WI'09/IAT'09 Joint Invited Talks

WI-IAT 2009

Data Mining for Malicious Code Detection and Security Applications

Bhavani Thuraisingham
The University of Texas at Dallas, U.S.A.

Abstract

Data mining is the process of posing queries and extracting patterns, often previously unknown from large quantities of data using pattern matching or other reasoning techniques. Data mining has many applications in security including for national security as well as for cyber security. The threats to national security include attacking buildings, destroying critical infrastructures such as power grids and telecommunication systems. Data mining techniques are being investigated to find out who the suspicious people are and who is capable of carrying out terrorist activities. Cyber security is involved with protecting the computer and network systems against corruption due to Trojan horses, worms and viruses. Data mining is also being applied to provide solutions such as intrusion detection and auditing. The first part of the presentation will discuss my joint research with Prof. Latifur Khan and our students at the University of Texas at Dallas on data mining for cyber security applications For example; anomaly detection techniques could be used to detect unusual patterns and behaviors. Link analysis may be used to trace the viruses to the perpetrators. Classification may be used to group various cyber attacks and then use the profiles to detect an attack when it occurs. Prediction may be used to determine potential future attacks depending in a way on information learnt about terrorists through email and phone conversations. Data mining is also being applied for intrusion detection and auditing. Other applications include data mining for malicious code detection such as worm detection and managing firewall policies. This second part of the presentation will discuss the various types of threats to national security and describe data mining techniques for handling such threats. Threats include non real-time threats and real-time threats. We need to understand the types of threats and also gather good data to carry out mining and obtain useful results. The challenge is to reduce false positives and false negatives. The third part of the presentation will discuss some of the research challenges. We need some form of real-time data mining, that is, the results have to be generated in real-time, we also need to build models in real-time for real-time intrusion detection. Data mining is also being applied for credit card fraud detection and biometrics related applications. While some progress has been made on topics such as stream data mining, there is still a lot of work to be done here. Another challenge is to mine multimedia data including surveillance video. Finally, we need to maintain the privacy of individuals. Much research has been carried out on privacy preserving data mining. In summary, the presentation will provide an overview of data mining, the various types of threats and then discuss the applications of data mining for malicious code detection, cyber security and national security. Then we will discuss the consequences to privacy.

Biography

Bhavani Thuraisingham joined The University of Texas at Dallas in October 2004 as a Professor of Computer Science and Director of the Cyber Security Research Center in the Erik Jonsson School of Engineering and Computer Science. She is an elected Fellow of three professional organizations: the IEEE (Institute for Electrical and Electronics Engineers), the AAAS (American Association for the Advancement of Science) and the BCS (British Computer Society) for her work in data security. She received the IEEE Computer Society's prestigious 1997 Technical Achievement Award for "outstanding and innovative contributions to secure data management". Dr Thuraisingham's work in information

security and information management has resulted in over 80 journal articles, over 200 refereed conference papers and workshops, and three US patents. She is the author of nine books in data management, data mining and data security including one on data mining for counter-terrorism and another on Database and Applications Security and is completing her tenth book on Secure Service Oriented Information Systems. She has given over 60 keynote presentations at various technical conferences and has also given invited talks at the White House Office of Science and Technology Policy and at the United Nations on Data Mining for counter-terrorism. She serves (or has served) on editorial boards of leading research and industry journals and was the Editor in Chief of Computer Standards and Interfaces Journal. She is also an Instructor at AFCEA's (Armed Forces Communications and Electronics Association) Professional Development Center and has served on panels for the Air Force Scientific Advisory Board and the National Academy of Sciences. Dr Thuraisingham is the Founding President of "Bhavani Security Consulting" - a company providing services in consulting and training in Cyber Security and Information Technology Prior to joining UTD, Thuraisingham was an IPA (Intergovernmental Personnel Act) at the National Science Foundation from the MITRE Corporation. At NSF she established the Data and Applications Security Program and co-founded the Cyber Trust theme and was involved in inter-agency activities in data mining for counter-terrorism. She has been at MITRE since January 1989 and has worked in MITRE's Information Security Center and was later a department head in Data and Information Management as well as Chief Scientist in Data Management. She has served as an expert consultant in information security and data management to the Department of Defense, the Department of Treasury and the Intelligence Community for over 10 years. Thuraisingham's industry experience includes six years of research and development at Control Data Corporation and Honeywell Inc. Thuraisingham was educated in the United Kingdom both at the University of Bristol and at the University of Wales.