Understanding the use of stolen account credentials by cybercriminals

Gianluca Stringhini
University College London
Abuse on online services

- Spam
  - [ACSAC2010]
  - [NDSS2013]
  - [USENIX2011]
  - [TDSC2016]

- Information stealing
  - [IMC2016]
  - [CSET2016]
  - [USENIX2015]

- Reputation manipulation
  - [IMC2013]
  - [WWW2017]
  - [WOSN2012]

- Fraud
  - [DIMVA2015]
  - [CCS2015]

- Hate/Bullying
  - [ICWSM2017]
  - [WebSci2017]

- Malware
  - [NDSS2017]
  - [AsiaCCS2017]
Compromised accounts

**Credentials** to online accounts get stolen by cybercriminals

- Phishing
- Data breaches

**Question:**
How are stolen credentials used by criminals in the wild?

- Steal sensitive information
- Send spam
- Sell the credentials on the black market
How can we answer this question?

No data available: we aren’t Google, Facebook etc.
➢ From the outside, we can only see spam

In 2014, a paper by Google shed some light on these topics, but their focus was narrow and they left many questions unanswered

We decided to collect data ourselves and to enable the research community to better understand the ecosystem of stolen account credentials
Gmail “honey” accounts

Google allows to enhance the functionality of accounts by setting up Google App Scripts

We can use this functionality to set up honeypots!
- Monitor which emails are opened
- Monitor which emails are sent
- Monitor durations of accesses, OS, browser
- Monitor locations of accesses

We can then leak credentials and have criminals use them
Our system (publicly available)

We asked Google to monitor the accounts on their side too.
How does the outlet of a leak influence criminal activity? [IMC2016]

Corporate webmail honey accounts (100 accounts)
- Belonging to a fictitious company
- Populated using the Enron dataset

We leaked the credentials through three outlets
- Paste sites
- Underground forums
- Information stealing malware

We monitored activity to the accounts for 7 months, receiving 329 accesses
Types of accesses

**Curious** – just check if accounts are real

**Gold Diggers** – look for sensitive information
  - Use the information
  - Set a price tag

**Spammers** – send spam

**Hijackers** – change the password locking the owner out
Influence of leak outlet on activity

Accesses of credentials leaked through malware are the “stealthiest”
Some accesses are stealthier than others
What are “gold-diggers” looking for?

Mostly financial or account information

Popular words are:
• Seller
• Account
• Payment
• Bitcoin
Timeline of account accesses
How does information on the location of the account owner influence accesses? [IMC2016]
How does the language of an account influence criminal activity? [arxiv]

We created 30 accounts in three different languages
• 10 English accounts
• 10 Greek accounts
• 10 Romanian accounts

We hid fake email invoices for banking institutions in each account

Summary of findings:
• Criminals are more likely to find the hidden sensitive information in the Greek accounts
• Criminals spend more time on Greek accounts

Possible explanation: if criminals do not understand the language of an account, they use automated translation tools
How does leaking credentials on the Dark Web affect criminal activity? [under submission]

We replicated the surface Web experiment on the Dark Web (paste sites and forums)

Some preliminary results:
• Dark Web accesses receive many more accesses than surface Web ones
• Dark Web accesses show a higher degree of sophistication
Conclusion

• We released a honeypot system that allows you to **design your own experiments** to better understand the modus operandi of cybercriminals

• We shed some light on the way that stolen Gmail accounts are used in the wild

• We also developed a honeypot version for Google Spreadsheets
References:

Questions?

g.stringhini@ucl.ac.uk
@gianluca_string

Code available at:
https://bitbucket.org/gianluca_students/gmail-honeypot