Combining Recurrence Quantification Analysis and Adaptive Clustering to Detect DDoS Attacks

Marcelo Antonio Righi
Cyber Defense Command
QGEx / SMU - Brazilian Army
Brasilia, DF, Brazil
righi@cdciber.eb.mil.br

Raul Ceretta Nunes
Applied Computing Department
Federal University of Santa Maria
Santa Maria, RS, Brazil
ceretta@inf.ufsm.br

***Abstract*—The high number of Distributed Denial of Service (DDoS) attacks executed against a lot of nations has demanded innovative solutions to guarantee reliability and availability of internet services in the cyberspace. In this sense, different methods have been used to analyze network traffic for denial of service attacks, such as statistical analysis, data mining, machine learning and others. However, few of them explore hidden recurrence patterns in nonlinear network traffic and none of them explores it together with the Adaptive Clustering. This work proposes a new method, called *DDoSbyRQA*, which uses the Recurrence Quantification Analysis (RQA) based on the extraction of network traffic dynamic features and the combination with an Adaptive Clustering Algorithm (A-Kmeans) to detect DDoS attacks. The experiments were made by using the CAIDA and UCLA databases and it has demonstrated the ability of the method to increase the accuracy of DDoS detection and to real-time applicability.**

Keywords— DDoS, RQA, Adaptive Clustering, A-Kmeans.

I. INTRODUCTION

Keeping the internet services is critical during conflicts and crises when nations need to be able to send information and to be increasingly resilient to the challenges that cyber conflict can provide. The ability to use and deploy Intrusion Detection Systems (IDS) to the networks can be crucial for enabling communication, especially in a hostile environment. The DDoS can stop the ability to maintain communication between interested actors. Military or civilian areas can be impaired and to lose the freedom to continue fighting in the cyberspace, putting at risk the security of a region or country.

Detection of DDoS can be an excellent solution to recover the security of a nation's cyberspace. For this to happen, it necessarily has mechanisms that can indicate in real time a possible attack and enable actions to be taken in a timely manner, because only in this way the success of the mitigation process may be satisfactory.

To detect DDoS attacks, different techniques are used. From [1], the detection techniques could be aggregated in at least four relevant methods: statistical-based, data mining-based, knowledge-based and machine learning-based. However, as noted in [2], many of them still have limitations and their quality of service can be affected due to an excessive number of false alerts. The existence of traffic with nonlinear dynamic behavior instead of just stationary behavior can be one of these limiting factors [3]. The network traffic contains properties as self-similarities [4], long-range dependence [5] and recurrence [3].

The Recurrence Quantification Analysis (RQA) [6] is a mathematic technique that allows analyzing the behavior of a nonlinear signal that repeats itself over a specific period. In the network security field, RQA already has been applied in other works [3, 7, 8]. However, in the current paper, we have changed the way that it is applied. To provide better results on DDoS attack detection this paper explores RQA to extract knowledge from dynamic features of the network traffic in a combination with an adaptive clustering method. The Adaptive Clustering method (A-Kmeans) [9], which automatically calculates the number of clusters rather than a fixed amount of these, is combined with RQA, which extracts dynamic features of a set of network flow attributes selected to effective express DDoS behavior [10].

Using RQA it is possible to extract various dynamic features of specific behaviors for each system called Recurrence Quantification Measures (RQM). Examples of RQMs are Recurrence Ratio (RR), Determinism (DET), Entropy (ENTR), Trend (TREND), Laminarity (LAM), among others. Developing the RQA over these RQMs allows us to obtain an analysis focused on the dynamic features of the traffic rather than an analysis focused on, for example, traffic statistical variability.

This work proposes the *DDoSbyRQA*, a new method for DDoS attack detection that combines the RQA based on extracting dynamic features (RQMs) with an adaptive clustering to classify DDoS network traffic (TCP Flood, UDP Flood, and ICMP Flood). Applied on CAIDA and UCLA databases, the *DDoSbyRQA* demonstrates the power of this combination. As result, we get a more accurate method when comparing to similar methods. In this way, the main contributions of this work are: (1) to demonstrate that the use of RQA can be applied on DDoS detection not only to analyze adopted network flow attributes but the dynamic features of them; (2) to demonstrate that an adaptive clustering method (A-Kmeans), which automatically calculates the number of clusters, can be a good partner of the RQA to increase the efficiency of DDoS detection method; and (3) to demonstrate that the method can be used in real time to take an effective action during DDoS attacks.

The remainder of this paper is organized as it follows: Section II presents related works and Section III presents a theoretical review of the RQA. Section IV presents details of the implementation of the proposed *DDoSbyRQA* method and Section V presents our experiments and results. Finally, Section VI presents the work conclusions.

II. RELATED WORKS

The traditional idea to characterize and detect DDoS attacks is to do attribute extraction based on network traffic behavior and construct an analysis of their behavior. For example, in [11] the authors propose a method to detect DDoS attacks using a classifier based on a decision tree algorithm (C 4.5). The authors use sixteen attributes to describe a normal network traffic pattern behavior. However, the rate of false positives is incremented when network

traffic increases [11], denoting a less effective method in situations in which there is increased flow on normal network traffic. Also, the choice of network traffic attributes did not consider important features for DDoS, since the chosen attributes do not contemplate the variance of the packets size and variance of the time arrival packets (time among received packets). These variances tend to zero during a DDoS attack [10].

In [12], the authors present a method for detection of DDoS attacks that explores different classifiers - the Apriori algorithm, FCM and K-Means clustering-, demonstrating that the combination of multiple classifiers can improve the accuracy of detection. From these works, it is easy to comprehend that the performance of a detector depends on extracted attributes and the chosen classifier. Different from the others, our work explores these factors when applying RQA combined with a self-adapter classifier (A-Kmeans) on a set of attributes of network traffic that could effectively characterize a DDoS attack.

The RQA was used in other works [3, 7, 8]. In [3] the authors demonstrate that RQA can be applied to offer qualitative and quantitative observations on detecting anomalous events in complex traffic (nonlinear). They suggested that the network traffic exposes itself to the omnipresent properties of self-similarity and long-range dependence, which are correlations in a wide range of time scales. In [7] the authors focus on demonstrating the visual analysis of Recurrence Quantification Measures (RQM) in Recurrence Plots (RP) and their power on detecting anomalies. Visual tools like web recurrence plot (http://www.recurrence-plot.tk/glance.php) and graphical API of the Weka data mining tool were used to conclude if changes visually indicate a DDoS attack or not. In [8] the authors extend the work performed in [7] to demonstrate that RQA can be applied to detect DDoS on VoIP networks but they maintain the empirical analysis based on visual tools of Recurrence Plots (RP). They did not consider the need for alert generation automatically and in real time. Differently from the above works, our approach looks to attributes and to a method that automatically analyzes the dynamic features (RQMs) over Recurrence Plots (RP). In addition, we also explore the use of Adaptive Clustering (A-Kmeans) in combination with RQA.

In [10], it is presented a method that characterizes DDoS attacks from seven attributes extracted directly from the network traffic. According to the authors, from these attributes, a classifier can effectively distinguish this kind of attacks. They have used the K-NN algorithm [13], which is a similarity-based supervised learning algorithm that makes the classification based on the nearest neighbor rule. The choice of *k* neighbor is fixed and determined by the researcher. However, the use of a classifier to operate directly on the attribute time series can be a significant limiter for obtaining a good efficiency of the DDoS detection method. In addition, manually setting the algorithm number of neighbors is a challenge and a limiter. In [14] the authors performed a combination of Wavelet Transform (WT) and RQA and clustering algorithm to classify the traffic. The used clustering method was the K-Means clustering, which has a predefined fixed number of clusters, and the use of wavelet preprocessing is a time consuming phase. Differently, adopting the set of attributes proposed in [10] our work explores the combination of RQA and Adaptive Clustering (A-Kmeans [9]) showing that it does not need a fixed number of clusters and get better results than non-adaptive one.

III. RECURRENCE QUANTIFICATION ANALYSIS (RQA)

RQA corresponds to the construction of the Recurrence Plots (RP), a visual graph of recurrence quantification of a given Time Series (TS), and its interpretation. The RP (see example in Fig. 1) was proposed in [15] as a technique of nonlinear dynamic analysis systems and provides a behavior visualization of the space trajectory of multidimensional phases [8]. In practice, RP is a two-dimensional square array that represents the evolution of dynamic system states and that is populated by black and white dots. The black dots indicate recurrence, namely the states of the dynamical system for these orbiting points in regions near each other in the trajectory of the phase space. Such regions are called the Recurrence Areas. A black dot marked at the coordinate (*i*, *j*) of the RP represents the recurrence of system states at time *i* and *j* [16, 6]. In other words, considering the RPs of Fig. 1, generated in the testing phase of this work, each state of the Average Packet Size (**AVG_PAC_SIZE**) in each moment (*i*) is compared with all other states in each moment (*j*, *j + 1*, ..., *n*). In case of recurrence, a black dot will be marked from each result of each comparison; otherwise, it will be a white dot. Now (*i + 1*) its state will again be compared with all other states (*j*, *j + 1*, ..., *n*) and so on until the end of the time series for each used attribute. The result is a square matrix of black and white dots that indicates the recurrence of the interesting attribute.

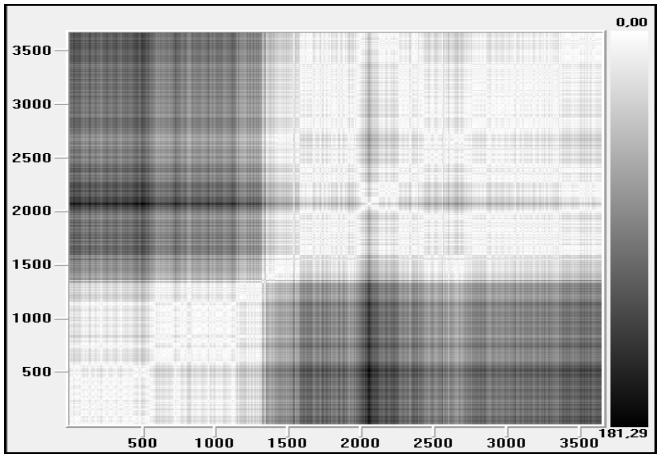

Fig. 1. RP of the Average Packet Size time series in one normal traffic. The axes correspond to the number of traffic system states considered in RQA, i.e., the RP demonstrates the recurrence over *N* states of the TS.

Given a network traffic time series $X\{x_i\}$, where $i = 1, 2,..., n$ [16, 17], the traffic system states can be expressed by X_j (see Equation (1)). In (1), *m* is the embedded dimension (represents how many delays are used in relation to the initial time series) and τ is the duration of the delay (time to wait between states). Note that *n* is the total number of samples in *X* and *N* is the number of states.

$$X_j = [x_j, x_{j+\tau}, ..., x_{j+(m+1)\tau}], j = 1, 2, ..., N \quad (1)$$

After collecting the traffic from pre-defined attributes, the RP is built to each one according to (2).

$$R_{ij} = \theta(\varepsilon - \| x_i - x_j \|) \quad (2)$$

R_{ij} corresponds to an element of the recurrence matrix (RP), where ε is the adopted threshold called Recurrence Radius, x_i and x_j are the states of the system in the *m-dimensional* phase space under analysis, N is the number of states considered and θ is the decision function defined in (3). According to (3), if the distance between the states x_i and x_j is smaller than the threshold ε, then the value of $\theta(\varepsilon)$ is 1 and there is a black dot in position (i, j) of RP; otherwise, the value of $\theta(\varepsilon)$ is 0 and there is a white dot (i, j) in RP.

$$\theta(f(\varepsilon)) = \begin{cases} 0(\varepsilon - \|x_i - x_j\| \leq 0 \\ 1(\varepsilon - \|x_i - x_j\| > 0 \end{cases} \quad (3)$$

It's to highlight that the ε is an important parameter in the RQA. This radius corresponds to a threshold that defines the recurrent points on the RP and it depends on each type of system that is being analyzed and their objectives [16]. The literature does not provide a specific method to establish the ideal Recurrence Radius to define recurrence points, taking it to be adjusted according to the type of application.

Despite RP allows visual analysis of recurrence, this type of analysis is human-based and can lead to different interpretations. Thus, to obtain more precision to the analysis, Recurrence Quantification Measures (RQM) [16] can quantify the behavior structures in the RP. RQMs can be computed and analyzed by algorithms. From [16], the main RQMs are Recurrence Ratio (RR), Determinism (DET), Average Length of the Diagonal Lines (L), Maximum Length of the Diagonal Lines (Lmax), Shannon Entropy (ENTR), Trend (TREND), Laminarity (LAM), Average Length of Vertical Structures (TT) and Maximum Length of Vertical Structures (Vmax).

The RQA can be applied in the analysis of short non-stationary series. However, compared to other techniques of nonlinear dynamic analysis, one of the main advantages offered by RQA is to enable the analysis of anomalies in non-stationary system minimizing the bias in the analysis when occurring overloads in parameters of the sampling system.

IV. THE *DDoSbyRQA* METHOD

This section presents the *DDoSbyRQA* anomaly detection method. It can distinguish between network traffic due to DDoS attacks vs. benign traffic. Fig. 2 shows the architecture of the detection solution, where the Attack Detection Module highlights the main functionalities of the proposed method.

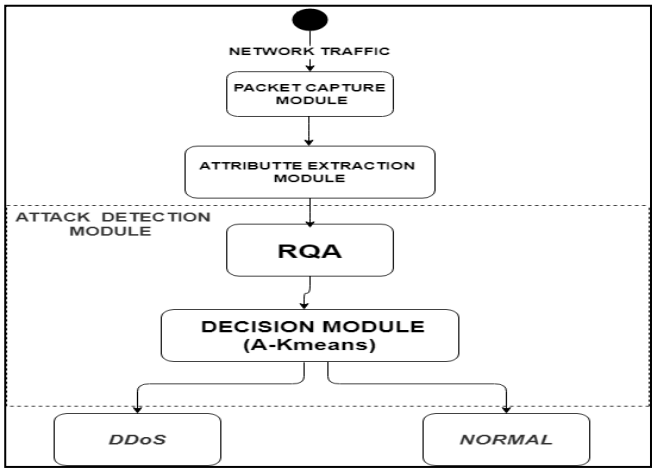

Fig. 2. The architecture of the attack detection by *DDoSbyRQA* method.

In general, the *DDoSbyRQA* is supported by a Packet Capture Module, which collects data on the network, and by an Attribute Extraction Module, which selects desired attributes for RQA. The Attack Detection Module encapsulates the method that combines the RQA and Adaptive clustering (A-Kmeans) to detect DDoS attacks. The Subsections A, B and C detail each module of the architecture and Subsection D presents the algorithm that implements the *DDoSbyRQA* method.

A. Packet Capture Module

The packet capture module is a module that corresponds to a network sniffer. It selects the inbound traffic to a network under analysis by *DDoSbyRQA*. After captured, the data is sent to the Attribute Extraction Module.

B. Attribute Extraction Module

The extraction of attributes corresponds to the phase of selection network attributes that potentially provide relevant information to the problem of interest (DDoS detection).

For detection of DDoS attacks, the RQA application requires attributes that characterize the anomalies of interest in a time series. From [10], it is known that seven attributes are enough to identify DDoS attacks. They are illustrated in Table I.

The function of the Attribute Extraction Module is to extract these seven attributes from network traffic and send them to the Attack Detection Module. The result value of each attribute corresponds to statistical values extracted from network traffic flow at each second, as described in Table I. At each sixty seconds, a new time series is formed and sent to the detection module. Thus, the output of this module is seven time series, one for each attribute described in Table I, at each minute.

TABLE I. ATTRIBUTES USED BY RQA. ADAPTED FROM [10]

Attributes used by RQA	
Attributes	***Description***
N_PAC	Number of packets
N_BYTES	Number of bytes
AVG_PAC_SIZE	The average packets size
VAR_T_PAC	The variance of the time arrival packets
VAR_S_PAC	The variance of the packets size
R_PAC	Total packets rate
R_BYTES	Total bytes rate

C. Attack Detection Module

The Attack Detection Module is the central module of the proposed solution (see Fig. 2). It is composed of (i) the RQA Module and by (ii) Decision Module centered in an adaptive clustering classifier.

1) RQA Module: It is important to highlight that in the RQA module the method also extracts dynamic features (RQMs) of the network traffic (for example, Entropy), which aim to enable the recurrence analysis through RPs.

This module is responsible by the RQA and RQMs computing and analyzing on the RPs. Each attribute received through Attribute Extraction Module is represented in RQA Module by a time series (60 seconds) modeled by samples held in equidistant periods. Every time series, one for each attribute that expresses DDoS attacks or normal traffic (Table I), result on Recurrence Plots with its RQMs extracted mathematically. From each time series, one RP is

built, as defined in Section III. After the formation of the RP, three dynamic features are extracted: Recurrence Ratio (RR), Entropy (ENTR) and Determinism (DET). These features correspond to RQMs used in *DDoSbyRQA* for DDoS detection. Our goal is to analyze anomaly occurrences over these RQMs and not over network traffic statistical attributes.

To extract the dynamic features from each network attribute, the quantification calculations (calculation of RR, DET and ENTR) applied to the RP in *DDoSbyRQA* are made as follows.

a) Recurrence Ratio (RR): Measures the density of recurrence points on the RP. See (4) for *RR* computation.

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j} \tag{4}$$

b) Determinism (DET): The ratio between the number of recurrence points that makes the diagonal structures and all points of recurrence.

$$DET = \frac{\sum_{l=l\min}^{N} lP(l)}{\sum_{i,j=1}^{N} R_{i,j}} \tag{5}$$

In (5), *P(l)* is the number of recurrence points for each diagonal formed and *l* is the RP diagonal length.

c) Shannon Entropy (ENTR): Represents the frequency distribution of the lengths of the diagonal lines.

$$ENT = \sum_{l=l\min}^{N} p(l)\log_2 p(l) \tag{6}$$

$$p(l)\frac{P(l)}{\sum_{l=l\min}^{N} P(l)} \tag{7}$$

Through these 21 dynamic features (3 for each of the 7 attributes), the RQA Module forms a set of data to express through the recurrence properties the network behavior. This set is then forwarded to the Decision Module to be clustered and classified.

2) Decision Module: the Decision Module has the function of classifying the set of dynamic features received from RQA Module. The data is first partitioned by similarity (clusters) and then classified as a DDoS attack (anomalous) or not (no anomalous).

In order to avoid the difficulty of defining the optimal number of clusters, the *DDoSbyRQA* method applies the A-Kmeans algorithm [9]. This algorithm works on a set of 21 RQMs derived from the values of Entropy, Determinism and Recurrence Ratio of seven network attribute (see Table I). The A-Kmeans automatically calculates the number of clusters (value of "k" is automatic) and compares each of them with pre-set thresholds during training phase with the normal traces databases.

The decision of the module is then centered on the calculation of centroids (central points of each cluster) of the set of dynamic features (RQMs) received from the RQA Module. If the majority of the formed clusters are classified as anomalous, then the traffic will be classified as a DDoS attack.

In short, the Decision Module is also enhanced with an adaptive clustering method to provide more flexibility in the calculation of the number of clusters used to classify the traffic. The A-Kmeans does it automatically. The automatic calculation improves the minimization of accuracy errors of the classifier. For example, in K-means [14] method the researcher determines the number of clusters.

D. DDoSbyRQA Algorithm

The following steps detail the algorithm that implements the *DDoSbyRQA* method.

Entry: time series traffic (seven attributes).

Output: an indication of DDoS attack or normal traffic.

Step 1: for each traffic series *X* (one for each of the seven attributes), to calculate the dynamic features (Recurrence Ratio, Entropy and Determinism) as described on subsection IV.C. This process is illustrated in (8), (9) and (10).

$$F_1 = f(X_{N\_PAC}) \quad F_2 = f(X_{N\_BYTES}) \quad F_3 = f(X_{AVG\_PAC}) \quad F_4 = f(X_{VAR\_T\_PAC}) \tag{8}$$

$$F_5 = f(X_{VAR\_S\_PAC}) \quad F_6 = f(X_{R\_PAC}) \quad F_7 = f(X_{R\_BYTES}) \tag{9}$$

$$F_n = \{RR_n, ENT_n, DET_n\}, \quad n = 1,2,3,4,5,6,7 \tag{10}$$

Step 2: to group the 21 dynamic features (from Step 1) to describe the dynamic patterns of network traffic behavior synthesized in *F* in (11).

$$F = \{[RR_n, ENT_n, DET_n]\} \tag{11}$$

Step 3: from A-Kmeans algorithm, it builds groups of dynamic characteristics in *F* within different clusters and classifies the traffic behavior as a DDoS attack (the majority of clusters anomalous) or Normal.

V. EXPERIMENTS AND RESULTS

This section presents the experiments setup (subsection A), the tests and results of *DDoSbyRQA* (subsection B), including the False Positive (FP) rates comparison with other methods (subsection C), and finally it demonstrates the performance tests (subsection D).

A. Experiments Setup

The experiments setup is twofold organized: first the choice of the dataset and then the setup of *DDoSbyRQA* operational parameters.

Doing real experiments with the DDoS attack is a challenge and the performance on laboratory needs good databases. Some authors [17, 18] have used databases like CAIDA 2008 [19] and CAIDA 2007 [20] to characterize normal traffic and DDoS attack traffic. In addition, UCLA CSD database [21] is well known and contains interesting datasets with and without attacks. The CAIDA 2007 database contains 1 hour of DDoS attacks (ICMP Flood and TCP Flood) divided into files of type "*pcap*" sanitized with 5 minutes each. The CAIDA 2008 database contains 16 hours of traffic without attack divided into files of type "*pcap*" sanitized with 1 hour each. The data was collected for sixteen days on the network in Chicago and San Jose, in the United States. The database UCLA CSD contains traces of 1 hour of DDoS attacks (UDP Flood) and traffic traces without attacks collected on ten different days. Assuming these databases contain workloads to test the *DDoSbyRQA*, the experiments in this paper used these three databases.

From these datasets, seven attributes were extracted as described in Table I, resulting in a time series *X* for each

attribute of interest. Thus, the experiments were arranged in two phases, one for training and other for tests. In the training phase was used the normal traffic (without DDoS attacks) and in the testing phase was used the anomalous traffic (with DDoS attacks). In the training phase, the goal of the experiment was to calibrate the threshold values of the *DDoSbyRQA* method. To a correct method operation, it was important to identify the behavior of each dynamic features (RQMs) in traces with and without attacks. To characterize the normal traffic, the experiments in this phase used 62 minutes of traces from CAIDA 2008 database and 152 minutes of traces from UCLA CSD database. All these traces without attacks. To characterize anomalous traffic, were used only data sets with traces containing DDoS attacks, one with 66 minutes from CAIDA 2007 database and other with 56 minutes from UCLA CSD database.

The *DDoSbyRQA* method was setup to work with time series corresponding to a sample of 60 seconds, containing a network traffic attribute to each one. Thus, without loss of generality we chose to set the duration of delay (τ) to one second and the embedded dimension (m) to 60. According to the experiments performed in [14, 15], in this work the RPs was generated with Recurrence Radius (ε) set to rate of 10%. Of course, these parameters of RQA could differ, but to demonstrate the power of the method we decided to fix the threshold ε (RQA most influencer parameter) on a value already used on similar works. The parameters τ and m have less influence on RQA [14] and thus our choice followed the chosen TS structure.

B. *Testing and Results*

The first test step was to evaluate the significance of the adopted MQRs. We highlighted the chosen MQRs derived from [14], a previous work on network anomaly detection with RQA. To be significant, a MQR must present different behavior to normal (training) and abnormal (testing) traces. Fig. 3 illustrates the results of the training phase for the dynamic features RR of the AVG_PAC_SIZE (one of the seven selected attributes). The analysis of dynamic features of other attributes follows the same methodology and its demonstration was removed to eliminate redundancy. In Fig. 3, the RR for the training dataset is shown to be stationary, with RR level around 25% (line 2). For the testing dataset, containing only traces with attacks, the stationary behavior still remains observable, but the level of RR was increased (line 1) to almost twice the observed in series without attack. These results demonstrate the feasibility of thresholds adoption to distinguish between normal traffic and DDoS attacks using the dynamic features (RQMs).

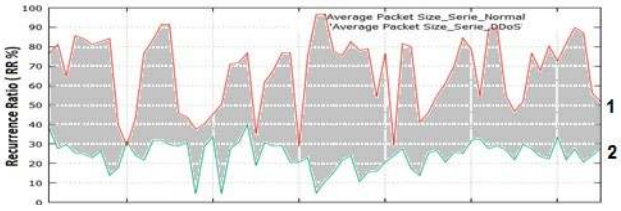

Fig. 3. Recurrence Ratio for Average Packet Size (AVG_PAC_SIZE).

The second test step was to evaluate the accuracy of the DDoSbyRQA. Table II (TCP flood / ICMP flood) and Table III (UDP flood) present the results of the testing phase. The experiment evaluated the proportion of True Positives (TP), False Positives (FP) and the resulting Accuracy (AC), with the accuracy defined as follows.

For purposes of comparison were conducted tests with (i) K-Means algorithms, (ii) RQA + K-Means, (iii) A-Kmeans, and (iv) RQA + A-Kmeans. The latter corresponds to the *DDoSbyRQA* method. The goal of these tests was to allow the evaluation of the impact of the RQA and Adaptive Clustering inclusion. These tests also considered two datasets: a dataset merging of databases CAIDA 2007 and CAIDA 2008 (see results in Table II) and other merging datasets from databases UCLA CSD Normal and UCLA CSD with DDoS (see results in Table III).

When comparing the results, in both cases (containing attacks and without attacks), showed in Tables II and III, it was possible to observe an improvement of the efficiency of classifiers when applied in conjunction with RQA. The True Positives (TP) when the RQA is associated with K-Means classifier improved more than 13% (13.88% for CAIDA dataset and 18.15% for UCLA CSD dataset) and more than 19% when associated with the A-Kmeans (20.03% for CAIDA dataset and 19.69% for UCLA CSD dataset). According to values in Tables III and IV, the reduction of the False Positives (FP) was also significant. As a result, with both datasets, there was an increase in the accuracy of classifiers when in conjunction with RQA, reaching an improvement of 10.54% for A-Kmeans on CAIDA dataset. The tests also demonstrated the association of RQA and A-Kmeans provided a more effective result when compared with RQA + K-Means. This result demonstrates the effectiveness of adaptive clustering proposed by *DDoSbyRQA* method. With CAIDA dataset the accuracy was improved by 12.42% and with UCLA CSD dataset the accuracy was improved by 8.62%, demonstrating better results in DDoS detection.

TABLE II. RESULTS FOR CAIDA 2007/2008 (TCP/ICMP FLOOD)

ALGORITHM	AC (%)	TP (%)	FP (%)
K-Means	70,96	69,23	28,33
RQA+K-Means	85,99	83,08	13,54
A-Kmeans	85,96	75,35	12,31
RQA+A-Kmeans	**98,41**	**95,38**	**1,54**

TABLE III. RESULTS FOR UCLA CSD NORMAL / DDOS (UDP FLOOD)

ALGORITHM	AC (%)	TP (%)	FP (%)
K-Means	84,34	60,63	11,25
RQA+K-Means	88,26	78,78	10,48
A-Kmeans	94,23	74,24	4,54
RQA+A-Kmeans	**96,88**	**93,93**	**3,03**

C. *Comparison with other methods*

The Table IV demonstrates that the False Positive (FP) rates of other similar DDoS detection methods are higher than the *DDoSbyRQA* method. It can be seen that the *DDoSbyRQA* method has an excellent performance when compared with others. Our method gets 1.54% of FP (see Table II) and the best other get 2.40% (see Table IV).

TABLE IV. FP RATES TO DDOS DETECTION METHOD CITED

REFERENCES	METHOD	FP
[11]	C 4.5 (Decision Tree)	**2,40**
[12]	Apriori+ FCM + K-Means	**2,45**
[13]	KNN	**8,11**
[14]	RQA+TW+ K-Means	**8,91**
[17]	Centroid-Based Rules	**3,23**

D. Performance test

The performance test of the *DDoSbyRQA* was executed on an Intel® Core ™ i7 4510U CPU 2.60GHz with 8 cores and 8GB of memory. The operating system was the Debian GNU / Linux 7 with kernel 3.2.0-4-amd64. The compiler used was the GNU C Compiler, version 4.7.2-5. The presented execution times represent the average of 20 executions.

The experiments measure three algorithm times. The time spent (*i*) to extract network traffic statistical attributes from data collected of a 60 seconds traffic window; (*ii*) to compute the Recurrence Plot graph and its Recurrence Quantification Measures; and (*iii*) to make a decision with the adaptive classifier. Table V shows the results of the performance test. The results demonstrate that the *DDoSbyRQA* can decide in less than one second. This performance result enables the proposed method to be applied in real time applications that operate over network traffic statistics collected with time windows higher than one second.

TABLE V. PERFORMANCE TEST RESULTS OF THE *DDOSBYRQA*.

DDoSbyRQA step	***Average Execution Time (ms)***
Extraction of network traffic attributes	**285**
Computation of RP and its RQMs	**324**
Adaptive clustering and decision	**325**
Total	**934**

VI. FINAL CONSIDERATIONS

In this paper, we discussed how DDoS detection on computer network could overcome many of the limitations and security challenges posed to cyberspace during conflicts and crisis that are exploited for adversary nations. To avoid important damages in the communication system of any country this paper presented an effective way to detect DDoS attacks in order to react accurately and quickly.

The effectiveness of anomaly-based DDoS detection methods has been a challenge for designers of detection algorithms. Thus, the use of RQA combined with A-Kmeans technique came as a new option to improve the quality of service of these algorithms. Until now in the context of detecting anomalies in network traffic, the RQA has been explored with limitations. This work has contributed evaluating it in conjunction with a small and known group of network traffic attributes and an Adaptive Clustering algorithm (A-Kmeans).

This work showed that from only seven network traffic attributes, which characterize DDoS, it is possible to extract relevant dynamic features (RQMs) that allows increasing the accuracy of DDoS detection. This method also aimed the anomaly detection with RQMs, making possible to overcome the negative influence of variability in traffic attributes that could lead to erroneous detections. We highlight that it is possible because the RQA looks for recurrence domain instead of a traffic domain.

The work experiments have shown that the use of the RQA increases the accuracy in identifying DDoS attacks mainly by two facts. First, the method classifies dynamic features of recurrence instead of traffic attributes (the tests evaluated classifiers with and without the RQA). The benefit, in this case, was an increment of up to 10.54% in the accuracy of detection. It is important to note that this result is associated with a significant increase in true positives and decrease in false positives. Second, without sudden variations in traffic, the method allows observing changes in behavioral patterns of recurrence that assist the classifiers to correctly generate clusters. With normal abrupt changes (not caused by DDoS attacks) the method allows observing the regularity of recurrence behavior.

The work also demonstrated that the use of A-Kmeans algorithm, an adaptive clustering algorithm that automatically calculates the number of clusters, fits well with DDoS detection based on RQA and that it improves the accuracy when combined with RQA. The improvement in detection accuracy was by 8.62% when compared with a non-adaptive cluster algorithm (K-Means). The worst performance of K-Means clustering reflects the difficulty of calibrating a non-adaptive cluster, which can be observed by the variability of accuracy when explored with two databases of different characteristics.

Not only effective for DDoS detection, the proposed *DDoSbyRQA* method can also be explored in other contexts of network behavioral analysis and other types of cybernetic attacks, mainly by its characteristic to enable the analysis in the domain of recurrence while minimizing the negative influence of variability that causes deviations in the analysis of traditional traffic statistics.

REFERENCES

[1] M. Gyanchandani, J. L. Rana, and R. N. Yadav, "Taxonomy of Anomaly Based Intrusion Detection System: A Review," In: International Journal of Scientific and Research Publications, v.2, n.12, 2012.

[2] A. S. Raut, and K. R. Singh, "Anomaly Based Intrusion Detection-A Review," Int. J. on Network Security, vol. 5, 2014.

[3] F. Palmieri, and U. Fiore, "Network anomaly detection through nonlinear analysis," Computers & Security, 29(7), pp. 737–755, 2010.

[4] W. Willinger, V. Paxson, and M. S. Taqqu, "Self-similarity and heavy tail: structural modeling of network traffic," A Practical Guide to Heavy Tails: Statistical Techniques and Applications, ISBN: 0-8176-3951-9, pp. 27-53, BirkhRäuser, Boston, USA, 1998.

[5] M. Grossglauser, and J. C. Bolot, "On the relevance of long-range dependence in network traffic," IEEE/M Transactions on Networking, 7(5): pp. 629-640, 1999.

[6] C. L. Webber, and N. Marwan, "Recurrence Quantification Analysis: Theory and Best Practices," Springer series: Understanding Complex Systems. Springer International Publishing, Cham Switzerland, 2015.

[7] N. Jeyanthi, J. Vinithra, S. Sneha, R. Thandeeswaran, and N.C.S.N. Iyengar, "A Recurrence Quantification Analytical Approach to Detect DDoS Attacks," In: Computational Intelligence and Communication Networks (CICN), Washington, DC, USA, pp. 58-62, 2011.

[8] N. Jeyanthi, R. Thandeeswaran, and J. Vinithra, "RQA based approach to detect and prevent DDoS attacks in VoIP networks," In: Cybernetics and Information Technologies. v.14, n.1, pp. 11-24, 2014.

[9] S. K. Bhatia, "Adaptive K-Means Clustering. American Association for Artificial Intelligence," Copyright. Palo Alto, California 94303 USA. Copyright, 2004.

[10] T. T. Oo, and T. Phyu, "A Statistical Approach to Classify and Identify DDoS Attacks using UCLA Dataset," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 2, Issue 5, 2013.

[11] Y. C. Wu, H. R. Tseng, W. Yang, and R. H. Jan, "DDoS detection and traceback with decision tree and grey relational analysis," International Journal of Ad Hoc and Ubiquitous Computing, 7, pp. 121–136, 2011.

[12] R. Zhong, and G. Yue, "DDoS detection system based on data mining," Proceedings of the 2nd International Symposium on Networking and Network Security, Jinggangshan, China, 2-4 April, pp. 062–065. Academy Publisher, 2010.

[13] H. Nguyen, and Y. Choi, “Proactive Detection of DDoS Attacks Utilizing k-NN Classifier in an Anti-DDoS Framework,” International Journal of Electrical and Electronics Engineering, vol. 4, nº 4, 2010.

[14] J. Yuan, R. Yuan, and X. Chen, “Network Anomaly Detection based on Multi-scale Dynamic Characteristics of Traffic,” INT J COMPUT COMMUN, ISSN 1841-9836, 9(1), pp. 101-112, 2014.

[15] J. P. Eckmann, S. O. Kamphorst and D. Ruelle, “Recurrence plots of dynamical systems. Europhys,” Lett, 56 (5), pp. 973-977, 1987.

[16] N. Marwan, and C.L. Webber Jr, “Mathematical and computational foundations of recurrence quantifications,” In: Recurrence Quantification Analysis: Theory and Best Practices. Springer Series: Understanding Complex Systems. Springer International Publishing, Cham, Switzerland, pp. 1-41, 2015.

[17] W. Bhaya, and M.E. Manaa, “The Proactive DDoS Attack Detection Approach Using Data Mining Cluster Analysis,” Journal of Next Generation Information Technology (JNIT), vol. 5, no. 4, 2014.

[18] M. Suresh, and R. Anitha, ”Evaluating Machine Learning Algorithms for Detecting DDoS Attacks,” In 4th international Conference on Advances in Network Security and Applications (CNSA), pp. 441-452, 2011.

[19] “The CAIDA UCSD Anonymized Internet Traces 2008,” Access in 05 may 2015 11:12h, https://data.caida.org/datasets/passive-2008/

[20] “The CAIDA "DDoS Attack 2007" Dataset,” Access in 15 may 2015 11:12h, https://data.caida.org/datasets/security/ddos-20070804/

[21] “UCLA CSD packet traces,” http://www.lasr.cs.ucla.edu/ddos/traces/public/usc