ФКСТЧ4 – НСЛ – 052
17.05.2022 г.

ТЕХНИЧЕСКИ УНИВЕРСИТЕТ ФКСТ СОФИЯ

С Т А Н О В И Щ Е

върху дисертационен труд за придобиване на образователна и научна степен „Доктор“

Автор на дисертационния труд: маг. инж. Ваня Димитрова Иванова

Тема на дисертационния труд: Откриване на IoT базирани ботнет атаки чрез анализ на мрежови трафик

Член на научното жури: доц. д-р инж. Иво Руменов Драганов / съгласно заповед № ОЖ-5.3-09 от 31.03.2022 г. на Ректора на ТУ-София /научно направление 5.3 Комуникационна и компютърна техника / специалност „Телевизионна и видеотехника“

1. **Актуалност на разработвания в дисертационния труд проблем в научно и приложно отношение**

Темата на депозирания дисертационен труд „Откриване на IoT базирани ботнет атаки чрез анализ на мрежови трафик“ е от изключително значение за практиката по детекция на разпределени мрежови атаки с цел извеждане на информационни услуги от достъпност. За 2020 година са регистрирани над 10 млн. ботнет атаки като броят на известните C&C сървъри надхвърля 15 000 само година по-рано. Географското разпространение на последните е изключително широко – в САЩ, Холандия, Германия, Франция, Китай, Русия и други страни. Засегнатите индустрии, развиващи производство и търговия, оценявано на милиарди долари, включват компютърни игри, хазарт, финансови услуги, онлайн магазини, доставка на Интернет и други телекомуникационни услуги. Финансовите загуби в отделни случаи за големите корпоративни представители надхвърлят 2 млн. щатски долара. Разнообразието по начина на организиране на ботнет атаките, включващи централизирани и разпределени архитектури, правят много трудно откриването на инициращите страни и изискват универсални способи за тяхното регистриране и последващ анализ. Целта на дисертационния труд е насочена именно към възможно най-устойчиво откриване на подобен род атаки като тя се постига чрез анализ на мрежовия трафик, обменян между атакуващи машини и засегнати такива.

2. **Степен на познаване на състоянието на проблема и творческа интерпретация на литературния материал**

За доброто познаване на състоянието на проблема от страна на докторанта говори използването на 207 литературни източника, публикувани в реномирани научни издания в световен мащаб. Всички те са на английски език като около три четвърти от тях са от последните 5 години. В анализа на състоянието на проблема докторантът е направил опит да представи единна таксономия на масово прилаганите IoT устройства, голяма част от които могат да бъдат използвани като ботове при изграждането на ботнет. Изследвано е разпределението и влиянието на последните от техническа и икономическа гледна точка. Разгледани са видовете мрежови атаки, осъществявани чрез IoT ботнет. Класифицирани са архитектурите за организация на подобни мрежи като е описан и принципът им на действие. Представени са основните видове ботнет атаки, в частност DDoS атаките като е обърнато внимание и на жизнения цикъл на ботнет. Особено място при разглежданията на докторанта заемат уязвимостите на IoT устройствата, по отношение на които има изграден списък от възможни действия за ограничаване на тяхното експлоатиране. Анализирано е действието на 19 от най-активните ботнет от последните десетина години като са набелязани мерки за навременно откриване на

тяхното действие. Направени са изводи, на чиято база са дефинирани целта и задачите на дисертационния труд. Всичко това свидетелства за задълбочените познания на докторанта в областта, в която попадат и последващите резултати, получени в рамките на дисертационния труд.

3. **Съответствие на избраната методика на изследване и поставената цел и задачи на дисертационния труд с постигнатите приноси**

Съгласно поставената цел и задачи на дисертационния труд, докторантът е насочил своите усилия към намирането на ефективни способи за класификация на агрегирани характеристики от мрежови трафик, обменян между атакуващи и атакувани машини, както и от периоди на нормалното опериране на поддържаните услуги. Приложени са методики и средства, характерни за машинното обучение като анализ на информативността на числено представени признаци с използване на параметри като информационно усилване и коефициент на Джини и последващо тяхно ранговане. Изградени са хистограми на разпределението на представящите характеристиките стойности след подходяща нормализация, търсено е сходство между тях на база на коефициентите на Пиърсън и Спиърмън. Информативната значимост на отделните характеристики е оценена и чрез FreeViz проектиране, както и чрез определяне на вероятностите за поява на изследваните видове атаки по зададена стойност на признак. Приложен е и метода с анализ на главната компонента върху отделните характеристики като е редуцирана размерността им от 10 на 8 компоненти. Последващо обучение на невронна мрежа с право разпространение на данните, машина с поддържащи вектори и ансамбъл от случайно генерирани дървета, както и комбинация между тях, е извършено с пълен набор от 10-комопнентни признаци и с редуциран набор от 8-компонентни такива, след което е оценена точността на класификация върху неизползвани при обучението тестови набори. Оценка е направена чрез параметри като Precision, Recall, Specificity, F1-measure и др. За всеки от тестваните класификатори е намерена оптимална конфигурация като най-точна се оказва тази на свързан двукомпонентен класификатор. Пълният комплекс от математически обработки, които докторантът е приложил, показва неговата способност да борави със средствата на машинното обучение с цел получаване на ефективни реализации на алгоритми за класификация на мрежови трафик при липса и наличие на ботнет атаки.

4. **Научни и/или научно приложни приноси на дисертационния труд**

Докторантът е заявил 2 научни приноса, изразяващи се в разработването на теоретични модели на комбиниран детектор и комбиниран многоцелеви класификатор на 10 вида ботнет атаки, работещи върху агрегирани записи от характеристики на мрежови трафик. Заявени са и 4 научно-приложни приноса, свързани с разработването на оптимизирани модели на еднослойна и многослойна невронна мрежа с обратно разпространение на грешката, на машина с поддържащи вектори и ансамбъл от дървета за вземане на решение с цел откриване и класификация като самостоятелни модули на същите по вид 10 типа ботнет атаки. Представени са и 5 приложни приноса, свързани с техническата реализация на класификаторите и обработката на данните от тяхната работа. На база на представените експериментални резултати в дисертационния труд, както и на математическите описания на отделните класификатори, би могло да се приемат направените от докторанта претенции за приноси за коректни.

5. **Преценка на публикациите по дисертационния труд: брой, характер на изданията, в които са отпечатани**

Към дисертационния труд има представени 7 научни публикации. Всички те са на английски език като в една от тях докторантът е единствен автор. Две са доклади от международна научна конференция, проведена у нас. Пет са статии от международни научни списания. Не са известни цитирания на така представените публикации. Основните резултати, представени в дисертационния труд, са изложени в основната си част във всички 7 публикации по дисертационните изследвания. Начинът на представяне на материала в тях съответства на необходимото научно ниво и е свидетелство за способността на докторанта да представя в обобщен вид получаваните от него резултати.

6. **Мнения, препоръки и бележки**

Препоръката ми към докторанта е да фокусира бъдещата си работа върху внедряване на така получените резултати в практиката по защита на публични и частни мрежи, предоставящи услуги с широк достъп срещу неоторизирани въздействия, в основата на които са разпределените атаки с използване на ботнет.

7. **Заключение с ясна положителна или отрицателна оценка на дисертационния труд**

В заключение смятам, че дисертационният труд на маг. инж. Ваня Димитрова Иванова, озаглавен „Откриване на IoT базирани ботнет атаки чрез анализ на мрежови трафик“ отговаря на изискванията за придобиването на образователната и научна степен „Доктор“ по научна специалност „Автоматизирани системи за обработка на информация и управление“ от направление 5.3 Комуникационна и компютърна техника и затова убедено предлагам на Уважаемите членове на Научното жури да подкрепят удостояването й с нея.

Дата: 16.05.2022 г.

Член на журито:

(доц. д-р Иво Драганов)

OPINION

on the PhD thesis for acquiring the educational and scientific degree "Doctor"
Author of the PhD thesis: MEng Vanya Dimitrova Ivanova
PhD thesis title: IoT based Botnet Attacks Discovery by Network Traffic Analysis
Member of the scientific jury: Assoc. Prof. Dr. Eng. Ivo Rumenov Draganov / according to order № ОЖ-5.3-09 from 31.03.2022 of the Rector of TU-Sofia / scientific direction 5.3 Communication and computer technics / specialty „Television and video technics"

1. **Relevance of the problem developed in the dissertation in scientific and applied terms**

The topic of the dissertation "Detection of IoT-based botnet attacks through network traffic analysis" is extremely important for the practice of detecting distributed network attacks in order to bring information services out of accessibility. In 2020, more than 10 million botnet attacks were registered, with the number of known C&C servers exceeding 15,000 just a year earlier. The geographical distribution of the latter is extremely wide - in the United States, the Netherlands, Germany, France, China, Russia and other countries. Affected industries worth developing billions of dollars in manufacturing and trade include computer games, gambling, financial services, online shopping, Internet delivery and other telecommunications services. Financial losses in some cases for large corporate representatives exceed $ 2 million. The variety of ways to organize botnet attacks, including centralized and distributed architectures, make it very difficult to identify initiators and require universal ways to register and analyze them. The aim of the dissertation is aimed at the most sustainable detection of such attacks as it is achieved by analyzing the network traffic exchanged between attacking machines and affected ones.

2. **Degree of knowledge of the state of the problem and creative interpretation of the literary material**

The good knowledge of the state of the problem by the doctoral student is indicated by the use of 207 literature sources published in renowned scientific journals worldwide. All of them are in English and about three quarters of them are from the last 5 years. In the analysis of the state of the problem, the doctoral student made an attempt to present a unified taxonomy of widely used IoT devices, many of which can be used as bots in building a botnet. The distribution and the influence of the latter from a technical and economic point of view have been studied. The types of network attacks carried out via IoT botnet are considered. The architectures for the organization of such networks are classified and their principle of operation is described. The main types of botnet attacks are presented, in particular DDoS attacks and attention is paid to the life cycle of the botnet. Vulnerabilities of IoT devices occupy a special place in the doctoral student's examinations, in respect of which there is a list of possible actions to limit their exploitation. The action of 19 of the most active botnets from the last ten years has been analyzed and measures for timely detection of their action have been identified. Conclusions are made, on the basis of which the purpose and tasks of the dissertation are defined. All this testifies to the in-depth knowledge of the doctoral student in the field in which the subsequent results obtained within the dissertation fall.

3. **Correspondence of the chosen research methodology and the set goal and tasks of the dissertation with the achieved contributions**

According to the set goal and objectives of the dissertation, the doctoral student has focused his efforts on finding effective ways to classify aggregate characteristics of network traffic exchanged between attacking and attacked machines, as well as periods of normal operation of supported services. Methods and tools typical for machine learning are applied as an analysis of the informativeness of numerically represented features using parameters such as information gain and Gini coefficient and their subsequent ranking. Histograms of the distribution of the values representing the characteristics were constructed after appropriate normalization, similarity was sought between them based on the Pearson and Spearman coefficients. The informative significance of the individual characteristics was assessed through FreeViz design, as well as by determining the probabilities of occurrence of the studied types of attacks at a given value of the attribute. The method of Principal Component Analysis on the individual characteristics is also applied, reducing their dimensionality from 10 to 8 components. Subsequent training of a neural network of feedforward type, a support vector machine and an ensemble of randomly generated trees, as well as a combination of them, was performed with a full set of 10-component features and a reduced set of 8-component ones. The accuracy of classification on test sets not used in training was assessed. Evaluation is made using parameters such as Precision, Recall, Specificity, F1-measure and others. For each of the tested classifiers an optimal configuration was found and the most accurate was that of a connected two-component classifier. The full set of mathematical treatments that the doctoral student has applied shows her ability to use the tools of machine learning in order to obtain effective implementations of algorithms for classifying network traffic in the absence and presence of botnet attacks.

4. **Scientific and / or scientifically applied contributions to the dissertation**

The doctoral student has stated 2 scientific contributions, expressed in the development of theoretical models of a combined detector and a combined multiclass classifier of 10 types of botnet attacks, working on aggregate records of network traffic characteristics. Four scientific and applied contributions related to the development of optimized models of single-layer and multilayer neural network with backpropagation of the error, a support vector machine and an ensemble of decision trees to detect and classify as separate modules of the same by type 10 botnet attacks. There are also 5 applied contributions related to the technical implementation of the classifiers and the processing of data from their work. Based on the presented experimental results in the dissertation, as well as the mathematical descriptions of the individual classifiers, the claims made by the doctoral student for contributions could be accepted as correct.

5. **Evaluation of the dissertation publications: number, nature of the publications in which they are printed**

Seven scientific publications are presented with the dissertation. All of them are in English and in one of them the doctoral student is the only author. There are two papers from an international scientific conference held in Bulgaria. There are five articles from international scientific journals. No citations to the publications presented in this way are known. The main results presented in the dissertation are given in the main part in all 7 publications on the dissertation research. The way of presenting the material in them corresponds to the required scientific level and is evidence of the ability of the doctoral student to present in summary form the results obtained by her.

6. Opinions, recommendations and remarks

My recommendation to the PhD student is to focus her future work on the implementation of the results obtained in the practice of protecting public and private networks providing wide access services against unauthorized influences, which are based on distributed attacks using a botnet.

7. Conclusion with a clear positive or negative assessment of the dissertation

In conclusion, I believe that the dissertation of MEng Vanya Dimitrova Ivanova, entitled "IoT based Botnet Attacks Discovery by Network Traffic Analysis" meets the requirements for obtaining the educational and scientific degree "Doctor" in the scientific specialty "Automated systems for information processing and control" from direction 5.3 Communication and computer technics and therefore I strongly suggest to the Distinguished Members of the Scientific Jury to support its award.

Date: 16.05.2022

Member of the jury:

(Assoc. Prof. Dr. Ivo Draganov)