

Data Mining for Intrusion Detection: Techniques, Applications and Systems *

Jian Pei Shambhu J. Upadhyaya Faisal Farooq Venugopal Govindaraju
Department of Computer Science and Engineering
State University of New York at Buffalo
Emails: {jianpei, shambhu, ffarooq2, govind}@cse.buffalo.edu

An intrusion is defined as any set of actions that compromise the integrity, confidentiality or availability of a resource. Intrusion detection is an important task for information infrastructure security. One major challenge in intrusion detection is that we have to identify the camouflaged intrusions from a huge amount of normal communication activities. Data mining is to identify valid, novel, potentially useful, and ultimately understandable patterns in massive data. It is demanding to apply data mining techniques to detect various intrusions.

In the last several years, some exciting and important advances have been made in intrusion detection using data mining techniques. Research results have been published and some prototype systems have been established. Inspired by the huge demands from applications, the interactions and collaborations between the communities of security and data mining have been boosted substantially.

This seminar will present an interdisciplinary survey of data mining techniques for intrusion detection so that the researchers from computer security and data mining communities can share the experiences and learn from each other. Some data mining based intrusion detection systems will also be reviewed briefly. Moreover, research challenges and problems will be discussed so that future collaborations may be stimulated. For data mining/database researchers and practitioners, the seminar will provide background knowledge and opportunities for applying data mining techniques to intrusion detection and computer security. For computer security researchers and practitioners, it provides knowledge on how data mining can benefit and enhance computer security. We will try to understand and appreciate the following technical issues.

- What is intrusion detection? Why is it challenging and why data mining techniques can really help?
- What are the major data mining techniques available for intrusion detection?

*This research is partially supported by NSF Grant IIS-0308001. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of NSF.

- Successful applications of data mining techniques in intrusion detection and the experiences.

Speakers

Jian Pei is an Assistant Professor of Computer Science and Engineering at State University of New York at Buffalo. He received his Ph.D degree in Computing Science from Simon Fraser University, Canada. His research interests include data mining, data warehousing, OLAP, database systems, bioinformatics and their applications. His research is supported in part by the National Science Foundation.

Shambhu J. Upadhyaya is an Associate Professor of Computer Science and Engineering at the State University of New York at Buffalo. His research interests are information assurance, computer security, fault diagnosis, fault tolerant computing, and VLSI Testing. He is the director of the Center of Academic Excellence in Information Assurance Education at Buffalo, accredited by the National Security Agency. His research on computer security has been funded by AFOSR, AFRL, DARPA, NSA and Telcordia Technologies. He is an Associate Editor of IEEE Transactions on Computers and is a senior member of IEEE.

Faisal Farooq received the B.Eng. in Computer Science from National Institute of Technology, Bhopal, India in 2001. He is currently working toward his M.S. degree in Computer Science at State University of New York at Buffalo. His research interests include information retrieval, databases, data mining and computer security.

Venugopal Govindaraju is a professor of Computer Science and Engineering at State University of New York at Buffalo, and Associate Director of CEDAR, the Center of Excellence for Document Analysis and Recognition at his university. He received his Ph.D degree in Computer Science at the University at Buffalo in 1992. His research is focused on Human Computer Interaction, Pattern Recognition, and Biometrics.