Short Paper Review of:
Peeking Into Spammer Behavior From a Unique Vantage Point

Citation:

Likes:

• I found the spam collection methods to be simple and easy to understand. While there may be other methods I believe that the authors approach of an open relay is the simplest when compared to sinkhole's, getting provider data, etc. I have some legal questions (see dislikes).

• You really can't get any simpler then an smtp open relay server. These are easy to setup and from my experience you'll be relaying junk/spam in no time. (See second dislike)

• I like the simple approach to identifying spam campaigns.

• I liked the hybrid method of an open-relay sinkhole, this mitigates the legal issues while still giving you most of the relavent data that you wouldn't get with just a sinkhole. The problem again however exists in the way the spammers test the relays. If you don't let through the port scan requests, etc. then the relay scanners will never detect you which means you won't get any spam to relay in the first place.

• The spam source classification was straight forward and the Type I / Type II approach was easy to follow.

Dislikes:

• To start this section, I'm curious to know what the legal implications are if you knowingly setup an open relay knowing full well that it's going to relay spam. I would have liked the authors to have addressed this is some way.

• I know of many people who have been contacted by their ISP's, including businesses, and told not to run an open smtp relay due to the volume of traffic identified as spam being relayed. In some cases (global crossing comes to mind) even the tier 1 service providers have statements in their service agreements that say you can't do this.
• The coordination of the LVS was more guesswork than actual knowledge, while the steps most likely are correct, the order of the approach may very greatly.

• Again I have the usual complaint about time frame for collection. The system could be completely automated even the statistical analysis and graphing so why didn't the authors run it for much longer. For that matter why wasn't it run on multiple disparate systems/IP's in physically/logically separate locations for better/larger dataset gathering.

Repeat / Add to Work?

• I could definitely repeat this work. I would however add known spam blacklists to the system to make it look more realistic to the relay testing software and to get rid of the known results. This would leave me with new unknown/unblacklisted results which could be shared easily. I am very interested in repeating this work, out of this first week of article reviews, this is by far my highest pick.

Synopsis:

• Overall I found the article to be interesting yet simple in approach. I would definitely be interested in repeating the work with some small changes and additions.