The Devil is in the details
Social Engineering by means of Social Media

BY

DAAN WAGENAAR
YANNICK SCHEELEN
Introduction

- **Online Social Networks**
  - LinkedIn (service data, disclosed data)
  - Facebook (entrusted data, incidental data)

- **Social Engineering**

- **Relevant information**

- **What else is new?**
Research Questions

How can Online Social Networks be used in the automated creation of a graphical view of the company hierarchy and its employees for the purpose of social engineering?

- How can current information gathering techniques be combined to achieve this goal?
- What are the consequences for companies?
- What can companies do to mitigate this process?
How did we start?

START ON LINKEDIN
CREATE FAKE PROFILE
LINKEDIN TIERS
GETTING CONNECTED WITH THE COMPANY
SEARCHING & FILTERING
CRAWLING THE RESULTS
Create fake profile

- Being a member is a necessity
  - Access to user profiles
  - Use LinkedIn’s search functionality
  - Etc...

- Create a false identity with information that conforms to the target company = zombie profile
LinkedIn tiers

- Getting information from other users depends on the tier:
  - 1\textsuperscript{st} tier
  - 2\textsuperscript{nd} tier
  - 3\textsuperscript{th} tier
  - Out of Network

- 2\textsuperscript{nd} tier show enough unobfuscated information
- Need at least one 1\textsuperscript{st} tier connection to get 2\textsuperscript{nd} tier results
LinkedIn tiers

- Getting information from other users depends on the tier:
  - 1\textsuperscript{st} tier
  - 2\textsuperscript{nd} tier
  - 3\textsuperscript{th} tier
  - Out of Network

- 2\textsuperscript{nd} tier show enough unobfuscated information
- Need at least one 1\textsuperscript{st} tier connection to get 2\textsuperscript{nd} tier results
LinkedIn tiers

- Getting information from other users depends on the tier:
  - 1<sup>st</sup> tier
  - 2<sup>nd</sup> tier
  - 3<sup>rd</sup> tier
  - Out of Network

- 2<sup>nd</sup> tier show enough unobfuscated information
- Need at least one 1<sup>st</sup> tier connection to get 2<sup>nd</sup> tier results
LinkedIn tiers

- Getting information from other users depends on the tier:
  - 1st tier
  - 2nd tier
  - 3rd tier
  - Out of Network

- 2nd tier show enough unobfuscated information
- Need at least one 1st tier connection to get 2nd tier results
LinkedIn tiers

- Getting information from other users depends on the tier:
  - 1st tier
  - 2nd tier
  - 3rd tier
  - Out of Network

- 2nd tier show enough unobfuscated information
- Need at least one 1st tier connection to get 2nd tier results
Getting connected with the company

- Company’s “followers” list

- List of partly obfuscated names
  - Current employment
  - First name + first letter of the last name
  - Hyperlink to the public profile
    - Public profile shows the full name...

- Crawl list of followers and send connection requests
  - Once the first connection was made, the company circle was infiltrated
Getting connected with the company

- Company's "followers" list
- List of partly obfuscated names
  - Current employment
  - First name + first letter of the last name
  - Hyperlink to the public profile
    - Public profile shows the full name...

- Crawl list of followers and send connection requests
  - Once the first connection was made, the company circle was infiltrated
Getting connected with the company

Company’s “followers” list

- List of partly obfuscated names
  - First name + first letter of the last name
  - Hyperlink to the public profile

Public profile shows the full name…

Crawl list of followers and send connection requests

Once the first connection was made, the company circle was infiltrated
Searching 2\textsuperscript{nd} tier connections
- Limit of 100 search results

Scoping the target company
- Define keywords

Reducing the LinkedIn dataset
- Apply filters
Crawling the results

- Final dataset was defined by the filtering process

- Our custom made crawler managed to:
  - Crawl all the names of 1st and 2nd tier connections
  - Crawl all the information these profiles put on their account
Now what?

CONTINUE ON FACEBOOK
Why Facebook?

- Data enrichment

- Getting to user’s private information
  - Not found on LinkedIn
Profile matching

- Unfortunately the profiles are not a 1-1 relation
- One user’s name on LinkedIn can appear many times on Facebook
  - ~901 million users...

- Matching profiles just by using the name won’t work
  - Social synergy is the key
Profile matching

- Unfortunately the profiles are not a 1-1 relation
  - One user's name on LinkedIn can appear many times on Facebook

Social synergy is the key
When do we have a match?

- Three ways to define when we have a certain match

1. Matching using public data
2. FLEMP
3. Zombie profiles
1) Matching using public data

- Using publicly available data on Facebook

- Can a match be found?
  - Same name, current employment, education, location, etc...
2) FLEMP

- “Friend List of Earlier Matched Profiles”
  - Why can this work?

- Search through the publicly available friend lists

- Compares names found in these lists to names of unidentified profiles in our dataset

- If a match is found, the profiles match
3) Zombie Profiles

- Use zombie profiles to spam friendship requests
  - When search returns multiple names and no match can be made
  - Spam friendship requests to all those profiles

- If the user accepts the friendship request
  - Crawl the data
  - Try to make a match with private data that is now accessible
How do we get the data?

- Public crawling
  - Collect all the information that is publicly available

- Zombie Profiles
  - Shotgun approach – friend as many people as possible
  - Undirected

- iCloner
  - Surgical approach
  - Directed
iCloner

- Take profile from one social network
- See if it doesn’t exist on the other social network
- Clone his details onto that social network
- Try to connect to his connections

- From LinkedIn ➔ Facebook
Which results did we get?
1 day of connecting
1 day of crawling
Resulted in...
LinkedIn Zombie Profile

- 106 invitations sent
- 39 accepted
- **36.7%**

Two degrees away
Friends of friends; each connected to one of your connections

11,400+
Defining the final dataset on LinkedIn

- **First filtering: 286 profiles**
  - Conformed to our initial search on the company
  - All information crawled
- **125 profiles were matched on Facebook**
  - 43%
- **After final filtering: 86 profiles defined on LinkedIn**
  - 37 on Facebook
  - Another 9 found using FLEMP
  - 0 found by using Zombie Profiles
  - 46 Facebook profiles in total
  - 55%
Information collected on LinkedIn

Crawling rate of LinkedIn fields

- First name
- Last name
- Headline
- Current Employment
- Job title
- Living location
- Industry
- Education
- Past Employment
- Summary
- Websites
- Interests
- Twitter
Information collected on Facebook:

- First name
- Last name
- Gender
- Friends
- Company
- Current City
- Wall viewable
- University
- Home town
- Company Position
- Degree
- Music
- Relationship
- Duration of employment
- Sports
- Activities
- Languages
- Birthday
- College
- Interest in
- Movies
- TV Programs
- High school
- Email
- Siblings
- Uncle & Aunt
- Children
- Political view
- Bio
- Religion
- Quotes
- Phones

Crawling rate of Facebook fields
Matching the information – Social Synergy

Fields used for profile matching in %

- **Current Employment, Education** - 20%
- **Current Employment, Education, Living location** - 13%
- **Found in Friend List of Earlier Matched Profiles (FLEMP)** - 11%
- Exact profile picture - 9%
- Education, Past education - 5%
- **FLEMP, Current Employment, Education** - 2%
- **Current Employment, Single result found** - 2%
- Education, Living location - 2%
- Education, Living location - 2%
- Current Employment - 2%
- **FLEMP, Living Location** - 2%
- Likes, Living location - 2%
- Past, education, Living location - 2%
Zombie Profiles and iCloner

- **Zombie Profiles**
  - 200 friendship requests sent
  - 13 accepted
  - 6.5%

- **iCloner**
  - 10 friendship requests sent
  - 6 accepted
  - 60%
  - 4 friendship requests received
What does it all mean?
Job function parsing

- Parse sub-departments in the targeted department
- Parse job function per sub-departments
- Assign weight to function
- Sort based on weight

Owner at Company
Senior Manager at Company Department
Junior IT function at Company
IT Security function at Company

Function
Company name
Department name
Why is this useful?
Information gathering

- **More** data can be gathered **faster**
- Data is automatically sorted
- Hierarchical structure of a company becomes visible
- Allows for social engineers to create attack scenarios easier
Creating a bond of trust

• Try and build a bond of trust with the target
  ○ Hey, I heard you just went on a holiday, how was it?
    ♦ Of course you know the target went on a holiday because you saw his Facebook wall posts...
  ○ I heard from a colleague you bought a new book, how is it?
    ♦ You know the colleague because you created a hierarchy of the company that puts them in the same function
    ♦ But in fact you just crawled the Facebook wall

• Get the target to tell you information that he/she would otherwise have never told you
Creating a bond of trust

- Try and build a bond of trust with the target
  - Hey, I heard you just went on a holiday, how was it?
    - Of course you know the target went on a holiday because you saw his Facebook wall posts...
  - I heard from a colleague you bought a new book, how is it?
    - You know the colleague because you created a hierarchy of the company that puts them in the same function
    - But in fact you just crawled the Facebook wall

- Get the target to tell you information that he/she would otherwise have never told you
Creating a bond of trust

- Try and build a bond of trust with the target
  - Hey, I heard you just went on a holiday, how was it?
    - Of course you know the target went on a holiday because you saw his Facebook wall posts...
  - I heard from a colleague you bought a new book, how is it?
    - You know the colleague because you created a hierarchy of the company that puts them in the same function
    - But in fact you just crawled the Facebook wall

- Get the target to tell you information that he/she would otherwise have never told you
Try and build a bond of trust with the target:

- Hey, I heard you just went on a holiday, how was it?
  - Of course you know the target went on a holiday because you saw his Facebook wall posts…
- I heard from a colleague you bought a new book, how is it?
  - You know the colleague because you created a hierarchy of the company…
  - But in fact you just crawled the Facebook wall...

Get the target to tell you information that he/she would otherwise never tell you.
Creating a false sense of authority

- Reference persons placed higher in the company hierarchy
  - Boss X just told me he needs access to those files, can you mail them to me?

- Create a false sense of authority

- Incline the target to comply faster to the social engineer
What can companies do?

MITIGATION
Creating Policies

- **Prevent social synergy**
  - Don’t put your work or education details on Facebook

- **Reduce the effect of data gathering techniques**
  - Set the right privacy settings on Facebook data
  - Verify that who you friend is that actual person

- **Be generic on LinkedIn**
  - Omit exact job function and department?
Generating user awareness

- Periodic testing of publicly available data
- Perform awareness sessions with concrete examples from our research
Conclusions
Conclusion

How can current information gathering techniques be combined to achieve our goal?
- Zombie profiles
- iCloning technique
- Efficient matching

What are the consequences for companies?
- Gathering data becomes easier and faster for social engineers
- Social engineering attacks can be created easier
- The company hierarchy can be visualized

What can companies do to mitigate this process?
- Create company policies for social media usage
- Generate user awareness
Creating a visualized hierarchy of a company and its employees in an automatic way is possible

- Automated
- Fast

Allowed by the wealth of information that is available online

People are generally not aware at how much information they share online and how easy it is to get access to it – if you really want it
Questions?

THANK YOU