Carnival Booth: An Algorithm for Defeating the Computer-Assisted Passenger Screening System

Samidh Chakrabarti
Aaron Strauss

6.806: Law and Ethics on the Electronic Frontier
May 16, 2002

Abstract

To improve the efficiency of airport security screening, the FAA deployed the Computer Assisted Passenger Screening system (CAPS) in 1999. CAPS attempts to identify potential terrorists through the use of profiles so that security personnel can focus the bulk of their attention on high-risk individuals. In this paper, we show that since CAPS uses profiles to select passengers for increased scrutiny, it is actually less secure than systems that employ random searches. In particular, we present an algorithm called *Carnival Booth* that demonstrates how a terrorist cell can defeat the CAPS system. Using a combination of statistical analysis and computer simulation, we evaluate the efficacy of *Carnival Booth* and illustrate that CAPS is an ineffective security measure. Based on these findings, we argue that CAPS should not be legally permissible since it does not satisfy court-interpreted exemptions to the Fourth Amendment. Finally, based both on our analysis of CAPS and historical case studies, we provide policy recommendations on how to improve air security.

Table of Contents

1 Introduction

On the morning of September 11, 2001, nineteen terrorists boarded four separate commercial airplanes across the northeastern seaboard of the United States. Moments after takeoff, they made their moves. Using the cardboard box cutters and razor blades they had smuggled through airport security checkpoints, the terrorists, working in teams of four or five, started slashing the throats of flight attendants and passengers in order to lure the pilots out of the cockpits. Before long, the terrorists had commandeered the controls of the aircraft, transforming their vehicles from passenger jets to guided missiles laden with sixty thousand pounds of explosive fuel each. By midmorning, the operation was complete. A field in rural Pennsylvania bellowed smoke. An entire facade of the Pentagon in Washington, DC lay in shambles. And the Twin Towers of the World Trade Center tumbled from the Manhattan skyline. In all, thousands of people perished on that morning. America, President Bush vowed, would never be the same.[1]

Almost immediately, the finger pointing began between federal agencies over who was responsible for not thwarting this attack—the most destructive ever carried out on the United States mainland. Some blamed the Central Intelligence Agency (CIA), calling it the largest intelligence failure since Pearl Harbor. Others assailed the Immigration and Naturalization Service (INS) for letting such dangerous individuals enter the country in the first place. But the one agency the public most criticized was the Federal Aviation Administration (FAA). How is it that they let nineteen people armed with knives slip right through security checkpoints? The grainy security camera pictures of Mohammed Atta, the terrorist ringleader, clearing security at the airport in Providence, RI, on September 11 now haunt America's consciousness. If only the metal detectors had been more sensitive, if only the security personnel had been more alert, if only Atta had aroused more suspicion, then perhaps this tragedy would never have happened.

[1] September 11, 2001 may have been the most widely covered American news story ever. For a more thorough treatment of the events of that day, see: http://www.time.com/time/911/

With the vulnerabilities of the nation's air transportation infrastructure made painfully clear, the FAA had a new sense of urgency in plugging the holes in a security system that was widely recognized as being as porous as a sieve. But the challenges in doing so were, and still are, daunting. Over 639 million passengers pass through airports annually in the USA.[2] Given this tremendous volume in traffic, unparalleled by any other nation, time-consuming security screening of every passenger is unfeasible. Since the number of security personnel available in airports is resource constrained, it would simply take too long to search everyone as meticulously as airports are able to in some other countries, such as Israel. While Americans love to fly, they also hate to wait.

To address these vexing security problems, the FAA has been trying in recent years to employ information technology to boost the overall efficiency of security screening. The kernel idea behind their approach is to be more intelligent about which passengers are selected for rigorous inspections. Intuitively, the FAA argues, if you only have the ability to scrutinize a small percentage of passengers, it seems best to spend the bulk of time carefully searching those who are likely to be terrorists and not waste much time searching those who have a small chance of posing harm. Why frisk Eleanor, the 80-year-old grandmother from Texas when you can stop Omar, the 22-year-old student fresh from Libya? If you can develop a profile describing who is likely to be a terrorist, then those are the people upon whom you should concentrate your security efforts.

Drawing from these intuitive underpinnings, the crown jewel of the FAA's information technology efforts is a system called the Computer Assisted Passenger Screening system (CAPS). The FAA contends that since CAPS uses profiles to pinpoint potential terrorists for closer inspection, it will not only result in the apprehension of more criminals, but will also make security screening more expedient for well-meaning citizens. Though in place since 1999, CAPS

---

[2] "Airport Activity Statistics of Certificated Air Carriers: Summary Tables," *U.S. Department of Transportation*. Washington, DC: 2001.

has gained much more attention as a promising counter-terrorism tool in the wake of September 11. The FAA already augmented the system in January, and plans for further expansion are underway.[3]

In our paper, we show that although these intuitive foundations might be compelling, their implementation in CAPS is flawed. That is to say that any CAPS-like airport security system that uses profiles to select passengers for increased scrutiny is bound to be less secure than systems that randomly select passengers for thorough inspection. Using mathematical models and computer simulation, we show how a terrorist cell can increase their chances of mounting a successful attack under the CAPS system as opposed to a security system that uses only random searches. Instinct may suggest that CAPS strengthens security, but it in fact introduces a gaping security hole easily exploitable by terrorist cells. It should be noted that CAPS has also received immense criticism from privacy advocates and civil libertarians[4], but in this paper we restrict our discussion to a purely technical perspective and the legal and policy implications of such an analysis.

In Section 2, we better define how the CAPS system operates. In Section 3, we present our algorithm for defeating CAPS. To evaluate the algorithm's efficacy, in Section 4 we present the results of a probabilistic analysis and computer simulation of airport security. In Section 5, we discuss a few case studies to understand what security techniques have been effective historically. In Section 6, we discuss the legal implications of our finding that CAPS performs worse than random search. And finally, we conclude in Section 7 with policy recommendations for how to improve air security.

---

[3] O'Harrow, Robert, Jr. "Intricate Screening Of Fliers In Works Database Raises Privacy Concerns." *Washington Post*. February 1, 2002. Page A01.

[4] Nojeim, Gregory. "Civil Liberties Implications Of Airport Security Measures." Statement to the White House Commission on Aviation Safety and Security. September 5, 1996. Available at: http://www.epic.org/privacy/faa/aclu_testimony.html

2 Defining CAPS

2.1 Government Guidelines

During the second term of his administration, President Clinton convened a panel to develop a set of recommendations to improve air transportation security. The resulting White House Commission on Aviation Safety and Security (chaired by Vice-President Gore) published its final report[5] in February of 1997, giving the federal seal of approval to automated passenger profiling. "Based on information that is already in computer databases," the Gore Commission wrote, "passengers could be separated into a very large majority who present little or no risk, and a small minority who merit additional attention." This statement also serves to articulate the federally supported two-stage architecture behind passenger profiling systems. First, the system should develop a secret profile describing characteristics of high-risk individuals (profile development). And then security manpower should be focused on those individuals who, based on available data, match the profile (profile evaluation). Such passengers are referred to as selectees.

The Gore Commission was cognizant of the constitutional fragility of such a system, so they invited testimony from civil liberty groups who were concerned that the derived profile might violate Fourteenth Amendment equal protection. To pacify these fears, the Commission dictated that "No profile should contain or be based on material of a constitutionally suspect nature." The Commission also insisted that the FAA should periodically consult the Department of Justice "to ensure that selection is not impermissibly based on national origin, racial, ethnic, religious or gender characteristics." Finally, to ensure that no one group is singled out, the Commission recommended that passenger profiling systems also choose random people who do not fit the profile as selectees.[6]

---

[5] "Final Report to President Clinton." White House Commission on Aviation Safety and Security. February 17, 1997. Available at: http://www.airportnet.org/depts/regulatory/gorecom.htm

[6] Ibid.

2.2 CAPS Architecture

Using the guidelines set forth by the Gore Commission, the FAA and Northwest Airlines jointly completed development of the first version of CAPS in 1998. The federal government closely guards the details of how the system operates, citing a compelling national security interest— if the specifications were to be released, it would be trivial for potential criminals to defeat the system. But drawing from an assortment of news articles[7], interviews with airport personnel, the Gore Commission report[8], leaks captured on the congressional record,[9] and prototypes written by software companies bidding to develop future versions[10], a cohesive if not comprehensive understanding of CAPS can be painted.

CAPS operates according to the same two-stage model described in the Gore Commission report: profile development followed by profile evaluation. First, based on a historical record of data pertaining to known terrorist activities, the software attempts to detect subtle patterns in the data that correlate with prior terrorist plots and anti-correlate with the activities of non-criminals. For instance, the software might find that those people who bought one-way tickets in cash and traveled abroad frequently had an elevated chance of being terrorists. CAPS then assembles these patterns into a secret profile suitable for inspection by the Department of Justice. Since the DOJ certification is only held periodically, the derived master profile is presumably static for long periods of time.

On a more technical level, CAPS likely accomplishes pattern detection through the use of a three-layer neural network. The first layer of the network contains hundreds if not thousands of nodes, the third layer contains a single output node, and the second layer contains an intermediate number of nodes. Each field of data available for profile development is fed into a separate node

---

[7] O'Harrow, Op. cit.
[8] "Final Report to President Clinton." White House Commission on Aviation Safety and Security. February 17, 1997. Available at: http://www.airportnet.org/depts/regulatory/gorecom.htm
[9] "Hearing on Aviation Security with a Focus on Passenger Profiling." *United States Congress Subcommittee on Aviation*. February 27, 2002. Available at: http://www.house.gov/transportation/aviation/02-27-02/02-27-02memo.html
[10] For more information, see HNC Software's website: http://www.hnc.com

of the first layer. Using a standard training procedure, such as back propagation, the weights of each connection are set such that when data from a terrorist is fed into the network, the output layer returns a value close to one. But when data from a non-criminal is fed into the network, the output layer returns a value close to zero. If training of the network is successful, the matrix of connection weights serves as the profile.

The CAPS system installed in 1999 and currently in use is only capable of doing profile development over data pertaining to the history of ticket purchases. Future versions of CAPS, however, will be able to incorporate a richer set of data, including driving history, credit card purchases, telephone call logs, and criminal records, among other information. Though allowing CAPS to access some of this data would require changes in privacy legislation, Congress, following on the heels of the PATRIOT act, is poised to facilitate.[11]

Once CAPS crafts a profile, it is incorporated into software that is accessible from every airline check-in counter nationwide. When a passenger checks in, the ticket agent enters the passenger's name into the CAPS console. Data mining software linked to government databases then scours for information about the passenger, retrieving data relevant to the profile. The software compares the similarity of the acquired data to the profile and computes a "threat index" assessing how much potential risk that passenger may pose.[12] In the technical scenario outlined above, the computed threat index would simply be the product of the profile matrix with the passenger's mined data vector.

If the passenger has one of the top 3-8% of threat indices relative to the other people on his flight, then CAPS flags him for "special treatment." To protect the integrity of the CAPS system, the precise percentage of people flagged by CAPS is unpublished, but it is known to be in that range.[13] To comply with the Gore Commission guidelines, a small percentage of people on

---

[11] O'Harrow, Op. cit.
[12] Ibid.
[13] "Computer-Assisted Passenger Screening and Positive Passenger-Bag Matching." *Assessment of Technologies to Improve Airport Security: First Report*. The National Academy Press: 2000.

each flight are randomly flagged as well. In all, the total number of people flagged by CAPS is limited by the security personnel resources available at each particular airport.

2.3 Special Treatment

What exactly do these "special treatment" flags entail? Before September 11, CAPS flags were only tied to the passenger's checked baggage, which were scanned using costly explosives-detection equipment. This represented the conventional wisdom of the time that a terrorist would try to smuggle explosives only in his checked luggage[14]. But now, the FAA is tying CAPS flags to individuals. Someone flagged by CAPS may have her carryon bags specially inspected, she may be subject to questioning, she may be asked to stand in a separate line, she may be asked to comply with a search of her body, or a guard may even escort her directly to the gate.[15]

In an article in *Slate* magazine, Microsoft Chief Architect Charles Simonyi related his experience of being flagged by CAPS.[16] During a routine business trip, security personnel insisted on completely unpacking and repacking all of his carryon bags. This happened time after time. "Then it hit me," Simonyi writes. "It was not that security was especially tight: It was only me they wanted. The label my friendly hometown airline had affixed to my bags had unexpectedly made me a marked man, someone selected for some unknown special treatment." The bottom line is that if CAPS flag you, you'll be treated differently, and you'll know it.

3 Defeating CAPS

This transparency is the Achilles' Heel of CAPS; the fact that individuals know their CAPS status enables the system to be reverse engineered. You, like Simonyi, know if you're carryons have been manually inspected. You know if you've been questioned. You know if

---

[14] On September 11, the version of CAPS in place failed to flag the luggage of 10 of the 19 hijackers. This underscores how easy it is for a terrorist to evade the CAPS profile.
[15] O'Harrow, Op. cit.
[16] Simonyi, Charles. "I Fit the Profile." *Slate Magazine*. May 25, 1997. See: http://slate.msn.com/?id=2058

you're asked to stand in a special line. You know if you've been frisked. All of this open scrutiny makes it possible to learn an anti-profile to defeat CAPS, even if the profile itself is always kept secret. We call this the "Carnival Booth Effect" since, like a carnie, it entices terrorists to "Step Right Up! See if you're a winner!" In this case, the terrorist can step right up and see if he's been flagged.

3.1 Carnival Booth Algorithm

We will now present an algorithm that a terrorist cell can employ to increase their probability of mounting a successful attack under the CAPS system as opposed to an airport security system that employs only random searches. The key idea is that a terrorist cell can probe the security system to ascertain which of their members have low CAPS scores. Then they can send these members on destructive missions. Since security manpower is disproportionately spent on people with high CAPS scores, and the operative has a low score, he will most likely face reduced scrutiny.

The algorithm, which we call *Carnival Booth*, then is as follows: (1) Probe the system by sending an operative on a flight. The operative has no intent of causing harm. He has no explosives. He has no weapons. He has nothing. He simply takes the flight and notes whether or not CAPS flags him. (2) If he is flagged, then send another operative in the same manner. (3) Repeat this process until a member who consistently eludes CAPS flags is found. (4) Now send this operative on a mission with intent to harm, complete with weapons or explosives. Since CAPS didn't flag him last time, he likely won't be flagged this time, so he incurs much less risk of special scrutiny.

To better understand how a terrorist cell using this algorithm stands a better chance of success under CAPS than under a random system, let's consider the numerical example illustrated in Figure 1. Suppose an airport only has the personnel resources to give 8% of people special scrutiny; the other 92% undergo standard screening through a metal detector. Under a system

where people are selected at random, this airport can afford to flag 8% randomly. This means that every time a terrorist attempts to go through security, he stands an 8% chance of increased scrutiny. This will be true no matter what tactic or algorithm the terrorist uses.

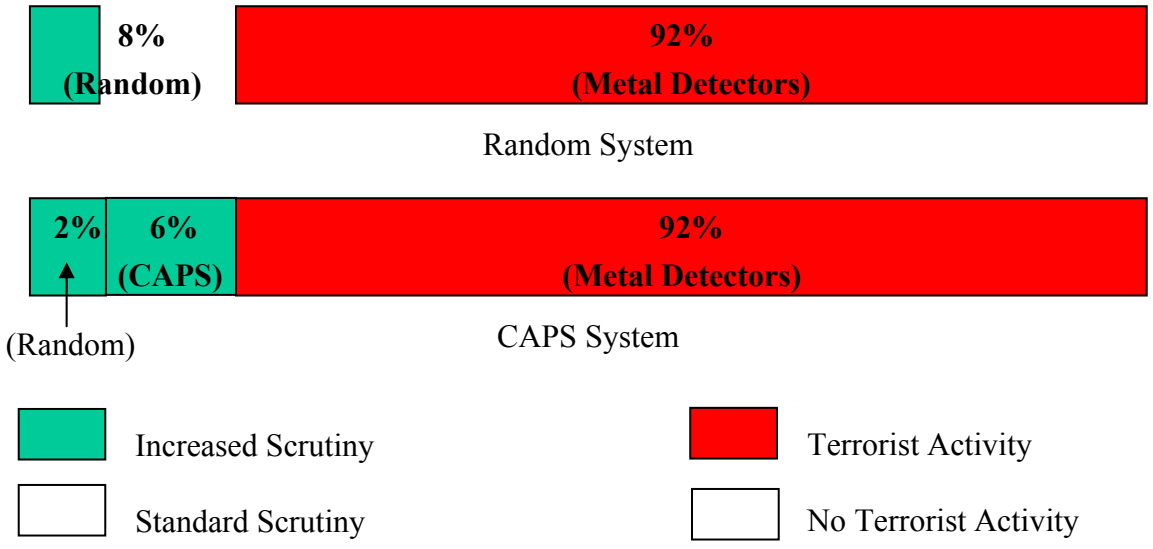

Figure 1: Regions of Terrorist Activity under CAPS and a Random System

Now compare this to the same airport using a CAPS system, which may for example flag the 6% of passengers with the highest threat indices and 2% randomly in order to equal their personnel-constrained 8% limit. By employing the algorithm described above, the terrorist cell knows that since their operative has previously probed the system without a flag, CAPS likely will not flag him again. In essence, the terrorist cell is able to relegate its harmful activities outside of the 6% CAPS flag zone. Now, their operative only has a 2% chance of calling up a thorough inspection. Compare this to the 8% chance the terrorist would incur under the random system. It's clear that terrorist cells would therefore prefer airports fortified by CAPS.

3.2 Cells vs. Individuals

Upon reading this analysis, it's natural to feel uneasy. How is it that such an intuitive system can actually weaken security? For clarity, it is useful to draw a distinction between individual terrorists and the coordinated activities of an entire terrorist cell. It is entirely probable

that even a rudimentary CAPS profile can flag many individual terrorists, shrinking the viable pool of recruits that a terrorist cell can send on a mission. But so long as the cell as a unit can identify those members who slip through CAPS, even if they are few in number, it has an enhanced probability of mounting a successful attack. It only takes one person to do harm. *Carnival Booth* makes the process of identifying such a person simple.

Evolutionary biology validates this perspective. The most famous example is the peppered moths of London. In the 1950s, Dr. H.B.D. Kettlewell, a physician, noticed that the peppered moth population living near industrialized areas of London started to change in color from light gray to dark gray. Through a series of experiments, Kettlewell showed that dark colored moths were more apt to survive near industrialized areas because, by matching the color of the smokestack soot that coated the ground, they could evade predators.[17] Individual light gray moths were eliminated from the gene pool, but since the population had a few dark gray moths (maybe only one), the species survived and proliferated. Even in nature, a non-conscious species can subvert static environmental conditions that expose them to danger. Conscious terrorists, one would expect, can likewise circumvent a profile.

3.3 Algorithm Assumptions

Terrorists, of course, are not moths. In fact, the evolutionary biology example points to several assumptions we must make in order to claim the effectiveness of our CAPS circumvention algorithm. Unlike a natural species, terrorists do not have, for example, infinite variation and boundless evolutionary time. This leads to three key questions: Do terrorist cells have a diverse enough membership to successfully use this algorithm? Do they have the money, patience, and planning skill to execute it? And above all, are they smart enough to know to use it? Using findings from our nation's "War on Terror" as a case study, we claim that it is safe to assume that the answers to all of these questions are affirmative.

---

[17] Ricklefs, Robert. *Ecology*. W H Freeman & Co. 1990.

The terrorists revealed in recent months are strikingly diverse. Most famously, John Walker Lindh— the "American Taliban"[18]— is a nineteen-year-old Caucasian boy from Marin County, the yuppie capital of California. He was fighting on the front lines for the Taliban in Afghanistan against American soldiers. If he had never been discovered, would it have been difficult for Al Qaeda to send him back to the United States on a terrorist mission? Then there is "shoe bomber" Richard Reid, accused of trying to detonate explosives in his sneakers during a transatlantic flight.[19] He is a British citizen with an English mother and Jamaican father. And just last week the FBI apprehended Lucas Helder, a 21-year-old art major at the University of Wisconsin-Stout.[20] Helder allegedly planted 18 pipe bombs in mailboxes in five different states to form a "smiley face" pattern. And who can forget Ted Kaczynski and Timothy McVeigh? Terrorists clearly have no shortage of diversity.

As the hijackers of September 11 showed, they also have no shortage of money, patience, and planning acumen. Mohammed Atta planned the attack years in advance, studying diagrams of the World Trade Center, going to flight school, and analyzing airplane specifications on the Internet. The money trail funding Atta's operation weaves through anonymous bank accounts in multiple countries. It is so intricate that the FBI still does not completely understand it. Terrorist cells like Al Qaeda have demonstrated that they have the resources required in terms of money, patience, and planning to use the algorithm.

What may be more alarming is that evidence from the September 11 investigation shows that Atta already knew the kernel idea behind this algorithm. *Newsweek* reported[21] that in the weeks before September 11, Atta and his conspirators practiced their attack by boarding the exact same target flights they intended to later hijack (same planes, same times, same origins and

---

[18] Tyrangiel, Josh. "The Taliban Next Door." *Time*. December 9, 2001.
[19] Canedy, Dana, et al. "Passenger With Shoe Bombs First Raised Only Eyebrows." *New York Times*. December 27, 2001.
[20] Eggen, Dan. "Pipe Bomb Suspect Arrested." *Washington Post*. May 8, 2002. Page A01.
[21] Reported in several issues during October 2001. See: http://www.msnbc.com/news/NW-attackonamerica_Front.asp?cp1=1

destinations). They wanted to ensure that they didn't raise any suspicions or red flags. This is a clear demonstration of Atta's cleverness. Like Atta, terrorists are smart. They already know this algorithm. And they are already using it.

4 Evaluating Carnival Booth

A combination of a probabilistic analysis and results from a computer simulation demonstrate in more concrete terms if it is possible for a terrorist cell to use Carnival Booth to defeat the CAPS system. In particular, we wanted to find out under what conditions CAPS outperforms or underperforms random search, and we wanted to quantitatively determine how much a terrorist cell could benefit from using the algorithm. To accomplish this, we compared three systems: (1) the CAPS system, (2) a system where passengers are randomly selected, and (3) a random system with more advanced administrative searching. The results are clear. The less a system relies on profiling and the more advanced its administrative searching, the more terrorists it will catch.

The analysis makes two reasonable assumptions. First, a terrorist must bring a weapon or bomb onboard the airplane to cause damage. Modifications to cockpit doors, sky marshals, and heightened passenger awareness after September 11th have forced potential terrorists to use more than cardboard cutters to gain control of a plane. Second, a future terrorist that is flying with no weapons and no criminal intent will not be apprehended by law enforcement. Even if this person has the highest threat index possible, there would be no reason to apprehend him or her (barring an outstanding warrant or alleged connection to a previous terrorist act).

4.1 Probabilistic Analysis

We model all three flavors of airport security with a two-stage architecture. The first stage is administrative screening, which includes the usage of metal detectors and basic questioning about luggage contents. All passengers are subjected to this level of screening. The

second stage is increased screening for those who are either flagged by CAPS or randomly selected.

Using probability theory, we developed a generalized function for these two-stage systems that determines the probability that a terrorist will be caught. First, the total probability of the terrorist being arrested is the sum of whether the terrorist is arrested during a 2nd level search prompted randomly, a 2nd level search prompted by a CAPS flag, or a 1st level administrative search:

Pr(Terrorist Arrested) =

Pr(Terrorist Arrested During Randomly-flagged 2nd Level Search)
+ Pr(Terrorist Arrested During CAPS-flagged 2nd Level Search)
+ Pr(Terrorist Arrested During Administrative 1st Level Search)

Introducing some notation, let A be whether the terrorist is arrested, R be whether the terrorist is randomly chosen for 2nd level screening, and C be whether the terrorist is selected by CAPS for 2nd level screening. Since the probability that the terrorist will undergo administrative search is equal to the probability that the he is not flagged for 2nd level search either by a CAPS flag or a random flag, the full equation becomes:

Pr(A) =

Pr(A | R) * Pr(R)
+ Pr(A | C) * Pr(C)
+ Pr(A | ¬(R U C)) * Pr(¬(R U C))

To analyze the dynamics of this equation, there are several constraints we can impose on these probabilities. First observe that the probability that 2nd level interrogation results in an

arrest should be the same whether the terrorist is flagged by the CAPS profile or randomly. Hence, Pr(A | R) must be the same as Pr(A | C). Also, since the 2nd level searches are more thorough than 1st level searches, logic dictates that Pr(A | ¬(R U C)) be lower than either Pr(A | R) or Pr(A | C). Finally, in all the simulations, the number of passengers that will be subjected to heightened security will be the same.

Although information about the CAPS system is limited, the total percentage of people flagged by CAPS, either through a profile or randomly, is reported to be between 2-8%. Accordingly, we choose Pr(R) to be 2% under CAPS, and 8% under a random system where all personnel resources can be devoted to random inspections. Since second level searches are presumably more effective than first level searches, our system arbitrarily sets Pr(A | R) at 75% and Pr(A | ¬ (R U C)) at 25% for the first two systems. For the third system employing better administrative searches, Pr(A | ¬(R U C)) is set to 40%. We can begin to compare the systems now that all values except Pr(C) are now known:

(CAPS system) Pr(A) = 0.75 * 0.02 +
0.75 * Pr(C) +
0.25 * (1 - 0.02 - Pr(C))

(Random system) Pr(A) = 0.75 * 0.08 +
0.75 * 0 +
0.25 * 0.92 = 29%

(Random w/ admin+) Pr(A) = 0.75 * 0.08 +
0.75 * 0 +
0.40 * 0.92 = 42.8%

The critical value Pr(C) depends on the effectiveness of the CAPS profile and whether terrorist cells are strategically using agents with low threat indices. If Pr(C) is greater than 6%, then the CAPS system will be more successful than a random system. If the probability a terrorist will be flagged by CAPS is less than 6%, then law enforcement should randomly select passengers for increased scrutiny. (The third equation demonstrates how effective even modest improvements in administrative searches can be. We will return to this topic later in our discussion.)

Clearly, we claim in this paper that by using our algorithm, a terrorist cell can push Pr(C) under 6%. This probability heavily depends on how many probes an agent completes before that terrorist cell sends the agent on a mission. Imagine that a terrorist cell sends a new recruit to probe a domestic flight. On average there could be 74 other passengers on that flight with him.[22] Watching the other passengers closely at security checkpoints, the agent notices that he has not been flagged. Although the cell can correctly conclude that 6% of his peers on the flight (approximately 6 passengers) had threat indices higher than his, this is no guarantee that on the next flight, the agent will still be in the bottom 94% of threat indices. Thus, the terrorist cell remains unsure whether they can "safely" send the agent on a mission. However, the more probes that the agent returns from without attracting a CAPS flag, the more confidence the cell has in this agent's ability to get through the system undetected.

With an infinite number of probes, the CAPS system will perform worse than a random system. If an agent has passed through security unflagged an infinite number of times, then the probability the agent will be flagged on his mission, Pr(C), becomes 0. Thus the overall probability of such an operative being apprehended is:

$$\text{(CAPS System)} \quad \lim_{\text{Probes} \to \text{Inf}} \quad \Pr(A) = 0.75 * 0.02 + 0.75 * 0 + 0.25 * 0.98 = 26\%$$

[22] "Airport Activity Statistics of Certificated Air Carriers: Summary Tables," *U.S. Department of Transportation*. Washington, DC: 2001.

This is four percentage points inferior to the random system. Obviously, a terrorist cell cannot launch an infinite number of probes. The chief question then becomes how many probes must a terrorist cell launch before they can be sure that their operative stands a smaller chance of apprehension than under a random system?

4.2 Computer Simulation

We developed a computer simulation to shed light on how much more confident the cell can be after each additional probe. Unfortunately, to definitively answer this question one would need the distribution of CAPS scores for both terrorists and the public at large. These distributions are, of course, confidential. Hence we ran our simulation several times using a wide variety of distributions, hoping that the real distributions are bounded by our models.

In crafting these distributions, we made a few general assumptions. The first is that terrorists, on average, have a higher CAPS index than the average airline flier. It would be unreasonable for the FBI to advocate a profiling-based system if this assumption were false. The second assumption is that the terrorist and general public distributions are Gaussian. This assumption stems from the fact that CAPS scores are generated from thousands of boolean tests. A combination of many random factors will usually create the bell-shaped Gaussian distribution curve. While the actual distributions probably exhibit non-zero skewness and kurtosis, we do not claim to know in which directions these statistical moments would manifest themselves; we therefore assume they are zero.

As for the more significant moments of mean and standard deviation, we varied these values in an attempt to cover many of the potential cases. Some distributions represented moderately effective profiles with a mean CAPS score of 50 out of 100 for terrorists and 30/100 for non-terrorists and standard deviations of 20. Other distributions modeled highly tuned profiles

with a mean CAPS score of 70/100 for terrorists and 20/100 for non-terrorists and standard deviations of 20.

For each distribution, and for each number of probes, 7,000 attacks were simulated. The computer tracked a simulated terrorist cell as it used the Carnival Booth algorithm to recruit new members and probe the system with each neophyte. During a probe, the computer randomly generated 75 other non-terrorist passengers. After finding a member that passed through the requisite number of such probes unflagged, it sent that agent on a mission. On that mission, the terrorist could either be flagged by CAPS, randomly selected, or neither, depending on the CAPS score probability distribution being tested. The percentage of terrorists arrested was then recorded.

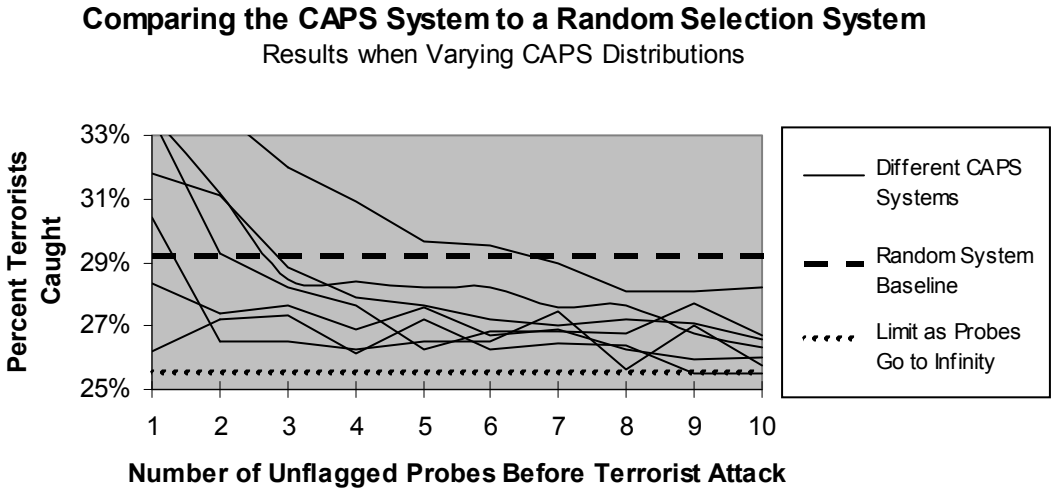

Figure 2: Number of Probes Required to Defeat CAPS for Varying Score Distributions

As shown in Figure 2, changes in the mean and standard deviation values make little difference in the end result. If a terrorist cell requires four unflagged probes prior to launching an attack, the CAPS model becomes worse than random. At six probes, the threat index becomes an entirely useless means of identifying terrorists. This immunity to the characteristics of the CAPS

distribution is a striking result— it means that even if CAPS utilized highly accurate profiles, a terrorist operative need only launch a few successful probes before he can beat the system.[23]

If CAPS is not a viable means of strengthening security, the question remains how to make our airports safer. With a random system, the paramount equation is:

(Random System) $\Pr(A) = \Pr(A \mid R) * \Pr(R) + \Pr(A \mid \neg R) * \Pr(\neg R)$

The most effective security measure would be one that increases the percentage of people receiving second level searches. But given that passengers would object to the prohibitive waiting times, the government has decided to inconvenience only 8 percent of airline passengers with increased screening. Thus, we consider Pr(R) fixed at 8% and Pr(¬R) fixed at 92%. The only two variables left to manipulate are the probability of catching a terrorist during administrative screening and the probability of catching a terrorist during heightened scrutiny. To maximize the equation, the term multiplying Pr(¬R) is the most advantageous to increase. This term corresponds to the administrative search accuracy. Since administrative searches cover everyone, investment in technologies used in those searches will return more bang for the buck.

5 Case Studies

Two detailed case studies buttress the proposal that investments should be diverted from profiling to administrative searches. First, we examine the failed attempt to bomb Los Angeles International Airport just before the millennium. Then, the security system of El Al is analyzed and used as a benchmark for great security. The net result of these case studies points both to flaws in the CAPS profiling system and the strengths of administrative searching.

[23] In the parlance of game theory, a terrorist cell using this algorithm would be said to have a *dominant strategy*. The Nash equilibrium would fall in favor of the terrorist cell.

5.1 Ressam's 1999 Terrorist Attempt

In August 1999, Ahmed Ressam began to plot a terrorist attack against the United States from his home in Montreal. Ressam decided to target Los Angeles International Airport since he had sensed lax security when he had flown through the airport in February 1999. By September, Ressam had found an accomplice who funded the project, which enabled him to begin to procure chemicals and timing devices. His plan was to plant an explosive in a suitcase, put the suitcase in an inconspicuous luggage cart, set the timing device, and leave the scene. With everything prepared, Ressam flew to Vancouver and then left for the United States in a rental car on December 14th. [24]

At approximately 6pm, Ressam boarded the ferryboat *Coho* in Victoria, British Columbia, which would take him to Port Angeles, Washington. Before being allowed to board the *Coho*, Ressam, like all other passengers, was required to pass through a U.S. customs pre-screening checkpoint. Ressam presented the customs official with a fake ID; he used the alias Benni Antoine Noris.[25] Even though Ressam was wanted by French and Canadian law enforcement, a background check on the name Noris cleared since it was not one of Ressam's known aliases.[26] Ressam even presented (in addition to a photo ID) a Costco card in the name of Noris.[27] Thus, Customs let Ressam board the ferry.

Upon arriving in Port Angeles, Ressam was the last person to drive his car off of the ferry.[28] A Custom's agent, Diana Dean, approached Ressam's car and began questioning him. Dean recalled, "I noticed during routine questioning that Ressam was acting in a nervous and strange manner while answering routine questions. I decided to perform a more thorough

---

[24] "Ahmad Ressam's Millennium Plot," *Public Broadcasting Service*. Available at: http://judiciary.senate.gov/oldsite/21020dd.htm

[25] *United States of America V. Ahmed Ressam*, Magistrate's Docket No. Case No. 99-547m, Complaint For Violation (1999)

[26] "Ahmad Ressam's Millennium Plot," *Public Broadcasting Service*. Available at: http://judiciary.senate.gov/oldsite/21020dd.htm

[27] *United States of America V. Ahmed Ressam*, Magistrate's Docket No. Case No. 99-547m, Complaint For Violation (1999)

[28] Ibid.

secondary examination."[29] The "secondary examination" required that Ressam exit his vehicle and that the car be fully searched. Ressam initially refused to cooperate with the search and only reluctantly exited his car. When he saw the officials had found his explosives, which were hidden in the trunk, Ressam ran.[30] Customs officials chased Ressam for several blocks before apprehending him.

Ressam's use of an alias demonstrates a weakness in profiling. Since the standard identification in the United States is a state-issued driver's license, there are fifty different sets of rules. Terrorists can attack the weakest link. For instance, in preparation for the September 11th attacks, seven of the nineteen hijackers abused Virginia's relaxed standards to obtain false licenses. Even if America adopts a national ID, potential terrorists can acquire falsified papers in other countries, as Ressam did.[31]

Saving the day in this case were the security standards of the US Customs Department and the professionalism of Diana Dean. No profile could have prevented Ressam from entering the country—all his papers appeared to be in order. The experience of Dean to perceive Ressam's uneasiness and the searching procedures that found the explosives concealed in the spare tire well of the truck prevented disaster from striking one of America's largest airports.

5.2 The El Al Standard

El Al is Israel's airline, and while it receives daily threats, they have not experienced a terrorist incident in over thirty years. Its success is mainly due to its tight on-site security. Each passenger each time he or she flies is psychologically evaluated. Carryon bags are checked multiple times. In essence, El Al's security system works from the assumption that every passenger is a threat, and treats him or her accordingly.

---

[29] Dean, Diana. "Statement Before Senate Judiciary Committee." February 10, 2000. Available at: http://judiciary.senate.gov/oldsite/21020dd.htm

[30] *United States of America V. Mokhtar Haouari,* S4 00 Cr. 15. Testimony Available at: http://www.pbs.org/wgbh/pages/frontline/shows/trail/inside/testimony.html

[31] Ressam had a Canadian passport and Quebec driver's license, both under the name Noris.

Every car that enters Ben Gurion International Airport in Tel Aviv is examined. Plain-clothed guards help secure each airport building.[32] As El Al's president, David Hermesh, stated, "If you're a passenger on El Al, most likely you will be observed from the minute that you left your car or you have been dropped off."[33]

Upon entering the airport, a passenger is asked a series of specific questions including "Who paid for your ticket?" "What is the purpose of your travels?" and "When did you book the flight?"[34] The questions are specifically designed to evoke an observable reaction. As passengers answer the questions, the officer carefully scrutinizes tone of voice, body language, and quickness of response. If the answers are unsatisfactory, a different officer will ask the ticket holder a different set of questions. Security personnel also might separate flight companions, making sure their stories match.[35] A psychological evaluation of the passenger is combined with information about the passenger previously obtained by El Al and Interpol.[36]

On the baggage side of the equation, El Al is just as strict. All luggage goes through a decompression chamber designed to trick bombs that are set to go off when the barometric pressure indicates that the plane is in the sky.[37] Sophisticated X-ray technology is used to detect liquid explosives.[38] Security personnel handle and examine every piece of carryon luggage. Every piece of luggage is also matched to its owner. And on layover stops, passengers must reclaim their luggage.[39]

---

[32] Perman, Stacy. "The El Al Approach: A look at the Israeli airline's security procedures," *Business 2.0*. November 2001. Available at: http://www.business2.com/articles/mag/0,1640,17508,FF.html
[33] "Model for air travel security may be El Al," *CNN*. September 26, 2001. Available at: http://www.cnn.com/2001/WORLD/meast/09/26/rec.el.al.security/
[34] Ibid.
[35] Perman, Op. cit.
[36] "Israeli-style security might have averted hijackings." *USA TODAY*. September 13, 2001. Avaiable at: http://www.usatoday.com/news/world/2001/09/12/israelisecurity.htm
[37] Ibid.
[38] Perman, Op. cit.
[39] Ibid.

While El Al does keep a database of individuals' nationalities, genders, criminal records and flight histories, this tracking should not be confused with the a CAPS-like profile.[40] Unlike CAPS, El Al's system only makes a determination of the risk of a passenger after a security agent has questioned him or her.[41] CAPS, on the other hand, is a prior system focused on predicting who *will become* a terrorist. El Al, through its system of psychological analysis and advanced baggage screening, has found success in determining who *is* a terrorist.

6 Legal Implications

Our analytic result that CAPS-based airport searches are less effective than random searches has important legal implications. While metal detectors and security checkpoints at airports may seem routine to us today, it was not always so. It was only after an extensive history of litigation that such searches were deemed permissible[42].

Initially, in light of the Fourth Amendment of the Constitution, it may not even be clear why any airport searches are authorized. The Fourth Amendment protects "the right of people to be secure in their persons, papers, and effects against unreasonable searches and seizures." The keyword is "unreasonable." Courts have generally recognized searches as being reasonable if they properly balance the degree of intrusiveness, the magnitude and frequency of the threat, and the efficacy of alternatives to the search.[43] Through all of the litigation, courts have established two major exceptions to the Fourth Amendment in relation to airport security: the administrative search exception and the stop-and-frisk exception. Any new security technique implemented at airports must conform to one of these two principles.

---

[40] Ibid.
[41] "Model for air travel security may be El Al," *CNN*. September 26, 2001. Available at: http://www.cnn.com/2001/WORLD/meast/09/26/rec.el.al.security/
[42] Most often the court cases were brought on by drug smugglers seeking to challenge the validity of airport searches that discovered their contraband.
[43]"Legal Issues." *Assessment of Technologies to Improve Airport Security: First Report*. The National Academy Press: 2000.

6.1 Administrative Search Exception

For a search to qualify for the administrative search exception, it must serve a societal purpose, be minimally intrusive, and all persons must be searched equally no matter what level of suspicion they may arouse.[44] An assortment of court cases establishes these criteria, but *United States v. Martinez-Fuerte*[45] (1976) provides the most direct example. In *Martinez-Fuerte*, the Supreme Court upheld the practice of border control agents stopping vehicles at checkpoints. First, the Court recognized that governmental interests may in some situations trump Fourth Amendment protections. The Court wrote, "Government or public interest in making such stops outweighs the constitutionally protected interest of the private citizen." Furthermore, the Court's opinion relied on the fact that the search was brief in duration, usually consisting of only a couple of questions and visual inspection. Since it did not greatly impede the flow of traffic or consume too much of the traveler's time, the Court ruled that the search was minimal in its intrusiveness given the government's compelling interest in border control. Finally, the Court held "that the stops and questioning at issue may be made in the absence of any individualized suspicion." Most importantly, what made the border control stops permissible was that everyone, regardless of suspicion, was democratically screened.[46]

Metal detectors at airports are therefore justified under the administrative search exception. By deterring people from bringing weapons onboard aircraft, they serve a societal purpose. They are also minimally intrusive. And security personnel require every passenger to go through them. CAPS-based screening, on the other hand, does not fit all of the administrative search exception criteria. In particular, the use of a profile to select particular individuals for scrutiny violates the requirement that the search be equally applied to all individuals regardless of suspicion.

---

[44] Ibid.
[45] *United States v. Martinez-Fuerte*, 428 U.S. 543 (1976)
[46] Ibid

6.2 Stop-and-Frisk Exception

The second type of exception under which an airport security technique may qualify is the stop-and-frisk exception. Under this type of exception, if security personnel have a minimal level of justification that a person may pose danger, they may conduct, without a warrant, a limited search of that person for the presence of weapons. Like the administrative search exception, the stop-and-frisk exception also has a complex history of precedent. *Terry v. Ohio*[47] (1968) is the case that perhaps best defines this standard. In this case, a Cleveland police officer patrolling his downtown beat observed two individuals stare into a storefront window a total of 24 times. Suspecting that they might be preparing to rob the store, the officer confronted the two men and frisked them for weapons and found each of them carrying a revolver. The Supreme Court was asked to decide whether the officer should have been permitted to frisk the two men since he did not have a warrant. The Court decided that although the officer had not established probable cause, his objective observations of the men's suspicious behavior provided enough justification.[48]

Subsequent court cases have further clarified what exactly constitutes a minimal level of objective justification. In *United States v. Cortez*[49] (1981), the Supreme Court decided that the "totality of circumstances" may be considered in determining if suspicion exists. Security personnel need more than a hunch to conduct a stop-and-frisk search, but they need not establish probable cause. More precisely, given all of the circumstances, the suspicion must be based upon an enhanced "likelihood" that the person poses harm.[50]

On the surface, CAPS-based security screening appears to qualify for the stop-and-frisk exception. By targeting people fitting a suspicious profile, the FAA would argue that CAPS establishes sufficient objective justification. However, based on our analysis of CAPS, we reach

---

[47] *Terry v. Ohio*, 392 U.S. 1 (1968)
[48] Ibid
[49] *United States v. Cortez*, 449 U.S. 411 (1981)
[50] Ibid.

the opposite conclusion. We showed that under the CAPS system a terrorist can reduce his chance of arrest to a level lower than a system that uses random search. This means that when security personnel conduct their second level search of passengers flagged by CAPS, they actually have a *reduced* probability of apprehending a terrorist carrying weapons or explosives. Since random searches can catch more terrorists, airport security cannot therefore legitimately establish that those passengers flagged by CAPS have an enhanced likelihood of harm. Consequently, CAPS fails to meet the standards of the stop-and-frisk exception.

Would it be possible to convince a judge that CAPS as currently implemented is not permissible for use in airports since it does not qualify for either the administrative search or the stop-and-frisk Fourth Amendment exceptions? Our analytic results fight an uphill battle against intuition. Although it is not certain that a judge would strike down CAPS as being unconstitutional, we hope that our analysis will encourage the judge to ask the FAA to be more convincing about why CAPS is constitutional.

7 Policy Recommendations

As the evidence has shown, the most efficient approach to reducing terrorism is to catch a terrorist in the midst of a mission. Attempting to determine who might become a terrorist in the future or who acts like a terrorist is error-prone and easy to defeat. As mathematically demonstrated in Section 4.1, the focus of the FAA's efforts to tighten airport security should be augmenting administrative searches that affect everyone that boards a flight.

The questions that airline employees or security personnel ask passengers must evoke more of a response. Currently, the standard questions are "Has your luggage been in your possession at all times?" and "Has anyone given you anything or asked you to carry on or check any items for them?" For the experienced traveler, these questions induce the reflex answers of "yes" and "no." By asking questions that require a slightly longer response, security personnel

would be able to gather information from body language and tone of voice. This strategy also requires the monetary investment to train the questioners so they know what signs to look for.

Techniques for screening checked and carryon luggage have the advantage that they are hard to probe. It is difficult to know whether an explosive will evade security unless one brings an explosive through airport security. If a terrorist is unsuccessful during a probe of such a system, they will be arrested for the weapons they are carrying. Thus, a terrorist cell cannot probe administrative searches of bags and persons. Increasing the effectiveness of object searching, perhaps through new X-ray technology, should be a top priority.

If the US government is determined to keep the CAPS system, then it is imperative that in the future terrorists are prevented from defeating the system. Any new versions of CAPS must reduce the Carnival Booth effect. In essence, it must be non-obvious if a passenger has been flagged.

For checked luggage this veil is easy to create since the passenger is unaware of what occurs below the concourse. Bags that belong to individuals not considered flight risks could be scanned with normal "medium x-ray" technology. On the other hand, the luggage of flagged passengers could be scanned with a slower, more detailed, and more costly "computer tomography" scanner.[51]

Disguising the two-level process when screening carryon bags and the passengers themselves is much more difficult. One possible solution could be to merge the new X-ray technology for screening individuals (such as America Science and Engineering, Inc.'s "Body Search"[52]) and the standard metal detectors into one device. This device could be a dual metal detector and X-ray scanner. Similar to the checked baggage system, for each ticket holder a security officer would discreetly turn on or off the X-ray scanner. Thus, only "flight risks" would be intrusively screened with the more costly X-ray scanner.

---

[51] Tyson, Jeff. "How Airport Security Works: Check Your Bags: CT Scanners." *Marshall Brain's How Stuff Works*. Available at: http://www.howstuffworks.com/airport-security4.htm.

[52] More information is available at American Science and Engineering's website: http://www.as-e.com/

Barring such sophisticated technology, CAPS must be dismantled. The carnival booth effect prevents CAPS from being an effective deterrent against terrorism. Effective deterrents are better-trained security personnel, more effective questioning, and more advanced administrative screening. To truly make our airways safer, these policies need to be applied to all passengers, not just the few that get flagged for extra scrutiny. In short, the financial resources our nation commits to counter-terrorism should not support a CAPS-like system, which may appeal to our intuitions, but is in fact more amenable to compromise.

[Aaron wrote Sections 4, 5, 7 and wrote the computer simulation.

Samidh served as the editor of this paper and wrote Sections 1, 2, 3, 6.]