

THREATRAPTOR

Motivation

- Traditional methods rely heavily on manual construction of queries, which are difficult to effectively query when the scale increases, and are easy to miss.
- IOCs are often fragmented, isolated, and lack a comprehensive enough description of the attack, so they cannot be used effectively. If you just rely on IOCs, the results will be very poor.
- OSCTI has more abundant and detailed information, but it is not used reasonably at present.

Innovation

- THREATRAPTOR uses unsupervised NLP to automatically extract IOCs from unstructured reports and generates a threat behavior graph to create a more complete attack chain.
- A special query language is used to support multiple query modes. A scheduling algorithm based on query semantics is also introduced to optimize query response time and execution efficiency.
- For more security, IOC protection mechanism is introduced. Then using the core anaphora resolution technology, the same IOC in different positions can be correctly identified and merged.

Pipelines

- Receive the unstructured OSCTI report.
- Through NLP, IOCs are extracted and dependencies are constructed.
- According to the extracted IOCs, a threat behavior graph is constructed.
- Based on the extracted threat behavior graph, the system automatically generates TBQL.
- Through the generated query, it looks for potential malicious attacks and returns the relevant log data.

Discussions

- THREATRAPTOR can use OSCTI, but it is extremely dependent on the quality of OSCTI. If OSCTI reports are inaccurate or not updated, then the threat cannot be identified.
- THREATRAPTOR can only handle a limited number of attack patterns and is difficult to cope with, for example, side-channel attacks.

Thoughts

- THREATRAPTOR innovates by using an automated threat behavior extraction and query mechanism, which greatly reduces human intervention and has higher flexibility and accuracy. On the other hand, the system may be able to use more data sources to make threat judgments, or train some model to make some degree of prediction, thereby reducing the impact of new types of threats.
- For me, network security and threats are directions that I have not touched at all. However, this kind of cyber threat intelligence gathering and knowledge base management seems to be easier to get started with.