SinkMiner: Mining Botnet Sinkholes for Fun and Profit

Technical Details/Contributions
The authors developed a tool called SinkMiner which is a forensics system which enables the discovery of previously unknown sinkhole IPs and the related sinkholed domains by efficiently mining large passive DNS databases. SinkMiner’s ability to locate this sinkhole domains is based on the authors observation that sinkhole operators often relocate C&C domains from a sinkhole IP to another. Based on this observation, given an initial seed of known sinkhole IPs, SinkMiner finds new sinkhole IPs by performing historical queries against a passive DNS database, searching for IPs that are related to the known sinkhole IPs by association (that is they at one point modified their DNS records to point to a known sinkhole IP).

SinkMiner divides these associated IP addresses into the following categories:

1. New Sinkhole: IP addresses owned by security operators and used for the purposes of taking over and/or studying botnets
2. Parking: Parking IPs are typically used as traffic vacuums. Used when a domain expires etc.
3. NX-Domain Rewriting: Domains used when an ISP receives a request for a non-existent domain.

The techniques used by the authors to label these IPs is very simplistic, they mainly just look for “sinkhole” and “parking” in the name server name.

Criticism
The author’s technique for labeling involves very basic string matching on name server names, such as looking for the word “park” and when found labeling that as a parking domain. They also look for the word “sink” to indicate a sinkhole. They state though, that sinkhole domains are often a secret, if I was a sinkhole operator attempting to keep my sinkhole domain a secret, I would not be naming it sinkhole.
Doability/Repeatability
If there exists an open source passive DNS database, I could do this work, and this would also allow this work to be repeated.

Possible Extensions
Strengthening the initial labeling phase would be a good extension. For example, perhaps there is a database of parking domains in existence that could be used for more accurate labeling.

Terms
Passive DNS Database

Interesting References