Chapter 4

Results and Observations

4.1 Results

4.1.1 Generating Results

The experiments were run on a Compaq Presario 5690 with a 500 megahertz Intel Pentium III Processor, a 19 GB hard drive, and 128 megabytes of RAM running Windows 98. The SAFE Model was run for three different GHC algorithms: Pure Local Search, Threshold Acceptance, and Simulated Annealing. Each algorithm was run for 30 replications, each with 22,500 iterations. While 22,500 iterations are theoretically not the maximum of the system, it is the practical limit for post-processing analysis by Microsoft Excel 97. Since Excel has a finite one-column limit of 62,500 lines, this limit represents the maximum number of iterations that it is possible to analyze using Excel. In order to create one Excel file containing all 30 replications, and considering that each replication represents one column in Excel, it was necessary to reduce the number of rows to accommodate the full 30 replications that are required to obtain statistically significant results. Since the output of the SAFE Model needs to be imported into Excel for post-processing analysis after each replication, the resulting Excel files increase in size with the importation of each replication. Eventually a point is reached where the files attain so great a size that it cannot be opened in Excel, requiring the system to be rebooted. Trial and error has demonstrated that the practical limit on the number of iterations (or lines in Excel) that can be managed without the system rebooting is at or around 22,500 iterations. Furthermore, this limit could not be overcome with the use of more capable hardware, since it is Excel's inherent capacity limitation that causes this problem.

When executed, the SAFE Model generates a series of seven text files for each replication. Four of the files contain, respectively, the number of false alarms, true alarms, false clears, and true clears generated by each detection system tested in each iteration. The fifth file contains the devices used in each detection system being examined during each iteration. The sixth and seventh files keep track of the total costs of the detection systems. In each iteration, the sixth file records the cost of the current detection system under evaluation during that iteration. The seventh file stores the optimal detection system cost generated up to that point during the execution of the SAFE Model. If, after the completion of an iteration, the SAFE model finds that the cost of the detection system for that iteration is lower than the cost of the optimal detection system, it would make the cost of the current detection system the new optimum.

All of the post-processing analysis was performed in Excel. After each replication, the text files that are generated by the SAFE model are read into Excel and then processed by Excel. The processing consists of taking the minimum, maximum, mean, and standard deviation for each replication. As discussed previously, there are thirty replications and each replication has 22,500 iterations. There are several other items of interest that are noted during post-processing. The iteration number indicating the best-found solution for each replication is recorded, so that the speed of each experimental configuration can be evaluated. The number of replications that find the best-found solution for each experimental configuration is also noted, since the more frequently the best-found solution is found, the better that experimental configuration may be. The detection system or systems that produce the best-found solution are themselves noted, so that if there is more than one solution that produces the best-found solution, commonalties between the detection systems can be examined.

The 22,500 iterations were divided into inner and outer loops as per the structure of a GHC algorithm. There were three separate cases run for each of the three algorithms. The first case had 150 inner and 150 outer loops, the second case had 100 outer and 225 inner loops, and the third case had 225 outer and 100 inner loops. Each of these cases used the same thirty randomly generated sequences for their initial conditions.

4.1.2 Initial Conditions

It was decided that the initial conditions for the SAFE Model would contain a randomly chosen sequence of eight devices. The sequence of devices was chosen by a random number generator designed in Visual Basic. This random number generator was run for each of the thirty replications. Once a sequence was chosen for a particular replication, it was always used for that replication regardless of the algorithm employed. The sequences used for each replication are listed in Table 4.1.

Table 4.1

Sequences for Initial Conditions

1	Active Prescreen	Hotel	India	Lima	Zulu	Foxtrot	Juliet	Mike
2	Mike	Lima	Romeo	India	Kilo	Hotel	Quebec	Peru
3	Uniform	Quebec	Baggage Match	Juliet	Romeo	October	Yankee	Hand Search
4	Peru	Xray	Romeo	Kilo	Sierra	Juliet	Whiskey	Quebec
5	Xray	Yankee	Mike	Whiskey	Golf	Uniform	Sierra	Romeo
6	Peru	Baggage Match	Yankee	Sierra	Xray	Romeo	Whiskey	Uniform
7	Zulu	October	Hotel	Quebec	Foxtrot	Uniform	Mike	Whiskey
8	Juliet	Yankee	Sierra	Uniform	Mike	Quebec	Whiskey	Kilo
9	November	Sierra	Whiskey	Romeo	Hand Search	Mike	Xray	Peru
10	Juliet	Novemb er	October	Active Profile	Xray	Peru	Zulu	Sierra
11	Quebec	Xray	Romeo	Peru	Zulu	Yankee	Hotel	Whiskey
12	Xray	Mike	India	Yankee	Whiskey	Juliet	Uniform	Kilo
13	November	Whiskey	Quebec	India	Sierra	Yankee	October	Xray
14	Foxtrot	Romeo	Uniform	Lima	October	Baggage Match	November	Whiskey
15	Whiskey	Golf	Kilo	Hotel	Yankee	Zulu	Peru	Mike
16	Zulu	Xray	Hand Search	Quebec	Golf	Sierra	Yankee	October
17	Peru	Romeo	Yankee	Kilo	Hand Search	Lima	Golf	Zulu
18	Romeo	Juliet	Yankee	October	Xray	Sierra	Whiskey	Lima
19	Zulu	Yankee	Active Profile	Golf	Juliet	Uniform	November	Quebec
20	Hotel	Golf	Lima	Mike	Xray	Peru	Foxtrot	Sierra
21	Uniform	Foxtrot	Mike	November	Kilo	Xray	Hotel	Whiskey
22	Romeo	India	Uniform	Hotel	Sierra	November	Foxtrot	Kilo
23	Xray	Lima	Yankee	Zulu	Romeo	Golf	India	Uniform
24	November	Baggage Match	Quebec	Yankee	Peru	Zulu	Xray	Romeo
25	Kilo	Whiskey	Uniform	Yankee	Hand Search	Xray	Sierra	November
26	Xray	Kilo	October	Golf	Whiskey	Mike	Yankee	Peru
27	Peru	Juliet	Quebec	Uniform	Foxtrot	Sierra	Active Profile	India
28	November	Quebec	Uniform	Zulu	Hand Search	Romeo	October	Xray
29	Sierra	Romeo	Foxtrot	Whiskey	Novemb er	Uniform	Lima	Mike
30	Romeo	India	Uniform	Zulu	Mike	Quebec	Whiskey	Kilo

It should be noted that the initial sequence of devices used in the SAFE model, as shown in Table 4.1, are a very small sample of the possible combinations of eight devices. In fact, the number of possible sequences can be computed using permutations as follows:

$$\begin{aligned}\text{Number of possible sequences} \\ &= {}_{25}P_8 \\ &= 25*24*23*22*21*20*19*18 \\ &= 4.36\text{x}10^{10}\end{aligned}$$

Given the enormous number of possible sequences, even when limiting the number of devices contained in the sequence to eight, it would be impossible to examine every possible combination of devices.

These conditions were used to test each detection system:

- 100,000 total entities entered the system,
- 100 out of the 100,000 were threat entities, and
- 99,900 out of the 100,000 were non-threat entities.

The objective function(s) that were used:

Min FA $\leq \varepsilon$, and

Min Total Cost $\leq \alpha$,

Where FC = 0 and ε, α are very small.

4.1.3 Computation Time of the SAFE Model

The computational time (i.e. time it takes to complete one run of the model) for each of the three search algorithms, Local Search, Threshold Acceptance, and Simulated Annealing, varies depending on the algorithm. A summary of the computational or running time is given in Table 4.2

Table 4.2

Computational Time by Search Algorithm (in seconds)

	Avg. Time	Min Time	Max Time	Sample Variance
Local Search	131	125	137	24
Threshold Acceptance	131	125	137	24
Simulated Annealing	133	127	136	33

It is apparent from this summary that the simulated annealing search algorithm takes slightly longer to run than either the Local Search or Threshold Acceptance algorithms. The difference in computational times appears to be small enough that it could be considered insignificant.

4.1.4 Results of the SAFE Model

The SAFE Model generates a vast amount of data from which many results can be drawn. Since the quantity of data generated is so enormous, it is difficult to analyze all possible data elements. Therefore, it is necessary to determine a methodology that can allow for a thorough analysis of those data points that are most interesting.

The first step in this methodology is to determine which, on average, of the nine experimental configurations eliminates the occurrence of false clears the most rapidly. The rationale for beginning the analysis from this perspective is the very large cost associated with a false clear. At 500 million dollars each, the cost of a single false clear is, by far, the greatest portion of the total cost for any detection system in which a false clear occurs. Figure D-1 (in Appendix D) shows the total cost, averaged over the 30 replications, for each iteration during the model run. The experimental configuration that minimizes total cost the most rapidly is the one that will be subjected to further examination.
From Figure D-1 and Figure D-2, (the latter being an inset view of Figure D-1 examining iterations 500 to 1,000), it may be determined that Simulated Annealing with 225 inner loops and 100 outer loops minimized the number of

false clears the fastest. On average, Simulated Annealing with 225 inner loops and 100 outer loops had no occurrences of false clears after iteration 704. All of the other experimental configurations minimize-out the occurrence of a false clear by iteration 1,405, on average. However, since Simulated Annealing with 225 inner and 100 outer loops eliminates false clears the fastest, it will be examined in detail, below.

The next step in the analysis is to determine which of 30 replications that used Simulated Annealing with 225 inner and 100 outer loops should be further examined. The same methodology that was used previously is once again implemented here. Figure D-3 shows the total cost for each of the 30 replications throughout the model run. Figure D-4 is an inset view of Figure D-3 showing the results of iterations one to 1,000. This segment of the model run is the most interesting because the number of false clears is zero in all of the replications by iteration 1,000. However, replication 3 eliminates false clears the fastest. In this replication, all false clears are eliminated by iteration 178 and therefore it is replication 3 of the Simulated Annealing algorithm with 225 inner and 100 outer loops that will be studied further.

The solution considered to be optimal by replication 3 for Simulated Annealing with 225 inner and 100 outer loops has a total cost of $150,375. The number of false alarms is 19,231, the number of true clears is 80,669, the number true alarms is 100, and the number of false clears is zero. (See Appendix C for detailed results of all replication from all experimental configurations.) The detection system that generates this solution contains only two devices. The first device is CTX5000AU and the second device is Hotel. This detection system produces a very good result and could be considered an best-found solution, unless a superior solution can be found.

4.1.5 The Best-found Solution

This potentially best-found solution is attractive in that it costs only $150,375. However, a better solution was found with a total cost of $150,077. (See Appendix C) This best-found solution has 9,740 false alarms, 90,160 true clears, 100 true alarms, and no false clears. The detection system that generates this solution contains only one device, Hotel. Further discussion of the best-found solution appears in Sections 4.2.1 and 4.2.4.

4.1.6 Verification of the Generalized Hill Climbing (GHC) Algorithm

At this point, it would be desirable to verify that the GHC algorithm did indeed perform as expected. For this verification, replication 3 for Simulated Annealing with 225 inner and 100 outer loops will once again be used.

Figure D-5 is a scatter plot of the actual number of false clears generated at each iteration. By plotting the actual number of false clears, the outcome without using GHC can be observed. By contrast, Figure D-6 shows the results of using GHC to minimize the number of false clears. There is an obvious and significant difference in the two graphs. Figure D-6 shows an initial solution containing only two false clears and as the SAFE Model runs, the occurrences of false clears drop to zero. Figure D-7 is an inset view of Figure D-6, the inset of which shows only the first 1,000 iterations. Figure D-7 gives a higher resolution of the area of interest, which better shows the rapid elimination of false clears.

A similar result can be seen for false alarms in Figures D-8, D-9, and D-10. Figure D-8 is a scatter plot of the actual number of false alarms generated at each iteration. By plotting the actual number of false alarms, the outcome without using GHC can be observed. By contrast, Figure D-9 shows the results of using GHC to minimize the number of false alarms, while simultaneously eliminating the number of false clears. Figure D-10 is an inset view of Figure 4.9, showing only the first 1,000 iterations. Both Figures D-9 and D-10 show that while the number of false alarms cannot be completely eliminated, they can be held to some finite limit. Furthermore, that finite limit is a small one; only 19,231 false alarms out of a possible 99,900.

It should be noted that the results observed in replication 3 for Simulated Annealing with 225 inner and 100 outer loops are typical of those found in the majority of replications for all of the various experimental configurations.

4.2 Observations and Discussion

4.2.1 Local Optima versus Global Optima

The results of the SAFE Model show that without question the solution space contains several local optima in addition to the global optima. The tables in Appendix C support this. Table 4.3 shows the total cost for each of the various local optima and the number of replications-- out of the thirty replications-- that report a particular local optima as the final result.

Table 4.3

Number of Occurrences of Local Optima

Total Cost (in dollars)	LS 100, 225	LS 150, 150	LS 225, 100	TA 100, 225	TA 150, 150	TA 225, 100	SA 100, 225	SA 150, 150	SA 225, 100	Total
150,375	19	12	18	20	21	12	19	23	24	168
151,226	4	14	5	7	4	5	7	5	3	54
152,293	0	0	1	0	0	0	0	0	0	1
154,621	3	1	2	0	3	6	0	0	1	16
154,823	1	1	1	0	1	0	0	0	1	5
155,372	0	1	1	0	0	2	0	0	0	4
155,890	0	0	0	0	0	1	0	0	0	1
156,916	0	0	0	0	0	1	1	0	1	3
Total	27	29	28	27	29	27	29	28	30	252

Tables 4.3 and 4.4 are a summary of the results that appear in Appendix C. The results in Table 4.3 can be compared with those in Table 4.4, where the number of replications-- out of the thirty replications-- that report the global optima as the final result.

Table 4.4

Number of Occurrences of Global Optima

Total Cost (in dollars)	LS 100, 225	LS 150, 150	LS 225, 100	TA 100, 225	TA 150, 150	TA 225, 100	SA 100, 225	SA 150, 150	SA 225, 100	Total
150,077	3	1	2	3	1	3	3	2	0	18

Several observations can be made from these results. The first is that taken as a group, the probability of reaching a global optimum is only 6.67%, assuming that no one experimental configuration is better than any other. It can also be noted that there is a slightly greater chance of reaching a global optimum if the experimental configuration has 100 outer loops and 225 inner loops.

From the results discussed above and the results presented in Section 4.1.4, it can be seen that there is no discernable advantage in using a particular experimental configuration to find the best-found solution for this problem. Table 4.2 shows that no one experimental configuration has a significantly faster computational speed; all are within seconds of each other. As illustrated by the graphs in Appendix D, while the Simulated Annealing with 225 outer and 100 inner loops minimizes-out false clears most quickly, it is not significantly faster than any of the other experimental configurations. Finally, the above analysis shows that no experimental configuration is significantly more likely to locate a global optimum. Therefore, it can be assumed that all of the experimental configurations are equally good for optimizing this problem.

4.2.2 Discussion of the Different Detection Systems chosen by the SAFE Model

From the discussion in Section 4.2.1, it is obvious that there are only a finite number of solutions that are chosen by any experimental configuration as either a local optimum or a global optimum. Each of these solutions has a total cost that can be considered very small, especially when taking into account that the total cost of a detection system in which every single threat entity clears would be greater than 50 billion dollars. The 50 billion dollar figure is calculated by multiplying the number of threat entities (100) by the amount of a false clear (500 million). Therefore, it can be safely said that any of the detection systems chosen as a solution would be sufficient to employ.

Table 4.5 shows the detection systems chosen by the SAFE Model as best-found solutions.

Table 4.5

Optimal Detection Systems and their associated Total Cost.

Total Cost (in dollars)	Detection System
150,077	Hotel
150,375	CTX5000AU, Hotel
151,226	India, Hotel
152,293	Foxtrot, India, Hotel
154,621	CTX5000AU, India
154,823	Hotel, India
155,372	Hand Search, Hotel
155,890	Golf, India, Hotel
156,916	CTX5000AM, Hand Search

As shown in Table 4.5, the solution that has the lowest cost ($150,077) also only has one device, Hotel. Deploying a system with only a single device could be a dangerous solution since it would offer no redundancy to the system. If the device failed or was tampered with, there would be no opportunity for a second device to detect security breaches. Since the SAFE Model does not deal with failure of devices, it is unknown what type of effect such an occurrence would have.

Another problem with deploying a detection system with only one device is that a single device may have difficulty detecting a particular type of threat. For instance, a device may have a very high probability of detecting a handgun, but have a lower probability of detecting explosive material. However, the device's overall probability of detection may still be high once the individual detection probabilities are combined. Since this work deals with only a device's overall probability of detection, there is no way to know if this type of situation exists.

Considering all of these issues and taking into account the fact that each of the solutions examined has a total cost within a range of seven thousand dollars, it may be worth recommending one of the local best-found solutions.

4.2.3 False Alarm and False Clear Rates

Table 4.6

False Alarm and False Clear Rates for Initial Solutions

Replication Number	False Alarm Rate (%)	False Clear Rate (%)
1	98.22	1
2	24.50	84
3	15.72	2
4	40.80	13
5	64.43	69
6	43.56	14
7	30.64	17
8	20.16	45
9	5.53	28
10	85.15	45
11	28.71	21
12	7.32	54
13	19.22	19
14	35.29	22
15	10.89	3
16	15.13	27
17	1.98	81
18	5.61	45
19	36.20	13
20	85.15	46
21	28.56	21
22	1.96	85
23	32.43	6
24	5.90	29
25	84.18	36
26	1.92	45
27	12.81	4
28	19.13	12
29	10.70	3
30	20.18	32

When considering which of the detection systems listed in Table 4.5 to recommend, the rate of false alarms and false clears should be considered. Table 4.6 shows the false alarm and false clear rates for the thirty detection systems that were used as initial solutions in the SAFE Model. The initial solutions were discussed in Section 4.1.2. Table 4.7 shows the false alarm and false clear rates for each of the best-found solutions that were discussed in Section 4.2.2.

Table 4.7

False Alarm and False Clear Rates for Best-found solutions

Total Cost (in dollars)	False Alarm Rate (%-age)	False Clear Rate (%-age)
150,077	9.75	0
150,375	19.25	0
151,226	14.50	0
152,293	14.50	0
154,621	23.50	0
154,823	14.50	0
155,372	19.25	0
155,890	14.50	0
156,916	32.00	0

It is obvious from Tables 4.6 and 4.7 that any of the best-found solutions is better at minimizing false clears than the initial solutions, since the false clear rate for all of the best-found solutions is zero. The false alarm rate is much more interesting, however. The false alarm rate is significantly lower in the best-found solutions than it is in the initial solutions. Since false alarms cost the airlines money, just as do false clears, any reduction in the number of occurrences of false alarms is beneficial. However, in the cases where the false alarm rate of the initial solution is less than the false alarm rate of any of the best-found solutions, the rate of false clears is very high. Therefore, initial solutions that fall into this case would be considered undesirable due to the high rate of false clears. In any case, the false alarm and false clear rates show that it is possible to limit the number of false alarms to a reasonable level while eliminating false clears entirely. This is an important result since the TSA desires to decrease the false clear rate without increasing the false alarm rate.

4.3 Final Recommendations

The analysis of this problem has uncovered many interesting results. However, the point of this work is to give the TSA a recommendation for their detection system. Using the results discussed above, there are three options that can be suggested, each with its own benefits and drawbacks.

Option 1: Deploy a detection system with only one device, Hotel. This system was the least expensive, having a total cost of $150,077 and a false alarm rate of 9.75%. However, there is obviously no redundancy within this system. If Hotel fails, the detection system is useless.

Option 2: Deploy a system with two devices. Use one of the following systems costing $151,226, $152,293, $154,823, or $155,890, since the difference in cost is not significant. The drawback with deploying this system is that its false alarm rate is 14.5%, which means that a great many passengers may be inconvenienced.

This brings up Option 3, which is something of a compromise between real world considerations and the best-found solution found by the SAFE Model.

Option 3: Deploy two separate detection systems, a primary and a back-up. The primary detection system would be the global optimum found by the SAFE Model, Hotel. Then deploy as the back-up detection system the next least expensive system that does not employ the device Hotel. Under these circumstances, the backup detection system would be a two-device detection system, the first device is CTX5000AU and the second device is India. This back-up detection system costs only $154,621 and has a false alarm rate of 14.50%. Therefore, this option is more reliable (in terms of false alarms) than Option 2, and maintains the use of the best detection system as discussed in Option 1.

The selection of any of these options by the TSA would represent an appropriate use of the SAFE Model.