policies are significantly better. Section 6 discusses and analyzes several extensions of our model. Section 7 concludes this paper and provides future research directions. Finally, the appendix provides proofs to the propositions in this paper.

2. The model

We consider a manager, which we refer to as the “approver,” facing a population of strategic potential applicants willing to enter a secure organization or facility. Note that in this paper we use “potential applicant” to represent the people who have intentions to apply but may or may not submit their applications. In contrast, we use “applicant” to represent the people who have already submitted their applications. The approver decides whether to screen each applicant based on observable attributes such as age, nationality, gender, occupation, and security history. The approver can immediately approve an application, in which case the applicant passes. On the other hand, if the approver decides to screen the applicant, the applicant will be placed in a queue to wait for the result of the screening process (either approval or rejection). Fig. 1 shows a flowchart of the process.

The approver’s goal is to reduce the risk of admitting bad applicants (attackers), while simultaneously reducing the inconvenience (congestion) to good applicants. Screening numerous applicants decreases the chance of admitting a bad applicant, but increases waiting times, which could decrease the willingness of good potential applicants to apply. All potential applicants are
assumed to decide whether to submit an application by weighing the costs and benefits, according to the announced screening policy. In the following subsection, we outline our assumptions regarding characteristics of potential applicants, cost structure, screening, and service characteristics, and then formulate the approver's optimization problem (taking into account the potential applicants' decisions).

2.1. Model setup

We assume: (1) the approver cannot immediately reject applicants; i.e., applicants can be rejected only after screening; (2) when the approver is indifferent between different levels of screening probabilities, she will choose the lowest level due to convenience/cost concerns. In Sections 2–5 we assume perfect screening; i.e., after screening, all of good applicants will pass and all of bad applicants will be caught; in Subsection 6.1, we relax this assumption by allowing non-zero screening errors. The elements of the model are described as follows.

Applicant characteristics. Each potential applicant can be one of two types, \( \theta \in \{g,b\} \), representing good and bad types, respectively. The type is known to the potential applicant himself, but unobservable to the approver. Each potential applicant also has an observable attribute that can take on one of two possible values, \( t \in \{1,2\} \). The probability that a potential applicant has attribute \( t = 1 \) is given by \( p \in [0,1] \). We assume that potential applicants are bad with probability \( x \) if \( t = 1 \) and with probability \( \beta \) if \( t = 2 \). The joint probability masses are \( p(1-x), px, (1-p)(1-\beta) \), and \( (1-p)\beta \), for a potential applicant of types 1g, 1b, 2g, and 2b, respectively.

Cost structure. We assume that the approver gains \( R \) for admitting each good applicant, and loses \( C \) for admitting each bad applicant. A good applicant receives a reward of \( r_g \) if he is approved and incurs a waiting cost of \( c_g \) per unit time if he is screened. A bad applicant receives a reward of \( r_b \) if he is approved and incurs a penalty of \( c_b \) if he is caught. To make analysis tractable, we assume that the reward \( r_b \) and the penalty \( c_b \) are sufficiently large such that bad applicants will pass eventually, regardless of whether they are screened or not, and the good applicants will pass only when they are not screened. Therefore the approver gets a reward of \( R \) from each good applicant and a penalty cost \( C \) from each bad applicant who is not screened.

Potential applicants’ strategy. Potential applicants behave as rational economic agents who maximize their expected utilities. Their utilities depend not only on their own decisions, but also on the decisions of other potential applicants and the decision of the approver. Each potential applicant has two options: to submit an application or not. Let \( p_t \in [0,1] \) denote the potential applicants' submission probability for \( t \in \{1,2\} \) and \( \theta \in \{g,b\} \) which is a pure strategy if \( p_t = 0 \), 1, and is a mixed strategy otherwise. The total traffic rate for screening is the summation over all screened applicants: \( \sum_{t=1}^{2} \sum_{\theta} \phi_t(p_{tg}A_{tg} + p_{tb}A_{tb}) \) where \( A_{tg} = p(1-x)A; \quad A_{tb} = pxA; \quad A_{2g} = (1-p)(1-\beta)A; \quad A_{2b} = (1-p)\beta A. \) We denote the potential applicants’ strategies by \( p = (p_{1g}, P_{1b}, p_{2g}, P_{2b}) \). Thus, the arrival rates of applicants are \( \lambda = (\lambda_{1g}, \lambda_{2g}, \lambda_{1b}, \lambda_{2b}) \), where \( \lambda_{1g} = p_{1g} A_{tg}, \text{for} \ t \in \{1,2\} \), and \( \lambda_{1b} = p_{1b} A_{tb} \).

For a good potential applicant, his utility of not submitting an application is zero and his utility of submitting an application is the reward minus the expected waiting cost. Thus, his expected utility is:

\[
W(\Phi, \bf{p}) = \frac{1}{\mu - \sum_{t=1}^{2} \sum_{\theta} \phi_t(p_{tg}A_{tg} + p_{tb}A_{tb})},
\]

Similarly, a bad potential applicant's expected utility is:

\[
u_{2b}(\Phi, \bf{p}) = p_{2b}[\mu(1 - c_b) + (1 - \Phi_t)r_b], \quad \text{for} \ t \in \{1,2\}.
\]

2.2. Definition of equilibrium

Combining good and bad potential applicants' utilities in Eqs. (2) and (4), for a given approver's strategy \( \Phi \), the objective function of a potential applicant with an observable attribute \( t \) and an unobservable type \( \theta \) is:

\[
\max_{p_t} u_{\theta}(\Phi, \bf{p}) = \begin{cases} 
\sum_{t=1}^{2} \sum_{\theta} \phi_t(p_{tg}A_{tg} + p_{tb}A_{tb}) & \text{if } \theta = g \ni \Phi_t \in [0,1] \\
\sum_{t=1}^{2} \sum_{\theta} \phi_t(1 - c_b) + (1 - \Phi_t)r_b & \text{if } \theta = b. 
\end{cases}
\]

Given \( \Phi \), we define the potential applicants' best response:

\[
\bar{p}_{\theta}(\Phi) = \arg \max_{p_{t}} u_{\theta}(\Phi, \bf{p}), \quad \text{for} \ t \in \{1,2\} \quad \text{and} \quad \theta \in \{g,b\}.
\]

In order to focus on the tradeoff between the congestion impact and the security impact on good/bad potential applicants, respectively, we ignore the approver’s screening cost which is proportional to \( \Phi \) due to labor, time or financial costs consumed during the screening process.
Having established our model, we now proceed to its analysis. In particular, we first derive best responses of potential applicants in Section 3, and then solve for the optimal strategy of the approver in Section 4.

3. Potential applicants’ best responses

3.1. Bad potential applicants’ best responses

**Proposition 1.** Bad potential applicants’ best responses are given by:

$$\hat{p}_b(\Phi) = \begin{cases} 1, & \Phi_t < s_b = \frac{s_b}{t_0+C_1}, \quad \text{for } t = 1, 2. \\ 0, & \Phi_t \geq s_b = \frac{s_b}{t_2+C_2}, \end{cases}$$

**Remark.** Bad potential applicants do not consider congestion in this section, so their strategies do not depend on the strategies of other potential applicants. Their strategies solely depend on a comparison between the probability of being screened and a threshold value $s_b$; they apply with probability one if the probability of being screened is sufficiently small (below $s_b$), and with probability zero otherwise. From the definition of $s_b$, we see that intuitively a bad potential applicant is more likely to apply when his reward after passing screening is high, or his penalty after being caught is low.

3.2. Good potential applicants’ best responses

A good potential applicant’s best response is more complex, because it depends on the decisions of other potential applicants through negative externality associated with congestion, as shown in Eqs. (2) and (3). If we fix the approver’s strategy $\Phi$ and bad potential applicants’ best responses $\hat{p}_b$, $t = 1, 2$, we are left with an equilibrium analysis problem on two classes of customers: good potential applicants with an attribute value of 1, and good potential applicants with an attribute value of 2. This equilibrium analysis problem has been solved in Balachandran and Schaefer (1980). As in Balachandran and Schaefer (1980), we first define the desired aggregate arrival rate of screened good potential applicants as follows:

Note that the traffic caused by bad applicants is:

$$\Phi_1 p_{1b} p_{1} A + \Phi_2 p_{2b} (1-p) \beta A.$$  

Assuming there are no good applicants with $t=2$, at the equilibrium, the maximum traffic of screened good applicants should be:

$$\hat{A}_{1g} = \mu - \Phi_1 p_{1b} p_{1} A - \Phi_2 p_{2b} (1-p) \beta A - \frac{\Phi_1 c_{w}}{r_g} + \frac{\Phi_2 c_{w}}{r_g}.$$  

which is derived from the fact that the good applicant of attribute 1 has zero utility at the equilibrium; $r_g - \Phi_1 c_{w} W - \frac{\Phi_1 c_{w}}{r_g} = 0.$

**Table 1** Major notations.

<table>
<thead>
<tr>
<th>Notation Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta \in (g,b)$ Unobservable type of a potential applicant</td>
</tr>
<tr>
<td>$t \in [1,2]$ Probability that a potential applicant has attribute 1</td>
</tr>
<tr>
<td>$p$ Probability that a potential applicant has attribute 1 is bad</td>
</tr>
<tr>
<td>$x$ Probability that a potential applicant with attribute 1 is bad</td>
</tr>
<tr>
<td>$\beta$ Probability that a potential applicant with attribute 2 is bad</td>
</tr>
<tr>
<td>$R$ Approver’s reward for admitting each good applicant</td>
</tr>
<tr>
<td>$C$ Approver’s penalty for admitting each bad applicant</td>
</tr>
<tr>
<td>$f_g$ Good applicant’s reward if passed</td>
</tr>
<tr>
<td>$f_b$ Bad applicant’s reward if passed</td>
</tr>
<tr>
<td>$c_w$ Good applicant’s cost per unit of waiting time</td>
</tr>
<tr>
<td>$c_b$ Bad applicant’s penalty if rejected</td>
</tr>
<tr>
<td>$A$ Poisson arrival rate of all potential applicants</td>
</tr>
<tr>
<td>$A_{0g}$ Poisson arrival rate of potential applicants with type $\theta$ and attribute $t$</td>
</tr>
<tr>
<td>$\mu$ Service rate of the screening process</td>
</tr>
<tr>
<td>$W(\Phi, p)$ Expected waiting time</td>
</tr>
<tr>
<td>$\Phi = (\Phi_1, \Phi_2)$ Approver’s strategy</td>
</tr>
<tr>
<td>$\Phi_1 \in [0,1]$ Probability of screening an applicant with attribute 1</td>
</tr>
<tr>
<td>$J(\Phi)$ Approver’s objective function</td>
</tr>
<tr>
<td>$\phi = (\phi_1, \phi_2)$ Arrival rates of applicants</td>
</tr>
<tr>
<td>$p = (p_1, p_2)$ Potential applicant’s submission probability</td>
</tr>
<tr>
<td>$u_0$ Potential applicant’s expected utility with attribute $t$ and type $\theta$</td>
</tr>
</tbody>
</table>

Similarly, if assuming there are no good applicants with $t=1$, then at the equilibrium, the maximum traffic of screened good applicants should be:

$$\hat{A}_{2g} = \mu - \Phi_1 p_{1b} p_{1} A - \Phi_2 p_{2b} (1-p) \beta A - \frac{\Phi_1 c_{w}}{r_g} + \frac{\Phi_2 c_{w}}{r_g} = 0.$$  

Thus, $\hat{A}_{1g}$ is the maximum demand rate “desired” by the good potential applicants with $t=1$. That is, when assuming there are no good applicants with $t=2$, the good potential applicants with $t=1$ would apply (utility is non-negative) unless their screened aggregate demand rate is more than $\hat{A}_{1g}$. When the demand rate reaches $\hat{A}_{1g}$, their utility becomes zero. Similarly, $\hat{A}_{2g}$ is the maximum demand rate “desired” by the good potential applicants with $t=2$, when assuming there are no good applicants with $t=1$.

Based on the above definitions of $\hat{A}_{1g}$ and $\hat{A}_{2g}$, Proposition 2 below specifies the best responses of good potential applicants. To simplify notation, we denote the Poisson arrival rate of all good potential applicants as $A_g$ and the Poisson arrival rate of all bad potential applicants as $A_b$. In other words, $A_g = A_{1g} + A_{2g} = [p(1-x) + (1-p)(1-\beta)]A$ and $A_b = A_{1b} + A_{2b} = [p x + (1-p)\beta]A$.

**Proposition 2.** Given the approver’s strategy $\Phi$ and bad potential applicants’ best responses $\hat{p}_b$ for $t=1, 2$:

(i) If $\Phi_1 = 0$, then $\hat{p}_{1g} = 1$; if $\Phi_2 = 0$, then $\hat{p}_{2g} = 1$.

(ii.a) If the approver uses a discriminatory policy and $0 < \Phi_1 < \Phi_2$, then the good potential applicants’ best responses are:

$$\hat{p}_{1g} = \frac{1}{\Phi_1 \hat{A}_{1g}} \max \left[ \frac{\hat{A}_{1g} - \Phi_1 \Phi_1 A_{1g} - \Phi_2 A_{2g}}{0}, \right]$$

$$\hat{p}_{2g} = \frac{1}{\Phi_2 \hat{A}_{2g}} \max \left[ \frac{\hat{A}_{1g} - \Phi_2 \Phi_1 A_{1g} - \Phi_2 A_{2g}}{0}, \right].$$

(ii.b) If the approver uses a discriminatory policy and $0 < \Phi_2 < \Phi_1$, then the good potential applicants’ best responses are:
Remark. First, note that the good potential applicants’ submission probabilities tend to increase in their corresponding desired aggregate arrival rates and decrease in screening probabilities and arrival rates of all potential applicants. Second, for the case \( \Phi_1 < \Phi_2 \), good potential applicants with \( t = 1 \) are less sensitive to congestion than those with \( t = 2 \). In other words, facing the same level of congestion, the good potential applicants with \( t = 1 \) have a higher utility and thus are more likely to apply than those with \( t = 2 \). Therefore, at the equilibrium, if none or a proportion of good potential applicants with \( t = 1 \) apply (i.e., \( A_1 < \Phi_1 A_1g \)), then none of those with \( t = 2 \) apply; if all of the good potential applicants with \( t = 1 \) apply (i.e., \( A_1 \geq \Phi_1 A_1g \)), then a proportion or all of the good potential applicants with \( t = 2 \) may apply. Similarly, if \( \Phi_1 > \Phi_2 \), good potential applicants with \( t = 2 \) are more likely to apply. When \( \Phi_1 = \Phi_2 \), the approver is indifferent between these two attribute values and thus good potential applicants behave the same.

### 4. Optimal screening policies

Substituting best responses of the potential applicants (as specified in Propositions 1 and 2) into the approver’s decision-making problem as specified in Eq. (1), we study the approver’s optimal strategy.

#### 4.1. Non-discriminatory screening policy

We start with a non-discriminatory screening policy that uses equal screening probabilities for applicants of two attributes: \( \Phi_1 = \Phi_2 = \Phi \).

**Proposition 3.** The approver’s optimal non-discriminatory strategy \( \Phi^* \) is specified in Table 2.

**Remark.** As shown in Table 2, the optimal non-discriminatory screening probability can only take one of these two values: \( \Phi^* = \frac{\alpha}{\alpha + \beta + \gamma + \delta} < s_b \) (in this case, all bad potential applicants apply, i.e., \( \mu_1 = 1 \)) and \( \Phi^* = s_b = \frac{\alpha}{\alpha + \beta + \gamma + \delta} \) (in this case, none of the bad potential applicants apply, i.e., \( \mu_1 = 0 \)). In general, the approver uses a screening probability smaller than \( s_b \) when the bad potential applicant has a high incentive to apply (high \( s_b \)) and when the reward of admitting a good applicant is relatively high compared to the penalty of admitting a bad one (high \( R \) and low \( C \)). In this case, \( \Phi^* \) increases in the arrival rate of all potential applicants \( A \); the approver screens a smaller proportion of applicants when the whole population expands. In contrast, when bad applicants are very dangerous to the approver (high \( C \)) and are deterred easily (low \( s_b \)), and the reward of admitting a good applicant is relatively low (low \( R \)) compared to the penalty of admitting a bad one (high \( C \)), the approver should use a screening probability which equals \( s_b \) to deter all the bad ones. In this case, \( \Phi^* \) increases in \( r_b \) and decreases in \( c_p \). Proposition 3 also implies that \( \Phi^* \) cannot be zero unless the bad applicant’s reward \( r_b \) or the service rate \( \mu \) equals zero.

#### 4.2. Discriminatory screening policy

Now we study the discriminatory policy. A general result is given in the following proposition.

**Proposition 4.** The approver’s optimal discriminatory policy is always better than any non-discriminatory policy.

**Remark.** Proposition 4 indicates that theoretically the approver’s optimal payoff under a discriminatory policy is higher than that under a non-discriminatory policy. The reason is that the approver can utilize observable information to disproportionately screen applicants. For example, if knowing applicants of attribute 1 are more likely to be bad, the approver will screen more of them and screen fewer applicants of attribute 2. Thus, on average, under a discriminatory policy, good applicants will experience a shorter waiting time, bad applicants will experience more strict screening, and the approver will get a higher payoff. In practice, however, there exist many reasons to justify a non-discriminatory policy. For example, due to equity/political concerns, the approver might not be allowed to implement a discriminatory policy. In addition, implementing this policy might be difficult due to a high cost of collecting and verifying observable data. Thus, to decide which policy to use, a good benefit/cost balance must be made.

Analysis of the optimal non-discriminatory screening policy in Section 4.1 implies the complexity to analytically derive the optimal discriminatory policy. Thus, we numerically calculate the optimal discriminatory policy \( (\Phi_1^*, \Phi_2^*) \) for a given set of parameter values in the next section.

### 5. Numerical experiments

In this section, we conduct numerical sensitivity analysis of our model and compare the discriminatory policy with the non-discriminatory policy. To highlight the most interesting cases, we provide sensitivity analysis only for parameters \( R, R_b, \) and \( A \). The baseline parameter values used in all numerical experiments in this paper are set as follows: \( p = .6; \alpha = .2; \beta = .02; R = 1; C = 10; r_b = 1; r_e = 2; c_p = 1; c_g = 1; A = 100; \) and \( \mu = 10 \). Note that in this baseline case, we have \( s_b = \frac{\alpha}{\alpha + \beta + \gamma + \delta} \approx 0.67 \).

#### 5.1. Changing the approver’s reward from admitting a good applicant \( R \)

For a discriminatory policy as shown in Fig. 2 (upper panel), when the approver’s reward from admitting a good applicant increases, the approver becomes more reluctant to screen applicants. In other words, the screening probabilities \( \Phi_1 \) and \( \Phi_2 \) are both non-increasing in \( R \). Thus, potential applicants (including the bad ones) become more likely to apply; i.e., \( p_{nv} \) (weakly) increases in \( R \) for \( t = 1, 2, \theta \in (g, b) \). The intuition is that when the reward from admitting a good applicant is higher, the approver is willing to reduce screening congestion and take a greater risk of admitting bad applicants.

For a non-discriminatory policy (middle panel in Fig. 2), the screening probability is also non-increasing in \( R \) and potential applicants (including the bad ones) become more likely to apply.
Comparing the approver’s payoffs for discriminatory and non-discriminatory policies (lower panel in Fig. 2), the difference is not significant for either large or small values of $R$, and only significant for medium values of $R$. The intuition is as follows. Recall that the benefit of a discriminatory policy is due to the reason that it can disproportionately screen applicants: screening fewer less risky ones to reduce congestion while screening more risky ones to lower the loss. When $R$ is sufficiently large ($>3$ in this example), it is more important to reduce congestion for attracting more good potential applicants than to reduce security risk for deterring bad potential applicants. Thus, the approver screens only a small percentage of applicants no matter under a discriminatory or a non-discriminatory policy. As a result, potential applicants’ submission behaviors are similar under these two policies. The benefit from a discriminatory policy is not obvious. When $R$ is sufficiently small ($<1$ in this example), reducing security risk by deterring bad potential applicants is relatively more important than reducing congestion to have more good potential applicants. Thus the screening probabilities are set high no matter under a discriminatory or a non-discriminatory policy. In summary, a discriminatory policy shows more advantage when the tension between congestion and security risk is more intensive.

### 5.2. Changing the parameter of bad potential applicants $r_b$

A high reward when approved (large $r_b$) motivates bad potential applicants to apply. First, look at the optimal discriminatory policy. In the example shown in Fig. 3, the optimal screening probability for applicants of attribute 1 is always $s_b = \frac{1}{r_b}$, which is increasing with $r_b$. Thus when $r_b$ increases, bad potential applicants of attribute 1 are all deterred. In contrast, the approver’s screening probability for applicants of attribute 2 is relatively small. Especially, when $r_b > 0.2$ in Fig. 3, the approver’s screening probability for applicants of attribute 2, $\phi_2$, is zero, and therefore, both good and bad potential applicants apply with probability 1: $p_{2g} = p_{2b} = 1$. This is because in this numerical example, an optimal discriminatory policy focuses on reducing the security risk from more risky applicants of attribute 1 while ignoring less risky applicants of attribute 2. Such a focus is more important when bad potential applicants are more likely to apply (for example, $r_b > 0.2$). Due to the increased congestion by screening more applicants of attribute 1, the submission probability of good potential applicants $p_{2g}$ is decreasing as $r_b$ increases.

Second, look at the optimal non-discriminatory policy. In the example shown in Fig. 3, the optimal screening probability for all applicants always follows $s_b = \frac{1}{r_b}$ Thus there are no bad applicants and submission probabilities of good potential applicants decrease as the screening probability increases.

Third, the difference in the approver’s payoffs under discriminatory and non-discriminatory policies is significant for large values of $r_b$. This is because when bad potential applicants have a greater incentive to apply, the approver in general screens more applicants, which results in higher congestion. Then it is more beneficial to screen fewer applicants with low risk to reduce congestion and to screen more applicants with high risk for security concern. This again confirms that a discriminatory policy shows more advantage when the tension between congestion and security risk is more intensive.

### 5.3. Changing the arrival rate of all potential applicants $\lambda$

A higher arrival rate of all potential applicants implies more congestion. First, look at the optimal discriminatory policy. In our numerical experiments on service capacity $\mu$. Since decreasing $\mu$ has the similar impact as increasing $\lambda$, we do not report its numerical results in this paper.
Numerical example, there are proportionally fewer bad applicants of attribute 2 than those of attribute 1. Thus, as shown in Fig. 4, when $A$ is sufficiently high ($A > 30$ in this example), in order to reduce congestion, the approver tends to ignore less risky applicants of attribute 2 by setting $\Phi_2 = 0$ and focus on screening more risky applicants of attribute 1 by setting screening probability $\Phi_1$ at the threshold level $s_b \approx 0.67$. Thus, at the equilibrium, potential applicants of attribute 2 apply and get approved immediately while bad potential applicants of attribute 1 are all deterred. In addition, good potential applicants of attribute 1 will apply with a positive submission probability which decreases as congestion ($A$) increases.

Second, look at the optimal non-discriminatory policy. Screening probabilities are always set high to fully deter all bad potential applicants. Similar with the discriminatory case, there are fewer good applicants as $A$ increases due to increased congestion.

Third, comparing the optimal discriminatory and non-discriminatory policies, we can see that the difference is small when congestion is low and the difference increases linearly with $A$ when congestion is high. The intuition is as follows. Under high congestion, the approver’s payoff using a non-discriminatory policy stays constant as $A$ increases; the traffic for screening is composed only by good applicants, and their aggregate demand rate is the desired aggregate arrival rate which is determined by service capacity and other parameters excluding $A$ (see Eq. (10) when $p_{1b} = p_{2b} = 0$). Using a discriminatory policy, since $\Phi_1 = s_b$ and $\Phi_2 = 0$, bad potential applicants of attribute 1 are all deterred and potential applicants of attribute 2 all submit applications and are admitted without being screened. Then the traffic for screening is composed only by the good applicants of attribute 1, and their aggregate demand rate is determined by parameters excluding $A$ (see Eq. (10))
when $\Phi_2 = 0$ and $p_B = 0$). Thus using a discriminatory policy, a payoff to the approver is composed by the total reward from admitting good applicants of attribute 1 which is a constant and the total reward net of penalty from admitting applicants of attribute 2 which is linearly increasing with $\lambda$. Therefore, as $\lambda$ increases, a discriminatory policy shows more advantage, which linearly increases.

### 6. Extensions of our model

In this section, we extend our model to the following situations: 

1. screening is imperfect, i.e., there exist errors that good applicants are identified as bad and bad applicants are identified as good (subsection 6.1). 

2. The approver considers that some proportion of good applicants may withdraw their applications before the approval/rejection decision, and this proportion depends on the waiting time (subsection 6.2). 

3. Bad potential applicants may factor waiting costs into their submission decisions (subsection 6.3).

#### 6.1. Imperfect screening

In this subsection we study imperfect screening. In particular, we let $e_g$ denote the error probability that good applicants are identified as bad and let $e_b$ denote the error probability that bad applicants are identified as good. Then the approver's objective function (1) becomes:

$$
\max_{\Phi} J(\Phi) = \max \{ \lambda g_11 - e_g \Phi_1 + \lambda b_2(1 - e_b \Phi_2) | R - \{ (1 - C_1) \Phi_1 \lambda g_1 + (1 - (1 - e_b) \Phi_2) \lambda b_2 | \}, \tag{12}
$$

where the approver gets the reward $R$ from each approved good applicant and gets the penalty $C$ from each approved bad applicant. The potential applicant's objective function (5) becomes:

$$
\max_{p_B}(\Phi, p) = \begin{cases}
  p_B [(1 - e_g \Phi_1) R - \Phi_2 C w(W | \Phi, p)] & \text{if } \theta = g, \\
  p_B [(1 - (1 - e_b) \Phi_2) R + (1 - e_b) \Phi_1 (C_0)] & \text{if } \theta = b.
\end{cases} \tag{13}
$$

#### 6.1.1. Bad potential applicants' best responses

With imperfect screening, Proposition 1 becomes:

**Proposition 5.** Bad potential applicants' best responses are given by:

$$
p_B(\Phi) = \begin{cases}
  1, & \Phi_t < s_b = \frac{\lambda g_1}{(1 - e_g \Phi_1) C_0} \\
  0, & \Phi_t \geq s_b = \frac{\lambda b_2}{(1 - e_b) \Phi_2 C_0}
\end{cases} \tag{14}
$$

**Remark.** Note that the error probability $e_b$ plays an important role in imperfect screening. Compared to the perfect screening model, with imperfect screening bad potential applicants are more likely to submit their applications as the threshold probability $s_b = \frac{\lambda g_1}{(1 - e_g \Phi_1) C_0}$ is higher than that of perfect screening, $s_b = \frac{\lambda g_1}{\Phi_1 C_0}$. When screening precision is low ($e_b > \frac{\lambda g_1}{\Phi_1 C_0}$), $s_b > 1$ and then all bad potential applicants submit their applications no matter what the screening probability is. So in order to deter some bad potential applicants, the error probability $e_b$ must be controlled within $[0, \frac{\lambda g_1}{\Phi_1 C_0}]$.

#### 6.1.2. Good potential applicants’ best responses

According to the updated utility functions defined in (13), similarly as the perfect screening case, we derive the desired aggregate arrival rate of screened good potential applicants:

$$
\hat{A}_{1g} = \mu - \Phi_1 \hat{p}_B \frac{\Delta g_1(1 - p) \beta A - \Phi_1 C_w}{(1 - e_g \Phi_1) C_0} \tag{15}
$$

and

$$
\hat{A}_{2g} = \mu - \Phi_1 \hat{p}_B \frac{\Delta g_1(1 - p) \beta A - \Phi_2 C_w}{(1 - e_b \Phi_2) C_0}. \tag{16}
$$

Then we obtain the following proposition about the good potential applicants' best responses.

**Proposition 6.** Based on the updated expression for $\hat{A}_{1g}$ and $\hat{A}_{2g}$ specified in Eqs. (15) and (16), Proposition 2 still holds for the model of imperfect screening.

**Remark.** Note that although under perfect screening and imperfect screening, the good potential applicants' best responses have the same structure, the screened aggregate demand rates $A_{1g}$ and $A_{2g}$ are different. Due to the screening error, bad potential applicants are more likely to apply while good potential applicants are less likely to apply. So good potential applicants' submission probabilities are smaller than those under perfect screening.

#### 6.1.3. The non-discriminatory screening policy

In the case of imperfect screening, the structure of the optimal non-discriminatory screening policy is more complex due to the impact of screening errors. To be concise, we denote $A = \mu e_b > 0$, $B \triangleq \frac{\lambda g_1}{\Phi_1 C_0}$ and define $f(y) \triangleq \frac{y - A \lambda g_1}{\Phi_1 C_0} - \sqrt{y - A \lambda g_1 - \Delta g_1}.$

**Proposition 7.** The approver's optimal non-discriminatory screening strategy $\Phi^*$ when screening precision is low ($e_b > \frac{\lambda g_1}{\Phi_1 C_0}$) is specified in Table 3, in which case all bad potential applicants submit their applications ($p_B = 1$). The approver's optimal non-discriminatory screening strategy $\Phi^*$ when screening precision is high ($e_b \leq \frac{\lambda g_1}{\Phi_1 C_0}$) is specified in Table 4, where under certain conditions, bad potential applicants may choose not to submit their applications ($p_B = 0$).

**Remark.** When $e_b > \frac{\lambda g_1}{\Phi_1 C_0}$ (Table 3), the optimal non-discriminatory screening probability can only take one of these three values: $0, f(A_3 + A_4) \in (0, 1)$, or 1. Different from perfect screening, $\Phi^*$ can be zero if errors are high (small $\frac{1}{1 - e_b}$), or the potential reward from good applicants is large relative to the potential penalty from bad applicants (high $\frac{\lambda b_2}{\Phi_2 C_0}$). All bad potential applicants submit their applications, i.e., $p_B = 1$, because they believe they are very likely to pass screening by error.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>$\Phi^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f(A_3 + A_4) &gt; 1$ and $\frac{\lambda b_2}{\Phi_2 C_0} = \frac{1}{1 - e_b}$ or $f(A_3 + A_4) &lt; 1 &lt; f(A_4)$ and $\frac{\lambda b_2}{\Phi_2 C_0} &gt; \max \left{ \frac{1 - e_b}{1 - e_b(1 - e_b) \Phi_2 C_0}, \frac{1 - e_b}{\Phi_2 C_0} \right}$ or $f(A_3) &lt; 1$ and $\frac{\lambda b_2}{\Phi_2 C_0} &gt; \frac{1}{1 - e_b}$</td>
<td>0</td>
</tr>
<tr>
<td>$f(A_3 + A_4) &lt; 1$ or $f(A_4)$ and $\frac{\lambda b_2}{\Phi_2 C_0} &lt; \frac{1}{1 - e_b}$ and $f(A_3 + A_4) \geq \frac{1 - e_b}{\Phi_2 C_0}$, or $f(A_3) &lt; 1$ and $f(A_3 + A_4) \geq \frac{1 - e_b}{\Phi_2 C_0}$</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 3**

The optimal non-discriminatory (imperfect) screening policy ($e_b > \frac{\lambda g_1}{\Phi_1 C_0}$ and $p_B = 1$).
When \( e_b \leq \frac{r_{1b} - 0.5}{c_{1b}} \) (Table 4), the optimal non-discriminatory screening probability can only take one of these three values: 0, \( f_2(A_x + A_g) \in (0, 0.5) \), or \( s^*_g \). When \( f_2(A_x + A_g) \leq 0.5 \), the screening probabilities approaches 0, because screening becomes useless. We observe similar patterns in the case of the non-discriminatory policy. Intuitively, the approver’s payoff using either a discriminatory policy or a non-discriminatory policy decreases in \( e_b \). In addition, a discriminatory policy shows more advantage when \( e_b \) is low. As \( e_b \) increases, screening in general becomes ineffective and thus the benefit of using a discriminatory policy decreases.

**Fig. 6** shows the impact of changing \( e_b \) on the error probability that good applicants are identified as bad.\(^6\) Since the approver’s strategy, either discriminatory or non-discriminatory, does not change with \( e_b \), the submission probabilities of bad potential applicants do not change in \( e_b \). Then \( e_b \) only influences the submission intention of the good potential applicants; their submission probability decreases slowly as \( e_b \) increases. As a result, the benefit of using a discriminatory policy does not vary much as \( e_b \) increases.

---

\(^5\) To separate the impacts of \( e_b \) and \( e_g \), we assume \( e_g = 0 \) in this case.

\(^6\) To separate the impacts of \( e_b \) and \( e_g \), we assume \( e_g = 0 \) in this case.
6.2. Non-zero abandon rate for good applicants

Now we study the scenario that good applicants could abandon the queue due to long waiting time. For example, in the visa context, after months of waiting, good applicants could give up and go elsewhere. This will change the objective function of the approver as the approver would optimally invest the screening effort expecting some approved applicants might not be available to come. To model this scenario, we define the abandon rate, \( A(W) \), as the percentage of good applicants who abandon the queue before the screening process ends. \( A(W) \in (0,1] \) is a non-decreasing function of the expected waiting time \( W \). Thus, the approver’s objective function (1) becomes:

\[
J(U) = \frac{1}{C_0} A(W) C_R \left( \frac{1}{C_0} U_1 \lambda_1 + \frac{1}{C_0} U_2 \lambda_2 \right) + (1 - A(W)) |\Phi_1 - \Phi_2| \lambda_{2b} C_1.
\]  

In our numerical experiments, we use a specific functional form, \( A(W) = \frac{1}{C_0} \exp \left( -\frac{c}{W} \right) \), where \( c \geq 0 \) measures the good applicants’ intention to abandon the queue. When \( c = 0 \), we have the same results as presented in previous sections. As shown in Fig. 7, for both discriminatory and non-discriminatory policies, when \( c \) increases, good applicants are more sensitive to congestion and thus the approver gets a lower reward from the good applicants who stay until screening ends. To alleviate congestion, the approver increases screening probabilities to deter all of the bad potential applicants and some of the good applicants. Some of our numerical results are given in Table 5.\(^{12}\)

---

\( J(\Phi) = |1 - A(W)| (\lambda_{1g} + \lambda_{2g}) R - |(1 - \Phi_1) \lambda_{1b} + (1 - A(W)) \lambda_{2b}| C. \) 

(17)
potential applicants. As a result, the benefit of using a discriminatory policy decreases in $\gamma$ until it approaches zero when $\gamma$ approaches infinity.

6.3. Non-zero waiting costs for bad applicants

In this subsection, we assume bad applicants have non-zero waiting costs in the screening process. In particular, the bad potential applicant’s expected utility in Eq. (4) becomes:

$$u_{tb}(\phi, p) = p_{tb}[\phi_t(-c_b) + (1 - \phi_t)c_b - \phi_t b_w W(\phi, p)],$$

where $b_w$ is the bad applicant’s cost per unit of waiting time and $c_b$ is the bad applicant’s cost per unit of waiting time. In this subsection, we assume $b_w = 0$, we have the baseline results as presented in previous sections. From Fig. 8, we have the following observations. The impact of increasing $b_w$ is similar to that of decreasing the threshold screening probability $\psi_b$, because the more sensitive to congestion, the less likely bad potential applicants submit their applications. For example, in the case of the discriminatory policy, when $0 < b_w < 1$, although the screening probability $\psi_1$ is very low, bad potential applicants of attribute 1 are deterred due to their sensitivity to congestion. When $1 < b_w < 2.5$, as $b_w$ increases, although the screening probability $\psi_2$ decreases, bad potential applicants of attribute 2 are deterred because a higher $b_w$ implies a lower threshold deterring all bad applicants. When $b_w > 2.5$, bad potential applicants of both attributes are deterred due to high waiting costs.

In the case of the non-discriminatory policy, the optimal screening probability is always set at the threshold to deter bad applicants and thus we see clearly this threshold (the optimal screening probability is always set at the threshold to deter bad applicants) is large when $\psi_b$ is small, and it is very small when $\psi_b$ is large. This again indicates that a discriminatory policy shows more advantage when the approver has more difficulties in balancing congestion and security risk.

7. Conclusion and future research directions

In this paper we study the approver’s optimal screening policies facing strategic potential applicants with private information. We provide analytical solutions for the optimal non-discriminatory screening policy and numerical illustrations for both the discriminatory and non-discriminatory policies. We find that although the optimal discriminatory screening policy is always better than the optimal non-discriminatory policy, the benefit of using a discriminatory screening policy is not significant when the tension between congestion and security risk is not intensive. These situations include: congestion is low (total potential arrival rate is low, or the service rate is high), or the approver’s reward from admitting a good applicant (or penalty from admitting a bad applicant) is very high or very low, or the bad potential applicant has a low incentive to apply (low reward if passed, or high penalty if caught, or high cost per unit of waiting time), or the screening error is high. In those cases, a non-discriminatory policy might be more appropriate in practice.

For future research, more elaborate decision models could be developed. First, this paper assumes that the parameters of the models are common knowledge, as most game-theoretic models do (Mas-Colei et al., 1995). However, in practice the approver and potential applicants may not know the system parameters, and thus a model with information would be an interesting direction. Second, this paper assumes that different types of screened applicants join a single queue, while alternatively, in the screening process different resources (multiple servers) could be used for different applicants with different observable characteristics; and as a result the waiting time performance could be different for applicants with different characteristics. Third, the service rate could be increased at a cost and thus optimization could be introduced to minimize the total costs. Finally, in our model we do not consider the fact that in some cases bad applicants of one attribute could pretend to be of the other attribute, in order to receive lower screening probabilities. For example, the bad applicants of attribute 1 could in principle pretend to be of attribute 2 to receive zero screening probabilities. For example, the bad applicants of attribute 1 could in principle pretend to be of attribute 2 to receive zero screening probabilities. For example, the bad applicants of attribute 1 could in principle pretend to be of attribute 2 to receive zero screening probabilities. For example, the bad applicants of attribute 1 could in principle pretend to be of attribute 2 to receive zero screening probabilities. For example, the bad applicants of attribute 1 could in principle pretend to be of attribute 2 to receive zero screening probabilities.
sequential games). In this situation, we expect that the discriminatory screening policy would become less effective, which again justifies the non-discriminatory policy.

Acknowledgements

This research was supported by the United States Department of Homeland Security through the National Center for Risk and Economic Analysis of Terrorism Events (CREATE) under Grant No. 2007-ST-061-000001. Any opinions, findings, and conclusions or recommendations in this document are those of the authors and do not necessarily reflect views of the United States Department of Homeland Security. Xiaofang Wang’s work is supported by the National Natural Science Foundation of China (No. 71001102) and the Fundamental Research Funds for the Central Universities, and the Research Funds of Renmin University of China (10XNK138). The authors thank Prof. Vicki M. Bier (University of Wisconsin-Madison) and Prof. Xisheng Shen (Beijing Jiaotong University) for helpful comments. We also thank Ms. Elizabeth Newell and Mr. Xiaojun Shan (both at University at Buffalo) for editorial help.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2011.01.019.

References


