

A SECURITY INCIDENT SHARING AND CLASSIFICATION SYSTEM FOR BUILDING TRUST IN CROSS MEDIA ENTERPRISES

Fillia Makedon¹, Song Ye¹, Tilmann Steinberg¹, Yan Zhao¹, Jamies Ford¹, Zhan Xiao¹, Basil Sudborough²

1. *The Dartmouth Experimental Visualization Laboratory (DEVLAB), Department of Computer Science Dartmouth College, 6211 Sudikoff Laboratory, Hanover, NH 03755, USA*

2. *Intralot Systems (www.intralot.gr)*

Abstract: Trust in cross-media applications is essential to successful collaboration. Cross media service delivery encompasses different types of security incidents and assumes a level of trust on the part of the participants of any one transaction. As enterprises and participants of cross media transactions become more susceptible to security risks facilitated by the heterogeneity of data being exchanged, it is important to develop protective infrastructures. Such infrastructures should enable reporting of security violations or misconduct on a regular basis with effortless incident submission, automatic classification of reported incidents, searching and collective knowledge extraction from similar incidents and sharing of information by authorized users. We report on such a system currently being developed. The Security Incident Sharing and Classification system (SISC), collects incidents in a database, through its incident submission interface, and classifies them according to different parameters. We demonstrate an automatic classification scheme based on the level of incident severity, where severe incidents are processed faster. The system builds trust through its monitoring and recommendation capabilities, thus preparing enterprises to encounter new security incidents that may arise. This is an open, customizable, self-standing risk monitoring system which can be built into any enterprise. The recommendation component of SISC extracts solution scenarios from the gathered knowledge of classified incidents and makes them available to SISC users.

Key words: Security Incidents, Severity Level, Incidents Classification,

INTRODUCTION

The explosive growth of the Internet [2, 3] has brought new cross media possibilities, especially in electronic commerce (e-commerce). This growth has also brought new security risks (incidents) that erode the trust of e-business users in cross media enterprises and can result in large-scale economic damage [4, 5, 6].

The future of cross media applications, (e.g., in e-commerce, e-media and e-publishing, etc.) depends on providing adequate security measures that minimize or anticipate security risks (incidents). Incidents may include fraud, denial of service, digital forgery, id theft, document falsification, copyrights infringement, quality of service violations and other small or large incidents of professional misconduct. We report on a protective infrastructure that can give users a better understanding of the risks as well as knowledge about previous incidents to gauge risks faced in a particular transaction [8, 12]. The Security Incident Sharing and Classification System (SISC) can be used by an e-business entity as an internal “whistle-blowing” or internal monitoring system. SISC standardizes incident collection with a uniform reporting scheme, provides automatic classification of incidents, and builds trust and human understanding into the system with a data sharing interface that pools different types of incidents in different entities.

SISC is also useful because security incidents are often distributed over multiple sites of an enterprise, where different aspects of an incident are encountered differently by different users and systems. SISC addresses the need for systematic methods of reporting, analyzing, educating and sharing information among cross media users. A taxonomy of observed incidents facilitates systematic analysis, ensures better understanding and can give a fast / accurate solution to counter the incident.

Our system is based on a common sense approach: When a new incident appears and is not yet fully recognizable and understood, it is critical to characterize this incident and as soon as possible, identify its key properties (e.g., level of severity, cause, domain (area it appears in), the predicted loss (profits or time), and possible solutions, etc.), based on available knowledge from similar incidents. The earlier the key properties of a new incident are identified and studied, the sooner it is possible to find a solution for it.

The current implementation of SISC classifies security incidents based on their severity level. We distinguish severe incidents from medium or trivial ones as early as possible. Severe incidents are immediately identified

by the system so that a timely response is possible. Although we have implemented SISC to classify incidents based on severity levels, it can be extended and customized to other types of measures.

SISC is a kernel component of EcomRISK.org [7, 20], which is a general security resource that contains a wide variety of interactive facilities designed to educate users and to motivate them to submit security incidents. SISC processes the submitted incidents and combines this information with other types of user input to achieve classification. For different types of users (e.g., advanced vs. novices), SISC has different incident submission forms. Users can search the incident database in multiple ways, by different parameters and features, such as “find me an incident like this” or “find me an incident that has these features”. SISC also includes a security solution database [22] that is linked to the submission form. Thus, once a user submits an incident, she is able to link to existing security tools and find out how, why and the popularity of using a particular security tool. Feedback from the usage of these tools is collected in this database and provides additional contextual information. By mining these data, security patterns and higher level correlations can be discovered. Due to space limitation, the details about the solution database are not covered in this paper.

The rest of this paper proceeds as follows. Section 2 provides related work. Section 3 outlines the approach of the incident sharing and classification system and describes the basic elements used to collect and disseminate information. Section 4 details the methodology and results. Section 5 describes the relations between SISC and cross media applications. Section 6 provides concluding remarks and future work.

RELATED WORK

We described a system for incident collection and sharing services that serves as a security resource to the online community. In **Table 1**, we present a short overview of several prestigious sites and contrast their offerings with those of the EcomRISK.org site, on which our incident sharing and classification system is hosted. The following sites are tightly related to EcomRISK.org and provide additional information potentially useful to EcomRISK users: GREeCOM.org, eJETA.org and ISTS.

Classification has been used as an efficient approach to organize and manage information. In this approach, statistical techniques are used to learn a model based on a labeled set of training data, which have been labeled to indicate their categories. This model is then applied to new data items to determine their categories. Several Internet-based information classification systems have been developed, for example, Chakrabarti et al. [17], Chekuri

et al.[18] developed automatic classifiers for subset of pages from Yahoo! Web directory; Chen and Dumais [19] developed a user interface that organizes Web search results into hierarchical categories. Although information classification systems have been used in other areas for several years, applying classification techniques to identify newly appeared security incidents is a relatively new approach.

Table 1. Comparison of EcomRISK.org to major, related web sites. We note that EcomRisk addresses the needs of a low-technical level audience and has many interactive facilities which enable the user to interact with the system and the system to learn from the user.

Web Site	Organization	Audience	Technical Level	Primary Focus	Component
EcomRISK.org	Education	general computer users professionals & experts	low	risk incident database; discussion forum; news; tutorials; tech reports	education
CERT.org	Education	professionals & experts system administrator	medium	risk incident database; security tutorials	technical
SANS.org	Industry	professionals & experts system administrator	high	discussing forum; vulnerabilities database	technical
SERIAS.purdue.edu	Education	general computer users professionals & experts	low	security seminar; post-secondary education	education
neohapsis.com	Industry	Clients	high	vulnerabilities database; articles; archives	consulting
securitysearch.net	Industry	professionals & experts	medium	articles & tutorials, security product reviews; software listings	technical
NIST.org	Government	policy makers federal government agencies	medium	information on different subjects	policy
NISER.org	Government	general computer users professionals & experts	Medium	news & events; services; articles; trainings; incidents submission	consulting

INCIDENT SHARING AND CLASSIFICATION

Incident sharing in SISC is done with (a) a uniform collection interface of actual e-business risk incidents and related solution solutions and (b) an incident database that can be easily browsed, searched, or extended and (c) classification of incidents based on their severity level (or based on other properties). By looking at the existing cases, the user can get an overview of

what risks have been identified so far, what the cost was, where the incident originated, and what tools or methods are available to counter or protect against these incidents.

1.1 Incident Collection

To make submissions of new content to the site as simple as possible, the incident collection system distinguishes between general users and experts and offers two versions of submission forms for each type respectively. An example of technical incident submission form is shown in Figure 1.

The full submission form collects from the user information about:

- the user and his or her relation to the incident;
- the incident and its type and history;
- the solution of the incident, if one was found;
- any consequences of the incident at the affected site.

In the *incident submission form*, a mix of multiple choice questions and free text answers allows for both automated classification and inclusion of case-specific information. This results in a database of common cases as contributed by a broad range of users ranging from experienced software programmers to young entrepreneurs.

After the user successfully submits the incident, feedback is given to the user to provide some additional information for the submitted incident and also as an incentive to attract submissions. For example, the statistical information about the submitted incident is shown, such as how many incidents with similar features are in the incident database already. If the user did not fix the incident yet, information about how to fix it will be provided. Furthermore, users can search for incidents by choosing taxonomy terms from the submission form, or specify terms to search in free text answers.

The *incident database* stores specific cases of e-business incidents. General categories can be derived from these cases, as described in [1], and are used as a foundation for the incident taxonomy, e.g., what caused the problem, when and where it was introduced, etc. The database includes cases of known incidents with previous modus operandi and thus serves as a repository of what is known about previous crimes of a similar nature. Statistical methods and stochastic analysis tools can be used to develop a *prognosis model*. The resulting *incident taxonomy* is customizable by an enterprise by ranking a list of desired parameters, such as, cost, type of industry impacted most, location, timeframe, and other parameters.

Incident Information

Where did the cause of this incident originate?

Internal External

What is the general classification of the incident?

Unauthorized Disclosure of Information

In which domain did the incident happen?

Finance and Investment

How did it enter the system?

Inadvertent – Implicit Sharing of Privileged/Confidential Data

Describe exactly how the cause entered your system, based on the classification above.

When did it enter the system?

During Operation

Describe exactly when the cause entered your system, based on the above classification.

Where did it enter the system?

Software – Operating System – Identification/Authentication

Describe exactly where the cause entered your system, based on the above classification.

How long did the incident effect the system?

72 hours or more

General summary of the incident.

Solution Information

Technically, what did the fix involve?

Applying a patch

Describe the solution, based on your classification above.

List the URLs of any Internet resources used for the solution.

In terms of policy, what did the fix involve?

Re-education on policy

Describe the policy changes, based on your classification above.

Quantifiable Ramifications

How many hours did it take to analyze and solve the problem?

25 hours or less

How much revenue was lost as a result of the incident?

10,000 U.S. Dollars or less

Figure 1. Technical Incident Submission Form

1.2 Incident Classification

We classify incidents into three groups based on their severity levels: ordinary, medium and severe. The reason we choose the severity level as the goal of the classification is two-fold. First, the severity level of an incident is very difficult to tell, especially when the incident is an unfamiliar one to the community. Most incidents are evaluated as severe or ordinary only after they have been solved and their damages have been repaired, which is not time efficient. Second, when an incident is encountered, it is helpful to know how severe it is in order to solve it. Limited resources should be used to solve any severe incidents rather than ordinary ones. This assures that severe incidents can be handled in a timely manner.

Basically, there are some intuitive rules to tell whether an incident is severe or not. For example, if a security hole is found in a widely used application, this can be considered as a severe incident; however, if a similar security hole is found in an application which is used only by very few users, the incident will be judged as an ordinary one. For most of the incidents happening, we can not find appropriate intuitive rules to tell its severity level. Furthermore, the rules themselves will always change over time, and an incident might have different severity levels for different user groups. Thus, we can not specify the method by which the correct severity levels can be computed from the input incidents and this task can not be solved by a traditional programming approach. We will use a classifier to analyze the pattern of incidents with different severity levels.

We use the linearly separable Support Vector Machine (SVM) classifier [15] to do the classification. Here we briefly introduce SVM and two classes of incidents that will be considered. Denote the tuple (x_i, y_i) , $i = 1, \dots, N$ as exemplars from a training set of incidents with two different security levels: low and high. The column vector x_i denotes the properties related to the severity level of incidents, $y_i = +1$ denotes incidents with high severity level, and $y_i = -1$ denotes incidents with low severity level. The linearly separable SVM classifier corresponds to a decision hyperplane that separates the two different classes of incidents. Incidents which lie on the hyperplane satisfy the following constraint:

$$w^t x_i + b = 0,$$

where w is normal to the hyperplane, $|b|/\|w\|$ is the perpendicular distance from the origin to the hyperplane, and $\|\cdot\|$ denotes the Euclidean norm. We define the *margin for any given hyperplane* to be *the sum of the distances from the hyperplane to the nearest positive and negative incidents*. The separating hyperplane is chosen so as to maximize the margin, which is any positive distance from the decision hyperplane. If a hyperplane exists that separates all the data then, within a scale factor:

$$w^t x_i + b \geq +1, \text{ if } y_i = +1$$

$$w^t x_i + b \leq -1, \text{ if } y_i = -1$$

These pair of constraints can be combined into a single set of inequalities which must be satisfied:

$$(w^t x_i + b) y_i \geq 1$$

For any given hyperplane that satisfies this constraint, the margin is $2/\|w\|$. The SVM classifier tries to separate two classes by finding the hyperplane that has the maximum margin to either class (minimize $\|w\|^2$ subject to the constraints in the above equation). The expectation is that the larger the margin, the better generalization of the classifier.

SVM classifiers have been receiving increasing attention and used successfully in many classification areas. Experiments [16] have shown that SVM classifiers outperform other popular classifiers. SVM classifiers can achieve relatively good performance with limited amount of training. SVM also provides an effective mechanism for nonlinear classification via a nonlinear mapping defined by kernel functions, which are used to project data to a higher dimension space if they are not linearly separable.

EXPERIMENTAL METHODOLOGY & RESULTS

1.3 Methodology

We first design and implement a *synthetic incident generator*, which can generate thousands of synthetic incidents according to some specific pattern in several minutes. These synthetic incidents provide a useful testbed for our incident classification research. Our synthetic incident generator can be customized to generate incidents, which have multiple properties in different distribution patterns.

We employ a publicly available SVM tool to implement our classifier. The tool we use is OSU SVM Classifier Matlab Toolbox version 3.00 [14], the core part of which is based on LIBSVM [13]. A linear SVM classifier (referred to as LinearSVC) with its default setting is applied in our experiments. Due to space constraints, we refer the readers to [13, 14, 16] for more technical and implementation details.

With the help of a synthetic incident generator, we obtain a total of 3,000 incidents to train and evaluate the SVM incident classifier. For all incidents, their severities are previously labeled by several security professionals. Based on *selected features*, an incident will be classified into one of three different classes: *ordinary*, *medium*, or *severe*, labeled 1, 2 and 3

respectively. To efficiently *train and evaluate the classifier*, all the incidents are randomly grouped into 3 groups, G1, G2 and G3, to ensure that each group has incidents labeled by different professionals.

1.4 Results

We use two of the three incident groups to train the SVM classifier and use the rest one to do the evaluation. The results are shown in **Table 2**.

Table 2. Performance of SVM Classifier:

	Classifier C _a	Classifier C _b	Classifier C _c
Training Groups	G1, G2	G1, G3	G2, G3
Evaluation Group	G3	G2	G1
Correct Prediction Rate	82.3%	82.5%	84.5%

The performance of our classifier is reliable, since using different combinations of training groups and evaluation group, we can achieve similar correct prediction rate, from 82.3% to 84.5%.

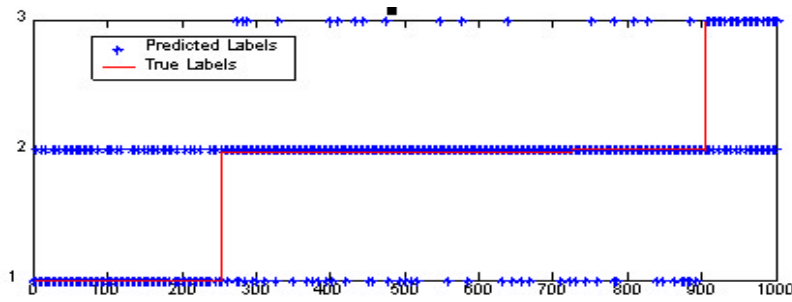


Figure 2. Prediction accuracy of Classifier Cc

Table 3. Prediction accuracy rate of the SVM Classifier

True \ Predict	Classifier Ca						Classifier Cb						Classifier Cc					
	Label 1		Label 2		Label 3		Label 1		Label 2		Label 3		Label 1		Label 2		Label 3	
		%		%		%		%		%		%		%		%		%
Label 1	216	75.8	46	7.3	0	0.0	202	76.8	42	6.5	0	0.0	185	73.4	49	7.5	0	0.0
Label 2	69	24.2	570	89.9	44	54.3	61	23.2	597	91.7	40	46.5	67	26.6	587	89.9	42	44.2
Label 3	0	0.0	18	2.8	37	45.7	0	0.0	12	1.8	46	53.5	0	0.0	17	2.6	53	55.8

Figure 2 and **Table 3** show the prediction results generated by a classifier when it is used to predict the labels for the evaluation incident group. In **Figure 2**, the incidents are sorted by actual labels: ranges 1-252 denote label 1, 253-905 denote label 2, and 906-1000 denote label 3. The asterisk for each incident shows the predicted label. Note that most incidents are classified correctly and incorrectly classified incidents are classified as the neighboring classes only. From the results, we believe our classifier is reliable to tell the severity level of an incident.

Although the SVM classifier can achieve reasonable accuracy rate when it is applied on the evaluate group, it can only correctly predict about 50% of the incidents which are actually severe ones. We are current working on this problem and several approaches have been applied to improve the accuracy rate for incidents which are actually severe. Due to the space limitation, we will not put the details of the approaches and the results in this paper.

APPLICABILITY: REPORTING VIOLATIONS

The incident classification system described in this paper is relevant to many cross-media applications, such as e-business systems. Below is a list of examples that apply.

- (1) The system can collect incidents which are also *alerts*, e.g.,
 - a. for denial of service attacks based on high-bandwidth media, e.g. multiple frivolous requests for large movie objects;
 - b. for obvious (and not-so-obvious) violations of copyrights, e.g. a user needs to get a pass from site X before they can access a media object on site Y, but Y gets requests for media objects using the same pass from multiple clients at different sites — meaning that the pass from site X has been shared among some users.
- (2) The system can be used to:
 - a. report copies of copyrighted objects on other sites, or report incidents of copying objects for non-personal use;
 - b. report instances where access to one object can lead to (unintended) access to another object, e.g. changing a URL of a preview of an image to retrieve the full-sized image, or escalating user privileges through security holes in order to access materials reserved for paying users;
 - c. investigate a claim of a content provider (e.g. a photographer) that a client (e.g. an online newspaper) has used a resource outside the previous agreement between the two (e.g. used a photo with an article in the daily edition, but

also in the Sunday supplement with another article when this was not agreed on before).

The *solutions database component* [22] can be extended similarly to include tools to prevent or detect the above problems, e.g. by offering watermarking tools for object identification.

CONCLUSION & FUTURE WORK

“Trust is the inverse of Risk” [21]. To build trust, it is important to incorporate the human element of perceiving different types of security risks. This paper describes SISC, a system which builds trust by collecting human input and providing a sharable incident collection and classification system. This system can be used not only publicly by security resource centers, but also internally by commercial entities to record, track and classify security incidents within their enterprise. SISC’s automatic classification of incidents is based on severity levels and this enables fast-response.

Since users may also input incidents description in the form of free text, we want to be able to extract valuable information that correlates to the fields of the submission form. An important improvement for the incident sharing and classification system is to make full use of free text that users input. A free text-based incident classification system is currently under development in our lab to provide a more useful and reliable classification of security incidents. Nonlinear SVM classifiers will be used to achieve higher accuracy in text-based classifications. A *Global Incident Collector System* is also under development to collect incidents from the Internet and connect them or correlate them to those in the SISC database. Finally, a *Recommender System* will connect to the submission process.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Li Shen’s help in building and testing the SVM classifier. This work has been supported by the Department of Justice contract 2000-DT-CX-K001.

References

- [2] Landwehr C.E., Bull A.R., McDermott J.P., Choi W.S., A Taxonomy of Computer Program Security Flaws. *ACM Computing Survey* 1994; 26:3, 211-254
- [3] Collins J.C., Lazier W.C. *Beyond Entrepreneurship: Turning your Business into an Enduring Great Company*. Prentice Hall, Eaglewood Cliffs, 1992, pp 95-134

- [4] Makedon, F. E-Commerce Security Resource: ECOMRISK Data Center. Grant Report, 2000
- [5] Linqvist U., Kaijser P., Jonsson E. The Remedy Dimension of Vulnerability Analysis. 21st National Information Sys Security Conf., 1998, pp 91-98
- [6] Bishop M., Bailey B. A Critical Analysis of Vulnerability Taxonomies. TR CSE-96-11, Dept. of Computer Science, University of California at Davis, 1996
- [7] Sahay A., Gould J., Barwise P. New Interactive Media: Experts' Perceptions of Opportunities and Threats for Existing Businesses. *European Journal of Marketing*, 1998, Vol. 32, No. 7/8, pp. 616-628
- [8] Marcuss A. EcomRISK.org The Source for E-Commerce Risk News & Assessment. TR2001-403, Dept. of Computer Science, Dartmouth College, 2001
- [9] U.S. Department of Commerce. The Emerging Digital Economy: Introduction. 1998 <http://www.ecommerce.gov/emerging.htm>
- [10] Hoffman D.L., Novak T.P., Chatterjee P. Commercial Scenarios for the Web: Opportunities and Challenges. *Journal of Computer Mediated Communications*, 1995, December, Vol. 1, No. 3
- [11] Institute for Security Technologies Studies at Dartmouth College, <http://www.ists.dartmouth.edu>
- [12] Vatis M.A., Cyberterrorism: The State of U.S. Preparedness. Statement before the House Committee on Governmental Reform, Wednesday, September 26, 2001.
- [13] Ashcroft, J., Remarks of Attorney General John Ashcroft, First Annual Computer Privacy, Policy & Security Institute, May 22, 2001.
- [14] Chang, C. and Lin, C., LIBSVM — A Library for Support Vector Machines <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [15] Ma, J., Zhao, Y. and Ahalt, S., OSU SVM Classifier Matlab Toolbox (ver 3.00) http://eewwww.eng.ohio-state.edu/~maj/osu_svm/
- [16] Burges, C.J.C. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2:121-167, 1998.
- [17] Cristianini, N. and Shwe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, 2000
- [18] Chakrabarti, S., Dom, B., Agrawal, R. and Raghavan, P. Scalable feature selection, classification and signature generation for organizing large text databases into hierarchical topic taxonomies. *The VLDB Journal* 7, 1998, 163-178
- [19] Chekuri, C., Goldwasser, M., Raghavan, P. and Upfal, E. Web Search using automated classification. In *Sixth International World Wide Web Conference*, Santa Clara, California, Apr. 1997, Poster POS725
- [20] Chen, H. and Dumais, S. Bringing order to the web: Automatically categorizing searching results. CHI2000, The Hague, Amsterdam
- [21] Makedon, F., Heckman, C., et al. EcomRISK.org: An E-Commerce Security Resource, Proceedings of the 4th International Workshop on Computer Science and Information Technologies, Patras, Greece, 2002
- [22] Kemp, J., Trust and Risk in Internet Commerce: Design for Trust, Dartmouth College Invited Presentation, 1/17/03.
- [23] Xiao, Z., A Consultation System for Cyber-Security, Master Thesis, Department of Computer Science, Dartmouth College, March, 2003