Mathematical Modeling and Simulation for Cyber Defense

Dinesh Kumar Saini

Article Info
Received:22-09-2025 Revised:02-11-2025 Accepted:09-11-2025 Published:19-11-2025

***Abstract*—** This essay provides an overview of a practical issue facing modern cyberspace: cyberattacks and defense. The objects used for cyberattacks are described at the beginning of the study, along with the ways in which they might be disseminated, including secondary memory, email attachments, instant messages, and malicious bots. The use of mathematical modeling and simulation to address the issue is described in this work along with a thorough assessment of the literature. In order to identify research needs, cyber defense architecture and potential components of a cyber defense model are also examined. This report concludes by identifying certain gaps and potential solutions.

***Index Terms*—Virus, worms, differential equation, instance messaging, FTP, E-mail, cyber-hardening**

INTRODUCTION

The way data is moved and information is exchanged has drastically changed in the last three decades due to the development of Internet/Network technology. Internet technology has expanded over time in tandem with technological advancements and demands, providing a wide range of features. The development of Internet technology has created significant obstacles, such as the need for an appropriate cyber security system to protect the sensitive data held on the system. In order to achieve this, it is suggested to research and comprehend the different malicious objects and create a mathematical model to depict their actions [1], [2]. First, we studied how harmful objects like viruses, worms, Trojan horses, bots, and others replicate and spread themselves [1], [2].

I. MATHEMATICAL MODELING

Modeling can be summed up as "the study of a system before actually it is built and implemented." Sometimes the enormous expense and time required to implement a system or real-world issue in the real environment make it impractical. Therefore, it is preferable to construct a model or prototype and analyze the system's behavior. In addition to being a replacement for the real system, a model also simplifies it [3], [4].

A. A Few Illustrations of Various Models

Physical Model: These are based on analogies between electrical, mechanical, or hydraulic systems.
Mathematical Model: Systems such as the supply and demand system that can be expressed as mathematical equations
Physical Static Model: Unlike the water-tank model, these physical models do not alter their behavior over time.
Physical Dynamic Model: These physical models, such as spring suspension systems or comparable electric systems, alter their behavior over time.
Mathematical Static Model: When a system is in an equilibrium state, such as a demand-supply system, these mathematical models provide a mathematical equation.
Mathematical Dynamic Model: These mathematical models permit system characteristics to change as a function of time, such as oscillatory motion.
Small static mathematical models that can be resolved using conventional mathematics are known as mathematical static analytical models.
Mathematical Static Numerical Model: Simulation can be used to solve these intricate static mathematical models.

Mathematical Dynamic Analytical Model: Traditional mathematics can be used to solve these small dynamic mathematical models.

Mathematical Dynamic Numerical Model: Simulation can be used to solve these intricate dynamic mathematical models.

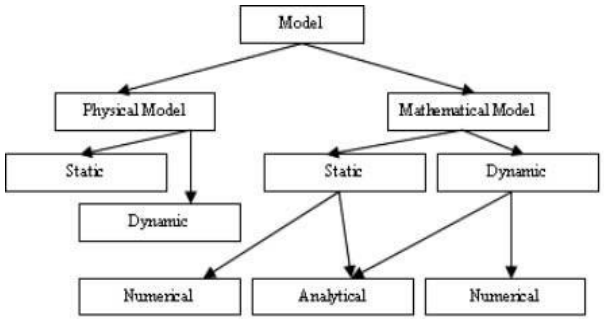


Fig. 1. Mathematical models

II. Principle of Modeling with Respect to Cyber Defense System

The biggest issue facing the modern world is cyberattacks. Understanding how harmful things behave is essential to solving this issue. Mathematical modeling can be crucial in this regard. It can assist in fixing the harmful object's potential parameters, which are crucial for determining how the malicious object can spread over the Internet. However, this mathematical modeling aids in the construction of the cyber protection system in the following cases [5,6]: Why? Cyber defense is caused by a variety of dangerous items, such as viruses, worms, Trojan horses, spam, etc., as well as specific technologies, such as P2P, instant messaging, bots, phishing, etc., that are abused for fraudulent purposes. We need to understand how these dangerous objects behave in order to defend against them, and modeling is the only way to do so.

Locate? We must be aware of any harmful objects or data that could cause our system to act strangely based on the network traffic that enters our system. The data's length, type, source, and methodology etc.).

Secondary memory (CD-ROM, hard drive, and floppy)

Email with attachments. Instant messaging, including chat, text messaging, and FTP. Programs for malicious bots [4]. can impact our system, among other things to be aware of. Given? the things that we are aware of and can be examined to forecast their future developments. results of the pertinent data, including time stamps, TCP dumps from certain ports, data accumulation times, etc. Presume? We must make certain important assumptions, such as the duration of the data, the time it takes to aggregate the data, the time it takes to aggregate the connections, the number of connections the data will be collected, etc. It is also necessary to determine the conditions under which assumptions can be used.

How? The physical principles that underpin the model include fuzzy logics, intelligent approaches, data/text mining, uncertainty principles, etc. Additionally, how should we search for the model—isolated, laired, spread, etc.?

Make a prediction? The malicious objects should be predicted by the model. Mathematical equations, the necessary computations, and the information's current state can all be used to describe it.

Is it valid? To assess the model's validity and consistency with regard to its underlying assumptions and principles, a number of tests must be conducted.

Verified? Tests must be found and run to determine whether the model is verified, which indicates that the predictions are accurate. In terms of the original motivations behind it, it is beneficial.

Get better? Simulated results assist in identifying factors that should be included, assumptions that may be lifted to improve the model, and parameter values that are not sufficiently known. Put into practice the "model-validate-verify-improve-predicate" iterative loop.

Utilize? What additional exercises can we find and incorporate to make the security model more flexible?

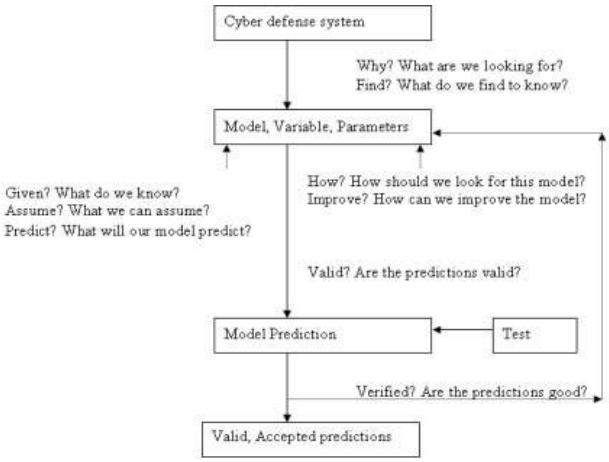


Fig. 2. Mathematical modeling process

III. Malicious Objects and Defense

The current Internet system, which had more than 233 million machines as of January 2004 [3], is vulnerable to threats from a variety of malicious items. These objects can be worms, viruses, Trojan horses, spam, etc., and they can propagate over the Internet through

The computer attacks are completely random. The precise moment of the next computer attack is unknown to us. However, we can determine the likelihood of an attack at a given moment in time using simulation's probability ideas.
The set of numbers P(xi) is referred to as a probability mass function if the stochastic variable (Time of attack) can take I distinct values, xi (i = 1, 2,..., I), and the probability of the value xi being selected is P(xi). Given that one of the values must be entered into the variable, it follows that

$$\sum_{i=1}^{I} P(X_i) = 1$$

Probability mass function can be defined as

$$P(x_i) = n_i / N$$

where N= total number of attacks and n_i number of attacks from a specified source.

A cumulative distributed function can also be found which gives the probability of stochastic attacks' being less than or equal to a given value. Different measures of probability functions can be used for the study of the stochastic system such as mean, mode, median, Standard deviation, etc. Models characteristic equations can be of two types – Linear and non-Linear. Non-linear system can be represented by Partial Differential Equations (PDE). Consider that malicious object has propagation property P, depends upon various other factors like- A, B, C …etc. It can be represented as

$$P=f(A, B, C...).$$

The velocity can be represented as

$$\partial P/\partial t = \partial f(A, B, C,...)/\partial t.$$

and the acceleration rate can be represented as

$$\partial^2 P/\partial t^2 = \partial^2 f(A,B,C,...)/\partial t^2.$$

Once the simulated results obtained by the use of certain approximation techniques mentioned below can be used for complementing the data generated by simulation as well as validation:-

Taylor series expansion: Any function that has derivatives can be expanded by Taylor's Formula, The value of the independent variable, x, in a region near x = a, a function f(x) can be approximated by the polynomial

$$F(x) = f(a) + f'(a)(x-a) + (f''(a)/2!)*(x-a)^2 + ... + (f^{(n)}(a)/a!)*(x-a)^n.$$

Finite difference approximation methods: This method transforms a partial deferential equation over small intervals. This is of two types-

Forward difference Approximation: It calculates the function gradient at various points by the formula:

$$f'(x_i) = (f(x_{i+1}) - f(x_i))/\Delta x$$

Backward difference approximation: It also calculates the function gradient at various points by the formula:

$$f'(x_i) = (f(x_i) - f(x_{i-1}))/\Delta x$$

Higher order derivatives: These can be calculated to describe the various important points in the distribution by the following formula-

$$f^{(n)} = (f^{(n-1)})'$$

Some regression tests such as Polynomial regression tests can also be used to validate the model. It finds that the values

can be fitted into a polynomial or not.

Once the characteristic equation is derived then results can be empirically/analytically validated on the basis of available standard mathematical hypothesis. The first thing for mathematical model validation is the dimensional homogeneity, which requires that each term has the same net dimensions [7]. Secondly, the models can be validated by checking qualitative and limit behavior. Except these some other things can also be considered, depending on how large the errors are? What is the accuracy and precision? Are the data fitted into the uniform curve? The data can be prepared by mean, mode, median or standard deviation. These data can be compared easily and help us to understand the behavior of malicious objects [8].

For example some definitions of virus can be modeled as follows:

Def-1: This definition defines the simple virus which infects the existing files and makes them able to behave in the similar way.

Let *P* is a set of programs, $v \in P$, and $p_i \in P$. If v is a virus then the following equation must be true. $f(v) = f(p)$, where function $f(v)$ and $f(p)$ gives the function $T(v, p_i, e, t, S) = 0$ then program v is a virus otherwise not.

Now, for modeling a cyber defense system one should know about the different components needed. Here is a good division of cyber defense components by O. Sami Saydjari in Fig. 3.

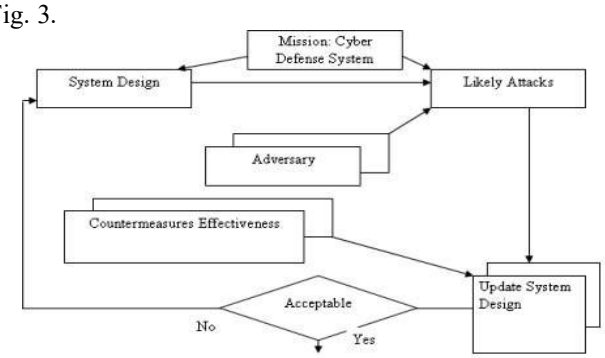


Fig. 3. Cyber defense system design model

These are sensors and exploitation, situation awareness, defensive mechanism, command and control, strategies and tactics, and science and Engineering [6]. He also tried to set an analogy in between an “art of war” to technology and he

i

behavior of programs v and pigot a success in doing such thing at some extent

Def-2: This is the definition of virus in respect of a fixed time t.

IV. Discussion and Summary

Although various models and solutions are provided, the

$$T(v, p_i, e, t, S) = \log \frac{f(v, e, t, S)}{f(p_i, e, t, S)} \quad (1)$$

following problems remain to be solved-

- All the characteristics of the malicious object’s propagation have to be represented in the form of mathematical equations so that it gives the right direction

where function

$f(v, e, t, S)$

and

$f(p_i, e, t, S)$ gives the to predict the future behavior. Another thing is to implement that characteristic equation in the existing

behavior of program v and p_i respectively at time t in the

system S at the occurrence of event e.

Now if function $T(v, p_i, e, t, S) = 0$ then program *v* is a virus otherwise not.

Def-3: This is the definition of virus in respect of a continuous time interval Δt .

$$T(v, p_i, e, t, S) = \log \frac{\int_{t=\tau_0}^{t=\tau_1} f(v, e, t, S)dt}{\int_{t=\tau_0}^{t=\tau_1} f(p_i, e, t, S)dt} \quad t=\tau_0 \quad (2)$$

environment (Existing Intranet and Internet), so to make a realistic model such that it can be implemented in real time environment without much change.

- Once a model is provided then its accuracy can be analyzed and improved on the basis of simulated results. This accuracy helps to protect the system by attack of malicious object. Like Code Red worms have been modeled with high accuracy and now it can be controlled

  easily [10], [11].Most of the models are with certain accuracy limits but still there is scope to improve them. So we will target our goals to improve these models.
- Most of the malicious objects having lot of characteristics, which is to be determined and because of

  this there is increase in the behavioral complexity [13]. We need to generalize each characteristic and narrow down the characteristic domain.
- Most of the cyber defense systems, when implemented create overheads like- slow down the existing system performance, increase the packet length or take more

where function $\int_{t=\tau 0}$ $f(v, e, t, S)$ and

time for comparison etc. So it is needed to provide such a cyber defense system, which does not create much $\int_{t=\tau 0}^{t=\tau 1}$ $f(p_i, e, t, S)$ gives the behavior of program v and p_i more extra overheads. Both the defensive systems (after attack recovery, and before attack) should be predictable

respectively in the time interval $\Delta t = \tau_1 - \tau_0$ in the system S at the occurrence of event e. Now if enough on the basis of mathematical modeling with high accuracy. .

- In cyber defense it is difficult to trace the attacker due to lower education and awareness [15], [16]. Only few countries have some inadequate policies, and education and awareness system. So, to restrict such type of malicious activities, our society needs to devise a proper set of policies.

REFERENCE

[1] D. K. Saini, "A Mathematical Model for the Effect of Malicious Object on Computer Network Immune System," Applied Mathematical Modeling, vol. 35, pp. 3777-3787 USA, doi:10.1016/.2011.02.025, 2011.
[2] B. K. Mishra and D. K. Saini, "Mathematical Models on Computer Viruses," Elsevier International Journal of Applied Mathematics and Computation, vol. 187, no. 2, pp. 929-936. USA, 2007.
[3] Elsevier International Journal of Applied Mathematics and Computation, vol. 188, no. 2, pp. 1476-1482, B. K. Mishra and D. K. Saini, "SEIRS epidemic model of transmission of malicious objects in computer network." USA, 2007.
[4] Design patterns and their impact on software quality, D. K. Saini and B. K. Mishra, ACCST Research Journal, vol. 5, no. 1, pp. 356-365 India, 2007.
[5] D. K. and H. Saini, "Proactive Cyber Defense and Reconfigurable Framework for Cyber Security," International Review on Computer and Software (IRCOS), vol. 2. no. 2, pp. 89-98, 2007.
[6] D. K. Saini and N. Gupta, "Fault Detection Effectiveness in GUI Components of Java Environment through Smoke Test," Journal of Information Technology, ISSN 0973-2896, vol. 3, no. 3, pp. 7-17 2007.
[7] "VAIN: A Stochastic Model for Dynamics of Malicious Objects," D. K. and H. Saini, ICFAI Journal of Systems Management, vol. 6, no. 1, pp. 14–28, 2008.

[8] H. Saini and D. K. Saini, "Malicious Object dynamics in the presence of Anti Malicious Software," European Journal of Scientific Research, ISSN 1450-216X, Vol.18 No.3 (2007), pp.491-499 © Euro Journals Publishing, Inc. [Online]. Available: http://www.eurojournals.com/ejsr.htm [9] D. K. Saini and N. Gupta, International Journal of Information Technology and Web Engineering, (IJITWE), vol. 3, no. 2, pp. 19-26 2008.
[10] H. Saini, D. K. Saini, and N. Gupta, "Cyber Defense Architecture in Campus Wide Network System," International Journal of Theoretical and Applied Information Technology (JATIT), April 2008 (E-ISSN 1817-3195, ISSN 1992-8645).
[11] ACM SIGSOFT, vol. 34, no. 2, 2009; D. K. Saini, "Testing Polymorphism in Object Oriented Systems for Improving Software Quality."
[12] "Integrating EduLearn Learning Content Management System (LCMS) with Cooperating Learning Object Repositories (LORs) in a Peer to Peer (P2P) architectural Framework," ACM SIGSOFT, vol. 34, no. 3, 2009, L. S. Prakash, D. K. Saini, and N. S. Kutti.
[13]"Enhanced Inquiry Method for Malicious Object Identification," D. K. Saini, J. H. Yousif, and W. M. Omar, ACM SIGSOFT, vol. 34, no. 3, 2009.
[14] D. K. Saini, L. A. Hadimani, and N. Gupta, "Software Testing Approach for Detection and Correction of Design Defects in Object Oriented Software," Journal of Computing, vol. 3, no. 4, April 2011.
[15] D. K. Saini, "Security Issues with Object-Oriented Software Architectures," International Journal of Computer Applications, vol. 40, no. 11, pp. 41-48, 2012.
[16] D. K. Saini and Y. Sharma, "Soft Computing Particle Swarm Optimization based Approach for Class Responsibility Assignment Problem," International Journal of Computer Applications, vol. 40, no. 12, pp. 19–24, 2012.