Methods for Detecting Data Exfiltration on Corporate Environments: A literature review

Gabriel Ferrari Wagnitz

[1]Departamento de Informática – Universidade Federal do Espírito Santo (UFES)

gabriel@wagnitz.com.br

Abstract. *Data exfiltration, that being the theft of private information or industrial secrets, is a very important topic of concern not only for professionals in IT but for society in general, as the leak of confidential data causes harm ranging from spam telemarketing calls to identity theft and fraud. This paper is a literary review on the subject of Data exfiltration detection, the existing techniques and how they address the many approaches used by malicious actors and also their limitations. This article focus on examining the approaches that uses exclusively network traffic data, analysing the metrics and algorithms used by the authors of the referenced articles, as well as the tools used to collect it.*

1. The revision process

One of the first challenges encountered was how to define the scope of the research in order to obtain from the academic paper repository the articles with the content I was looking for. There are a lot of papers written about the subject of Data Exfiltration and Cybersecurity Threats in a broader sense but there's only a small subset that explores a technical solution to the problem, specially sharing techniques and results.

The solution was trying to focus on finding those that uses the network traffic analysis, and also remove from the scope the ones that works on detecting attacks through the DNS protocol, since this approach differs from the general Data Exfiltration attacks.

The resulting search string used was:

(TITLE (exfiltration) OR TITLE (threat)) AND (TITLE (detection OR prevention OR detecting OR protection OR protect) AND TITLE-ABS-KEY (network AND traffic) AND NOT TITLE (dns)) AND PUBYEAR >2010 AND (LIMIT-TO (SUBJAREA , "COMP"))

That reduced the search to 78 documents in Scopus including articles between conference papers and articles. Since I didn't have much success due to time constrains on work more on the filtering and articles, and since the most relevant articles on the top of the search passed all the criteria. This review comprises of a summary of knowledge gathered by the 4 control articles plus the top 5 articles from the search.

We chose articles form 2011 onwards since it was around that year that cloud computing, working from personal VPNs, corporate applications going public, startups gathering big data started to be relevant on the main stream media.

2. Research Summary

The problem of data exfiltration, specially from an insider threat agent, has been acknowledged as a challenge in the field of cybersercurity. That is because it is a multi-step process, each one demanding a different approach to detect and prevent. On the next chapters we will define each step and provide an overview of the solutions proposed by the analysed literature. On Table 1 we describe the steps of a typical exfiltration attack.

The process described here is similar to the one proposed by Bertino and Ghinita (2011), with the addition of an earlier ”Network Intrusion” step in the case of external threat actors, since the cited article focuses on internal threats.

Dimension of Activity	Action
(A) Network intrusion (External Attacker)	1. Knowledge gathering about the environment
	2. Obtain privileged access on user machine
	3. Escalate privilege to read access on database, file storage or exchange server
(B) Identify data sources	1. Learn schema, File System structure, Application Storage structure
	2. Find objects that contain sensitive data
	3. Learn authorization settings for database entities, folders, file storage buckets, etc.
	4. Issue fake interrogations/commands/access to conceal tracks
(C) Retrieve data from source	1. Transfer preparation (i.e., narrow-down) queries
	2. Bulk transfer queries
	3. Issue decoy queries
(D) Lateral Movement	1. Transfer of results within organization
	2. Changes to authorization permissions
	3. Hide/deactivate staging resources
	4. Encrypt data
(E) Exfltration (proper)	1. Transfer of data outside organization
	2. Hide data transfer within complex operations
	3. Initiate fake transfers to check for surveillance signs

Table 1. Summary of Dimentions and Actions within Exfiltration Mission

2.1. Network Intrusion

This article wont go deep on review detections techniques for network intrusion as this is a broad topic within cybersecurity. The many ways in which a malicious agent can infiltrate a network, gaining ungranted privileged access to services, servers and applications includes [Chapple et al. 2021]:

- Exploitaition of application vulnerabilities
- Phishing
- Social Engineering
- Brute-force authentication
- Exploitation of insecure authentication (weak passwords, no MFA, repeated passwords, etc.)
- Phisical Intrusion

Most of the security work done in organizations are focused on mitigating these types of attack, preventing the ways in which potential attacker can enter the network. The techniques applied for that varies from cybersecurity awareness programs to the staff, to deter users from clicking sketchy emails or downloading software from unknown sources, to zero trust network architectures, that tries to reduce the number of paths an attacker could take to enter the network and to propagate the attack once inside.

The step of network intrusion is not specific to data exfiltration attacks but for almost all types of attacks: Ransomware, malware that disrupt systems functionality, bitcoin mining on the companies computers, accessing private emails and data exfiltration. Without gaining access to the environment, only a few malicious activities can be performed such as denial of service or feeding false data to online forms.[1]

One big challenge though is preventing insider attacks, since exfiltrated data are often piggybacked on top of conventional traffic, such as web, email or instant messaging. And access control also does not solve the problem of malicious insiders who have the proper credentials to read data. [Bertino and Ghinita 2011] Hence the present day perimeter security solutions such as IDS/IPS systems, firewall, etc. are not capable of detecting insider attacks.[Suresh et al. 2012]

2.2. Identify Data Sources

The article by Bertino and Ghinita (2011), focuses on data being exfiltrated from a DBMS, and the one by Surresh, Malhotra, Kumar and Thanudas (2012) proposes a solution for FTPS servers, but the idea of identifying anomalous data access still holds true if the storage to be monitored is an Exchange Server, Code Repository, Credential Vault, etc.

One of the initial steps an attacker may take once they have read access on the data is trying to understand how the data is structured to identify how relevant data is stored, the folder structure, content of database tables, naming conventions etc.

A technique proposed by Bertino and Ghinita (2011) is to implement an anomaly detection system to identify this exploratory behavior on data access and identify early stages of exfiltration attempts.

[1]One recent example of this type of attack was on the mass surge in fake reservations in Bolsonaro's campaign events on the website, mirroring the ones done against Trump in 2020

2.3. Retrieve data from source and Lateral movement

After an initial survey on the data, an attacker will execute bulk extrations on the data and move it through the network to a location with less oversight (i.e. user workstation).

An efficient solution to detect data exfiltration attempts must monitor the activity of the data source, acting on understanding what comprises legitimate data access vs malicious exfiltration activity. During the process of retrieving larges amounts of data there often are spikes in network traffic or CPU usage on the servers that can be used to detect that an attack is occurring.

In a more general note we can list two distinct approaches on how to work on the data source level, that comprises the application (FTP, DBMS, Exchange) and also the server that hosts it.

The first technique is based of an access policy and the second one on abnormal behaviour detection:

2.3.1. Confine and Mark Policy

One proposed solution for this issue is a ”confine-and-mark” approach for protection against exfiltration of sensitive data, that consists in [Bertino and Ghinita 2011]:

1. Restrict the segment(s) of the network from which sensitive data can be accessed.
2. Label sensitive data with cryptographic authorization tokens, such that only data that have been approved for transfer can leave the protected network.

That way the traffic the attempts of exfiltration data can be detected with some form of deep package inspection. This approach requires the implementation of a process of authorization for individual batches of data, for individual users. The network diagram of an environment implementing this technique is demonstrated on Figure 1.

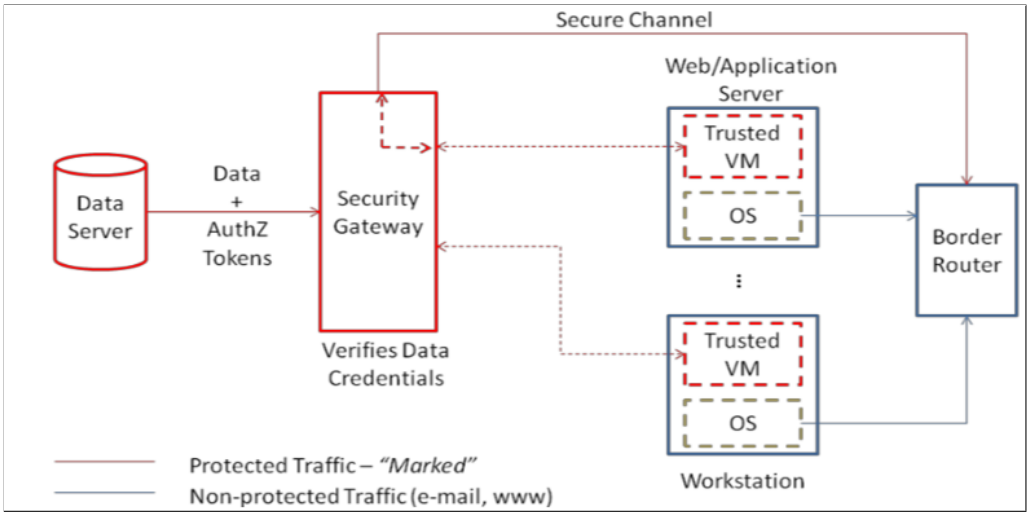

Figure 1. mark-and-confine network diagram

2.3.2. Anomaly Behaviour Detection

On the Surresh et al. (2012) article it names this approach IDEMT (Integrated data exfiltration monitoring tool). The idea behind it is implement two layers of anomaly detection, one for network traffic, that correlates learnt normal behaviour with current measured to raise alerts, and one for data access, understanding access pattern on the application level.

The implementation of this technique comprises of two phases: learning and detection. On the first stage data is captured both by application and network traffic logs and fed to a Traffic Behaviour Analyser and a File Access Pattern Analyzer, to create the base confidential values.

Traffic Behaviour Analyser utilizes metrics related to Inbound vs Outbound traffic volume. The normal correlation between traffic in and out is above 0.8, but can vary in different environments.

File Access Pattern Analyzer utilizes metrics related to the number of files a given user access, both counting the number of access files and also which are the frequently access files.

Once the determined period for the learning phase ends, the application now will evaluate the traffic based on the flow shown in Figure 2

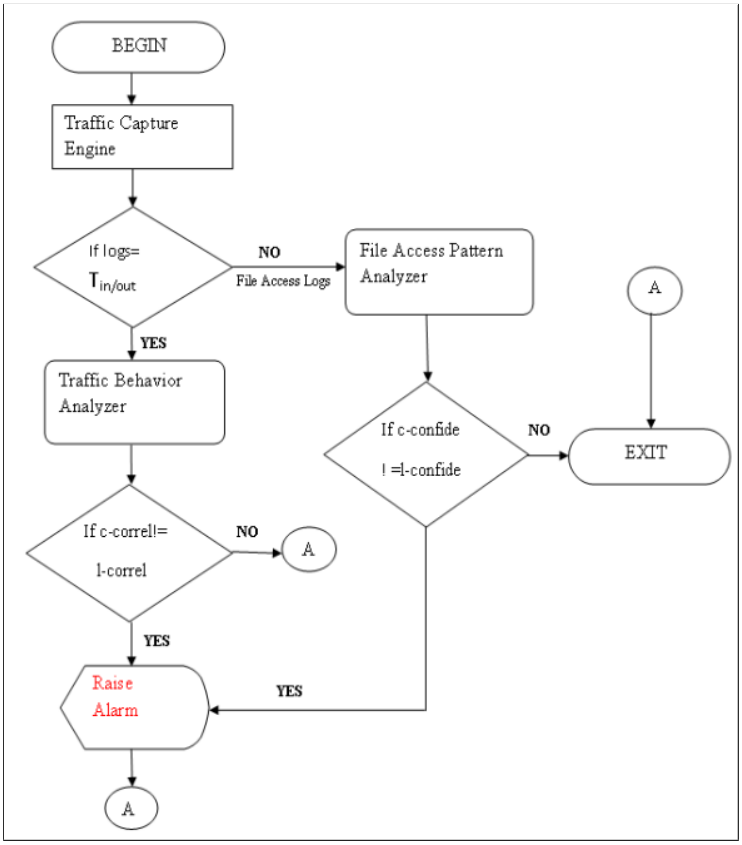

Figure 2. IDEMT flow chart

2.4. Exfiltration

The proper exfiltration process is when the malicious agent succeeds in transferring the confidential information to outside the organization premises. This step can be achieved using many different methods, some of the more common are [Suresh et al. 2012]:

- Native Remote Access Applications
- FTP capable Malware
- Native FTP client
- SQL injection
- SMTP capable Malware
- HTTP File upload

Detecting data exfiltration at this stage, specially when the data has already been manipulated and moved inside the network can be even more challenging since it can now be moved from a place where is not expected to be and often in a different format.

A more intelligent attacker may use sophisticated methods to disguise exfiltrated data as ordinary traffic such as sending data through DNS protocol or image stenography [King et al. 2021]. And that's why the literature focus it's efforts on the internal network connection of the data repositories.

3. Conclusion

There are many challenges on detecting data exfiltration, specially because the scenarios varies much between organizations. Some organizations may have large volumes of personal or patient data stored that cannot be easily moved without notice, others have small very valuable assets that needs to be secured such as a script for a new movie or show, the patent for a product, legal documents, etc. The field of cybersecurity still needs a lot of research on the best practices and techniques to prevent those types of attacks and as most cases in this area, the best approach may be a combination of multiple tools and procedures.

We need to understand and measure what are the best data points to be extracted from the environment and combine it with governance to implement business practices to create a security framework to promote confidentiality without harming usability of legitimate users.

References

Campanha anti-bolsonaro nas redes reserva ingressos para convenção do pl no rj; partido diz que tomará medidas cabíveis. *G1 Rio de Janeiro*.

Bertino, E. and Ghinita, G. (2011). Towards mechanisms for detection and prevention of data exfiltration by insiders. *ACM Conference on Computer and Communications Security*.

Brindha, T. and Shaji, R. (2015). An analysis of data leakage and prevention techniques in cloud environment.

Chapple, M., Stewart, J., and Gibson, D. (2021). *(ISC)(2) CISSP certified information systems security professional official study guide, 9th edition*. Sybex, Indianapolis, IN, 9 edition.

Hou, H., Xu, Y., Chen, M., Liu, Z., Guo, W., Gao, M., Xin, Y., and Cui, L. (2020). Hierarchical long short-term memory network for cyberattack detection. *IEEE Access*, 8:90907–90913.

King, J., Bendiab, G., Savage, N., and Shiaeles, S. (2021). Data exfiltration: Methods and detection countermeasures. pages 442–447. Institute of Electrical and Electronics Engineers Inc.

Suresh, N. R., Malhotra, N., Kumar, R., and Thanudas, B. (2012). An integrated data exfiltration monitoring tool for a large organization with highly confidential data source. pages 149–153.

Tsikerdekis, M., Waldron, S., and Emanuelson, A. (2021). Network anomaly detection using exponential random graph models and autoregressive moving average. *IEEE Access*, 9:134530–134542.