**Response to the Handling Editor and Reviewers**

**Submission No.:** JNCA-D-16-00617R1

**Title:** Model Order Selection and Eigen Similarity based Framework for Detection and Identification of Network Attacks

**Journal of Network and Computer Applications (JNCA)**

**Research Paper**

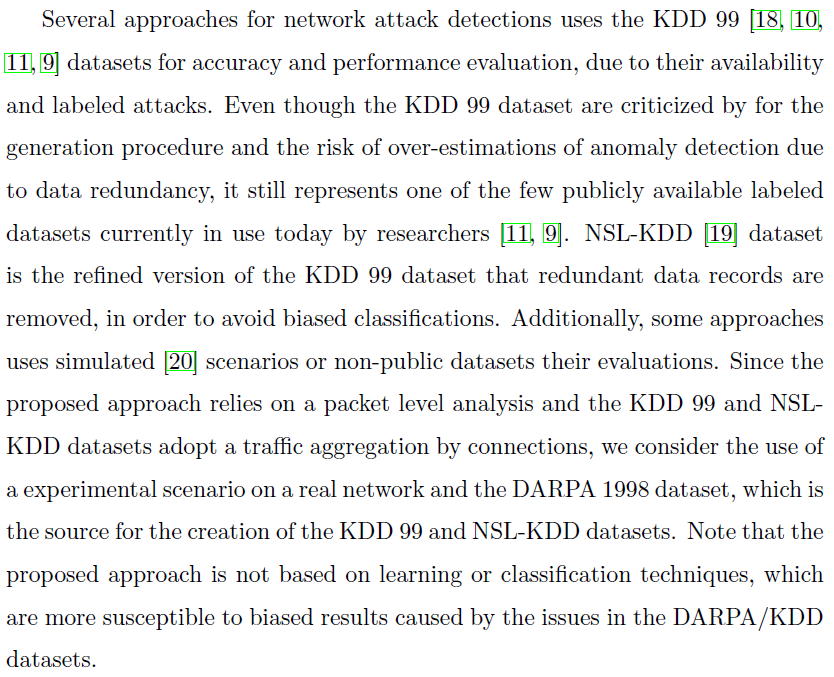
Dear Mohammed Atiquzzaman, Ph.D., Editor-in-Chief, and reviewers,

The authors would like to thank the Associate Editor-in-Chief and the Reviewers for volunteering their time in reviewing our manuscript and providing us with valuable comments which allowed us to significantly improve the paper. We have carefully revised the manuscript based on the reviewer’s comments and suggestions. All changes in the paper have been marked with a color. The following is our point-by-point responses to the raised comments.

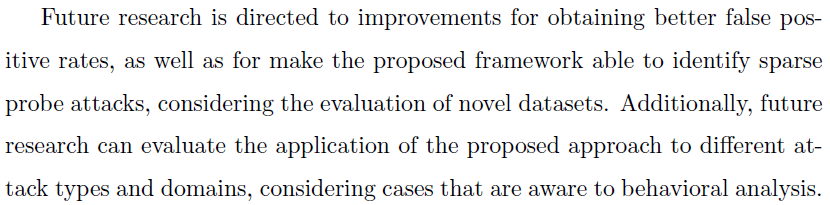
Kind regards,

The authors.

1. **EDITOR**
   1. **COMMENT: *The referees have suggested that certain aspects of the paper should be re-examined, before it can be considered further for publication.***
      1. **REPLY:** We thank the Editor for the time, work, and positive evaluation of our paper.
2. **REVIEWER #3**
   1. **COMMENT: *The reviewer appreciates the efforts that the authors have made to revise this paper. The revised version has improved a lot compared to the previous version. Most of the questions raised from the reviewers have been clarified. The reviewer has three additional comments regarding the revised version.***
      1. **REPLY:** We thank this reviewer for the stimulating comments and for appreciating the contribution of our paper, as well as for the suggestions.
   2. **COMMENT: *The mentioned dataset used for additional evaluation is almost twenty years ago, and it might be able to reflect the current state of arts of modern attacks.***
      1. **REPLY:** We are thankful for your observation. The proposed work focus on flooding and probe attack detection through a packet level analysis. Therefore, we evaluated the main datasets publicly available, according to [9, 10], regarding the attack types, attack identification (label) and data granularity necessary to evaluate our proposal. The DARPA 98 dataset still represents one of the few publicly available labeled datasets currently in use today by researchers [11, 9]. Since the proposed approach relies on a packet level analysis and the KDD 99 and NSL-KDD datasets adopt a trace aggregation by connections, we consider the use of a experimental scenario on a real network and the DARPA 1998 dataset. Additionally, we propose the evaluation of novel datasets as future work.

**

(...)

**

* 1. **COMMENT: *The authors have provided complexity and running time evaluation in the revision. However, given that the EVD complexity in O(M^2\*N+N^3), it looks like the proposed approach might be DoS'ed if an attacker attempts to flood using randomized port numbers, e.g., (M=) 64K network ports. This is because performing bandwidth consumption flooding attacks does not need to work with a dedicated (or a monitored) port number.***
     1. **REPLY:** Thank you very much for the valuable observation. The proposed framework focus and is evaluated for selected ports of interest of an internal network, therefore, it is possible to reduce the number of monitored ports and to reduce the impact of bandwidth consumption flooding attacks. The data collection and parsing may be more vulnerable to high bandwidth attack. However, the proposed approach relies on the data aggregation by port and time, which can be impacted by large amount of traffic, but less than the data collection and parsing. The proposed approach can also be applied for higher number of ports, considering distributed or parallel approaches. We added this possibility as possible future work.
  2. **COMMENT: *The authors have revised the paper to focus exactly on probe and scanning attacks. This may greatly limited the contribution of the proposed work, as we already have a number of commercial solutions to deal with the problem.***
     1. **REPLY:** The proposed work focus on flooding and probe attack detection through a packet level analysis based on signal processing concepts and schemes. We revised the paper to clarify that the proposed approach focus on DoS attacks that attempts to block access to system or network resources consuming resources through large amount of TCP or UDP traffic, not considering exploiting of system vulnerabilities or lower level traffic. The main contributions of the proposed framework are the capability to blindly detect time frames under network attack via MOS and eigen analysis, and the detailed identification of the network attack via eigen similarity analysis, without training. Taking into account this information, as well as the provided literature review and comparison to the current work, we tried to highlight the contribution of this work in relation to the state-of-the-art.

1. **REVIEWER #4**
   1. **COMMENT: *The authors have satisfactorily addressed the queries raised by the reviewers. However, only DARPA 1998 dataset has been considered to conduct the experiments. The authors can also conduct the experiments on other publicly available datasets such as CAIDA (https://www.caida.org/data/) and UNB-ISCX (http://www.unb.ca/research/iscx/dataset/iscx-IDS-dataset.html). Such datasets contain more recent attack vectors and types. Conducting experiments on such datasets and discussion of analysis on such datasets can provide in depth understanding of the proposed work.***
      1. **REPLY:** We thank this reviewer for appreciating our efforts and for the suggestions. During the first review we evaluated the main publicly available dataset regarding network attack detection, according to [9, 10], including the Caida and UNB-ISCX datasets, but we find some limitations on these datasets for the evaluation of the proposed approach and scenario. The Caida dataset contains only attack traffic to the victim and responses to the attack from the victim, non-attack traffic has as much as possible been removed. Therefore, the traffic is composed by malicious traffic, without legitimate and noise traffic, what differs of real network traffic and differs of the proposed data modeling as signal superposition of legitimate traffic, noise and malicious traffic. The Caida dataset contains DoS attacks, but to the best of our knowledge, it does not contains port scan attacks. The UNB-ISCX is publicly available for researchers, but requires a dataset request to be provided access to the content. Therefore, due to time restrictions for response to reviewers, we have decided focus on DARPA dataset and propose to evaluate other datasets in future works.
2. **REVIEWER #5**
   1. **COMMENT: *Since only one user (computer) was used, the amount of traffic noise (DHCP) would be very limited.***
      1. **REPLY:** We are thankful for your observation. Considering the first review, we have added the evaluation of our proposal applied to the DARPA 98 dataset, which contains mixed traffic that can reproduce more realistic legitimate and noise traffic.
   2. **COMMENT: *There is no description of the systems used, because attacks such as the fraggle attack are easily eliminated by both systems and routers. Plus, the network is very limited with the number of devices used and type of traffic observed.***

## **REPLY:** We are thankful for your helpful observation. Considering the experimental scenario, we used the hping for reproducing the fraggle attack, nmap for port scan and metsploit for synflood attack. Considering the number of devices and traffic, we extended our evaluation to the DARPA 98 dataset, which includes the evaluation considering external and internal machines.

* 1. **COMMENT: *The authors stated """The UDP scan technique sends UDP packets to the destination port, and if it responds with a ICMP port 230 unreachable message, then it indicates that the scanned port is closed. On the other hand, if a message is not received, then the port is considered as open."""\* lots of systems today disable this by default, that is why it is unclear which systems were used for testing.***

## **REPLY:** We thank the reviewer for this constructive comment. We extended the evaluation of our proposal to the DARPA 98 dataset, which implements different methods of port scan.

* 1. **COMMENT: *What if the port scan was not sent in a sequence fashion and sent with time frames higher than the samples selected for evaluation and checking? The information considered in the proposed framework will not help.***

## **REPLY:** We are thankful for this important observation. The evaluation of our approach applied to the DARPA 98 dataset showed that cases with a delay between port scanning was not detected by the proposed approach, according to discussion presented in section 5.6. Addicionally, we proposed to consider the solution of this limitation as future work.

## 

## 

* 1. **COMMENT: *Why didn't the authors use flow analysis or even compare to such related work. Flow analysis has been used a lot for such detection, especially when you're looking at time and ports.***

## **REPLY:** We thank the reviewer for this constructive comment. Our proposal aims to rely on packet count by port and time, without efforts for flow identification, deep packet inspection (DPI) or state analysis. We extended the evaluation of our proposal to compare to a signal processing approach based on network flow [2].

## 

## (...)

## 

* 1. **COMMENT: *How can such module detect new novel attacks, when the only experiments tested were SYN(flood), fraggle (rare today), and two port scanning techniques? ex: traffic like DHCP are considered noise, so how would this framework detect an ARP Scan?***

## **REPLY:** Thank you very much for your comment. The proposed work focus on the detection of flood and probe attacks through techniques that do not require previous knowledge about the environment or training for classification of normal or abnormal behaviors. In the literature, signal processing techniques have been applied to attack detection due to their capability to blindly detect anomalies, since these anomalies are previously unknown. In our proposal, we focus on flood and probe attacks, regarding the behaviors that characterize these types of attacks. However, our approach can be evaluated regarding different attack types and domains, considering cases that are aware to behavioral analysis, according to our proposal for future works.

## 

* 1. **COMMENT: *The Eigen Similarity Analysis for Detection of Ports Under Port Scan Attack shows that port 80, 443, and 53 were not an attack, even though all of these traffic are currently used for exfiltration or covert communication. So, how will the proposed framework detect such attacks? Detecting such attacks could be considered novel if they could achieve such detection.***

## **REPLY:** We are thankful for this constructive comment and suggestion. The proposed work focus on flooding and probe attack detection through a packet level analysis. Therefore, the proposed approach was not evaluated to detect exfiltration or covert communications that does not present flooding or scanning behaviors. The DARPA 98 dataset includes DoS attacks based on exploitation and flooding, but we focus our evaluation on cases that produces flooding or probe behaviors. Additionally, we consider the detection of exfiltration or covert communication as future works.

## 