Malware Attribution

Jennie Parrott

Cybersecurity Service Provider Internship Program, ITS 293, Spring 2023

Advisors: Dr. Curtis Arnold, Chief Scientist – Cyber, & Dr. Robert Adamove, Sr. Associate, (Core4ce)

Problem Statement:

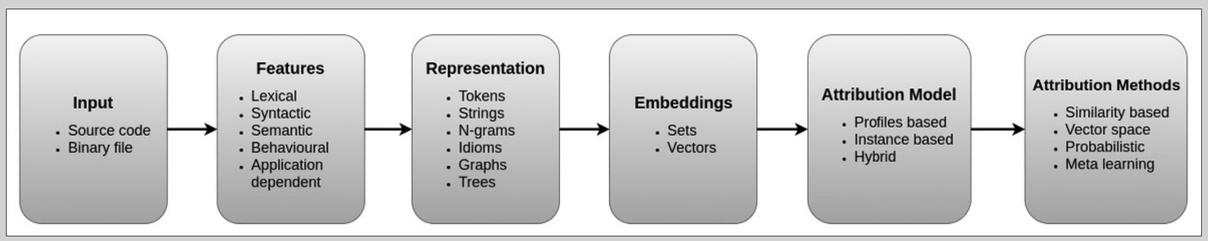
Malware attacks used by state and non-state cybercriminals are on the rise. Malware attribution and coding signature analysis try to determine who was behind a cyberattack, why they did it, and how to identify harmful attacks. However, the availability of malware code generators and the increasing use of sophisticated obfuscation techniques, which enable malware developers to produce malware variants at unprecedented rates, make this process more complex and call for sophisticated methodologies. Lack of research limits cyberattack attribution and countermeasure development. Advanced attribution and coding signature analysis methods are needed to detect cyberattack sources and types.

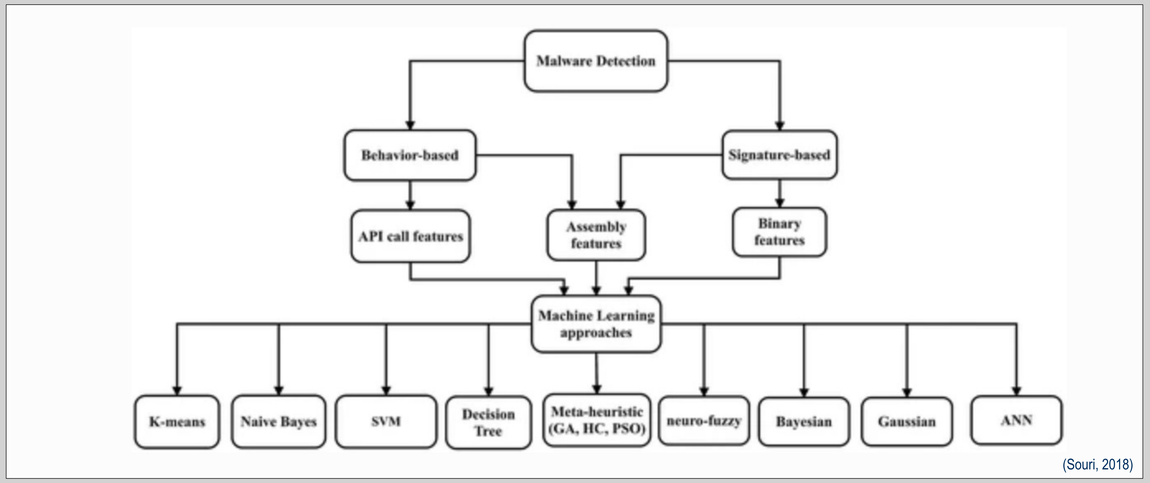
Purpose:

The study aims to explore a small sample of malware attribution techniques and evaluate the effectiveness of each approach—limitations and challenges associated with malware attribution. The significance of the malware attribution study is enhancing understanding of the techniques for identifying individuals, groups, or nations responsible for creating and distributing malware. The significance of malware attribution lies in its ability to help defend against future attacks by providing valuable intelligence about the motives, tactics, techniques, procedures, and capabilities of attackers. Some key reasons malware attribution is significant include response and countermeasures, legal action, diplomacy, and intelligence gathering.

Approach:

This study uses a qualitative approach using secondary research from academic journals and literature published within the past five years. The study focuses malware attribution methodology and reviews coding signature analysis, similarity-based, vector space, probabilistic, and meta-learning methods, and the challenges and limitations of each method.





(Souri, 2018)

Discussion & Conclusion:

This research study presents a sample of malware attribution methods. Each method could be used independently or in combination of methods. A combination of methods increases accuracy and reliability. However, this is not a "silver bullet," as each method has associated challenges that must be solved. First, a lack of source code is shared between organizations and entities, making it challenging to aggregate and analyze data from multiple sources, thus impeding the accuracy and effectiveness of attribution efforts. Another challenge is the use of obfuscation techniques by attackers. Obfuscation techniques make computer code more difficult to understand or reverse-engineer, making it difficult for analysts to attribute an attack to a specific source accurately. Furthermore, the evolving programming styles can make establishing a clear and accurate picture of the attack landscape challenging.

Path Forward:

Recommendations for future research include direct extraction of constructed binary code, extraction of obfuscation-resistant characteristics, and a self-learning feature in the system to understand malware authors' growing programming style. Recommendations to improve the effectiveness of malware attribution include establishing industry-wide standards, collaboration, data sharing, and developing joint initiatives.

References:

Ahmadian, A., & Maleki, H. (2018). Malware attribution techniques: A survey. Journal of Network and Computer Applications, 112, 1-19. doi: 10.1016/j.jnca.2018.05.017.

Kalgutkar, V., Kaur, R., Gonzalez, H., Stakhanova, N., & Matyukhina, A. (2019). Code authorship attribution. ACM Computing Surveys, 52(1), 1–36. https://doi.org/10.1145/3292577

UNCLASSIFIED/APPROVED FOR PUBLIC RELEASE