INF6422 – Concepts avancés en sécurité informatique

Travail pratique 1 – Hiver 2016

Propagation de logiciel malveillant

	Section	Date d'exécution	Date de remise
Travail pratique 1	01 (B2)	19 janvier 2016	2 février 2016

Consignes:

- La remise se fait par équipe et via Moodle seulement.
- Le rapport remis doit être en format PDF. Libre à vous de le construire dans le format de votre choix (.docx, .odt, .tex, etc.).
- Le rapport doit inclure une page titre contenant le titre du cours, du laboratoire, vos noms et matricules.
- Le rapport doit être rédigé en français.
- Le rapport doit être remis avant 12h00 le jour de la remise. Une pénalité de 10% sera appliquée par jour ouvrable de retard.

1. Mise en contexte

1.1 L'épidémiologie pourrait être définie comme l'étude des rapports existant entre les maladies ou tout autre phénomène biologique, et divers facteurs susceptibles d'exercer une influence sur leur fréquence, distribution et évolution. Entre d'autres mots, l'épidémiologie s'intéresse aux facteurs qui influencent la santé des populations. Plus particulièrement, l'épidémiologie s'intéresse, entre autre, à étudier la dynamique de propagation des maladies infectieuses afin d'établir des stratégies de prévention et d'intervention permettant de diminuer l'impact sur la santé publique. À cet effet, la modélisation mathématique s'est révélée particulièrement intéressante afin de simuler des scénarios épidémiologiques, d'évaluer les risques associés et de quantifier l'efficacité et l'impact de différentes méthodes d'intervention et de prévention. Plusieurs approches peuvent être retenues, telles que les simulations numériques, les modèles déterministes ou encore les modèles stochastiques. Chaque approche présente des avantages et des inconvénients. Il convient donc de choisir la méthode la plus appropriée en fonction des questions de recherche auxquelles vous souhaitez répondre.

1.2 Appliquée à la sécurité informatique, l'épidémiologie pourrait être vue comme l'étude des différents facteurs qui influencent la fréquence, la distribution et l'évolution des logiciels malveillants. Plus particulièrement, l'approche épidémiologique a inspiré de nombreux travaux de recherche [1] portant sur l'étude de la propagation des logiciels malveillants. Le présent laboratoire vous permettra de vous familiariser avec certaines approches mathématiques fréquemment utilisées afin de modéliser la propagation de logiciels malveillants au sein d'un réseau.

2. Modèle déterministe [5 points]

Une approche très répandue dans l'étude de la propagation des logiciels malveillants consiste à développer un modèle déterministe basé sur les concepts de compartiments et de règles [2]. Les compartiments servent à diviser la population étudiée en différentes classes et les règles à définir les conditions de transition entre chacune des classes.

2.1 En vous basant sur l'article « Optimising Networks Against Malware » [3], quel modèle comportemental (SI, SIS, SIR) s'appliquerait et pourquoi? Justifiez votre réponse en expliquant quel modèle s'applique et pourquoi les autres modèles ne s'appliquent pas.

2.2 Donnez les équations différentielles de votre modèle déterministe en fonction du temps. Chaque compartiment doit avoir sa propre équation différentielle. Vos équations doivent au minimum contenir les paramètres suivant : S, I, N et λ, où N représente la taille totale de la population et λ le nombre de contacts par machine par unité de temps.

3. Simulation numérique [5 points]

Lors de la question précédente, vous avez développé un modèle théorique basé sur un système d'équations différentielles. Heureusement, il existe une solution analytique à ce système afin de représenter le nombre de machines infectées en fonction du temps :

$$I(t) = \frac{I_0 N}{(N - I_0)e^{-\lambda t} + I_0}$$

3.1 Considérez que *t* varie de 1 à 1200 secondes avec les conditions suivantes : $I_0 = 1$, N=3000 et $\lambda = 0.02$. Représentez graphiquement I et S en fonction du temps sur une période de 1200 secondes. Identifiez clairement les courbes I(t) et S(t).

3.2 Faites varier le paramètre λ (0.01, 0.025, 0.05). Représentez graphiquement les différentes courbes de I(t) sur le même graphique. Commentez les différents graphiques obtenus en expliquant la signification quantitative du paramètre λ et son impact sur la vitesse de propagation d'un vers.

4. Modèle stochastique [2 points]

Dans l'article « Optimising networks Against Malware » [3] l'auteur utilise un modèle stochastique basé sur les chaînes de Markov afin de modéliser la propagation d'un vers dans un réseau

4.1 Expliquez les caractéristiques d'un modèle stochastique et pourquoi ce type de modèle s'applique dans le contexte de l'article. Est-ce qu'une approche déterministe aurait été préférable?

5. Performance et optimisation [8 points]

Toujours dans l'article « Optimising networks Against Malware » [3], l'auteur étudie l'effet de la topologie du réseau sur la vitesse de propagation d'un vers informatique.

5.1 Appliquez le concept du double-tétraèdre étudié en classe à l'article. Situez votre double-tétraèdre dans un contexte d'optimisation, toujours en vous référant à l'article. Expliquez quels sont les aspects de votre double-tétraèdre qui sont fixes, et ceux qui sont modifiés.

5.2 Quels autres aspects/méthodes (autres que la topologie du réseau) pourraient être étudiés dans le cadre d'un problème d'optimisation où l'objectif est de limiter la vitesse de propagation d'un logiciel malveillant dans un réseau? Donnez au minimum deux exemples.

5.3 Il a été démontré qu'une plus grande biodiversité au sein d'un écosystème permettait de ralentir la propagation des virus ou des bactéries dangereuses. Comment pourriez-vous appliquer le concept de diversité au sein d'un réseau informatique? Quelle(s) forme(s) prendrait cette diversité?

Références

[1] Wang, Y., Wen, S. and Xiang, Y. Modeling the Propagation of Worms in Networks: A Survey. IEEE Communications Surveys & Tutorials, Vol.16, No.2, 2014.

[2] http://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology

[3] Bureau, P.-M., Fernandez, J.M. Optimising Networks Against Malware. IEEE, IPCCC, 2007.