Terrorist Detection System

Yuval Elovici[1], Abraham Kandel[2,3], Mark Last[1], Bracha Shapira[1], Omer Zaafrany[1], Moti Schneider[4], and Menahem Friedman[1,5]

[1] Department of Information Systems Engineering, Ben-Gurion University
Beer-Sheva 84105, Israel
Phone: +972-8-6461397, Fax: +972-8-6477527
{mlast,bshapira,zaafrany}@bgumail.bgu.ac.il,
elovici@inter.net.il

[2] Department of Computer Science and Engineering, University of South Florida,
4202 E. Fowler Ave. ENB 118, Tampa, FL, 33620, USA

[3] Currently at the Faculty of Engineering, Tel-Aviv University, Israel
kandel@csee.usf.edu

[4] School of Computer Science, Netanya Academic College, Netanya, Israel
motis@netanya.ac.il

[5] Department of Physics, Nuclear Research Center – Negev

Abstract. Terrorist Detection System (TDS) is aimed at detecting suspicious users on the Internet by the content of information they access. TDS consists of two main modules: a training module activated in batch mode, and an on-line detection module. The training module is provided with web pages that include terror related content and learns the typical interests of terrorists by applying data mining algorithms to the training data. The detection module performs real-time monitoring on users' traffic and analyzes the content of the pages they access. An alarm is issued upon detection of a user whose content of accessed pages is "too" similar to typical terrorist content. TDS feasibility was tested in a network environment. Its detection rate was better than the rate of a state of the art Intrusion Detection System based on anomaly detection.

1 Introduction

The Internet is an efficient communication infrastructure that is increasingly used by terrorist organizations to safely communicate with their affiliates, coordinate action plans, spread propaganda messages, raise funds, and introduce new supporters into their networks [1]. Governments and intelligence agencies are calling to invest major efforts in development of new methods and technologies for identifying terrorist activities on the web in order to prevent future acts of terror. TDS presents an example to such an effort.

By means of content monitoring and analysis of web pages accessed by a group of web users, it is possible to infer their typical areas of interest [2, 3, 4]. It is also possible to identify users that access specific, potentially illegitimate information on the internet [2, 3, 4]. Using this approach, real time web traffic monitoring may be per-

J.-F. Boulicaut et al. (Eds.): PKDD 2004, LNAI 3202, pp. 540–542, 2004.

formed to identify terrorists as they access typical terror-related content on the internet. Terror Detection System (TDS) described in [2, 3, 4] implemented this approach.

2 Terrorist Detection System (TDS): Overview

TDS is a content-based detection system recently developed to detect users who are interested in terror-related pages on the Web by monitoring their online activities. The reader is referred to [4] for a detailed description of TDS. The system is based on real-time monitoring of internet traffic of a defined group of Web users (e.g. students in a specific University campus). The group is suspected to include hidden individual terrorists and the system aims at detecting them. The current version of TDS refers only to the textual content of the accessed web pages. It consists of two main modules: *a training module* activated in batch, and a real-time *detection module*.

The *training module* receives as input a set of web pages that include terror related content. It applies cluster analysis on the textual representation of each page resulting with a set of vector that efficiently represents typical terrorists' areas of interest.

The *detection module* performs on-line monitoring of all traffic between the users being monitored and the Web. The content of the pages they access is analyzed, transformed to a form of a vector, and added to the vector representing the user profile. The profile for each user is kept during a period of time and number of transactions defined by operative system parameters. Similarity is measured between each user profile and the typical terrorist areas of interests. A consistent high similarity between a specific user and terror-related content would raise an alert about that user. Each user related to the monitored group is identified by a "user's computer" having a unique IP address. In case of a real-time alarm, the detected IP can be used to locate the suspicious computer and hopefully the suspected user who may still be logged on to the same computer. In some intranet environments or cooperative ISPs and according to legal privacy issues users may be identified by their user names to enable fast location upon an alert.

The detection module, being activated in real-time, is required to efficiently capture the textual content of Web pages from the Internet traffic. Actually, the detection efficiency is crucial to TDS effectiveness; skipped pages or inaccurate analysis of pages due to slow handling of traffic might result in unreliable detection. TDS detection rate was compared to performance of ADMIT, a state of the art Intrusion Detection System [5], and obtained higher results as described in [2].

3 TDS Significance

An important contribution of TDS lies in the unique environment of its application. The detection component is planned to run in a real-time wide-area network environment, and should be capable of on-line monitoring of many users. Therefore, a crucial design requirement was high-performance which called for enhancement of the algorithms involved, especially the mining algorithm, to high-performance and scalability. We believe that the adjustment of the algorithm to high-performance

might serve many other applications. In addition, TDS is an example for a successful application of data mining and machine learning techniques to international cyber-war effort against world-wide terror.

Acknowledgement

This work is partially supported by the National Institute for Systems Test and Productivity at University of South Florida under the USA Space and Naval Warfare Systems Command Grant No. N00039-01-1-2248 and by the Fulbright Foundation that has granted Prof. Kandel the Fulbright Research Award at Tel-Aviv University, Faculty of Engineering during the academic year 2003-2004.

References

1. Birnhack M. D. and Elkin-Koren, N.: Fighting Terror On-Line: The Legal Ramifications of September 11. Internal Report, The Law and Technology Center, Haifa University. [http://law.haifa.ac.il/faculty/lec_papers/terror_info.pdf]. (2002)
2. Elovici,Y., Shapira, B., Last, M., Kandell, A., and Zaafrany, O.: Using Data Mining Techniques for Detecting Terror-Related Activities on the Web. Journal of Information Warfare, Vol. 3, No.1, (2003) 17-28
3. Shapira, B., Elovici, Y., Last, M., Zaafrany, O., and Kandel, A.: Using Data Mining for Detecting Terror-Related Activities on the Web. European Conference on Information Warfare and Security (ECIW), (2003) 271-280
4. Last, M., Elovici, Y., Shapira, B., Zaafrany, O., and Kandel, A.: Content-Based Methodology for Anomaly Detection on the Web. Advances in Web Intelligence, E. Menasalvas et al. (Editors), Springer-Verlag, Lecture Notes in Artificial Intelligence, Vol. 2663, (2003) 113 – 123
5. Sequeira, K., Zaki, M.: ADMIT: Anomaly-based Data Mining for Intrusions. Proceedings of SOGKDD 02, ACM. (2002) 386-395