Socialkk privacy score through vulnerability contagion process

Aghiles DJOUDI, Guy PUJOLLE

Sorbonne University

June 8, 2019

Outline

1. Introduction

2. Contribution

3. Conclusion

Context

Introduction

	World (2018)	World (2022)	France (2018)
Users	3.8 billion	4.2 billion	25.9 million
Email accounts	4.4 billion	5.6 billion	68 million
Email accounts per user	1.7	1.9	2.1
Emails received each day	281 billion	333 billion	1.4 billion
The email market	9.8 Mrds $	20.4 Mrds $	?

Table 1: Email statistics [1].

- Email:
  - More than 50% of the world population use email
  - Usage: 75% personal, 25% professional
- Facebook:
  - 2.2 billion active users (29% worldwide)
  - 10 billion messages are sent every day
  - 8.051 billion $

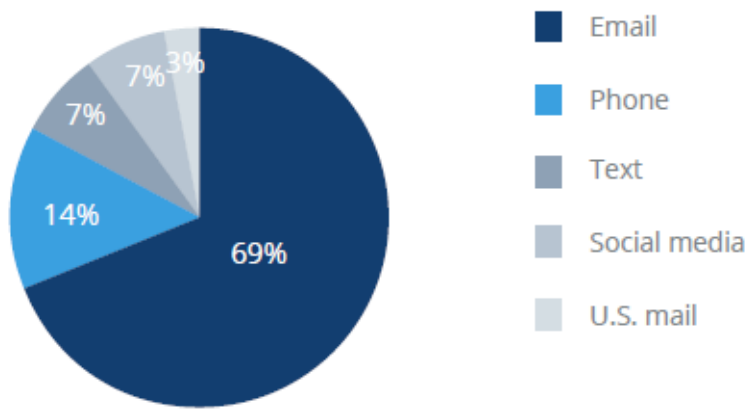


Figure 1: Communication tools [1].

Context

Introduction

	World (2018)	World (2022)	France (2018)
Users	3.8 billion	4.2 billion	25.9 million
Email accounts	4.4 billion	5.6 billion	68 million
Email accounts per user	1.7	1.9	2.1
Emails received each day	281 billion	333 billion	1.4 billion
The email market	9.8 Mrds $	20.4 Mrds $	?

Table 1: Email statistics [1].

- Email:
  - More than 50% of the world population use email
  - Usage: 75% personal, 25% professional
- Facebook:
  - 2.2 billion active users (29% worldwide)
  - 10 billion messages are sent every day
  - 8.051 billion $

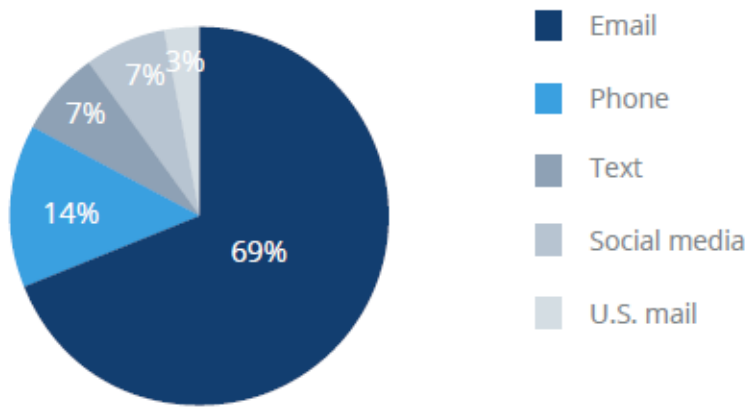


Figure 1: Communication tools [1].

Problematic

Introduction

- Privacy issues:

1. **External vulnerabilities**
   - ➟ Network technologies: 4G, 5G, Wifi, Ethernet.
   - ➟ Security protocols: HTTPS, SMTPS, IPsec
2. **Internal vulnerabilities**
   - ➟ Service providers
     - ✳ General Terms and Conditions of Use (GTC)
   - ➟ 3^{rd} parties
     - ✳ Privacy settings
     - ✳ Permission management
   - ➟ Users
     - ✳ Configuration of user accounts
     - ✳ Users list management

Figure 2: Privacy, Am I concerned ?.

Problematic

Introduction

- Privacy issues:

1. **External vulnerabilities**
   - ➟ Network technologies: 4G, 5G, Wifi, Ethernet.
   - ➟ Security protocols: HTTPS, SMTPS, IPsec
2. **Internal vulnerabilities**
   - ➟ Service providers
     - ✳ General Terms and Conditions of Use (GTC)
   - ➟ 3^{rd} parties
     - ✳ Privacy settings
     - ✳ Permission management
   - ➟ **Users**
     - ✳ **Configuration of user accounts**
     - ✳ **Users list management**

Figure 2: Privacy, Am I concerned ?.

Motivation

Introduction

- Give users a way to measure their vulnerabilities
- Help users to better configure their email accounts.
- Alert users of a new vulnerability.
- Make users aware about the level of threat diffusion.

Figure 3: Privacy index
[**maximilien_privacyasaservice_20**

Challenges

Introduction

- Recommend customized security measures
  - New password every time period
  - Secure the exchange with vulnerable accounts
  - Adapt permissions to changes
- Vulnerability measurement of the social environment
  - Measure the level of vulnerability of interactions
  - Measure the level of influence between users.
- Calculate the vulnerability of the message path
  - Identification of MTA servers
  - Assign a trust score to each server
  - Calculate the average confidence of the path.

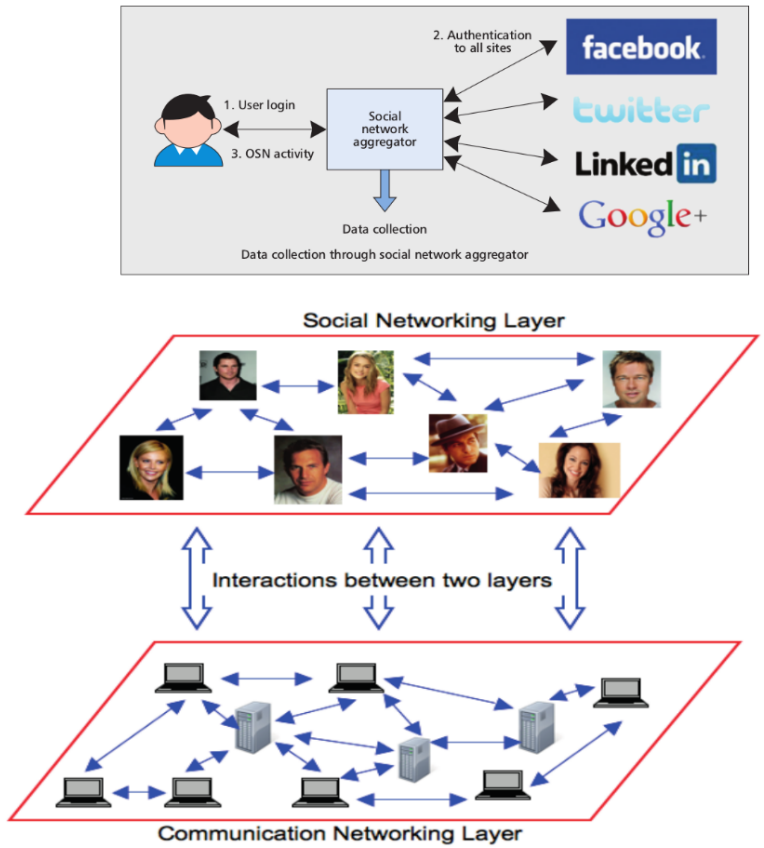


Figure 4: Social interaction.

Challenges

Introduction

- Recommend customized security measures
  - New password every time period
  - Secure the exchange with vulnerable accounts
  - Adapt permissions to changes
- **Vulnerability measurement of the social environment**
  - **Measure the level of vulnerability of interactions**
  - **Measure the level of influence between users**
- Calculate the vulnerability of the message path
  - Identification of MTA servers
  - Assign a trust score to each server
  - Calculate the average confidence of the path.

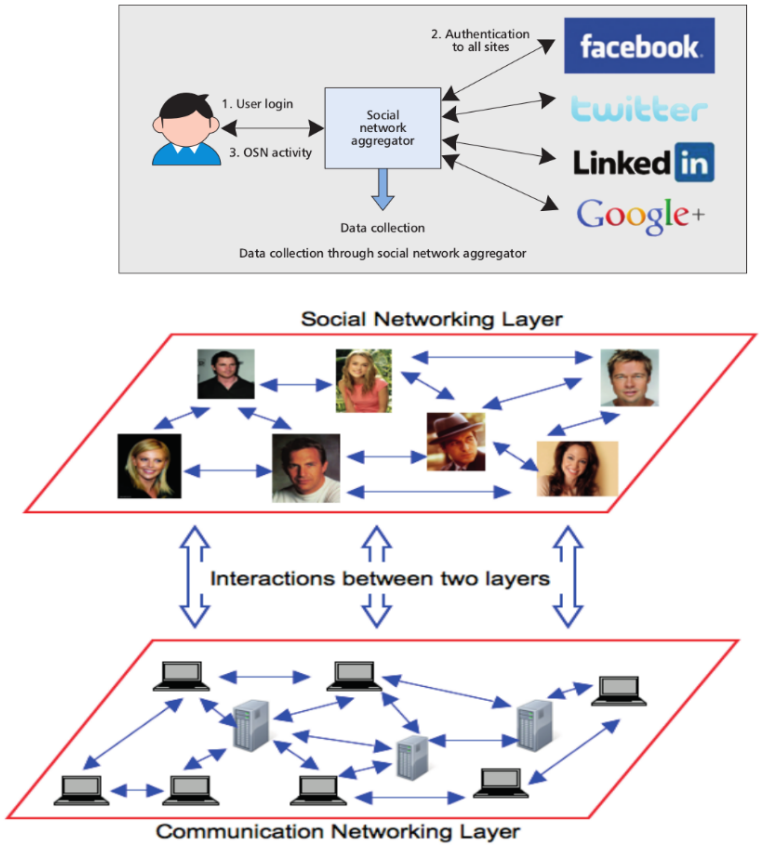


Figure 4: Social interaction.

Contributions

Introduction

- Social vulnerability estimation.
  - Individual vulnerability -> Social vulnerability.
  - Vulnerability diffusion process.
  - Relationship between trust and vulnerability.
  - Data: Enron & Caliopen emails.

Figure 5: The vulnerability of one user is the vulnerability of all users.

Outline

1. Introduction

2. Contribution

3. Conclusion

Outline

1. Introduction

2. Contribution
   1. Related work
   2. Diffusion process
   3. Experimentation
   4. Results exploitation

3. Conclusion

Outline

1. Introduction

2. Contribution

3. Conclusion

1. Related work
2. Diffusion process
3. Experimentation
4. Results exploitation

Related work

Comparison

Works	Contribution	Goal
[**gandouz_protect_2012**] **Protect U**	Classification of interlocutors	Friends lists management
[**fang_privacy_2010**] **Privacy Wizard**	Friends Classification	Permission Configuration
[**frey_social_2011**] **SocialMarket**	Common Interests	Assessment of Trust Relationshi
[**yongbozeng_study_2015**] **PARE**	Information Leakage	Evaluation of Information Dissem
[**hameed_lens_2011**] **LENS**	Spam Protection	Trusted Emitters Evaluation
[**tran_social_2010**] **SocialEmail**	Classify msg by paths	Evaluate message reliability
[**nepali_sonet_2013**] **Privacy Index**	Visibility, sensitivity	Msg exposure assessment

Table 2: Contributions from existing work.

Outline

1. Introduction

2. Contribution

3. Conclusion

1. Related work
2. Diffusion process
3. Experimentation
4. Results exploitation

Step 1: Individual vulnerability measurement

Method

Parameter	Value	
Network connection	Private, Public	1,2
Technology	Ethernet, 5G, 4G, Wifi	[1:4]
Operating system	Windows, Unix, Mac	[1:3]
Web browser	Firefox, Chrome, Opera, ...	[1:10]
Password strength	low, medium, strength	[1:3]
Sessions opened	counter	[1:10]
TLS version	v1.0, v1.1, v1.2, v1.3	[1:4]

Table 3: Individual Vulnerability parameter

$$Y = \sum_{i}^{n} \frac{w * V}{n} \quad (1)$$

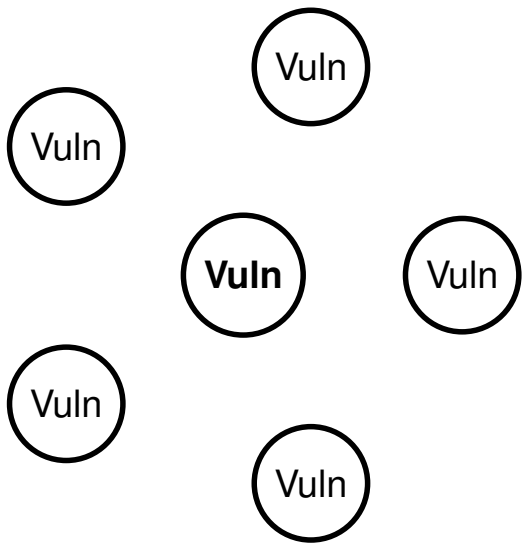


Figure 6: Individual vulnerability level.

- **Y:** Individual vulnerability
- **w:** Weight of each vulnerability
- **V:** Scores mentioned above

Step 2: Users reputation estimation

Method

Parameter	Value	
Frequency of msg exchanged	continuous	N
Discussion time	continuous	N
% of messages exchanged	cipher, signed or clear	[1:3]
Message type exchanged	Text, images, videos, script	[1:4]

Table 4: Trust grant features

$$\alpha = P(reputation) = P(X \geq 1) = 1 - (1 - P(trust))^n \qquad (2)$$

- Where,
  - **X**: trust grant, random variable, X ~B(n,p)
  - **n**: deg(node)
  - **P(X=1)**: The probability of being assigned one trust grant by an interlocutor

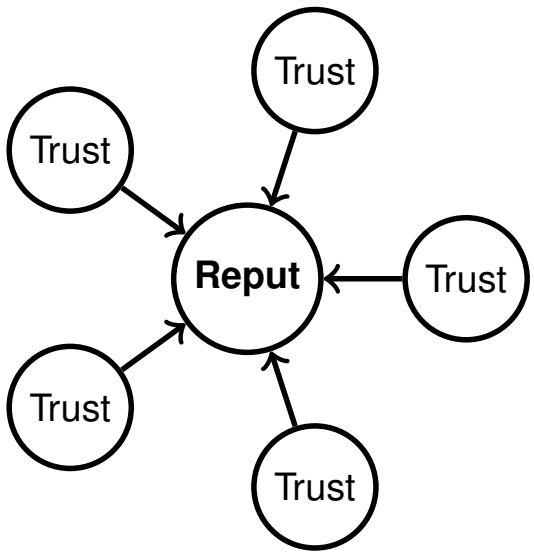


Figure 7: Reputation level.

Step 3: Social vulnerability measurement

Freidkin's theory of social influence

- Input (Features):
  - $Y^{(1)}$ = Vector of the individual vulnerabilities of N users (eq 1)
  - α= The level of reputation (influence) of each user (eq 2)
  - M = Adjacency matrix N x N
- Model:

$$Y^{(t)} = \alpha M Y^{(t-1)} + (1-\alpha) Y^{(t-1)} \tag{3}$$

- Output:
  - $Y^{(t)}$ = Vector of the social vulnerabilities of the N users

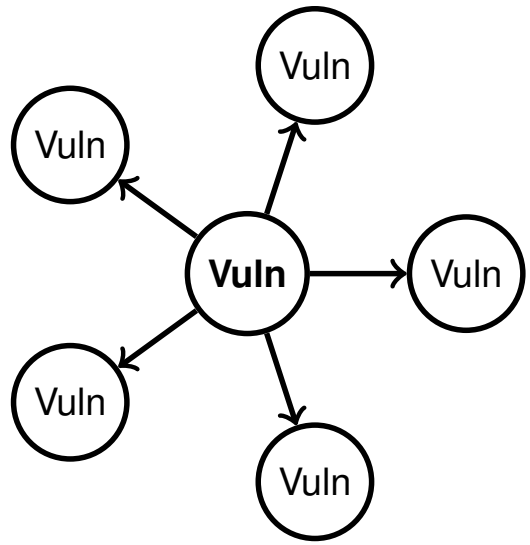


Figure 8: Social vulnerability.

Step 3: Social vulnerability measurement

Freidkin's theory of social influence

Formal properties of the model:

- When a user's influence is high, the model is reduced to:
  - average vulnerabilities of his friends weighted by their trust levels.

$$Y^{(t)} = 1 * MY^{(t-1)} + (1-1)\,Y^{(t-1)} \tag{3}$$

$$Y^{(t)} = MY^{(t-1)}$$

- In the absence of influence, the model is reduced to:
  - his own vulnerability weighted by the level of mistrust of his friends

$$Y^{(t)} = 0 * MY^{(t-1)} + (1-0)\,Y^{(t-1)} \tag{3}$$

$$Y^{(t)} = Y^{(t-1)}$$

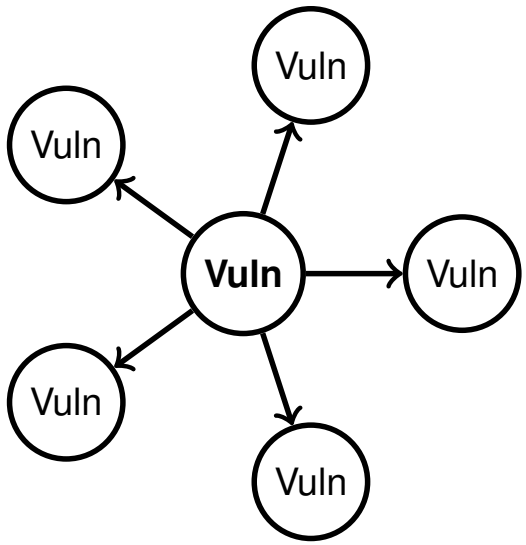


Figure 9: Social vulnerability.

Outline

1. Introduction

2. Contribution
   1. Related work
   2. Diffusion process
   3. Experimentation
   4. Results exploitation

3. Conclusion

Email datasets

Experimentation

Parameter	Value
Users	958
Messages	6966
Diameter	958
# of msg on average	2.413361
Msg density	0.00252
Modularity	0.654600
Average distance	3.042114

Table 5: Enron dataset properties.

Figure 10: Enron logo.

Parameter	Value
Users	5885
Messages	26547
Diameter	2096
# of msg on average	9.02192
Msg density	0.001533
Modularity	0.86526
Average distance	3.914097

Table 6: Caliopen dataset properties.

Figure 11: Caliopen logo.

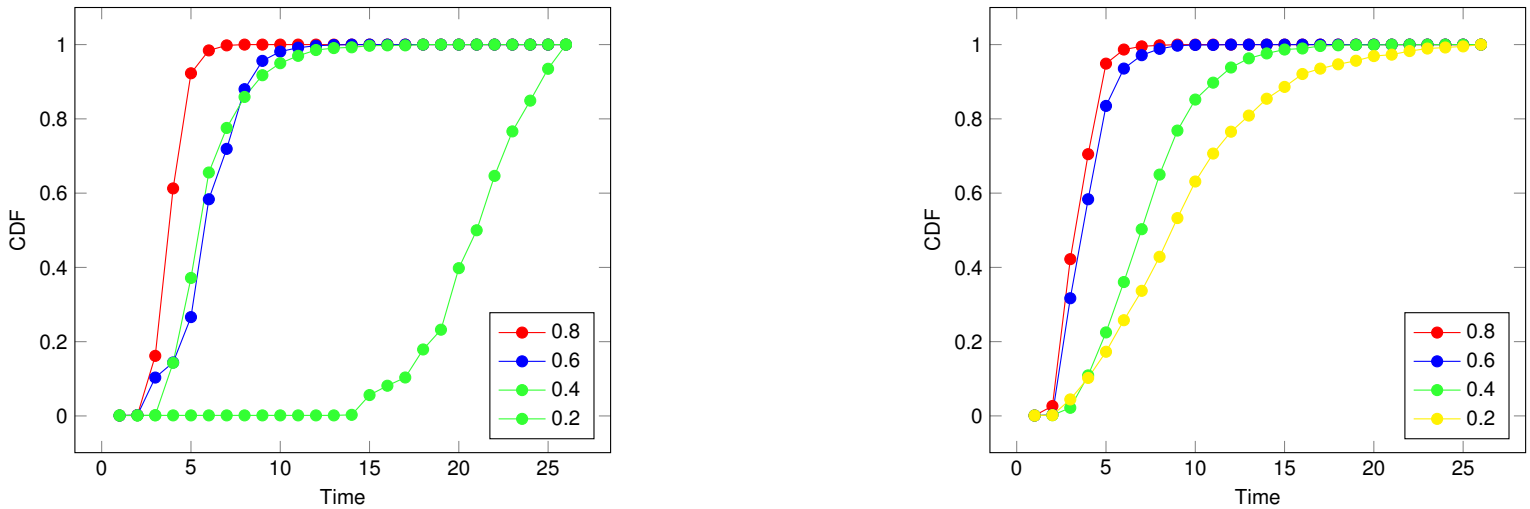
Outline

1. Introduction

2. Contribution
   1. Related work
   2. Diffusion process
   3. Experimentation
   4. Results exploitation

3. Conclusion

Results exploitation



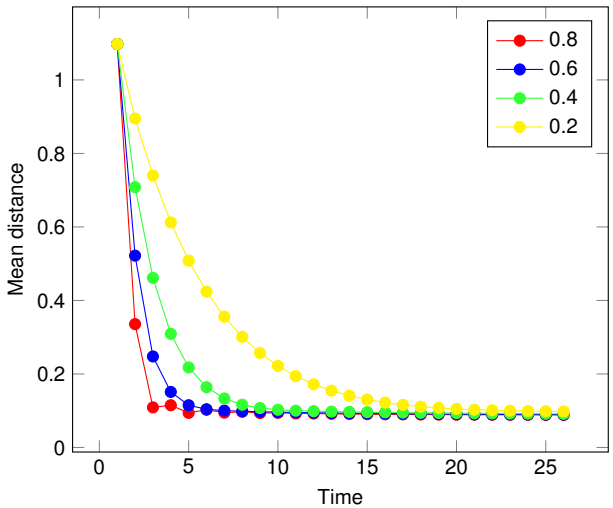
(a) Enron dataset.

(b) Caliopen dataset.

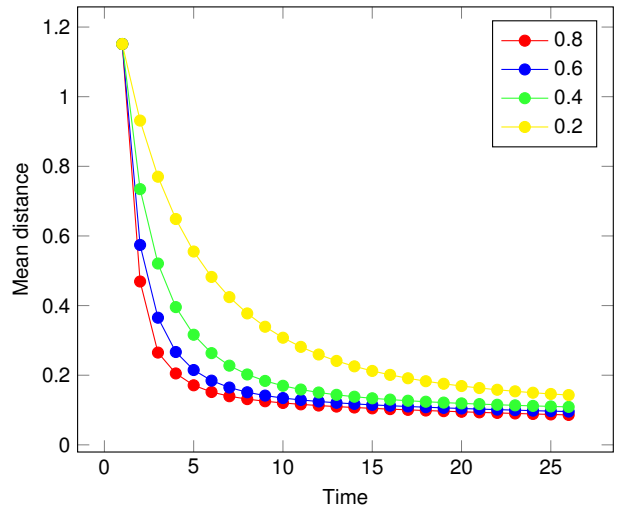
Figure 12: Cumulative distribution function of infected users.

- Figures shows the CDF of the vulnerability diffusion process.
- The vulnerability diffusion process increases as the reputation level of vulnerable users increases.
- Users with high reputation values contribute significantly to the diffusion
  - They spread their vulnerabilities quickly and widely through the network.

Results exploitation



(a) Enron dataset.



(b) Caliopen dataset.

Figure 13: Convergence of the diffusion process.

- The process converge when the mean distance between social vulnerability scores is the minimum.
- Assigning trust to vulnerable users allows them to achieve a high level of reputation.
- Consequently, they infect all other vulnerability values.

Results

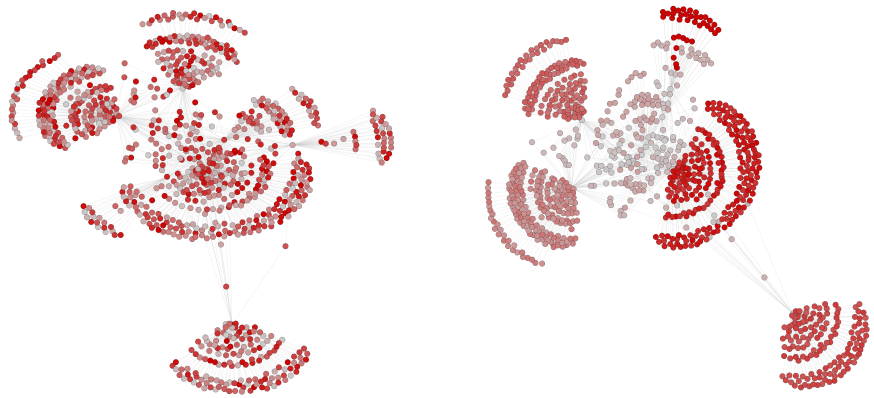
Comparison

Initial values:

- generated randomly (normal distribution)
- represent individual vulnerabilities.
- dark color = highly infected

Final values:

- obtained after convergence.
- represent social vulnerabilities.



(a) Individual privacy vulnerability.

(b) Social privacy vulnerability.

Figure 14: Individual & Social privacy vulnerabilities.

User ID	Individual Vul	Social Vul
34	0.84	0.67
67	0.12	0.87
206	0.76	0.33
588	0.23	0.78

Table 7: Individual and social privacy vulnerabilities.

Outline

1. Introduction

2. Contribution

3. Conclusion

Conclusion

- The purpose of this work is to simulate a diffusion process of individual vulnerabilities.
  - The vulnerability of one user is the vulnerability of all users.
  - At the end of the diffusion (convergence), all users gets their social vulnerability scores.
- Future work
  - To propose mechanisms to improve the reputation of non-vulnerable users.
    - Suggest well known interlocutors with acceptable vulnerability scores.
  - To propose mechanisms to improve the vulnerability of reputed users.
    - recommend configurations and softwares.

Conclusion

- The purpose of this work is to simulate a diffusion process of individual vulnerabilities.
  - The vulnerability of one user is the vulnerability of all users.
  - At the end of the diffusion (convergence), all users gets their social vulnerability scores.
- Future work
  - To propose mechanisms to improve the reputation of non-vulnerable users.
    - Suggest well known interlocutors with acceptable vulnerability scores.
  - To propose mechanisms to improve the vulnerability of reputed users.
    - recommend configurations and softwares.

Thank you !

References

[1] *Nombre d'e-mails envoys par jour dans le monde 2017-2022 | Prvision*. [Online; accessed 1. Mar. 2019]. Mar. 2019 (p. 3, 4).