

ПЕТЯ ИВАНОВА ПЕТРОВА

**МОДЕЛИРАНЕ НА МРЕЖОВИ АТАКИ И АЛГОРИТМИ
ЗА ЗАЩИТА**

**АВТОРЕФЕРАТ
ЗА ПРИСЪЖДАНЕ НА ОБРАЗОВАТЕЛНА И НАУЧНА
СТЕПЕН „ДОКТОР”
ПО ПРОФЕСИОНАЛНО НАПРАВЛЕНИЕ
4.6 ИНФОРМАТИКА**

НАУЧНИ РЪКОВОДИТЕЛИ:

**ПРОФ. Д.Т.Н. АНДОН ДИМИТРОВ ЛАЗАРОВ
ВВМУ „Н. Й ВАПЦАРОВ“**

**ПРОФ. Д-Р ГЕОРГИ ГЕОРГИЕВ ДИМИТРОВ
УНИВЕРСИТЕТ ПО БИБЛИОТЕКОЗНАНИЕ И
ИНФОРМАЦИОННИ ТЕХНОЛОГИИ**

**БУРГАС
2022**

Дисертационният труд е обсъден и насрочен за защита с решение на Научен съвет на Бургаски свободен университет - Бургас от 15.04.2022 г.

Докторантът е зачислен в задочна форма на обучение със Заповед ЛС-38 от 07.12.2015 г. и отчислен с право на защита със Заповед УМО 66 от 31.01.2020 г. на Ректора на БСУ.

Дисертацията се състои от съдържание, списък на съкращенията, списък на фигурите, увод, цел и задачи, четири глави, използвана библиография, благодарности и научни публикации към дисертационния труд. Записката е с общ обем от 114 страници и съдържа 7 фигури и 2 таблици. Към дисертационния труд са оформени 2 приложения. Всяка глава завършва с изводи. Библиографският списък включва 114 източника, от които 95 на латиница, 3 са на кирилица (български език) и 16 интернет източника. Изложението е онагледено с 7 фигури и 2 таблици. Означенията на формулите и графиките в автореферата съответстват на тези от дисертационния труд. Защитата на дисертационния труд ще се състои на 24.06.2022 г. в Център по информатика и технически науки – Бургаски свободен университет – Бургас пред научно жури.

УВОД. ЦЕЛ И ЗАДАЧИ НА ДИСЕРТАЦИОННИЯ ТРУД

1. Обща характеристика на процесите в компютърните информационни системи

Инцидентите в областта на кибернетична сигурност са ежедневие. Примерите са многобройни, от загуба на информация за отделни потребители, атаки със зловреден софтуер и компютърни вируси, до мащабно престъпно поведение, провокирано от организираната престъпност. Доминиращото използване на botnet за разпространение на електронна поща - спам, Distributed Denial of Service Attack (разпределени атаки за отказ на услуга) и разпространението на злонамерен софтуер доведе до активиране дейността на експертите по компютърна сигурност в световен мащаб. Това е причина да бъдат насочени огромен интелектуален потенциал и изчислителни ресурси за изследвания и решаване на проблеми свързани с кибернетична сигурността на компютърните мрежи и информационни системи.

В този смисъл, изграждането на математически модели на функционирането на компютърните мрежи при злонамерени въздействия, изграждането на високоефективни методи и алгоритми за прогнозиране и противодействие на кибератаките, са от съществено значение за решаване на проблемите по осигуряване на устойчива киберсигурност в компютърните мрежи.

Моделиране на поведението на компютърните мрежи включва дефиниране на пространството от инвариантни параметри на поведението на компютърните мрежи, подход, който позволява адекватно моделиране на сложната структура и многообразие от компоненти на Internet.

Телекомуникационен трафик се моделира като статистически процес с вероятностното разпределение на заявките и отговорите на Poisson (Пуасон). В действителност, изследванията на трафика при обмена на данни по Internet показват, че простите модели на Poisson не отразяват адекватно поведението на реален мрежови трафик, включително трафик на локална и глобална мрежа. Поведението на локалната мрежа може да бъде моделирано като

self-similar (само-подобен) процес, point process (точков процеси), marked point process (маркиран точков процеси и time series (времеви редове).

5. Цел и задачи на дисертационния труд

На базата на анализа на процесите, свързани с киберсигурността на компютърните мрежи може да се дефинира целта на дисертационния труд.

Математическо моделиране на процеси в компютърните мрежи при въздействие със злонамерен, изграждане и приложение на генетичните алгоритми за откриване на прониквания в компютърната мрежа и защита на данни.

В съответствие с формулираната цел могат да бъдат дефинирани следните основни задачи:

- Моделиране на процеси при въздействие на компютърна мрежа със злонамерен софтуер.
- Изграждане и приложение на генетични алгоритми за откриване на прониквания в компютърната мрежа.
- Определяне на функцията на пригодност в генетичния алгоритъм за откриване на проникване в компютърните мрежи.
- Реализация на мрежова сигурност чрез криптиране с генетичен алгоритъм.

ГЛАВА I

МОДЕЛИРАНЕ НА ПРОЦЕСИ ПРИ ВЪЗДЕЙСТВИЕ НА КОМПЮТЪРНА МРЕЖА СЪС ЗЛОНАМЕРЕН СОФТУЕР

Предмет на настоящата глава е математическият анализ и моделът на процесите на податливост, експозиция, инфекция и възстановяване на компютърните мрежи в случай на въздействие със злонамерен софтуер. Поведението на мрежата описва със система от диференциални уравнения. Анализират се два случая: случай на равновесие в компютърната мрежа и случай на не равновесие в компютърната мрежа. Получени са аналитични изрази за изчисляване на мрежовите характеристики в случай на

податливост, експозиция, инфекция и възстановяване на компютърни възли по време на атака със злонамерен софтуер.

1.1. Въведение в проблема и постановка на задачата

Задачата на настоящата глава е изграждане на математически модел на поведението на компютърната мрежа при предразположеност (податливост), експониране, инфектиране и възстановяване след атака със зловреден софтуер, т.е. определяне на броя на звената в компютърната мрежа предразположени към атака, експонирани на въздействие, инфектирани и възстановени след въздействие.

Математически модел на кибернетична атака: Решение на система от диференциални уравнения при равновесно състояние на компютърната мрежа

Основни величини и параметри:

S – броят на предразположените към въздействие на злонамерен софтуер;

E – броят на експонираните към инфектиране със злонамерен софтуер, т.е. боят на инфектирани компютри, които не инфектират други компютри,

I – броят на инфектираните със злонамерен софтуер компютри, т.е. боят на инфектираните компютри, които инфектират други компютри,

R – броят на възстановените компютри,

k изчислителен капацитет на пренасяне, $k = S + E + I + R$

β е степента на инфекциозен контакт,

δ е степента на пропадане на възлите (nodes) в мрежата в резултат на инфекция,

τ е степента на инфектиране на експонирания незащитен клас, изложен на инфекция,

μ е степента на смъртност, дължаща се на атака на злонамерения софтуер.

r – естествен темп на нарастване на броя на компютрите от клас S .

- Функция на възстановяване след атака със злонамерен софтуер

$$\text{Rec}(I) = \begin{cases} \rho I, & 0 \leq I \leq I_{\min} \\ m, & I \geq I_{\min} \end{cases} \quad (1)$$

- I_{\min} е минималният брой инфектирани възли, след което се включва антивирусната програма, $m = \rho \cdot I_{\min}$

- Динамиката на възлите от възприемчивия S клас се дефинира със скоростта на изменение на S

$$\frac{dS}{dt} = r \cdot S - (r \cdot S) \cdot \left(\frac{S}{k}\right) - \frac{\beta \cdot I}{k} \cdot S - \delta \cdot S \quad (2)$$

- Динамиката на възлите от експонирания E клас се дефинира със скоростта на изменение на E

$$\frac{dE}{dt} = \frac{\beta \cdot I}{k} \cdot S - (\tau + \delta) \cdot E \quad (3)$$

- Динамиката на възлите от инфектирания I клас се дефинира със скоростта на изменение на E (експозиция)

$$\frac{dI}{dt} = \tau \cdot E - (\mu + \delta) \cdot I - \text{Rec}(I) \quad (4)$$

- Динамиката на възстановяване на възлите от клас R се дефинира със скоростта на изменение на R

$$\frac{dR}{dt} = \text{Rec}(I) - \delta R \quad (5)$$

Решение на система от диференциални уравнения при равновесно състояние на компютърната мрежа, т.е.

$$\frac{dS}{dt} = 0, \quad \frac{dE}{dt} = 0, \quad \frac{dI}{dt} = 0$$

- при $I < I_{\min}$

$$S' = \frac{a(\delta + \tau)}{\beta \cdot \tau} \quad E' = \frac{a \cdot r \cdot (k\beta - a(\delta + \tau)) - k \cdot \delta \cdot \beta \cdot \tau}{\beta^2 \cdot \tau^2 \cdot k} \quad I' = \frac{r \cdot (k\beta - a(\delta + \tau)) - k \cdot \delta \cdot \beta \cdot \tau}{\beta^2 \cdot \tau \cdot k} \quad (11)$$

където $a = \rho + \mu + \delta$

- при $I > I_{\min}$

$$I_{1,2} = \frac{n \pm \sqrt{t}}{2 \cdot \tau \cdot k \cdot \beta^2}, \quad (16)$$

$$S_{1,2} = \frac{2 \cdot \tau \cdot k \cdot \beta (r - \delta) - n \pm \sqrt{t}}{2 \cdot \tau \cdot r \cdot \beta} \quad (17)$$

$$E_{1,2} = \frac{(\mu + \delta)(n \pm \sqrt{t}) + 2 \cdot m \cdot \tau \cdot k \cdot \beta^2}{2 \cdot k \cdot \tau^2 \cdot \beta^2} \quad (18)$$

където $n = \beta \tau k (r - \delta) - r (\tau + \delta) (\mu + \delta)$, $t = n^2 + 4[\beta^2 \tau k m r (\tau + \delta)]$.

- Условието за ендемично равновесие при включена антивирусна програма

$$\sqrt{t} \geq I_{\min} \cdot 2 \cdot \tau \cdot k \cdot \beta^2 - n \quad (20)$$

Решение на системата от диференциални уравнения при неравновесно състояние на компютърната система

$$\frac{dS}{dt} \neq 0 \quad \frac{dE}{dt} \neq 0 \quad \frac{dI}{dt} \neq 0$$

- Системата диференциални уравнения за S , E , I , R се записва във вида

$$\begin{aligned} \frac{dS}{dt} &= r \cdot S - (r \cdot S) \cdot \left(\frac{S}{k} \right) - (\beta \cdot I) \cdot S - \delta \cdot S \\ \frac{dE}{dt} &= (\beta \cdot I) \cdot S - (\tau + \delta) \cdot E \\ \frac{dI}{dt} &= \tau \cdot E - (\mu + \delta) \cdot I - \rho \cdot I \\ \frac{dR}{dt} &= \text{Re c}(I) - \delta R \end{aligned} \quad (40)$$

Решението на системата диференциални уравнения има вида:

- За класа предразположени към атака компютърни възли

$$S(t) = \frac{E_0 \cdot e^{-a \cdot t}}{1 + E_0 \cdot (b/a) \cdot e^{a \cdot t}} \quad (51)$$

където $a = (r - \beta \cdot I_{\min} - \delta)$, $b = \left(-\frac{r}{k} \right)$

- За класа експонирани към атака компютърни възли

$$E = E_0 \cdot e^{-\omega t}, \quad (70)$$

където

$$\omega_{1,2} = \frac{-(\tau + 2\delta + \mu + \rho) \pm \sqrt{(\tau + 2\delta + \mu + \rho)^2 - 4 \left[(\tau + \delta)(\mu + \delta + \rho) - \frac{\tau\beta E_0 e^{-a.t}}{1 + E_0.(b/a).e^{a.t}} \right]}}{2} \quad (66)$$

където $\omega > 0$.

- За класа инфектирани компютърни възли

$$I = I_0 \cdot e^{-\omega.t}$$

- За класа възстановени компютърни възли

$$R(t) = m(1 - e^{-\delta.t}),$$

където $m = \rho \cdot I_{min}$.

1.6. Числен експеримент на оценка от въздействието със злонамерен софтуер върху компютърната мрежа

Численият експеримент се провежда при следните начални стойности на основните параметри уязвимост $S_0 = 93$, експозиция $E_0 = 5$, инфектиране $I_0 = 2$, възстановяване $R_0 = 0$, скорост (степен) на контактно инфектиране $\beta = 0.05$, скорост (степен) на унищожение (пропадане) на възлите $\delta = 0.02$, скорост на инфектиране $\tau = 0.04$, скорост на проникване $r = 0.2$, изчислителен капацитет на пренасяне с нарастване $k = 100$, скорост на възстановяване $\rho_0 = 0.03$, скорост на унищожение в резултат на атака $\mu = 0.01$.

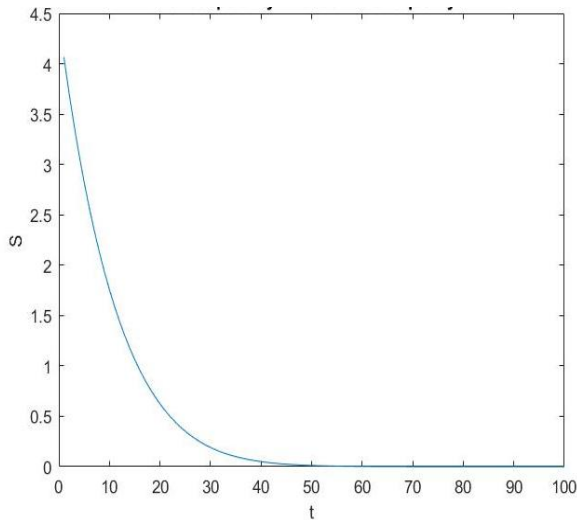
Със символа b се означава степен на включване на нови възли (nodes) във възприемчивия клас, μ е степента на смъртност, дължаща се на атака със злонамерения софтуер (вирус), β е степента на инфекциозен контакт, δ е степента на пропадане възлите (nodes) в мрежата в резултат на инфекция, τ е степента на инфектиране на експонирувания незащитен клас, изложен на инфекция.

Визуализация на резултатите от експеримента

Инфектиране при ендемично равновесие нараства с времето на експозиция, което в началото е стръмно, след което скоростта

на нарастване намалява (Фиг.1). Този ход на кривата на инфектирането $I_1(t)$ следва закона за изменение на податливост (уязвимост) $S_1(t)$ (Фиг. 2) и експозицията (изложението) $E_1(t)$, а (Фиг. 3) на мрежата от зловреден софтуер.

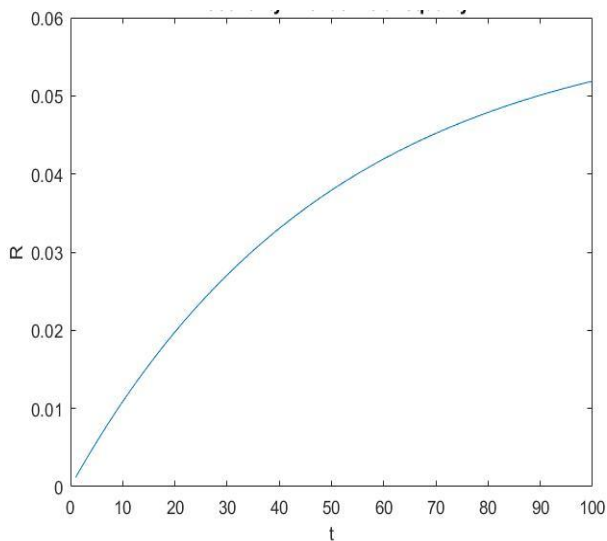
Зависимост на податливост (уязвимост) от времето при ендемично не равновесие е показана на Фиг.1.



Фиг. 1

Фиг. 1. Зависимостта на податливостта (уязвимостта) от времето при ендемично не равновесие.

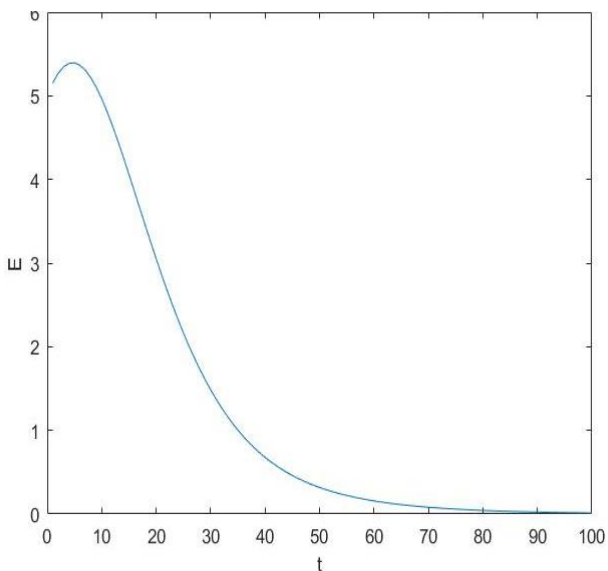
Зависимост на възстановяемост от времето при ендемично не равновесие, т.е. при действие на защитен софтуер е показана на Фиг. 2.



Фиг. 2

Фиг. 2 Зависимост на възстановимост от времето при ендемично не равновесие, т.е. при действие на защитен софтуер.

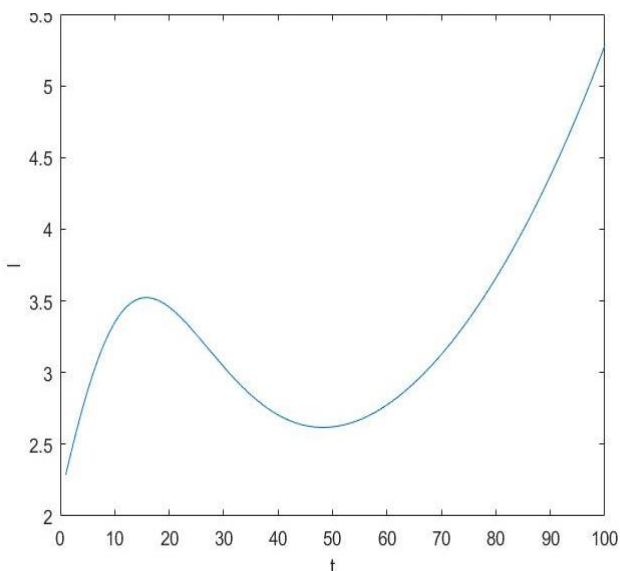
Зависимост на експозиция (изложение) от времето при ендемично не равновесие, т.е. при действие на защитен софтуер е показана на Фиг. 3.



Фиг. 3

Фиг. 3. Зависимост на експозиция от времето при ендемично не равновесие, т.е. при действие на защитен софтуер.

Зависимост на инфектиране от времето при ендемично не равновесие, т.е. при действие на защитен софтуер е показана на Фиг. 4.



Фиг. 4

Фиг. 4. Зависимост на инфектиране от времето при ендемично не равновесие, т.е. при действие на защитен софтуер.

Анализът на резултатите показва, че при ендемично равновесие, т.е. отсъствие на защитен софтуер, показателите, характеризиращи поведението на компютърната мрежа нарастват с изменение на времето, което в началото на кривата е с висока скорост.

При ендемично не равновесие, т.е. наличие на защитен софтуер, параметрите имат сложна зависимост от времето. Интерес представляват зависимостите на експозицията и инфектирането от времето. В началото на интервала на наблюдение има краткотрайно експоненциално нарастване на експозицията, когато защитния софтуер не е в активна фаза ($t = 5$). Инфектирането нараства до момента $t = 15$, след което намалява до момента $t = 60$, когато настъпва момента за обновяване на софтуера. Нарастването на инфектирането е поради не обновените защитни функции на защитния софтуер.

1.7. Изводи

Приведени са основните диференциални уравнения, описващи състоянието на компютърната мрежа при въздействие със злонамерен софтуер. Получени са решения на диференциалните уравнения при равновесие на компютърната система и при отсъствие на равновесие при процесите на податливост, експозиция, инфектиране и възстановяване след атака със злонамерен софтуер. Приложени са оригинални решения на нехомогенните диференциални уравнения и системата от нехомогенни диференциални уравнения.

Разработената методика за оценка на поведението на компютърната мрежа може да бъде приложена при априорно известен закон на въздействие, подчиняващо се, като правило, на разпределението на Poisson.

ГЛАВА II ИЗГРАЖДАНЕ НА ГЕНЕТИЧЕН АЛГОРИТЪМ ЗА ОТКРИВАНЕ НА ПРОНИКВАНИЯ В КОМПЮТЪРНАТА МРЕЖА

Предмет на настоящата глава е анализът на основните функции и структура на генетичния алгоритъм за откриване на прониквания в компютърните мрежи. Последователността от характеристики на мрежовите комуникации се интерпретира като хромозома, която определя правило за откриване на проникване, докато самите характеристики се разглеждат като гени на хромозомата. Основните мрежови характеристики на IP v-4 и IP-v6 и тяхната структура са предоставени и илюстрирани в примери за мрежови връзки с проникване и без проникване. Задачата на настоящата глава е изграждане на синтаксиса на генетичния алгоритъм и анализира мрежовата структура и характеристики (гените и техните кодове) на правилата (хромозомите) за определяне вида на комуникационния обмен.

В съответствие с така дефинираната задача се прави обща характеристика на система за откриване на проникване в компютърните мрежи. Разглеждат се видовете прониквания в

компютърните мрежи. Направена характеристика и се дефинират параметрите на генетичния алгоритъм, използван за откриване и разпознаване на атаки в компютърните мрежи. Дефиниран е синтаксисът и структурата на различни правила (хромозоми) в генетичния алгоритъм. Подробно е описан и илюстриран с примери процесът на кодиране на мрежовите характеристики (гените).

2.5. Дефиниране на синтаксиса и структурата на правилото (хромозомата) в генетичния алгоритъм

Генетичните алгоритми се прилагат за генериране на мрежови характеристики (правила) за оценка на трафика в компютърната мрежа. Тези правила се използват за определяне и диференциране на нормалните и аномалните мрежови свързвания. Аномални връзки са тези, които се отнасят до събития, оценени с висока вероятност като злонамерено проникване в мрежата. Правилата, съхранени в базата от правила (знания) на системата за откриване на проникване, са със синтаксис на клауза от предикатната логика [86, 94]:

If {condition} then {act}

В честта *if* се дефинира състояние (условие), описано с мрежовите характеристики, като IP адреси на източник и местоназначение и номера на портове (използвани TCP/IP мрежови протоколи), продължителност на комуникацията и т.н., включително и индикация за вероятност от проникване. Тази част от правилото генетичният алгоритъм сравнява с мрежовите характеристики от правилата, съхранени в базата от правила на системата за откриване на проникване. Характеристиките в условната част са свързани чрез логически оператор AND. Частта *act* (действие) се отнася до действие, дефинирано от правилата за сигурност, като доклад за предупреждение към системния администратор, спиране на комуникацията, регистриране на съобщение в системни наблюдавани (проверявани) файлове или всички заедно, изброени по-горе. Някои мрежови

характеристики имат по-голям относителен принос при дефиниране на мрежовите свързвания и комуникационен обмен.

2.6. Кодирание на мрежови характеристики (гените) на правилата (хромозомите)

Гените в хромозомите могат да бъдат представени с различни типове данни, двоични числа (байтове) десетични числа или числа с плаваща запетая. Това се обуславя от различния формат и диапазон от стойности на данните за различните мрежови характеристики. В Таблица 1 се привежда названието на мрежовите характеристики - атрибути на хромозомите, броят на гените, дефиниращи атрибутите на хромозомите и техните формати, съответно в първата, втората и третата колона.

Например, характеристиката „Duration“ има три компонента (часове, минути и секунди), всеки от които е представен от един ген от тип байт (Табл. 1). По същия начин, всяка от характеристиките “Protocol” („Протокол“), “Source Port Number”, „Порт за източник“, „Destination port number“ е кодирана с помощта на един ген от тип цяло число, а всяка от характеристиките „Source IP-v4“ и „Destination IP-v4“ има четири компонента. (a, b, c и d), всеки от които е представен от един ген от тип байт, „Source IP-v6“ и „Destination IP-v6“ имат осем компонента. (a, b, c, d, e, f, g, h), всеки от които е представен от един ген от тип байт.

Наименование на атрибутите на хромозомата	Брой на гените	Формат на кода
Source IP-v 4 address	4	a.b.c.d
Source IP-v 6 address	8	a.b.c.d.e.f.g.h
Destination IP -v 4 address	4	a.b.c.d
Destination IP -v 6 address	8	a.b.c.d.e.f.g.h
Source Port Number	1	Integer
Destination Port Number	1	Integer
Duration	3	h:m:s
State	1	Integer
Protocol	1	Integer

Number of Bytes sent by Originator	1	Integer
Number of Bytes sent by Responder	1	Integer
Attack_name	1	Integer

Таблица 1. Атрибути, генна структура и кодове на мрежовите характеристики на хромозомата (правилото) в генетичния алгоритъм.

Атрибутът “Attack name” (име на атака) се намира в частта на правилото (*act*) – *действие*, която класифицира мрежовите характеристики на етап обучение или определя характера на комуникацията на етап откриване на проникване, когато (*condition*) – *състоянието* или *условието* на дадено правило съвпада с това от етап обучение.

Пример на правило, което класифицира мрежова комуникация като атака Denial of Service - DoS (отказ на услуга) *Neptune* е следната хромозома [30]:

*if (duration=“0:0:1” and protocol=“finger” and
source_port=18982 and destination_port=79 and
source_ip=“9.9.9.9” and destination_ip=“172.16.112.50”) then
(attack_name=“neptune”).*

Правилото показва, че ако мрежовият пакет произхожда от IP адрес 9.9.9.9 и порт 18982, и се изпраща на IP адрес 172.16.112.50 и порт 79 с помощта на протокола *finger*, а продължителността на комуникацията е 1 секунда, тогава най-вероятно е мрежова атака от тип *neptune*, която може да доведе до изключване на хоста на дестинацията (местоназначението на пакета). Всяко правило се кодира като хромозома, като се използва вектор с фиксирана дължина, където всяка мрежова характеристика се кодира, използвайки един или повече гени от различни типове (втората и

seconds; the connection is stopped by the originator; the protocol used is TCP; the originator sent 7320 bytes of data; and the responder sent 38891 bytes of data} then {stop the connection}

В този случай кодираната хромозома има вида [11]:

```
/d 1 0 b -1 -1 -1 -1 8 2 1 2 b -1 -1 -1 4 2 3 3 5 0 0 0 8 0 0 0 0 0 0 4 8  
2 1 1 9 0 0 0 0 0 7 3 2 0 0 0 0 0 3 8 8 9 1/
```

Правилото може да бъде интерпретирано по следния начин: ако мрежова комуникация с изходния IP адрес 209.11. ?? . ?? (209.11.0.0 ~ 209.11.255.255), IP адрес на местоназначение 130.18.176. ?? (130.18.176.0 ~ 130.18.255.255), номер на порта на източника 42335, номер на порт за местоназначение 80, времетраене 482 секунди, завършва с състояние 11 (комуникацията е прекратена от създателя), използва протокол тип 9 (TCP), а източникът изпраща 7320 байта данни, отговарящите изпращат 38891 байта данни, тогава това е подозрително поведение и може да бъде идентифицирано като потенциално проникване.

Валидността на това правило се оценява чрез съвпадение на предварителен набор от данни, съставен от свързвания (комуникации), маркирани като аномални или нормални. Ако правилото е в състояние да намери аномално поведение, „награда“ ще бъде даден на текущата хромозома. Ако правилото отговаря на нормална комуникация, ще бъде наложено „наказание“ върху хромозомата. Очевидно нито едно правило не може да се използва за разделяне на всички аномални връзки от нормални връзки. Популацията се развива, за да намери оптималният набор от правила.

Използват се символите (* и ,?) за означение на wildcards (мрежови части на IP адресите), като съответните гени в хромозомата са представени с -1. Тези wildcards се използват за представяне на диапазон от специфични мрежови адреси, т.е. представяне на мрежов блок (диапазон от IP адреси или номера на портове) в правило. След като информацията за полето на

характеристиките е включена в правилата, способността на системата за откриване на проникване може да бъде подобрена, тъй като проникването може да започне от много различни адреси [86]. Включването на времетраенето на мрежовата комуникация в хромозомата осигурява включване на времева информация за мрежови връзки. Максималната стойност на продължителността е 99999999 секунди, което е повече от една година. Това е необходимо за идентифициране на сложни прониквания, които могат да обхванат часове, дни или дори месеци. Ще се подчертае още веднъж, че генетичният алгоритъм стартира с начална популация, която се състои от произволно избрани правила и се развива чрез използване на операторите на кръстосване и мутации. В съответствие с ефективността на оценъчната функция на пригодност, следващите популации са подчинени на правилата, които съответстват на проникващите свързвания. При end (край) на генетичния процес, алгоритъмът спира, правилата се селектират и добавят в базата на системата за откриване на проникващи свързвания.

Правилата за класифициране на атаки DoS (Smurf, Mailbomb), R2L Warezmaster, multihop), U2R (Snmpguess, Buffer-overflow), Probing (ip-sweep, saint), изведени от базата с данни за проникванията на DARPA-USA, имат следната структура [91]:

DOS:

Rule 1 – *if duration = 0 and protocol_type = tcmp and dst_host_srv_count = 255 and then Smurf*

Rule 2 – *if duration = 1 V 5 V 11 and protocol_type=tcip and dst_host_srv_count >= 2^ <= 247 and then Mailbomb*

R2L:

Rule 3 – *if duration = 0v duration <=289 and protocol_type = tcp and dst_host_srv_count >=1^ <= 128 and then waremaster*

Rule 4 – *if duration = 0 and protocol_type = icmp V top V udp and dst_host_srv_count >= 1^ <= 20 and then multihop*

U2R:

Rule 5 – if duration = 0 V duration <= 289 and protocol_type = udp and src_bytes-I > and then Snmpguess

Rule 6 – if I and = 0 and protocol_type = tcp and dst_host_srv_count <= 100 and then buffer-overflow

Probe:

Rule 7: if duration = 0 and protocol_type – icmp and dst_host_srv_count >= 1^ <=255 and then ipsweep

Rule 8: if duration = 0 and duration <= 11 and protocol_type – icmp V tcp V udp and dst_host_srv_count >=1^ <=255 and then saint

В дефиницията на горните правила са използвани следните означения „<=“ (по-малко или равно), „=>“ (равно или по-голямо), ^ (логическо „и“).

2.7. Основни операции върху хромозомите от дадено поколение и експериментална оценка на функцията на пригодност на Firas Alabsi

Селекция

При генериране на всяко следващо поколение част от получената популация се избира да създаде ново поколение. Решенията за избор на индивидите (хромозомите или правилата) се извеждат чрез базиран процес, което гарантира да бъдат избрани с висока вероятност индивиди с по-високи стойности на функцията на пригодност. От популацията се избират двойки хромозоми, които да бъдат родители на следваща популация, т.е. да се извърши рекомбинация на двойки хромозоми.

Кръстосване (рекомбинация)

Кръстосването или рекомбинацията създава едно или повече нови поколения от родителските хромозоми, за да се получат подобри хромозоми с високи стойности на функцията на пригодност.

Мутация

Мутацията променя произволно новото потомство. Това се прави, за да се предотврати попадането на всички решения в

популацията от хромозоми в локален оптимум при вземане на решение.

На Фиг. 5 е показана експериментална оценка на еволюцията на текуща генерация от хромозоми чрез функция на пригодност (FF) и тестова хромозома, рекомбинация и мутация, и нова генерация от хромозоми в бинарен формат; със знак ♀ и ♂ са означени хромозоми за рекомбинация; със знак ● са означени невалидна хромозома, със знак ○ е означена хромозома без изменение, със знак ◌ е означена мутираеща хромозома.

Current Generation	FF	Crossover&Mutation	New Generation
1011101010001010	0.63 ♀		1011101010001010
1000011110101000	0.43 ●	1011100100001001	1011100100001001
1100101000101011	0.50 ●	0101111010001010	0101111010001010
0101111010001001	0.75 ♂		0101111010001001
0101000111010101	0.56 ◌	0101000011010101	0101000011010101
0010111000101011	0.56 ○		0010111000101011

↑

Test Chromosome
0101111010001101

Фиг. 5.

Фиг. 5. Оценка на текуща генерация от хромозоми и Test Chromosome, чрез функция на пригодност при рекомбинация и мутация, и нова генерация от хромозоми в бинарен формат; ♀ и ♂ хромозоми за рекомбинация; ● невалидна хромозома, ○ хромозома без изменение, ◌ мутираеща хромозома.

Функцията за текуща хромозома и тестова хромозома се изчислява като отношение на броя на съвпадащите битове в хромозомите към броя на битовете в тестовата или съпоставимата хромозоми, т.е.

$$Fitness = M_b / N_b,$$

където M_b е броят на съвпадащите битове в тестовата и съпоставимата хромозоми, N_b е броят на битовете в хромозомата – тестова и текуща съпоставима.

В случая общият брой на битовете в хромозомата е 16.

2.9. Изводи

Разработена е идеята за прилагане на генетични алгоритми в системите за откриване и превенция от прониквания в компютърните мрежи. За оценка на текущи и нови прониквания се предвижда използване на интегрирани и взаимно-допълващи се оценъчни функции на пригодност, с което ще се повиши точността на оценките и ефективността на превенция на прониквания в компютърните мрежи.

Приложен е функцията на пригодност оценка на мрежовите характеристики (гени) при откриване на проникване в компютърните мрежи, базирано на генетичен алгоритъм. Приведени са обобщен генетичен алгоритъм и псевдо-код. Направена експериментална оценка на метода на Firas Alabsi за изчисление на функцията на пригодност на генерираните правила (хромозоми).

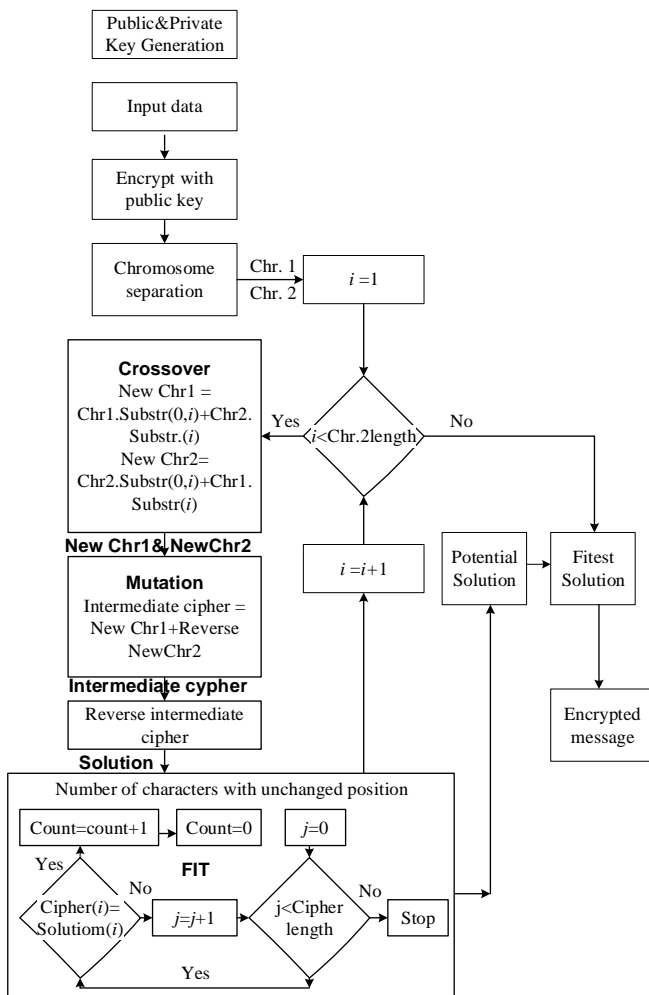
Като бъдеща активност на автора в областта на откриване на злонамерени прониквания в компютърните мрежи чрез прилагане на подхода на генетичните алгоритми се предвижда изграждане на база от данни (правила, знания) за нови неизследвани структури от мрежови характеристики на компютърни атаки, както и методи за тяхното противодействие и превенция.

ГЛАВА III

РЕАЛИЗАЦИЯ НА МРЕЖОВА СИГУРНОСТ С ПОМОЩТА НА КРИПТИРАЩ ГЕНЕТИЧЕН АЛГОРИТЪМ

3.1. Обща характеристика на комбинирания криптиращ алгоритъм

Сигурността в предаване на данни е основен проблем в комуникационните системи. Предмет на настоящата глава е сигурността на предаване на поверителна информация и на данни с разработване на публичен алгоритъм за криптографирано на обикновен текст с използване на генетичен алгоритъм, за да се осигури поверителност, автентичност, цялостност и безотказно предаване на информацията. Защитата на информацията се постига с приложение на различни криптиращи техники. Един от мощните методи за криптиране е асиметричното RSA ((Rivest-Shamir-Adleman)) криптиране, което може допълнително да бъде усилено с приложение на инструментите на изкуствения интелект, като генетичният алгоритъм. За разлика от RSA криптиращ алгоритъм, който стартира избора на две прости числа, в дисертационния труд се разработва криптиращ алгоритъм, който се базира на случайния избор на две взаимно не прости числа, който се усилва чрез използване на основни операции от генетичния алгоритъм – рекомбинация и мутация. Мутацията се постига чрез конкатенация на двете хромозоми, едната от които е с инверсен запис на шифровите символи. Схемата на криптирането на съобщението в предавателя е представена на (фиг. 6). В отделния блок публичният и частният ключ се генерират на базата на две неравноправни числа, генерирани на случаен принцип в интервала 10 - 1000. Във втория блок се представят входните данни, т.е. предаваното съобщение, което трябва да бъде криптирано. Криптирането в трети блок се извършва операция с публичен ключ над входните данни.



Фиг. 6 Блок-схема на криптиране на съобщението и повишаване на устойчивостта на шифъра с генетичен алгоритъм

Низът на криптираното съобщение е разделен на две равни части с дължина $n/2$. Първата част е първата хромозома, втората част е втората хромозома. Създава се цикъл с индекс в диапазона от 0 до дължината на втората хромозома $n/2$. На всяка i -та стъпка от цикъла се прилагат следните оператори:

Описание на алгоритъма

Първо се генерират публични и частен ключ като се прилага алгоритъм с използване на две не взаимно прости числа. При RSA шифриране се използват две прости числа. Подава се текстово съобщение, което се криптира с публичен ключ. Получава се криптирано съобщение (шифър). Шифриранията текст се разделя на хромозоми.

Стрингът на криптираното съобщение се разделя на две равни части от символи с дължина $n/2$ символа. Първата част е първа хромозома, втората част е втората хромозома. Създава се цикъл с индекс в интервала от 0 до дължината на втората хромозома $n/2$. На всяка i -та стъпка на цикъла се прилагат операциите от генетичния алгоритъм:

Кръстосване

От двете хромозоми се генерират две нови хромозоми чрез конкатенация на две отделни части от двете хромозоми, определени от i -та точката на рекомбинация (кръстосване):

New Chromosome 1 = Chromosome 1 Substring (0, i) + Chromosome 2 Substring (i)

От първата хромозома се взема substring от символи с дължина от 0 до i . От втората хромозома се взема substring от символи с дължина i .

New Chromosome 2 = Chromosome 2 Substring (0, i) + Chromosome 1 Substring (i)

От втората хромозома се взема substring от символи с дължина от 0 до i . От първата хромозома се взема substring от символи с дължина i ,

където i е точката на рекомбинация или кръстосване (crossover).

Мутация

Мутацията се реализира чрез конкатенация на New Chromosome 1 с Reverse New Chromosome 2.

Получава се междинен шифър (междинно шифрирано съобщение), което след инверсия е потенциално решение за шифровано съобщение.

Операторът за мутация се прилага, за да се разместят позициите на символите в криптираното съобщение, без да се променят самите символи, което увеличава ентропията на шифъра. Мутацията се извършва чрез конкатенация на New chromosome 1 с инверсната New chromosome 2.

Полученият стринг на шифъра, манипулиран от мутационния оператор, се генерира междинен шифър, т.е. междинното криптирано съобщение. Междинното криптирано съобщение се подлага на инверсия. Инверсията на шифъра допълнително размества позициите на знаците и увеличава ентропията на шифъра. Инверсията на междинния шифър е потенциално решение за криптирано съобщение.

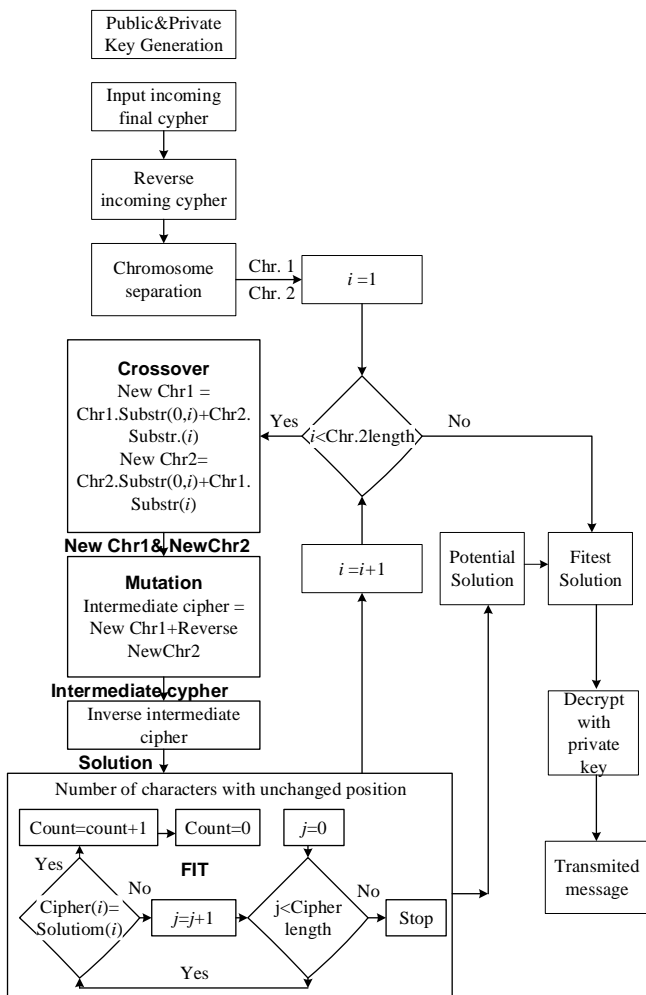
Оценка на шифъра с функция на пригодност, операции в блок FIT

Дали потенциалното решение за шифрираното съобщение е решение за крайния шифър се оценява, като се определят колко символа в шифрираното съобщение остават с непроменени позиции в междинния шифър чрез специална функция на пригодност. Сравняват се шифрираното с публичен ключ съобщение с потенциалното решение на шифровано съобщение, получено след приложение на операциите от генетичния алгоритъм рекомбинация и мутация. Минимална стойност на функцията на пригодност е критерий за решение на шифъра.

Формира се цикъл за оценка на позицията на всеки j -ти символ от шифрованото съобщение, дали е с променена позиция по отношение на шифрираното с публичен съобщение. Операторът counting определя колко символа от шифрираното съобщение и манипулираното с генетичен алгоритъм остават с непроменени позиции.

Потенциалното решение на шифъра, низът с минимален брой символи с непроменени позиции, т.е. низът с минимална функция на пригодност за годност е решението на шифъра.

В приемника, т.е. на етапа на декриптиране, операторите се прилагат в обратен ред (Фиг. 7).



Фиг. 7. Блок диаграма на декриптиране на шифъра и получаване на предаденото съобщение

Първо, операторите на генетичния алгоритъм се използват за възстановяване на шифровото съобщение. След това се използва частен ключ за възстановяване на съобщението. Схемата на дешифрирането на шифъра в приемника е представена на (фиг. 7). Публичните и частните ключове са представени в първия

отделен блок. Във втория блок се представят входните данни, т.е. шифрованото съобщение, което трябва да се декриптира. Първата операция над низа от знаци на шифъра е инверсията на низа на шифъра. След това низът на шифъра се разделя на две хромозомни последователности с еднаква дължина и процедурата продължава както в етапа на криптиране, включващ оператори, crossover на хромозомите, mutation с конкатенация на първите хромозоми с инверсната версия на втората хромозома, инверсия на шифър и дефиниране на потенциално решение чрез функция на пригодност. Крайният шифър с минимална функция на пригодност се дешифрира с частен ключ за получаване на предаденото съобщение.

3.2. Структура на криптиращ алгоритъм, усилен с операторите на генетичния алгоритъм

Структурата на криптиращия алгоритъм и методите за реализация могат да се дефинират по следния начин.

- Да се генерират двойка ключове /публичен и частен/. В асиметричната-ключова криптография е в сила правилото, че когато даден текст е шифрован с публичен ключ той се дешифрира със съответния личен ключ и обратно. Обикновеният текст е шифрован с публичен ключ за производство на междинен шифър. Междинният шифър е шифрован с генетичен алгоритъм за производство на последния шифър.

- Дешифриращ метод:

Окончателният шифър е дешифриран с помощта на генетичен алгоритъм, за да получите междинен шифър, който отново се дешифрира с помощта на съответния личен ключ, за да се получи обикновеният текст.

- Първоначално шифриране:

Междинен шифър = шифроване /обикновен текст, публичен ключ/
или

ИЛИ

Междинен шифър = шифроване /обикновен текст, частен ключ/

И

Генетично шифроване: окончателен шифър = шифроване/междинно шифроване/

- Генетично дешифриране:

Междинно шифроване = дешифриране /окончателен шифър/

Окончателен дешифриращ: обикновен текст = дешифриране/междинен шифър, частен ключ/

ИЛИ

Обикновен текст = дешифриране /междинен шифър, публичен ключ/

Основни стъпки и компоненти на генетичния алгоритъм

1. Генериране на ключове:

Стъпка 1: Generate two non-coprime number /a, b/

Стъпка 2: Set $l = \text{L.C.M} /a, b/$; $g = \text{G.C.D} /a, b/$

Стъпка 3: Set /a, g/ as private key

Стъпка 4: Set $x = ((a-1)/g) + ((a-1)\%g)$; $y = g$

Стъпка 5: Store x, y

Стъпка 6: Set (b, l) as public key

Стъпка 7: Set $p = ((l-1)/b + ((l-1)\%b))$; $q = b$

Стъпка 8: Set p, q

Стъпка 9: Stop

Функция:

Input: text, key type, public key (p, q) or private key (x, y)

Output: Intermediate Cipher

Стъпка 1: if key type = private key Read public key (p, q) from database Set m: = 2 else Read private key (x, y) from database Set m: = 1

Стъпка 2: Set $x = ((x-(y-1))^y)+y$

Стъпка 3: Set $p = ((p-(q-1))^q)+q$

Стъпка 4: if $(x/y = p/q)$ and $(x*q = p*y)$ then Set n: = $x*q$

Стъпка 5: Set key_arr []: = n

Стъпка 6: if (length of text = odd number) Then text: = text + @

Стъпка 7: Set i: = 0, c: = "", j: = 0

Стъпка 8: while (i < length of text) Repeat

Съпка 9: if in = EVEN number Set c: = c+text [i] + (m*key_arr[j])
Else
Съпка 10: Set c: = c+text [i] – (m*key_arr[j]) End if
Съпка 11: Set j: = j+1
Съпка 12: if j = key_length then set j: = 0
Съпка 13: Print c
Съпка 14: Stop

- **Евристично криптиране:**

Прилага се генетичен алгоритъм:

Input: Intermediate Cipher (c)

Output: Final Cipher

Съпка 1: Set st1: = substring of c (1 to c/2)

Съпка 2: substring of c (c/2 to c)

Съпка 3: while (i<length of text) Repeat Step 4 to Step 8

Съпка 4: perform crossover at mate point I with st1 (0, i) and st 2 (I, n) and st 2 (0, i) and st 1 (I, n)

Съпка 5: Set st: = st 1 + Reserve (st 2)

Съпка 6: Reserve st and set gst = st

Съпка 7: compare c and gst to find out fit value

Съпка 8: select the gst having minimum no of fit value

Съпка 9: Return gst

Съпка 10: Stop

- **Декодиращ евристичен алгоритъм:**

Реализация на генетичния алгоритъм:

Input: Final Cipher (c)

Output: Intermediate Cipher

Съпка 1: Set st: = reserve (cipher)

Съпка 2: Set n = length of cipher

Съпка 3: Set st 1: = substring of c (1 to n/2) st 2: = substring of c (n/2 to n)

Съпка 4: Set st 2: = Reverse of st 2 i: = 1

Съпка 5: while (i<length of st 2) Repeat Step 6 to Step 9

Съпка 6: perform crossover at mate point i with $st\ 1(0, i)$ and $st\ 2(i, n)$ and $st\ 2(0, i)$ and $st\ 1(i, n)$

Съпка 7: Set $gst = st\ 1 + st\ 2$

Съпка 8: compare gst and cipher to find out fit value (no of character position Remain unchanged)

Съпка 9: select the gst having minimum no fit value

Съпка 10: Return gst

Съпка 11: Stop

- **Дешифриращ евристичен алгоритъм с пример:**

Функция:

Input: Intermediate cipher, keytype, public key (p, q) or private key (x, y)

Output: Plain Text

Съпка 1: if key type = public key Set $m = 2$ and Read private key (x, y) from atabase else Set $m = 1$ and Read public key (p, q) from database

Съпка 2: Set $x = ((x-(y-1))*y)+y$

Съпка 3: Set $p = ((p-(q-1))*q)+q$

Съпка 4: if $(x/y = p/q)$ and $(x*q = p*y)$ then Set $n = x*q$

Съпка 5: Set $key_arr [] = n$

Съпка 6: Set $i = 0, c = "", j = 0$

Съпка 7: while $(i < \text{length of plmcipher})$ Repeat Step 4 to Step 7

Съпка 8: if $i = \text{EVEN number}$ then Set $c = c + \text{plmcipher}[i] - (m * \text{key}[j])$ Else

Съпка 9: Set $c = c + \text{plmsiphert}[i] + (m * \text{key}[j])$ End if

Съпка 10: Set $j = j + 1$

Съпка 11: if $j = \text{length of key} []$ then Set $j = 0$

Съпка 12: Print c

- **Криптографски алгоритъм:**

Съпка 1: Start

Съпка 2: Call Proposed Encryption Heuristic

Стъпка 3: Call Proposed Encryption Heuristic using Genetic Algorithm

Стъпка 4: Call Proposed Encryption Heuristic using Genetic Algorithm

Стъпка 5: Call Proposed Decryption Heuristic

Стъпка 6: Stop

3.6. Изводи

Разработеният криптиращ алгоритъм, усилен с операторите на генетичния алгоритъм, генерира кодови шифри с висока степен на ентропия, оценена с минималния брой на съвпадащите позиции на символите в криптираното съобщение, което осигурява висока степен на защита на текста, предаван по компютърната мрежа.

Чрез разработеният криптиращ алгоритъм със случайно избрани две взаимно-не прости числа, усилен с операторите на генетичния алгоритъм, може да се интерпретира, като разширение на областта на приложение на криптиращия RSA алгоритъм, реализиран със софтуерните инструменти C#.

Разработеният криптиращ алгоритъм с прилагане на операторите на генетичния алгоритъм препотвърждава криптиращите свойства на известни подобни алгоритми чрез разработване на нов софтуерен продукт.

Разработеният и имплементиран в среда C# генетичен алгоритъм е с надеждно и устойчиво криптиране с висока степен на защита на предаваното съобщение и може да бъде приложен при обмен на данни, изискващи високо ниво на сигурност.

ГЛАВА IV ЗАКЛЮЧЕНИЕ-РЕЗЮМЕ НА ПОЛУЧЕНИТЕ РЕЗУЛТАТИ

В заключение следва да се отбележи, че в съответствие с целта и поставените задачи в дисертационния труд е извършено математическо моделиране на процеси в компютърните мрежи при въздействие със злонамерен софтуер, дефиниране на

синтаксиса на генетичен алгоритъм за откриване на прониквания на компютърната мрежа, предложен е модифициран криптиращ алгоритъм, усилен с генетичен алгоритъм при изграждане на шифрованото съобщение. Получени са следните научни, научно-приложни и приложни резултати, които се представят в резюме според изискванията на чл. 27 (2) от Правилника за приложение на ЗРАСРБ:

4.1. НАУЧНИ РЕЗУЛТАТИ

4.1.1. Предложен е модел на процесите на податливост, експозиция, инфекция и възстановяване на компютърна мрежа в случай на въздействие на злонамерен софтуер, описани със система от диференциални уравнения за моментна и прогнозна оценка на състоянието на компютърната мрежа – система от уравнения (6).

4.1.2. Предложено е оригинално решение на системата от диференциални уравнения в два случая - на равновесие при константни променливи – аналитичен израз (37) - глава I и отсъствие на равновесие при време зависими променливи - аналитичен израз (65) - глава I, дефиниращи класовете състояния на машините в компютърната мрежа.

4.1.3. От решението на системата от нехомогенни диференциални уравнения са изведени аналитични изрази за изчисляване на мрежовите характеристики в случай на податливост, експозиция, инфекция и възстановяване (реконструкция) на машините в компютърната мрежа по време на атака със злонамерен софтуер.

4.2. НАУЧНО-ПРИЛОЖНИ РЕЗУЛТАТИ

4.2.1. Разработен е софтуерен инструмент за защита на компютърна мрежа чрез криптиране на предаваната информация с приложение на операторите на генетичен алгоритъм, реализиран с програмния език C# (глава III).

4.2.2. Разработени са софтуерни продукти, реализирани в среда Matlab за илюстрация на решенията на системата диференциални уравнения за моментна и прогнозна оценка.

4.3. ПРИЛОЖНИ РЕЗУЛТАТИ

4.3.1. Разширен е списъкът с атрибути, генната структура и кодовете на мрежовите характеристики на хромозомата (правилото) в генетичния алгоритъм, разширяваща обхвата на неговото действие по откриване на мрежови прониквания (Таблица 1, глава II).

4.3.2. Оценена е структурата на функцията на пригодност на Firas Alabsi и са илюстрирани основните операции върху хромозомите от дадено поколение с данни, получени от числен експеримент.

4.3.3. Резултатите от оценката са приведени на Фиг. 5 (глава III). Направена е експериментална оценки на параметрите A, AB на функцията на пригодност на Firas Alabsi с данни, получени от симулиране на мрежови комуникации от типа Normal, DoS, R2L, U2R, Probe, реализирани със случайно генерирани мрежови характеристики на пет хромозомни структури за всяка категория. Резултатите са представени в Таблица 2 (глава III).

4.3.4. Тези научни, научно-приложни и приложни резултати са добра основа и предпоставка за бъдещи изследвания в технологиите: Индустрия 4.0-изкуствен интелект, Интернет на нещата, роботиката, киберсигурността, както и на методите за вземане на решения, многокритериалния синтез, анализа, превенцията и управлението на риска, големите данни (big data), размитите оценки и Soft Computing.

Определени концентрирани усилия и творчески подходи са необходими в обучението с акцент върху e-learning за създаване на знания, компетентности и умения в широк кръг от студенти, изследователи и специалисти.

СПИСЪК НА АВТОРСКИТЕ ПУБЛИКАЦИИ ПО ТЕМАТА НА ДИСЕРТАЦИЯТА

1. Lazarov, A., P. Petrova. Genetic algorithm in computer network protection. Engineering Sciences, No. 1, pp. 80-95, ISSN: 1312-5702. E-ISSN: 2003-3542, DOI: 10.7546/EngSci.LIX.22.01.07, LIX, 2022.
2. Lazarov, A., P. Petrova, Modelling activity of a malicious user in Computer Networks. Cybernetics and information technologies, Volume, No 2, Sofia 2022. Print ISSN: 1311-9702. On-line ISSN: 1314-4081. SJR 027, Q2.
3. Lazarov A., P. Petrova. Crypto genetic approach in information security, XXII International Symposium on Electrical Apparatus and Technologies SIELA 2022, 1-4 June 2022, Burgas, Bulgaria (in print).
4. Лазаров, А., П. Петрова. Концепция на генетичния алгоритъм за откриване на прониквания в компютърната мрежа, Електронно списание на център по Информатика и технически науки на Бургаски Свободен Университет, Том 8, бр. 1, 2019, pp. 3-12.
5. Лазаров, А., П. Петрова. Определяне на функцията на пригодност в генетичния алгоритъм за откриване на проникване в компютърните мрежи, Списание „Компютърни науки и комуникации”, Том 8, No1 (2019), БСУ, Бургас, pp. 13-22.
6. Лазаров, А., П. Петрова. Моделиране на процесите при въздействие на компютърна мрежа със злонамерен софтуер,

БЛАГОДАРНОСТИ

За развитието и резултатите на настоящата дисертация определено и съществено място заемат ръководството и преподавателите от Център по информатика и технически науки на БСУ и моите научни ръководители проф. д.т.н. Андон Димитров Лазаров от ВВМУ „Н. Й. Вапцаров“ и проф. д-р Георги Георгиев Димитров от Университет по библиотекознание и информационни технологии.

За завършената форма и съдържание на дисертацията имат своя дял критичните конкретни бележки и конструктивни предложения на доц. д-р Пенка Георгиева и доц. д-р Веселина Жечева.

**На всички поднасям моите най-искрени
благодарности!**

*На моето семейство благодаря за пълната
безкористна подкрепа и съпричастност.*

ПЕТЯ ПЕТРОВА