Federal Aviation Administration

Mining for Information in Accident Data

Nelda J. Milburn
Lena Dobbins
Julia Pounds
Scott Goldman

Federal Aviation Administration
Civil Aerospace Medical Institute
Oklahoma City, OK 73125

November 2006

Final Report

NOTICE

This document is disseminated under the sponsorship
of the U.S. Department of Transportation in the interest
of information exchange. The United States Government
assumes no liability for the contents thereof.

_____

This publication and all Office of Aerospace Medicine technical reports are
available in full-text from the Civil Aerospace Medical Institute's publications
Web site:
www.faa.gov/library/reports/medical/oamtechreports/index.cfm

Technical Report Documentation Page

1. Report No. DOT/FAA/AM-06/26	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Mining for Information in Accident Data		5. Report Date November 2006
		6. Performing Organization Code
7. Author(s) Milburn N, Dobbins L, Pounds J, Goldman S		8. Performing Organization Report No.
9. Performing Organization Name and Address FAA Civil Aerospace Medical Institute P.O. Box 25082 Oklahoma City, OK 73125		10. Work Unit No. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agency name and Address Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591		13. Type of Report and Period Covered
		14. Sponsoring Agency Code

15. Supplemental Notes

Work was accomplished under approved task AHRR522.

16. Abstract

Purpose. This project evaluated WinMine, an analytic tool developed by Chickering, Heckerman, Meek, Platt, and Thiesson (2000) to determine its usefulness for identifying higher-order relationships in research data from dynamic and high-consequence aviation events. Traditionally, researchers have relied on several types of analyses to better understand the relationships between factors related to an outcome. However, researchers need an analytic approach that can clearly illustrate the interactions among causal factors as probabilities associated with the chain of events. **Method.** A convenience sample of aviation accident data previously classified using the Human Factors Analysis and Classification System (HFACS; Shappell & Wiegmann, 2000; 2001) was used to evaluate WinMine in contrast to traditional methods, such as bar graphs, contingency tables, and odds ratios. **Results.** WinMine showed an advantage when compared with other methods because it graphs quantifiable interrelationships between factors and illuminates the underlying hierarchical structure of variables. **Discussion.** Each technique examined contributed toward understanding the causal factors; however, WinMine provided a better picture of the factor interrelationships than the other methods.

17. Key Words WinMine, HFACS, Accident Data, Decision Tree, Analytic Tool, Conditional Probability		18. Distribution Statement Document is available to the public through the Defense Technical Information Center, Ft. Belvior, VA 22060; and the National Technical Information Service, Springfield, VA 22161	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 11	22. Price

Form DOT F 1700.7 (8-72)   Reproduction of completed page authorized

Mining for Information in Accident Data

INTRODUCTION

Researchers have used a variety of analytic techniques to better understand factors related to aviation incidents. The *format of the data,* the *type of analysis,* and the *method of presenting* the results were all of interest for this project. For comparison, 3 reports involving aviation accident data were reviewed.

1) *Review of Aviation Accidents Involving Weather Turbulence in the United States: 1992-2001* NASDAC (2004) analyzed aviation accident/incident data extracted from the National Transportation Safety Board (NTSB) database. The report was a 10-year review in which there were "a total of 20,332 accidents that occurred in the United States" (p. 1). The NTSB "cited weather as a cause or factor in 4,326 accidents. Of these weather events, the NTSB cited weather turbulence as a cause or factor in 509 accidents, or eleven percent of the total weather accidents" (p.1). The report summarized the total number of accidents per year that were weather-related compared to weather turbulence accidents (in a bar chart), the percentage of weather turbulence accidents from all weather-related accidents per year (in a tabular format), the total weather accidents by phenomenon (in a pie chart), the number of general aviation weather turbulence accidents by month (in a bar chart), and many other graphs and charts that are limited by the number of variables that can be presented at once. Each of these methods permits an understanding of the frequencies of each factor, a comparison of frequencies of one factor (on the x-axis) with another factor (on the y-axis), or a particular factor (such as the number of injuries) as a function of another variable (such as type of turbulence). However, these graphs do not allow the researcher to understand complex relationships between variables.

2) *Review of Aviation Accidents Occurring in the State of Alaska* Office of System Safety (ASY, 2003) analyzed 20,325 accidents that occurred in the United States between 1992 and 2001 extracted from the NTSB Aviation Accident and Incident Database. The database was filtered to include only accidents that occurred in the state of Alaska—a total of 1,569 accidents. The frequency data were summarized in tabular form as counts and percentages, bar charts, pie charts, and as a line graph depicting the number of accidents per year that was overlaid with a trendline. The frequency data were also used to compute monthly averages. The pie charts were easy to interpret—for example: as the percentage of accidents that occurred as a function of light condition. However,

this type of analysis makes it difficult for the researcher to discern complex relationships between variables.

3) Work by Shappell and Wiegmann (2003) examined general aviation (GA) accidents classified as controlled flight into terrain (CFIT) –an "in-flight collision with terrain, water, or obstacle without indication of a loss of control" using the definition provided by the ICAO (International Civil Aviation Organization) Common Taxonomy Team (cited in Shappell & Wiegmann, 2003).The data were dichotomous, that is, coded as either 1 or 0, depending upon whether the factor was judged by analysts to be present or absent in the accident. If present, the factor was classified into 1 of the 17 Human Factors Analysis and Classification System (HFACS) categories (Table 1). Odds ratios were used to

Table 1. List of the 17 Human Factors Analysis and Classification System (HFACS).

Organizational Influences

1. Resource Management
2. Organizational Climate
3. Organizational Process

Unsafe Supervision

1. Inadequate Supervision
2. Planned Inappropriate Operations
3. Failed to Correct Problem
4. Supervisory Violations

Preconditions for Unsafe Acts

1. Substandard Conditions of Operators
   a) Adverse Mental States (AMS)
   b) Adverse Physiological States (APS)
   c) Physical/Mental Limitations (PML)
2. Substandard Practices of Operators
   a) Crew Resource Mismanagement (CRM)
   b) Personal Readiness (PR)

Unsafe Acts of Operators

1. Errors
   a. Decision Errors (DE)
   b. Skill-Based Errors (SBE)
   c. Perceptual Errors (PE)
2. Violations (V)
   a. Routine
   b. Exceptional

(Source: Shappell & Wiegmann, 2000, 2001)

examine the likelihood of a causal factor resulting in a CFIT, and the results were presented in a tabular format (Table 2). The analysis showed the relationship of each factor independently with CFIT, but it could not show the interaction between the factors, nor the probability associated with each factor along a chain of events.

Clearly, a better way of examining and reporting the affect of each factor is needed. Potentially, a probabilistic relational analysis could improve our understanding of the interaction between causal factors in dynamic aviation events. Therefore, the purpose of this project was to evaluate WinMine (Chickering, Heckerman, Meek, Platt, & Thiesson, 2000) as an analysis method to determine its usefulness for identifying higher-order relationships in an archival aviation database. To test this, a convenience sample of data was borrowed from an analysis of dynamic and high-consequence aviation events (the third study reported above). Because the main focus of this paper is an examination of the WinMine tool rather than a traditional hypothesis test, the paper will not interpret the results of the data analysis or the specific probabilities associated with the factors in GA accidents but will instead limit the scope of the research to an evaluation of the WinMine software.

METHOD

Database

Accident data previously classified using HFACS (Shappell & Wiegmann, 2000; 2001; 2003) were used to demonstrate the functionality of the WinMine software. Five certified flight instructors served as subject matter experts and analyzed 16,500 accidents to create the database of which 16,278 were used in this analysis. Two pilot-raters independently coded each accident as CFIT or not (coded 1 or 0, respectively) according to the definition provided by the ICAO Common Taxonomy Team. The pilot-raters classified the GA accidents into 17 causal categories defined by the NTSB, 9 of which were used in this study: Adverse Mental States (AMS), Adverse Physiological States (APS), Physical/Mental Limitations (PML), Crew Resource Mismanagement (CRM), Personal Readiness (PR), Decision Errors (DE), Skill-Based Errors (SBE), Perceptual Errors (PE), and Violations (V). Notice that the category "Violations" is divided into 2 sub-categories (routine and exceptional) on the complete list (Table 1), but the higher-level classification was used for this analysis. Of the 4 levels of failure, only the factors described in the sub-categories under "unsafe acts of operators" and "preconditions for unsafe acts" were used for this analysis.

Apparatus

The WinMine Toolkit (Chickering, Heckerman, Meek, Platt, & Thiesson, 2000) is a "set of tools for Windows 2000/NT/XP that allow you to build statistical models from data and graphically represents the results. Development of the toolkit was performed by the same team that has contributed to the data-mining technologies in Microsoft's SQL Server database product" (Chickering, 2002). Some of the tools are DOS command-line executables that can be run in scripts. For the most basic analyses, WinMine accepts dichotomously coded,

Table 2. Chi-Square and odds ratio for CFIT for each HFACS causal category.

HFACS Causal Category	Chi-square		Odds Ratio	95% Confidence Interval		
				Lower	Upper	
Unsafe Acts of Operators						
Decision Errors	1.792	ns	0.923	0.822	1.038	ns
Skilled-Based Errors	6.229	ns	1.178	1.036	1.341	ns
Perceptual Errors	50.404	p<.001	1.847	1.555	2.193	p<.001
Violations	380.748	p<.001	3.264	2.883	3.695	p<.001
Substandard Conditions of						
Operators						
Adverse Mental States	146.069	p<.001	2.907	2.427	3.482	p<.001
Adverse Physiological States	7.097	ns	1.497	1.110	2.017	ns
Physical/Mental Limitations	29.826	p<.001	0.639	0.543	0.751	p<.001
Crew Resource Management	18.916	p<.001	0.631	0.512	0.778	p<.001
Personal Readiness	136.486	p<.001	4.089	3.168	5.276	p<.001

(Source: Shappell and Wiegmann, 2003)

categorical data. More advanced techniques allow both discrete and continuous variables to be used.

WinMine graphically represents probabilistic dependencies found within a dataset in a network that is very similar to a Bayesian network and is referred to as a dependency network. The networks are similar because both methods use a graph and have a probability component. It is important to keep in mind that in its consistent form "the graph component is a cyclic directed graph such that a node's parents render that node independent of all other nodes in the network. As in a Bayesian network, the probability component consists of the probability of a node given its parents for each node—the local distributions" (Heckerman, Chickering, Meek, Rounthwaite & Kadie, 2000, p. 70). Refer to that original document for a full, in-depth description of the algorithms, assumptions, and intricacies of the WinMine product that are beyond the scope of this paper.

RESULTS

Researchers have counted, summarized, categorized, graphed, plotted, and tabulated accidents in innumerable ways—often with the goal of trying to make sense of accidents—such as what were the causes, how could the accident have been prevented, what can be learned from the accident to prevent others from happening, and countless other questions. Although researchers have employed a variety of data analysis techniques, the methodologies have failed to clearly illustrate the interactions among the causal factors and to provide probabilities associated with the factors. Therefore, in quest of a better tool to highlight or expose any relationships between possible causal factors, the WinMine product was evaluated.

WinMine has several component parts (Toolkit), with one or more modules designed to 1) read the data files, 2) merge two files together (if necessary), 3) select variables for use (or to be ignored), 4) code missing data, 5) identify the roles of variables (in the Plan Phase), 6) view the model as a function of the strength of relationships between variables, and finally 7) view a tree diagram with the detailed probabilities associated with each variable. Each of those component parts is valuable, and most are similar to functions available in other analytic software; however, items 5 through 7 are fairly unique to this software, therefore their usefulness was the focus of this evaluation.

The Plan Phase

Prior to performing an analysis, the user should specify whether a variable is to be modeled as an input or output variable, or both. That distinction essentially establishes how variables will be modeled, the role they will play in the analyses, and how they will be displayed in WinMine's graphic output called a decision tree. An "input" variable can only predict other variables. An "output only" variable is one that can only be predicted by other variables. For example, CFIT is actually an outcome or presumably the last event in a chain and could not precede a V, or a DE, or an AMS. (Granted, if the pilot walked away from the CFIT, it is likely that the pilot will experience an AMS and may contemplate possible DE or a V.) But the categories we are considering are precursors to an accident, not consequences. Logically then, CFIT should be established within the *Plan Phase* of WinMine as an "output only" variable with the goal of determining whether WinMine could show how the other factors might be related to this output variable.

In contrast, some variables may be an internal *link* in a chain of events that are therefore best defined as *both* input and output variables. Thus, an input-output variable is one that is both *predicted by* other variables and can also serve as a *predictor*. Consequently, some knowledge of the role of the variables contained within a specific data set may aid in defining their *expected* roles. However, imposing a specific structure on the variables restricts the placement of the nodes within the graphic representation of the data but will not inhibit the calculated probabilities associated with each node. Rather, defining the role of a variable (in the Plan Phase) is a way of specifying the model to test. In some cases, it makes sense to predefine the placement or ordering of variables within the model. For example, a store owner wants to determine the likelihood that someone will purchase a new computer, given that potential buyers meet the following defined set of conditions: The purchaser is currently employed, lives within a 25-mile radius of the store, has an annual income greater than $40 thousand per year, has a high-speed Internet connection such as DSL (digital signal link), and had purchased a computer more than 2 years ago. In this example, the *decision to purchase* variable should be assigned as an *output* variable. However, because the business owner is unsure about the role that each of the other variables may play or the order of sequencing of those variables with regard to the output variable, each should be coded as *input-output* variables.

Once the researcher specifies the structure, a stochastic analysis can be executed to reveal the relationships between variables and identify probabilistic dependencies. A further use of the Plan Phase will be reserved until other component parts, vital to understanding that aspect, have been described and evaluated. Items 6 and 7 are WinMine output screens, the DNet (dependency network) Viewer and the Decision Tree.

The DNet Viewer

WinMine's output is displayed in DNet Viewer. The initial output screen contains an interactive graph composed of nodes connected by arrows that indicate relational links between the variables. An impressive aspect of the graphical output is that it is dynamic with respect to presentation of the strength of relationships between the variables. The user can evaluate the model as a function of the strength of the dependencies by using the mouse to move a slider bar along its vertical length. In DNet Viewer, a slider bar is available on the left side of the screen. The top endpoint of the bar is labeled "All Links." When the bar is moved to this position, all relationships between the nodes are shown. The bottom endpoint of the bar is labeled "Strongest Links." When the bar is moved to this position, only the strongest links (dependencies) between the nodes are shown. (See Heckerman et al. [2000] for a complete description of the method used for ranking the strength of links.)

WinMine's output in DNet Viewer is a graph containing nodes and arrows indicating relational links between the variables. WinMine graphically displays the structure of the data and also redundantly color-codes the variables

(as structural nodes) for easy interpretation of their role in the structure. Figure 1 shows the result when the slider bar is set to show all relationships between all variables. An analyst interested in this type of data might interpret this pattern of links to mean that there is at least a weak relationship between most variables. When the slider bar is set at approximately 30% from the top position, the variables V, PE, and SBE show a relationship to CFIT (Figure 2). When the slider bar is moved to approximately 75% from the top position (Figure 3), only V shows a link to CFIT.

The Decision Tree

More specific details can be obtained from the tool by examining the decision tree. The decision tree shows the variable nodes including an itemized list of preliminary conditions that led to each specific point within the chain and can be examined by double clicking on each leaf node. A branch on the decision tree is formed from conditional probabilities and is a graphical way of displaying dependencies found within the data. The number of observations (accidents) is shown on each branch of the decision tree. Specific interpretation of the decision path
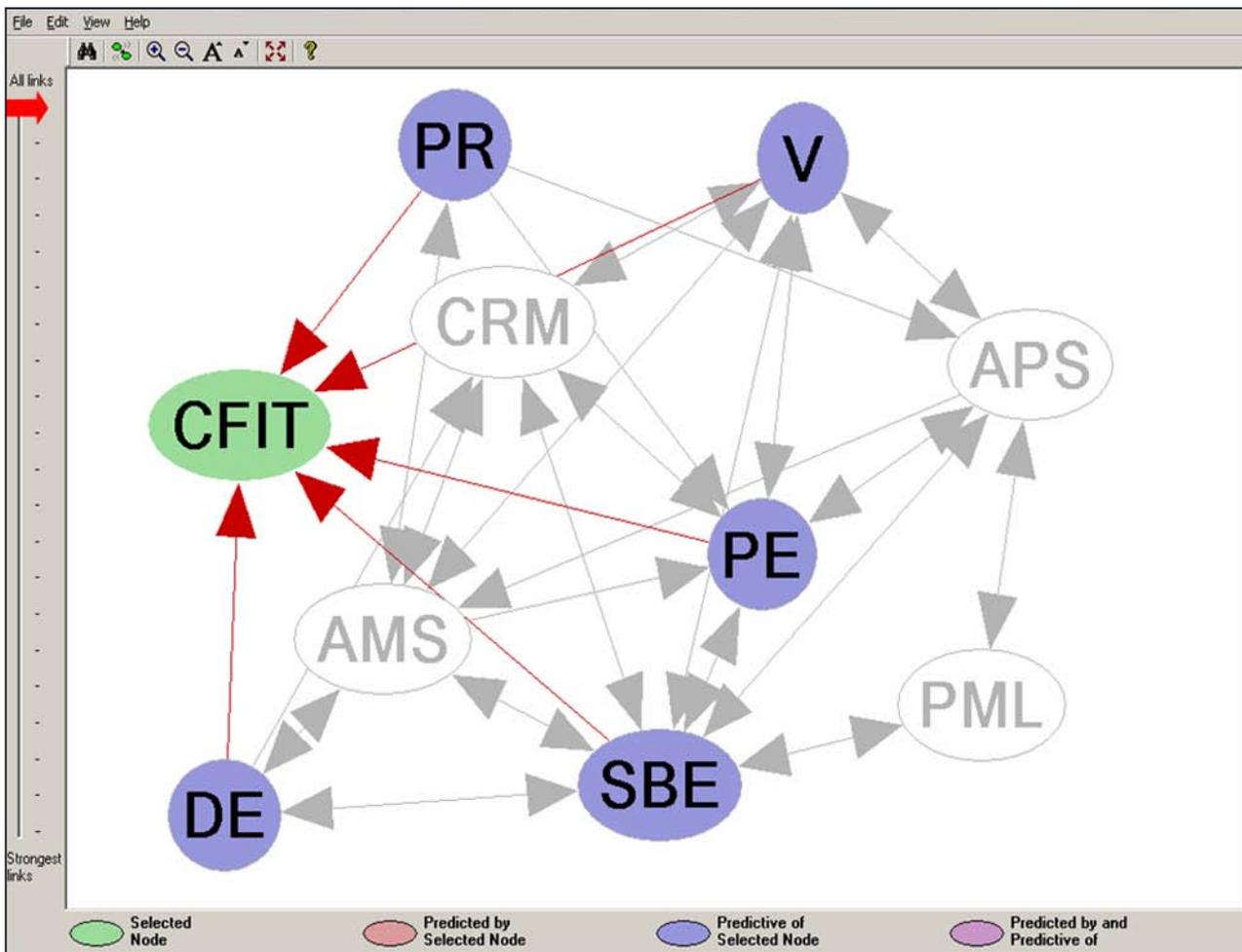

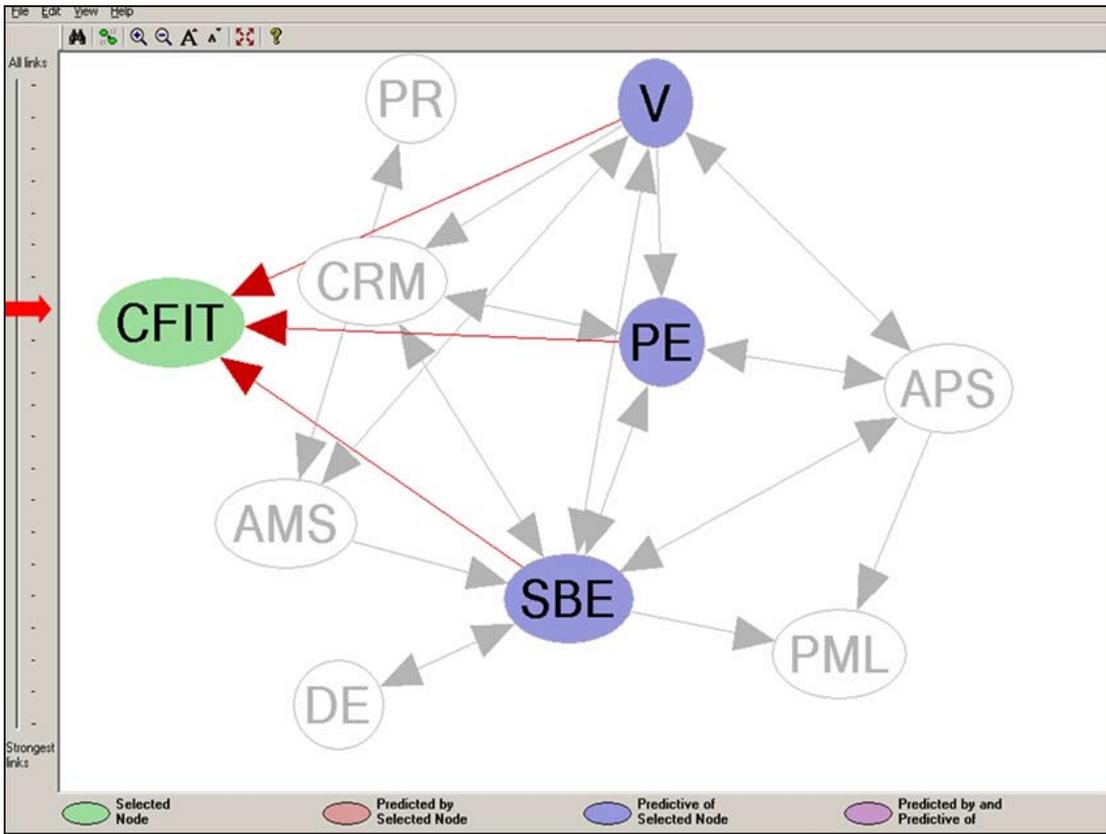**Figure 1.** Sample output when slider bar is set to show all relationships between all variables.

4

Figure 2. Example of output when slider bar is set at approximately 30% from the top position.
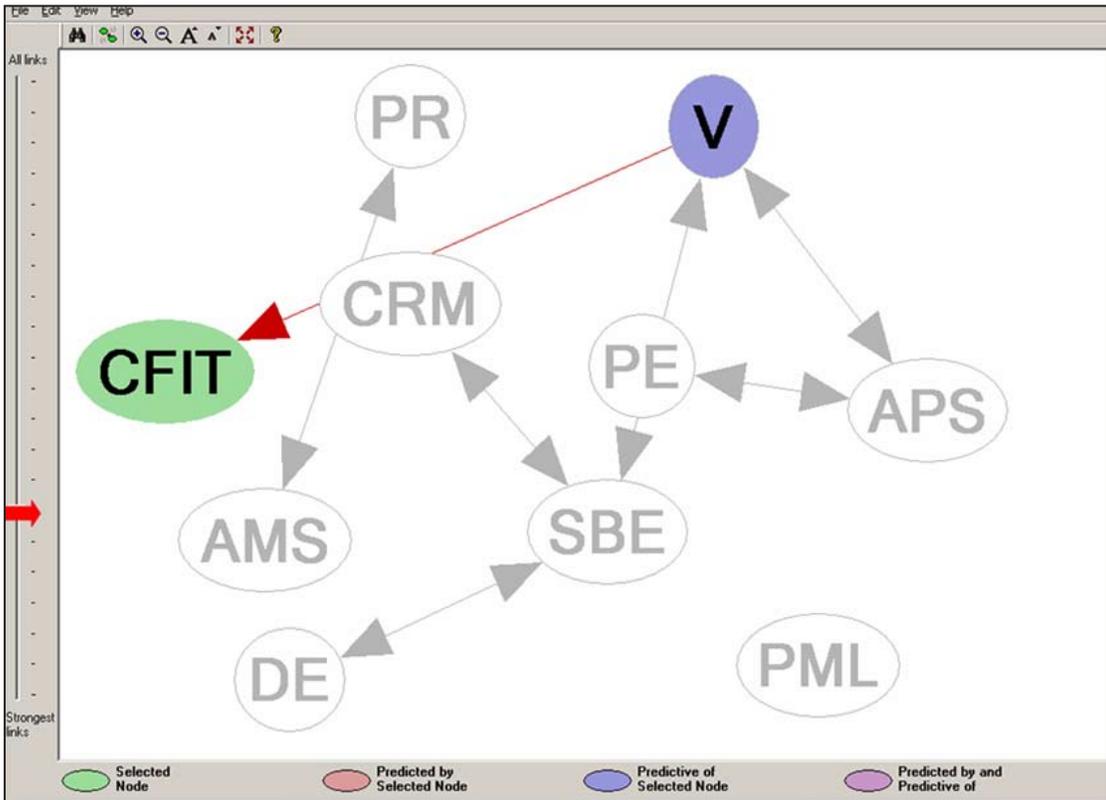

Figure 3. Example of output when the slider bar is approximately 75% from the top position.

(formed from the Bayesian conditional probabilities) and based on the specific conditions that led to a particular point) is also available by double clicking on the variable name within the decision tree. Double clicking on the variable name will display a dialog box that lists the pre-conditions—the *given(s)* taken into consideration in the probability. The dialog box may indicate that the program is providing probabilities (likelihoods) for purchasing a computer (the selected output variable) or in our dataset, the likelihood of a CFIT *if* certain other variables were coded in a specific way—meaning that certain other conditions co-exist. This may be more clearly understood within the scenario of the likelihood of rolling a *4* on a die after having rolled a *1*, a *2*, and a *3* (in that order).

For example, the dialog box in Figure 4 shows the specifics for a CFIT in the sample of accident cases used for this test. The number of cases in the data set analyzed was 16, 278; the probability that an accident case was missing data for the CFIT variable was 0.000123. The probability that the coders did not identify a CFIT in an accident case was 0.909. The probability that the

coders identified a CFIT in an accident was 0.0905. This type of information can be used to make inferences about the data sample. For example, most accident cases apparently did not include a CFIT. That fits the percentages reported by Shappell and Wiegmann (2003); their report also compares the "percentage of CFIT and non-CFIT accidents associated with at least one instance of each particular causal category" (p.12). All of that is good information. But, one of the benefits of WinMine is that it allows the researcher to see other associated probabilities. For example, Figure 4 indicates that, of the accidents that did not involve a violation (n=14,431), 9,728 did involve a skill-based error and 154 did involve both personal readiness and skill-based errors. Although this same information is obtainable using a series of "select if" statements in SPSS, this is a level of analysis that the other techniques rarely explore. The point being that the researcher can delve into the data without spending hours conjuring up many "what if" scenarios because these are readily available in the many other output screens within WinMine by double-clicking on a selected variable node.
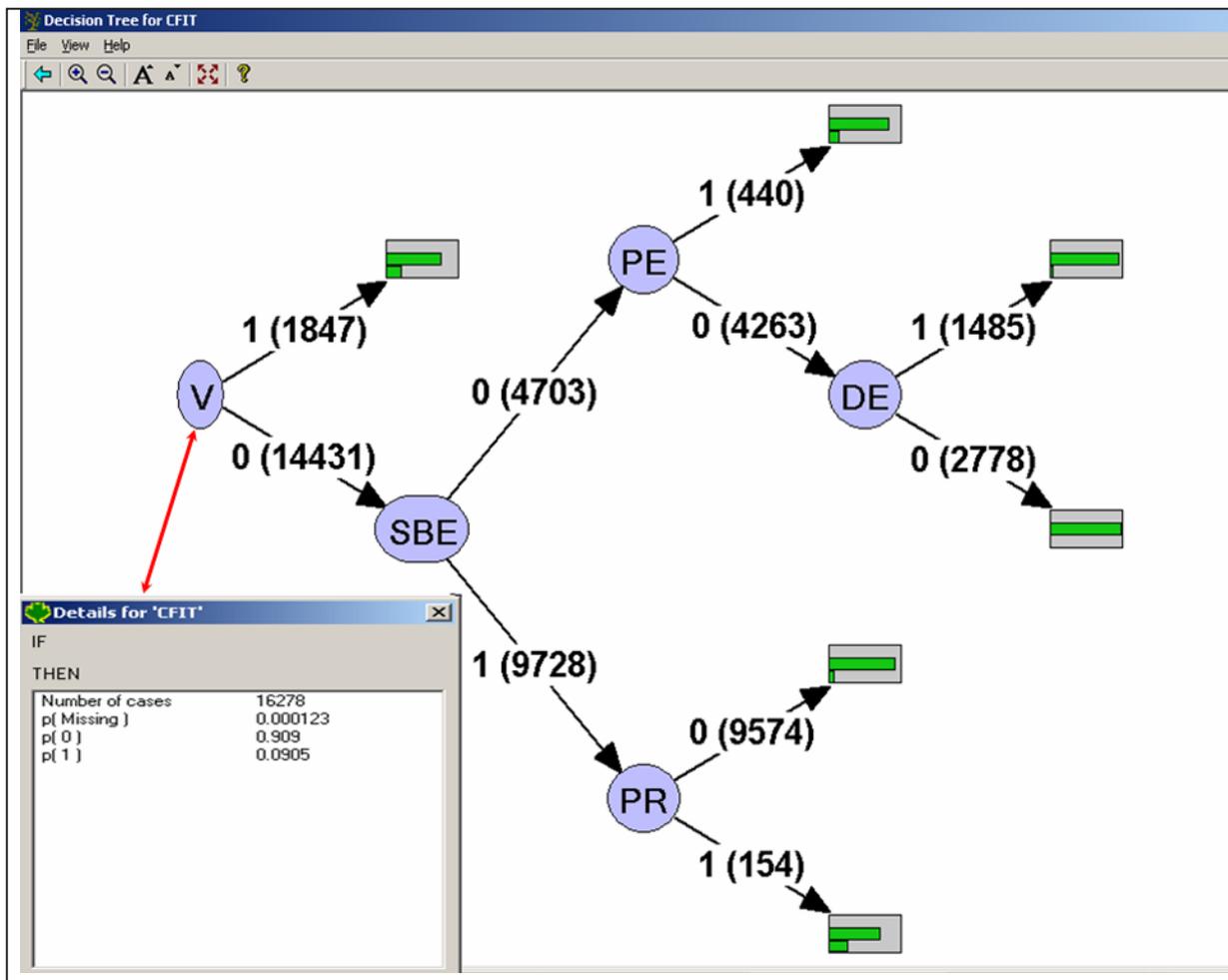

Figure 4. Example of a decision tree and the dialog box with specifics for a CFIT.

DISCUSSION

Each method of data analysis has strengths and weaknesses. The challenge for the researcher is to use an appropriate analysis to get the most meaningful information from the data. For example, if the researcher wanted to get an initial look at the frequencies within each category, the bar chart is a perfect tool, but if a more in-depth examination of the data is desired, similar to the one previously described, then conventional bar charts, correlations, pie charts, and line graphs cannot convey that level of information.

Analysis using WinMine reveals interrelationships between variables in ways that other methods cannot do in a direct manner—to accomplish the same task using other software, the researcher must select cases that fit specifically, pre-defined criteria, analyze that sub-set of data, and continue to select sub-sets of cases until a specific chain of "select if" statements leads to a desired path. However, that type of analysis assumes that the direct path is known in advance; otherwise, given 9 variables with 2 possible states for each (1=yes, and 0=no), the exercise would be especially time consuming if all possible permutations were tested.

Caution

Interpreting relational data should be done with care. When interpreting the graphical output, we can assume that, if there is an arrow from one variable to another, the previous variable is a predictor of the latter. Figure 1 suggests that PR,V, PE, SBE, and DE are predictors of a CFIT, that is consistent with Shappell and Wiegmann (2003, p. 12). On Figure 3, the arrow from V (Violation) to CFIT indicates that a violation predicts a CFIT. When interpreting uni-directional links between variables, one might conclude that each variable helps to predict the other. However, Heckerman et al. (2000) state that the relationships are "significant only in whatever sense was used to learn the network with finite data" (p. 67). One must cautiously interpret the results because they are only appropriate in the context of the other variables also present in the network and, thus, only within the context of those included in the analysis.

Challenges

Two things make the WinMine program difficult to use. One is related to the user's level of expertise using DOS to do data manipulations. It is clumsy to switch back and forth between the user-friendly dialog boxes of a Windows environment to DOS-level interactions. Also, the software has a somewhat limited amount of documentation available for users. A search of the research literature did not find any documented use of the software that might have facilitated the interpretation of some of the features, options, output, and possible applications.

WinMine requires some investment of time to learn; however, there is a WinMine users group available via E-mail that is very helpful if problems are encountered or questions arise. Frequently, the lead programmer personally answers those inquiries.

Strengths

The decision tree is the greatest advantage of the WinMine program, because it can be used to quantify the likelihood of an event (for a specific dataset), given that other circumstances exist, such as quantifying the HFACS hierarchical structure using coded accident cases. The same analytical method used by WinMine can also be obtained using traditional analytic software but would require many more steps. For example, to use SPSS to quantify the likelihood of an event, you must first determine the path or chain of conditions that you want to test. Suppose that a computer storeowner, needing some information to plan a marketing campaign, wanted to determine the likelihood that someone who purchased a computer from the store more than 2 years ago still lives within a 25-mile radius from the store and has DSL service. The owner must write a series of "select if" statements (e.g., select if DSL = yes and 2+years = yes and miles_25 = yes) to match the predetermined filtering restrictions. Then, from that reduced data set, the owner can determine the percentage of purchasers that fit that specific description. For every set of conditions to explore, a separate analysis must be run. That is, you must start over to define the analysis to know the proportion of purchasers who do not have DSL but satisfy the other set of conditions. In contrast, WinMine would calculate the probabilities given that the variable DSL was coded *yes*; and, on a separate branch of the tree, given that the variable DSL was coded *no*. After searching for dependencies between the variables, each branch (e.g., each HFACS causal factor in our dataset) is placed within the hierarchy; and its structure is based on calculable, verifiable conditional probabilities found within the data rather than on a subjective (however well-informed) structure. Therefore, rather than performing analyses that simply answer questions such as how often each type of error occurs, the data can be exhibited as probabilities of an event, given a certain set of events.

WinMine allows the user to construct a graphical model either as a dependency network or a Bayesian network. When comparing the two approaches, Heckerman et al. (2000) admitted that "a dependency network is not useful for encoding causal relationships.... Nonetheless, there are straightforward and computationally efficient algorithms

for learning both the structure and probabilities of a dependency network from data; and the learned model is quite useful for encoding and displaying predictive (i.e., dependence and independence) relationships. In addition, dependency networks are well suited to the task of predicting preferences—a task often referred to as collaborative filtering—and are generally useful for probabilistic inference, the task of answering probabilistic queries" (p. 49). Furthermore Heckerman et al. discussed the possibility that a causal interpretation of the graph may be suspect if "one uses a computationally efficient learning procedure that excludes the possibility of hidden variables" (p. 50). In these situations, the graphed relationships should be considered predictive or correlational and cannot be interpreted as causal.

The WinMine Toolkit is made up of several distinct units (tools), each with a specific function, much like tools within a carpenter's toolkit. Not all of the tools were discussed in this paper because they were not used in the present analysis; however, the toolkit includes a feature that allows the user to link (match) and then merge two separate data files that are keyed on a common variable, such as an identification number. Another feature allows the user to predefine an imposing structure on the data (using the Plan Phase) to determine whether the variables within the dataset adhere to a specifically defined structure. For example, one use of this feature would be to examine the HFACS hierarchical structure proposed by Shappell and Wiegmann (2003) to determine the fit of that structure to the data. The next step would be to contrast the hierarchical-structured model with a model in which all variables (except CFIT) were defined as "input-output variables" (as was the case in this evaluation, thus allowing the data to define the model) and subsequently compare the fit of the models using a metric similar to those used in advanced statistics such as structural equation modeling products like LISREL, AMOS, and CALIS. Furthermore, a follow-on study should evaluate whether the HFACS hierarchical structure also appears in other datasets, thereby providing evidence of the robustness of the HFACS-proposed taxonomic structure. Consequently, such evidence could be used to further develop future taxonomies.

The WinMine Toolkit is useful for visualizing relationships between variables, such as the 9 HFACS categories examined in this study. In summary, this evaluation found that the WinMine tool is a good visual aid to exploring, interpreting, and perhaps even making predictions from or finding structure within dichotomously-coded categorical data.

REFERENCES

Chickering, D. (2002). The WinMine Toolkit. Redmond, WA. Microsoft technical report no. MSR-TR-2002-103.

Chickering, D.M., Heckerman, D., Meek, C., Platt, J.C., & Thiesson, B. (2000). Goal-oriented clustering. Technical report, MSR-TR-2000-82, Microsoft, Redmond, WA.

Heckerman, D., Chickering, D., Meek, D. Rounthwaite, R. & Kadie, C. (2000). Dependency networks for inference, collaborative filtering, and data visualization. Journal of Machine Learning Res 1:49-75.

Review of Aviation Accidents Involving Weather Turbulence in the United States: 1992-2001 (2004) NASDAC National Aviation Safety Data Analysis Center, FAA Office of System Safety, reference no. 04:551.

Review of Aviation Accidents Occurring in the State of Alaska Office of System Safety (ASY, 2003). NASDAC National Aviation Safety Data Analysis Center, Federal Aviation Administration. April 2003.

Shappell, S. & Wiegmann, D. (2000). The human factors analysis and classification system (HFACS). Washington, DC: Department of Transportation/Federal Aviation Administration; FAA report no. DOT/FAA/AM-00/7.

Shappell, S. & Wiegmann, D. (2001). Applying reason: The human factors analysis and classification system (HFACS). *Human Factors and Aerospace Safety,* 1(1), 59-86.