Honey Sheets: What Happens to Leaked Google Spreadsheets?

Martin Lazarov, **Jeremiah Onaolapo**, and Gianluca Stringhini
University College London, UK

9th USENIX Workshop on Cyber Security Experimentation and Test
Austin, TX
August 8, 2016

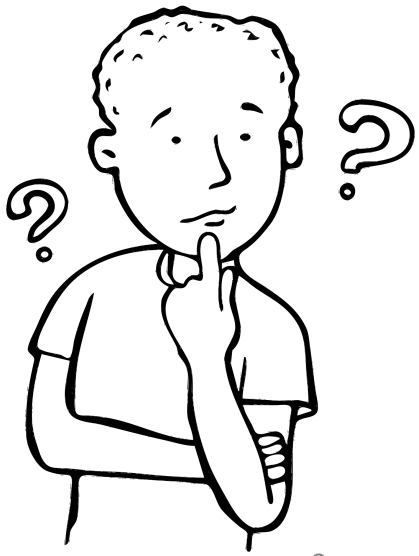
Introduction

- Many useful services are cloud-based
  - Dropbox, OneDrive, etc.
- Valuable content in online accounts
- Cybercriminals attack online accounts and sell credentials

(Bursztein et al. 2014; Herley and Florencio 2010; Stone-Gross et al. 2011)

Question

- What happens to online accounts and documents after compromise?



Previous work

- Malicious activity in webmail accounts
  (Bursztein et al. 2014, Stringhini and Thonnard 2015)
- Emphasize spearphishing as primary attack vector
- No publicly available infrastructure to monitor compromised accounts

Aim

- Study **actions** and **access patterns** of cybercriminals on leaked online spreadsheets

- We developed an infrastructure to help researchers understand what happens to compromised cloud documents

Our infrastructure

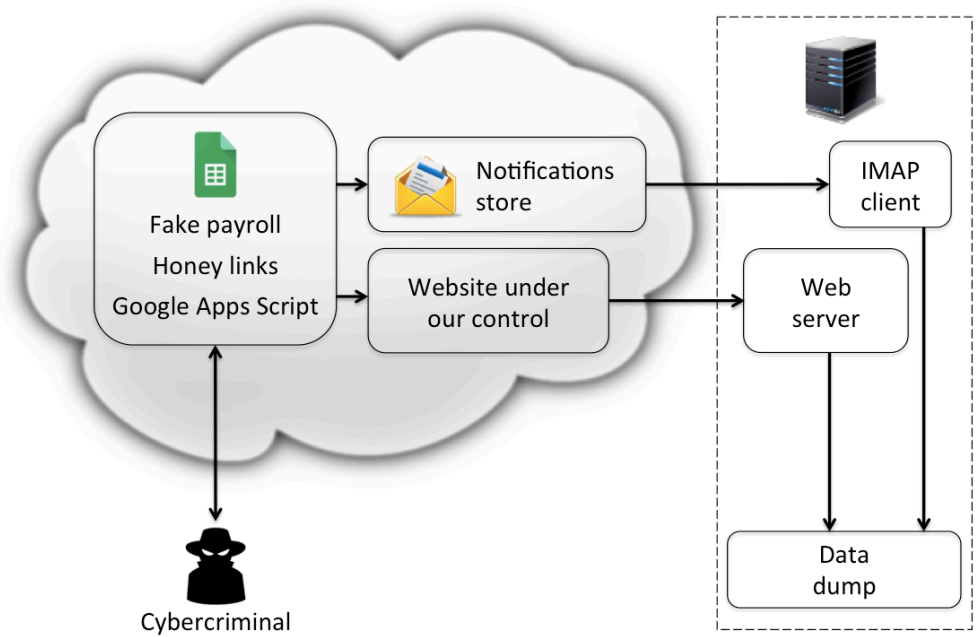
- **Honeypot system** comprising honey spreadsheets and monitoring infrastructure

- We developed proof-of-concept to test our ideas

System components

- **Honey spreadsheets** containing fake information, including honey links
- **Web server** to monitor clicks on honey links
- **Notification store** to receive messages about activity in honey spreadsheets
- **IMAP client** to retrieve those messages

System overview



Scenarios

We tested 2 scenarios using our proof-of-concept

1. Scenario 1– Hacker leaking financial information
2. Scenario 2 – Naïve user sharing spreadsheet info with colleagues

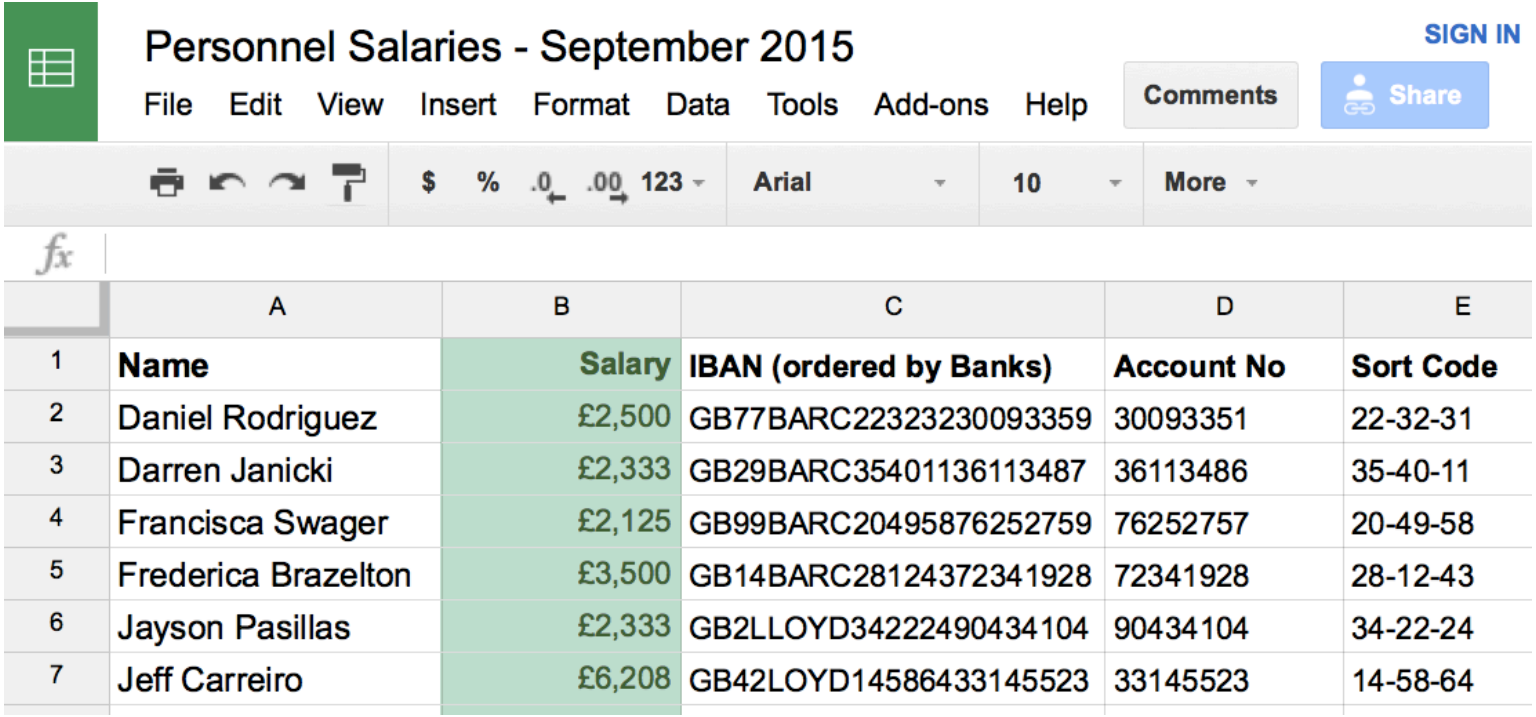
Quick note

- The system is flexible and can be adapted to many scenarios
- The scenarios to set up depend on the questions that the researcher intends to find answers to

Proof-of-concept

- Created 5 spreadsheets with fake payroll info
- Inserted `goo.gl` honey links in spreadsheets
  - 3 honey links point to our website
  - 6 honey links point to nonexistent bank pages
- To track location, browser info, IP addresses etc. of visitors

Spreadsheet example



	A	B	C	D	E
1	**Name**	**Salary**	**IBAN (ordered by Banks)**	**Account No**	**Sort Code**
2	Daniel Rodriguez	£2,500	GB77BARC22323230093359	30093351	22-32-31
3	Darren Janicki	£2,333	GB29BARC35401136113487	36113486	35-40-11
4	Francisca Swager	£2,125	GB99BARC20495876252759	76252757	20-49-58
5	Frederica Brazelton	£3,500	GB14BARC28124372341928	72341928	28-12-43
6	Jayson Pasillas	£2,333	GB2LLOYD34222490434104	90434104	34-22-24
7	Jeff Carreiro	£6,208	GB42LOYD14586433145523	33145523	14-58-64

Honey links example

G	
Transfer Accounts	**Payment Link**
Barclays	http://goo.gl/TmzFcm
HSBC	http://goo.gl/epbE0Q
Natwest	http://goo.gl/p9yti9
Lloyds	http://goo.gl/uYrZ6V
Halifax	http://goo.gl/AHGjKy
TSB	http://goo.gl/nIsJLQ
Paypal	http://goo.gl/PHr1wE
TransferWise	http://goo.gl/k0Gf35

Leaking the spreadsheets

- We leaked URLs pointing to the spreadsheets on `pastebin.com`

- Known mode of operation of cybercriminals leaking credentials and documents

leaked payr0lls

A GUEST FEB 25TH, 2016 52 NEVER

text 0.12 KB raw download

```
1. leaked payr0lls..
2.
3. https://docs.google.com/spreadsheets/d/1e0KA7OhvCwog6oDyE8--IgG8ma3tkOim80aO2Rdpwpc/edit?usp=sharing
```

October 2015 payrolls

A GUEST MAR 17TH, 2016 48 NEVER

text 0.17 KB raw download

```
1. hi Richard, here is the spreadsheet with the payrolls for October 2015
2.
3. https://docs.google.com/spreadsheets/d/1-yrO5JORFCdmBYBqcBPJAT7xBTAwOPyBY08NslpD4sQ/edit?usp=sharing
```

Ethical considerations

- No real information in the spreadsheets
- We did not leak credentials of the accounts hosting the spreadsheets
- We obtained IRB approval from our institution

Summary of results

- Scenario 1 (Hacker): 46 days
  - 112 accesses, 17 modifications
- Scenario 2 (Naïve user): 26 days
  - 53 accesses, 11 modifications

Summary of results

- Differences in accesses not statistically significant

- Dataset available online

Examples of modifications

- Decoy bank account number deleted
- C++ code snippet inserted
- Insult in spreadsheet
- Defacement of spreadsheet
  - Our infrastructure could potentially attract trolls and cyberbullies

Activity on goo.gl links

- **39 unique IPs** visited the 3 honey links pointing to our webserver
- **44 visits** to those 3 honey links
- **174 clicks total** on all 9 honey links
- Accesses from **35 countries**

Locations of accesses



Limitations

- Visitors likely not sophisticated cybercriminals
- Attackers could copy the honey sheets and interact with them offline
- Google Apps Script tracking limited for visitors that are not logged in

Future work

- Make spreadsheets more believable
- Scale up experiments
- Devise taxonomy of attackers targeting cloud documents
- Build comprehensive infrastructure for monitoring compromised webmail accounts and spreadsheets

Another piece of the puzzle

- Jeremiah Onaolapo, Enrico Mariconti, Gianluca Stringhini. "What Happens After You Are Pwnd: Understanding The Use Of Leaked Webmail Credentials In The Wild."
  - To be presented at the *ACM Internet Measurement Conference 2016 (IMC 2016)*, Santa Monica, California.

- Honeypot infrastructure that monitors actions and accesses to compromised **webmail** accounts

Overarching idea



- Publicly available comprehensive infrastructure
- To help the research community “see” further into the underground ecosystem of compromised accounts and documents
- Criminologists are already using the system

Thanks

Questions?

j.onaolapo@cs.ucl.ac.uk

References

Cormac Herley and Dinei Florencio. “Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy”. In: *Economics of Information Security and Privacy*. 2010.

Brett Stone-Gross et al. “The underground economy of spam: A botmaster's perspective of coordinating large-scale spam campaigns”. In: *USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)*. 2011.

References

Elie Bursztein et al. “Handcrafted Fraud and Extortion: Manual Account Hijacking in the Wild”. In: *ACM SIGCOMM Conference on Internet Measurement*. 2014.

Stringhini, Gianluca, and Olivier Thonnard. “That ain’t you: Blocking spearphishing through behavioral modelling.” *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*. Springer International Publishing, 2015.