IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

A Novel Approach to Detect Intruder for Hierarchical WSN Network

Manoj Kumar Gupta[1], Lokesh Singh[2]

M. Tech Scholar, Computer Science & Engineering, TIT, Bhopal, India [1]

Assistant Professor, Computer Science & Engineering, TIT, Bhopal, India [2]

Abstract: As of late, with wide utilization of PC frameworks, web, and fast development of PC systems, the issue of interruption discovery in system security has turned into a critical issue of concern. In such manner, different interruption recognition frameworks have been produced for utilizing abuse recognition and inconsistency discovery strategies. These frameworks attempt to make strides identification rates of variety in assault sorts and diminish the false positive rate. In this paper, another interruption discovery strategy has been presented utilizing Min Max K-means clustering algorithm, which defeats the lack of affectability to starting focuses in K-means algorithm, and expands the nature of clustering. The investigates the NSL-KDD information set demonstrate that the proposed strategy is more effective than that in view of K-means clustering algorithm. Additionally, the strategy has higher discovery rate and lower false positive recognition rate.

Index Terms: K-means algorithm, NSL-KDD, clustering algorithm.

I. INTRODUCTION

Lately, the security of PC systems has turned out to be progressively vital and global need in view of the wide utilization of PCs, fast development of PC systems, and the development of electronic trade. These days, PC systems have turned into the objective of interruptions and malignant gets to. The spread of web assaults has expanded significantly [3]. These security issues make a test of painstakingly distinguishing assaults and interruptions from typical practices of real clients

The gatecrashers misuse vulnerabilities of PC frameworks, while in fact building of a framework with no defenselessness is verging on unthinkable. Moreover, a few aggressors may utilize obscure examples of interruption, and it makes the issue more troublesome. Along these lines, interruption identification frameworks have turned into a fundamental and discriminating part of PC hosts what's more, PC systems [2]. Now that it's out in the open, notwithstanding static protecting systems, for example, firewalls and access controls, interruption recognition frameworks implement the security approaches and systems to fulfill the prerequisites of PC security.

Two noteworthy classifications of interruption identification philosophies are Misuse Detection and Anomaly Detection. Abuse recognition recognizes interruption in view of known examples while Anomaly location concentrates on obscure examples. Indeed, Anomaly recognition characterizes the ordinary conduct examples of framework, with the presumption that an interruption will for the most part incorporate some deviation from typical conduct. In the mean time, there are different scientific classifications of interruption location frameworks and related procedures.

Since interruption identification framework gathers data from distinctive information sources like sorts of bundles, hosts, convention subtle elements, and so forth the information examination is exceptionally entangled and hard.

On the other hand, the information mining procedures have prepared to do recognizing examples in expansive measure of information and extricating helpful data from extensive databases. Consequently, by utilizing of information mining strategies, interruption recognition frameworks can create a typical conduct model [6]. Thus, assault examples and typical information examples are adequately isolated. In this respect, different information mining strategies, for example, Affiliation Rules, Classification, and Clustering are connected in interruption identification frameworks.

Interruption location techniques taking into account arrangement information mining models are frequently ineffectual in identifying obscure interruption designs in light of the fact that order is a directed learning system and it can't deal with unlabeled information. Thusly, clustering is utilized to order adequately unlabeled information and distinguish obscure interruptions. Joining clustering procedures with Anomaly–based interruption recognition frameworks is a vital issue of exploration for discovering new and obscure interruption design.

K-means, a normal bunching calculation, has been demonstrated for applying to the interruption recognition frameworks. K-means grouping calculation puts the comparative information from distinctive information sources in the same bunch and the divergent ones in diverse bunch, then marks the bunches as either typical or an assault to assess the anomaly of information. In any case, K-means grouping calculation is normally chosen arbitrarily starting bunch focuses, which influence the last bunching results. Actually, K-means grouping calculation

relies on upon initial positions of the bunch focuses and effectively falls in neighbourhood essentials.

II. RELATED WORK

In [1], early Internet (identified with the lovely plan and development of structures, and so on.) outline objectives did not put security as a high need. Be that as it may, today Internet security is a rapidly developing concern. The quantity of Internet assaults has expanded fundamentally [6] , yet the test of recognizing such assaults by and large falls on the end has and administration suppliers, requiring/requesting framework directors to distinguish and piece assaults all alone. Particularly, as interpersonal organizations have ended up focal center points of data and correspondence, they are more the objective of consideration and assaults. This makes a test of deliberately recognizing underhanded and coldblooded associations from typical ones. Past work [3] [7] has demonstrated that for an assortment of Internet assaults, there is a little subset of association estimations that are great markers of whether an association is a piece of an assault or not. In this paper we take a gander at the adequacy of utilizing two diverse co-bunching arrangements of PC guidelines to both gathering associations and additionally check which association estimations are solid pointers of what makes any given gathering unusual and startling with respect to the aggregate information set. We run tries different things with these co-grouping arrangements of PC directions on the KDD 1999 Cup information set. In our trials we find that delicate co-bunching, running on tests of information [2], discovers steady cutoff points/rules that are solid markers of irregular and startling identifications and makes gathers that are profoundly unadulterated. At the point when running hard co-bunching on the full information set (more than 100 runs), we by and large have one gathering with 92.44% assault associations and the other with 75.84% typical associations. These outcomes are equivalent to the KDD 1999 Cup winning section, demonstrating that co-bunching is an in number, unsupervised technique for isolating ordinary associations from odd and sudden ones. At last, we trust that the thoughts introduced in this work may motivate research for (unusual, startling thing) identification in informal communities, for example, recognizing spammers and culprits who trick individuals.

Presently/as of late, tested by malevolent and barbarous utilization of system and (deliberately) assaults on PC framework, intrusion discovery framework has turned into a crucial and infrastructural (machine/technique/route) for securing basic helpful thing/profitable supply and data. Most present intrusion identification frameworks concentrate on mix of two things/gas-electric vehicle administered and unsupervised machine learning advances. The related work has exhibited that they can get prevalent execution than applying single machine learning arrangement of PC directions in identification model. Furthermore, with the nearby consideration of related works[8] [11], highlight selecting and speaking to methods for doing things are additionally critical in quest for high (squandering next to no while working or creating something) and adequacy [9]. Execution of indicated assault sort recognition ought to likewise be enhanced and (made sense of the value, sum, or nature of). In this paper, we join data pick up (IG) technique for selecting more discriminative elements and triangle region based bolster vector machine (TASVM) by consolidating k-means bunching arrangement of PC guidelines and SVM classifier to identify assaults. Our framework (finishes or picks up with exertion) (nature of being near reality or genuine number) of 99.83%, identification rate of 99.88% and false alert rate of 2.99% on the 10% of KDD CUP 1999 (procedure of making sense of the value, sum, or nature of something) information set [5]. We likewise (fulfill or pick up with exertion) a superior recognition execution for particular assault sorts concerning (high) quality and review.

Taking into account the considered hybrid operation and change operation in (identified with minor concoction gathering guidelines within living things) set of PC directions, this paper enhances molecule swarm streamlining arrangement of PC directions.

The enhanced molecule swarm advancement set of PC directions is utilized to enhance (however much as could reasonably be expected) punishment limit/rule c and piece capacity limits/rules g of SVM and the highly enhanced model named new-PSO-SVM is set up. KDD Cup 99 intrusion identification information set is utilized to do test. The outcomes demonstrate that PSO advancement enhances the order (nature of being near reality or genuine number) rate of SVM

III. PROPOSED SYSTEM

Another interruption identification technique is proposed utilizing Min Max K-means grouping calculation. The test on NSL-KDD information set demonstrates that the proposed interruption recognition strategy can improve the Detection Rate and decreases the False Positive Rate. K-implies calculation begins with arbitrarily picking of the beginning focuses of bunches and every data point is appointed to its closest focus. At that point, the ideal focus of every group is figured by essentially averaging the focuses. Min Max K-means is another bunching calculation that tries to take care of the K-implies introduction issue. The calculation begins with arbitrarily picking of the starting focuses of bunches, then endeavors to apply minimizing of the greatest inside change of groups as opposed to minimizing the total of inner fluctuation of bunches in K-implies calculation. Utilizing k-implies a viable approach to figure out dark and worm gaps assaults can likewise be recognized in groups.

- **Topology Module**

This section contains description of functionality of the scripts used in building topology. This module involves building Wireless Network topology, topology consisting of mobile nodes, each node working with multiple channels.

DOI 10.17148/IJARCCE.2016.5219

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

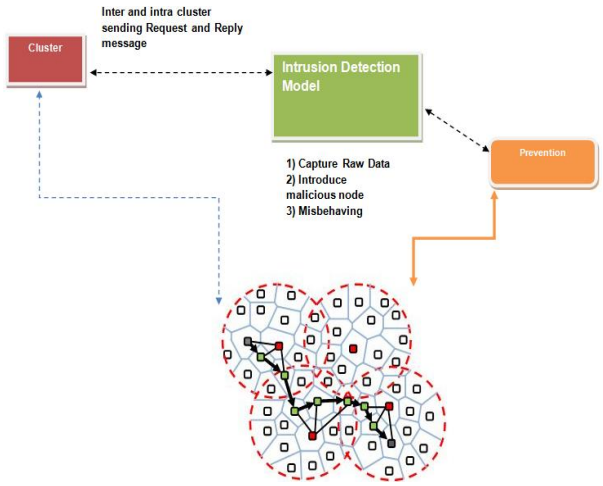

Fig1: Architecture Diagram

Algorithm for Detection of Malicious Node

INPUT: A RREQ packet to node
OUTPUT: Detection Status of Node for all control packets to this node do
if the packet is neither from nor to this node itself then
if request is duplicate RREQ then
isduplicateRREQ =TRUE
end if
Step 1 :if isDuplicateRREQ =TRUE AND Reward Timer is pending then
message "Not a New Request" and skip all the next steps
else
Drop Counter = Drop Counter + 1
end if
Step 2 Set the timers
Sense Timer = CURRENT TIME
Reward Timer = CURRENT TIME
Step 3 Start the sense timer such that
Sense Timer = CURRENT TIME + SENSE TIME
Drop Counter = Drop Counter + 1
Calculate Time To Send for this packet
if Time To Send > Sense Timer then
Drop Counter = Drop Counter + 1
else
Start the reward timer such that
Reward Timer = CURRENT TIME + REWARD TIME
Drop Counter = Drop Counter - 1
end if
end if
if Drop Counter increases then
"Mark the Node as Malicious" and stop

- **K-means Model**

K-means algorithm starts with randomly choosing of the initial centers of clusters and each input point is assigned to its nearest center. Then, the optimum center of each cluster is computed by simply averaging the points. MinMax K-means is a new clustering algorithm that tries to solve the K-means initialization problem.

The algorithm starts with randomly choosing of the initial centers of clusters, then attempts to apply minimizing of the maximum internal variance of clusters instead of minimizing the sum of internal variance of clusters in K-means algorithm. Each cluster is weighted so that higher weights are assigned to the cluster with larger internal variance. By applying this method, the results become less dependent to initialization and quality of clustering increases even when the initial centers of clusters are not selected optimally.

- **Attack Model**

The attackers' goal is to discover the traffic patterns among mobile nodes.

- The adversaries are passive signal detectors, i.e., they are not actively involved in the communications. They can monitor every single packet transmitted through the network.
- The adversary nodes are connected through an additional channel which is different from the one used by the target MANET. Therefore, the communication between adversaries will not influence the MANET communication.

IV. CONCEPT AND RESULT DISCUSSION

Concept Used

K-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.

The main idea is to define k centres, one for each cluster. These centres should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centre.

Since k means is only concerned in the centre of the cluster...the disadvantage of this is...if the intruder node is not in the centre of the cluster. It is difficult to find the malicious node. Science the malicious node can be anywhere in the cluster as in existing system

In our proposed system, we have proposed and black hole attack in the clusters, .in which we are not concerned with the centre node in cluster. If the malicious node is present anywhere in cluster we can easily identify the intruder node, since it won't forward the route request to the destination node. Thus creating a LOOP for packet drop

Result and Analysis

Experiment shows that our proposed system can save energy as shown in Fig 2. Our proposed architecture provides energy efficient system that proves the efficiency of the overall system.

DOI 10.17148/IJARCCE.2016.5219

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

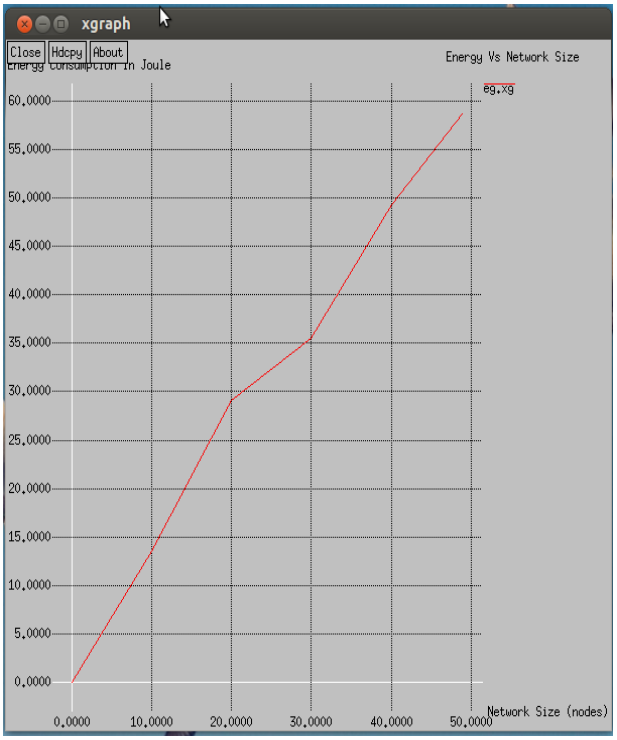

Fig 2: Energy

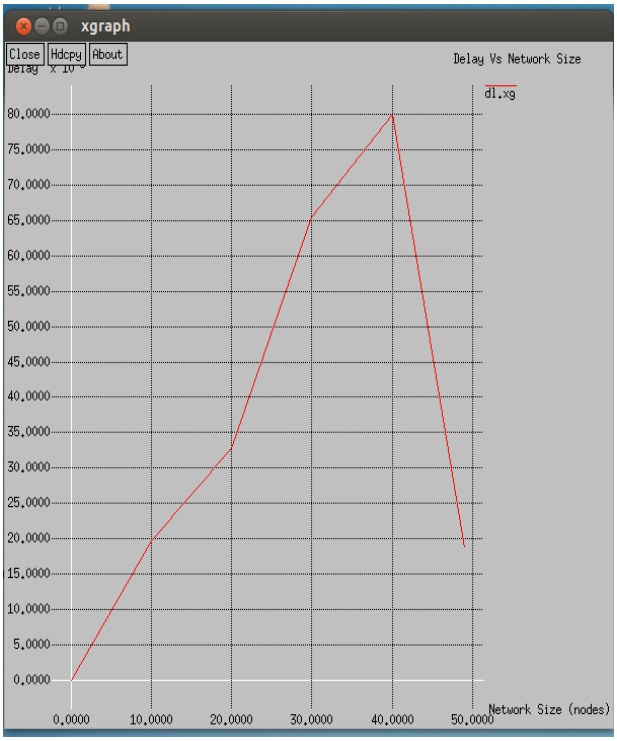

Fig 3: Delay

End to End Delay: As shown in fig 3, we can see the end to end delay is reduced as the number of nodes increases in the network.

The overall overhead in the network is very essential as the performance degrades as the number of nodes in the network increases. As per the fig 4, we can conclude that the overhead is minimal in our proposed scheme.

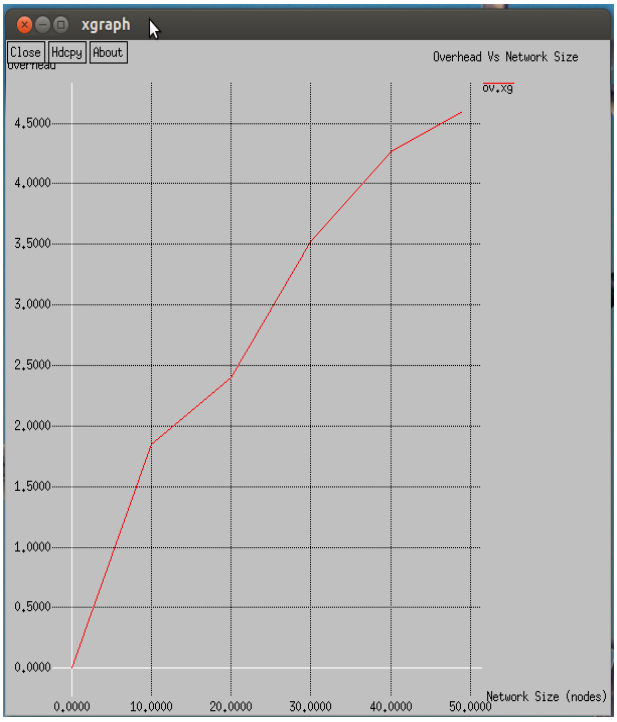

Fig 4: Overhead

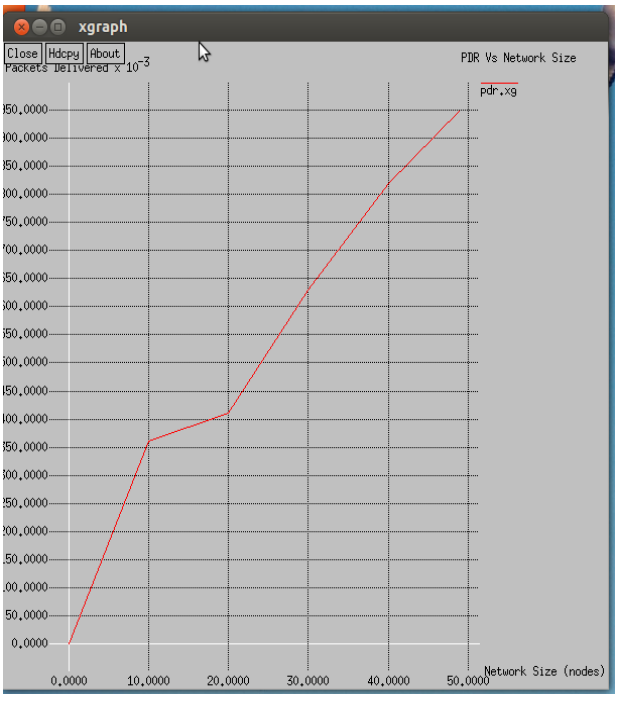

Fig 5: Packet Delivered

Packet Delivery Ratio: The packet delivery ratio is an important parameter in the network that any network cannot be compromised. The packet delivery ratio does not degrades as the network size increases as in fig 5.

Throughput: It can be defined as ratio of total number of packets received by the destination from the source to the time it takes for the destination to receive the last packet. It is clear from the fig 6 that the throughput of the network is high.

DOI 10.17148/IJARCCE.2016.5219

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

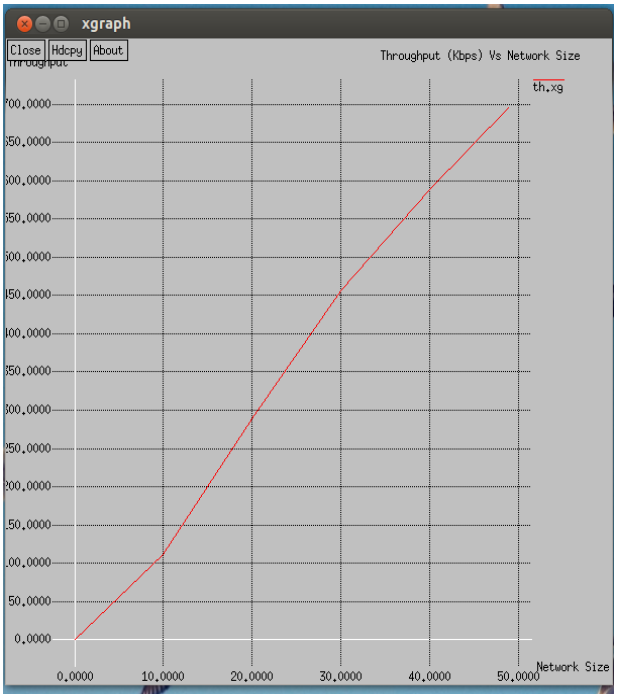

Fig: Throughput

V. CONCLUSION

In this paper, another interruption discovery technique is proposed taking into account a Min Max K-means bunching calculation which conquers the lack of affectability to introductory focuses in K-implies calculation, and builds the nature of bunching. The proposed technique is contrasted and the interruption discovery technique in view of K-means calculation. The test on the NSL-KDD information set demonstrates that the proposed technique is more compelling than that in view of K-means grouping calculation. Likewise, demonstrates the system has higher identification rate and lower false positive discovery rate. In as much as, the quantity of obscure interruption is expanding in certifiable system environment, later on, we will concentrate on doing the examination on the best way to expand the identification rate for identifying obscure assaults or new assaults adequately.

REFERENCES

[1] E. E. Papalexakis, A. Beutel, and P. Steenkiste, "Network anomaly detection using co-clustering," in Proc. 2012 Int. Conf. Advances in Social Networks Anal. and Mining (ASONAM 2012), 2012, pp. 403-410.

[2] P.Tang, R.-a. Jiang, and M. Zhao, "Feature selection and design of intrusion detection system based on K-means and triangle area support vector machine," in Proc. 2010 2nd Int. Con. Future Networks (ICFN '10), 2010, pp. 144-148.

[3] S. Wang, "Research of Intrusion Detection Based on an Improved Kmeans Algorithm," in Proc. 2nd Int. Conf. Innovations in Bio-inspired Computing and Applicat. (IBICA), 2011, pp. 274-276.

[4] J. Luo and S. M. Bridges, "Mining fuzzy association rules and fuzzy frequency episodes for intrusion detection," Int. J. Intell. Syst., vol. 15, pp. 687-703, 2000.

[5] K. Wankhade, S. Patka, and R. Thool, "An Overview of Intrusion Detection Based on Data Mining Techniques," in Proc. 2013 Int. Conf. Commun. Syst. and Network Technologies, 2013, pp. 626-629.

[6] H.-J. Liao, C.-H. Richard Lin, Y.-C. Lin, and K.-Y. Tung, "Intrusion detection system: A comprehensive review," J. Network and Comput. Applicat., vol. 36, pp. 16-24, 2013.

[7] D. K. Denatious and A. John, "Survey on data mining techniques to enhance intrusion detection," in Proc. Int. Conf. Comput. Commun. And Inform. (ICCCI - 2012), 2012, pp. 1-5.

[8] Y. Li, Z. Li, and R. Wang, "Intrusion detection algorithm based on semisupervised learning," in Proc. Int. Conf. Inform. Technology, Comput. Eng. and Manage. Sci. (ICM), 2011, pp. 153-156.

[9] Y.-F. Zhang, Z.-Y. Xiong, and X.-Q. Wang, "Distributed intrusion detection based on clustering," in Proc. 4th Int. Conf. Mach. Learning and Cybern., 2005, pp. 2379-2383.

[10] L.-B. Qiao, B.-F. Zhang, R.-Y. Zhao, and J.-S. Su, "Online Mining of Attack Models in IDS Alerts from Network Backbone by a Two-Stage Clustering Method," in Cyberspace Safety and Security, Switzerland: Springer Int. Publishing, 2013, pp. 104-116.

[11] Y.-q. XU, B. ZHANG, and X.-t. QIN, "Clustering intrusion detection model based on grey fuzzy K-mean clustering," J. Chongqing Normal University: Natural Sci., vol. 1, pp. 81-83, 2013.

[12] Z. Li, Y. Li, and L. Xu, "Anomaly intrusion detection method based on K-means clustering algorithm with particle swarm optimization," in Proc. 2011 Int. Conf. Inform. Technology, Comput. Eng. and Manage. Sci. (ICM), 2011, pp. 157-161.

[13] G. Tzortzis and A. Likas, "The MinMax K-Means clustering algorithm," Pattern Recognition, vol. 47, pp. 2505-2516, 2014.

[14] M. R. Ackermann, M. Märtens, C. Raupach, K. Swierkot, C. Lammersen, and C. Sohler, "StreamKM++: A clustering algorithm for data streams," J. Exp. Algorithmics (JEA), vol. 17, p. 2.4, 2012.

[15] B. Bahmani, B. Moseley, A. Vattani, R. Kumar, and S. Vassilvitskii, "Scalable K-means++," Proc. VLDB Endow., vol. 5, pp. 622-633, 2012.

[16] M. Tavallaee, E. Bagheri, W. Lu, and A.-A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in Proc. 2nd IEEE Symp. Computational Intell. for Security and Defence Applicat., 2009.