www.ijaceeonline.com ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, October 2017, pp. 7–14

Intelligent Soft Computing Techniques for Providing Network Security: A Survey

Riyaz B
Research Scholar, School of Computing Science and Engineering,
VIT University- Chennai Campus, Chennai, India
Email: briyaaz@gmail.com

Ganapathy S
Assistant Professor (Sr.), School of Computing Science and Engineering,
VIT University- Chennai Campus, Chennai, India
Email: sganapathy@vit.ac.in

Abstract: *Network security is necessary to safeguard the intellectual data from intruders in computer networks during the transmission. Data mining techniques are used to perform the feature selection and classification tasks to find the new patterns, regularities and irregularities in the large data sets. Data summarization, visualization, Clustering, Classification and Feature Selection are the techniques to accomplish the task. In this paper, a survey of effective and intelligent soft computing techniques based feature selection and classification algorithms which are proposed to perform effective classification and intrusion detection. These methods are proposed for effective intrusion detection that uses fuzzy logic, rough sets, fuzzy-rough sets, neural network and neuro-tree classifiers are presented. Moreover, these intelligent data mining techniques include feature selection and classification techniques used to identify and prevent the intrusions on the network. In addition, these techniques are also used to provide high security over the network and strengthen the quality of service.*

Keyword: *Classification; Feature Selection; Fuzzy Rough-sets; Fuzzy sets; Intelligent Intrusion Detection.*

1. INTRODUCTION

Information processing systems [1] used through internet are disposed to many kinds of threats and leads to different kinds of instance in significant losses. The basic intention of the information security is to expand the protective information systems that secure form un-authorized users. Intrusion Detection systems (IDS) was mainly used to identify different kinds of threats on the network and the computer systems. IDSs are categorize into two systems namely misuse and anomaly detection system. Misuse/Signature based detection systems detects the disruption of permissions effectively and Anomaly detection systems establish whether it has deviation from the normal usage pattern or not. In IDSs the intrusions were identified and that can be averted effectively by developing an intrusion prevention system. The Transparent Intrusion Detection system [2] has provided the scalable, proactive and secure system which builds on a valuable system hardware. This system provides the functionality of preventing attacks for the non- distributed system attacks. The security feature of this system attained through its transparency of different nodes within the network by maintaining the full control over the traffic through the system, as the result of making the attacks resistant that target against to the system itself.

Cite this paper:

Riyaz B, Ganapathy S, "Intelligent Soft Computing Techniques for Providing Network Security: A Survey", International Journal of Advances in Computer and Electronics Engineering, Vol. 2, No. 10, pp. 7-14, October 2017.

1.1 Intelligent Intrusion Detection Systems

Intelligent IDSs [1] are the intelligent computer programs located in either the network or a host which inspects the surroundings and that acts more flexible to attain high range of detection accuracy. Intelligent IDSs has the capability of constraint checking and decision making. Intelligent IDSs has been developed by the proposed intelligent preprocessing techniques and the effective classification. Such type of IDSs has delivered good detection rate when compared with other systems.

1.2 Intelligent Preprocessing Techniques

Preprocessing or Feature selection is detecting relevant features and gets rid of irrelevant attributes and has the objective of producing some feature subsets that illustrate the stated problem with a minimum breakdown of performance. Feature selection has different advantages, such as improvement in the performance, data understanding, data reduction, less storage capacity requirements and processing cost was reduced. Feature selection classified into two models:

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, October 2017, pp. 7–14

Wrapper and Filter methods. Wrapper method evaluates the subset of variables and optimizing a predictor. The filter method select features with independent of any predictor and does not consider the relationship between variables.

1.3 Classification

Classification was the data mining function that allocates the item in a collection of targeted classes. Classification predicts the certain outcomes, which based on the given input. For predicting the outcome result, a training set that contains the set of attributes and the outcomes for the respective attribute that called prediction attribute. There are many algorithms that are available for classification. The most relevant work for Intrusion Detection systems are discussed in this section.

2. RELATED WORK

There are many works have been proposed and implemented in this direction by various researchers in the past [1-7][14-19][28-35]. Among them, Al-Jarrah [3] proposed the Botnet Intrusion Detection used for reducing the large scale datasets that has high complexity and exponential growth of computational requirements. For detection of botnet attacks which based on the characteristics of network traffic flow at the transport layer, a novel algorithm was designed. The proposed model performed better to the packet payload inspection that was approached in terms of computational time and immunes to packet encryption. To neglect the redundant and irrelevant features, a novel forward selection ranking technique and Voronoi-based data partitioning in large scale datasets was used. The result shows that, very high detection accuracy was achieved and computational cost was reduced.

2.1 Feature Selection

Attribute selection and tuple selection were used in feature selection algorithm. For attribute selection this algorithm uses some rules and information gain value. For tuple selection the rule based approach was used. Gradually feature removal method: The gradually feature removal (GFR) method [17] gradually removes the less important features in the 41 features of KDD'99 cup dataset. This algorithm selects only 19 features and by using these features best accuracy was obtained.

Feature selection based on Modified mutual information: The feature selection based on Modified Mutual information algorithm (FSMMI) [18] has analyzed the selected features and their relationship with different types of attacks. This algorithm selects some of the features for identifying Probe, U2R, DoS and R2L attacks by determining the mutual information. The FSMMI [19] algorithm increases the importance between the input features and output features so that it decrease the selected features redundancy. This algorithm increases the information with output by selects only one feature at a time. The mutual information was adjusted by neglecting a quantity which was proportional to the average MMI within that selected features. This algorithm has selected some features for DoS, Probe, U2R and R2L which are desirable for classification.

Feature selection based on Conditional Random Field: Conditional Random Field (CRF) [20] presented the layered approach, in which each and every layer was considered as one type of attack. So, the probability value has measured for every relevant feature and for each and every type of attack, different features was selected. Here both domain knowledge and practical significance has used and the feasibility analysis was performed for every feature before selecting it in a particular layer. Since every layer was independent to other layers.

Comparison: Using the KDD cup dataset it removes the repeated data or less important data and removing the next set of features by using k-means clustering and the clusters are predetermined in the gradually feature removal method. Mutual information was obtained for the feature selection and became more flexible in the modified mutual information based selection. In the CRF method, the probability values are measured for the selection of relevant features. From the comparative analysis of these methods, it was perceived that the mutual information and information gain ratio values that produced a better method for the feature selection to perform best attribute selection.

2.2 Wrapper Method Based Genetic Feature Selection

For an each and every iteration of this genetic feature algorithm that produces a decision tree. For n iterations, a series of decision trees were obtained and the best decision tree was used to generate the rules. Best trees were obtained by highest sensitivity and specificity.

Its selects only the most important and providing features of classification was the main advantage of this algorithm. Wrapper method [4] for feature reduction is based on cascading of GA weighted sum and neuro-tree of genetic-X-means. Using this method different features subsets are obtained for neuro-tree or genetic-X-means. Here, 13 relevant features are obtained by GA weighted sum from the 41 features using genetic-X-means to control the selection of features. Using these features, feature selection was reduced the training and testing time and at the same time that produced similar accuracy of that feature set.

Neuro-tree classifier: Sindhu et al. [4] proposed a Neuro-tree classifier for the intrusion detection, genetic based approach selects some features that used

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, October 2017, pp. 7–14

for classification. The goal of the neuro-tree classifier was to removing the redundancy, identification of suitable subset features by using wrapper method and realizing these proposed IDS to achieve high detection accuracy. A new means that preventing the over fitting and the new fitness evaluation structure for sensitivity maximization and specificity are the major contributions of IDS. The neuro-tree classifier has the main advantage that reduces the fast convergence and false alarm rate.

2.3 Filter Method

Bolon et al. [5] proposed the cost based feature selection based on the filter method. This framework added a new parameter called λ, to adjust the cost influence into the evaluation function. Two filter methods are used to test the adequacy of the framework: CFS (Correlation based feature selection) belonged to subset feature selection and mRMR (Minimal Redundancy Maximal Relevance) belonged to ranker feature selection method. Using these methods it allows the user to reduce the cost without compromising the classification error which can be used in real world applications. But need to study other feature selection using the wrapper method to test the framework.

Wang et al. [6] proposed the feature selection algorithm for any large scale hybrid data set which based on the method of decomposition and fusion. Hybrid method combines the filter method and the wrapper method for better feature selection algorithm. Large scale data set means the given parts of records are collected and that forms the small data sets called subset. Decomposition means fragmenting the large number of dataset into a subset family that has the same distribution of that large data set. Fusion means combining all the estimates that obtained from the subset together and producing the final feature subset of the large data set. This hybrid method save the computational time over the large scale data and that produced the feature subset with effective and efficient.

Liu et al. [7] presented the feature selection based on effective distance that used to compute the uniformity between a pair of samples. Author proposed the Sparse representation based algorithm to measure the effective distance by constructed the bi-directional network. Then developed the three novel unsupervised filter method feature selection by using the effective distance, included the Effective Distance based Laplacian Score (EDLS) and two Effective Distance based Sparsity Scores (EDSS-1, EDSS- 2). The main advantage of the sparse representation based algorithm was distance measures can mirror the global and dynamic property of the data. Then unsupervised filter methods for the effective distance can be used in dimension reduction and graph based learning methods. Here the proposed algorithm has the better results compared to several supervised methods. In future the sparse representation based algorithm can be used in dimension reduction and spectral clustering.

Parham Moradi et al. [8] proposed the feature selection method that has been developed by combining ant colony optimization (ACO) and graph clustering (GC) with the searching process. The GCACO has three steps for working: First, a graph was represented by the given problem space, in which every node that indicates the feature and the edges which indicates the similarities between the features. Second, the features were split into different clusters by make use of efficient community detection algorithm. Finally at the last step, a novel ACO based search was proposed for the selection of final feature subset. In this ACO search, an ant was placed on a randomly selected cluster and in each and every step that ant decides to selects the next position of the present cluster or move to the next cluster. This proposed method used for both irrelevant and redundant features because, every ant from the clustered graph attempted for searching the features with less number of similarities and maximizing the dependency on the final target class. This method obtained good classification accuracy and improved in execution time.

Teisseyre [9] presented that Multi-label classification (MLC) has the challenge about reducing the dimensionality of the feature vector. For this a novel method, Feature ranking (FR) was proposed and it consist of two steps. First, need to fit the Ising model by using only labels that was performed by fitting K logistic models. Second, need to test how much addition of single feature that affects the initial network based on score statistics. The importance final feature measure was based on the score statistics averaged values. Now consider the FR procedure which was based on the Ising model using regularized logistic regressions and this procedure incorporates all the features simultaneously, but it has the act of computing expensive more for large number of labels. Markov network was proposed and applied in MLC. Markov network has used the Ising model, done relatively simply by building an initial network that contained only labels that was especially for some moderate number of labels. Then score statistic which was computationally efficient and that was not necessary to refit any initial networks when adding some features. The significance was tested very quickly by the added features which was essential in FR methods. The final ranking was based on the score statistic that by added the features together.

Sasikala et al. [10] proposed the memetic algorithm by combining the Shapely value and Genetic algorithm for multi-class classification through handling different dimensional data. Shapely value Analysis (SVA) was the game theory for localization causal function that addressed the interactions among the

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, October 2017, pp. 7–14

number of group elements in a dataset with the multiple features. A Shapely value Embedded Genetic Algorithm (SVEGA) model was devoted as the feature selection tool for the classification engine. This SVEGA has minimum running time and produced the best classification accuracy rate for all small, medium and huge dimensional dataset with minimum selected features. This algorithm based on specificity and sensitivity which was used by learning algorithm for finding the leading position and efficiency.

Joldzic et al. [2] has presented the scalable and distributed solution for the lower level denial of service detection attacks which was executed by transfering huge amount of data with the aim of interrupting regular network service. Active traffic was balanced among multiple traffic processors to achieve the scalability. Device polling was used for packet processing and make use of flexibility and network programmability. Traffic processors were added elastically into the pool which depends upon the traffic volume. This complete detection system was transparent to the external viewers that improves the reliability of intrusion detection.

Zhang et al.[24] proposed the feature selection for heterogenous data or mixed data. In real word data entities, attributes were different types of values i.e., mixed data. Example in medical diagnosis, age represented in integer value, weight and BP are real value and sex in categorical value. Fuzzy set generate uncertainty because they are imprecise. Fuzzy sets are sets, whose elements have degree of membership function in real value interval[0,1]. This real value precisely describes whether the object possibly belongs to a set or not. Fuzy relations is a fuzzy set defined on tuples (x_1…….x_n) that may have varying degrees of membership within the relation. This membership grade indicates strength of the present relation between elements of the tuple. This author proposed the Fuzzy rough set method which deals with the real valued data. This fuzzy rough sets are based on the fuzzy relations which are defined for different kinds of attributes that to measure similarities between the objects. Distance function was utilized to define the fuzzy relations for real value attributes i.e., smaller distance between objects has high similarity attributes. A conditional entropy was used to select features of mixed data.

Guo et al. [25] presented the two level hybrid approach for intrusion detection that composed of two anomaly detection and one misuse detection. In the first stage,the anomaly detection method builds the detection component with low computing complexity. In second stage, the k nearest neighbor algorithm builds two detection components. That detection component of stage 1 takes part into the development of two detection component of stage 2. Using this proposed methods, the false positive rate has reduced and the intrusion detection accuracy rate was increased.

Li et al. [26] proposed a practical multi-view based false alarm reduction system (MVPSys) for reducing the Network intrusion detection system(NIDS) false positive more effectively. MVPSys prototype has conducted the false alarm system in either offline mode or in real time mode. A Semi-supervised learning algorithm was implemented to automatically exploit the unlabled and labeled data without any human intervention. This system extracts and organized the features automatically into two features sets from an incoming alarm, they are destination feature set and source feature set. Here the target enironment has the features of former and the source environment has the features of latter. This proposed system has achieved 95% stable accuracy by reducing the alse alarm.

Al-yaseen et al. [27] has proposed the Multi-level hybrid intrusion detection system for real time intrusion detection problems in data analytics and classified the network data into abnormal and normal behaviours. This proposed model used the extreme learning method and SVM that improves the detection efficiency of unknown and known attacks. The K-means algorithm was modified to build the high-quality training datatsets for improving the performance of classifiers. This modified K-means algorithm builded some small training datasets from the entire original datasets that reduces the classifiers training time and the intrusion detection performance was improved. The proposed algorithm produced high efficiency of the attack detection and achieved best performance accuracy.

2.4 Classification

H. Liu et al. [11] presented that kNN(k Nearest Neighbors) used for multilabel classification. The number of nearest neighbors, i.e., optimal value of k, in kNN classifier is different for every dataset and assigning also difficult to assigned. Further, kNN has poor performance of prediction and sensitive to noisy data. Author proposed the lazy learning algorithm for the multi-labeled data and that has two strategies. At first Shelly nearest neighbor (SNN) that gets reliable and right neighbor information for prediction. Second was Certainty factor (CF) that used to handle the unbalanced problems and uncertainty of data.

One of the advantage of using SNN method was number of nearest neighbors selected was a variable which determined by data, whereas the k-NN method used fixed k. SNN method was free from the parameter k. The CF measurement was used to representing how accurate and truthful to be in making a decision. The CF is a value in the range -1 to 1, which representing the change of degree in belief. The belief increases if the CF value is positive and belief decreases if the CF value is negative. If the CF value zero

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, October 2017, pp. 7–14

means, there is no change in the belief on hypothesis i.e., the CF was neither a probability value nor a truth value. Here missing data cannot be handled and that is the future work.

Ensemble learning is a learning method where multiple learners are trained to solve the same problem to obtain better predictive performance to find the suitable hypothesis. Bolon et al. [12] presented an ensemble which combined filters rather than a single method for their increase in individual strength and stability of the feature selection process, overcome their weak points of the datasets. Two approaches are presented based on the classifier. Ensemble 1 classifies more number of times as there are filters and Ensemble 2 classifies only once in the result of combining the different subsets by the filters. Ensemble has achieved the best performance by tested the all scenarios and produces the better result dataset.

2.5 Support Vector Machines

Support Vector Machines (SVMs) [13] is an effective classifier that provides the representation which depends upon few parameters and adequate generalization of new objects. Resampling, cost-sensitivity learning, feature selection and one class learning are the solutions that handle the classifying imbalanced data sets problem. The author proposed a new feature selection method that reduces the dimensionality for subgrouping of features. This proposed approach performs the feature ranking for predictive performance and achieved better results on the imbalanced data sets, minimized the error numbers in the minority class. Backward feature elimination approach used for feature ranking to remove less impact features. This approach was more flexible and allows some different kernel functions for classification and nonlinear feature selection using the SVM.

Shen et al. [14] proposed a new Fruit fly Optimization Algorithm (FOA) for effective parameter optimization of the machine learning algorithm. FOA has the ability, which addressed the SVM's model selection problem for classification. SVM has shown that too performed very well in the task of medical diagnosis. FOA-SVM was proposed effectively and successfully detected the problem of medical data classification. The objective function was designed to explore the SVM's maximum generalization ability by considered the FOA-SVM cross validation classification accuracy. The FOA method's efficiency and effectiveness was observed in terms of the classification accuracy, specificity, CPU time and sensitivity on the medical datasets that was taken from the UCI machine learning repository. This proposed work has produced high predictive accuracy with minimum processing time.

2.6 Fuzzy Sets

Fuzzy logic [1] used to improve the detection accuracy and a fuzzy that supports SVM for network intrusion detection. SVM detection agents were used to detect the TCP attacks, UDP attacks, content- based detection and ICMP attacks separately. This multi-agent collaborative detection method has increases the detection accuracy and speed.

Yu-Ping et al. [21] proposed a Hierarchical based neuro–fuzzy inferencing intrusion detection (HFIS), in which the principal component analysis of neural network was used to reduce the data size. The main merits of HFIS were it has the capability to perform the misuse detection and anomaly detection. So this detection method has better performance and higher speed.

Li et al. [22] has proposed the Hidden Markov Model (HMM) and the fuzzy logic which depends on hybrid intrusion detection method. This method was more efficient for classifying the anomaly data profile from the normal data profile. This HMM method needs only less storage amount without the profile data and the training time was reduced, which needed less testing data.

2.6.1 Fuzzy Cognitive Map

Papageorgiou et al. [15] presented a new soft computing approach called Fuzzy cognitive map (FCM) which is used for medical decision support system. FCM used to model the variations in complex systems using the relationship and set of concepts. FCM has inherited the quality of neural networks and fuzzy logic in the graph based structure modeling complex decision making problem. Due to this decision system that supports uncertain knowledge, flexibility and adaptable to any complex problems, comprehensible modeling philosophy was near to the human reasoning and it has the capability to handle the complex issues in different domains, this FCM has found large applicable to many different kinds of scientific areas from knowledge modeling to decision making and prediction and became popular.

FCM consist of many concepts and edges that were represented the casual relationship between the concepts. This concept has mirrored the characteristics, key factors and the qualities of the system. Here, each concept has the activation value and signed fuzzy weights of the edges. The positive sign has indicated the direct relationship between the concepts and the negative sign has indicated the indirect relationship. To increase the effectiveness of the FCM, this method adjusted the weights which assigned initially and then extract the hidden valuable knowledge that given by the experts by introducing the learning algorithms. The classification accuracy of this proposed Nonlinear Hebbian Learning (NHL-FCM) method was high compared to other learning methods and the FCM was transparent, simple and comprehensive to access the

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, October 2017, pp. 7–14

accurate asses of risk level of the concepts and decision making.

2.7 Rough Sets

Zihui Che [23] has presented the anomaly detection model that was based on a rough set method and the HMM concept and this method has several advantages:

- Number of attributes and the complexity was reduced, so that the training time of the HMM has decreased after the reduction of the redundant information.
- Rough set method reduction generates decision conditions, which to make further detection after the HMM evaluation. The accuracy of anomaly detection was improved by revised the detection results.
- Misuse and malicious intrusions were identified by means of attribute reduction.

Incremental Dependency classes (IDCs) [16] was proposed to calculate rough set dependency and feature selection algorithm was recommended for the benefits:

- IDCs produced 100% accurate dependency measures.
- IDCs neglect the calculation of the positive regions and that can be used for larger datasets.
- It requires 68% less runtime memory when compared to their counterparts using the positive region based method.
- Time has reduced by using IDCs and that used to performing other tasks, e.g., classification, clustering and pattern recognition.

3. COMPARISON METHODOLOGIES

3.1 Fuzzy Rough-Set Based Method for Feature Selection in Mixed Data

Fuzzy Rough-set based method deals with the real valued data and also processed the mixed data. This method was based on some fuzzy relations that are defined for different kinds of attributes that measured the similarities between objects. The distance function was employed to some fuzzy relations for the real-value attributes. This feature selection based on fuzzy rough set was basically called as attribute reduction. Conditional entropy was used for selecting best features of mixed data. The proposed fuzzy rough set based feature selection algorithm was performed well in terms of the classification accuracy when compared with existing feature selection methods. The overall classification accuracy of the proposed feature selection was 80% only.

3.2 Multilabel Classification

The k-NN classifier was used for the predicting the dataset and it uses SNN and certainty factor for reliable data, predicting the neighbor information and handling unbalanced problems. SNN was free from the parameter k that compared to k-NN. In this method number of the selected nearest neighbors was a variable which determined by data. The certainty factor has the value of -1 to 1, which represents the change degree in belief. The belief increased if it is positive and negative when the belief was decreased. Binary relevance was used to transform the multilabel data into single labeled data. This proposed BRSC takes the SNN as reliable information for prediction and this certainty factor determined the final class labels of the unknown/new instance. This proposed method does not handle the missing values for multilabel classification.

In the feature selection wrapper method, which uses a decision tree to remove the redundant features subsets and the genetic-X-means was used for the relevant feature selection, so the testing time was reduced to improve the accuracy. In the filter method, CFS and mRMR was used for the subset feature selection and the Sparse representation was based on the effective distance that used for dimension reduction and spectral based clustering. The GCACO reduces the irrelevant features and the dependency was maximized. Ensemble learning method that uses for better predictive performance. FCM model used for the complex decision making support and it was transparent, simple and comprehensive for the decision support in the personalized risk assessment

4. RESULT ANALYSIS

Using the fuzzy rough set method the classification accuracy was gained which compare to other fuzzy classification methods. The SNN and certainty factor predicted the known/unknown labels by the degree of belief. The machine learning feature selection was effective and efficient UCI datasets saves 90% computation time and produces 98% effective feature selection results. From the Sparse representation the proposed EDLS, EDSS-1and EDSS-2 produces 80% clustering results of different unsupervised feature selection and the LS, SS-1 and SS-2 produces to 98% classification results of different unsupervised feature selection method. In the GCACO method the classification accuracy for different datasets that produces 95.90% accuracy.

Through the FCM method, 95% classification accuracy was obtained from the personalized risk assessment model with the medical guidelines for the decision support. Table 1 shows the performance analysis of the various methods. From Table I, it can be seen that the performance of the IREMSVM (Ganapathy et al. 2013) is very high when it is compared with other existing methods. This is due to the use of intelligent agents for decision making.

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, October 2017, pp. 7–14

TABLE I PERFORMANCE ANALYSIS

S. No.	Author(s)	Methods	Overall Accuracy (%)
1	Zhang et al. (2016)	Fuzzy Rough Set	87.00%
2	Liu et al. (2016)	BRSC	99.40%
3	Liu et al. (2016)	EDSS-2	83.00%
4	Parham et al. (2015)	GCACO	95.90%
5	Papageorgiou et al. (2015)	FCM	95.00%
6	Al-Yaseen et al. (2017)	Modified k-means	95.70%
7	Guo et al. (2016)	Hybrid approach	95.76%
8	Ganapathy et al (2013)	IRESVM	99.60%
9	Li et al. (2016)	MVPSys	95.00%

5. SUGGESTION PROPOSED

End of the above analysis, no system has been proposed for identifying the novel attacks effectively. For improving the detection accuracy, we can use intelligent fuzzy rules with spatio-temporal nature. These rules are capable of identifying the new attacks dynamically and efficiently. Moreover, we can include some new features which are useful for identifying the novel attacks. New features must be finalized based on the recent attacks behaviors. These features and intelligent agents are helpful to identify the new available attacks and also predict and detect the upcoming attacks.

6. CONCLUSION

In this paper, we made a survey on the intelligent soft computing techniques based feature selection and classification algorithms for Intrusion detection. The advantages of soft computing techniques and intelligent agents on feature selection and classification algorithms to perform effective classification were analyzed. The soft computing technique based classification algorithms include Fuzzy classifier, Fuzzy Rough set, Neural Network and neuro-fuzzy classifiers based are discussed in this paper. Comparative analysis is also performed in this paper for the recent classifiers which are proposed in this direction. In addition, the performance analysis is also carried out to identify the best classifier. End of this survey, we have found that the lack of availability for identifying the novel attacks in internet. For this purpose, we have suggests to introduce new fuzzy rules with temporal constraints for identifying the novel attacks effectively.

REFERENCES

[1] S. Ganapathy, K. Kulothungan, S. Muthurajkumar, M. Vijayalakshmi, P. Yogesh, and A. Kannan, "Intelligent Feature Selection and Classification Techniques for Intrusion Detection in Networks: A Survey," *Commun. Networking, Eurasip journals, Springer*, Vol. 2013, No. 1, pp. 271, 2013.

[2] O. Joldzic, Z. Djuric, and P. Vuletic, "A transparent and scalable anomaly-based DoS detection method," *Comput. Networks, Elsevier*, Vol. 104, pp. 27–42, 2016.

[3] O. Y. Al-Jarrah, O. Alhussein, P. D. Yoo, S. Muhaidat, K. Taha, and K. Kim, "Data Randomization and Cluster-Based Partitioning for Botnet Intrusion Detection," *IEEE Trans. Cybern.*, Vol. 46, No. 8, pp. 1796–1806, 2015.

[4] S. S. Sivatha Sindhu, S. Geetha, and A. Kannan, "Decision tree based light weight intrusion detection using a wrapper approach," *Expert Syst. Appl., Elsevier*, Vol. 39,No. 1, pp. 129–141, 2012.

[5] V. Bolon-Canedo,I. Diaz, N.Marono and A. Alonso, "A framework for cost-based feature selection," *Pattern Recog., Elsevier*, Vol. 47, pp. 2481–2489, 2014.

[6] F. Wang and J. Liang, "An efficient feature selection algorithm for hybrid data," *Neurocomputing, Elsevier*, Vol. 193, pp. 33–41, 2016.

[7] M. Liu and D. Zhang, "Feature selection with effective distance", *Neurocomputing, Elsevier*, 2016, http://dx.doi.org/10.1016/ j.neucom.2015.07.155

[8] M. Parham. and R. Mehrdad., "Integration of graph clustering with ant colony optimization for feature selection," *Knowledge-Based Syst.,Elsevier*, Vol. 84, pp. 144–161, 2015.

[9] P. Teisseyre, "Feature ranking for multi-label classification using Markov networks," *Neurocomputing, Elsevier*, Vol. 205, pp. 439–454, 2016.

[10] S. Sasikala, S. Appavu, and S. Geetha, "A novel adaptive feature selector for supervised classification," *Inf. Process. Lett.,Elsevier*, Vol. 1, 2016.

[11] H. Liu, X. Wu, and S. Zhang, "Neighbor selection for multilabel classification," *Neurocomputing, Elsevier*, Vol. 182, pp. 187–196, 2016.

[12] V. Bolón-Canedo, N. Sánchez-Maroño, and A. Alonso-Betanzos, "Data classification using an ensemble of filters," *Neurocomputing, Elsevier*, Vol. 135, pp. 13–20, 2014.

[13] S. Maldonado, R. Weber, and F. Famili, "Feature selection for high-dimensional class-imbalanced data sets using Support Vector Machines," *Inf. Sci. (Ny)., Elsevier*, Vol. 286, pp. 228–246, 2014.

[14] L. Shen, H. Chen, Z. Yu, W. Kang, B. Zhang, H. Li, B. Yang, and D. Liu, "Evolving support vector machines using fruit fly optimization for medical data classification," *Knowledge-Based Syst., Elsevier*, Vol. 96, pp. 61–75, 2016.

[15] E. I. Papageorgiou, J. Subramanian, A. Karmegam, and N. Papandrianos, "A risk management model for familial breast cancer: A new application using Fuzzy Cognitive Map method," *Comput. Methods Programs Biomed., Elsevier*, Vol. 122, No. 2, pp.123–135, 2015.

[16] M. S. Raza and U. Qamar, "An incremental dependency calculation technique for feature selection using rough sets," *Inf. Sci. (Ny)., Elsevier*, Vol. 343–344, pp. 41–65, 2016.

[17] Y. Li, J Xia, S Zhang, J Yan, X Ai, K Dai, "An efficient intrusion detection system based on support vector machines and gradually feature removal method." *Expert Syst. Appl* ,Vol. 39, pp. 424–430, 2012

[18] F. Amiri, MMR Yousefi, C Lucas, A Shakery, N Yazdani, "Mutual information based feature selection for intrusion

ISSN: 2456 - 3935
International Journal of Advances in Computer and Electronics Engineering
Volume 2, Issue 10, October 2017, pp. 7–14

detection systems." *J. Network Comput. Appl* , Vol. 34, pp. 1184–1199, 2011

[19] R. Battiti, "Using mutual information for selecting features in supervised neural net

learning." *IEEE Trans. Neural Netw.*,Vol. 5, pp. 537–550, 1994

[20] KK Gupta, B Nath, R Kotagiri, "Layered Approach using Conditional Random Fields for Intrusion Detection.", *IEEE Trans. Dependable Secure Comput.*,Vol. 7, 2010

[21] Y.P Zhou, J.A Fang, Y.P Zhou, "Intrusion Detection Model Based on Hierarchical Fuzzy Inference System", *Second IEEE International Conference on Information and Computing Science* , Vol 2 (IEEE Computer Society, Washington), pp. 144–147, 2009

[22] Y. Li, R Wang, J Xu, G Yang, B Zhao, "Intrusion detection method based on fuzzy hidden Markov model." *Sixth IEEE International Conference on Fuzzy Systems and Knowledge Discovery*, Vol 3 (IEEE, Piscataway), pp. 470–474, 2009

[23] ZCX Ji, "An efficient intrusion detection approach based on hidden Markov model and rough set", *IEEE International Conference on Machine Vision and Human machine Interface (IEEE Computer Society, Washington)*, pp. 476–479, 2010.

[24] X. Zhang, C. Mei, D. Chen, and J. Li, "Feature selection in mixed data: A method using a novel fuzzy rough set-based information entropy," *Pattern Recognit., Elsevier* , Vol. 56, pp. 1–15, 2016.

[25] C. Guo, Y. Ping, N. Liu, and S. S. Luo, "A two-level hybrid approach for intrusion detection," *Neurocomputing, Elsevier*, Vol. 214, pp. 391–400, 2016.

[26] W. Li, W. Meng, X. Luo, and L. F. Kwok, "MVPSys: Toward practical multi-view based false alarm reduction system in network intrusion detection," *Comput. secur., Elsevier*, Vol. 60, pp. 177–192, 2016.

[27] W. L. Al-yaseen, Z. Ali, M. Zakree, and A. Nazri, "Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system," *Elsevier*, Vol. 67, pp. 296–303, 2017.

[28] S Ganapathy, K Kulothungan, P Yogesh, A Kannan, "A Novel Weighted Fuzzy C–Means Clustering Based on Immune Genetic Algorithm for Intrusion Detection", *Procedia Engineering*, vol. 38, pp. 1750-1757, 2012.

[29] S.Ganapathy, N.Jaisankar, P.Yogesh, A.Kannan, " An Intelligent System for Intrusion Detection using Outlier Detection", *2011 International Conference on Recent Trends in Information Technology (ICRTIT)*, pp. 119-123, 2011.

[30] S Ganapathy, R Sethukkarasi, P Yogesh, P Vijayakumar, A Kannan, "An intelligent temporal pattern classification system using fuzzy temporal rules and particle swarm optimization", *Sadhana*, vol. 39, no. 2, pp. 283-302, 2014.

[31] N Jaisankar, S Ganapathy, P Yogesh, A Kannan, K Anand, "An intelligent agent based intrusion detection system using fuzzy rough set based outlier detection", *Soft Computing Techniques in Vision Science*, pp. 147-153, 2012.

[32] K Kulothungan, S Ganapathy, S Indra Gandhi, P Yogesh, A Kannan, "Intelligent secured fault tolerant routing in wireless sensor networks using clustering approach", *International Journal of Soft Computing*, vol. 6, no. 5, pp. 210-215, 2011.

[33] S Ganapathy, P Vijayakumar, P Yogesh, A Kannan, "An Intelligent CRF Based Feature Selection for Effective Intrusion Detection", *International Arab Journal of Information Technology*, Vol. 16, No. 2, pp. 44-50, 2016.

[34] S. Ganapathy, P. Yogesh, and A. Kannan, "Intelligent Agent-Based Intrusion Detection System Using Enhanced Multiclass SVM", *Computational Intelligence and NanoScience*, Vol. 2012, pp. 1-10, 2012.

[35] P. H. Rajmahanty, S.Ganapathy, "Role of Decision Trees in Intrusion Detection Systems : A Survey," *International Journal of Advances in Computer and Electronics Engineering*, Vol. 2, no. 4, pp. 9–13, 2017.

Authors Biography

Riyaz. B, is a Research Scholar of Department of CSE in VIT University, Chennai Campus, Chennai, Tamilnadu, India. He completed his B.Tech in Information Technology at Anna University, Chennai. He completed his M.E in Computer Science and Engineering at Karpagam University, Coimbatore. His research interests are Computer Networks, Network Security and Data Mining.

Ganapathy. S is currently working as Assistant Professor (Senior) in the School of Computing Science and Engineering, VIT University-Chennai Campus. He has received Ph.D and M.E degrees in Computer Science and Engineering from Anna University, Chennai, India. He has published more than 40 papers including reputed journals. His areas of interests are Data Mining Techniques and Network Security.

Cite this paper:

Riyaz B, Ganapathy S, "Intelligent Soft Computing Techniques for Providing Network Security: A Survey", International Journal of Advances in Computer and Electronics Engineering, Vol. 2, No. 10, pp. 7-14, October 2017.