Vol.11.Issue.2.2023 (April-June)

http://www.bomsr.com

Email:editorbomsr@gmail.com

RESEARCH ARTICLE

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH

A Peer Reviewed International Research Journal

INTERNATIONAL STANDARD SERIAL NUMBER 2348-0580

A MATHEMATICAL MODELLING ON MACRO VIRUS SPREAD IN THE COMPUTER NETWORK

Manoj Kumar M[1], Padmavathi Ramamoorthi[2]

[1,2] PG & Research Department of Mathematics

Sri Ramakrishna College of Arts & Science(Autonomous) Coimbatore-641006, India

DOI:10.33329/bomsr.11.2.36

ABSTRACT

Macro virus is a nano machine program that creates software programs such as MS Excel or word. When it enters into a system it creates a thread file and spreads into other program or software files by triggering the file. There is a chance to identify the Macro virus spread in the computer system if it has an end-user anti-malware. Microscopic- Markov chain approach is used to model the Macro virus spread in a Computer Network. The basic reproduction number R_0 is found using the Next Generation Matrix (NGM). The model is verified via Numerical Simulation.

keywords: Markov chain approach, Basic reproduction number, Macro virus spread, Anti- malware.

1 INTRODUCTION

A micro organism is a nano machine program that is expanded by inserting copies of itself into other practicable code or documents. A macro virus is a computer virus written in the same macro language used to create software programs such as Microsoft Excel or Word. It centers on software applications and does not depend on the operating system (OS). As a result, it can infect any computer running any kind of OS, including Windows, macOS and Linux. Macro viruses work by adding their code to the macros associated with documents, spreadsheets and other data files. They target software rather than systems and can infect any OS. Macro viruses have been around since 1995 during the Concept virus first appeared. It was accidentally included on a CD-ROM (compact disc read-only memory) called Microsoft Compatibility Test and shipped by Microsoft to hundreds of corporations.

With the release of Microsoft Office 2000 and all subsequent versions, Microsoft disabled macros by default. After that, it's become more difficult for bad actors to easily launch macro viruses. However, as long as macros are available to users, the risk of macro viruses remains serious. In fact, if a macro virus infects a file, it can potentially damage not only the document, but the system and other applications. Hence, security teams should not ignore the risk.

How does a macro virus work and spread?

The possible ways for macro viruses may spread are

- phishing emails containing malicious attachments. The macro virus spreads quickly as users share infected documents, often by forwarding the infected email
- its code may entered to users computers after they click on malicious links in banner ads or URLs (Uniform Resource Locators)
- when files are shared over a network
- when infected files are placed on a removable disk and shared among multiple users
- when an infected file is downloaded via a Modem or the internet or intranet and opened

Any program that uses macros can act as the virus host, and any copy of an infected program, regardless of where it resides email, hard disk, USB (Universal Serial Bus) drive, etc. or which OS supports it, can contain and spread the virus. Since, the virus is dormant until the infected macro is run, it's difficult to detect. Hence, it is like a Trojan horse, a malicious program. But, unlike a Trojan, a macro virus can replicate automatically and infect other computers quickly.

What can macro viruses do?

When a macro virus infects an application, it executes a series of commands and a sequence of actions that begin automatically when the application is opened. An infected macro that is executed typically infects every document on a user's computer.

The virus can also do the following:

- cause irregularities in text documents, such as inserting or deleting words or pictures
- create new files
- corrupt or erase stored data
- format hard drives
- access email accounts
- send out copies of infected files to everyone on the user's contact list

How have macro viruses evolved?

Initially, macro viruses mostly infect Word or Excel documents two applications with powerful macro languages and features. And they almost exclusively target the Windows OS.Macro viruses evolved into a dominant type of virus, affecting all kinds of applications, files and OS that use macros. Since a macro virus is cross-platform, it can infect both Windows and Mac computers using the same code.

For example, in 2017, MacDownloader, the first Word macro virus for Apple's macOS, was discovered. MacDownloader enabled hackers to use malicious macros in Word documents to install malware on Mac computers to steal user's data, such as browser history logs, webcam files, passwords and encryption keys.

What are signs of a macro virus infection?

A macro virus infection may be difficult to detect. However, it does leave some footprints that could indicate an infection. For example, the system speed may be affected, or the com- puter may display strange error messages. The Nuclear macro virus display the message: "And finally I would like to say: STOP ALL FRENCH NUCLEAR TESTING IN THE PACIFIC." Other possible red flags of a macro virus infection include the following:

- if the computer asks the user for a password on a file when the password wasn't enabled.
- if a document was saved as template file.

2 MOTIVATION AND METHOD

- Macro virus usually infect Microsoft Office applications like word and excel.
- One of the most common way for macro viruses spread is through phishing emails.
- Once a macro virus has infected a document, it typically infects all other similar documents on the computer.
- Macro viruses tend to spread easily and silently.
- Macro viruses also spread through shared networks or discs.

3 SYMPTOMS - SYNDROME

- A device that's running slowly.
- Strange error messages appearing.
- The computer asks the user for a password on a file when the password wasn't enabled.
- A document was saved as template file.

4 PREVENTION

At present, Microsoft Word and Excel disable macros automatically, triggering infected file,enable the macros. That the macro viruses are avoided by not enabling macros. Before accessing an unknown document or spreadsheet collect the information about the document to verify whether it is phising document or not.

5 MODEL DESCRIPTION

Based on the spreading dynamics of macro virus a computer has six categories(CAMSFI).

- C - Computer System - A Computer free from Macro Virus.
- A - End user Anti Malware - A Computer with Anti Malware and free from Macro virus.
- M - Microbe overwhelmed - A Computer with Macro virus and it is unidentified.
- S - syndrome - A Computer with Macro virus and its shows the symptoms of Macro virus.
- F - Flip through - A Computer with Macro virus and it has identified.
- I - Idle - A Computer System which is idle state.

Consider a network of N computer the matrix B denotes the contacts among the computers.

Probability of i^{th} computer system is affected by macro virus is

$$r_i(t) = \pi_{j=1}^{N}[1 - \beta(M_j + S_j)b_{ji}]$$

Probability of i^{th} end user Anti Malware system is affected macro virus is

$$r_i^a(t) = \pi_{j=1}^{N}[1 - \gamma\beta(M_j + S_j)b_{ji}]$$

6 MICROSCOPIC MARKOV CHAIN APPROACH

The transmission flow among the states of a computer system is shown in Figure 1.

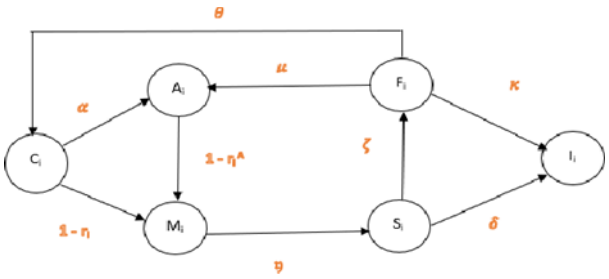

Figure 1: Transmission flow among the state

Table 1: Parameters and their description

Parameters	Description
α	Probability of using end user Anti Malware (Anti Virus awareness)
1-r_i	Probability of a computer system is affected by macro virus
1-r_i^a	Probability of a end user anti malware system is affected by macro virus
η	Rate of transmission from Microbe overwhelmed state to Syndrome state
ζ	Rate of transmission from Syndrome state to Flip through state(Scanning)
δ	Rate of transmission from Syndrome state to Idle state(Death State)
κ	Rate of transmission from Flip through state to Idle state(Death State)
μ	Rate of transmission from Flip through to end user Anti Malware (aware state)
θ	Rate of transmission from Flip through to computer system (unaware state)

$X_i(t)$ denotes probability of the i^{th} system is in the state X, $X = C, A, M, F, S, I$. The transmission flow among the possible states are shown in figure 1. The mathematical equations representing the probability of i^{th} system in various states at time $t + 1$ are given by

$$C_i(t+1) = C_i(t)(r_i - \alpha) + F_i(t)(\Theta) \quad (1)$$

$$A_i(t+1) = C_i(t)(\alpha) + A_i(t)(r_i^a) + F_i(t)(\mu)) \quad (2)$$

$$M_i(t+1) = C_i(t)(1 - \gamma_i) + A_i(t)(1 - \gamma_i^a) + M_i(t)(1 - \eta) \quad (3)$$

$$S_i(t+1) = M_i(t)(\eta) + S_i(t)(1 - \zeta - \delta) \quad (4)$$

$$F_i(t+1) = S_i(t)(\zeta) + F_i(t)(1 - \mu - \kappa - \vartheta) \quad (5)$$

$$I_i(t+1) = S_i(t)(\delta) + F_i(t)(\kappa) + [I_i(t)(1)] \quad (6)$$

$$r_i(t) = \pi_{j=1}^{N}[1 - \beta(M_j + S_j)b_{ji})] \quad (7)$$

$$r_i^a(t) = \pi_{j=1}^{N}[1 - \gamma\beta(M_j + S_j)b_{ji})] \quad (8)$$

$$C_i(t+1) + A_i(t+1) + M_i(t+1) + S_i(t+1) + I_i(t+1) + F_i(t) + R_i(t) = 1 \forall t, \forall_i = 1, 2, \ldots N \quad (9)$$

Solving the system of equations 1 to 9 iteratively, the time evolution of the system for any initial condition can be found.

6.1 Continuous time Markov chain approach

Linearising r_i and r_i^a, $r_i \approx 1 - \sum_{j=1}^{N} \beta\left(M_j + S_j\right)B_{ji}$ and $r_i^a \approx 1 - \sum_{j=1}^{N} \gamma\beta\left(M_j + S_j\right)B_{ji}$.The continuous time linearised mathematical equation representing the model is

$$\frac{dC_i}{dt} = C_i(t)\left[1 - \sum_{j=1}^{N}\left[\beta\left(M_j + S_j\right)b_{ji}\right] - a - 1\right) + F_i(t)\theta$$

$$\frac{dA_i}{dt} = C_i(t)(a) + F_i(t)(\mu) + A_i(t)\left[1 - \sum_{j=1}^{N}\left[\gamma\beta\left(M_j + S_j\right)b_{ji}\right] - 1\right)$$

$$\frac{dM_i}{dt} = C_i(t)\left[\sum_{j=1}^{N}\left[\beta\left(M_j + S_j\right)b_{ji}\right]\right] + A_i(t)[1 - \sum_{j=1}^{N}\left[\gamma\beta\left(M_j + S_j\right)b_{ji}\right]] - M_i\,(t)(\eta)$$

$$\frac{dS_i}{dt} = M_i\,(t)(\eta) - S_i(t)(\zeta - \delta)$$

$$\frac{dF_i}{dt} = S_i(t)(\zeta) - F_i(t)(1 - \mu - \kappa - \theta - 1)$$

$$\frac{dI_i}{dt} = S_i(t)(\delta) - F_i(t)(\kappa) + I_i\,(t)(1 - 1) \tag{10}$$

Definition 1

The feasible region Γ is defineΣd as Γ = {C_1, A_1, M_1, S_1, F_1, I_1, C_2, A_2, M_2, S_2, F_2, I_2, .., C_N;A_N;M_N; S_N; F_NN; I_N)$\forall \epsilon R^{6N}/\sum_{j=1}^{N}(C_i + A_i + M_i + S_i + F_i + I_i) = N\}$

The feasible region Γ is positively invariant with respect to equation 1 to 9 and the disease free equilibrium (DFE) P_0 = (C_1^0, 0, 0, 0, 0, 0, C0, 0, 0, 0, 0, 0, ..., C_N^0 , 0, 0, 0, 0, 0) with C_i^0 = 1 always exits in Γ Theorem 1 finds the basic reproduction number R_0 using Next Gen eration Matrix (NGM).

Theorem 1

The basic reproduction number R_0 of the system is given by $\Lambda_{\max}(H)$, where $h_{ij} = \left(\frac{1}{\eta} + \frac{1}{\zeta+\delta}\right) K_i b_{ij}$ & $K_i = \beta\,[C_i(t) + \gamma A_i(t)]$.

Proof

The equations corresponding to M_i Microbe overwhelmed and S_i - Syndrome states are

$$\frac{dM_i}{dt} = C_i(t)\sum_{j=1}^{N}\beta\, M_j b_{ji} + C_i(t)\sum_{j=1}^{N}\beta\, S_j b_{ji} + A_i(t)\sum_{j=1}^{N}\gamma\beta M_j b_{ji} + A_i(t)\sum_{j=1}^{N}\gamma\beta S_j M_i b_{ji}\;(t)(\eta)$$

$$\frac{dS_i}{dt} = \eta M_i\,(t) - (\zeta - \delta)\,S_i(t) \tag{11}$$

This can be written as

$$\frac{dM_i}{dt} = K_i \sum_{j=1}^{N} M_j b_{ji} + K_i \sum_{j=1}^{N} S_j b_{ji} \; - M_i \; (t)(\eta)$$

by taking $K_i = \beta[C_i(t) + \gamma A_i(t)]$&$X = (M_1, M_2, \ldots M_N, S_1, S_2, \ldots S_N)^T$

The above equations can be written as

$$\frac{dx}{dt} = \; (F \; - \; V \;)X$$

where,

$$F = \begin{bmatrix} K_1b_{11} & K_1b_{21} & \cdots & \cdots & K_1b_{N1} & K_1b_{11} & K_2b_{21} & \cdots & \cdots & K_1b_{N1} \\ K_2b_{12} & K_2b_{22} & \cdots & \cdots & K_2b_{N2} & K_2b_{12} & K_2b_{22} & \cdots & \cdots & K_2b_{N2} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ K_Nb_{1N} & K_Nb2N & \cdots & \cdots & K_Nb_{NN} & K_Nb_{1N} & K_Nb2N & \cdots & \cdots & K_Nb_{NN} \\ 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 \end{bmatrix}_{(2N \times 2N)}$$

$$-V = \begin{bmatrix} -\eta & 0 & \cdots\cdots & 0 & 0\cdots & 0\cdots\cdots & 0 \\ 0 & -\eta & \cdots\cdots & 0 & 0\cdots & 0\cdots\cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots\cdots & -\eta & 0\cdots & 0\cdots\cdots & 0 \\ & & & & & & \\ \eta & 0 & \cdots\cdots & 0 & -(\zeta+\delta)\cdots & 0\cdots\cdots & 0 \\ 0 & \eta & \cdots\cdots & 0 & 0\cdots & -(\zeta+\delta)\cdots\cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots\cdots & \eta & 0\cdots & 0\cdots\cdots & -(\zeta+\delta) \end{bmatrix}_{(2N \times 2N)}$$

$$\mathrm{V} = \begin{pmatrix} D(\eta)_{N\times N} & D(0)_{N\times N} \\ D(-\eta)_{N\times N} & D(\zeta+\delta)_{N\times N} \end{pmatrix}$$

$$R_0 = \lambda_{max}(FV^{-1})$$

$$V^{-1} = \begin{bmatrix} D(\frac{1}{\eta}) & 0 \\ D(\frac{1}{\zeta+\delta}) & D(\frac{1}{\zeta+\delta}) \end{bmatrix}$$

$$R_0 = \lambda_{max}(FV^{-1}) = \lambda_{max}(H)$$

Hence,

$$FV^{-1} = \begin{bmatrix} \frac{K_1 b_{11}}{\eta} + \frac{K_1 b_{11}}{\zeta+\delta} & \frac{K_1 b_{21}}{\eta} + \frac{K_1 b_{21}}{\zeta+\delta} & \cdots & \frac{K_1 b_{N1}}{\eta} + \frac{K_1 b_{N1}}{\zeta+\delta} & \frac{K_1 b_{11}}{\zeta+\delta} & \frac{K_1 b_{21}}{\zeta+\delta} & \cdots & \frac{K_1 b_{N1}}{\zeta+\delta} \\ \frac{K_2 b_{12}}{\eta} + \frac{K_2 b_{12}}{\zeta+\delta} & \frac{K_2 b_{22}}{\eta} + \frac{K_2 b_{22}}{\zeta+\delta} & \cdots & \frac{K_2 b_{N2}}{\eta} + \frac{K_2 b_{N2}}{\zeta+\delta} & \frac{K_2 b_{12}}{\zeta+\delta} & \frac{K_2 b_{21}}{\zeta+\delta} & \cdots & \frac{K_2 b_{N2}}{\zeta+\delta} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{K_N b_{1N}}{\eta} + \frac{K_N b_{1N}}{\zeta+\delta} & \frac{K_N b_{2N}}{\eta} + \frac{K_N b_{2N}}{\zeta+\delta} & \cdots & \frac{K_N b_{NN}}{\eta} + \frac{K_N b_{NN}}{\zeta+\delta} & \frac{K_N b_{1N}}{\zeta+\delta} & \frac{K_N b_{2N}}{\zeta+\delta} & \cdots & \frac{K_N b_{NN}}{\zeta+\delta} \\ 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 \end{bmatrix}_{(N\times N)}$$

where,

$$H = \begin{bmatrix} \frac{K_1 b_{11}}{\eta} + \frac{K_1 b_{11}}{\zeta+\delta} & \frac{K_1 b_{21}}{\eta} + \frac{K_1 b_{21}}{\zeta+\delta} & \cdots\cdots & \frac{K_1 b_{N1}}{\eta} + \frac{K_1 b_{N1}}{\zeta+\delta} \\ \frac{K_2 b_{12}}{\eta} + \frac{K_2 b_{12}}{\zeta+\delta} & \frac{K_2 b_{22}}{\eta} + \frac{K_2 b_{22}}{\zeta+\delta} & \cdots\cdots & \frac{K_2 b_{N2}}{\eta} + \frac{K_2 b_{N2}}{\zeta+\delta} \\ \vdots & \vdots & \cdots\cdots & \vdots \\ \frac{K_N b_{1N}}{\eta} + \frac{K_N b_{1N}}{\zeta+\delta} & \frac{K_N b_{2N}}{\eta} + \frac{K_N b_{2N}}{\zeta+\delta} & \cdots\cdots & \frac{K_N b_{NN}}{\eta} + \frac{K_N b_{NN}}{\zeta+\delta} \end{bmatrix}_{(N\times N)}$$

The epidemic threshold is given by $R_0 < 1 \Rightarrow \lambda_{max}(H) < 1; h_{ij} = \left(\frac{(K_i)(b_{ji})}{\eta} + \frac{(K_i)(b_{ji})}{\zeta+\delta}\right) K_i\, b_{ji} \& K_i = \beta\,[C_i(t) + \gamma A_i(t)]$

7 Numerical Simulation

A scale free network with 1000 nodes are used to represent the contacts among the Computers. The parameter values are taken as α= 0:4; γ = 0:3; β = 0:4; η = 0:4; = 0:5; ζ = 0:02; κ = 0:005; t_m = 50; θ = 0:3; μ = 0:4. Using MATLAB, the virus spread among the networks based on equation 1 to 9 are calculated. The propagation of virus spread in the network is shown in Figure 2 , Figure 3 shows the propagation of Microbe Overwhelmed (M) and Syndrome (S).

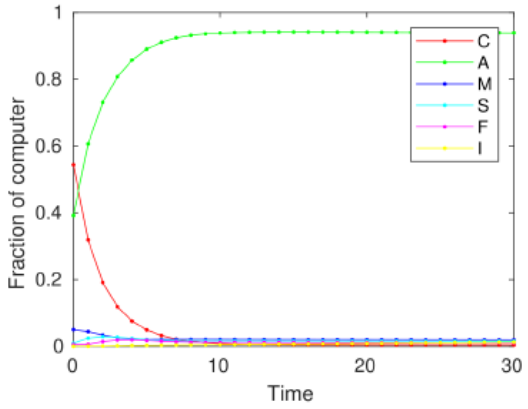

Figure 2: The propagation of virus spread in the network

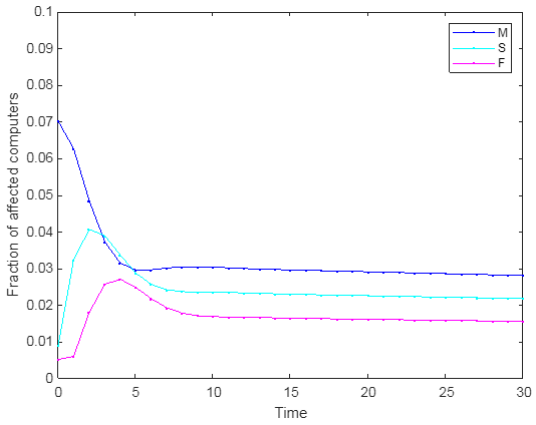

Figure 3: The propagation of M - Microbe overwhelmed and S – Syndrome

8 Conclusion

Most of the spreading virus have both symptomatic and asymptomatic cases in their spreading dynamics. In this paper the spreading dynamic of Macro virus over the complex network is framed. The mathematical model corresponding to the model is framed using Microscopic Markov Chain approach. The basic reproducing number R_0 is found using the Next Generation Matrix (NGM). The model is verified through Numerical Simulation.

References

[1] Y.Shang, Int. J. Biomath. 6(2),135007(2013).

[2] C.Granell, S. Gomez and A. Arenas, On the dynamical interplay between awareness and epidemic spreading in multiplex networks, arXiv:1306.4136v2[physics.soch-p].

[3] C.Granell, S. Gomez and A. Arenas, Phys. Rev. E90(1).012808(2014).

[4] C. Zheng, C. Xia, Q. Guo and M. Dehmer, J. Parallel Distrib. Comput. 115, 2028 (2018).

[5] C. Xia, Z. Wang, C. Zheng, Q. Guo, Y. Shi, M. Dehmer and Z. Chen, Inform. Sei. 471,185200 (2019).

[6] Bozo, Word Macro Viruses, An Electronic Document available on: http://www.xs4all.nl/ cicatrix/index.html.

[7] Chengi Jimmy Kuo, Free Macro Anti Virus Technique, An Electronic Document available on: http://www.xs4all.nl/ cicatrix/index.html.

[8] Joe Wells, Concepts: Understanding The Virus and its Impacts, An Electronic Document available on: http://www.xs4all.nl/ cicatrix/index.html

[9] McAfee, Macro Viruses, An Electronic Document available on: http://www.xs4all.nl/ cicatrix/index.html.

[10] Method for emulating an executable code in order to detect maliciousness, document available on http://www.freepatentsonline.com/20040133796.html.