Building a phishing net

A short exercise
About us

University of Oslo:

- Approximately 70k users
- Faculty, staff, students, others

E-mail system based on open source products:

- Exim/Postfix, SpamAssassin, ClamAV, Cyrus, Nginx, Keepalived, Sympa etc.
- Last summer a dramatic decrease in spam volume
  ➔ phishing volume not affected
Relevant statistics

- Monthly average of 75++ users contacting about phishing
- 0-4 compromised accounts
Fed up

Common observation:

- Vast majority of end users do not understand why these phishing messages can't be intercepted.

Question:

- Well, can they?

Common observation:

- Phishes resemble each other, with less than 10 basic layouts.

Question:

- How much would a small custom made filter catch?
A first approach

Prevent password leakage

- Small perl script hooked into exim
  - Scans first part of message for possible passwords
  - Words satisfying complex passwords requirements
  - Authenticates envelope sender with passwords
- Works, not false positives
- Only internal messages
  - Need local user name
- Fewer and fewer phishes rely on reply-to
Second approach

Prevent phishes from reaching users

• Small perl script, hooked into Exim
• Scans first n bytes of messages after known patterns:
  - Reply-to/Sender addresses
  - Urls in body
  - Specific combinations of words
• < 5 hours development to working copy
• False positives
  - Automated messages: blogs, site passwords, registration notices
  - Building a whitelist based on results
Results – then what

- Just log
  - ✔ Fast, non-intrusive
  - ✗ No impact
- Send a warning to recipients
  - ✗ Experience tells us it doesn't help
  - ✗ Users read their messages chronologically
  - ✗ Users pay more attention to the scary message from their friendly phishers than to our warning.
- Reject
  - ✗ False positives are mostly automated messages = no one will ever notice that they weren't delivered
  -> Likely violation of first rule of e-mail administration: always deliver legitimate messages!
... then what #2

- Quarantine for inspections by admin
  - Privacy-issues
  - Scalability-issues
- Quarantine for inspections by recipient
  - Non-intrusive
    - Needs a end user framework
    - More work, but doable
Measuring the impact

Monthly average

- \(~7k\) number of phishing messages intercepted
  - May has been a lot slower – \(~1.5k\)
- 15 user reports of phishing attacks (12 unique)
- 1 compromised account after filter (in 2 months)
Other solutions?

Yes, of course!

- Black lists stop some
- AV stops some
- Spam scanners detect some
  - still delivered = problematic
- Antispam appliances such as Barracuda and Ironport both detect a fair amount and has a built-in quarantine functionality
  - Expensive
  - Black box
Conclusion

It is feasible to build custom made filters with substantial impact while spending relatively little time and money.

They won't catch everything, but neither do products costing €€€. When running a system suitable for such plugin extensions, it may very well be a task worth doing.