PROBLEM STATEMENT FOR WEB PHISHING DETECTION

PROBLEM

Phishing is a cyber attack that uses email as its method of attack. The term “phish” is a reference to the act of fishing, throwing a hook and hoping to catch something juicy. The objective is for the recipient to believe the message is legitimate and to click a link, open an attachment. Malicious links will lead to a website that often steals login credentials or financial information like passwords, account IDs or credit card details.

BACKGROUND

The Internet is a network of computers filled with valuable data, so there are many security mechanisms in place to protect that data. But there's a weakest link: the human. If the user freely gives away their personal data or access to their computer, it's much harder for security mechanisms to protect their data and devices. A phishing attack typically starts with an email that claims to be from a legitimate website, like a banking website or online store. The goal of the email is to obtain private data from the user, so it either asks the recipient to reply with personal information or it links to a website that looks remarkably like the original site. If the user is convinced and enters private details on the site, that data is now in the hands of the attacker! If the user filled in login details, they can then use those credentials to log in to the real website, or if the user provided credit card details, they can use the credit card to make purchases anywhere. Otherwise phishing was responsible for more than 80% of reported security incidents. It's one of the most common vectors for ransomware, which encrypts data and renders computers useless. An analysis of more than 55 million emails by cloud security provider Avanan found that one email in 99 contains a phishing attack. The 2020 Verizon Data Breach Investigations Report found that 22% of all data breaches involved phishing, and dark web monitoring firm ID Agent estimates that phishing attacks have increased more than 600% since the start of the COVID-19 pandemic.

RELEVANCE

Phishers can use public sources of information to gather background information about the victim's personal and work history, interests and activities. Typically through social networks like LinkedIn, Facebook and Twitter. These sources are normally used to uncover information such as names, job titles and email addresses of potential victims. This information can then be used to craft a believable email. Typically, a victim receives a message that appears to have been sent by a known contact or organization. The attack is then carried out either through a malicious file attachment, or through links connecting to malicious websites. In either case, the objective is to install malware on the user's device or direct the victim to a fake website. Fake websites are set up to trick victims into divulging personal and financial information, such as passwords, account IDs or credit card details.

OBJECTIVE

The vast majority of the time, the purpose of a phishing attack is to steal data, money—or both.

Data—The type of data that cybercriminals are most often interested in are usernames and passwords, identity information (e.g., social security numbers), and financial data (e.g., credit card numbers or bank account information). Login credentials can be used to breach the victim’s systems to steal intellectual property or inject malware for other malicious purposes. Data of any kind can also be monetized by selling it on the dark web to other criminals.

Money—If the intent of the phishing attack is to steal money, the cybercriminals may send a fake invoice, try to convince the victim to wire money, or ask the victim to input financial account information into a fake website.

CONCLUSION

Phishing attacks remain one of the major threats to individuals and organizations to date. The main reason is the lack of awareness of users. Preventing these huge costs can start with making people conscious in addition to building strong security mechanisms which are able to detect and prevent phishing domains from reaching the user.