Mathematical Modeling in Cyber Defense

Muharrem Tuncay GENÇOĞLU*‡

*Fırat University Technical Sciences Vocational School, Elazığ, Türkiye

(mtgencoglu23@gmail.com)

‡Muharrem Tuncay GENÇOĞLU; Fırat University Technical Sciences Vocational School, Elazığ, Türkiye,

mtgencoglu23@gmail.com

Received: 10.09.2020 Accepted: 27.12.2020

Abstract- With the development of technological developments in addition to internet technology, the necessity of cyber defence systems has emerged for the protection of valuable or valuable information stored in the cyber environment. It is important to understand the behaviour of malicious objects with the increasing threat of cyber-attacks in recent years. Mathematical modeling is required for this. In this study, a mathematical modeling process and a cyber defence system modeling principle are given. Since the attacks on the computer are completely stochastic, a Cyber Defence System Design Model based on the detection of the behaviour of malicious objects with the help of the probability distribution function and differential equations has been proposed.

Keywords Cyber Security, Mathematical Modelling Process, Cyber Security System Design.

1. Introduction

Advances in network technology in recent years have caused serious changes in data transfer and information exchange. Along with technological developments, internet technology has become more functional and developed. With these developments in internet technology, the necessity of cyber defence systems to protect valuable or insignificant information stored in the cyber environment has emerged seriously. To create cyber defence systems, some mathematical models have been suggested to search and figure out diverse malicious objects and to detect and represent their behaviour [2]. These models concentrate on the self-replication of some malicious objects.

2. Mathematical Modelling

Sometimes it may not be possible to implement a system in real environment due to high cost and excessive time requirement. In this case, it would be more appropriate to create a model and examine the manner of the system. Modeling is to replace and simplify a real system. If we take a look at some different models from here;

- Physical Model: Based on Mechanical, Electrical or Electric and Hydraulic systems.
- Mathematical Model: Systems body forth by mathematical equations.
- Physical Static Model: Systems whose behavior does not change over time.
- Physical Dynamic Model: Systems whose behavior changes over time.
- Mathematical Static Model: These are models that give a mathematical equation when the system is in balance.
- Mathematical Dynamic Model: Models that allow the change of system eigenvalues depending on a time function.
- Mathematical Static Numerical Model: Complex static mathematical models that can be resolved by simulation.
- Mathematical Static Analytical Model: These are small static mathematical models that can be resolved with basic mathematical methods.
- Mathematical Dynamic Analytical Model: These are minor dynamic mathematical models that can be resolved with basic mathematical methods.
- Mathematical Dynamic Numerical Model: Complex dynamic mathematical models that can be resolved by imitation[4].

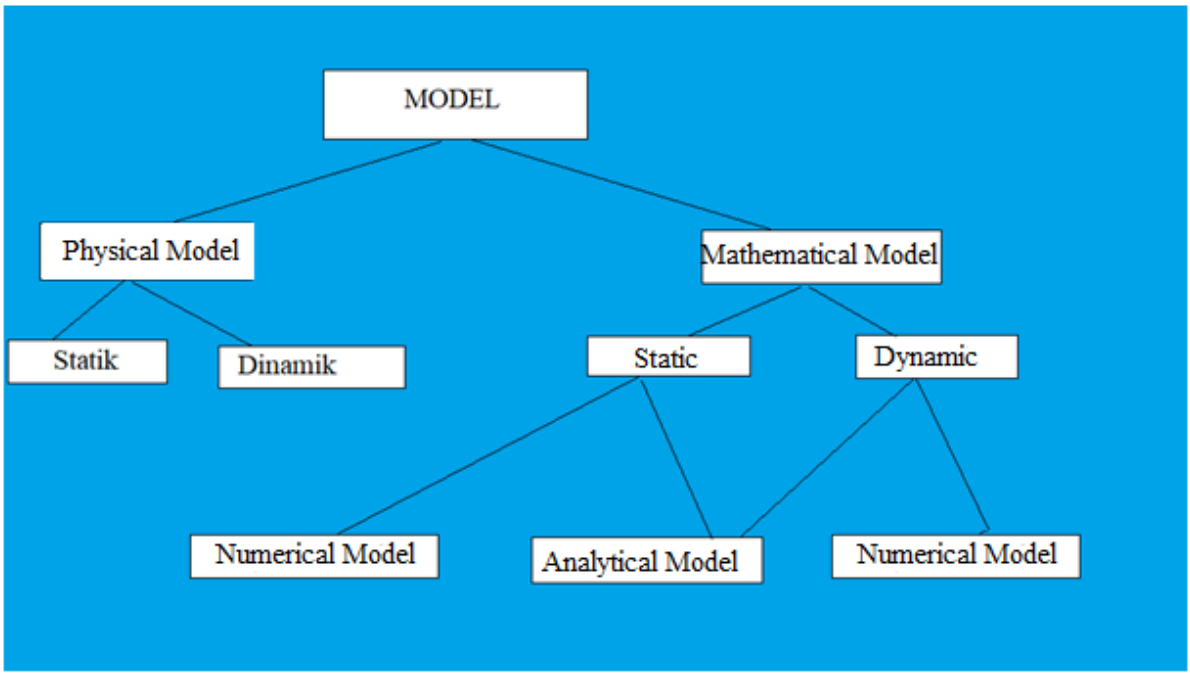

Fig. 1. Mathematical Modeling.

3. Cyber Defense System Modeling Principle

Cyber assaults are the biggest issue in the modern world. Understanding the behavior of malicious objects is essential to overcome this problem. This is very important in mathematical modeling. Malicious objects such as Viruses, Worms, Trojans, Spam and certain technologies such as instant messaging, bots, phishing can be understood and defended against them using modeling. Situations can be determined where some necessary assumptions such as the lifetime of the data, the time of collection, the number of connections can be applied. Malicious objects; It should be estimated with the help of calculations to be made and mathematical equations representing the state of information. More exercises need to be done to make the security model adaptable[5].

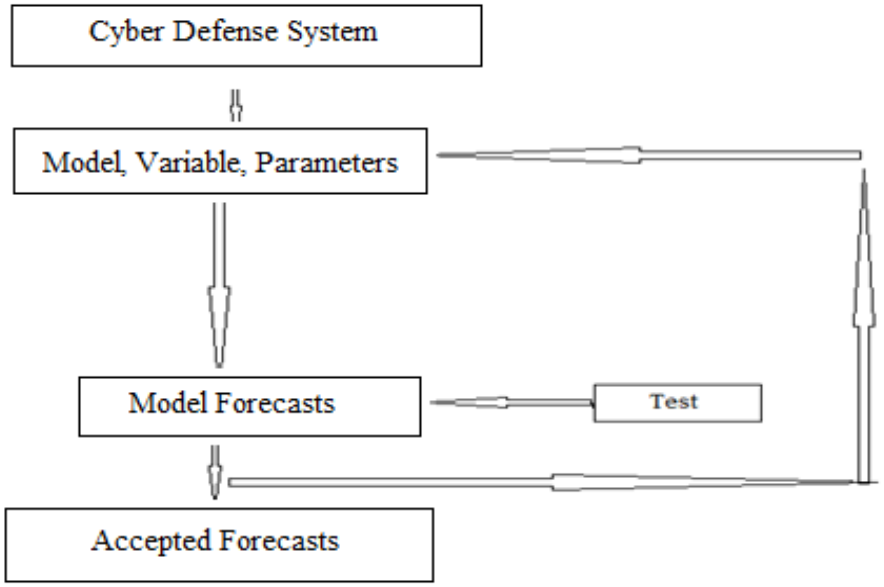

Fig. 2. Mathematical Modeling Process.

4. Malware Objects and Defensive

Since the assaults on the computer are completely stochastic, we cannot know the real-time of the subsequent assault. Only through probability concepts in modeling the probability of the attack can we find it.

To express the attack time of the stochastic variable x_i $(i = 1,2,3, \dots)$ the probability of x_i is expressed by $P(x_i)$.

where n_i is the number of attacks from a certain source and N is the total number of attacks.

$$P(x_i) = \frac{n_i}{N} \quad (1)$$

Here, considering $P(x_i)$ as a set of numbers,
$\int_R P(x_i)dx_i = 1$, that is, the area under the curve is 1.

We can express $\sum_{i=1}^{\infty} P(x_i) = 1$, which is the probability density function. There can also be a probability distribution function that gives the probability of small or equal stochastic attacks to a given value.

$$F(x_i) = \sum_{x_i \le x} P(x_i) \quad (2)$$

Different measurements of the probability function such as mean, mode, median, and standard deviation can be used to examine the stochastic system. Characteristic equation models can be Linear and Non-Linear. The non-linear system can be expressed by partial differential equations.

Let's assume that the malicious object has *P* spread feature due to diverse other factors such as A, B, C,…. In this case, it can be body forth by $P = f(A, B, C, \dots)$. From here, taking the 1st and 2nd derivatives, respectively;

velocity $\frac{\partial P}{\partial t} = \frac{\partial f(A,B,C,\dots)}{\partial t}$, acceleration $\frac{\partial^2 P}{\partial t^2} = \frac{\partial^2 f(A,B,C,\dots)}{\partial^2}$ is calculated.

The imitated conclusions acquired by using particular approximation techniques mentioned down can be used to complement the simulation-created data as well as verify.

4.1. Taylor Series Expansion

Any function with a derivative can be widened by the Taylor formula. In an area close $X = a$, the function $f(x)$ can be approached using the polynomial for the value of the independent variable x.

$$F(x) = f(a) + f^{\iota}(a)(x-a) + \frac{f^{\iota\iota}(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n. \quad (3)$$

4.2. Finite Difference Approach Methods

This method divides partial differential equations into small intervals. It does this in two ways;

I. Advanced Difference Approach:
Calculates the gradient of the function at diverse points with the formula

$$f^{\iota}(x_i) = \frac{(f(x_{i+1}) - f(x_i))}{\Delta x}. \quad (4)$$

II. Backward Difference Approach:
Calculates the gradient of the function at various points with the formula

$$f^{\iota}(x_i) = \frac{(f(x_i) - f(x_{i-1}))}{\Delta x}. \quad (5)$$

4.3. Higher-Order Derivatives

These derivatives can be used to express diverse significant points in the dispersion according to the formula below;

$$f^{(n)} = (f^{(n-1)})^{\iota} \quad (6)$$

Tests can be used to verify the model like polynomial regression tests. Thus, it can be determined whether the values can be placed in a polynomial. The characteristic equation is obtained, the conclusions can be verified experimental or analytical, according to existing standard math hypotheses. To validate the mathematical model, it is the first examination of dimensional homogeneity that requires each term to have the same mesh size. The second is to verify the quality and limit behavior of the models by checking. Apart from these, depending on how large the errors are, some issues such as accuracy and precision, the ability to prepare the data with mean, mode, median or standard deviation can be examined. These data are with ease comparable and can aid us to understand the attitude of malicious objects. For example, some definitions of the virus can be modeled as follows;

I. Let v be a simple virus that affects existent files and make them behave in the same way, a set of programs P, $v \in P$ and $p_i \in P$, $f(v)$ and $f(p_i)$ is the attitude of virus v and p_i, respectively;

$$f(v) = f(p_i) \quad (7)$$

II. Definition of the virus at a constant t

$$T(v, p_i, e, t, S) = \log\left(\frac{f(v,e,t,S)}{f(p_i,e,t,S)}\right). \quad (8)$$

Here $f(v, e, t, S)$ and $f(p_i, e, t, S)$ gives the attitude of programs v and p_i, in order of, at time t in S the system where e the event occurs.

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

III. Definition of the virus based on a continuous-time interval Δt

$$T(v, p_i, e, t, S) = \log\left(\frac{\int_{\tau_0}^{\tau_1} f(v,e,t,S)dt}{\int_{\tau_0}^{\tau_1} f(p_i,e,t,S)dt}\right). \quad (9)$$

Here, the functions $\int_{\tau_0}^{\tau_1} f(v, e, t, S)dt$ and $\int_{\tau_0}^{\tau_1} f(p_i, e, t, S)dt$ respectively It gives the attitude of programs v and p_i in the time range $\Delta t = \tau_1 - \tau_0$ in S the system at the occurrence of e event. If $T\ (v, p\_i, e, t, S) = 0$ then program v is a virus otherwise it is not.

Now, to model a cyber defensive system, it is necessary to combine different components such as sensors and abuse, situational awareness, defense contraption, command and control, strategies and tactics, and science and engineering.

A cyber defense system design model using these components is given in figure 3[6,3].

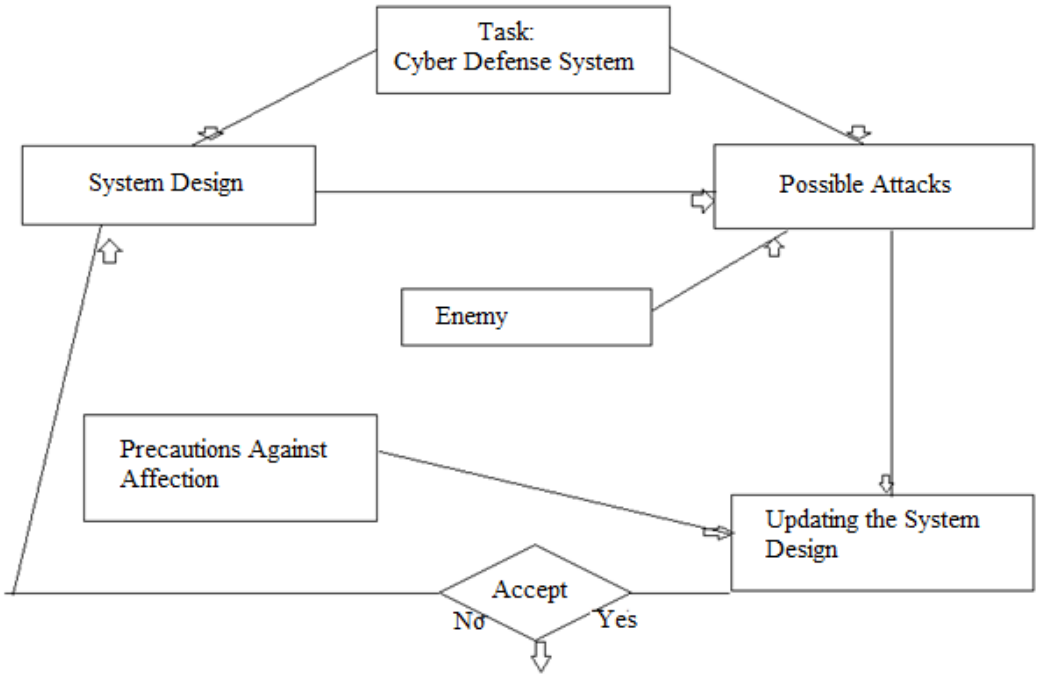

Fig. 3. Cyber Defensive System Design Model.

5. Conclusion

Although various models and solutions are presented, some problems still await solutions. These problems are mainly:

- All properties of malicious objects must be body forth in the shape of mathematical equations to predict their future behavior. Also, a realistic model should be created to apply these characteristic equations in the current environment.

- The accuracy of a model can be ensured after it has been analyzed and developed based on imitated conclusions. This verification assists to preserve the system against assault by malignant objects. Worms similar code red are modeled with high correctness and are now easily controllable. Although most models have certain verification limits, it is still possible for them to be developed[1].

- Most malicious objects have an increase in their behavioral complexity, with a lot of features to be identified. Therefore, each feature should be generalized and the feature area should be narrowed[7].

- Most cyber defense systems incur fixed expenses, such as packet length, increasing data length, or more expensive for comparison, which slows down the current system when implemented. For this reason, it is necessary to provide such a cyber defense system that does not create any more fixed costs. Defense systems that provide recovery for both pre-attack and post-attack must be highly verified and sufficiently predictable based on mathematical modeling.

- It is hard to track down the attacker because of low cognizance in cyber defense. Therefore, we need to develop an appropriate policy and awareness to limit such malicious activity.

References

[1] A. O. Kalashinkov, (2014). Example of using game-theoretic approach in problems, *Cybersecurity Issues*, 1(2), ss. 49-54.

[2] D. K. Saini, (2011). A Mathematical Model for the Effect of Malicious Object on Computer Network Immune System, Applied Mathematical Modeling, 35, ss. 3777-3787. DOI:10.1016/.2011.02.025

[3] D. K. Saini, (2011). A Mathematical Model for the Effect of Malicious Object on Computer Network Immune System, Applied Mathematical Modeling, 35, ss. 3777-3787. DOI:10.1016/.2011.02.025

[4] D. K. Saini, B. K. Mishra, (2007). Design Patterns and their effect on Software Quality, ACCST Research Journal, 5(1), ss. 356-365.

[5] D. K. Saini, N. Gupta, (2007). Fault Detection Effectiveness in GUI Components of Java Environment through Smoke Test, Journal of Information Technology, ISSN 0973-2896, 3(3), ss. 7-17.

[6] D. K. Saini, H. Saini, (2008). VAIN: A Stochastic Model for Dynamics of Malicious Objects, Journal of Systems Management, 6(1), ss. 14- 28.

[7] O.I. Stasuk, L.L. Goncharova, (2017). Differential mathematical models to investigate the computer network architecture of an all-mode Systems of control over a distance of railways, *Cybernetics and Systems Analysis*, 53(1), ss. 157-164.