Running Head: Sharing Cybersecurity Data

Implications of an interconnected cybersecurity landscape

A SURVEY OF LITERATURE

JACOB BLAZINA | BLAZINAJ@CWU.EDU | MAY 25, 2021 | CS 325

2021

Contents

[Abstract 2](#_Toc73690156)

[Keywords Cybersecurity, Information Security, Incident Response, Privacy, Information Sharing, Risk 2](#_Toc73690157)

[Introduction 3](#_Toc73690158)

[Background on Cybersecurity Information Sharing 3](#_Toc73690159)

[Technical Solutions to Cybersecurity Sharing 5](#_Toc73690160)

[Shortcomings of closed-door Incident Response Investigations 6](#_Toc73690161)

[Quantifying the Risks cybersecurity data sharing 7](#_Toc73690162)

[Privacy Concerns 8](#_Toc73690163)

[Current Events Related to Cybersecurity Data Sharing 9](#_Toc73690164)

[Conclusion 10](#_Toc73690165)

[References 11](#_Toc73690166)

[Appendix 1: Figures used for reliably determining data leakage 12](#_Toc73690167)

# Abstract

Cybersecurity data is inherently sensitive and may contain private information. In the context of security, information must be quickly and thoroughly analyzed for detection and response. Whenever data is shared across organizational boundaries, there is risk that the data may become compromised. Compromised cybersecurity data is doubly sensitive. It contains private information and it also contains information that may allow attackers to find lateral attack vectors to then compromise even more data. Cybersecurity data is shared in ever greater amounts as the international dependence on information systems grow. Starting with threat-information-sharing for anti-malware programs and evolving into multi-source information sharing for Intrusion Detection and Prevention systems, the library of sensitive datatypes that may be compromised is expanding. There are many technological approaches that both facilitate this transmission of data while also securing it, such as the DFAX and CASE frameworks. There are risks inherent in sharing cybersecurity data and there is no obvious way to mitigate the privacy concerns involved in sharing sensitive information. Organizations that do not want to share cybersecurity data are at a disadvantage, and the recent presidential executive order may be the final blow for institutions that want to keep that information contained inside their boundaries. The balance of pros and cons related to sharing sensitive cybersecurity data across organizations continues to evolve at a rapid pace.

## Keywords Cybersecurity, Information Security, Incident Response, Privacy, Information Sharing, Risk

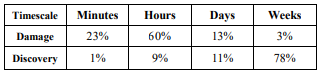
# Introduction

Cybersecurity, also commonly referred to as Information Security, is a concern that grows at the same accelerated rate as the continued global growth of Information Technology (IT) infrastructures. Upwards of 50% of economic growth in the ten years leading up to 2005 was related directly to information system adoption (Bhatia, 2016, p. 57). As more private companies and government institutes lean on IT infrastructure, the importance of cybersecurity grows. In recent times, information systems continue to become ever more interconnected with other systems across organizational boundaries.

This paper addresses the pros and cons of information sharing in general, with a specific look at the newly emerging field of **cybersecurity** information sharing. Cybersecurity information is inherently sensitive, and the illicit use of this data can passively expose and lead to the compromise of existing information systems. The implications of sharing cybersecurity data across organizational boundaries are dramatic. Sharing this data helps organizations detect and respond to incidents, but also introduces new attack vectors that can become incidents in and of themselves.

# Background on Cybersecurity Information Sharing

Threat Information sharing is key to many preventative cybersecurity measures. A case study performed in association with the Arizona Cyber Threat Response Alliance (ACTRA) by Haas, Ahn, and Grimmelmann in 2015 sheds some light on the current landscape of threat information sharing. Companies that provide anti-malware services utilize threat information sharing to rapidly respond to a growing list of cyber vulnerabilities (Haas et al., 2015, p. 2). This case study specifically addresses the benefits of sharing cybersecurity data and attempts to alleviate the concern of sharing this data across organizational boundaries.

One of the main arguments for sharing cybersecurity data relates to the timeline of detection and response. In Table 1, the authors illustrate the importance of rapid detection of an incident. Notice the inverse relationship between discovery time and damage done. This table shows the percentage of damage done within minutes, hours, days, or weeks of an incident as well as the percentage depicting the amount of information that is discovered regarding an incident.

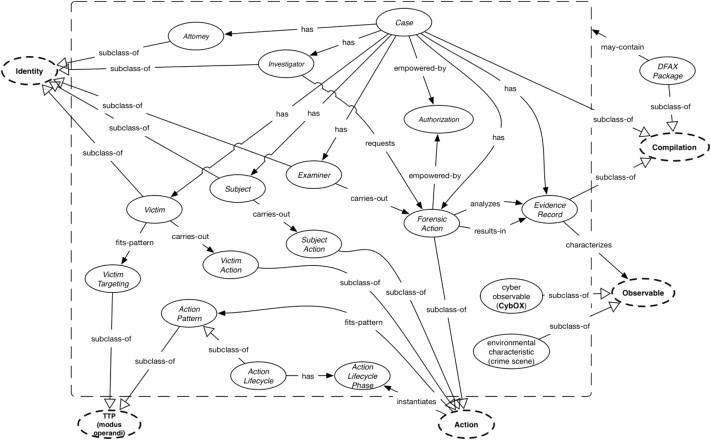
*Table 1: Time scale disparity: compromise vs discovery (Haas et al.)*

Upwards of 83% of the damage done happens within the first minutes or hours of an incident, while around 89% of the discovery does not happen until days or weeks later (Haas, 2015, p. 3). At the point of discovery, it is far too late to mitigate the current incident, and any recovery efforts are only for the benefit of not allowing the attack vector to be used again.

The timescale of incident detection and response can be improved via threat information sharing. In their section on Alerts, Haass et al. describes Intrusion Detection Systems (IDS) that can pick up anomalous data by pattern matching with previously shared incident data from other organizations. By sharing this data across organizational boundaries, an automated system can provide near-instantaneous detection of an incident; which could potentially mitigate a majority of the damage. The researchers continue on to highlight the flip side of this rapid detection: alerting the attackers that their attack has been noticed. When an attacker knows that the organization is aware of the attack, they will quickly go dark to avoid being caught (Haass, 2015, p. 3). Because of this, the best designed detection measures quietly alert the victims. This way victims can gather important data about the attacker before and during their response.

# Technical Solutions to Cybersecurity Sharing

There are emerging technologies that provide functionality for sharing cybersecurity data across organizational boundaries. Casey, Back, and Barnum in 2015 discuss one prevalent technology to foster the sharing of cybersecurity data. The authors describe a programming language called Cyber Observable eXpression (CybOX). This language is an open source effort for standardizing digital data that is fostered by the US Department of Homeland Security (Casey, 2015). As depicted in figure 1, CybOX (circled in red) is one small piece of a framework called Digital Forensic Analysis eXpression (DFAX).



*Figure 1: DFAX Architecture*

Another article headed by the same author addresses the same framework from the API perspective; specifically, the Cyber-investigation Analysis Standard Expression Portion (CASE) (Casey, 2017). CASE works similarly to DFAX and still utilizes CybOX, but is more flexible and can be used in a more general setting (Casey, 2017). These technologies show that there is a large network of interconnected systems that are designed to facilitate the sharing of information security data. Many of these technologies are open source and designed to grow and evolve with the changing cybersecurity landscape. Because of the interconnected nature of these technologies, organizations that continue to push for in-house response procedures may not be benefitting.

# Shortcomings of closed-door Incident Response Investigations

There are current examples of information technology that do not foster collaborative sharing of cybersecurity data. In a 2011 article, David Bennett addresses one of the common cases of closed-door cyber investigations: getting data off mobile devices. Mobile devices are notoriously hard for investigators to crack and device manufacturers have historically been reticent to provide data to investigators (Bennett, 2021).

Like any legal investigation, even if investigators retrieve data from a device, the data may not be admissible in court if the means of obtaining the data were unsound (Bennett, 2021). This highlights the fact that the decision to share or not share data across organizational boundaries may have a major impact. There is a blurry line between cybersecurity data and personal data as well. A blurry line may be investigators presenting their data request in a cybersecurity context. For example, investigators could be investigating a hacker group connected to a drug cartel, and request device manufacturers to release network logs that could potentially relate to the criminal investigation.

When organizations are averse to sharing incident data, it is important as well to consider the human element of incident investigations. An organization that is trying to fully handle an incident in-house can be subject to internal biases in their methodology that could hamper an investigation (Spring, J & Illari). The security ramifications of a cybersecurity data breach can be very hard to quantify. An incident can be virtually impossible to directly link back to a previous breach; for example, an attacker may have gleaned from a compromised Information Security Policy gap that an organization had in their change management process. The attacker could use that knowledge to exploit a legacy employee account for malicious purposes. In this scenario, it would be virtually impossible for internal investigators to determine how the specific attack vector was constructed. Outside investigators may identify these gaps and find the attack vector much quicker than an internal team.

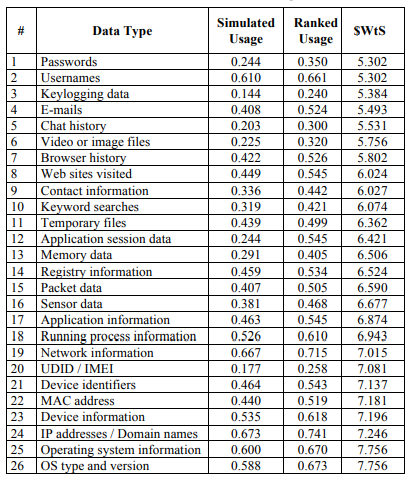
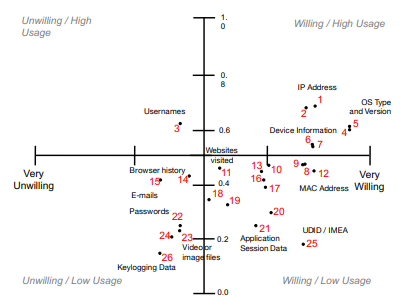
# Quantifying the Risks cybersecurity data sharing

The sharing of data has ushered in a revolution in user experience; allowing technology to quickly add value to the daily lives of many people. Bank account information may now be simultaneously shared between a home lender, credit card issuer, and accounting software, which allows individuals to analyze their financial footprint and quickly apply for loans. Educational information moves between universities, testing centers, and high schools to streamline the process of getting a formal education.

There are examples of information sharing providing immense benefits both for organizations and for consumers. On the flip side, there are examples of information sharing becoming a huge problem. In 2017, the national credit bureau Equifax experienced a data breach that affected over 145 million people (Zou 2018, p. 2). This number is only a rough estimate, and the implications reach much further when considering lateral security issues that stem from an incident of this size. The sole purpose of credit bureaus is specifically to facilitate information sharing to streamline the process of consumer loans and other applications. This information sharing by nature produces substantial risk that would otherwise be avoided without these data streams.

It is hard to accurately predict the risk involved with cybersecurity data streams. A 2016 article written by Riccardo Bortomaleotti, Andreas Peter, Maarten Everts, Willem Jonker, and Pieter Hartel depicts the efforts involved when trying to quantify the effects of an incident. Data is considered to be leaked if an attacker was able to access it (Andreas, p. 484). While the authors present an excellent approach of setting up information systems in such a way that data leakage can be quantified, the very nature of their approach is complicated and unapproachable for most organizations. Their proposed solution involves engineering systems to accurately calculate how much data was leaked in an incident. The authors even address that their approach is difficult to scale and is computationally expensive (Andreas, p. 493), which may prohibit smaller cybersecurity teams from effectively setting up information systems.

# Privacy Concerns

 It is easy to talk about cybersecurity data in the abstract, but one must always remember that cybersecurity data often consists of private information that can offer an unflattering snapshot into individuals’ personal and professional lives. Bhatia, Breaux, Friedberg, Hibshi and Smullen in a 2016 article directly address the privacy concerns involved in sharing cybersecurity data. In figure 2, the researchers define 26 cybersecurity data types that can be considered private data, and map them on a quadrant grid that shows the usage of particular data types in incident investigations mapped against user’s willingness to share that data.

*Table 2: Estimates for Incident Data Usage*

Figure 2: Distribution of Willingness to Share Data

*Figure SEQ Figure \\* ARABIC 2: Distribution of Willingness to Share Data*

Figure 2 illustrates that some data types like usernames get used a lot in incident investigations, yet users generally are unwilling to share that data. Likewise, many of the data types that users are willing to share have a lower usage in investigations (Bhatia 62). The nature of this spread shows that people do not necessarily agree on what kind of information should be shared even in the event of an incident. The proliferation of incident information sharing across organizational boundaries exacerbates this issue.

# Current Events Related to Cybersecurity Data Sharing

The topic of cybersecurity data sharing became even more relevant with President Biden’s 2021 Executive Order which outlines policies regarding the sharing of private cybersecurity data with the Federal government. In light of recent national cybersecurity incidents involving government systems, the President has determined that the best way forward is to expand the reach of federal systems to facilitate the sharing of cybersecurity data between organizations. Organizations must now rapidly respond to government requests for cybersecurity data related incidents.

The benefits of this program are illustrated by Haass et al.: the quicker an organization can respond to an incident, the more damage mitigation that can be done (2015). The risks of this program are: the more data that is being shared across organizational boundaries, the higher the risk of data leakage and privacy compromise. According to the executive order, incident response is a top priority. The executive order does say that this sharing of information must still comply with all current privacy laws, but does not set any guidelines.

# Conclusion

Further case study research around individual and organizational perception on the privacy implications of sharing cybersecurity data would help illuminate the trade-offs between risk mitigation and risk escalation. The federal government getting involved in this debate has legal and moral implications that would benefit from being addressed for constitutionality by the Supreme Court. Cybersecurity information sharing has been a topic since the beginning of the Information System age. Yet in the last decade, information sharing has become ever more important in light of the privacy and risk factors that inherently come with sharing sensitive data across organizational boundaries. There are many pros and cons of sharing sensitive cybersecurity data in the context of incident detection and response. As the information technology landscape continues to evolve, it will be ever more important to keep addressing the risk vs reward of voluntarily (or involuntarily) sharing cybersecurity data.

# References

Andreas, P., Bortolameotti, R., Everts, M., Jonker, W., & Hartel, P. Reliably determining data leakage in the presence of strong attackers. (2016). *ACSAC ’16: Proceedings of the 32nd Annual Conference on Computer Security Applications*

Bhatia, J., Breaux, T., Friedberg, L., Hibshi, H., & Smullen, D. Privacy risk in cybersecurity data sharing. (2016). *WISCS ’16: Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security*

Bennet, D. (2021). The challenges facing computer forensics investigators in obtaining information from mobile devices for use in criminal investigations. *Information Security Journal: A Global Perspective, 21(3)*

Casey, E., Back, G., & Barnum, S. (2015). Leveraging cybox to standardize representation and exchange of digital forensic information. *Digital Investigation: The International Journal of Digital Forensics Incident Response,* 21(S1)

Casey, E., Barnum, S., Griffith, R., Snyder, J., Beek, H., & Neson, A. (2017). Advancing coordinated cyber-investigations and tool interoperability using a community developed specification language. *Digital Investigation: The International Journal of Digital Forensics Incident Response*, 22("C")

Exec. Order No. 14028, 3 C.F.R. (2021). https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/

Haass, J., Ahn, G., & Grimmelmann, F. (2015). Actra: A case study for threat information sharing. *WISCS ’15: Proceedings of the 2nd ACM Workshop on Information Sharing and Collaborative Security*, 21

Meyer, P., Hiesgen, R., Schmidt, T., Nawrocki, M., & Wahlisch, M. Towards distributed threat intelligence in real-time. (2015). *SIGCOMM Posters and Demos ’17: Proceedings of the SIGCOMM Posters and Demos*

Spring, J. & Illari, P. Review of human decision-making during computer security incident analysis. (2019). *Digital Threats: Research and Practice*

Zou, Y. & Schaub, F. Concern but no action: Consumers’ reactions to the equifax data breach. (2018). *CHI EA ’18: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*

# Appendix 1: Figures used for reliably determining data leakage

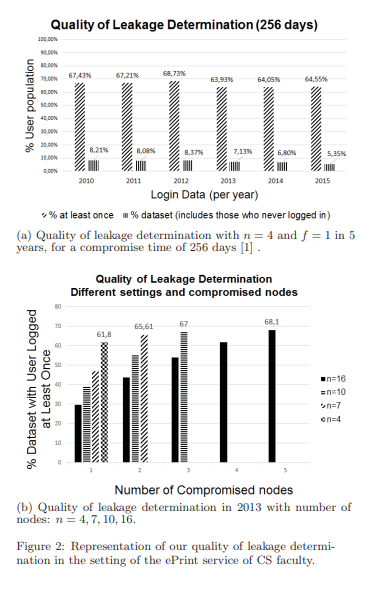
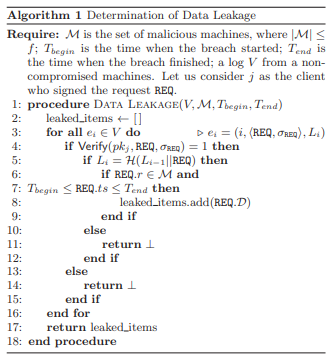
Excerpt from Andreas et al. (2016)

Figure 3: Algorithm Code (Andreas et al. p.490. 2016.)

Figure 4: Quality of Data leakage (Andreas et al. p.492. 2016.)