Chapter 7 Forecasting Software Vulnerabilities Using Time-Series Techniques

Baidyanath Biswas IIM Lucknow, India

Dileep Kumar G. Adama Science and Technology University, Ethiopia

ABSTRACT

This chapter discusses the concepts of time-series applications and forecasting in the context of information systems security. The primary objective in such formulation is the training of the models followed by efficient prediction. Although economic and financial forecasting problems extensively use time-series, predicting software vulnerabilities is a novel idea. The chapter also provides appropriate guidelines for the implementation and adaptation of univariate time-series for information security. To achieve this, the authors focus on the following techniques: autoregressive (AR), moving average (MA), autoregressive integrated moving average (ARIMA), and exponential smoothing. The analysis considers a unique data set consisting of the publicly exposed software vulnerabilities, available from the U.S. Dept. of Homeland Security. The problem is presented first, followed by a general framework to identify the problem, estimate the best-fit parameters of that model, and conclude with an illustrative example from the above dataset to familiarize readers with the business problem.

DOI: 10.4018/978-1-5225-3534-8.ch007

Copyright © 2019, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

1. LEARNING OBJECTIVES

- This chapter will guide the readers to learn four advanced forecasting techniques Autoregressive (AR), Moving Average (MA), Autoregressive Integrated Moving Average (ARIMA), and Exponential Smoothing.
- The academic context of this study is set in the background of cybersecurity and software vulnerability disclosures. It makes the chapter a unique topic for researchers.
- The chapter will aid students to solve forecasting problems dealing with various Information Security issues virus attacks, malware, and data breaches.
- Throughout the chapter, we extensively use R to model the forecasting problems for our analysis. R is a freeware, open-source language and commonly used to analyze statistical research problems.
- In a real business scenario, the organization needs to minimize the IT risk of future software breaches and exploits. A proactive way of doing this involves the selection and use of software with common functionality but lowest vulnerability count among a set of available players in the market. Timeseries forecasting methods proposed in this chapter will help academicians as well as top management in organizations to achieve this goal.

2. INTRODUCTION AND MOTIVATION

Security vulnerabilities are weaknesses that are resident in each software through which attackers can intrude into the IT systems and cause severe damage to the users. Often hackers exploit publicly disclosed vulnerabilities by targeting information systems, which lead to severe financial losses in those organizations (Telang & Wattal, 2007). The Common Vulnerabilities and Exposures (CVE) Project jointly with the US National Vulnerability Database (NVD) acts as a ready repository for publicly reported software vulnerabilities. While there were only 900 publicly reported vulnerabilities till 1999, a drastic increase in their disclosure each year has led to more than 82,000 CVEs reported in the NVD till date¹. The year 2017 is witnessing a massive rise in the number of vulnerability exploits with maximum exposure count of more than 5400 till May, compared to 6400 for the entire year of 2015 and 2016. Figure 1 reports the approximate annual number of security vulnerabilities till date.

39 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: <u>www.igi-</u> <u>global.com/chapter/forecasting-software-vulnerabilities-</u> <u>using-time-series-techniques/207383</u>

Related Content

Predicting NFL Point Spreads via Machine Learning

Daniel M. Brandon (2024). *International Journal of Data Analytics (pp. 1-18)*. www.irma-international.org/article/predicting-nfl-point-spreads-via-machine-learning/342851

Improvisation of Cleaning Process on Tweets for Opinion Mining

Arpita Grover, Pardeep Kumarand Kanwal Garg (2020). *International Journal of Big Data and Analytics in Healthcare (pp. 49-59).*

www.irma-international.org/article/improvisation-of-cleaning-process-on-tweets-for-opinionmining/253845

Cold Calling as a Strategic Tool to Improve the Recruitment and Selection Processes

Shivani Agarwaland Priyansh Teotia (2021). Using Strategy Analytics to Measure Corporate Performance and Business Value Creation (pp. 1-15). www.irma-international.org/chapter/cold-calling-as-a-strategic-tool-to-improve-the-recruitmentand-selection-processes/285843

Using Data Science Software to Address Health Disparities

Jose O. Huerta, Gayle L. Prybutokand Victor R. Prybutok (2021). *International Journal of Big Data and Analytics in Healthcare (pp. 45-58).* www.irma-international.org/article/using-data-science-software-to-address-health-disparities/277647

Learning Analytics: Application on Curriculum Design and Assessment

Melquiades R. Hayag Jr. (2018). *Impact of Learning Analytics on Curriculum Design and Student Performance (pp. 113-124).* www.irma-international.org/chapter/learning-analytics/198666