Comparative Analysis of Cybersecurity Metrics to Develop New Hypotheses

D. Fisher, S. Madnick, N. Choucri, X. Li, and J. Ferwerda, Massachusetts Institute of Technology

Explorations in Cyber International Relations
Massachusetts Institute of Technology Harvard University

Workshop on
People, Power, and CyberPolitics
MIT, December 7 and 8, 2011

Abstract

Few Internet security organizations provide **comprehensive, detailed, and reliable quantitative metrics,** especially in the international perspective **across multiple countries, multiple years, and multiple categories.** Organizations ask why they should spend valuable time and resources collecting and standardizing data. This report aims to provide an encouraging answer to this question by demonstrating the value that even limited metrics can provide in a comparative perspective.

We present some findings generated through the use of the Explorations in Cyber Internet Relations (ECIR) Data Dashboard. In essence, this dashboard consists of a simple graphing and analysis tool, coupled with a database consisting of data from disparate national-level cyber data sources provided by governments, Computer Emergency Response Teams (CERTs), and international organizations. Users of the dashboard can select relevant security variables, compare various countries, and scale information as needed.

In this paper, we present **an example of observations concerning the fight against cybercrime, along with several hypotheses attempting to explain the findings.**

We believe that these preliminary results suggest valuable ways in which such data could be used and we hope this research will help provide the incentives for organizations to increase the quality and quantity of standardized quantitative data available.

The Data Dashboard

The Data Dashboard consists of a simple graphing and analysis tool, coupled with a database consisting of data from disparate national-level cyber data sources provided by governments, Computer Emergency Response Teams (CERTs), and international organizations. The dashboard was developed to provide historical trend data and news to policymakers, academics, IT professionals and other stakeholders. By consulting the Dashboard, the user can compare trends in various categories in national-level cybersecurity threats and vulnerabilities among several countries and/or regions over time. The Dashboard provides data in five diverse categories:

- **Demographic Data:** Population, GDP (2000 US $), Electric Power Consumption, Software Piracy Losses, Tertiary School Enrollment
- **IT Attributes:** # Personal Computers, # Users with Internet Access, # Secure Internet Servers, # Hosts
- **Political Attributes:** Political Stability Index, Government Effectiveness Index, Rule of Law Index, Polity Index, Militarization Index
- **Threat Attributes:** Total CERT Reported Incidents, Virus/Worm/Malicious Code/Malware, Phishing/Personal Data Abuse, Scanning, DoS & Integrity Attacks
- **Cyber Crime Attributes:** Total Cyber Crime Cases, Cyber Crime Damage Dollar Loss, Cyber Crime Arrests, % Cyber Crimes Reported to Police

These data can be manipulated by the user through an online interface available to the general public, and can be manipulated to scale data sets by other attributes.

Case Study: Software Piracy Losses

The data available through the Data Dashboard system are particularly useful in cross-country metric comparisons. In Figure 1 below, the software piracy losses of seven countries are compared; at first sight, it appears that China and the United States have the highest piracy rates.

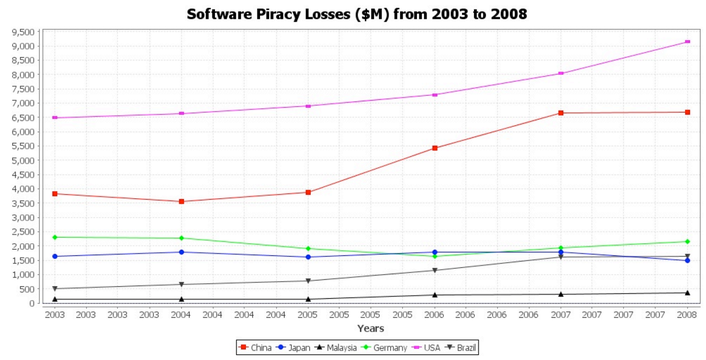

Figure 1: Software piracy losses of seven countries from 2003 to 2008.

The Data Dashboard also **allows users to scale existing data by different metrics such as number of personal computers, number of users with Internet access, or other relevant data sets.** In the case of software piracy, a more relevant data set may scale piracy losses by number of Internet users, as can be seen in Figure 2 below.

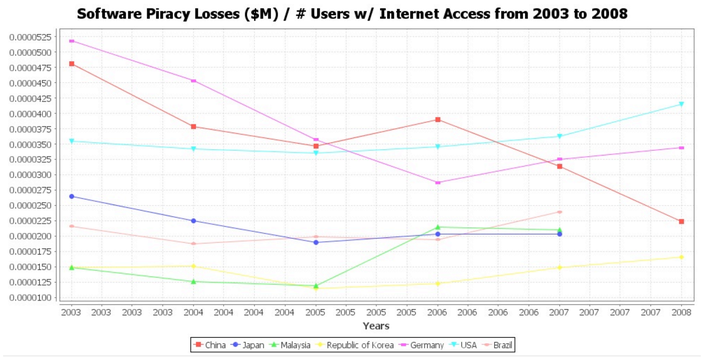

Figure 2: Piracy losses scaled by number of Internet users, 2003 to 2008.

We might expect that different nations would have different levels of enforcement vis-à-vis piracy that may affect the volume of illegal activity. One candidate method to measure this variation is the Rule of Law index, compiled by the World Bank . **Figure 3 makes use of the Data dashboard by utilizing three diverse types of information: software piracy losses, numbers of Internet users, and Rule of Law.** Nations with higher Rule of Law should theoretically display lower rates of cyber crime versus countries with less developed legal systems.

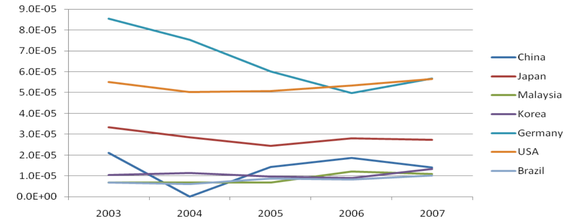

Figure 3: Software piracy losses adjusted for Rule of Law per Internet user.

These results suggest that the absolute volume of software piracy has risen over the past decade, with the bulk of activity in a few nations such as the United States and China. The trends affecting piracy rates are more ambiguous, however, with different countries exhibiting different levels and trends. Surprisingly, a country's rate of piracy does not seem to be conditional on whether it is a developed or a less developed country. In almost all of the metrics examined here, Korea and Japan have outperformed, while the United States and Germany have exhibited less positive results. Although isolating the national events that lead to these different trends is difficult, this analysis suggests fruitful avenues for further research should data remain consistently available.

Next Steps

This analysis has sought to accomplish two distinct tasks. First, although Internet security data is scant and scattered across the Internet, we were nevertheless able to identify and illustrate several interesting trends from existing data that shed light on how different nations and organizations are coping with the global rise of cyber crime. Second, our results highlight several inconsistencies and ambiguities in data provided by different organizations, and we believe that relying on these published statistics without the benefit of a rigorous comparative analysis could lead to fundamentally misleading conclusions regarding the international efforts to improve cybersecurity and fight against cybercrime.

Although the data challenges may be interpreted as an argument against using such data sources, our research implies that the dilemmas could be ameliorated through careful analysis and by promoting a minimum level of standardization across Internet security statistics and the contexts within which they are collected. It is also clear through our analysis that an increase in data availability – be it from governments, CERTs, or international organizations – could yield positive effects in both supporting existing scholarship and the expansion into new fields of knowledge related to cyber crime. In spite of the identified shortcomings of the data, it is possible to learn and identify interesting trends using the Data Dashboard.

In the future, as part of the ECIR project, more data sources will be identified and incorporated to increase the effectiveness of the Dashboard project, so that further interesting international cybersecurity relations can be investigated.

This work is funded by the Office of Naval Research under award number N000140910597. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author (s) and do not necessarily reflect the views of the Office of Naval Research.