VOL. 10, NO. 18, OCTOBER 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

SWARM BASED CLASSIFIER MODEL USING ENSEMBLE FEATURE RANKING METHODS

P. Amudha[1], S. Karthik[2] and S. Sivakumari[1]
[1]Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India
[2]Department of Computer Science, SNS College of Technology, Coimbatore, India
E-Mail: amudharul@gmail.com

ABSTRACT

Intrusion Detection System (IDS) is a security support mechanism which has become an essential component of security infrastructure to detect attacks, identify and track the intruders. In intrusion detection, the quantity of data is huge that includes thousands of traffic records with number of various features. Selecting a subset of informative features can lead to improved classification accuracy. In this paper ensemble of feature ranking techniques are used to select the most relevant features that can represent the pattern of the network traffic. The efficiency of the presented method is validated on KDDCUP'99 dataset using hybrid swarm based classifier, Simplified Swarm Optimization (SSO) with Ant Colony Optimization (ACO). The performance of the proposed method is compared with the SSO and hybridization of SSO with Support Vector Machine (SVM). It is shown that the hybridization of SSO with Ant Colony Optimization using hybrid feature ranking method outperformed other algorithms and can be efficient in the detection of intrusive behaviour.

Keywords: intrusion detection, simplified swarm optimization, ant colony optimization, support vector machine, KDDCUP'99.

INTRODUCTION

In recent years, Data Mining techniques on network traffic data provides a potential solution that helps develop better intrusion detection systems. Feature selection is an important pre-processing tool in data mining that helps in increasing the performance of classification models [1]. The purpose of this paper is: to select the most informative feature using an ensemble feature ranking technique. The ensemble approach determines a feature's importance or score from multiple feature ranking techniques which are combined to generate three ranking lists: fusion, selection and hybrid; to introduce a hybrid classifier model of Simplified Swarm optimization (SSO) with Ant Colony Optimization (ACO) to maximize the classification accuracy. The purpose of this hybridization is to improve the performance of SSO for mining the intrusion pattern of the network traffic.

The remaining part of this paper is organized as follows: Section 2 gives related work; Section 3 explains the theoretic aspects of techniques. Section 4 explains the methodology used in this work. Section 5 provides experimental results and discussions. Finally section 6 gives the conclusion.

RELATED WORK

Feature selection is a method which selects features from the original dataset that can produce maximum classification accuracy for the predicted target class. Early studies on ensemble of feature ranking techniques were performed by Rokach *et al*. [2]. The experiments in this study are performed to check whether ensemble of feature subsets improve classification accuracy over individual rankers. Kees Jong *et al.* [3] proposed an ensemble approach by combining feature rankings extracted from different iterations of an algorithm named ROGER. Saeys *et al*. [4] performed a study on ensemble of feature selection techniques. The study proves that ensemble methods provide more robust and stable results for high dimensional datasets when compared to individual feature selectors. Sri HarshaVege [5] proposed an ensemble of multiple feature ranking techniques.

Evolutionary computation and swarm intelligence techniques are great examples that nature has been a continuous source of inspiration. The behavior of bees, ants, glow-worms, fireflies, fishes and other organisms have motivated researchers to develop new optimization algorithms [6]. Dorigo *et al*. [7] [8] [9] presented an algorithmic implementation of the ant behaviour for solving minimum cost path problems on graphs known as simple Ant Colony Optimization. ACO is set apart from the other approaches, as it is primarily applied to combinatorial optimisation. ACO based classification algorithm called Ant-Miner was developed by Parpinelli *et al*. [10] which support only nominal attributes and generates ordered rule set. This approach provides only less information to the classifier which can have a negative impact on the accuracy of the discovered knowledge. Vittorio Maniezzo *et al.* [11] discussed that the Ant Colony Optimization (ACO) can be used as a model for developing meta-heuristic algorithms for combinatorial optimization problems. Nicholas Holden and Alex A. Frietas [12] proposed hybridization of the PSO/ACO algorithm to benchmark the performance of classification algorithms. Nada M. A. Al Salami [13] proposed a hybridized approach using Ant Colony and Genetic programming algorithms to solve combinatorial optimization problem and to find the optimal solution, by accumulating the most effective sub-solutions. Noorhaniza Wahid [14] proposed a Simplified Swarm Optimization

ARPN Journal of Engineering and Applied Sciences

(SSO) algorithm for feature selection and classify the audio data. Revathi S and A Malathi [15] preprocessed the data using a hybrid SSO and compared the proposed approach with PSO with Random Forest.

THEORETIC ASPECTS OF TECHNIQUES

Simplified swarm optimization

Particle Swarm Optimization (PSO) is a population based stochastic optimization technique and it solves data mining problems by simulating the social behaviors observed in animals or insects, such as birds flocking or fish schooling. PSO has been successfully dealt with continuous variables and it suffers from early conjunction particularly in high dimension multimodal problems [16]. SSO [14] was proposed to solve classification problem which can deal with the dataset containing both the discrete and continuous variables.

The steps involved in SSO algorithm are:

a) Initialize randomly the positions X_i and V_i of each particle in the swarm (it must be within the boundaries defined to the positions and velocities).
b) The parameters such as swarm size, the number of maximum generation, maximum fitness value and three predetermined constants: C_w, C_p, C_g are determined initially. Generation and initialization of pbest and gbest is done with the random position.
c) For each particle in the swarm, evaluate the fitness value using the problem's objective function.
d) Modify the best particle position among all the swarm particles based on step 3.
e) Update the particle's velocity and position vectors using equations (1),

$$\begin{aligned} & if\ (0 \le R \le C_w), then\ \{x_{id} = x_{id}\}; \\ & elseif\ (C_w \le R < C_p), then\ \{x_{id} = p_{id}\}; \\ & elseif\ (C_p \le R < C_g), then\ \{x_{id} = g_{id}\}; \\ & elseif\ (C_g \le R \le 1), then\ (x_{id} = new(x_{id})\}; \end{aligned} \tag{1}$$

f) Iteratively repeat the steps 2, 3 and 4 until the best fitness value is obtained or the maximum
g) generation is met by equation (2),

$$x_{id}^t = \begin{cases} x_{id}^{t-1} & if\ rand() \in [0, C_w) \\ p_{id}^{t-1} & if\ rand() \in [C_w, C_p) \\ g_{id}^{t-1} & if\ rand() \in [C_p, C_g) \\ x & if\ rand() \in [C_g, 1) \end{cases} \tag{2}$$

In SSO, velocity of each particle is updated using equation (3),

$$v_{id}^t = w \cdot v_{id}^{t-1} + c_2 \cdot rand_2 \cdot \left(p_{gd} - x_{id}^{t-1}\right) \tag{3}$$

where, w is inertia weight, c_2 denote the acceleration coefficient, d = 1,2,…,D, $rand_2$ is random number uniformly distributed in the range of [0, 1]. Then each particle moves towards a new potential position as follows:

$$x_{id}^t = x_{id}^{t-1} + v_{id}^t, \qquad d = 1,2,...,D \tag{4}$$

Ant colony optimization

Ant Colony Optimization (ACO) is one of the recent techniques for approximate optimization which is inspired by ants foraging behavior [9]. ACO is applied for discovering classification rules and each classification rule has the form, IF<*term1* AND *term2* AND…> THEN <*class*> [10]. In Ant algorithm, each ant constructs a solution incrementally for the target problem. Ant-Miner consists of three major steps namely rule construction, rule pruning and pheromone updating. The probability that the term is added to current partial rule which is under the construction of ants is given by equation (5),

$$\rho_{ij} = \frac{\tau_{ij}(t) \cdot \eta_{ij}}{\sum_{i}^{a} \sum_{j}^{b_i} \tau_{ij}(t) \cdot \eta_{ij}, \forall i \in I} \tag{5}$$

where, η_{ij} is a problem-dependent heuristic vale for $term_{ij}$, $\tau_{ij}(t)$ is the amount of pheromone between attributes i and j, a is the total number of attributes, b_i is the total number of values of attribute i, I is the set of attributes that are not yet used by the ant. After the rule construction, rule pruning is performed to improve the quality of a rule by removing the irrelevant terms that was added during rule construction. The quality of a rule is measured by equation (6),

$$Q = (TP / TP + FN) * (TN / FP + TN) \tag{6}$$

where, TP is the number of cases covered by the rule and having the class predicted by the rule, TN is the number of cases that are not covered by the rule and having a class different from the class predicted by the rule, FP is the number of cases covered by the rule and having a class different from the class predicted by the rule, FN is the number of cases that are not covered by the rule while having the class predicted by the rule. The pheromone is updated after each constructs a rule and the amount of pheromone deposited is proportional to the goodness of the solution an ant has built.

$$\tau_{ij}(t+1) = [\tau_{ij}(t)] + [\{\rho\} \cdot \Delta\tau_{ij}(t)] \tag{7}$$

ARPN Journal of Engineering and Applied Sciences

METHODOLOGY

Ensemble feature ranking methods

In this experiment, the benchmark intrusion detection dataset, KDDCup'99[17] which is obtained from UCI machine Learning Repository [18] is used. After eliminating the duplicate instances, instances with the same attack category and 10% normal instances are included in each dataset. After the pre-processing step, ensemble of feature ranking techniques is obtained by combining various feature ranking techniques to generate a single ranking list. This technique is used to improve the performance of the classifiers. The two steps in obtaining ensemble list are:1) Create a set of n ranking lists using different feature ranking methods 2) Create a single ranking list using the three types of permutation methods: fusion based, selection based, and hybrid. Fusion based makes use of all the features obtained from individual feature selectors and selects fixed number of highly ranked features to form a final result. Selection based method calculates the occurrence count of each feature in all ranking lists. Then it sorts the features according to occurrence count. Finally selects the features whose occurrence count is more than a user defined threshold value [5]. The working principle of the algorithm is:

1. Let a= (a_1, a_2, a_3,… ,a_n) be the feature set of a dataset D.
2. R is a set of n ranking lists obtained, where R = {R_1, R_2,… ,R_n}
3. $\forall$ $R_i \in$ R, where i = 1 to n

$\forall$ f_i^j //each feature i in list j

if ($f_i^j \notin$ F) then, F $\longleftarrow$ (f_i^j , r[f_i^j])

if ($f_i^k \in$ F) then, F $\longleftarrow$ r[f_i^k] ,where k = j+1 to n-1

4. Repeat the above step until all features in n ranking lists are checked.
5. Calculate the frequency count of each feature in the list F and sort it.
6. Create list E with top ranked features.

The features selected using ranking methods: Gain ratio, Chi-squared and Relief using Weka tool [19] are given in Table-1. In the hybrid method, the classification algorithms [20]: J48, Naïve Bayes, BayesNet (Hill climbing and Tabu search methods) which shows maximum accuracy rate on all datasets are identified to select the best features. The important features are chosen by finding the common features using 10-fold cross validation method. The features selected using ensemble feature ranking methods are given in Table-2-4.

Table-1. Features selected using ranking methods.

Dataset	Feature selection methods		
	Gain ratio	Chi-squared	Relief
DoS+10% normal	3,4,5,6,8,10,12,13,23,25,26, 29,30,31,33,34,35,37,38,39	5,6,3,23,30,29,33,34,12,4,35,38, 25,26,39,37	3,4,13,29,39,26,25,38,33
Probe+10% normal	2,3,4,5,6,12,25,26,27,28,29, 30,33,34,35,37,38,39,40,41	5,3,6,33,35,12,40,34,27,37,29	3,12,2,36,33,34,35,32,31,37
U2R+10% normal	9,10,11,13,14,16,17,18,25,29	14,3,13,10,17,1,16,33,37, 5,32,36	1,33,36,32,34,12,31,2,35,27
R2L+10% normal	3,4,5,6,9,10,11,13,14,17,18, 19,22,25,26,33,38,39	5,3,6,33,10,36,22,37,1,24,32,23	3,33,36,34,14,32,2,31,12,4

Table-2. Features selected by fusion method.

Dataset	Selected features	# features
DoS+10%Normal	3-6,8,10,12,13,24-26,32,33,38,39	15
Probe+10%normal	2-4,6,8,12,34,35,37	9
U2R+10%normal	1-3,7,9,10,11,14	8
R2L+10%normal	4,5,7,9,10,12,22,33,36	9

VOL. 10, NO. 18, OCTOBER 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

Table-3. Features selected by selection.

Dataset	Selected features	# features
DoS+10%normal	3-6,13,23,25,26,29,30,33-35,37-39	16
Probe+10%normal	2,3,5,6,27,33-35,37,40	10
U2R+10%normal	1,10,13,14,16,17,32,33,36	9
R2L+10%normal	3-5,6,10,14,22,32,36	9

Table-4. Features selected by hybrid ensemble method.

Dataset	Selected features	# features
DoS+10%normal	3-6,12,23,25,26,29,30,33,34,35,37-39	15
Probe+10%normal	3,5,6,12,29,33,34,35,37,40	10
U2R+10%normal	10,13,14,17	4
R2L+10%normal	3,5,6,10,22,33	6

Intrusion detection system using hybrid SSO-ACO algorithm

The proposed SSO-ACO algorithm is used to solve the classification problem and can cope with dataset containing both discrete and continuous variables. The working procedure of hybrid SSO-ACO algorithm is as follows:

a) Initialize parameters and population randomly
b) Initialize each particle position vector and velocity vector
c) Evaluate the fitness for each Particle using the ACO algorithm
d) Find the personal best and the Global best
e) While stopping criterion not satisfied do
   Update each particle's velocity and Position
   Evaluate the fitness for each Particle using the ACO algorithm and update the personal best and the global best
f) End while

In SSO-ACO, each particle is considered as a set of *n*pheromone matrices where *n*is the number of features in a data set. Each particle can be interpreted into a rule with a predefined class. During the evaluation of particle, the vector is transformed to a set of terms or rule conditions and added to the rule produced by the algorithm for fitness evaluation. The position of each particle contains *N* dimensions of features and the predictive class *X*. The IF-THEN rule generated by SSO as in equation (8) will be performed in all dimensions:

IF $LowerBound \leq x_{ij} \leq UpperBound$ is true, THEN prediction is Class X (8)

If the two bounds cross over, both terms will be omitted from the decoded rule. This seeding procedure will likely produce some seeding positions outside the range of the values seen within the dataset [12] [16]. In SSO rule mining, the value of *Lower Bound* and *Upper Bound* will be updated during and after the generation process. The equation (4) will be used to update the new current value of the corresponding position. The SSO algorithm optimizes the upper bound and lower bound values of the terms. After generating the best rule, it is then pruned by Ant-Miner algorithm and added to the rule set. Ant-Miner's pruning procedure involves finding the term which, when removed from a rule, gives the biggest improvement in rule quality. When this term is found, it is permanently removed from the rule. This procedure is repeated until no terms can be removed without loss of rule quality. The rule quality is evaluated using equation (13). The highest fitness value of the individual within the range will be searched in every optimization process which is significantly important to obtain the best quality of the rule to solve the classification problems [12].

After the pruning procedure, the instances covered by the rule are removed from the training set. A While loop is executed provided that, the number of uncovered instances of the class C in the training set is larger than the maximum number of uncovered instances per class. When this threshold has been reached, the training set will be reset by adding the previously covered samples. The unordered rules are ordered by their quality and the whole rule set is pruned. This involves removing terms from each rule that does not affect the accuracy [12]. Also all the redundant rules which do not contribute to the classification accuracy are removed. A series of testing data are used to measure its classification accuracy according to the rule set obtained. For each instance, a prediction value is computed by examining every element in the rule set for the corresponding class if it is covered by the rule. The prediction value is calculated as in equation (9),

ARPN Journal of Engineering and Applied Sciences

Prediction value += α *rule quality + β *percentage of the rule covered (9)

where, α *and* βare two parameters corresponding to the importance of the rule quality and the percentage of the rule covered. The prediction value for each class is accumulated and the final result is calculated from the class with the highest prediction value [16]. The parameters set in the algorithm are given in Table-5.

Table-5. Parameter settings.

Parameter	Value
Number of Particles (m)	50
Maximum number of Iteration	10
Maximum uncovered cases	5
Number of rules covered	5
Minimum number of cases per rule	5
Quality weight (α)	0.5
Coverage Weight (β)	0.5

In this work, the basic pheromone updating rule in the ant algorithm is changed as,

$$\tau_{ij}(t+1)=[\{\rho+(1-\rho/1+\rho)\}\cdot\tau_{ij}(t)]+[\{\rho-(\rho/1+\rho)\}\cdot\Delta\tau_{ij}(t)] \quad 10)$$

Also the problem specific heuristic function is designed using Laplace correction formula [21],

$$\eta_{ij}=\frac{|\,term_{ij},K\,|+1}{|\,term_{ij}\,|+No\_of\_classes} \quad (11)$$

where,

$|term_{ij}|$ indicates the number of training sets that consists of $term_{ij}$ in the current k class

No_of_classes defines the numbers related to the variable class.

ANALYSIS OF EXPERIMENTAL RESULTS

This paper report values for classification accuracy (CA), detection rate (DR), false alarm rate (FAR), are used to evaluate the performance of intrusion detection tasks. The performance of the classifier has been evaluated using 10-fold cross validation method and algorithm runs 10 times for each fold. The performance of proposed hybrid approach is compared with SSO and hybridization of SSO with Support Vector Machine (SVM) [22]. Table-6 presents the average classification accuracies of all datasets and ranking based on average values of each classifier on comparing the three ensemble ranking methods. The ranking of each technique is given in the parenthesis.

Table-6. Comparison of average classification accuracies in % and ranking of techniques.

Dataset	fusion			selection			hybrid		
	SSO	SSO-ACO	SSO-SVM	SSO	SSO-ACO	SSO-SVM	SSO	SSO-ACO	SSO-SVM
DoS+10% normal	67.34(3)	81.3(3)	76.64(3)	67.71(2)	94.52(2)	82.46(2)	80.41(1)	96.49(1)	93.34(1)
Probe+10% normal	71.22(2)	74.4(3)	67.82(3)	67.24(3)	87.85(2)	78.23(2)	80.47(1)	94.49(1)	86.05(1)
U2R+10% normal	68.79(2)	85.93(3)	76.64(3)	67.71(3)	91.34(2)	86.92(2)	84.32(1)	95.32(1)	91.53(1)
R2L+10% normal	71.59(2)	84.4(3)	78.37(3)	67.71(3)	88.65(2)	85.26(2)	85.42(1)	94.15(1)	87.4(1)
Avg.acc (%)	69.67(8)	81.50(6)	74.86(7)	67.57(9)	90.59(2)	83.21(4)	82.62(5)	95.11(1)	89.58(3)
rank	2	3	3	3	2	2	1	1	1

The Table shows that the classification accuracy of hybrid ensemble feature ranking method outperforms both fusion and selection methods in all the four datasets. Among the hybrid combination, SSO-ACO shows the best performance and is competitive with SSO-SVM. The overall rank is specified along with average classification accuracy within the parenthesis. It is noted that hybrid-SSO-ACO achieves 95.11% and scores high among other classifier combinations followed by selection-SSO-ACO and hybrid-SSO-SVM. The sum of the ranks of each dataset is calculated from Table-6 and from this ranking, SSO based classifier with the hybrid ensemble can be seen as the best approach among other combination of swarm intelligence algorithms.

ARPN Journal of Engineering and Applied Sciences

The accuracy rate of SSO-ACO using hybrid ensemble feature selection method is also compared with the benchmark classifier, Naïve Bayes and the result is plotted in Figure-1. It is evident that the swarm based hybrid classifier SSO-ACO outperforms Naïve Bayes for three datasets, where U2R dataset obtains maximum accuracy (99.58%) with Naïve Bayes. The experimental results of the ensemble feature ranking methods and swarm based classifiers for all four datasets on various metrics are presented in Figures 2, 3 and in Table-7.

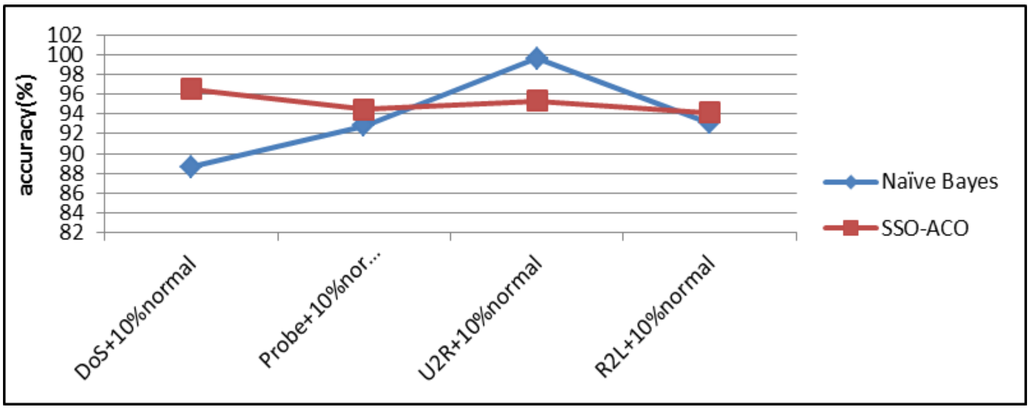

Figure-1. Comparison on accuracy rate with Naïve Bayes classifier.

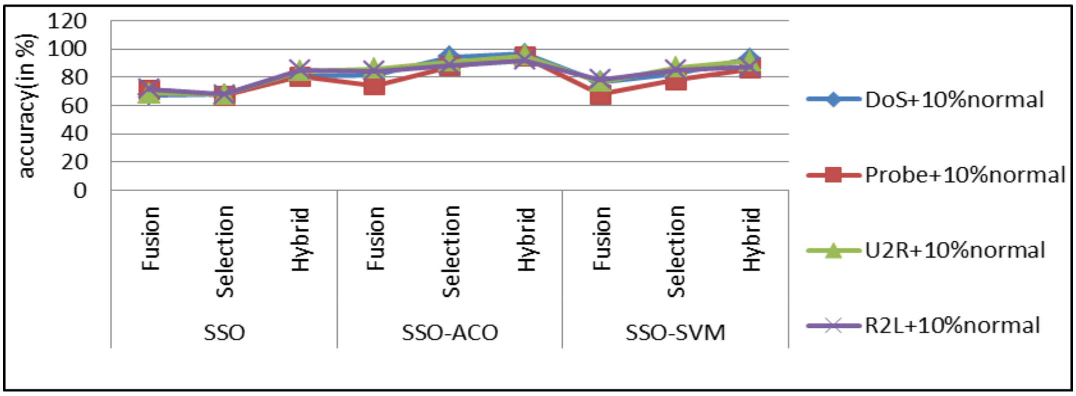

Figure-2. Accuracy of different methods.

Analyzing the Figure-2, it is apparent that SSO-ACO achieved higher accuracy than the other pesented classifiers using the three ensemble feature ranking approaches. These reported values indicate that SSO-ACO applied with hybrid feature ranking method has the effect of improving the classifiaction accuracy on all datasets.The improvement was generally higher for the DoS+normal (96.49%) and U2R+normal (95.32%) datasets. SSO-SVM classifier showes a comparable performance whereas, SSO gives similar results for all datasets. Figure-3 reveals that SSO-ACO achieves high detection rate compared to other classifiers for all ensemble feature ranking methods. Fusion method with SSO-ACO obtains detection rate of 97.55% on DoS+normal dataset and similar results on Probe, U2R and R2L datasets.Also, this hybrid swarm based classifier with the combination of selection and hybrid methods outperform other classifiers. On comparing, SSO-ACO classifier with hybrid ensemble method achieves highest detection rate(98.67%) on DoS+normal dataset; the results of SSO and SSO-SVM are comparable to with the hybrid ensemble.

ARPN Journal of Engineering and Applied Sciences

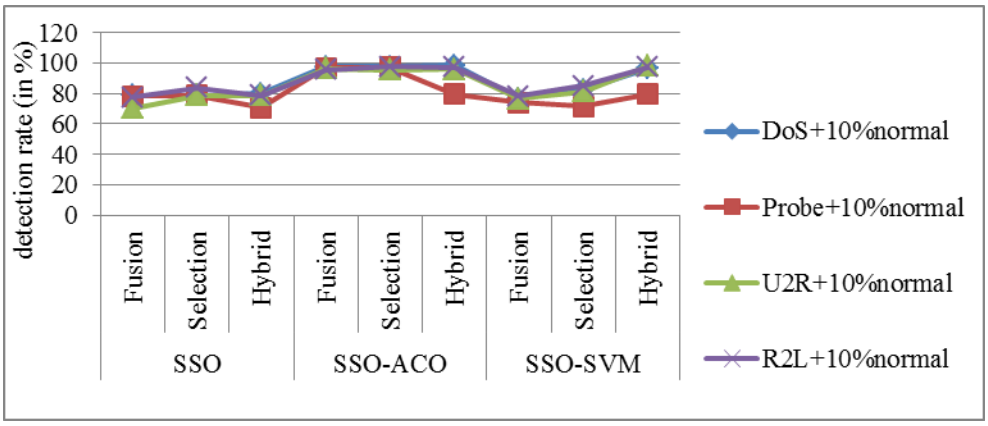

Figure-3. Detection rate of different methods.

Recently false alarm rates and detection rate are the main issues and challenges in designing an effective intrusion detection system. From Table-7, it is seen that false alarm rate of proposed SSO-ACO is low on all datasets used. However, the FAR using hybrid ensemble method obtains the lowest false alarm rate ranging between 0.0043 to 0.0489,where 0.0043% is obtained by DoS+normal dataset. Note that false alarm rate of SSO classifier is more for the datasets used on all ensemble methods. For the dataset with probe frequent attack, false alarm rate given by SSO-SVM classifier is more for both fusion and selection ensemble methods.

Table-7. FAR using ensemble feature selection methods.

Dataset	Fusion			Selection			Hybrid		
	SSO	SSO-ACO	SSO-SVM	SSO	SSO-ACO	SSO-SVM	SSO	SSO-ACO	SSO-SVM
DoS+10%normal	0.3928	0.0475	0.2335	0.3928	0.0477	0.1753	0.3928	0.0477	0.1753
Probe10%normal	0.3846	0.0468	0.385	0.3948	0.0472	0.3694	0.3948	0.0472	0.3694
U2R+10%normal	0.4043	0.0485	0.2335	0.3892	0.0487	0.1004	0.3948	0.0472	0.3694
R2L+10%normal	0.4034	0.0487	0.162	0.3928	0.0487	0.143	0.3928	0.0487	0.143

CONCLUSIONS

This work has presented three ensemble feature ranking methods from the individual feature selectors. The proposed methods selected subset of features from the intrusion detection dataset. The feasibility of selected features is evaluated on the hybrid swarm based classifiers: SSO-SVM and SSO-ACO and are compared with SSO. All the classifiers have good classification accuracy with hybrid ensemble method. The experimental results also demonstrate that the proposed hybrid feature ensemble algorithm has a superior performance on the proposed SSO-ACO classifier and outperforms over both fusion and selection methods and other classification algorithms in terms of accuracy, detection rate and false alarm rate. The impact of ensemble hybrid feature ranking method is analyzed on the benchmark classifier, Naïve Bayes. According to the good performance of SSO-ACO with hybrid ensemble feature ranking method, the proposed swarm intelligence algorithm can be used to solve intrusion detection as classification problems.

REFERENCES

[1] H. Liu, H. Motoda, R. Setiono, Z. Zhao. 2010. Feature Selection: An Ever Evolving Frontier in Data Mining. In proceedings of JMLR: Workshop and Conference. 4: 4-13.

[2] L. Rokach, B. Chizi, O. Maimon. 2006. Feature selection by combining multiple methods. Advances in Web Intelligence and Data Mining. pp. 295-304.

[3] Kees Jong, Jérémie Mary, Antoine Cornuéjols, Elena Marchiori, MichèleSebag. 2004. Ensemble Feature Ranking. University of Paris-SudOrsay, France.

[4] Y. Saeys, T. Abeel, Y. V. Peer. 2008. Robust feature selection using ensemble feature selection techniques. W. Daelemans *et al.* (Eds.): ECML PKDD 2008, Part II, LNAI. 5212, pp. 313-325.

[5] Vege, Sri Harsha. 2012. Ensemble of Feature Selection Techniques for High Dimensional Data. Masters Theses and Specialist Projects. Paper 1164.

[6] Fries T.P. 2008. A fuzzy-genetic approach to network intrusion detection. Proceedings of the 2008 GECCO Conference companion on Genetic and evolutionary computation. pp. 2141-2146.

[7] Dorigo M, Di Caro G. 1999. The ant colony optimization meta-heuristic. New ideas in Optimization. pp. 11-32.

[8] Dorigo M, Di Caro G, Gambardella L.M. 1999. Ant Algorithms for Discrete Optimization. Artificial Life. 5: 137-172.

[9] Dorigo M, Stutzle T. 2004. Ant colony optimization. MIT Press.

[10] Parpinelli RS, Lopes HS, Freitas AA. 2002. Data mining with an ant colony optimization algorithm. IEEE Transactions on Evolutionary Computation. 6(4): 21-32.

[11] Vittorio Maniezzo, Luca Maria Gambardella and Fabio de Luigi. 2004. Ant Colony Optimization. Chapter 5, Lecture notes in Computer Science.

[12] Nicholas Holden and Alex. A. Freitas. 2008. A hybrid PSO/ACO algorithm for discovering classification rules in data mining. Journal of Artificial Evolution and Applications.

[13] Nada M. A. Al Salami. 2009. Ant Colony Optimization Algorithm. UbiCC Journal. 4(3).

[14] Noorhaniza Wahid. 2011. A Novel Approach To Data Mining Using Simplified Swarm Optimization. Ph.D Thesis.

[15] S. Revathi, A. Malathi. 2013. Optimization of KDDCup99 dataset for Intrusion detection using Hybrid Swarm Intelligence with Random Forest Classifier. International Journal of Advanced Research in Computer Science and Software Engineering. 3(7): 1382-1387.

[16] Changseok Bae, wei-Chang Yeh, Noorhaniza Wahid, Yuk Ying Chung, Yao Liu. 2012. A New Simplified swarm Optimization (SSO) Using Exchange Local search Scheme. Intl. Journal of Innovative Computing, Information and Control. 8(6): 4391-4406.

[17] http://kdd.ics.uci.edu/ Databases/kddcup99/10 percent.gz.

[18] UCI Repository of Machine Learning Databases, University of California, Irvine, Dept of Information and Computer Science. http://www.ics.uci.edu/~mlearn/ML Repository.html.

[19] www.cs.waikato.ac.nz/ml/weka/.

[20] Jaiwei Han, Micheline Kamber. 2006. Data Mining Concepts and Techniques. 2nd Edition, Morgann Kaufmann.

[21] Babak Fakhar. 2014. A Novel Method for Extracting Classification Rules based on Ant-Miner. Journal of mathematics and Computer Science. (8): 377-386.

[22] Vapnik V.N. 1995. The Nature of Statistical Learning Theory. Springer-Verlag.