Luigi Gallo

Tutor: Prof. Alessio Botta

XXXIV Cycle - II year presentation

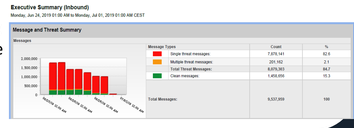
Identifying threats in a large company's inbox

Motivation

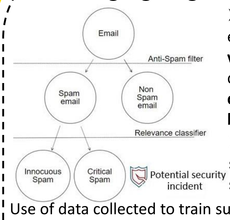
1)Email is the **most used channel** for making cyber attacks.
2)Email attacks are **the primary infection vector** in 78% of cyber espionage incidents.
3)(Spear) Phishing, (CEO) financial fraud and malware propagation with emails are **increasing in number and in malignance** ($1.8 billions of monetary losses in USA in 2019).

Context and challenges

1) People increasingly publish **personal information**, typically used to make email attacks **trustworthy and captivating**.
2) Email attacks are very sophisticated and mingle with **a lot of noise** (marketing, advertising, errors, newsletters, sex photos etc.).
3) In large companies, the number of employees who may **fall victim of phishing or download malware** is considerable.
4) The number of spam emails, among which the attacks are hidden, **is huge**!!



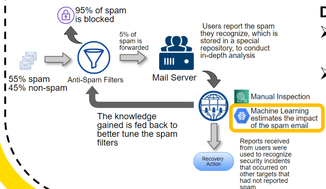
IDEA : highlighting "the needle in the haystack"



- A system that analyses spam emails and on the basis of **technical, visual and cognitive properties**, determines which of them **requires defensive action to prevent security breaches**. (ideally, acts automatically)
- A framework that **collects spam emails and the actions of security analysts**.

Use of data collected to train supervised machine learning models, obtaining **automatic classifiers** to support analysts and **specific guidelines** on how to design **effective awareness campaigns**

The life cycle of a spam email in the company

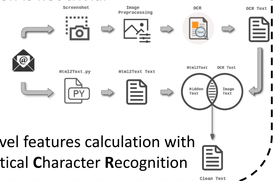


Dataset collected:
- **21.932 real spam** email
- **3.931** of them have created a security incident or at least required a defensive action to prevent future infections.

Feature set design

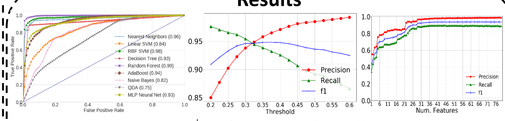
Feature rationale: have a discriminating power to verify two necessary conditions as long as a security incident occurs
- **The recipient is deceived by the email**
- **The "payload" of deception is not trivial**

The full set of features comprises **79 features** grouped in 8 *feature field*: 1)General; 2)Content; 3)View; 4)Contet_view; 5)Subject; 6)Attachments; 7)Links; 8)Other



Novel features calculation with **O**ptical **C**haracter **R**ecognition

Results



The best ML algorithms are **Random Forest** and Support Vector Machine (RBF Kernel)

Performance achieved
F1-Score: 94,8%
Recall: 93.4%
Precision: 96.3%

Selecting the best **36 features** is enough to capture the problem!

Contacts
Email: luigi.gallo3@unina.it
Telephone: +39 335-791-9892

Comics research group
http://comics.unina.it/

Participations (H2020 project)
CONCORDIA
Cyber security cOmpeteNCe fOr Research anD innovAtion

Collaborations
TIM

Future steps
- Better understand the phenomenon of phishing using the results from the feature ranking evaluation process, in order to design an **awareness campaign** to train company employees on the specific cognitive vulnerabilities they have shown in the data (an empirical large experiment with 40.000+ people is in roadmap)
- Testing the robustness of the model in **adversarial environments** (Evasion and Poisoning attacks). Answer to the following question: given a phishing email that has obtained a high success rate (i.e. has very good chances of misleading a human), if we perturb a little the features such that it is no longer considered relevant by the automatic classifier, does it keep its effectiveness on human minds?

References
[1] **Luigi Gallo**, Alessio Botta, and Giorgio Ventre. 2019. Identifying Threats in a Large Company's Inbox. In Proceedings of the 3rd ACM CoNEXT Workshop on Big DAta, Machine Learning and Artificial Intelligence for Data Communication Networks (Big-DAMA '19, Orlando USA)
[2] Emmanuel Gbenga Dada, Joseph Stephen Bassi, Haruna Chiroma, Shafi'i Muhammad Abdulhamid, Adebayo Olusola Adetunmbi, and Opeyemi Emmanuel Ajibuwa. 2019. Machine learning for email spam filtering: review, approaches and open research problems. Heliyon 5, 6 (2019)
[3] Enrico Blanzieri and Anton Bryl. 2008. A survey of learning-based techniques of email spam filtering. Artificial Intelligence Review 29, 1 (01 Mar 2008), 63–92.

UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

Ph.D
INFORMATION TECHNOLOGY
ELECTRICAL ENGINEERING