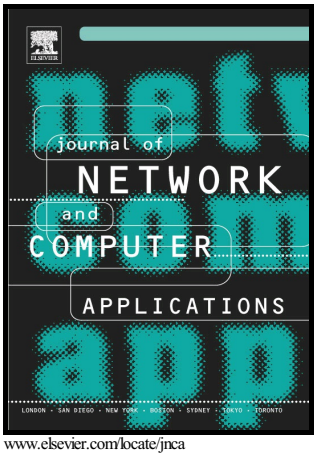
Data Exfiltration: A Review of External Attack Vectors and Countermeasures

Faheem Ullah, Matthew Edwards, Rajiv Ramdhany, Ruzanna Chitchyan, M. Ali Babar, Awais Rashid



www.elsevier.com/locate/jnca

PII: S1084-8045(17)30356-9
DOI: https://doi.org/10.1016/j.jnca.2017.10.016
Reference: YJNCA1996

To appear in: *Journal of Network and Computer Applications*

Received date: 8 August 2017
Revised date: 18 October 2017
Accepted date: 28 October 2017

Cite this article as: Faheem Ullah, Matthew Edwards, Rajiv Ramdhany, Ruzanna Chitchyan, M. Ali Babar and Awais Rashid, Data Exfiltration: A Review of External Attack Vectors and Countermeasures, *Journal of Network and Computer Applications,* https://doi.org/10.1016/j.jnca.2017.10.016

Data Exfiltration: A Review of External Attack Vectors and Countermeasures

Faheem Ullah[1], Matthew Edwards[2], Rajiv Ramdhany[2], Ruzanna Chitchyan[3], M. Ali Babar[1, 4], Awais Rashid[2]

[1]University of Adelaide, Australia & [4]IT University of Copenhagen, Denmark

[2]Security Lancaster, School of Computing and Communications, Lancaster University, UK,

[3]Department of Computer Science, University of Leicester, UK

ABSTRACT

Context: One of the main targets of cyber-attacks is data exfiltration, which is the leakage of sensitive or private data to an unauthorized entity. Data exfiltration can be perpetrated by an outsider or an insider of an organization. Given the increasing number of data exfiltration incidents, a large number of data exfiltration countermeasures have been developed. These countermeasures aim to detect, prevent, or investigate exfiltration of sensitive or private data. With the growing interest in data exfiltration, it is important to review data exfiltration attack vectors and countermeasures to support future research in this field. *Objective:* This paper is aimed at identifying and critically analysing data exfiltration attack vectors and countermeasures for reporting the status of the art and determining gaps for future research. *Method:* We have followed a structured process for selecting 108 papers from seven publication databases. Thematic analysis method has been applied to analyse the extracted data from the reviewed papers. *Results:* We have developed a classification of (1) data exfiltration attack vectors used by external attackers and (2) the countermeasures in the face of external attacks. We have mapped the countermeasures to attack vectors. Furthermore, we have explored the applicability of various countermeasures for different states of data (i.e., in use, in transit, or at rest). *Conclusion:* This review has revealed that (a) most of the state of the art is focussed on preventive and detective countermeasures and significant research is required on developing investigative countermeasures that are equally important; (b) Several data exfiltration countermeasures are not able to respond in real-time, which specifies that research efforts need to be invested to enable them to respond in real-time (c) A number of data exfiltration countermeasures do not take privacy and ethical concerns into consideration, which may become an obstacle in their full adoption (d) Existing research is primarily focussed on protecting data in 'in use' state, therefore, future research needs to be directed towards securing data in 'in rest' and 'in transit' states (e) There is no standard or framework for evaluation of data exfiltration countermeasures. We assert the need for developing such an evaluation framework.

Keywords: Data Exfiltration, Data Leakage, Data Theft, Data Breach, External Attack Vector, Countermeasure

1. Introduction

Data theft (formally referred to as data exfiltration) is one of the main motivators for cyber-attacks irrespective of whether carried out by organised crime, commercial competitors, state actors or even "hacktivists". Preventing data exfiltration is increasingly becoming a challenging task due to two main reasons. First, over the span of last few years, cyber-crime has transformed from an individual's act to an organizational act. This transformation has provided attackers (or often called hackers) with high budget, resources, and sophistication to become more professional in data exfiltration. Second, the existing data infrastructure contains several means (as shown in Fig. 1a) that are originally designed for the legitimate exchange of data but can be used for data exfiltration.

Fig. 1b shows one scenario of how these legitimate means can be leveraged for data exfiltration. *Stage-1*: The attacker researches the available means within an enterprise to find some weaknesses that can be exploited to exfiltrate data. *Stage-2*: Once the weakness has been identified, the attacker launches the attack to exploit the identified weakness. The attack can either be network-based or physical-based. In a network-based attack, an attacker may use different techniques such as phishing email, malware, or some kind of injections. In a physical attack, an attacker may leverage techniques like copying data to a removable device or even get hold on some printed documents. *Stage-3*: Once the attacker gains the required access to the sensitive data, the process