Offre de stage Master 2

Intitulé : **Exfiltration des données à travers la stéganographie réseau**

Etablissement d'accueil : ENSSAT, Lannion

Laboratoire d'accueil : Equipe SOTERN, IRISA

Durée de stage : 06 Mois

PERIODE : 2025 -2026

ENCADRANTS

Mohammed NAFI (mohammed.nafi@irisa.fr), SOTERN, IRISA, Lannion

Gilles Guette (gilles.guette@imt-atlantique.fr), SOTERN, IRISA, Rennes

Guillaume DOYEN (guillaume.doyen@imt-atlantique.fr), SOTERN, IRISA, Rennes

CONTEXTE ET PROBLEMATIQUE

L'exfiltration des données est un domaine de recherche très actif à l'heure actuelle. Elle désigne le transfert illicite et non autorisé des informations sensibles ou privées à partir d'un réseau d'une organisation vers une destination externe contrôlée par un attaquant. Cette attaque peut être réalisée par un membre interne ou externe à l'organisation. Elle pourrait avoir des conséquences désastreuses pour l'organisation ciblée (réputation, etc.). Pour mener ce type d'attaque, un attaquant peut utiliser différents vecteurs comme les canaux cachés, l'injection SQL, l'écoute clandestine, le phishing, la stéganographie, etc. La stéganographie est une technique qui vise à dissimuler les informations sensibles dans un média de couverture (texte, image, audio, vidéo) de manière à ce que l'existence même de ces informations soit invisible et difficile à détecter. Parmi les techniques de stéganographie, on trouve la stéganographie réseau qui implémente la stéganographie en exploitant les protocoles réseau existants tels que TCP/IP, DNS, HTTP, etc. Elle décrit la transmission des données sur un réseau sans qu'elles ne soient détectées. L'exfiltration de données par stéganographie réseau est très difficile à détecter par les méthodes de détection traditionnelles.

Divers outils, techniques et stratégies ont été développées dans le but de se prémunir contre l'exfiltration des données. Ces techniques visent à détecter, prévenir et/ou investiguer la fuite de données. Il existe des techniques proactives et réactives. Parmi les approches proactives, on peut citer la classification des données (sensibles ou non sensibles), le contrôle d'accès, le chiffrement, le stockage distribué, etc. Parmi les techniques réactives, on trouve les méthodes qui inspectent le contenu des paquets, et celles qui se basent sur les anomalies. En effet, ces dernières analysent le trafic réseau pour tenter de détecter des anomalies et comportements suspects.

Plusieurs travaux de recherche ont étudié la stéganographie et l'exfiltration des données. A titre d'exemple, les auteurs dans [1] ont exploité l'entête des paquets IP pour dissimuler des informations secrètes. Une approche de détection d'exfiltration des données en temps réel exploitant le protocole DNS a été présentée dans [2]. Les auteurs dans [3] ont proposé un Framework qui permet la détection de codes malicieux cachés dans les protocoles TCP/IP, tels que DNS et HTTP. Une classification des méthodes d'exfiltration de données en trois catégories, à savoir les méthodes basées sur le contenu, les en-têtes et les métas données a été présentée dans [4]. Dans [5], les auteurs ont proposé un Framework de détection de logiciels malveillants qui utilisent la technique de stéganographie (stegomalwares). Leur modèle combine à la fois les techniques d'apprentissage supervisé et non supervisé pour analyser les codes malveillants. Les auteurs dans [6] ont proposé une méthode de prévention d'exfiltration des données basée à la fois sur le modèle de machine Learning et le chiffrement pour protéger les données en transit. Dans leur modèle, seules les données confidentielles seront chiffrées avant d'être transmises. Une méthode de détection d'exfiltration des données utilisant le modèle de machine Learning a été présentée dans [7]. Les auteurs dans [8] ont proposé une technique basée sur le chiffrement par attributs qui combine à la fois la détection et la prévention d'exfiltration des données.

Ce stage de Master s'inscrit dans le cadre de l'étude de la problématique de sécurité dans les réseaux. Il concerne principalement l'exfiltration des données à travers la stéganographie. L'accent sera particulièrement mis sur la stéganographie réseau qui exploite les protocoles réseaux disponibles, tels que DNS, HTTP, etc.

OBJECTIFS

L'objectif principal de ce stage est d'étudier, dans un premier temps, les différentes méthodes d'exfiltration des données. L'accent sera particulièrement mis sur les approches utilisant la stéganographie réseau [9, 10]. Dans un second temps, les techniques de prévention et de détection des attaques d'exfiltration des données par stéganographie proposées dans littérature seront passées en revue et classifiées.

Le déroulement de ce stage est organisé comme suit :

1. Identifier et étudier les différents vecteurs d'attaque d'exfiltration de données et les techniques de défense (contre-mesures) proposées dans la littérature
2. Proposer une classification à la fois des méthodes d'exfiltration et les contre-mesures
3. Faire une étude comparative des techniques étudiées
4. Etudier l'applicabilité des techniques de défense aux différents vecteurs d'attaques
5. Tester quelques scenarios d'attaques et de défense

REFERENCES

[1] Ganivev, A., Mavlonov, O., & Turdibekov, B. (2021, November). Improving data hiding methods in network steganography based on packet header manipulation. In 2021 International Conference on Information Science and Communications Technologies (ICISCT) (pp. 1-5). IEEE.

[2] Ozery, Y., Nadler, A., & Shabtai, A. (2024). Information based heavy hitters for real-time DNS data exfiltration detection. In Proc. Netw. Distrib. Syst. Secur. Symp (pp. 1-15).

[3] Zillien, S., Petrov, D., Ruffing, P., & Gross, F. (2024, June). A Development Framework for TCP/IP Network Steganography Malware Detection. In Proceedings of the 2024 ACM Workshop on Information Hiding and Multimedia Security (pp. 95-100).

[4] King, J., Bendiab, G., Savage, N., & Shiaeles, S. (2021, July). Data exfiltration: methods and detection countermeasures. In 2021 IEEE International Conference on Cyber Security and Resilience (CSR) (pp. 442-447). IEEE.

[5] Kadebu, P., Shoniwa, R. T., Zvarevashe, K., Mukwazvure, A., Mapanga, I., Thusabantu, N. F., & Gotora, T. T. (2023). A hybrid machine learning approach for analysis of stegomalware. International Journal of Industrial Engineering and Operations Management, 5(2), 104-117.

[6] Apau, R., Asante, M., Twum, F., Ben Hayfron-Acquah, J., & Peasah, K. O. (2024). Image steganography techniques for resisting statistical steganalysis attacks: A systematic literature review. PloS one, 19(9), e0308807.

[7] Das, A., Shen, M. Y., Shashanka, M., & Wang, J. (2017, December). Detection of exfiltration and tunneling over DNS. In 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 737-742). IEEE.

[8] Sasikala, V., & CH, B. S. (2024, March). Data Leakage Detection and Prevention Using Ciphertext-Policy Attribute Based Encryption Algorithm. In 2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 1-5). IEEE.

[9] Seo, J. O., Manoharan, S., & Mahanti, A. (2016, August). A discussion and review of network steganography. In 2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech) (pp. 384-391). IEEE.

[10] Lubacz, J., Mazurczyk, W., & Szczypiorski, K. (2014). Principles and overview of network steganography. IEEE Communications Magazine, 52(5), 225-229.