Patented Method for **Detecting** and **Preventing** Data Exfiltration Attacks

U.S. Patent - US 11,457,025 B2

Inventors:
Balaji Venkat Venkataswami, Suri Parthasarathy, Chandramouleeswaran Sundaram,
Ragavendran Selvaraj, Mohamed Ismail Ibrahim, Chandrasekar Balasubramanian

Data exfiltration refers to the malicious transfer of data from a computer or other device. It may be done manually by someone with physical access to a computer or by an automated program over a network. Such exfiltrated data streams closely mimic routine network traffic, and hence detecting these attacks can be challenging. Since endpoints are one of the easiest access points for hackers, an all-encompassing approach to data security that monitors and protects every endpoint in an organization's network is critical.

Challenges in Detection and Prevention

Systems that rely on vendor-set, common, easy-to-crack passwords are most vulnerable to data exfiltration. Hackers can gain access to target machines through remote applications or by installing a removable media device, in cases where they have physical access to the target machine. In cyberattacks such as Advanced Persistent Threats (APTs), malware gains access to the network and remains undetected as it stealthily seeks out target data. Since APTs typically rely on social engineering techniques or phishing emails to launch them, user education is a preventative measure. However, it is difficult to adequately block the download of such malware without restricting access to applications that users need.

To effectively compromise an endpoint, malware communicates externally with a Command and Control (C&C) server to receive instructions or exfiltrate data. Hence, continuous monitoring of all communication in the network, early detection and blocking of any new or unauthorized communication is critical. Current solutions to

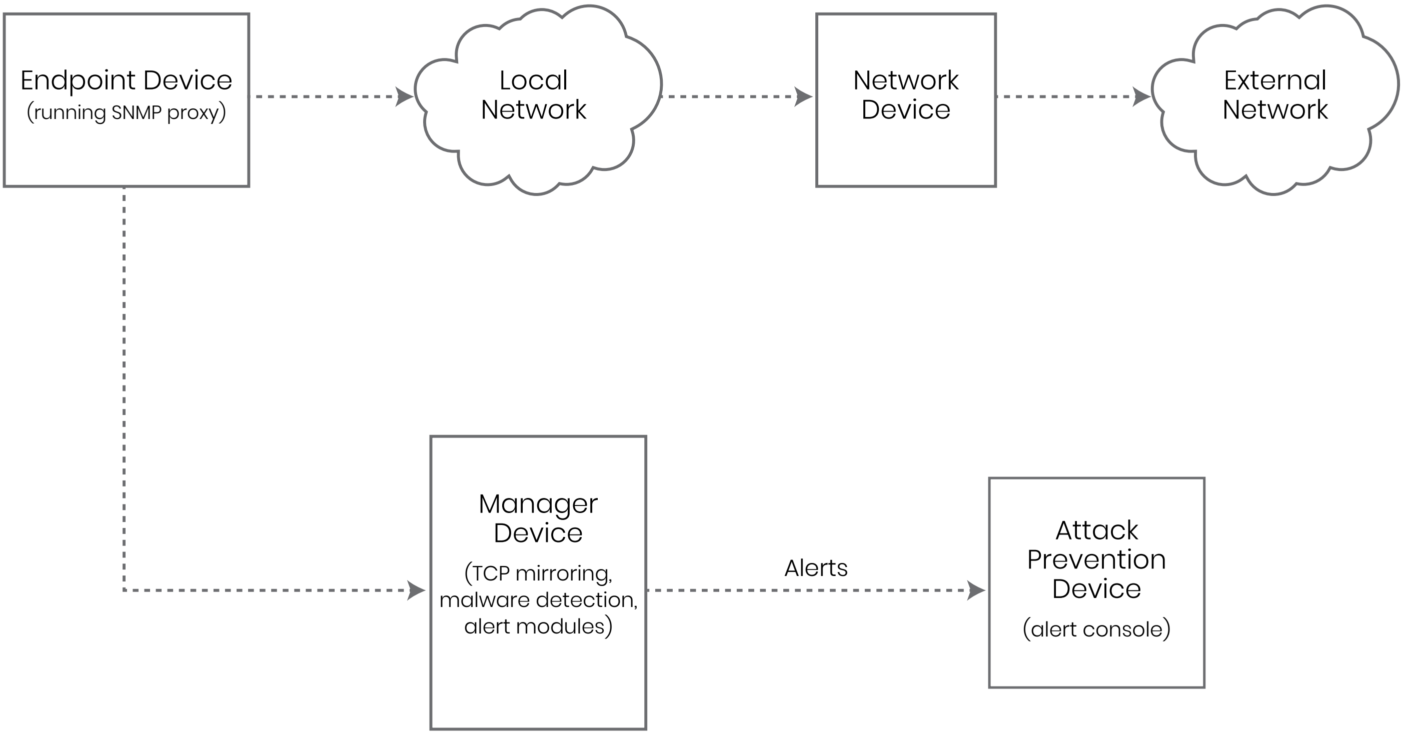
monitor, detect, and prevent/mitigate such attacks focus on different approaches such as identifying and mitigating malicious network threats, monitoring device data and gateway data, dynamic device clustering using device profile information, and so on. These solutions rely on either rule-based or anomaly-based methods to detect behavioral deviations. While rule-based methods are generally quite accurate with current attack techniques, they are unable to handle new types of attacks.
With anomaly-based methods, the effectiveness depends on how *normal* and *abnormal* behavior are identified. But here again, this determination is based on specific attack techniques. Since cyber criminals use a variety of techniques to bypass network surveillance, these solutions do not effectively address all scenarios.

The Patented Solution

The patented method and system to detect and prevent data exfiltration attacks uses Deep Learning (DL) algorithms. The method includes detecting downloads at one or more endpoint devices using an SNMP (Simple Network Management Protocol) proxy installed in them. The proxy detects download activity based on MIB (Management Information Base) data. The network traffic associated with the downloads is mirrored and streamed to the SNMP manager device. These mirrored streams of data are provided as input to a DL model that detects presence of malware in the downloads and alerts and stops the data transfer. The method includes creating a training dataset, a validation dataset, and a test dataset from historical network traffic. The DL model is trained using the training dataset to detect presence of malware, continuously tuned using the validation dataset, and evaluated using the test dataset.

The solution addresses many scenarios like:

- Data flowing out through port 80
- Static list of blacklisted websites
- Data sent as an attachment in an email
- Not able to read encrypted data sent thru https
- Internal attacks like TCP SYN flood or ICMP Smurf attack
- Man in the Middle (MitM) attacks



ZIF™ (Zero Incident Framework™), is an award-winning AIOps platform for IT Operations. ZIF^{TM} delivers business outcomes by leveraging unsupervised pattern-based machine learning algorithms. Infrastructure and application telemetry data are aggregated, correlated, and potential failures are predicted. To enable faster resolution and better user experience, ZIF^{TM} deploys intelligent bots for proactive remediation. Developed by GAVS Technologies (www.gavstech.com), ZIF^{TM} is available as an on-premises and SaaS solution.

To find out more about ZIF^{TM},
please visit **www.zif.ai** or write to **inquiry@zif.ai**