An Improved Worm Mitigation Model for Evaluating the Spread of Aggressive Network Worms

C. Onwubiko, A. P. Lenaghan and L. Hebbes

***Abstract* — An enhancement to existing epidemiological worm models is proposed which is used to simulate the spread of aggressive worms within computer networks. The proposed model presents worm propagation dynamics in five state transitions in a finite state machine model. The results obtained from the simulation are used to compare the dependability of previous worm quarantine models.**

***Keywords* — aggressive worm, computer networks, finite state machine, modelling, simulation**

I. INTRODUCTION

IN this paper we model aggressive worms, provide appropriate countermeasures and examine the accuracy of worm quarantine models by comparing results from our simulation with the recorded data of Code Red II.

In the wake of recent network worm proliferation, especially its impact on business service networks, service level agreements and information security concerns, questions for more realistic models of worm defence have been raised. Recent worm attacks are becoming more aggressive, exploiting vulnerabilities in systems to infect other susceptible systems; some worms use compromised systems as vehicle to infect more systems. In contrast, the defences of existing worm models are inadequate in providing appropriate countermeasures to worm outbreak. This is because existing worm models lack detailed understanding of the different states of worm propagation dynamics, assuming decreasing worm infection rate and often lags behind the threat providing a diminishing level of protection as a result. Responding swiftly to evolving worm incidents requires both a detailed understanding of worm propagation states and coordinated deployment of adequate countermeasures, which include human and automated processes.

To adequately provide appropriate countermeasures, a realistic and improved worm mitigation model is required that demonstrates understanding of worm propagation dynamics, which helps predict and hence respond efficiently to threats. Fundamental to an improved worm mitigation model is the identification of the characteristics of worms, their attributes and observable effect on computer networks and information systems.

C. Onwubiko is the author to whom correspondence should be sent and with the Networking and Communications Group, Kingston University, Penrhyn Road, Kingston Upon Thames, KT1 2EE, UK (e-mail: k0327645@kingston.ac.uk)

A. P. Lenaghan and L. Hebbes are with the Networking and Communications Group, Faculty of Computing, Information Systems and Mathematics at Kingston University, Penrhyn Road, Kingston Upon Thames, KT1 2EE, UK, (Phone +44 (0)20 8547 2000; email: {A.Lenaghan, L.Hebbes}@kingston.ac.uk)

Aggressive computer worms possess the capability to take over the Internet in few seconds as shown by recent incidents. For instance the January 2003 SQL Slammer incident infected 75,000 computers in 10 minutes, while Code Red is reported to have infected 359,000 computers in about 14 hours in July 2001 [1]. Comparing statistics, it shows that SQL Slammer compromised 6,300,000 computers in about the same time duration as Code Red, making SQL Slammer a more aggressive worm-type. In modelling aggressive worms we carried out a series of simulations and analyses using the simple epidemic Susceptible-Infectious (SI) model, Kermack-McKendrick (SIR) epidemic model [2] before extending the SIR Model to include the classes of quarantine and removed systems. The extended model is referred to as the Improved Worm Mitigation Model (IWMM), which presents a five state worm propagation model: *susceptible*, *removed*, *infected*, *quarantined* and *recovered* but not strictly in this order.

Section II discusses related work, definition and classification of aggressive worms; section III explains our methodology and modelling domain. In section IV we propose the improved worm mitigation model (IWMM) for modelling aggressive self-propagating malicious logic. We conclude the discussion with a summary in section V.

II. RELATED WORK

Research on self-replicating, self-propagating malicious logic, its effect, how it propagates, its nature and how it can be monitored have attracted a lot of attention at present because of the huge impact recent viruses and worms have had on networks and computing systems. A good number of contributions in this area exist [3,4,5,6].

A. Definition

Worms are programs that exploit faults on vulnerable system resources, which are classified as operational, external, human-made, software, malicious, deliberate and permanent according to Avizienis [7]. Worms as programs are deliberately crafted to exploit known specific vulnerabilities in certain applications. For instance, Code Red scanned the Internet but could only infect the Windows 2000 operating system with Internet Information Services (IIS) server installed, but could not infect Windows NT [8]. When worms are released to the Internet, they quickly scan systems they get in contact with, especially those in the 'hitlist' – list of target systems – looking for vulnerable systems to infect; when they find and infect some systems, they use these systems as vehicle to infect other susceptible and connected systems. Here we define the following terms:

Susceptible systems S(t): systems that have the propensity to be infected by the worm. This class of system have not developed immunity to the worm. *Infected systems I(t):* systems that are already under the influence of the infection; they are contaminated with the worm and require treatment. *Recovered systems R(t):* systems that were once infected, but now treated to develop immunity to the infection. *Removed systems U(t):* a subset of the susceptible systems that are disconnected from the network for immunisation, but not previously infected.

B. Classification of Aggressive Worms

An aggressive worm in this paper is characterised as: i) spreading through the network; ii) continuously activated; iii) having early saturation propagation; iv) having no topological constraints; v) their rate of spreading increases proportionately with the number of infected systems.

By *worm spreading through the network*, we mean, worm infection that propagates by exploiting vulnerable systems connected to the network (Internet for instance) and using these infected systems as vehicle to infect other susceptible systems with exploitable vulnerabilities and connected to the network. Not every system on the Internet is susceptible; susceptible systems are systems that possess exploitable vulnerability for a particular worm attack. By *continuously activated*, we mean, worms that randomly and continuously scan connected systems to find and infect susceptible systems. Again, the scanning and infecting rates are metaphors for early saturation. The account of W32.Blaster incident, exploits a flaw in Microsoft Windows' Remote Procedure Call (RPC) Distributed Component Object Model (DCOM) interface. RPC DCOM enables Microsoft software components, including HTTP, to communicate, which means any of such systems connected to the Internet is susceptible to this particular worm. The point is that W32.Blaster exploits susceptible connected systems as it spreads, thus there were no topology constraints with this worm; the same as the Code Red and Melissa worms.

Epidemiological studies show the infection rate of an epidemic disease to be proportionate to the degree of infected people, but modelling aggressive worms, we present the infection rate to be increasingly proportionate to the number of infected systems for an unprotected network; an unprotected network is a network without worm countermeasures at the given time.

III. Methodology & Modelling

The Kermack-McKendrick (SIR) model is an epidemic model that assumes each host to exists in one of three possible states: *susceptible*, *infected* or *recovered*. It assumes that susceptible hosts can be infected, and infected hosts can recover due to treatment, and that recovered host will develop immunity to the worm. It can be argued that a single approach of treating systems will not provide adequate countermeasures to combating worms. Although the SIR model provides better approximate results compared to the SI model that has no built-in countermeasures, there is still doubts if the SIR model will successfully and adequately defend against aggressive worm attacks.

The "two factor" is a variant of SIR model, again, the two factor model considered also three state worm (*susceptible*, *infected* and *recovered*) propagation dynamics. We argue that a three state worm propagation model does not suitably model worm propagations and therefore believe that the model has not considered the different and appropriate countermeasures required for combating each stage of a worm incident.

We propose a change to existing models by introducing the Improved Worm Mitigation Model (IWMM) a refined variant of the SIR model. The IWMM model presents five state worm (*susceptible*, *removed*, *infected*, *quarantine*[†], *recovered*) propagation dynamics and provides additional countermeasures. The model assumes that *susceptible* systems can be *infected*; and *susceptible* systems can also be *removed*. *Infected* systems can be *quarantine* and then *recovered*. *Removed* systems (system unplugged from the network) can be immunised[‡]. When a system is immunised it develops immunity and will no longer be infected by that particular worm.

What the IWMM demonstrates is that, susceptible systems can move to infected state or to removed state; infected systems will move to quarantine and finally recovered; while removed systems are immunised as a preventive control since not previously infected. It is pertinent to note that previous worm defence models did not take into consideration the combined effect and benefit of removal and recovery the way we did; none presented finite state machine model of worm propagation dynamics or demonstrated a five-state worm defence model. We emphasis that "removal" is not the same as "recovery" and cannot be used interchangeably.

Countermeasures considered with the IWMM include: patching, access control list (ACL) filtering at the edge, disconnecting systems, halting or stopping processes and removing software, IP NAT and IP table firewall.

Our practice is: a) to treat infected systems (systems that manifest traits of the worm) so they can longer be used as vehicle to infect other susceptible systems; and b) to remove susceptible systems from the network to stop further spreading of the infection.

The study of epidemic outbreaks can be modelled stochastically or deterministically, however, large-scale Internet worm incidents have been modelled using deterministic models in epidemiological studies [2]. To model aggressive worms we used deterministic epidemic models and compared results obtained from these models against IWMM and recorded data of Code Red incident.

IV. Improved Worm Mitigation Model

Kermack-McKendrick (SIR) model

Using the Kermack-McKendrick (SIR) model we have:

[†] Control measures to treat an infected system
[‡] Preventive controls applied to healthy systems to prevent infection

$$\frac{dS(t)}{dt} = -bS(t)I(t) \quad (1)$$

$$\frac{dI(t)}{dt} = bS(t)I(t) - gI(t) \quad (2)$$

$$\frac{dR(t)}{dt} = gI(t) \quad (3)$$

Equations (1–3) are three coupled non-linear ordinary differential equations, referred to as the Kermack-McKendrick SIR model; and used to model aggressive worms, results obtained from the simulation are compared to the observed data of Code Red. To analyse how realistic a model is, we used the same parameters for all the models: SI, SIR and IWMM (TABLE 1).

The SIR model considered only a single countermeasure - treating infected systems - which are the class of recovered systems. With the Code Red II or the SQL Slammer incident, it is evident that applying only a single countermeasure was inadequate in combating the worm in the shortest possible time. Although the Kermack-McKendrick (SIR) model improves on the classical simple (SI) model by considering a countermeasure; but the SIR model did not adequately control the spread of aggressive worms in the shortest possible time (Fig. 3).

The Improved Worm Mitigation Model (IWMM)

Using the IWMM model we have:

$$\frac{dS(t)}{dt} = -bS(t)I(t) \quad (4)$$

$$\frac{dI(t)}{dt} = bS(t)I(t) - gI(t) - mI(t) \quad (5)$$

$$\frac{dR(t)}{dt} = gI(t) \quad (6)$$

$$\frac{dU(t)}{dt} = mS(t) \quad (7)$$

Equations (4–7) are four coupled non-linear ordinary differential equations, referred to as the Improved Worm Mitigation Model. The improved worm mitigation model can be modelled in discrete event (DE) and hybrid models; with the hybrid model embedding finite state machine (FSM) on a continuous time (CT) model. We believe that to obtain a dependable worm model capable of mitigating aggressive worms, the model should present a possibility for combinations of models. Epidemiological models use continuous time, while Analytical Active Worm Propagation (AAWP) model is based on a discrete time model [5]; but IWMM uses both DE and CT models, a case that illustrates its strong formulation underpinning.

The following factors were considered in the formulation of IWMM model: ***i)*** Treating ***infected systems*** by human countermeasures, which include: patching, cleaning, filtering access control list (ACL), netflow sampled and IP accounting; ***ii)*** Treating ***susceptible systems***, by human countermeasures, include: IP NAT, filtering, patching, disconnecting systems from the network (disconnect, halt and shutdown processes and systems); ***iii)*** Immunising ***removed systems***, these are subsets of the susceptible systems that are disconnected from the network before they are infected; human countermeasures applied include: remediation services such as patching and filtering. Deciding on a single countermeasure to employ in controlling the spread of worms can be quite challenging, however, to adequately control the spread of aggressive worms, we employed a combination of measures.

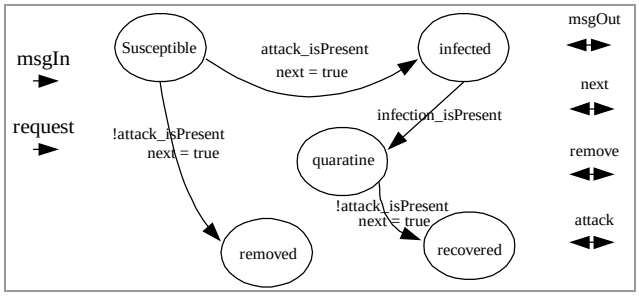

Fig. 1. IWMM 5-state FSM worm transition model

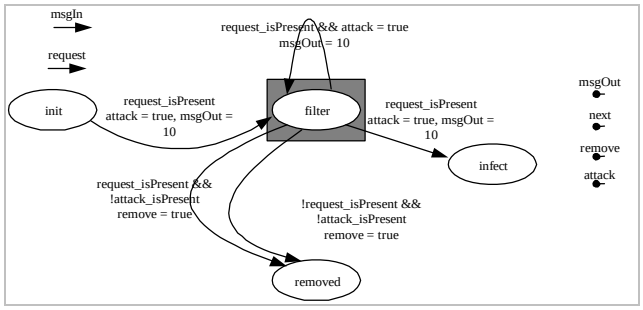

Fig. 2. IWMM model filter selection for FSM design

A. Analysis

The IWMM model produces the most accurate result to combat worm spread when compared to both the SI and SIR models as shown in Fig. 3. Also, the IWMM model yields the closest correlation to the recorded data of Code Red incident (Fig. 3 Fig. 4); and considered worm state transition as follows: a) *susceptible* "to" *infected*; b) *susceptible* "to" *removed* and c) *infected* "to" *quarantine* and d) *quarantine* "to" *recovered* (Fig. 1Fig. 2). However, the mechanism is, infected systems are quarantine (treatment after infection), while removed systems are immunised (preventive controls before infection) where they are isolated from the network while applying the necessary patches. Contrarily, neither the SIR nor the "two factor" models considered elaborate worm state transitions but presented state transition as "*susceptible* to *infectious* to *recovered*".

The class of *removed systems* accounts for a huge reduction in the total number of infected systems and towards shortening worm propagation life cycle.

B. Simulation experiments

Simulation was carried out in four scenarios, but in two separate domains, namely, continuous time (CT) and hybrid domains; with the hybrid comprising of CT and FSM in Ptolemy II modelling software. Ptolemy II [9] is the current software infrastructure of Ptolemy Project; it is open architecture and open source which encourages researchers to build their own model, leveraging and extending the core software infrastructure of Ptolemy II.

The first of the four scenarios is the classical simple (SI) epidemic model although its full result is not reported because of its trivial nature, the second scenario is the Kermack-McKendrick (SIR) model of equations (1-3), with the third scenario been the IWMM model of

equations (4-7) and finally the IWMM model is used to model Code Red. In the third and fourth scenarios, we considered the effect of a combination of countermeasures and also considered the rate of infection, which we assumed to be increasingly proportionate to the number of infected systems. Each simulation was conducted using a CT domain except the IWMM model that was conducted in two separate experiments; one on a CT domain and the other on a hybrid model (FSM of the hybrid experiments is shown in Fig. 1Fig. 2).

The IWMM result accounts for the countermeasures considered in the model; the class of remove system - $\frac{dU(t)}{dt} = mS(t)$ as shown in equation (7), and the number of systems quarantined from the class of infected systems as shown with a square bracket in this equation $\frac{dI(t)}{dt} = bS(t)I(t) - g(t) - [mI(t)]$.

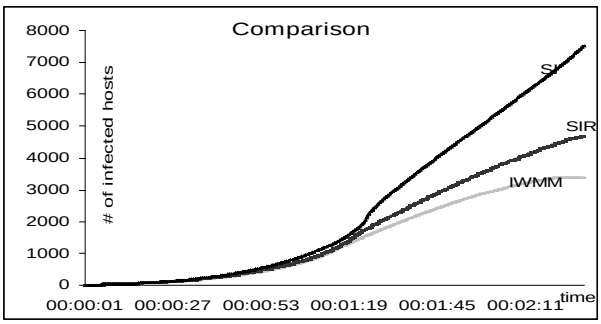

Fig. 3. Comparing SI, SIR and IWMM models

TABLE 1: PARAMETRIC DEFINITION OF VALUES[§]

Infection rate β = 1.8 in 1800 seconds; sample population N = 10000; and initial infection (I_0) = 10; recovery rate γ = 2; initial susceptible (S_0) = 10000; initial recovery (R_0) = 1; removal rate μ = 1; initial removal (U_0) = 0. Where t is the time; S(t) is the number of susceptible systems in time t; I(t) is the number of systems infected in time t; R(t) is the number of systems that have recovered and develop immunity to the infection; U(t) is the number of removed systems (disconnected); β is the infection rate; γ is the recovery rate and μ is the removal rate

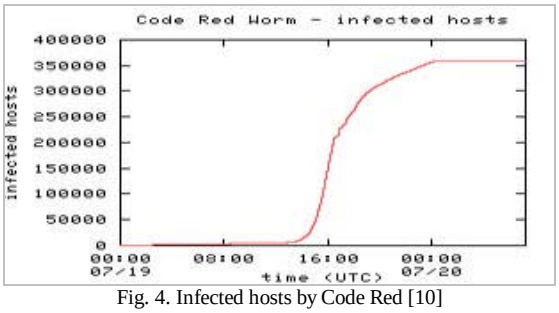

Fig. 4. Infected hosts by Code Red [10]

At the beginning of a worm outbreak (day zero outbreak), every system in the given population sample (Internet) is susceptible to the outbreak if they posses exploitable vulnerability. The initial systems to be infected are vulnerable systems that are in the attacker's "hitlist" (target systems); this accounts for the initial infected system (I_0 = 10) (TABLE 1). Also at the beginning of a worm outbreak when adequate patches have not been released and distributed, it is highly unlikely that any appropriate countermeasures have been deployed, which is why the initial removed systems (U_0 = 0) is zero.

There are two observations made when comparing results obtained with the IWMM and that of the recorded data of Code Red as follows: Firstly, with Code Red, at the beginning of the incident, the infection was gradual because target hosts were those on the hitlist and infection rates are proportional to the number of infected hosts. But as soon as the infection progressed, the infection rate increased exponentially, two things account for this: a) at the early stages of a worm incident, adequate patches or countermeasures have not been widely known and distributed; b) the worm was spreading very fast exploiting target systems compromised via infected hosts in the hitlist. Secondly, with the IWMM model, the worm spread was controlled by the modelled countermeasures of equations (6,7). From our recorded data, over 100,000 systems were infected in 8 hours, while with the IWMM considering additional countermeasure, less than 50,000 where infected within the same time duration.

V. SUMMARY

This paper discusses the characteristics of aggressive worms and proposes an epidemiological worm quarantine model (IWMM) to model aggressive worms. The proposed worm model demonstrates 5-state finite state machine model that captures detailed transitions of worm propagation mechanics; also produced closely related correlation to recorded data of Code Red incident. Comparing other defence models discussed in the paper with the IWMM model; it shows that the IWMM model presented more elaborate worm state transitions than the other models and also provided additional countermeasures as well. We conclude that IWMM is a dependable model for the analysis of aggressive worm spread.

REFERENCES

[1] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford: Slammer Worm Dissection. Inside the Slammer Worm; Cooperative Association for Internet Data Analysis and University of California, San Diego; *Published in IEEE Security & Privacy, Vol. 1 No. 4*, Jul. - Aug 2003.

[2] J.C. Frauenthal, Mathematical Modeling in Epidemiology. New York: Springer-Verlag, 1980, ch.2.

[3] P. Porras, L. Briesemeister, K. Skinner, K. Levitt, J. Rowe, Y. A. Ting: A Hybrid Quarantine Defense, *in Proc. of ACM working on rapid malcode WORM'04, pp.73-82, October 29, 2004, USA,*

[4] E. Spafford: "The Internet worm program: an analysis," *ACM Computer Communications Review*, Vol. 19, pp. 17–57, Jan. 1989

[5] Z. Chen, L. Gao, K. Kwiat: "Modeling the Spread of Active Worms," presented *in the IEEE INFOCOM*, 2003.

[6] C. C. Zou, W. Gong, D. Towsley: Code Red Worm Propagation Modeling and Analysis; *in Proc. of 9th ACM Conference on Computer and Communications Security, pp.138-147,* Nov.2002.

[7] A. Avizienis, J. Laprie, B. Randell, C. Landwehr, "Basic concepts and Taxonomy of Dependable and Secure Computing," *IEEE Transactions on Dependable and Secure Computing, Vol. 1, No.1,* Jan. – March 2004

[8] eEye Digital Security: Code Red II Worm Analysis. http://www.eeye.com/html/Research/Advisories/AL20010804.html

[9] Ptolemy II: http://ptolemy.eecs.berkeley.edu/ [Accessed 22/07/2005]

[10] CAIDA, the Cooperative Association for Internet Data Analysis: http://www.caida.org/analysis/security/code-red/coderedv2_analysis.xml [Accessed 14/07/2005].

[§] These values are test values only; any realistic parametric values can be used provided the basic assumption of use is plausible.