RSAConference2016
San Francisco | February 29 – March 4 | Moscone Center

Connect to Protect

SESSION ID: HUM-W02

Trends in Social Engineering: How to Detect and Quantify Persuasion

Markus Jakobsson
CTO
ZapFraud Inc
@JakobssonMarkus

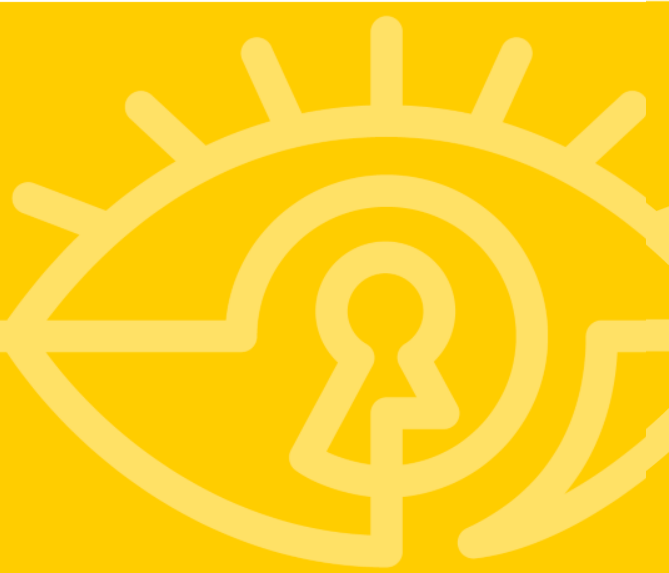
#RSAC

My collaborators

- Ana Ferreira, Cintesis, University of Porto
- Damon McCoy, NYU
- Elaine Shi, Cornell University
- Youngsam Park, University of Maryland + ZapFraud
- Ting-Fang Yen, DataVisor
- Arthur Jakobsson

Overview

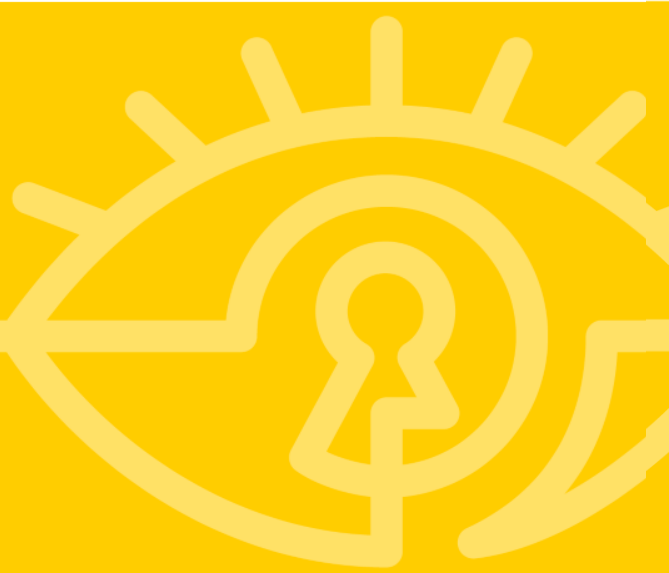
- Scams and persuasion – why we care
- Data sampling and datasets
- Scam trends – what is happening?
- Persuasion – how and why does it work?
- Case study – Business Email Compromise (BEC) examples
- Using insights into persuasion to improve filtering
- Action items / recommendations



Datasets

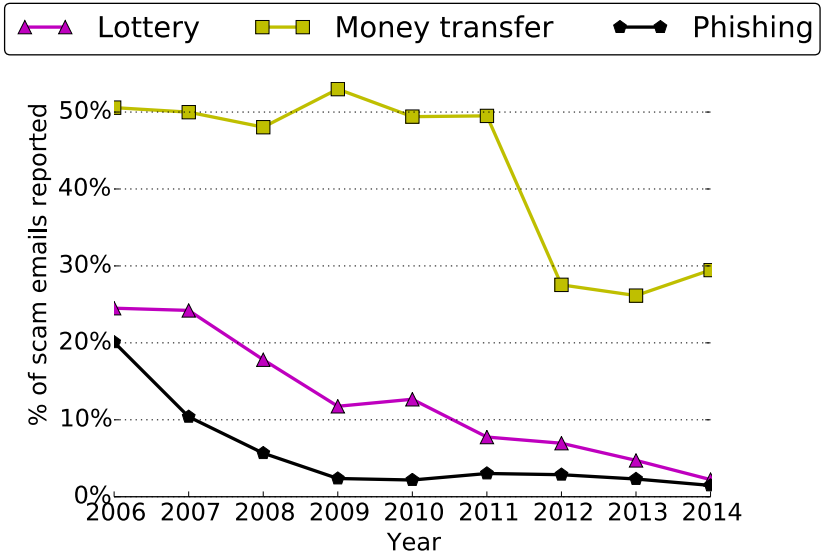
Selecting (reasonably) unbiased datasets

- Complaints from people who lost money (FBI/IC3)
  - Not everybody who loses money files complaints
- Submissions to scam reporting websites
  - Maybe mostly “average sneaky” scams are submitted?
- Spam benchmark dataset (“untroubled” spam archive)
  - Reflects what spam filters blocked, not what people received
- Ham datasets
  - Enron, Jeb Bush, subscriber inboxes, Amazon reviews



Trends in 419 Scams

Untargeted in decline – except authority

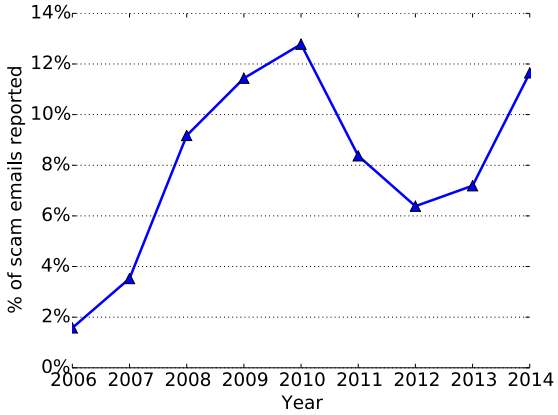


Money transfer scams includes:

- Next of kin scams
- Commodity scams
- Charity/dying person scams
- Widow/orphan/refugee scams

Authority scams include:

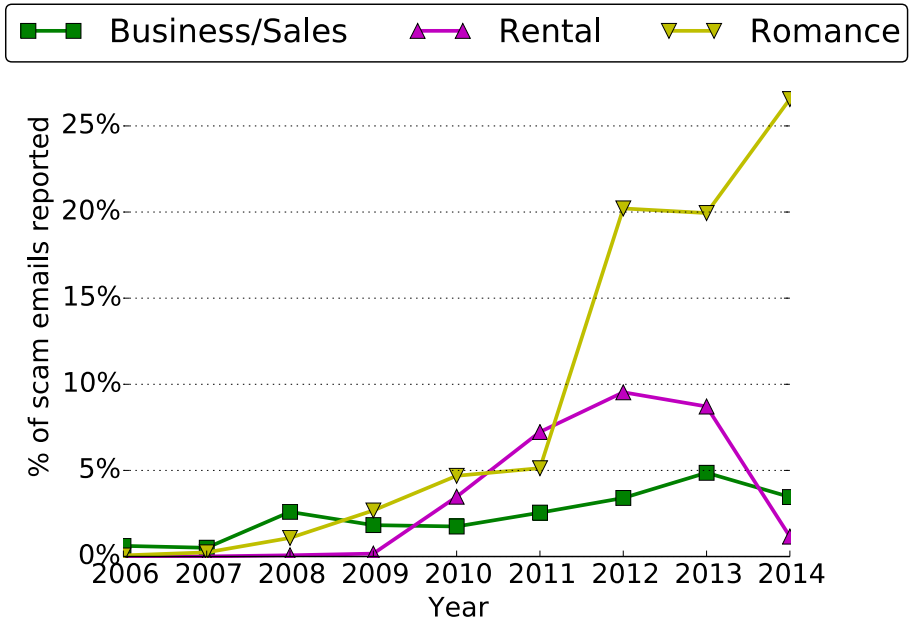
- Government scams
- Bank scams

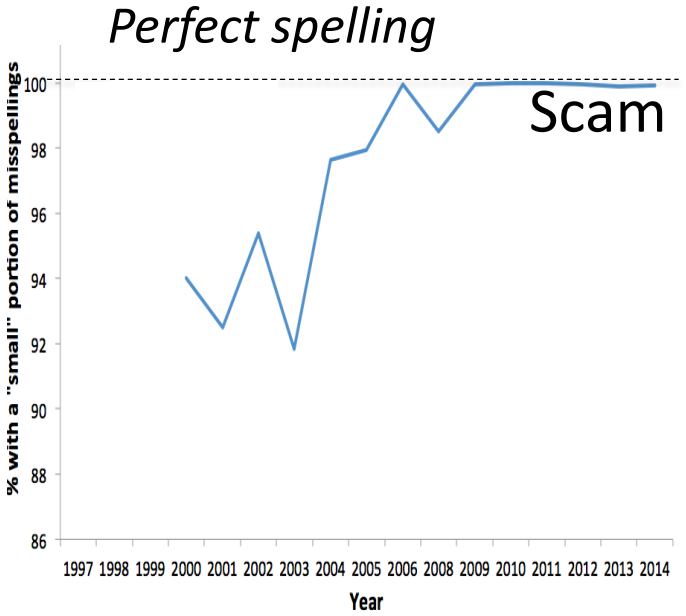


Comparing to FBI/IC3 findings

- ~50% increase 2013-2014
- most common fraud against elderly

Targeting and sophistication *up*

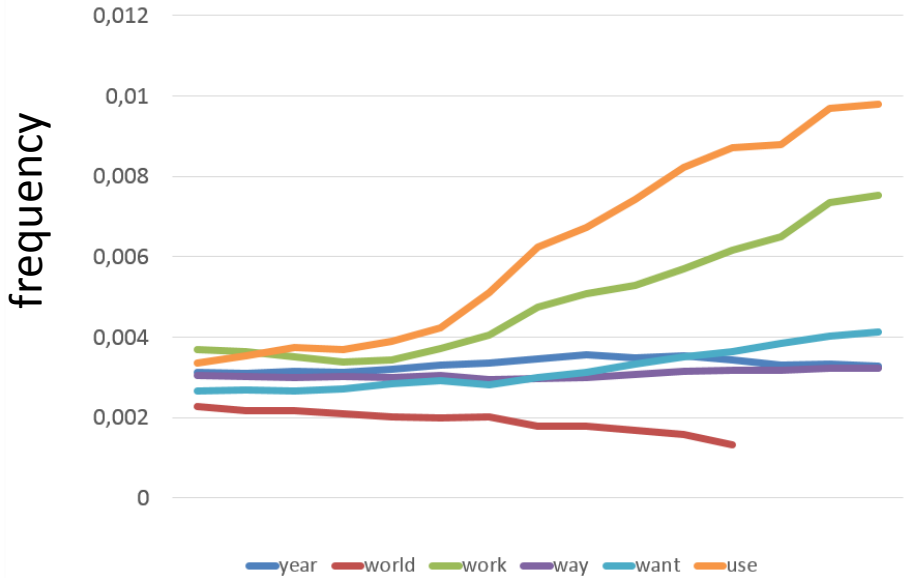




Ham remains the same, scam changes

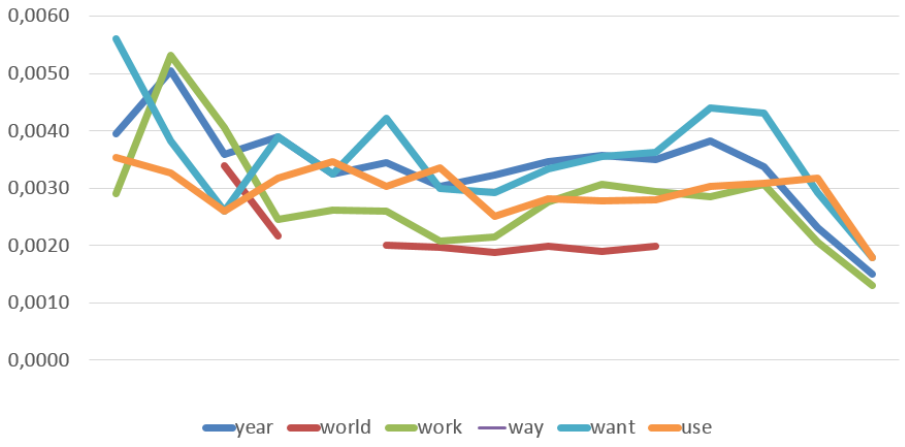
Words commonly used in scams …

… in Amazon reviews …

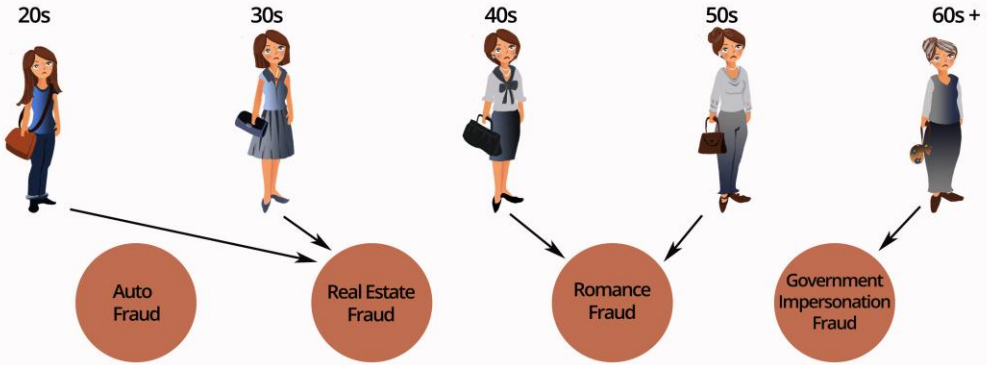


… and in scam messages …

(2000-2014)

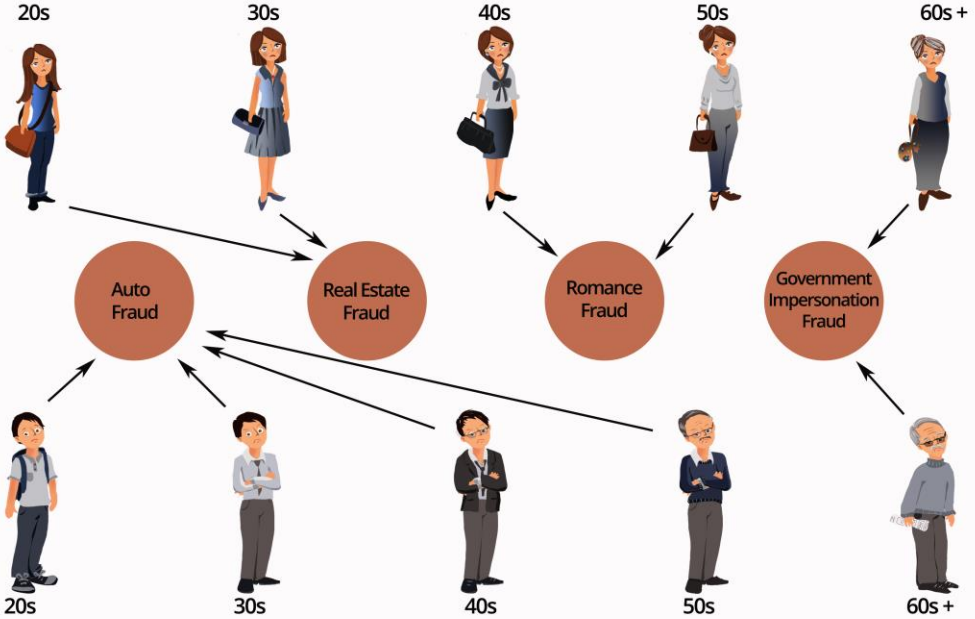


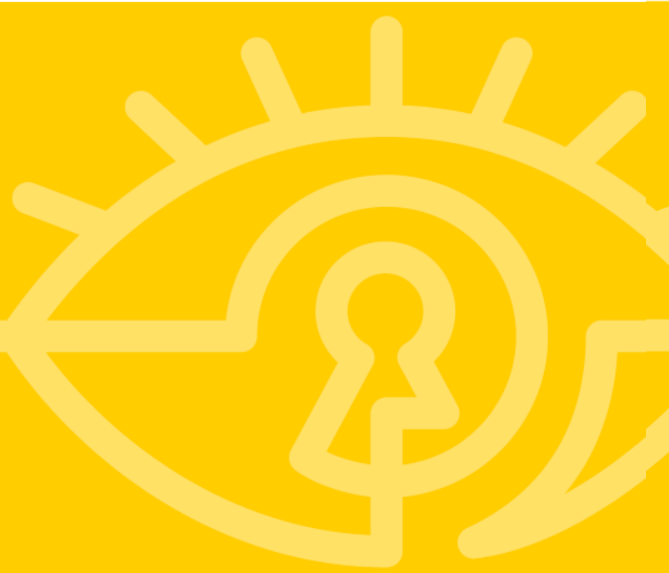
What is persuasive … is personal



Zapfraud.com

What is persuasive … is personal





What is Persuasion?

Persuasion is about convincing arguments

Appeal to greed and opportunism

Appeal to a wish to comply

Appeal to weakness

Appeal to empathy

And just appeal

Persuasion is about structure

What happens makes sense

Persuasion is about knowledge

Gather contextual knowledge from:

- Breaches
- Account take-overs
- Social networks
- Other public sources

Quantifying persuasion

Degree risk perception

Perfection of targeting

Persuasion = Credibility * Fit

Yield ~ persuasion in absence of protection

Measuring credibility

You *cannot* ask “Does this look risky to you?”

Measuring credibility

But you *can* ask
"What type of risk is this *primarily* associated with?"

Example scam to evaluate

You have exceeded your mailbox quota.

Your account will be blocked 8 AM tomorrow unless you request more space. You can request more space by clicking here.

Type of risk is primarily associated with?

- The recipient may get a computer virus.
- The recipient may lose his password.
- This may be a scam aimed at stealing your money.
- There is no risk.
- The recipient may get unwanted advertisements.
- The recipient's account may be blocked if she does not pay attention.

Correct answer

- The recipient may get a computer virus.
- The recipient may lose his password.
- This may be a scam aimed at stealing your money.
- There is no risk.
- The recipient may get unwanted advertisements.
- The recipient's account may be blocked if she does not pay attention.

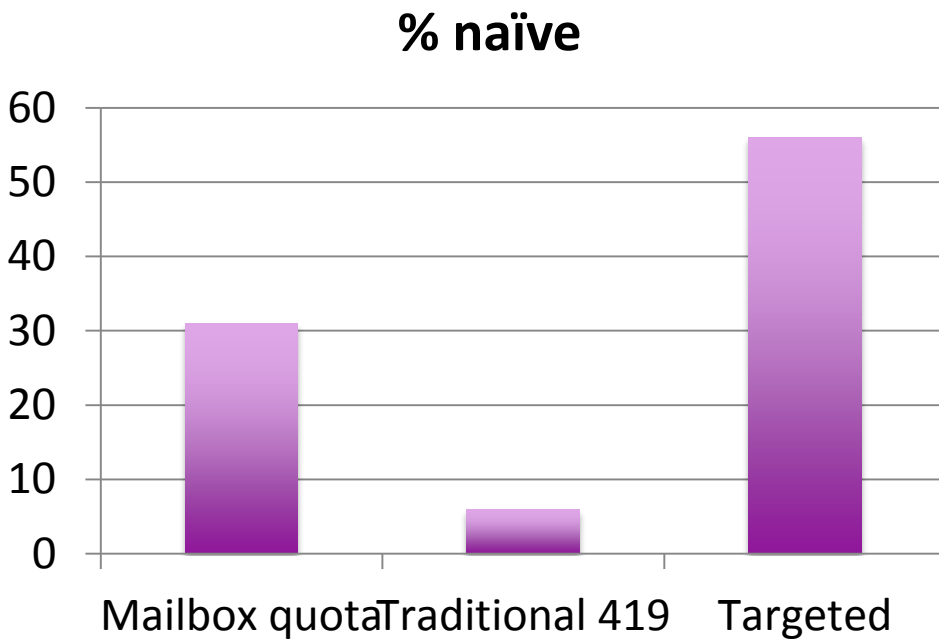
Reasonable answer

- The recipient may get a computer virus.
- The recipient may lose his password.
- This may be a scam aimed at stealing your money.
- There is no risk.
- The recipient may get unwanted advertisements.
- The recipient's account may be blocked if she does not pay attention.

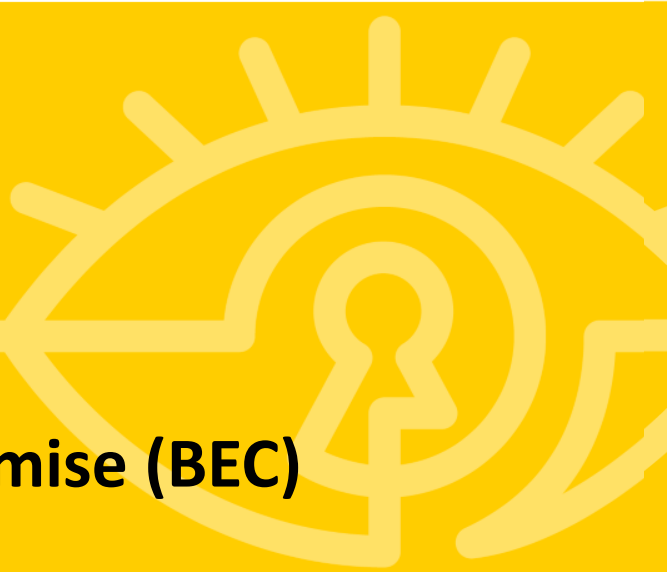
Naïve answer

- The recipient may get a computer virus.
- The recipient may lose his password.
- This may be a scam aimed at stealing your money.
- There is no risk.
- The recipient may get unwanted advertisements.
- The recipient's account may be blocked if she does not pay attention.

Comparing credibility

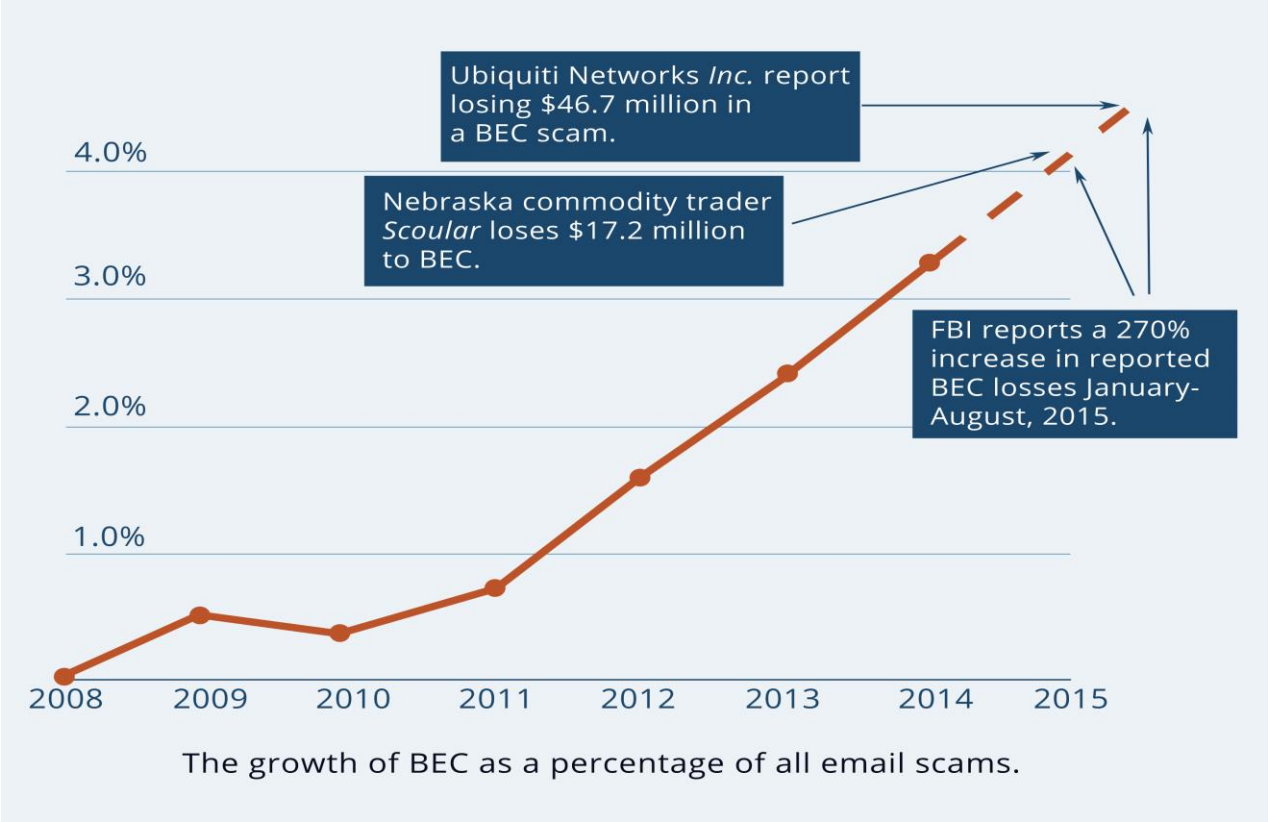


[Jakobsson, Yen, 2015]



Case Study: Business Email Compromise (BEC)

BEC: Targeted and on the rise



The growth of BEC as a percentage of all email scams.

BEC: A first example

From: Liz, Gonzales <EGonzalez@media-produtcion.com>
Subject: October invoice
Date: October 29, 2015 at 9:10:17 AM GMT+8
To: Rudy.McCoy@glitz.com

Dear Rudy,

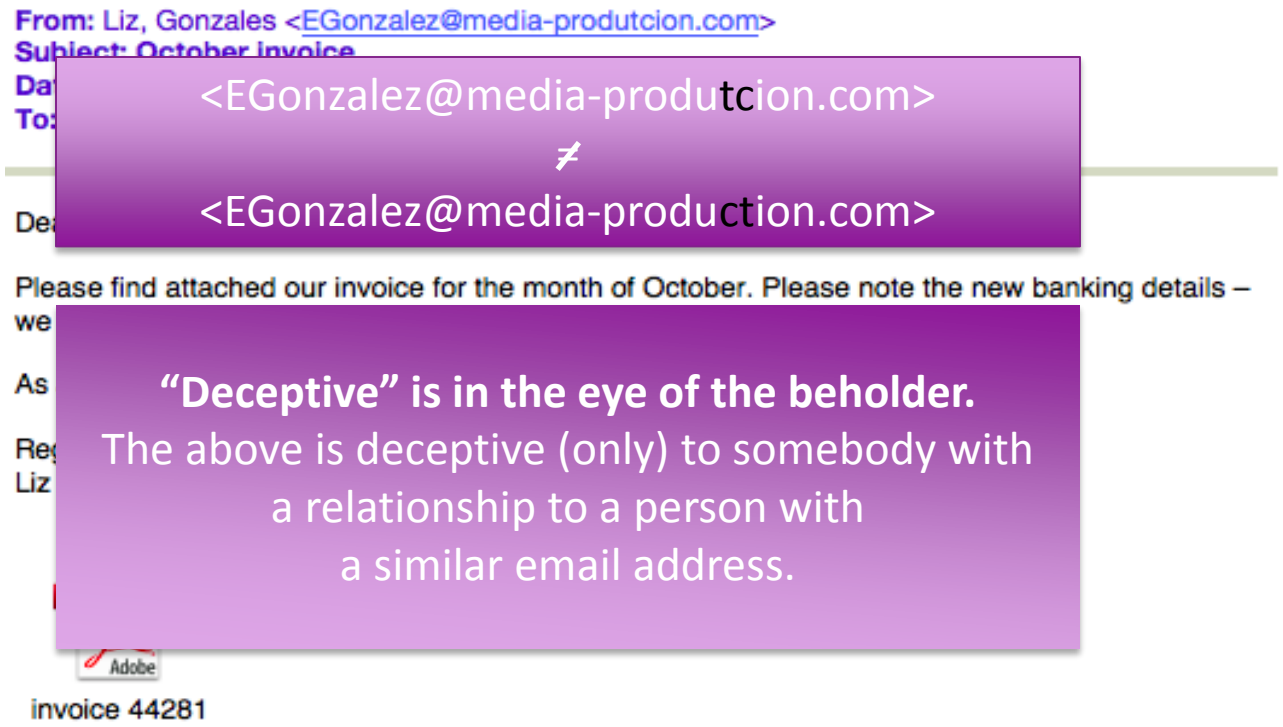
Please find attached our invoice for the month of October. Please note the new banking details – we are staying with US Bank, but the bank updated our account number.

As always, we appreciate your business.

Regards,
Liz

invoice 44281

BEC: A first example



"Deceptive" is in the eye of the beholder.
The above is deceptive (only) to somebody with a relationship to a person with a similar email address.

BEC: A second example

From: Jonathan Blackwell <JBlackwell@blackwellfinancial.com>
Subject: Have a few minutes? Need your help
Date: February 12, 2016 at 2:16:03 AM GMT+8
To: Jim Anderson <JAnderson@blackwellfinancial.com>

Hi Jim,

Are you at your desk? I need to ask for your help.

Jon

**** sent from my tablet, please forgive typos ****

Why recipients fall for BEC

- Persuasive structure:
  The sender is – or looks like – somebody the recipient knows.
  (Everybody want to be nice to friends and colleagues.)
- Persuasive content:
  The request will relate to “normal business” – no Libyan princesses.
  (Why not comply when it makes sense?)

Why spam filters fail to block BEC

- Not high-volume messaging
  Volume-based detection fails
- No typical spam keywords, but normal business conversation
  Content-based detection fails
- Sent by a trusted party (that has been corrupted)
  Trusted parties can send pretty much anything
- … or by a party with no bad reputation (account just created)
  Reputation-based detection fails

Detecting BEC based on persuasive structure

Detecting BEC based on persuasive structure

From: Jonath <JBlackwell@blackwellfinancial.com>
Subject: Ha e a few minutes? Need your help
Date: Febru y 12, 2016 at 2:16:03 AM GMT+8
To: Ji

Hi Jim

Are y

Jon

**** s

JBlackwell@blackwellfinancial.com has a reply-to address to JBlackwell682@gmail.com

and

JBlackwell@blackwellfinancial.com is a trusted party

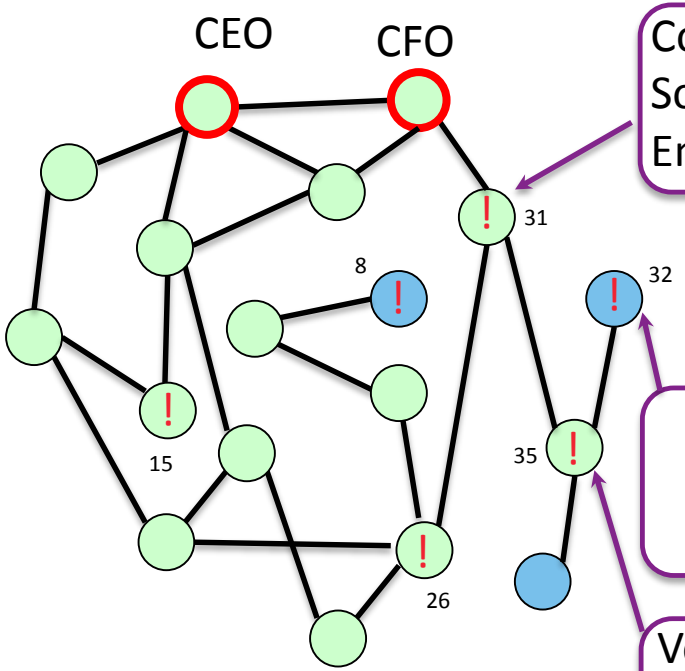
and

JBlackwell682@gmail.com is a never-seen reply-to address with a deceptively similar user name to the trusted party.

Observation/Classification/Action

Deceptive Sender	Cousin name	Discard
New reply-to Deceptive reply-to	Spoof Passive ATO	Ask apparent sender to confirm; deliver
High-risk content	Active ATO	Ask sender on *other channel* to confirm

Quantifying Exposure and Risk



High trust

External

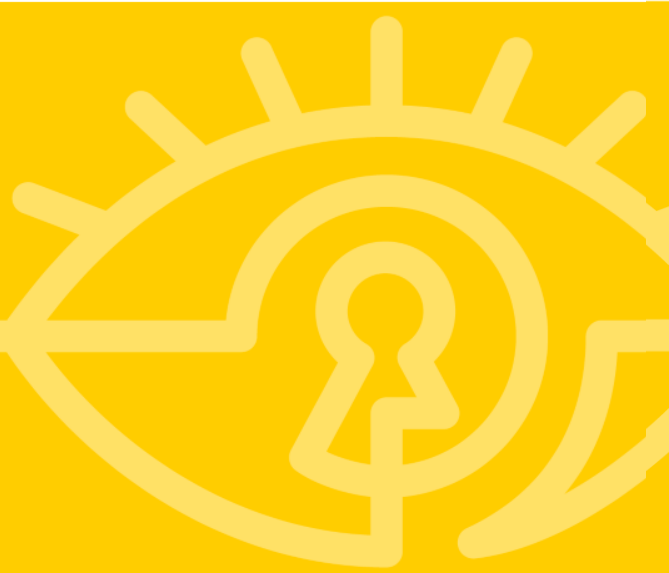
Internal

Commonly cc-ing ‘external self’
Sometimes sending from wrong acct
Emails w/ high-risk keywords (wire)

Has used ‘non-self’ reply-to
Org. does not support DMARC

Very large number of contacts
Has responded to high-risk email
Email w/ high-risk keyword (invoice)

ZapFraud



Action Items

What you should do

- Recognize that spam filters do not address scam/BEC
- Best practices: Assess your organization's exposure to BEC
- Consider internal awareness campaigns
- Be aware of your exposure to targeting
- Review and improve processes for making payments