DARD – A DECOY APPROACH FOR ENHANCING INTELLECTUAL PROPERTY THEFT PROTECTION AGAINST AUTOMATED CYBERATTACKS SECURING USB DRIVES

# Mrs. S.Kaviya 1, J. Rizwana2, M. Sathya3, M. Nargies4

1M.E., Assistant Professor Department of ECE P.A. College of Engineering and Technology, Pollachi, Coimbatore- 641 202, India.

2,3,4UG Scholar, Department of ECE P.A. College of Engineering and Technology, Pollachi, Coimbatore-641 202, India.

rizwanaj04@gmail.com,sathya57673@gmail.com,nargies14@gmail.com,skaviyaece@gmail.com

# ABSTRACT

The threat of cyberattacks is steadily increasing due to the advent of smart functioning and current global events. Attackers sometimes aim to steal valuable data, such as intellectual property (IP), by exfiltrating a large number of IP documents from a target company. They then use automated methods to identify individuals of interest. In order to fool attackers who employ automated techniques for categorising stolen documents, this article presents the DARD (Deceptive Techniques for Resilient Defence against IP theft) system. DARD automatically generates a new deceptive repository from a given original repository of documents, which misleads popular automated algorithms and produces clusters of documents that are very different from the real ones. By using this approach, DARD stops attackers using automated techniques from effectively grouping and topic-identifying documents. In order to create a misleading repository, DARD employs four deceptions, which are introduced in the article: Basic Shuffle, Shuffle increment, Shuffle decrease, and Change topic. We evaluate our system's performance by taking into account three different kinds of attackers with varying levels of expertise. Through extensive tests, we show that the DARD system can deceive document clustering techniques as well as autonomous topic modelling, including well-known commercial systems like Amazon Comprehend. Thus, our approach provides a strong defence against theft of intellectual property (IP).

**Index Terms— Deceptive repository, clustering, topic modeling, adversarial setting.**

# INTRODUCTION

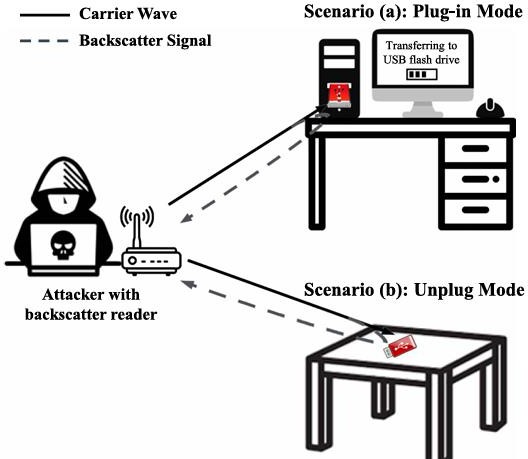
USB drives are widely used for data storage and transfer due to their portability, ease of use, and compatibility. However, these devices are highly vulnerable to security risks, such as unauthorized access and malware attacks. Traditional approaches to mitigating these risks often fall short in detecting and responding to such sophisticated attacks. To address these challenges, this project introduces an integrated framework that combines advanced malware detection using Deep Neural Networks (DNNs).

Fig.1. SpyUSB enables an attacker to launch attacks in two scenarios.

while external and portable hard drives are real hard disks. As such, they have internal moving parts and don't serve well as MP3 players on a long-distance run. Fig. 1 depicts the attack scenarios of SpyUSB. The attacker employs a backscatter reader to provide RF carrier signals for SpyUSB back scattering and can read the back scattered data from SpyUSB. When SpyUSB is plugged into a host computer (i.e., plug-in mode), the attacker can eavesdrop on instant bit stream from the communication between the host computer and data storage of SpyUSB.

# METHODOLOGY

**A. Threat Model**

*Manufacturing and distribution of SpyUSB: SpyUSB* can be distributed through two channels: supply chain attacks and retrofitting a normal USB flash drive. In the first channel, *SpyUSB* can be manufactured as a whole and sold or given out as a normal USB flash drive. For example, it may be given away as a shopping giveaway or a souvenir for conference attendees. In the second channel, we design the spy module as an independent part that can be implanted into a USB flash drive by only connecting through two ports: a USB interface with four lines and an SPI interface with four lines. In this condition, users trust *SpyUSB* and use them as normal USB flash drives.

**B. Overview of SpyUSB Attack**

We achieve data theft by directly inserting a tiny spy module with a backscatter transmitter into a USB flash drive, which is treated as a benign device. The main purpose of an attacker is to steal data from a USB flash drive. He can use a remote transceiver to send activation commands and carrier waves to the spy module to initiate an attack. The entire process is completed with the cooperation of a hardware spy module and a remote attacker, which can be split into three steps: device activation, data collection, signal modulation, and transmission,

***Device activation:*** In order to ensure the attacker can receive unabridged messages and reduce the power consumption of *SpyUSB*, data theft is not always ongoing. *SpyUSB* is in sleep mode before being activated by a remote attacker’s instruction, which is a predetermined binary code. The *SpyUSB* device contains an ultra-low power wake-up receiver circuit to identify the activation command.

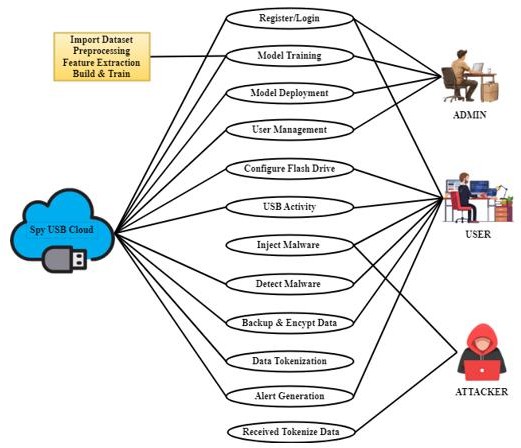
***Data collection:*** As for data collection, *SpyUSB* chooses to work in plug-in or unplug mode depending on whether the USB flash drive connects or disconnects with a computer. In plug-in mode, the spy module works as a slave of the data bus in order not to disturb communications on the data bus. It never sends data or instructions on its own initiative but collects data by eavesdropping on all data flow on the bus.

Fig. 2. SpyUSB enables an attacker

As described, we expect that Grey Box and Enhanced Grey Box advertisements may remove misleading keywords from the repository R′. Additionally, we assume that all of the adversaries in this section are unable to access the mapping of keyword replacements, employ K-means as the clustering algorithm, and fig.2 discribe about estimate the number of clusters in the exfiltrated repository using the Silhouette Coefficient, Calinski-Harabasz Index, and Davies-Bouldin Index. It should be noted that as the three adversaries lack the ground truth of the repository R, they are unable to self-estimate values such as the Purity or the ARI on their clustering.

* 1. **PROPOSED METHODOLOGY**

1. The Deceptive Repositories

We created six distinct misleading repositories from the repository Rd presented in order to compare the opponents' performances. Basic Shuffle, Shuffle Increment, and Shuffle Reduction are the deceptive operations that were applied to Rd in order to generate the six repositories. Every misleading operation has been carried out twice: once using the limited version of K-means and once using random selection to divide the documents. All of the subsets have been created using roughly the same number of documents using both approaches. According to their TF-IDF weight, the keywords have been replaced in descending order. Finally, for comparison, the number of m-terms- replacements made by each operation was the same. Based on our experiments, we have set the number of m-terms- replacements at 60.

1. Attacking the Deceptive Repositories

We use the Adjusted Rand Index (ARI) to gauge the adversaries' performance. The ARI calculates the ratio of agreement between a predicted grouping (the one the adversary, in our case, achieved) and the clustering given by the documents' genuine labels. An ARI of 1 indicates full agreement between the two clusterings, an ARI around 0 indicates a random labelling of the anticipated clustering, and an ARI with a negative value indicates a labelling that is worse than a random one. The ARI ranges from -1 to 1. The Black Box attackers believe they have gained access to the unmodified repository. As described in Black Box attackers examine the exfiltrated repository.

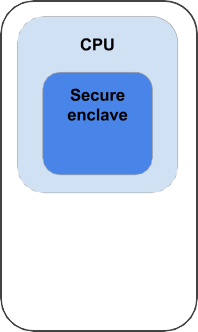


Fig. 3. Server process

DARD is no more complicated than traditional encryption in terms of implementation ease and end-user usability. Both techniques entail keeping plaintext in RAM while the operation is running and using secure storage for secrets (such encryption keys or keyword-to-mapping keys). Fig.3 server process of cpu. Thus, DARD can be easily integrated into an organization's current security infrastructure with minimal complexity or additional burden.

DARD's misleading stores offer attackers a lag before they realise they have come across designed content, unlike encrypted information, where attackers may identify encrypted material instantly upon discovering it. Because of this delay, security teams have valuable time to detect and address issues before hackers move on to other sites or resources in search of unencrypted data. Additionally, in some cases, the DARD system can be used in conjunction with encryption to further enhance data security and privacy. Organisations will have a stronger and more comprehensive security posture when the two approaches are strategically combined. DARD can supplement encryption in the following scenarios.

# RESULTS AND DISCUSSION

Data encryption and obfuscation are only two of DARD's deception tactics. In order to give attackers the impression that they have gained access to important data, the repository's structure—that is, the number of clusters and their subjects—must be carefully planned. Attackers are enticed to conduct additional analysis as a result, squandering time and money on pointless material. In addition to enabling intrusion detection, this active approach actively deceives attackers by distracting them from the real data.



Fig. 4. Securing USB Drive

Fig. 4. Securing USB Drive describe about Hardware modification is the pre-requisite of SpyUSB. A SpyUSB can be distributed through two channels. The first channel is supply chain attacks. Spy modules are integrated into USB flash drives during manufacturing, or SpyUSB can be manufactured as a whole.



Fig. 5. Training Phase

Fig. 5. Training Phase is a critical component of the DARD approach, where machine learning models are trained to detect and prevent automated cyber attacks on USB drives. In this phase, we collect and preprocess data.

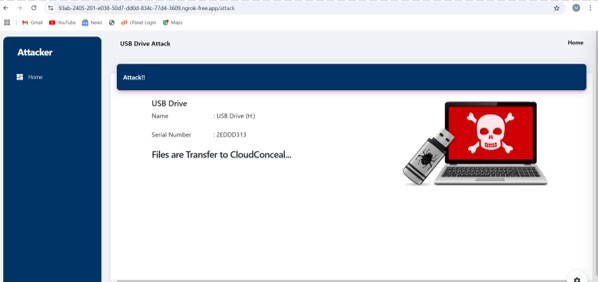


Fig. 6. Classification

# CONCLUSION

The 4-deceptive-operations framework DARD, which may be used to manipulate the clusters that are created by a document repository, is presented in this study. According to our research, dishonest operations completely fool adversaries who are not aware of this effort (0% ARI). In the worst-case scenario, they have an average ARI of 53.5% against Enhanced Grey Box adversaries, demonstrating their great effectiveness (average ARI of 25%) against adversaries who are aware of how the deceptive tactics operate. Additionally, we looked into how misleading operations affected the topic modelling task. We found that LDA only provides false keywords among the issue's most characteristiCS when the opponents use it to model the topic on the deceptive repositories. Furthermore, we show that even commercial systems like as Amazon Comprehend are effectively misled by our approach against topic modelling.

# REFERENCES

1. Deloitte. (2021). *Impact of COVID-19 on Cybersecurity*. [Online]. Avail- able:

https://www2.deloitte.com/ch/en/pages/risk/articles/impact-covid- cybersecurity.html

1. Interpol. (2021). *COVID-19 Cyberthreats*. [Online]. Available: [https://www.interpol.int/Crimes/Cybercrime/COVID-19-](http://www.interpol.int/Crimes/Cybercrime/COVID-19-cyberthreats) [cyberthreats](http://www.interpol.int/Crimes/Cybercrime/COVID-19-cyberthreats)
2. J. Tidy. *The Three Russian Cyber-attacks the West Most Fears*. Accessed: Apr. 21, 2023. [Online]. Available: [https://www.bbc.com/](http://www.bbc.com/) news/technology-60841924.
3. M. L. Morgia, A. Mei, E. N. Nemmi, S. Raponi, and J. Stefa, “Nation- ality and geolocation-based profiling in the Dark(Web),” *IEEE Trans. Services Comput.*, vol. 15, no. 1, pp. 429–441, Jan. 2022.
4. M. L. Morgia, A. Mei, A. M. Mongardini, and J. Wu, “It’s a trap! Detection and analysis of fake channels on telegram,” in *Proc. IEEE Int. Conf. Web Services (ICWS)*, Jul. 2023, pp. 97– 104.
5. CISA. *Impacket and Exfiltration Tool Used To Steal Sensitive Information From Defense Industrial Base Organization*. Accessed: Apr. 21, 2023. [Online]. Available: [https://www.cisa.gov/uscert/ncas/](http://www.cisa.gov/uscert/ncas/) alerts/aa22-277a
6. S. AG. (2020). *Software Ag Adhoc: Disruption of Services Due To Malware Attack*. [Online]. Available: [https://www.softwareag.com/](http://www.softwareag.com/) encorporate/company/news
7. T. Group. (2020). *Toll It Systems Update*. [Online]. Available: [https://www.tollgroup.com/toll-it-systems-updates](http://www.tollgroup.com/toll-it-systems-updates)
8. Verizon. (2021). *Dbir 2021 Data Breach Investigations Report*. [Online].

Available: [https://www.verizon.com/business/resources/reports/dbir/](http://www.verizon.com/business/resources/reports/dbir/)

1. Wikipedia. (2020). *United States Federal Government Data Breach*. [Online]. Available: https://en.wikipedia.org/wiki/+ 2020UnitedStatesfederalgovernmentdatabreach
2. G. Pagnotta, F. De Gaspari, D. Hitaj, M. Andreolini,

M. Cola- janni, and L. V. Mancini, “DOLOS: A novel architecture for moving target defense,” *IEEE Trans. Inf. Forensics*

pp. 5890–5905, 2023. Intel. *Intel Software Guard Extensions*. Accessed: Apr. 21, 2023. [Online]. Available: [https://www.intel.com/content/www/us/en/](http://www.intel.com/content/www/us/en/) developer/tools/softwareguardextensions/overview.html

1. Amazon. (2021). *Topic Modeling*. [Online]. Available: https://docs. aws.amazon.com/comprehend/latest/dg/topic- modeling.html
2. M. Steinbach, G. Karypis, and V. Kumar, “A comparison of document clustering techniques,” Dept. Comput. Sci. Eng., Univ. Minnesota, Tech. Rep. 00-034, 2000.
3. M. Marcin´czuk et al., “Text document clustering: Wordnet vs. TF- IDF vs. word embeddings,” in *Proc. 11th Global Wordnet Conf.*, 2021,

pp. 207–214.

1. *Open-Access Archive. Arxiv*. Accessed: Apr. 21, 2023. [Online]. Avail- able: https://arxiv.org/
2. M. S. G. Karypis, V. Kumar, and M. Steinbach, “A comparison of document clustering techniques,” in *Proc. TextMining Workshop (KDD)*, 2000, pp. 1–2.
3. T. Hofmann, “Unsupervised learning by probabilistic latent semantic analysis,” *Mach. Learn.*, vol. 42, nos. 1–2, pp. 177–196, 2001.
4. D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent Dirichlet allocation,”

*J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Jan. 2003.

1. A. K. Uysal and S. Gunal, “The impact of preprocessing on text classification,” *Inf. Process. Manage.*, vol. 50, no. 1, pp. 104–112, Jan. 2014.
2. T. Joachims, “A probabilistic analysis of the Rocchio algorithm with tfidf for text categorization,” in *Proc. 14th Int. Conf. Mach. Learn. (ICML)*, vol. 97, Jul. 1997, pp. 143–151.