Blocklists: What should you block and why?

Blocklists are often used in research to label malicious hosts and to train models to identify similar malicious targets. A quick web search reveals that a multitude of blocklists exist. Often, their entries differ massively. Some differ in respect to the actual targets but focus on the same threat. Others differ in the types of targets, e.g.:
- malicious ASes
- Spammers
- C&C Server

Efficient blocklists are a valuable product that is often monetized, thus information is limited. Therefore, an analysis of blocklists requires a detailed analysis of listed targets.

The goal of this thesis is to identify different blocklists and analyze their content. The analysis should be done passively based on information from blocklist providers and additional data sources, but also actively based on an analysis of entries. Possible approaches are whether they are routed, reachable or their behavior. Furthermore, the churn of blocklists and differences between lists should be analyzed.

- Identify and analyze different blocklists
- Analyze differences between blocklists
- Analyze the effectiveness/value of blocklists

Caution!

Almost all allocations change over time. Please check regularly to ensure you have the latest version of the DROP lists. They should not be imported into your networks filters and forgotten about. If you do not keep this type of filter data up to date, over time you will eventually encounter problems reaching areas of the Internet if allocations listed in an old version of these lists get reassigned to new networks. Before applying any filters or blocks to your network always carefully consider the ramifications of such filters.

https://www.spamhaus.org/drop/

Your Task

- Basic programming knowledge in Python or Go
- Familiarity with GIYF-Based work approaches

Contact

Johannes Zirngibl  zirngibl@net.in.tum.de
Patrick Sattler  sattler@net.in.tum.de
Markus Sosnowski  sosnowski@net.in.tum.de

https://net.in.tum.de/members/zirngibl/