Poster: Apate: Anti-Phishing Analysing and Triaging Environment

Elmer Lastdrager
Faculty of Electrical Engineering, Mathematics and Computer Science
University of Twente
Enschede, The Netherlands
Email: e.e.h.lastdrager@utwente.nl

Pieter Hartel
Faculty of Electrical Engineering, Mathematics and Computer Science
University of Twente
Enschede, The Netherlands
Email: p.hartel@utwente.nl

Marianne Junger
School of Management and Governance
University of Twente
Enschede, The Netherlands
Email: m.junger@utwente.nl

I. Introduction

Phishing is a scalable act of deception whereby impersonation is used to obtain information from a target [1]. Phishing emails are a threat to the general population, since 89% of all residents of the Netherlands, aged 12-74, reported having used email in the past three months [2]. Many methods to reduce phishing have been proposed and deployed. These can be categorised as social or technical. Firstly, social countermeasures against phishing, such as training [3], [4], aim to improve users' decisions. Secondly, technical solutions, such as blacklists or filtering [5], try to prevent phishing emails from reaching a user's inbox. However, blacklists and filtering techniques are not perfect and users occasionally receive phishing emails in their inbox.

The companies whose name has been abused in phishing attacks differ per country [6]. Particularly, countries where English is not the native language, such as the Netherlands, require customised phishing emails. We investigated the characteristics of phishing emails that were received by citizens of the Netherlands, both from a sender and from a receiver perspective. To do this, we used an email corpus that was provided by the Dutch anti-fraud agency Fraud Helpdesk [7]. We were granted access to both the historical archive (since 2013) and the live stream of incoming emails.

We developed a system for analysing reported phishing emails, called Apate (Anti-Phishing Analysing and Triaging Environment). Apate was deployed in November 2014 and has since been analysing all emails in the archive, as well as the ones that are currently being forwarded to the Fraud Helpdesk. Potential victims of a phishing and/or malware email are encouraged to forward received emails to the Fraud Helpdesk. At the moment, the resulting dataset contains 143,674 emails that were received since January 2013.

II. Design

Apate was build in Python 3 and consists of a basic framework that imports emails from the corpus, hashes each email using SHA256 and stores the original source on disk, while putting meta data in a PostgreSQL database. The actual functionality is provided by a series of modules that extract information from the email or enrich existing meta data.

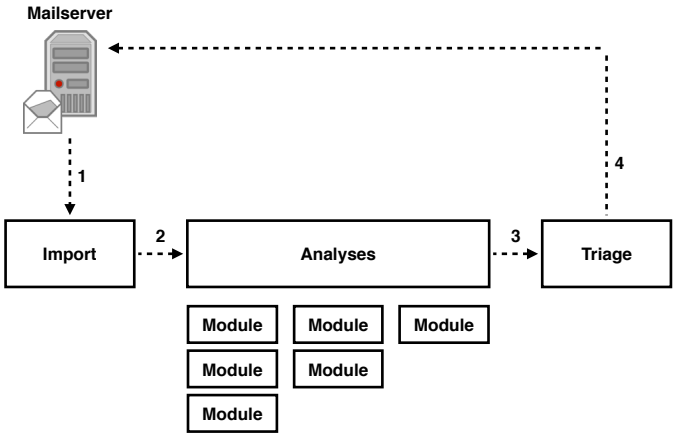

Fig. 1. Flow of a single email through the system. (1) fetching from email server. (2) start analyses. (3) mark as analysed. (4) send feedback to user.

Finally, the triager of Apate uses the information from the modules to decide whether an email is phishing, contains malware or is something else. The resulting label (phishing, malware, unknown) will be used to send feedback to the user who forwarded the email. Unknown emails can be regular spam, false positives (the user forwarded a legitimate email) or unrecognised phishing. To reduce the workload, automatically sent emails such as out-of-office messages are discarded.

The contents of each forwarded email were processed to extract information from both plain text and HTML emails, as well as extracting information about the reporter. URLs were extracted from the forwarded emails and subsequently analysed, whereby the domain was extracted and resolved. All URLs were sent for analysis to VirusTotal [10], where they were checked against 60 blacklists. A URL was marked as suspicious in our results, if it was listed in at least one blacklist. The attachments of reported emails were also scanned using VirusTotal [10]. At the moment of scanning the dataset, 55 different anti-virus products were in use by VirusTotal. For our analysis, a file is classified as a virus when at least one of the anti-virus solutions marks the file as infected.

An annotated list of first names [9] was used to identify

the gender of the submitter of an email. If a given name was tagged differently in several languages, the Dutch gender-name association was used. We extracted the sender information out of the name part in the "From" header of the email.

III. Results

The analysis lead to three lessons learned:

1) Potential victims open their email mostly during office hours, with a peak on Monday morning.
2) Users express doubts about their decision to flag an email as phishing. Additionally, some explicitly ask for an expert judgement.
3) Phishing emails are becoming more sophisticated. Phishers include details such as reference numbers, disclaimers, and phone numbers.

Other findings include:

- 41% of the reports originated from an email address that customers get from their internet service provider together with their internet access. A further 22% originates from the large international email providers (Microsoft, Google, Apple, Yahoo). The remaining reporters used various addresses, for example from the company they work for or from their own domain.
- Attempting to determine the gender of the reporters based on their name gave results for 57.3% of the reports. Out of the successful gender identifications, 57.2% were male, 40.2% female and 2.6% could be either of the two.
- 0.5% of the emails were sent automatically and 7.0% of the emails were not forwarded emails, but spam or other communication. The remaining 92.5% were further analysed.
- In order to find patterns in reporting and therefore understand when potential victims process their mail, the day and time of receiving the forwarded email is plotted in Fig. 2. Emails are forwarded mostly during office hours.
- Most of the emails (75.5%) contain at least one URL, totalling to 146,669 URLs. These URLs are hosted on 20,854 unique domains and 484 unique IP addresses (i.e. the URL contains the IP address). The most common top-level domain is .com (41%), followed by .nl (17%) and .net (5%).
- 45% of the forwarded emails contain a comment from the user who reports it. Almost all of these comments are greetings, such as "Regards, John". Some user comments demonstrate that they are agitated from having received a phishing email. Other users included questions or even their address details.
- The language of the phishing email itself was mostly Dutch (78%), followed by English (16%), German (3%) and other languages (3%).
- Classifying emails using blacklists resulted in 38.2% of the emails being categorised as phishing for containing at least one blacklisted URL. 17% of all email attachments were classified as virus. The remaining 54.0% of the emails could not be labelled.

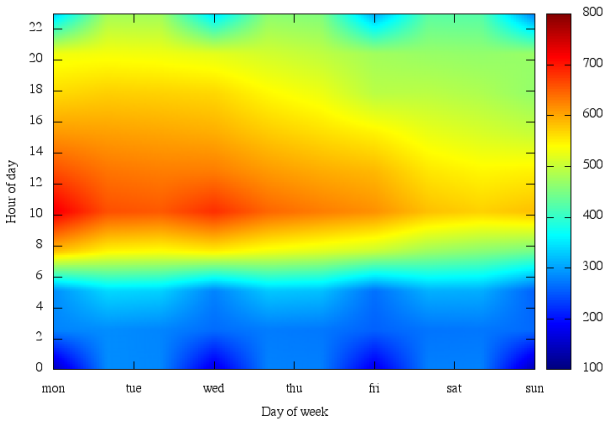

Fig. 2. Distribution of received emails by the Fraud Helpdesk per day (horizontal) and time (vertical). The graph has been smoothed.

IV. Acknowledgements

The authors would like to thank Fleur van Eck, John Kellij and Jos Kerssen from the Fraud Helpdesk. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 318003 (TREsPASS). This publication reflects only the authors views and the Union is not liable for any use that may be made of the information contained herein.

References

[1] E. E. H. Lastdrager, "Achieving a Consensual Definition of Phishing Based on a Systematic Review of the Literature," *Crime Science*, vol. 3, no. 9, 2014.
[2] Statistics Netherlands. (2013) ICT usage within the Netherlands. [Online]. Available: http://statline.cbs.nl/Statweb/
[3] S. Sheng, B. Magnien, P. Kumaraguru, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, "Anti-Phishing Phil," in *Proceedings of the 3rd symposium on Usable privacy and security - SOUPS '07*. New York, New York, USA: ACM Press, 2007, p. 88.
[4] P. Kumaraguru, J. Cranshaw, A. Acquisti, L. Cranor, J. Hong, M. A. Blair, and T. Pham, "School of phish: a real-world evaluation of anti-phishing training," in *Proceedings of the 5th Symposium on Usable Privacy and Security - SOUPS '09*. New York, New York, USA: ACM Press, 2009, p. 1.
[5] M. Khonji, Y. Iraqi, and A. Jones, "Phishing Detection: A Literature Survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 2091–2121, 2013.
[6] Kaspersky. (2013) The evolution of phishing attacks: 2011-2013. [Online]. Available: http://media.kaspersky.com/pdf/Kaspersky_Lab_KSN_report_The_Evolution_of_Phishing_Attacks_2011-2013.pdf
[7] Fraud Helpdesk. (2015) Fraud helpdesk: The dutch national anti-fraud hotlinedesk. [Online]. Available: http://www.fraudehelpdesk.nl
[8] A. Savand, A. Swartz, Y. Barkan, A. Musayev, M. Cepl, S. Rivera, and I. Gromov. (2014, November) html2text. [Online]. Available: https://alir3z4.github.io/html2text/
[9] J. Michael, "40000 namen, anredebestimmung anhand des vornamens," *c't magazin für computer technik*, no. 17, p. 182, 2007.
[10] VirusTotal. (2014, November). [Online]. Available: http://www.virustotal.com
[11] T. French. (2014, November) Netherlands broadband overview. [Online]. Available: http://point-topic.com/free-analysis/netherlands-broadband-overview/