

Data-Based Detection of Potential Terrorist Attacks on Airplanes

Karen Kafadar

Department of Mathematics
University of Colorado-Denver
Denver, Colorado 80217-3364

Max D. Morris

Department of Statistics &
Department of Industrial Engineering
Iowa State University
Ames, Iowa 50011-1210

Executive Summary

Statistical algorithms have been developed to detect unusual events or aberrations that may indicate serious problems; e.g., high incidence of infectious diseases (possible epidemics), excessive charges on a credit card or telephone (possible stolen card), or random operations on a computer (attack by a virus). Similar algorithms can be developed to detect unusual combinations of characteristics of passengers booked on an airplane. These algorithms will require availability, coordination, and summarization of relevant information from various data sources. When used in combination in an intelligent way, these data can provide real-time feedback to airline system personnel and security agents that can be used to identify potential threats to airline passenger safety before the threat has a chance to materialize. This problem has features that distinguish from the other applications noted above, and its solution will require new methodology and a different combination of resources.

Problem

The terrorist attack of September 11, 2001 is dramatic evidence that the security of the national airlines system is inadequate. The reaction to this event has been to increase security at airports: (1) allow only ticketed passengers past the security gates; (2) screen carry-on luggage more carefully for possible weapons. Neither measure would have flagged the terrorists on September 11; all were legitimately ticketed, and the “weapons” were no more sophisticated than box cutters and razor blades. Current plans for enhanced

physical security measures may help reduce the risk of a similar future event. However, the development of new information-based security systems will offer a valuable addition to our defense against such attacks.

After the fact, some common elements of the suspected terrorists are obvious: none were U.S. citizens, all had lived in the U.S. for some period of time, all had connections to a particular foreign country, all had purchased *one-way* tickets at the gate with *cash*. The statistical odds that five out of 80 revenue passengers (in the case of one of the four hijacked flights on September 11) fit this profile might, by itself, be unusual enough to warrant concern. However, even if this were not the case, the sophistication and extent of databases *not* owned by the airlines given the present access arrangements, may be needed to provide additional critical information on how “unusual” such events may be.

The credit card and telecommunications industries have developed and applied, with considerable success, statistical algorithms to identify unusual events. Such fraud detection algorithms use profiling techniques to develop a comprehensive “profile” of the user, based on numerous characteristics (e.g., average length or price of purchase, percentage of calls or purchases exceeding a given threshold, gender, and profession). From this profile, an unusual charge is flagged, and the user is called for verification; if the user is unavailable, the charge is instantly denied and the card is temporarily suspended. These algorithms must fast enough to operate in real time, and be carefully tuned to utilize the most relevant information in the most efficient way possible. A similar approach can be used to design tools to ensure the safety of air travel.

Approach

Statistical methods of multivariate outlier detection have been used in numerous environments; e.g., in public health (identifying possible disease epidemics early before they spread), credit card industry (detection of fraudulent usage), telecommunications (intercepting cellular phone numbers), computer systems (break-ins), industrial manufacturing (monitoring a process for defective units in the production of thousands of parts with multiple measurements on each part). Given perhaps thousands of characteristics on perhaps millions of items, which ones are most important in capturing the distinctive aspects of the individual items relevant to detecting problems of concern? A common scenario in statis-

tical process/quality control involves part production; as a part enters the system, several of its characteristics are measured, and, if any of them (or combination of them) is out of specification, the part is flagged as a defect. We expect a certain fraction of supposed “defects” just by chance alone, but too many false-alarms would have adverse effects on the overall process (e.g., shut-downs). The quantification of the types and numbers of defects is the subject of SPC and SQC techniques (cf. the Westinghouse rules for identifying aberrant observations in the *Statistical Quality Control Handbook*, 1984, Section 1B). Similarly, the Centers for Disease Control and Prevention (CDC) monitors the number of cases of various diseases in the National Notifiable Diseases Surveillance System, which assists the CDC in identifying potential public health outbreaks (Stroup et al. 1993; Kafadar and Stroup 1992). In this application, attention is focused on identifying “clusters” of new cases that are sufficiently atypical of “background” to warrant action.

The same principles, adapted and modified to the specific nature of the problem, can be used in the context of screening passengers on an aircraft. A large quantity of historical information on the characteristics of airline customers may be available from travel agents, booking agencies, or the airlines’ marketing departments. It may be possible to obtain a very precise estimate of the proportion of airline tickets that that are purchased with cash (versus a credit card or purchase order) as well as some other typical characteristics of cash-paying customers (e.g., widows over the age of 60). If the historical data suggest that only 1.0% of all airline tickets are purchased with cash, then one or two cash sales on an airplane of 100 passengers would not be unusual, particularly if these cash-paying customers were elderly women. Five such cash-paying customers, none of whom is female nor elderly, might be unusual enough – relative to what is typical of air traffic – to trigger a signal as a suspicious event. The challenge in data mining is to identify which characteristics, among the thousands that are available through various data bases, are most useful in flagging the unusual, possibly terrorist, customers, without unduly detaining or violating the civil rights of the innocent.

The present solution to screen everyone is helpful but may not, by itself, reduce the risk of future terrorist attacks sufficiently. The kind of data-screening approach described here can be applied to identifying unusual profiles, or combinations of them, among passenger

lists, and would bring complementary information to the process, thus enhancing the overall efficiency of security screening. Replacing or augmenting some of the physical search effort with automated computer searches will limit the number of flights that will require an intensive physical screening effort.

Distinctive features

Three key features of this problem distinguish it from the data mining efforts applied in these other contexts. First, credit card and phone companies usually have numerous transactions on an individual from which a reasonably precise profile can be derived. In a given year, perhaps 100–300 charges have been placed; some companies used to require a deposit until such historical data on the customer can be collected. In contrast, a given year of travel for an individual might consist of less than a dozen flights. Second, summarizing information across flights throughout the year is complicated by the fact that flight numbers change and flights themselves are replaced or canceled. Third, the algorithm for identifying fraud in the credit card industry is developed by relying on large numbers of known fraudulent events. In contrast, we have only a handful of suspicious events represented in the data collected on air traffic over the past decade. As a result of these distinctions, new tools must be developed that rely on more diverse forms of information.

Tasks

Several tasks would have to be completed in the process of developing the methodology envisioned here. First, available data resources must be reviewed to characterize relevant available passenger information. Given a passenger’s name, address, and a contact phone number, various data bases (public or private) can identify the social security number (SSN), from which much information will be readily available (credit history, police record, education, employment, age, gender, etc.). A large effort will be required to locate these data bases, obtain necessary permissions to use these data, tabulate the variables that they contain, and develop algorithms to permit linkages across these data bases. Second, given the huge number of characteristics available on both individual passengers and general industrial trends, the most useful approaches for identifying “signals” within the natural variability or “noise” must be developed. This problem requires dimension reduc-

tion techniques; presently, some form of principle components is most often used, but other combinations might be more effective. Third, acceptable “error rates” in (a) falsely detaining an innocent passenger, and (b) failing to detain a plane that carries a terrorist, must be characterized in light of the risks and relevant policy constraints. Ideally, both “sensitivity” (probability of correctly identifying a suspicious event) and “specificity” (probability of correctly passing a benign event) are high; we need to balance civil liberties and limited investigation resources with safety and national security. Fourth, because this will require the use of sensitive and protected data, a determination of which characteristics are allowed for purposes of identifying unusual passenger lists must be made. Various legal issues will need to be considered in this regard. Fifth, the proper integration of these methods with overall security operations will require the development of policies that direct security personnel on the appropriate response to such detected unusual patterns (e.g., five cash-paying customers, all male, all foreign, and all under the age of 30).

The first three tasks involve the development of statistical and data mining tools. The airlines already collect much data on various flights (e.g., percent occupancy, percent on-time arrival, etc.). Additional data, summarized by flight, is needed; e.g., percentage of males/females; employed/unemployed; young/old. Given the historical data of this sort, for various flight types (e.g., domestic; within region; across region; jumbo jet; etc.), we can begin to identify departures from a typical pattern. The historical data must be relevant to a specific flight. For example, a United flight leaving San Francisco for Seoul, Korea, could be expected to carry a much larger fraction of Asian passengers than one might see on a flight from, say, Des Moines to Denver.

Data mining techniques derive from methods in *exploratory data analysis*, or *EDA* (Tukey 1977); whereas *EDA* was developed on rather small data sets, data mining today uses the same philosophy but on much larger data sets. When the data come in the form of multiple characteristics on a single item, exploratory tools for multivariate data are needed, such as classification (Ripley 1995), regression trees (Brieman et al. 1984), multivariate adaptive regression splines/trees (Friedman 1991/2001), and methods of statistical learning (Hastie, Tibshirani, and Freidman 2001).

This discussion emphasizes the need for statistically based analysis techniques based on

linked resources in support of real-time decisions. Issues regarding the restrictions on legal access to databases, and strategies on how detected anomalies should be treated, are also critical to the successful development and deployment of the system envisioned here, but will require expertise from outside of the mathematical sciences.

Impact

The potential impact of this work is profound. Most importantly, the security of the air transportation can be improved substantially through modern, intelligent use of pattern recognition techniques applied to large linked databases. Second, the reduction in security risk is important, not only for the physical safety of citizens, but also for our economic, social, and political systems as well. Finally, additional impact may be realized through connections between this work and other event-screening methodologies, such as are used in the diverse applications areas noted above.

References

- Brieman, L., Friedman, J.H., Olshen, R., Stone, C. (1984), *Classification and Regression Trees*, Wadsworth.
- Friedman, J.H. (1991), "Multivariate Adaptive Regression Splines", *Annals of Statistics*, 1-61.
- Friedman, J.H. (2001), "Multivariate Adaptive Regression Trees", preprint.
- Kafadar, K., Stroup, D.F. (1992), "Analysis of aberrations in public health surveillance data: estimating variances on correlated samples," *Statistics in Medicine* 11: 1551–1568.
- Ripley, B.D. (1995), *Pattern recognition and classification*, Cambridge University Press.
- Statistical Quality Control Handbook, 10th ed.*, 1984.
- Stroup, D.F., Wharton, M., Kafadar, K., Dean, A.G. (1993), "Evaluation of a method for detecting aberrations in public health surveillance data," *American Journal of Epidemiology* 137(3): 373–380.
- Tukey, J.W. (1977), *Exploratory Data Analysis*, Addison-Wesley.