

# Évaluation de la confidentialité par un processus de diffusion de vulnérabilité

Aghiles DJOUDI

Sorbonne Université

April 19, 2019

# Outline

1. Introduction

2. Développement

3. Conclusion

# Context

## Introduction

	Monde (2018)	Monde (2022)	France (2018)
Nombre d'utilisateurs	3,8 milliards	4,2 milliards	25,9 millions
Nombre de comptes email	4,4 milliards	5,6 milliards	68 millions
Nombre d'adresses email par utilisateurs	1,7	1,9	2,1
Nombre de mails reçus chaque jour	<b>281 milliards</b>	<b>333 milliards</b>	<b>1,4 milliard</b>
Le marché de l'email	9,8 Mrds de \$	20,4 Mrds	?

Table 1: Les chiffres 2018 de l'email BibEntry2014Sep

- Email:
  - 50% de la population mondiale utilise l'email
  - Utilisation: 75% personnels, 25% professionnels
- Facebook:
  - 2,2 milliards d'utilisateurs actifs (29% mondiale)
  - **10 milliards de messages sont envoyés chaque jour**
  - 8,051 milliards de dollars



Figure 1: Services de messagerie.

# Context

## Introduction

	Monde (2018)	Monde (2022)	France (2018)
Nombre d'utilisateurs	3,8 milliards	4,2 milliards	25,9 millions
Nombre de comptes email	4,4 milliards	5,6 milliards	68 millions
Nombre d'adresses email par utilisateurs	1,7	1,9	2,1
Nombre de mails reçus chaque jour	281 milliards	333 milliards	1,4 milliard
Le marché de l'email	<b>9,8 Mrds de \$</b>	<b>20,4 Mrds</b>	?

Table 1: Les chiffres 2018 de l'email BibEntry2014Sep

- Email:
  - 50% de la population mondiale utilise l'email
  - Utilisation: 75% personnels, 25% professionnels
- Facebook:
  - 2,2 milliards d'utilisateurs actifs (29% mondiale)
  - 10 milliards de messages sont envoyés chaque jour
  - **8,051 milliards de dollars**



Figure 1: Services de messagerie.

# Motivation

## Introduction

- Donner un moyen aux utilisateurs de mesurer leur vulnérabilités
- Aider les utilisateurs à mieux configurer leur messagerie.
- Alerter les utilisateurs d'une nouvelle vulnérabilité.
- Sensibiliser les utilisateurs du niveau de diffusion des menaces.



Figure 2: Indice de confidentialité  
**maximilien\_privacyasaservice\_20**

# Défis

## Introduction

- ➡ Recommander des mesures de sécurité personnalisés
  - Nouveau mot de passe chaque période de temps
  - Sécuriser l'échange avec des comptes vulnérables
  - Adapter les permissions aux changements
- ➡ Calculer la vulnérabilité de l'environnement social
  - Calculer le niveau de vulnérabilité des interactions
  - Calculer le niveau d'influence entre les utilisateurs.
- ➡ Calculer la vulnérabilité du chemin des messages
  - Identification des serveurs MTA
  - Attribuer une note de confiance à chaque serveur
  - Calculer la confiance moyenne du chemin.

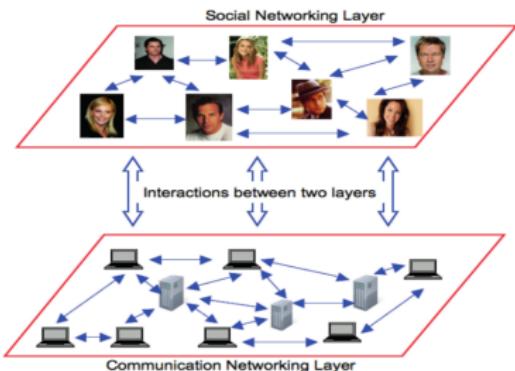
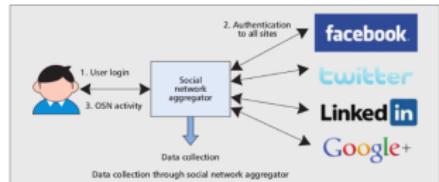


Figure 3: Interaction sociale.

# Défis

## Introduction

- ➡ Recommander des mesures de sécurité personnalisés
  - Nouveau mot de passe chaque période de temps
  - Sécuriser l'échange avec des comptes vulnérables
  - Adapter les permissions aux changements
- ➡ Calculer la vulnérabilité de l'environnement social
  - Calculer le niveau de vulnérabilité des interactions
  - Calculer le niveau d'influence entre les utilisateurs
- ➡ Calculer la vulnérabilité du chemin des messages
  - Identification des serveurs MTA
  - Attribuer une note de confiance à chaque serveur
  - Calculer la confiance moyenne du chemin.

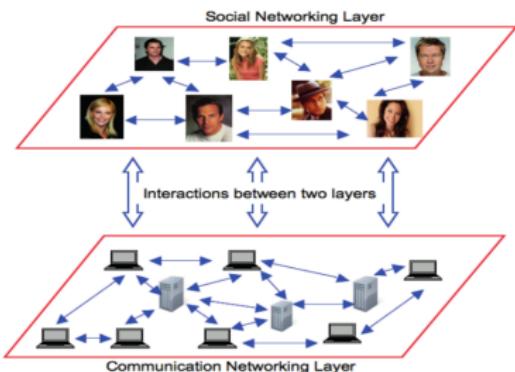
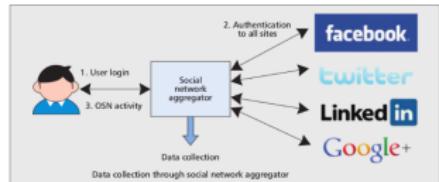


Figure 3: Interaction sociale.

# Contributions

## Introduction

- ➡ Estimation de l'indice de confidentialité social.
  - Vulnérabilité individuelle -> Vulnérabilité sociale.
  - Processus de diffusion de vulnérabilité.
  - Relation entre confiance et vulnérabilité.
  - Données: É-mails de Enron & Caliopen.



Figure 4: La vulnérabilité d'un utilisateur est la vulnérabilité de tous.

# Outline

1. Introduction

2. Développement

3. Conclusion

# Outline

- 1. Introduction
- 2. Développement
  - 1. Travaux connexes
  - 2. Processus de diffusion
  - 3. Expérimentation
  - 4. Exploitation des résultats
- 3. Conclusion

# Outline

1. Introduction

2. Développement

3. Conclusion

**1. Travaux connexes**

2. Processus de diffusion

3. Expérimentation

4. Exploitation des résultats

## Travaux connexes

### Comparaison

Travaux	Contribution	Performance
<b>gandouz_protect_2012</b> Protect U	Classification des interlocuteurs	Configuration des listes d'amis
<b>fang_privacy_2010</b> Privacy Wizard	Classification des interlocuteurs	Configuration des permissions
<b>frey_social_2011</b> SocialMarket	Intérêt communs	Évaluation des relations de confiance
<b>yongbozeng_study_2015</b> TAPE	Fuite d'information	Évaluation de la diffusion de l'information
<b>hameed_lens_2011</b> LENS	Protection anti-spam	Évaluation des émetteurs de courriels
<b>tran_social_2010</b> SocialEmail	Classer les chemins des msg	Évaluation de la fiabilité du message
<b>nepali_sonet_2013</b> Privacy Index	Visibilité, sensibilité	Évaluation de l'exposition des messages

Table 2: Contributions des travaux existants.

# Outline

1. Introduction

2. Développement

3. Conclusion

1. Travaux connexes

**2. Processus de diffusion**

3. Expérimentation

4. Exploitation des résultats

# Etape 1: Calcule de la vulnérabilité individuelle

## Méthode

### → Entrée:

- Vulnérabilité de la machine utilisée:
  - \* Connexion réseaux (privé (1) ou publique (2))
  - \* Type d'architecture: Ethernet, 5G, 4G, Wifi (1:4)
  - \* Système d'exploitation (Windows, Unix) (1:2)
  - \* Navigateur web (1:10)
- Vulnérabilité du compte utilisé
  - \* Mdp utilisé, mode de récupération des mdp (1:5)
  - \* Nombre de sessions ouvertes en même temps.(1:nbr)
  - \* Mode de chiffrement, signature, version TLS

### → Sortie:

$$Pi = \sum_i^n \frac{w * V}{n}$$

(1)



Figure 5: Vulnérabilité individuelle.

- \* Pi: Vulnérabilité individuelle
- \* w: Poids de chaque vulnérabilité
- \* V: Les vulnérabilités citées au dessus

## Etape 2: Calcule de la réputation des utilisateurs

### Méthode

→ Entrée:

- Fréquence d'utilisation de la messagerie.
- Horaire, durée des échanges (1:5)
- % des échanges chiffrés, signés, claires (1:3)
- Importance des interlocuteurs: Liste favoris (2), noir(1)
- Type de données: Texte, images, vidéos, script (1:4)

→ Méthode:

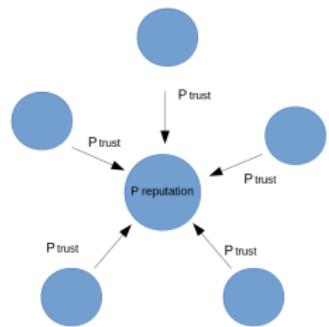
- Loi binomiale

→ Output:

$$P(\text{reputation}) = P(X \geq 1) = 1 - (1 - P(\text{trust}))^n \quad (2)$$

→ Where,

- \* X: Niveau de confiance,  $X \sim B(n,p)$
- \* n: deg(noeud)
- \*  $P(X=1)$ : La probabilité de se faire attribué une confiance par un interlocuteur



$$P \text{ reputation} = 1 - (1 - P \text{ trust})^n$$

Figure 6: Niveau de réputation.

## Etape 3: Calcule de la vulnérabilité sociale

Théorie de l'influence sociale de Freidkin

➡ Entrée:

- $Y^{(1)}$  = Vecteur des vulnérabilités individuelles de N utilisateurs (eq 1)
- $\alpha$  = Le niveau de réputation (d'influence) de chaque utilisateur (eq 2)
- $M$  = Matrice d'adjacence  $N \times N$

➡ Modèle:

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

➡ Sortie:

- $Y^{(t)}$  = Vecteur des vulnérabilités sociales des N utilisateurs

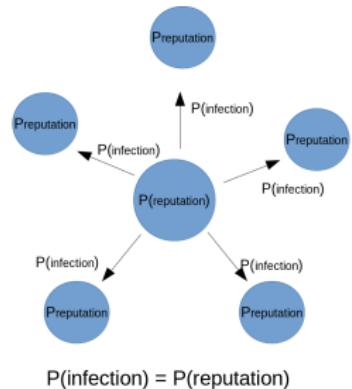


Figure 7: Vulnérabilité Sociale.



# Etape 3: Calcule de la vulnérabilité sociale

Théorie de l'influence sociale de Freidkin

Propriétés formelles du modèle:

- Lorsque l'influence d'un utilisateur est élevé, le modèle se réduit aux:
  - vulnérabilités moyennes de ses amis pondérées par leur niveaux de confiances.

$$Y^{(t)} = 1 * M Y^{(t-1)} + (1 - 1) Y^{(t-1)} \quad (3)$$

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

- En absence d'influence, le modèle se réduit à:
  - sa propre vulnérabilité pondérée par le niveau de méfiance de ses amis

$$Y^{(t)} = 0 * M Y^{(t-1)} + (1 - 0) Y^{(t-1)} \quad (3)$$

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

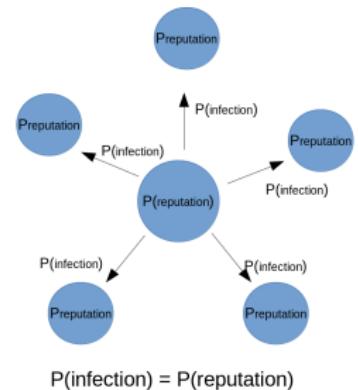


Figure 8: Vulnérabilité sociale.

# Outline

1. Introduction

2. Développement

3. Conclusion

1. Travaux connexes

2. Processus de diffusion

3. **Expérimentation**

4. Exploitation des résultats

# Expérimentation

## Expérimentation

Paramètre	Valeur
Utilisateurs	958
Messages	6966
Diamètre	958
# de msg en moyenne	2.413361
Densité des msg	0.00252
Modularité	0.654600
Distance moyenne	3.042114

Table 3: Propriétés des données Enron.



Figure 9: Enron logo.

Paramètre	Valeur
Utilisateurs	5885
Messages	26547
Diamètre	2096
# de msg en moyenne	9.02192
Densité des msg	0.001533
Modularité	0.86526
Distance moyenne	3.914097

Table 4: Propriétés des données Caliopen.



Figure 10: Caliopen logo.

# Outline

1. Introduction

2. Développement

3. Conclusion

1. Travaux connexes

2. Processus de diffusion

3. Expérimentation

4. Exploitation des résultats

# Résultats

## Comparaison

Valeurs initiales:

- ▶ générées aléatoirement (distribution normale)
- ▶ représentent les vulnérabilités individuelles.
- ▶ couleur foncée = vulnérabilité élevé

Valeurs finales:

- ▶ obtenu après convergence.
- ▶ représentent les vulnérabilités sociales.



(a) Vulnérabilité individuelle.

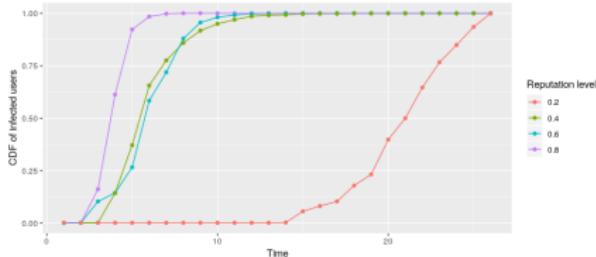
(b) Vulnérabilité Sociale.

Figure 11: Vulnérabilité individuelle & sociale.

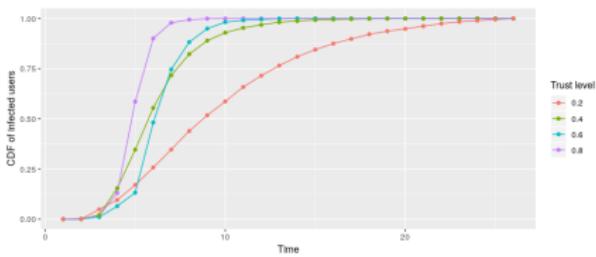
User ID	Vul individuel	Vul sociale
34	0.84	0.67
67	0.12	0.87
206	0.76	0.33
588	0.23	0.78

Table 5: Différence entre les vulnérabilités individuelles et sociales en matière de protection de la vie privée.

# Exploitation des résultats



(a) Les données de Enron.

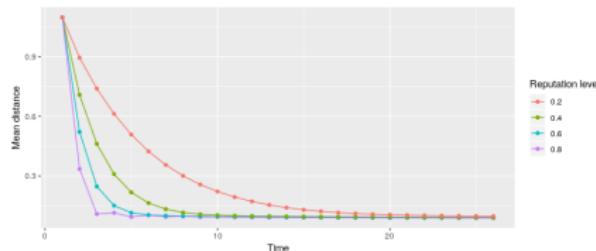


(b) Les données de Caliopen.

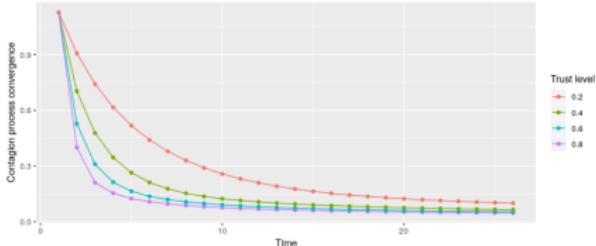
Figure 12: Fonction de distribution cumulative des utilisateurs infectés.

- Les utilisateurs avec des valeurs de réputation élevées contribuent considérablement à la diffusion
- Ils diffusent leur vulnérabilités rapidement et largement dans la messagerie.

# Exploitation des résultats



(a) Les données de Enron.



(b) Les données de Caliopen.

Figure 13: Convergence du processus de diffusion.

- Attribuer une confiance à des utilisateurs vulnérables leur permet d'obtenir un niveau de réputation élevé.
- Par conséquent, infecter l'ensemble des valeurs de vulnérabilité de la messagerie.

# Outline

1. Introduction

2. Développement

3. Conclusion

# Conclusion

- ➡ Le but de ce travail est de simuler un processus de contamination des vulnérabilités individuelles.
  - La vulnérabilité d'un utilisateur est la vulnérabilité de tous.
  - A la fin de la diffusion, tous les utilisateurs auront un indice de vulnérabilité social.
- ➡ Travaux futures
  - Proposer des mécanismes pour améliorer la réputation des utilisateurs non-vulnérables.
    - \* Suggérer des interlocuteurs bien réputés avec des indices de vulnérabilité acceptables.
  - Proposer des mécanismes pour améliorer la vulnérabilité des utilisateurs réputés.
    - \* Recommander des configurations et des logiciels.

## Conclusion

- ➡ Le but de ce travail est de simuler un processus de contamination des vulnérabilités individuelles.
  - La vulnérabilité d'un utilisateur est la vulnérabilité de tous.
  - A la fin de la diffusion, tous les utilisateurs auront un indice de vulnérabilité social.
- ➡ Travaux futures
  - Proposer des mécanismes pour améliorer la réputation des utilisateurs non-vulnérables.
    - \* Suggérer des interlocuteurs bien réputés avec des indices de vulnérabilité acceptables.
  - Proposer des mécanismes pour améliorer la vulnérabilité des utilisateurs réputés.
    - \* Recommander des configurations et des logiciels.

Thank you !

## References