Evaluating the Provision of Botnet Defences using Translational Research Concepts

Dilara Acarali, Muttukrishnan Rajarajan and Nikos Komninos
School of Mathematics, Computer Science and Engineering, City, University of London, London, U.K.

Keywords: Cyber-security, Botnets, Translational Research, Implementation Science, Botnet Propagation.

Abstract: Botnet research frequently draws on concepts from other fields. An example is the use of epidemiological models when studying botnet propagation, which facilitate an understanding of bot spread dynamics and the exploration of behavioural theory. Whilst the literature is rich with these models, it is lacking in work aimed at connecting the insights of theoretical research with day-to-day practice. To address this, we look at botnets through the lens of implementation science, a discipline from the field of translational research in health care, which is designed to evaluate the implementation process. In this paper, we explore key concepts of implementation science, and propose a framework-based approach to improve the provision of security measures to network entities. We demonstrate the approach using existing propagation models, and discuss the role of implementation science in malware defence.

1 INTRODUCTION

Botnets are malware-based platforms built illicitly on vulnerable networks to serve a cyber-crime agenda. In botnet research, propagation modelling often appropriates epidemiological models of diseases (Brauer, 2008) to identify dissemination factors, to devise immunisation strategies (Yong et al., 2012), and to predict reach and speed. However, findings may not be applied appropriately in day-to-day practice. Networks are highly variable, meaning that any proposed security measure will fit some scenarios but fall short in others. Translational research (TR) is a field of public health care for converting experimental results into care for patients (Rubio et al., 2010) (Woolf, 2008). We believe that botnet research would benefit from a similar approach. Standard methods for observing and evaluating security provisions would ensure that a) they are fully utilised in practice, and b) they are applied consistently and effectively for maximum impact.

In this paper, we aim to start a discussion about using TR and implementation science (IS) to effectively apply botnet research findings to real-life networks. This is particularly relevant to botnet propagation, where mitigation can prevent worsening outcomes. To our knowledge, this topic has not previously been explored. Our contributions are 1). an adaptation of the Dynamic Sustainability Framework (DSF) to apply IS methods to botnet defence, 2). an analysis of propagation-based usage scenarios, demonstrating how DSF can help deliver better protection, and 3). a discussion of the potential role of IS in malware defence. Section 2 provides a background on TR. Section 3 introduces the DSF and demonstrates its use, whilst Section 4 provides a discussion on the usage of IS. Related works are covered in Section 5, and we conclude in Section 6.

2 BACKGROUND

Translational research (TR) investigates the process of converting scientific knowledge into practical solutions for standard practice, described as "the interface between basic science and clinical medicine" (Woolf, 2008). Basic science is the pursuit of knowledge with no practical consideration (Rubio et al., 2010) (analogous to bot case studies). Clinical research introduces patients for behavioural analysis (Rubio et al., 2010) (analogous to botnet propagation simulations). TR has 2 parts (Figure 1). The first (called *T1*) applies lab-based studies to clinical trials, whilst the second (*T2*) uses obtained results to aid practical decision-making (Rubio et al., 2010) (Woolf, 2008). *T1* requires in-depth scientific knowledge and the relevant tools for innovative research, whilst *T2* requires understanding of communities, culture, and work envi-

Acarali, D., Rajarajan, M. and Komninos, N.
Evaluating the Provision of Botnet Defences using Translational Research Concepts.
DOI: 10.5220/0006884405140519
In *Proceedings of the 15th International Joint Conference on e-Business and Telecommunications (ICETE 2018) - Volume 2: SECRYPT*, pages 514-519
ISBN: 978-989-758-319-3

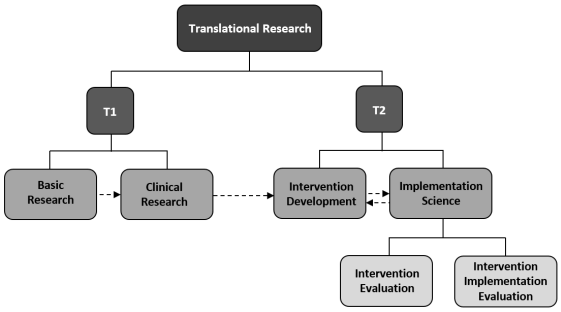

Figure 1: Breakdown of translational research into subfields.

ronments (Woolf, 2008) (Rubio et al., 2010). *T2* has 2 stages: 1) the translation of results to practice, and 2) the evaluation of this process (Woolf, 2008), also known as implementation science (IS). For botnets, we focus on *T2*, with the development, application, and evaluation of security policies.

To reap maximum benefit from *T1* results, there must be a systematic, well-observed process of implementing new ideas into existing systems (Woolf, 2008) (Khalil, 2016) (Bauer et al., 2015). IS broadly refers to the application of new approaches (Woolf, 2008) with performance evaluation and due consideration of surrounding influential factors (Khalil, 2016). A single new method is an intervention, and a package of interventions (with a common goal) is a strategy (Bauer et al., 2015). IS considers the possibility of inconsistent application leading to "quality gaps" (Bauer et al., 2015), and as a result, enables us to observe and evaluate natural variation and responses to interventions (Bauer et al., 2015). Health care literature organises this process into frameworks.

3 FRAMEWORK

3.1 Outline & Definitions

We use the Dynamic Sustainability Framework (DSF), proposed by Chambers et al. (2013). It incorporates internal and external context environments, as well as temporal change, allowing us to evaluate the intervention over time. This means that DSF focuses on sustainability to address "voltage drops" and "program drift" (Chambers et al., 2013). It enables the continuous improvement of results by exposing the intervention to changing populations, and adapting it to better fit emerging conditions (Chambers et al., 2013). In networks, segmentation results in internal and external spaces. User behaviours, business practices, and new technologies add variability that can affect intervention roll-out. These factors impact the uniform implementation of new policies, highlighting the need for spatio-temporal sustainability. Evolving malware also means interventions must be robust and adaptable.

DSF splits the process into 3 layers; the intervention, the practice setting, and the ecological system (Chambers et al., 2013). We split the practice setting into 2 parts for a layered inner environment to mimic the layers of an organisational network (Figure 3). Influencing factors at each layer are defined as a set of constructs. Our adaptations of these are listed in Table 1. The intervention includes actionable steps, desired outcomes, delivery platforms, and practitioners. It is deployed within the practice setting, which represents all internal influences (e.g. systems, resources, and staff) (Chambers et al., 2013). This sits within the ecological system, representing external factors like legislation, regulation, and market forces (Chambers et al., 2013). DSF operates over periods, denoted as (T_0, T_1, ... T_n). Hence, the intervention process covers each layer across all periods, revealing changes in implementation. This approach is unique to DSF, designed to improve patient health by adapting interventions as required (Chambers et al., 2013). Observing how the network state changes with time under our defensive interventions is key to botnet mitigation.

3.2 Usage Scenarios

We now outline how the framework may be applied to impede botnet propagation. Three epidemic scenarios are considered, in: 1) an enterprise network, 2) a mobile network, and 3) a social network. We use epidemics as our practical example because propagational success plays a large role in determining a botnet's eventual impact. The propagation stage also provides opportunities for mitigation of future attacks. Additionally, epidemic modelling is widely used in botnet literature to test behavioural theories and spread parameters. Lastly, epidemic modelling fits into the clinical research stage of *T1*, and we can carry the results over to the real-life settings of *T2*.

3.2.1 Enterprise Network Epidemics

Yong et al. (2012) used a modified *SIRS* model to define reproductive ratio R_0 in relation to 2 types of *I* (infected) node; hidden and active. They recommended keeping $R_0 < 1$ by a) minimising *S* (susceptibles), b) increasing removal rate, and c) decreasing transmission rate. Success depends on defence architecture, location, and deployment time. Dagon et al. (2006) used the *SIR* model to capture the diurnal nature and regional bias of botnets. Observing a "natural quarantine effect" (Dagon et al., 2006),

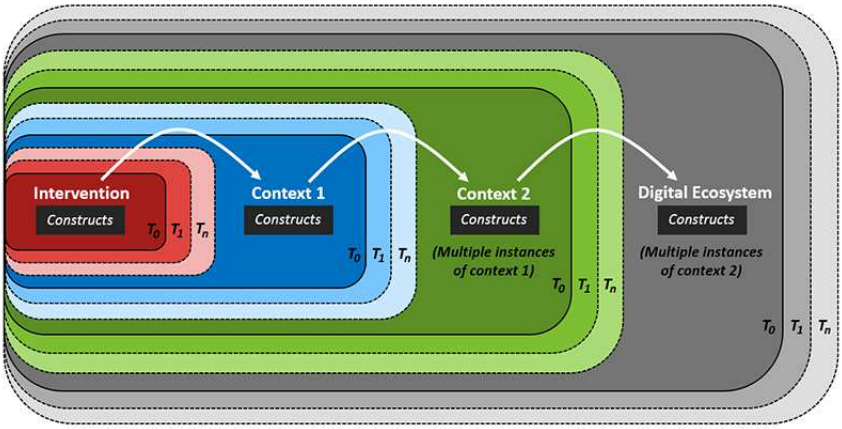

Figure 2: Diagram of updated Dynamic Sustainability Framework (Chambers et al., 2013), showing the 4 layers and their constructs, along with increments in time (T_X).

they found that worms have optimal release times and suggested appropriate surveillance and geographical/temporal prioritisation for defenders.

We now apply the framework to bolster defences against a possible epidemic outbreak. Based on the results above, this network's intervention includes improving user awareness and patching policies (to minimise *S*), reducing R_0, identifying risky malware deployment spots and times, centralising security systems across departments, and deploying a resilient infrastructure (for quarantining to minimise contact). The patients are network nodes, and practitioners are the users, admins, policy makers, and security directors. The 1st internal context is a network segment (e.g. a departmental LAN), whilst the 2nd is the wider enterprise, which is made up of multiple segment contexts. Meanwhile, the external context is the connected community around this enterprise. For evaluation, we need to consider the quality of training and user receptiveness, adherence to policies, and system performance. Collecting data on these factors then provides a feedback loop, allowing both the strategy and the implementation to be improved (Chambers et al., 2013).

3.2.2 Mobile Network Epidemics

The *SIR* model is used by Khouzani et al. (2012) in their proposed malware quarantine scheme based on regulating each node's communication range. When the reception gain of *S* nodes is decreased, contact frequency also drops. This should be balanced against the lowest acceptable QoS for the network. Lu et al. (2016) employs epidemic principles stochastically to demonstrate that spread in WiFi networks is dependent on node proximity, and specifically on node density, wireless transmission range, and the node mobility radius. Propagation eventually stops if the mobility radius is limited enough (Lu et al., 2016).

This time, an epidemic is already underway. The intervention involves responsive measures like tracking of *I* nodes, grounding (where possible) of mobile devices (reducing their mobility radius), reducing the gains of *S* nodes near to *I* nodes, and performing targeted clean-up. Users should be notified and may assist in grounding. As before, nodes represent patients, whilst users, admins, and owners of WiFi infrastructure are the practitioners at various levels. The 1st internal context is the specific cell or geographic location of the infection, whilst the 2nd is the wider cellular/geographic region surrounding it. The external context is the wider WiFi network, and the Internet. The effectiveness of this implementation will be impacted by users' attitudes to the event and to the measures taken. For example, the reduction in network quality might garner a negative reaction. The practicality of the grounding policy, ability to accurately detect and track *I* nodes, and the efficiency of the recovery process also need to be measured for evaluation, and can provide opportunities for improvement.

3.2.3 Social Network Epidemics

Sanzgiri et al. (2012) used the *SI* setup to model Twitter-based botnets, with propagation impacted by degree of followers, activity levels, and the level of response/interest in tweets, contributing to the probability of user clicks on malicious URLs. They suggest limiting the sharing of links, or better control and monitoring of short-URLs, noting that individuals may not be aware of potential risks. Yan et al. (2011) used epidemiological concepts to demonstrate that propagation rate is influenced by friendship networks, activity, and the infection source. They sug-

Table 1: Outline of network constructs at each layer of updated DSF (Chambers et al., 2013).

Constructs	Description
Intervention	
Components	Collection of individual elements that make up the intervention.
Practitioners	Individuals who will implement the intervention across all levels.
Characteristics	Collection of descriptive characteristics for the intervention.
Aims	Targeted outcomes of the intervention.
Delivery Vectors	Means of delivery of the intervention.
Practice Setting	
Context 1 - Local Network	
Infrastructure	Topology and design of the network at this level, including defences.
Systems	Collection of specialised systems or technical resources used.
People	Collection of local staff, including count, roles, and hierarchy.
Culture & Climate	Social environment created by management, individuals, and group attitudes.
Business Function	Function of this context in the wider network.
Training	Nature and availability of knowledge and skills transfer to practitioners.
Supervision	Collection of regulators and enforcement methods.
Context 2 - Wider Network	
Infrastructure	Topology and design of the wider network, interconnection of local contexts.
Systems	Collection of systems or resources used organisation-wide, e.g. email services.
People	Collection of management staff, including count, roles, and hierarchy.
Culture & Climate	Social environment created by management based on brand/organisation ethos.
Business Model	Goals of organisation, plus functions/services, financial and implementation plans.
Digital Ecosystem	
Regulation & Legislation	Industry standards and national legislature influencing operations.
Market Trends	Current popular technologies, activities of competitors.
Populations	Characteristics of populations engaging with devices and services.
Usage Culture	Attitudes towards technology and current trends in behaviour.
Partners	Collaborating third-parties, shared infrastructure.
Upcoming Technology	Collection of technologies in the pipeline for mass deployment.

gest the use of an early warning system that propagates alerts amongst users to reduce susceptibility, and the use of centralised servers to monitor and sanitise suspicious URLs (Yan et al., 2011).

To defend a social network against a potential epidemic, the intervention incorporates early detection, suspension of infected accounts to reduce activity, increased monitoring for URLs shared by highly-active highly-connected users, and an in-built URL shortening service. Users should be notified of risks and ongoing security events. The internal context is the user community or friendship group, encapsulated by the wider social network, made up of multiple similar communities. The external context is the rest of the Internet. Patients could be user accounts or users themselves, with practitioners being the platform admins and security staff. Evaluation needs to consider the culture/brand of the social networking platform, user engagement with security advice, public reception to new policies (including privacy concerns), and the effectiveness of early detection systems. Suggested data collection includes user feedback, activity logs over time and region, frequency and distribution of URL and short-URL use, and performance metrics for detection systems and URL-shortening services.

4 DISCUSSION

The scenarios in Section 3.2 show how the framework can be applied to different situations, populations, and infrastructures. It enables us to identify specific actions via the intervention strategy, and then to determine areas where implementation may be insufficient. The provided suggestions are based on literature and experience, but real-life application will flag significant, unpredicted problem areas. Consideration of these factors over time means that improvements can be made to the intervention incrementally, providing sustained long-term solutions. Being able to achieve this with a standardised framework is highly beneficial, as it eases the process and makes different interventions comparable. Sustainability of an intervention should be a top priority. Networks become vulnerable due to changes in technology and usage over time, a process mirrored by evolving malware threats. Therefore, it is difficult to predict what the future will bring. Furthermore, when deploying new technology, it is not possible to consider a priori every scenario surrounding its use. Hence, interventions must be flexible enough to incorporate new insights. The aim of DSF is the improvement of interventions

Table 2: DSF applied to 3 usage scenarios, with interventions (Strategy) derived from epidemic modelling results. Impact Factors determine intervention success, and Measures denote evaluative considerations.

	Scenario 1	**Scenario 2**	**Scenario 3**
Settings	-Enterprise	-Mobile	-Social
Strategy	-Train users for response -Shrink *S* via patching & updates -Centralise defensive systems -Find optimal spread spots, times -Add network contingencies to take segments offline	-Inform users of situation -Ground infected devices -Drop reception gains of close *S*'s -Targeted cleaning of detected *I*'s	-Add early detection systems -Send regular advice to users -Suspend infected accounts -Track URLs of active users -Add in-built short-URL service
Impact Factors	-Engagement of users in training -Uniform delivery of training -Adequate supervision -Adherence to patching policy -Efficiency of recovery process -Robustness of contingencies	-Culture & attitudes of users -Ability to ground devices -Ability to monitor movement -Efficiency of gains reduction -Efficiency of recovery process	-Culture & attitudes of users -Reception of new policies -User response to security advice -Effectiveness of early detection -Uptake of short-URL service
Measures & Feedback	-User skills levels, experience -Vulnerability status -Usage of spread spots -Activity at optimal times -Defence system performance	-User experience -Current QoS -Current estimated *S* count -Current estimated *I* count -Current *R* count	-User feedback -Activities over time & region -Freq. & distribution of URL use -Detection system performance -Short-URL service performance

- through constant measurement from a dynamic context, interventions can be optimised (Chambers et al., 2013).

In both health care and cyber-security, a lot of resources are put into *T1*. However, *T2* efforts may actually have the larger practical impact (Woolf, 2008). Therefore, the proper translation and delivery of new research should receive greater funding and focus. Well-realised *T2* endeavours can bring return-on-investment for *T1* processes (Woolf, 2008). Additionally, a more pragmatic outlook may be beneficial in epidemiological studies. Galea (2013) argues that epidemiology should shift to become more consequentionalist (where models' worth are measured based on results rather than theory), suggesting that, given limited time and resources, this would help to prioritise actions to maximise overall health (Galea, 2013). Finally, TR is concerned with the overall improvement of public health, including the observed environment, plus the overall distribution of health across different regions (Galea, 2013). We suggest that a similar mentality be adopted against botnets. The threat is worldwide and successful defence in one environment is insufficient - botnets formulated elsewhere can still attack this environment. Therefore, we should work collaboratively on global systems to mitigate bot malware at the highest level, collectively protecting national systems, businesses, homes, NGOs, and the shared Internet infrastructure. Only when this distribution of sustained network health can be achieved will the botnet threat truly lose its potency.

5 RELATED WORK

Dedeke (2017) discussed the NIST Cyber-Security Framework (CSF) (NIST, 2014), designed for organisations to identify and reduce security risk within their systems. CSF aims to create a paradigm shift from compliance-based to risk management-based defence, which should yield higher quality precautions (Dedeke, 2017). CSF has 3 sections, with the core housing 5 functions. These are 1). identification of requirements, 2). development of safety measures, 3). monitoring, 4). development of action plans, and 5). deployment of recovery strategies (Dedeke, 2017) (NIST, 2014). This aligns with some actions in our suggested intervention strategies. CSF outlines a coordination scheme between executive, business and implementation levels (NIST, 2014), but does not incorporate dynamic delivery contexts. Dedeke (2017) highlights the use of implementation tiers to track progress over time, as changes are not otherwise considered.

Bassam and Deborah (2010) presents an ESM (Enterprise Security Management) framework for building secure, well-structured enterprises with adaptive processes. The framework consists of 12 steps, starting with an analysis of requirements, followed by: identification of capability gaps, prioritisation of tasks, development of architecture, monitoring, and continuous realignment - described as the "periodic reassessment of requirements, capabilities and updating the architecture" (Bassam and Deborah, 2010). This allows the system to keep pace

with emerging threats. Layers within the enterprise are considered, with emphasis on information sharing across domains (Bassam and Deborah, 2010). The inclusion of change, and the layered view of enterprises, makes ESM similar to DSF. However, DSF is not enterprise-specific, allowing it to be used flexibly in many different scenarios.

These frameworks approach security implementation predominantly from a business perspective. However, they do not formalise a process for turning pure research (*T1*) into applicable methods (*T2*). As such, they do not focus on getting the best out of available knowledge. They also do not integrate time and internal/external contexts, and so cannot improve interventions against changing trends and populations. We believe that taking inspiration from the well-established field of health care provides a unique angle to exploit in generating new ideas for the maintenance of network health.

6 CONCLUSIONS

In health care, TR delivers research knowledge to patients. We have proposed that a similar approach be applied for botnet mitigation - bringing technical knowledge and innovative solutions more effectively to users and networks. To demonstrate this approach, we utilised the Dynamic Sustainability Framework, applying it to epidemic modelling scenarios for bot propagation. We suggested measurement techniques, highlighted key constructs, and discussed the evaluative process. IS deals with the impact and implementation of an intervention, allowing us to consider how we develop multi-faceted approaches, how/where we deliver them, what meaningful impact they have, and why they may be lacking. This is vital for improving and sustaining the health of networks. We hope that this work contributes towards a discussion about how we deliver new solutions and how TR may play a role in this.

REFERENCES

Bassam, S. F. and Deborah, L. F. (2010). Cyber Security Framework for Enterprise System Development: Enhancing Domain Security through ESM. In *Military Communications Conference, 2010-MILCOM 2010*, pages 924–929. IEEE.

Bauer, M. S., Damschroder, L., Hagedorn, H., Smith, J., and Kilbourne, A. M. (2015). An Introduction to Implementation Science for the Non-Specialist. *BMC Psychology*, 3(1):32.

Brauer, F. (2008). Compartmental Models in Epidemiology. In *Mathematical Epidemiology*, pages 19–79. Springer.

Chambers, D. A., Glasgow, R. E., and Stange, K. C. (2013). The Dynamic Sustainability Framework: Addressing the Paradox of Sustainment Amid Ongoing Change. *Implementation Science*, 8(1):117.

Dagon, D., Zou, C. C., and Lee, W. (2006). Modeling Botnet Propagation Using Time Zones. In *NDSS*, volume 6, pages 2–13.

Dedeke, A. (2017). Cybersecurity Framework Adoption: Using Capability Levels for Implementation Tiers and Profiles. *IEEE Security Privacy*, 15(5):47–54.

Galea, S. (2013). An Argument for a Consequentialist Epidemiology. *American Journal of Epidemiology*, 178(8):1185–1191.

Khalil, H. (2016). Knowledge Translation and Implementation Science: What is the Difference.

Khouzani, M., Altman, E., and Sarkar, S. (2012). Optimal Quarantining of Wireless Malware through Reception Gain Control. *IEEE Transactions on Automatic Control*, 57(1):49–61.

Lu, Z., Wang, W., and Wang, C. (2016). On the Evolution and Impact of Mobile Botnets in Wireless Networks. *IEEE Transactions on Mobile Computing*, 15(9):2304–2316.

NIST (2014). Framework for Improving Critical Infrastructure Cybersecurity. Last accessed 5th January 2018.

Rubio, D. M., Schoenbaum, E. E., Lee, L. S., Schteingart, D. E., Marantz, P. R., Anderson, K. E., Platt, L. D., Baez, A., and Esposito, K. (2010). Defining Translational Research: Implications for Training. *Academic Medicine: Journal of the Association of American Medical Colleges*, 85(3):470.

Sanzgiri, A., Joyce, J., and Upadhyaya, S. (2012). The Early (Tweet-ing) Bird Spreads the Worm: An Assessment of Twitter for Malware Propagation. *Procedia Computer Science*, 10:705–712.

Woolf, S. H. (2008). The Meaning of Translational Research and Why it Matters. *Jama*, 299(2):211–213.

Yan, G., Chen, G., Eidenbenz, S., and Li, N. (2011). Malware Propagation in Online Social Networks: Nature, Dynamics, and Defense Implications. In *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, pages 196–206. ACM.

Yong, W., Tefera, S. H., and Beshah, Y. K. (2012). Understanding Botnet: From Mathematical Modelling to Integrated Detection and Mitigation Framework. In *Software Engineering, Artificial Intelligence, Networking and Parallel & Distributed Computing (SNPD), 2012 13th ACIS International Conference on*, pages 63–70. IEEE.