Mitigating the Threat of Phishing Attacks with Personalized Context-Aware Learning

Bindu Thota

W210 Capstone Proposal
Fall ‘23

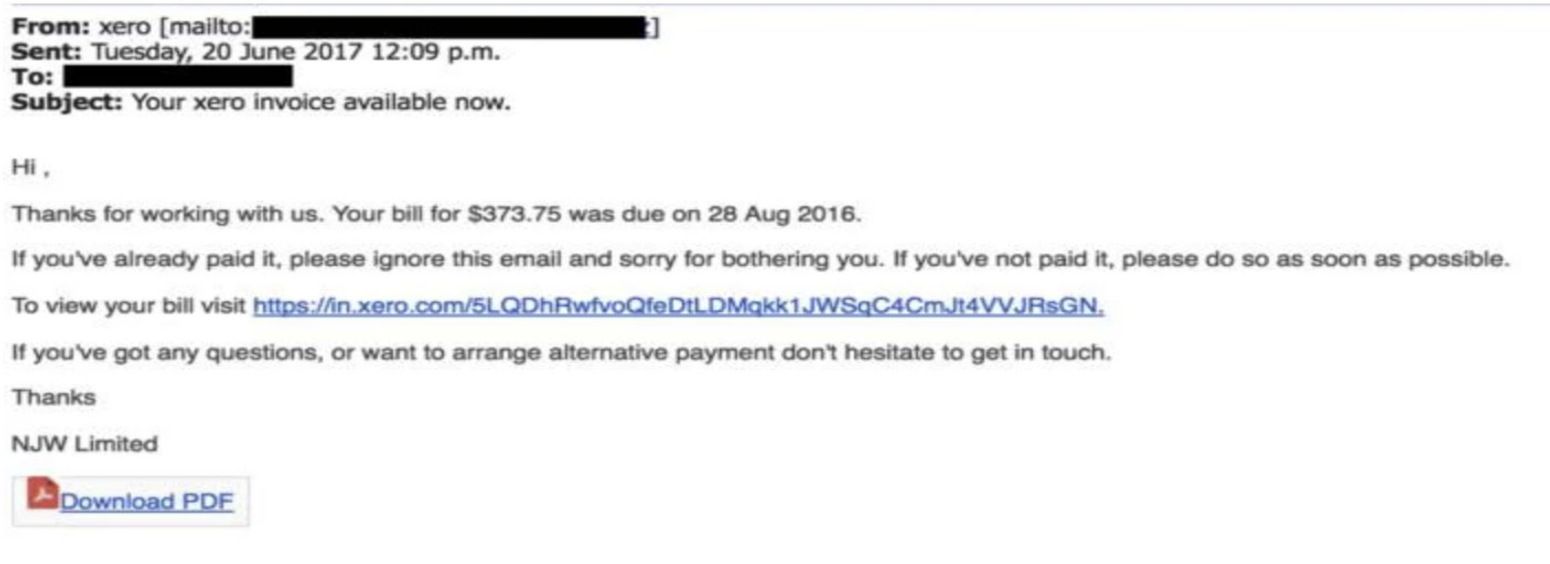
Berkeley
UNIVERSITY OF CALIFORNIA

What is a Phishing Attack?

Invoke fear or urgency to retrieve confidential information

Fake Invoice

Source:
https://blog.usecure.io/the-most-common-examples-of-a-phishing-email#The-fake-invoice-scam



From: xero [mailto:[redacted]]
Sent: Tuesday, 20 June 2017 12:09 p.m.
To: [redacted]
Subject: Your xero invoice available now.

Hi ,

Thanks for working with us. Your bill for $373.75 was due on 28 Aug 2016.

If you've already paid it, please ignore this email and sorry for bothering you. If you've not paid it, please do so as soon as possible.

To view your bill visit https://in.xero.com/5LQDhRwfvoQfeDtLDMqkk1JWSqC4CmJt4VVJRsGN.

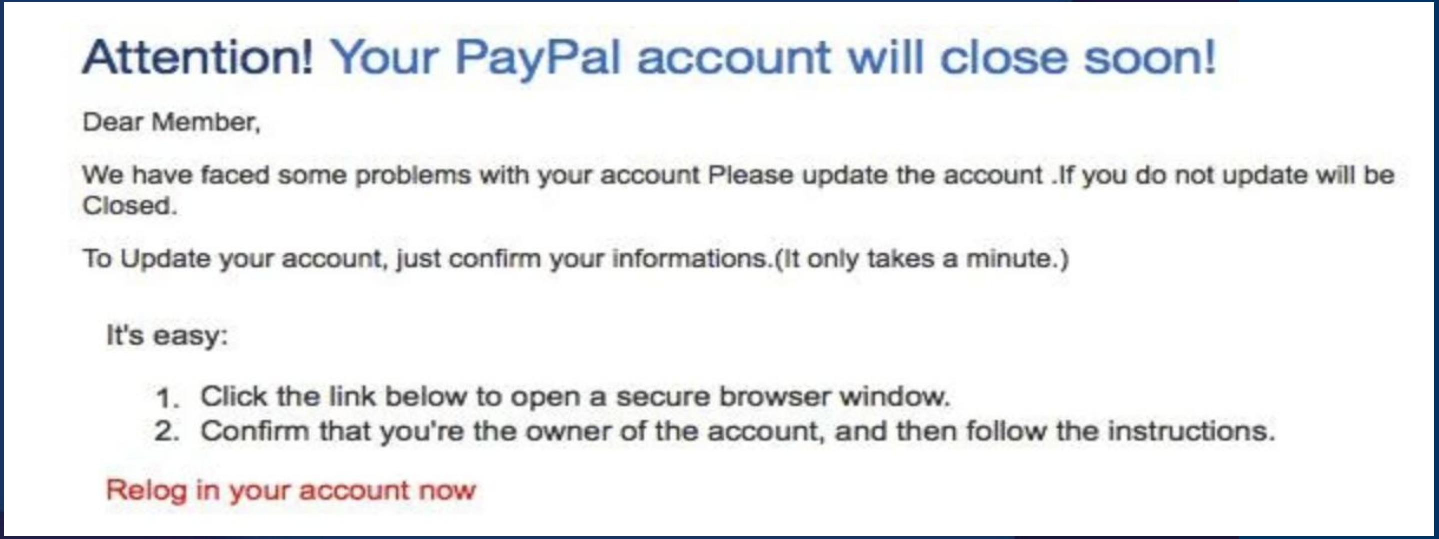
If you've got any questions, or want to arrange alternative payment don't hesitate to get in touch.

Thanks

NJW Limited

Impersonating PayPal

Source: https://blog.usecure.io/the-most-common-examples-of-a-phishing-email#The-fake-invoice-scam



Attention! Your PayPal account will close soon!

Dear Member,

We have faced some problems with your account Please update the account .If you do not update will be Closed.

To Update your account, just confirm your informations.(It only takes a minute.)

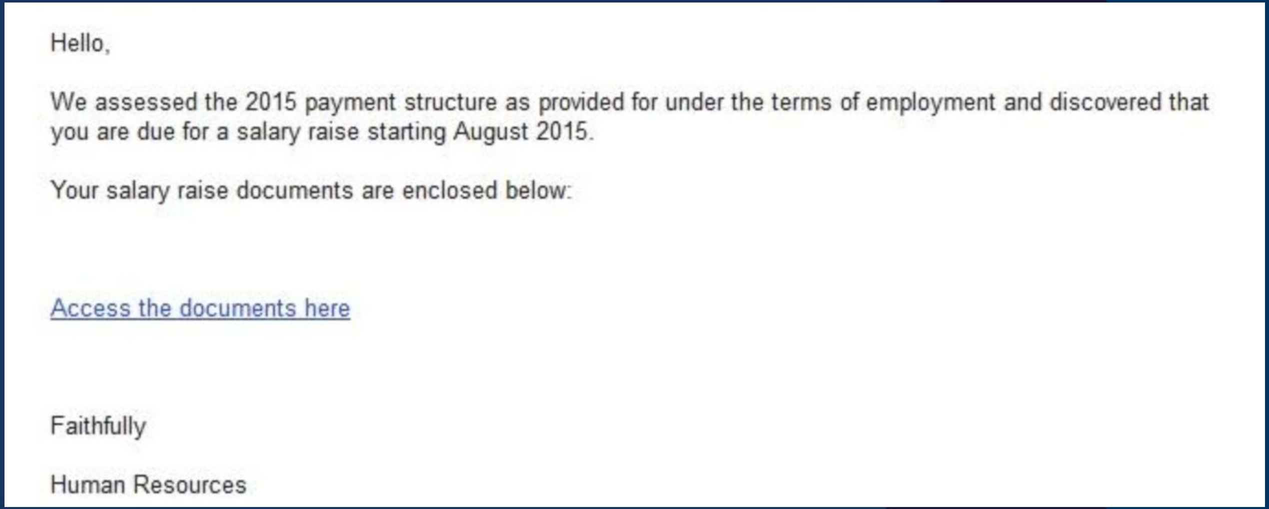
It's easy:

1. Click the link below to open a secure browser window.
2. Confirm that you're the owner of the account, and then follow the instructions.

Relog in your account now

Fake Message from HR

Source:
https://blog.usecure.io/the-most-common-examples-of-a-phishing-email#The-fake-invoice-scam



Hello,

We assessed the 2015 payment structure as provided for under the terms of employment and discovered that you are due for a salary raise starting August 2015.

Your salary raise documents are enclosed below:

Access the documents here

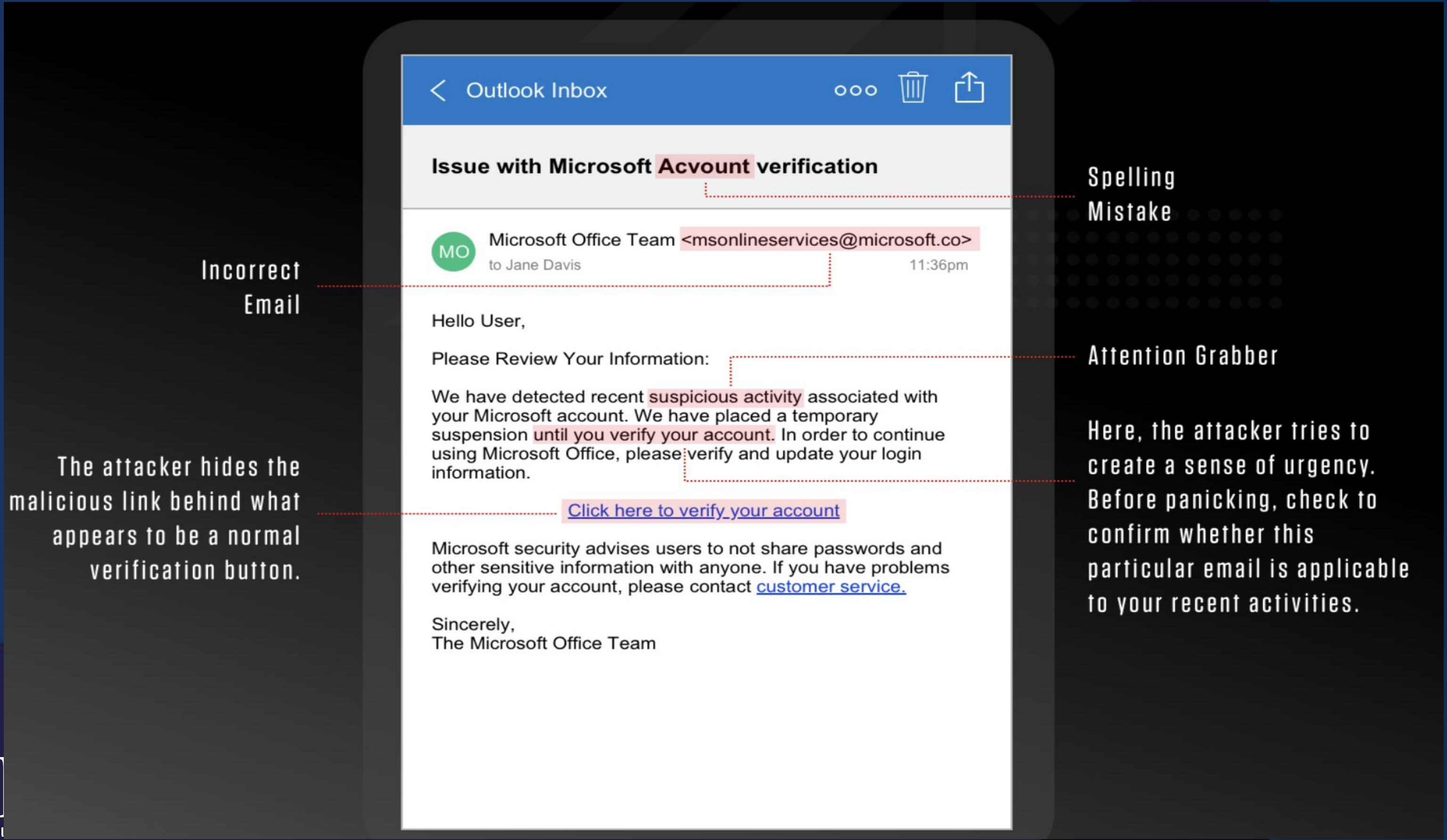
Faithfully

Human Resources

Berkeley
UNIVERSITY OF CALIFORNIA

How to Spot a Phishing Email

Source: https://www.crowdstrike.com/cybersecurity-101/phishing/how-to-spot-a-phishing-email/



Issue with Microsoft Acvount verification

MO Microsoft Office Team <msonlineservices@microsoft.co>
to Jane Davis 11:36pm

Hello User,

Please Review Your Information:

We have detected recent suspicious activity associated with your Microsoft account. We have placed a temporary suspension until you verify your account. In order to continue using Microsoft Office, please verify and update your login information.

Click here to verify your account

Microsoft security advises users to not share passwords and other sensitive information with anyone. If you have problems verifying your account, please contact customer service.

Sincerely,
The Microsoft Office Team

Spelling Mistake

Incorrect Email

Attention Grabber

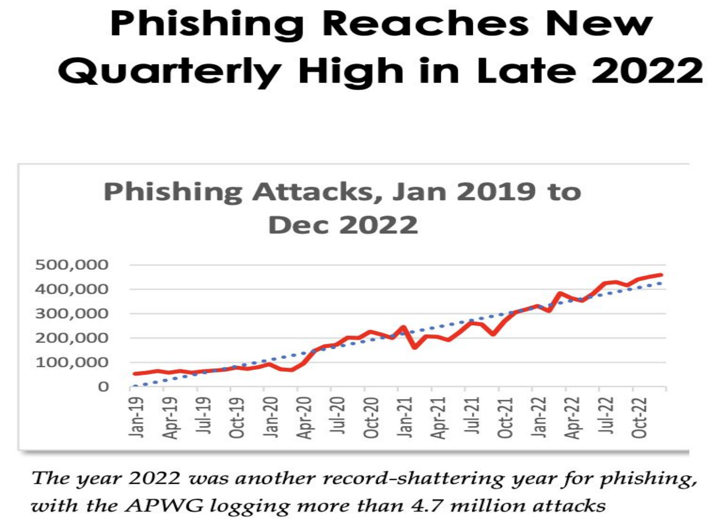
The attacker hides the malicious link behind what appears to be a normal verification button.

Here, the attacker tries to create a sense of urgency. Before panicking, check to confirm whether this particular email is applicable to your recent activities.

Impact

Source: APWG.org

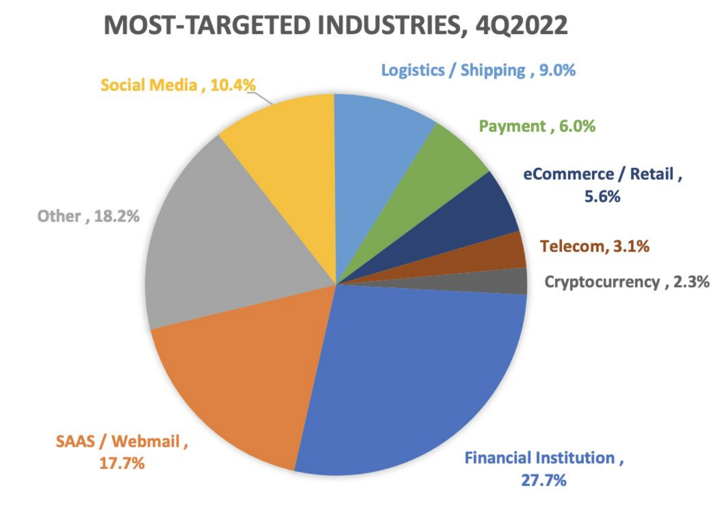
FBI estimated victims reporting over $52 million in losses in 2022



Phishing Reaches New Quarterly High in Late 2022

Phishing Attacks, Jan 2019 to Dec 2022

The year 2022 was another record-shattering year for phishing, with the APWG logging more than 4.7 million attacks



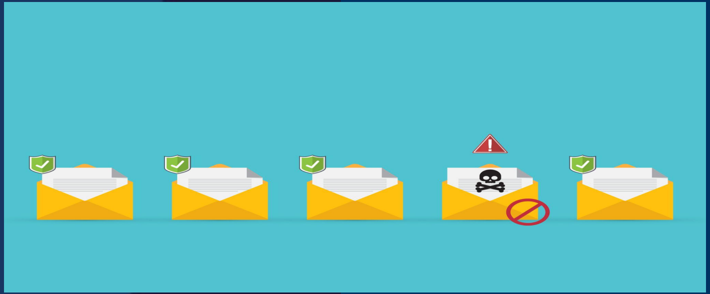
MOST-TARGETED INDUSTRIES, 4Q2022

State-of-The-Art

- Anti-phishing filters have evolved to use ML/AI and deep learning and NLP in recent years
- Metric
  - Accuracy

What is the problem?

User still needs to validate manually to avoid false positives and false negatives



Toy Example

You won the lottery!

Real ?

False Positive ?

False Negative ?

Toy Example

Msg: you to CA Lottery on Aug 1: CA Lottery Commision, I'd like to buy a lottery ticket

Msg: CA Lottery to you on Aug 2: Here's your ticket

Mag: CA Lottery to you on Sept 1: You won the lottery!

Personalized Context-Aware Learning

Document Extraction for generating context-aware embeddings

Options:

K-nearest neighbors

Neural Models

→

Phish Prediction

Fine-tuned BERT, GPT models

Data Sources:

- APWG Phishing Repository
- PhishTank
- PhishLoad
- Nazario
- For context based data, we might need to generate and curate data.
  - This generation can probably be done using LLMs.

Thank You!

Berkeley
UNIVERSITY OF CALIFORNIA