Supporting decision making in IT security – An information-driven approach

NIKOLA VLAHOVIC
Faculty of Economics and Business
University of Zagreb
Trg J. F. Kennedyja 6, Zagreb,
CROATIA

Abstract: - As threats for information systems (IS) from information environments (i.e. Internet) become more frequent additional investment in providing security of IS is required. In order to make adequate decision on appropriate commercial solution for protection of the core IS, management executives should first decide on the minimum criteria for evaluation. In this paper we propose an information-driven approach to determining the relevancy of different features of commercial solutions that suppliers of Internet security make available to managers. These are the criteria that should be evaluated in order to make the decision that suites the requirements of a particular organization. The approach is based on the modified ID3 algorithm implemented in the DoctuS expert system shell. Results of the developed knowledge model on selecting relevant criteria for leading commercially available solutions is presented and discussed.

Key-Words: - IT management, IT Security, Decision-making, Decision support systems, Expert systems

1 Introduction

Investing in IT has always been one of most controversial issues for management executives. Often these investments are treated as costs due to the inability to adequately predict and measure their benefits on company performance. Despite contradicting results of recent surveys about the influence of investment in IT [1][5][6], information systems (IS) of companies have become highly dependant on information and communication technology (ICT) [10]. This is why security issues have become the primal concern of the information officers and managing executives.

For the purpose of maintaining security of the organization company leaders enact security policies. These policies contain directives and mechanisms for various types of security issues. One of the most important security issues in e-business are program intrusions by programmes specially developed for this purpose – *malware*. The *malware* problem has spread so rapidly that it has created an industry to counter it and try to keep businesses clean of infection – the anti-virus industry. The anti-virus solutions market contains a great number of products that address some or all *malware*-related security issues in a variety of degrees and success. This is the result of a fierce competition between brands for larger market shares, but it also makes the decision about choosing a brand more complicated for company executives. Most of the relevant features of these *anti-malware* solutions are qualitative and thus hardly comparable between different brands. Other, quantitative features such as ratios and percentages may be misleading if compared across brands due to the differences in calculation methods used.

In order to avoid these pitfalls for decision makers and to support decision-making process we propose an information-driven approach to investment in IT security as well as to general ICT investments. The information-driven approach eliminates the need to convert qualitative sets of properties into quantitative measures of performance. In this way no additional and false information is incorporated in the decision making process allowing for better modelling of the problem at hand thus enabling reaching a better decision.

The structure of the rest of the paper is as follows: In Section 2 current approaches to decision making about Enterprise IS and IT security will be presented. Advantages and disadvantages will be discusses and we will propose our model that can enhance the quality of

this process. In Section 3 we will present the modified ID3 algorithm for inductive logical programming that is the basis of the proposed information-driven approach. This algorithm is implemented using DoctuS Expert System Shell which will also be described. In the following section a knowledge base developed using DoctuS Expert System Shell will be described. In Section 5 results will be presented, followed by a Conclusion and final remarks.

2 IS Investment and IT Security

According to the third annual information security survey conducted by Information Week and Ernst & Young, nearly half of more than 1,290 respondents representing information systems chiefs and security managers suffered security-related financial losses in the past two years [7]. Most companies hesitate to develop a structured and detailed disaster recovery plan (DRP) or security policies until a security incident arises. According to another survey [8], only 85 per cent of the Fortune 1,000 companies had disaster recovery plans a decade ago.

This is why Hawkins et al [3] suggest a framework of disaster recovery planning (DRP) consisting of three functional areas: management, information technology and human resources. Here the selection of anti-virus software is an imperative for protecting the data contained in company LAN as well as regular and timely sweeps in order to ensure system integrity at all times.

There is a number of specific parameters that need to be taken into consideration by the managing executives when making decisions about improving or maintaining the security of the IS. Costs that should be considered, are not only monetary amounts of implementation and support but also costs of functionality of the overall information system after the security solution is implemented. On the other hand, the estimation of the adequacy of the 'out-of-the-box' solutions is prone to error because it is highly dependent of the current organization of the enterprise and subjective.

Decision making about adequate security software should be supported by IT in order not only to allow unbiased comparison of different solutions but also to enable the managing executives to get a clearer picture of the criteria and their relevance on the final outcome. This is why information-driven approach seems most appropriate.

3 ID3 algorithm and DoctuS

As we have shown, the existing forms of evaluation of security solutions, especially *ex ante* methods of predictive and forecasting models, are not entirely appropriate. Hochstrasser and Griffiths [4] argue than many evaluation techniques are inadequate, particularly those which are based purely on numeric measures.

Information-driven approach to decision support systems is based on the qualitative features of the decision criteria instead on interpolation of quantitative values that may not adequately describe properties of suggested solutions.

Decision tree learning is an example of this kind of approach. Some of the most successful applications of decision tree learning use ID3 algorithm. ID3 algorithm [9] is an attribute-based induction learning system. It is an iterative algorithm used to construct decision trees based on a training set of example cases. For each node in the decision tree attribute with the highest information gain is chosen to split the tree. The algorithm prefers simpler decision trees by eliminating attributes without positive information gain in accordance with the principle known as the Occam's razor. The original ID3 algorithm is a heuristic, though, because of unordered set of attributes used for decision tree construction. The derived decision tree represents a model of knowledge contained in starting examples and therefore can be used to construct production rules.

In the case of decision making, each attribute A of a possible solution $S(j)$ can have a number of values v.

$$S(j) = \{A_i(j): \{v_1(i), v_2(i), \ldots v_m(i)\}, i=1,2,\ldots n\}$$

These values v have different level of desirability function $D(x)$ so that they form an ordered list:

$$A_i=\{[v_1, v_2, \ldots v_n], \quad \forall v_n: D(v_{n-1}) < D(v_n)\}.$$

If these values are known for a set of proposed solutions we can calculate information gain for each attribute used, using information entropy for each attribute

$$E(A_i) = - \Sigma f(v_n(i)) \log[f(v_n(i))]$$

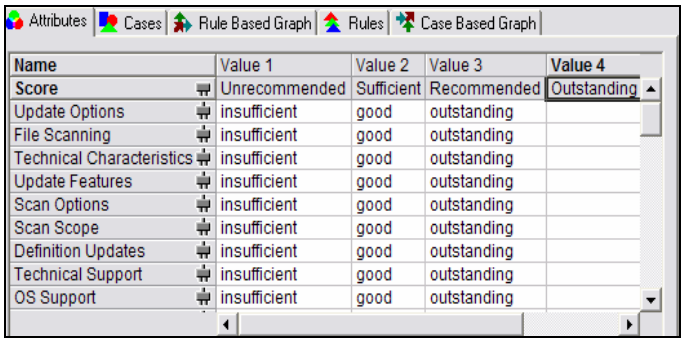
Information gain is measured by the expected reduction of entropy $E(i)$ for the given set of attributes A_i and their values $v_n(i)$.

Due to desirability function $D(x)$ which is intuitively included in the model as intrinsic knowledge (decision makers only estimate comparative relations between values for each set of attribute) heuristics of the original method is modified to fit the preferences of the decision makers.

DoctuS Expert System Shell implements the described modified ID3 algorithm and allows for analysis of the informativity of the decision attributes in accordance to decision makers preferences [2]. It enables consideration of a great number of attributes that may affect the final outcome of decision making process. Proposed solutions may include complex relations between the attributes originating from a number of knowledge domains i.e. interdisciplinary problems. This is why DoctuS can be used to enhance understanding of the problem between experts in different domains of knowledge and provide unbiased insight in the decision making process.

4 Knowledge base model

For the purpose of supporting the decision on choice of available anti-malware software solutions, a knowledge base model was produced in collaboration with the IT security experts, users of the IS and management. The produced model was derived according to the technical specifications for 10 leading anti-malware software solutions available. These specifications yielded 53 different decision attributes with at least two values. Some of these attributes are shown in Fig. 1. Values of all attributes are order form the most undesirable towards most desirable ones, according to expert knowledge. A special attribute that contains the final outcome of the decision process is called decision attribute and for the purpose of this analysis it is called '*Score*'. It represents the suitability of a particular anti-malware solution. The most unfavourable value is '*unrecommended*', intermediate values are '*sufficient*' and '*recommended*' and the most favourable value is '*outstanding*'. Similarly all of the other attributes are defined in the Attributes section of the DoctuS knowledge base (Fig. 1).



Name	Value 1	Value 2	Value 3	Value 4
Score	Unrecommended	Sufficient	Recommended	Outstanding
Update Options	insufficient	good	outstanding	
File Scanning	insufficient	good	outstanding	
Technical Characteristics	insufficient	good	outstanding	
Update Features	insufficient	good	outstanding	
Scan Options	insufficient	good	outstanding	
Scan Scope	insufficient	good	outstanding	
Definition Updates	insufficient	good	outstanding	
Technical Support	insufficient	good	outstanding	
OS Support	insufficient	good	outstanding	

Figure 1. A segment of Decision Attributes Table with attribute values

After most of the attributes that seemed relevant were introduced into the knowledge model, their interconnections and interdependencies were defined using Rule-based graph (Fig.2).

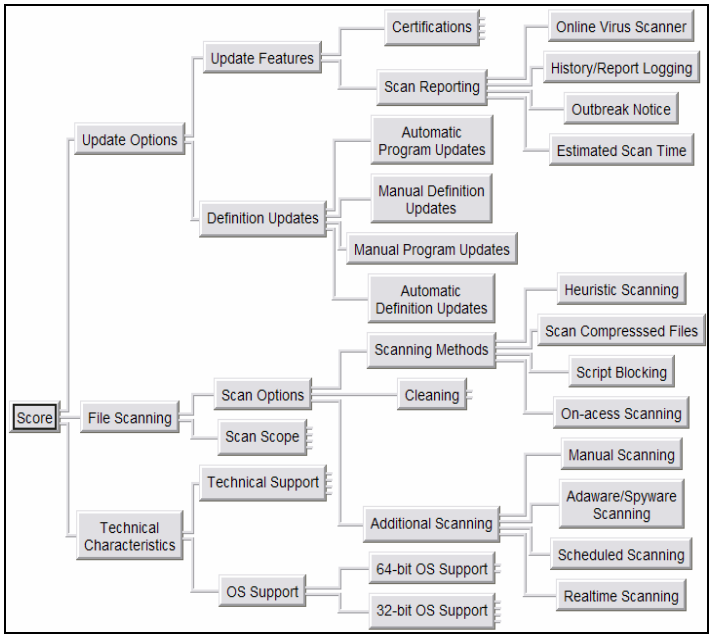


Figure 2. Rule-based graph

This graph is used to define the immediate interdependencies between attributes. On the left-hand side of the Fig. 2. there is the decision attribute '*Score*'. Its value is determined by the values of attributes '*Update Options*', '*File Scanning*' and '*Technical Characteristic*'. For each of these, a separate set of attributes that affect their values was defined on the next level of the rule-based graph. This was an iterative process known as knowledge extraction, and it involved a number of interviews with experts. Also it involved

merging different parts of the tree because each segment of the tree referred to different expert domains. I.e. users of the system were concerned with availability of technical support while IT experts were concerned with channels of definition updates, for example.

During the modelling of the rule-based graph some additional attributes were added to the knowledge base (such as available certifications etc…).

When the rule-based graph was finished all of the dependant attributed required definition of rules. These rules specify in what way is the dependant attribute affected by different combination of values of the bound attributes. All combinations of values of bound attributes can be presented in the form of a grid - 2D view (Fig. 3).

	Update Opti	insufficient	good	outstanding
File Scannin	Technical Ch			
insufficient	insufficient	Unrecommended	Unrecommended	Unrecommended
insufficient	good	Unrecommended	Unrecommended	Unrecommended
insufficient	outstanding	Unrecommended	Unrecommended	Unrecommended
good	insufficient	Unrecommended	Unrecommended	Unrecommended
good	good	Unrecommended	Sufficient	Sufficient
good	outstanding	Unrecommended	Sufficient	Sufficient
outstanding	insufficient	Unrecommended	Unrecommended	Unrecommended
outstanding	good	Unrecommended	Recommended	Recommended
outstanding	outstanding	Unrecommended	Recommended	Outstanding

Figure 3. 2D View of the rules for decision attribute *'Score'*

In Fig. 3 rules for decision attribute *'Score'* are shown. In the headers of rows and columns there are all of the combinations of bound attributes' values. In the centre of the table at every intersection of rows and columns there is a cell with an outcome of the decision for each combination of affecting attributes. For example, the *'Score'* will be *'outstanding'* only if *'Update Options'*, *'File Scanning'* and *'Technical Characteristics'* are also evaluated as '*outstanding'*.

Due to the desirability function D used for defining the order of attribute values all defined rules can be checked for inconsistencies. For each pair of rules R_1 and R_2 values v of bound attributes A_i must be ordered in the same way as the desirability of decision attribute values v:

$$\forall[R_1,R_2]: \{D(v(R_1)) \leq D(v(R_2)): \quad D(v(A_i)) \leq D(v(A_j)), i \neq j, \forall i=1,2,..n, \quad \forall j=1,2,...n\}.$$

After all of the rules are included in the knowledge base and after all of the consistency issues are resolved, the knowledge base is prepared for consultation and testing. For this purpose a training set of example cases is used. These examples represent different IT security solution brands. For each brand input attributes (shown in Fig. 2. as leaf nodes of the rule-based graph) are added in 'Cases' section of the knowledge base.

Decision attributes are derived by the system using knowledge model by activating 'Reason' function. The results are given in Fig. 4.

5 Discussion

All of the used example cases yielded evaluation scores at least *'sufficient'* and none returned *'unrecommended'* value. This is due to the fact that only top 10 commercially available brands were used.

In order to differentiate between the brands a case-based graph can be created from defined example cases. Since case-based graph is induced using ID3 algorithm, only most informative attributes are used to classify given cases. These attributes represent the only significant factors that may differentiate evaluated brands. As we can see from the Case-based graph shown in Fig. 5 the availability of script blocking feature in combination with Windows NT support and ISA 2005 Certification makes a distinction between outstanding solutions and the rest, according to the model of expert knowledge of the consulted specialists.

Another advantage is the capability of the model to provide the whole picture of the decision process without inconsistencies. It can be used to generate explicit explanations why some products are better than the others.

Finally, knowledge base can be used for benchmarking dependant attributes to reveal interdependencies between different decision factors. In this way the knowledge of the experts can be enhanced as well as the understanding of different features and their influence on the decision making process.

Attributes | Cases | Rule Based Graph | Rules of Score | Case Based Graph

	Bit Defende	Kaspersky	F-Secure 2006	PC-cillin 2006	Eset NOD32	McAfee 2007	Norton 2007	AVGPro 7	CA AV 8.	Norman 5
Score	Outstanding	Outstanding	Recommended	Recommended	Outstanding	Recommended	Recommended	Sufficient	Sufficient	Sufficient
ISA 2005	yes	yes	no	yes	yes	yes	yes	yes	yes	yes
WB100% 2005	yes	yes	yes	no	yes	yes	yes	yes	yes	yes
W.C.L Level 1	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
W.C.L Level 2	yes	yes	yes	yes	yes	yes	yes	no	yes	yes
Online Virus Scanner	available	available	not available	available	not available	not available	not available	not availa	available	not availat
Estimated Scan Time	available	available	not available	not available	not available	not available	not available	not availa	not availa	not availat
History/Report Logging	available	available	available	available	available	available	available	not availa	available	available
Outbreak Notice	available	available	available	available	available	available	available	available	available	available
Automatic Definition Update	available	available	available	available	available	available	available	available	available	available
Automatic Program Updates	available	available	available	available	available	available	not available	not availa	available	available
Manual Definition Updates	available	available	available	available	available	available	available	available	available	available
Manual Program Updates	available	available	available	available	available	available	available	not availa	available	available
On-acess Scanning	provided	provided	provided	provided	provided	provided	provided	provided	provided	provided
Realtime Scanning	provided	provided	provided	provided	provided	provided	provided	provided	provided	provided
Scheduled Scanning	provided	provided	provided	provided	provided	provided	provided	provided	provided	provided
Heuristic Scanning	provided	provided	provided	provided	provided	provided	provided	provided	provided	provided
Manual Scanning	provided	provided	provided	provided	provided	provided	provided	provided	provided	provided
Adaware/Spyware Scanning	provided	provided	provided	provided	provided	provided	provided	provided	not provic	provided
Script Blocking	available	available	available	available	available	available	available	not availa	not availa	not availat
Scan Compresssed Files	provided	provided	provided	provided	provided	provided	provided	provided	provided	provided
Auto-Clean Files	available	available	available	available	available	available	available	available	available	available
Quarantine Files	available	available	available	available	available	available	available	available	available	available
Email-POP3 Protection	provided	provided	provided	provided	provided	provided	provided	provided	provided	provided
P2P/File Sharing Protection	provided	provided	not provided	not provided	not provided	not provided	not provided	not provic	not provic	not provide
Registry Startup Protection	provided	not provide	provided	not provided	not provided	not provided	not provided	not provic	not provic	not provide
WebMail Protection	not provide	not provide	not provided	provided	not provided	not provided	not provided	provided	not provic	not provide
Live Chat	yes	no	no	no	no	no	no	no	yes	no
Phone Support	available	available	available	available	not available	available	not available	not availa	not availa	available
Manual/FAQ/Knowlage Base	available	available	available	available	available	available	available	available	available	available
User Form	no	yes	no	no	no	no	no	no	no	yes
Windows XP	supported	supported	supported	supported	supported	supported	supported	supporte	supporte	supported
Windows 2000/2003	supported	supported	supported	supported	supported	supported	supported	supporte	supporte	supported
Windows ME	supported	supported	supported	supported	supported	supported	not supported	supporte	supporte	supported
Windows NT	supported	supported	supported	not supported	supported	not supported	not supported	supporte	not supp	supported
Windows 98	supported	supported	supported	supported	supported	supported	not supported	supporte	supporte	supported
Windows 95	not supporte	not support	not supported	not supported	not supporte	not supported	not supported	supporte	not supp	supported

Figure 4. Cases Table with properties of 10 leading anti-malware software solutions

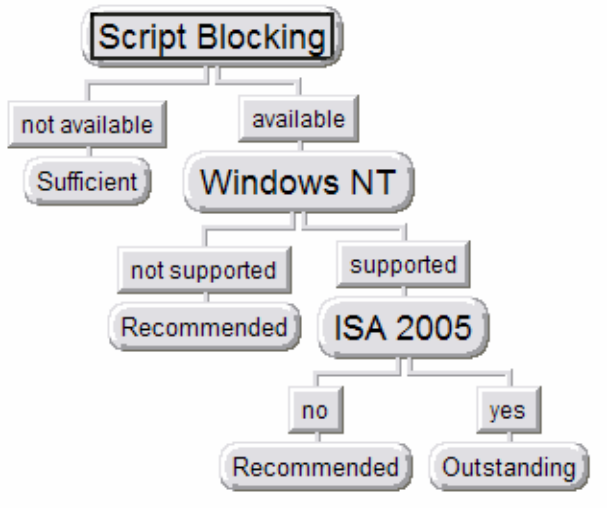


Figure 5. Case-based graph

Benchmarking between example cases using attribute *'OS Support'* revealed that the main concern for providing OS Support is the possibility of supporting Windows NT operating system (Fig. 6).

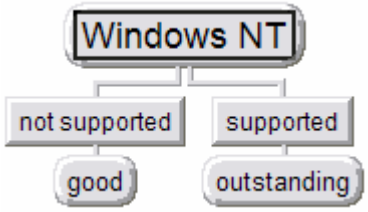


Figure 6. Benchmarking between cases (anti-malware solution brands) by *'OS Support'* attribute

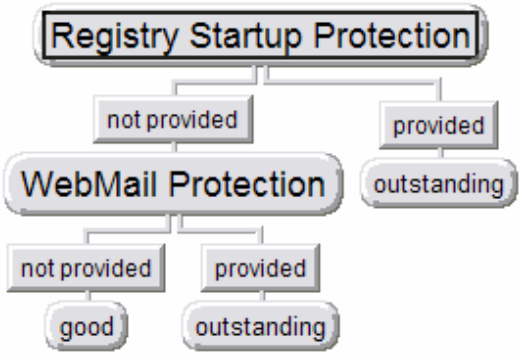


Figure 7. Benchmarking between cases (anti-malware solution brands) by *'Scan Scope'* attribute

Benchmarking between example cases using attribute *'Scan Scope'* revealed that best solutions must provide at least registry startup protection or webmail protection within their file scan options (Fig. 7).

6 Conclusion

Managing executives are faced with decision-making about investment in IT and security of their company's information systems. Results of recent surveys, though, testify questionable influence of these investments on the overall performance of the organisation. The reasons lie in the fact that many evaluation techniques, especially those based on numeric measures, are inadequate.

In this paper we presented an approach that does not use interpolation of numeric measures to describe the problem domain but instead uses information-driven approach. This approach is based on the decision tree model that is induced using modified ID3 algorithm implemented in DoctuS Expert System Shell. The basis of the model was the model of expert knowledge acquired from the experts involved in the decision making process.

The model was trained using actual data about top commercially available security brands. At the current level of sophistication of anti-malware solutions decision tree model revealed the properties that make the most significant diversification between tested brands. It was concluded that these features are: availability of script blocking, support of old OS systems and ISA 2005 certification.

Benchmarking other attributes provided explicit knowledge about different features of anti-malware solutions. First of all the main concern in providing OS support is that the solution provides support for Windows NT operating system. Also best brands provide either registry startup protection or webmail protection within their scanning options.

References:

[1] Andresen J, The unidentified value of IT in construction industry, *Proceedings of the International Conference on Information Technology in Construction*, Hong Kong, pp. 93-105.

[2] Baracskai Z, Velencei J and Dörfler V, In Expert System After Inductive Reasoning Comes Reductive Reasoning, *DoctuS White papers* available at http://doctus.info/white [17-09-2007]

[3] Hawkins SM, Jen DC and Chou DC, Disaster recovery planning: A strategy for data security, *Information Management & Computer Security*, Vol. 8, No. 5, 2000, pp. 222-229.

[4] Hochstrasser B and Griffiths C, *Controlling Information Technology Investment: Strategy and Management*, Chapman & Hall, London, 1991.

[5] Li H, Love, PED and Irani Z, The relationship between the use of IT/IS and the productivity of consulting firms in construction, *International Journal of Construction Information Technology*, 2000.

[6] Love PED and Irani Z, An exploratory study of information technology evaluation and benefits management practices of SMEs in the construction industry, *Information & Management*, Vol. 42, pp. 227-42, 2004.

[7] Panettieri JC, Security, *Information Week*, 27 November 1995, pp. 32-40.

[8] Patrowicz LJ, A River Runs through IT – Disaster Recovery, *CIO Magazine*, April 1, 1998, http://www.internetwright.com/drp/disaster_content.html, [17-09-2007]

[9] Quinlan JR, Induction of Decision Trees, *Machine Learning*, No 1, 1986, pp. 88-106.

[10] Ward J and Peppard J, *Strategic Planning for Information Systems*, Wiley, Chichester, 2002.