Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 85 (2016) 503 – 510

Procedia Computer Science

International Conference on Computational Modeling and Security (CMS 2016)

Feature selection using relative fuzzy entropy and ant colony optimization applied to real-time intrusion detection system

Ravi Kiran Varma P[a] *, Valli Kumari V[b], Srinivas Kumar S[c]

MVGR College of Engineering, Vizianagaram, AP, India
University College of Engineering, Andhra University, Visakhapatnam,, AP, India
University College of Engineering, JNTU Kakinada, Kakinada,, AP, India

Abstract

Intrusion Detection System (IDS) is one of the most important component of network defense mechanism. In an attempt to detect network attacks, network traffic features need to be identified and both attack and normal data need to be profiled. This paper proposes a set of network traffic features that can be extracted for Real-Time Intrusion Detection. This paper also proposes Fuzzy Entropy based heuristic for Ant Colony Optimization (ACO) in-order to search for global best smallest set of network traffic features for Real-Time Intrusion Detection Data set. The proposed feature reduction algorithm was tested on standard bench-mark UCI data sets, and found to be efficient. Further the algorithm was applied to Real-Time IDS data set and found to produce promising results.

Keywords: *Feature selection, Fuzzy Rough sets, Fuzzy Entropy, Ant Colony Optimization, Real-Time Intrusion Detection.*

1. Introduction

Intrusions into computer networks and systems is a major threat in today's network centric world. Few most prevalent intrusive attacks include Denial-of-Service (DoS) attacks, Distributed-Denial-of-Service (DDoS) attacks, probing based attacks, and account takeover attacks. Intrusion Detection System (IDS) is one of the main defence component that can detect such attacks from entering the network perimeter. There are few works in the literature which has mentioned and addressed the need for a real-time IDS, which makes use of network traffic features in-order

[*] Corresponding Author email: ravikiranvarmap@gmail.com

doi:10.1016/j.procs.2016.05.203

to capture the signature of the attack in the incoming data. The contributions of this work is two-fold: i) several continuous and discrete network packet features were identified to detect recent, common and novel intrusive attacks. ii) Using relative fuzzy-entropy as a heuristic, ACO was utilized to perform a global best optimized attributes search.

The author of[1] has used audit records and profiles in detecting abnormalities of system usage. John McHugh et al.[2] discussed the importance of IDS, detection approaches and methods. The work of Lee and Stolfo[3] has discussed the importance of network traffic features and treated IDS as a data analysis engine. The author[4] proposed to make use of TCP/IP protocol features to collect real-time data for IDS. The authors[5] proposed a real-time anomaly detection system, which uses incremental mining techniques to detect large scale attacks. In this paper protocol based traffic features including flags were used to construct the data base of DoS attack traffic. The work[6] proposes a practical real time IDS that makes use of machine learning techniques. In this work DoS, Probe attacks and normal traffic was profiled using several network traffic features. The authors have identified twelve features, and obtained a detection accuracy of above 98%. They have used Information Gain (IG), to select the most important features. In another similar work related to real time IDS[7] 35 features were identified for data profiling and classification of DoS attacks. Genetic Algorithm (GA) along with K-Nearest Neighbour (KNN) was used for feature selection and later identified 28 top features and a detection accuracy of 78% was reported. SYN Flag based DoS attacks were detected using threshold technique for a Rabbit processor related Embedded Web Server[8]. Rough Sets and Ant Colony Optimization (ACO) is used towards feature reduction of any kind of data sets particularly suitable for discrete data is proposed in[9]. In another work[10], Principal Component Analysis (PCA) and Artificial Neural Networks (ANN) were used to classify different types of intrusions & Q-Learning and Rough Set hybrid was used for feature selection[11].

Rough set based attribute selection requires the real-valued data to be discretized[12] as a pre-processing phase. The advantage of Fuzzy-Rough technique over Rough sets based attribute selection is that, Fuzzy based techniques do not require an additional phase of data discretization for real-valued data. Due to Fuzzification of data set, the semantics or the originality of the data are preserved better. Fuzzy-Rough techniques were used for attribute selection[13] applied to gene expression data and for classification of cancer. Jensen and Shen has proposed combination of Fuzzy Rough sets and ACO for feature selection[14]. Fuzzy-Entropy technique was proved to be effective for feature reduction[15]. A comprehensive review of IDS data pre-processing is given in[16]. This paper contributes towards identifying several discrete and continuous network traffic features which can help in detecting different types of real-time intrusion attacks. Fuzzy-Entropy was used as a heuristic factor for ACO based feature selection among the 21 identified features. The selected features were evaluated using different classifiers. Rest of the paper is organized as follows: Section 2 lists the set of network traffic features identified and types of attacks tested for this study, Section 3 gives the algorithm of fuzzy-entropy related ACO based feature selection. The experiments conducted and the result analysis was presented in Section 4, and Section 5 concludes the paper.

2. Network Traffic Feature Identification for Attack Detection

To identify network attacks in real-time, it is essential to analyse the network traffic for signature features which will help to identify attacks. Earlier studies[3,4,6,7,&16] proved that network traffic based features can be extracted and used for analysis of anomalies from network traffic. This belongs to a category of IDS called as misuse detection. Network traffic consists of continuous stream of packets, over which an attack will be spread. Barely any activity can be detected using one or two packets, rather a group of packets belonging to a particular session only can reveal any information. The algorithm for the network traffic data collection is as shown in Algorithm 1.

A session in computer networks can be uniquely identified by using source (Internet Protocol) IP and source port pair (socket) and a destination IP and destination port pair. The question is that how much time has to be allocated for recording a session's data in order to identify any attacks or normal activity information. In order to maintain a real-time nature of attack detection a 2 seconds time frame is ideal to detect any meaningful network activity. 21 essential network traffic features were identified as shown in Table 2, which can help to detect most common types of attacks as shown in Table 1. Denial of Service (DoS) and Distributed Denial of Service (DDoS) occupies the major share in today's cyber threats. Account hijacking through *tftp brute force* and Metasploit exploits like the *SMB 040* and *XP2Vncinject* were also considered. Figure 1 shows the block diagram of the process of data construction, feature selection, and evaluation using a classifier. Four types of decision tree classifiers was used to evaluate the proposed feature selection and IDS evaluation. As shown in Table 1, 8368 samples were randomly selected for training and testing and to generate a model for the proposed IDS. Low Orbit Ion Canon (LOIC) tool was used to generate the DoS and DDoS attacks. Metasploit framework was used to generate the account break-in attack.

Table. 1: Common Network Intrusions Considered in this work.

S. No	Attack Type	No of Random Samples Selected for Evaluating NFERTID
1	Normal Data	1500
2	DOS TCP	946
3	DOS UDP	986
4	DOS Normal	793
5	DOS HPingFlood	499
6	DDOS TCP	974
7	DDOS HTTP	999
8	Account Hijacking Through Brute Force	243
9	Probe attacks using Nmap	496
10	Account Hijacking Through SMB 040 Exploit	432
11	Account Hijacking Through XP2Vncinject	500
	Total	8368

Algorithm 1:
Objective: Network-Traffic-Feature-Data-Collection. To create a session table within a *n* Sec time frame
Input: Incoming packets from the network.
Output: A session table that continuously updates with the Extracted features

```
Begin
    Initialize timer = 0 seconds.
    ST = Ø
    ∀ packet ∈ network traffic do {
        SID = computeSessionID(packet)

        if (session exists in ST) do {
            // update session features in ST
            updateSessionTable(ST, packet)
        }
        else do {
            session = createSession(packet)
            addToSessionTable(ST, session)
        }
    }

    if timer seconds > n do {
        flushToFile(fileName, ST)
        timer = 0
    }
End
```

Table.2: Network Traffic Features Extracted

Feature No	Feature name	Description	Type (D: Discrete, C: Continuous)
1	Source Port	Source Port	D
2	Destination Port	Destination Port	D
3	Source Packets	No. of Packets from Source to Destination	C
4	Destination Packets	No. of Packets from Destination to Source	C
5	Source Bytes	No of bytes from Source to Destination	C
6	Destination Bytes	No of bytes from Destination to Source	C
7	Protocol	Type of Protocol (IP/TCP/UDP/ICMP)	D
8	SYN	No packets with SYN flag Set	C
9	ACK	No packets with ACK flag Set	C
10	FIN	No packets with FIN flag Set	C
11	PSH	No packets with PSH flag Set	C
12	RST	No packets with RST flag Set	C
13	URG	No packets with URG flag Set	C
14	SYN+ACK	No packets with SYN+ACK flag Set	C
15	Duration	Time in ms the session is alive	C
16	Is Alive	Set to 1 if the session did not end within the 2 Sec Time Frame.	D
17	Service	Type of Service being used	D
18	Unique Destinations	No of Unique Destinations (IP's) from a source as that of in the Current packet	C
19	Unique Sources	No of Unique Sources(IP's) with a destination as that of the current packet	C
20	Unique Destination Sockets	No of Unique Destination Sockets (IP + Port Pair) from a source as that of in the Current packet	C
21	Unique Source Sockets	No of Unique Source Sockets (IP + Port Pair) with a destination as that of the current packet	C

3. Fuzzy-Entropy & Ant Colony Optimization for real-time IDS feature selection

3.1 Fuzzy-Entropy

Fuzzy-Entropy is based on fuzzy-rough equivalence relation, which is different from the rough set equivalence relation. A small background is necessary in-order to understand the fuzzy-rough equivalence relation based fuzzy-entropy. Let $\mathcal{J}$ be an “Information System”, which is constituted by set of Objects $\mathcal{O}$ and a set of attributes $\mathcal{A}$, $\mathcal{J} = (\mathcal{O}, \mathcal{A})$. The fuzzy-equivalence of any attribute $\mathscr{r} \in \mathcal{A}$ is given by $\mathcal{O}/IND(\mathscr{r})$. The set $\mathcal{O}/IND(\mathscr{r})$is a partition induced by $\mathscr{r}$, and defined as $\mathcal{O}/IND(\mathscr{r}) = \{\mathscr{r}_1, \mathscr{r}_2\}$, where $\mathscr{r}_1 \& \mathscr{r}_2$ are the 2 fuzzy sets on 2 membership functions[15]. Say if, $\mathcal{Z}$ is a set subset of attributes such that $\mathcal{Z} = \{z_1, z_2\}$, the fuzzy equivalence class $\mathcal{O}/IND(\mathcal{Z})$, also represented in simple term as $\mathcal{O}/\mathcal{Z}$, consists of objects that are indiscernible with respect to z_1 & z_2.

$$\mathcal{O}/IND(\mathcal{Z}) = \mathcal{O}/\mathcal{Z} = \otimes \{ \mathscr{b} \in \mathcal{Z} \mid \mathcal{O}/IND(\{\mathscr{b}\}) \} \quad (1)$$

If $\mathcal{Z} = \{\mathscr{g}, \mathscr{h}\}$, $\mathcal{O}/IND(\mathscr{g}) = \{\mathscr{g}_1, \mathscr{g}_2\}, \mathcal{O}/IND(\mathscr{h}) = \{\mathscr{h}_1, \mathscr{h}_2\}$, then

$$\mathcal{O}/\mathcal{Z} = \{\mathscr{g}_1 \cap \mathscr{h}_1, \mathscr{g}_1 \cap \mathscr{h}_2, \mathscr{g}_2 \cap \mathscr{h}_1, \mathscr{g}_2 \cap \mathscr{h}_2 \} \quad (2)$$

According to fuzzy set operations, $a_1 \cap a_2 \cap a_3 = \min(a_1, a_2, a_3)$.

Shannon[17] has defined the Information-Entropy (IE), for ‘m’ outputs, of and event $\mathcal{E}$, as

$$\mathfrak{H}(\mathcal{E}) = -\sum_{k=0}^{m} p_k \, log_2 \, p_k \quad (3)$$

The attribute set $\mathcal{A}$ is a union of both conditional as well as decision attributes. $\mathcal{A} = \{\mathcal{A}_\mathcal{C} \cup \mathcal{A}_\mathcal{D}\}$. The Fuzzy-Entropy of any fuzzy membership function $\mathcal{M}_i$, among the subsets of fuzzy membership functions $\mathcal{M}_1, \mathcal{M}_2, \ldots . \mathcal{M}_m$ of an attribute is given by:

$$\mathfrak{H}(\mathcal{M}_i) = -\sum_{\mathbb{Q} \in \mathcal{O}/\mathcal{A}_\mathcal{D}}^{m} p(\mathbb{Q}| \, \mathcal{M}_i) \, log_2 \, p(\mathbb{Q}|\mathcal{M}_i) \quad (4)$$

Whereas, $p(\mathbb{Q}|\mathcal{M}_i)$ is the decision-relative probability $p(\mathbb{Q}|\mathcal{M}_i\,) = \;|\mathbb{Q} \cap \mathcal{M}_i|/|\mathcal{M}_i\,|$

$\text{Fuzzy} - \text{Entropy of } \mathcal{Z}, a\ subset\ of\ attributes\ is\ given\ by$

$$\mathbb{FE}(\mathcal{Z}) = \sum_{\mathcal{M}_i \in \mathcal{O}/\mathcal{Z}} \frac{|\mathcal{M}_i|}{\sum_{Y_i \in \mathcal{O}/\mathcal{Z}} |Y_i|} \mathfrak{H}(\mathcal{M}_i) \quad (5)$$

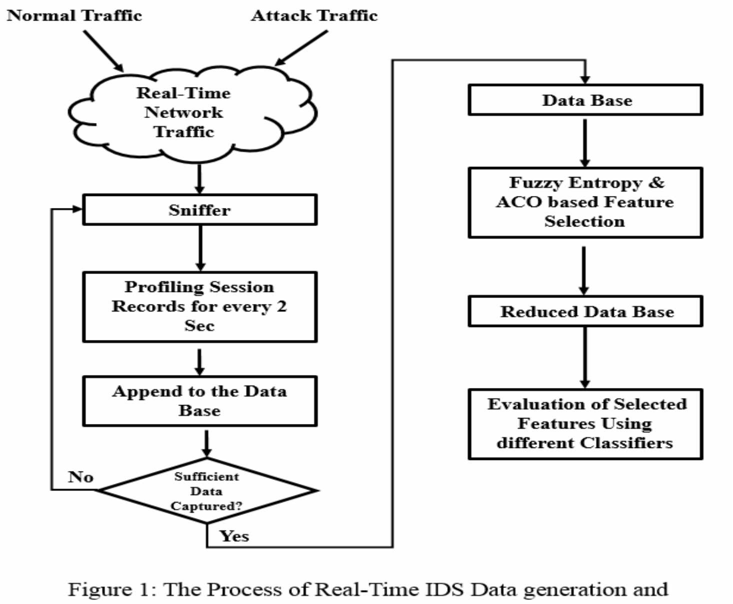
Normal Traffic
Attack Traffic
Real-Time Network Traffic
Sniffer
Profiling Session Records for every 2 Sec
Append to the Data Base
Sufficient Data Captured?
No
Yes
Data Base
Fuzzy Entropy & ACO based Feature Selection
Reduced Data Base
Evaluation of Selected Features Using different Classifiers

Figure 1: The Process of Real-Time IDS Data generation and evaluation.

3.2 Ant Colony Optimization with Fuzzy-Entropy as Heuristic

Artificial Ant Colony Optimization (ACO) developed by Dorigo et al.[18, 19] was inspired by the natural behaviour of ant swarms. Ants in search of their food source travel from one place to another in groups by using their special ability called pheromone search. Ants while travelling will release a chemical substance called pheromone in their path. The ants that follow will select their path based on the concentration of this pheromone, more of it, more the chance to select that path. Finally the pheromone concentration will be gathered towards the best path and all ants follows that path. In artificial ACO two important parameters are used in the solution construction, the pheromone as well as the heuristic. Heuristic depends on the problem domain. In this paper the problem is to select the global best real-time IDS attributes, using fuzzy-entropy. Therefore the relative-fuzzy-entropy is employed as heuristic factor for this domain.

A set of ants are released to find a local best solution collectively for an iteration for x number of iterations. At the end of all iterations the global top solution is identified. In an iteration, an ant selects next node, which is nothing but a data attribute for our problem domain, from the current position based on a probability equation.

$\mathcal{P}^{\mathrm{n}}_{a,\mathscr{b}}(\mathrm{itr}_i)$ is the probability that n^{th}ant in i^{th}iteration chooses an attribute from $a\ to\ \mathscr{b}$. $\gamma$is a constant that can tune contribution of pheromone, and σ is a constant that can tune contribution of heuristic. The heuristic value from an attribute $a\ to\ \mathscr{b}$is given by:

$$\eta(a,\mathscr{b}) = \mathbb{FE}(\mathcal{A}_{\mathrm{n}}) - \mathbb{FE}(\mathcal{A}_{\mathrm{n}} \cup \mathscr{b}) \tag{6}$$

$\mathcal{A}_{\mathrm{n}}$is the set of attribute already covered by ant n.

$$\mathcal{P}^{\mathrm{n}}_{a,\mathscr{b}}(\mathrm{itr}_i) = \frac{\tau^{\gamma}_{a,\mathscr{b}}\eta^{\sigma}_{a,\mathscr{b}}(\mathrm{itr}_i)}{\sum_{\mathcal{B}\in(\boldsymbol{\mathcal{A}_{C}}-\mathcal{A}_{\mathrm{n}})}\tau^{\gamma}_{a,z}\eta^{\sigma}_{a,z}(\mathrm{itr}_i)}, \mathscr{b} \in (\boldsymbol{\mathcal{A}_{C}} - \mathcal{A}_{\mathrm{n}}) \tag{7}$$

Algorithm.2: ACO feature selection using relative fuzzy-entropy as heuristic:
Input: $\mathcal{I} = (\mathcal{O},\mathcal{A})$, $\mathcal{A} = \{\mathcal{A}_{C} \cup \mathcal{A}_{D}\}$, Information/Data table of objects and attributes
Output: The smallest Attribute-Reduct, $\mathbb{r}$ and its cardinality $\mathbb{l}$

```
Fuzzification done for the Information/Data table $\mathcal{I} = (\mathcal{O},\mathcal{A})$
$\mathbb{r} = \{\mathcal{A}_{C}\}$ and $\mathbb{l} = |\mathcal{A}_{C}|$
$\mathbb{c}$ = 0, initialize iteration-counter.
while($\mathbb{c} < \mathbb{c}_{max}$ )
{
    $\forall\ ant$, n
    {
        $initialize\ \mathcal{K}_{\mathrm{n}} = \{0\}$, $\mathbb{L}_{\mathrm{n}} = 0$;
        if($\mathbb{c}$ == 0)
            Random-select attribute $\mathfrak{r}_{\mathrm{n}} \in \{\mathcal{A}_{C}\}$
        else
                Use equation (7) to select attribute $\mathfrak{r}_{\mathrm{n}}$
        $\mathcal{K}_{\mathrm{n}} = \{\mathfrak{r}_{\mathrm{n}}\}$, $\mathbb{L}_{\mathrm{n}} = 1$;
        do
         {
              $\mathfrak{F} = False$ , $\mathfrak{X}$ = 0
              Say the last attribute of $\mathcal{K}_{\mathrm{n}}$ is $\mathfrak{l}_{\mathrm{n}}$
              $\forall$ attribute $\mathfrak{a}_{\mathrm{n}} \in \{\mathcal{A}_{C}$ -$\mathcal{K}_{\mathrm{n}}\}$
              {
                    compute $\eta(\mathfrak{l}_{\mathrm{n}}, \mathfrak{a}_{\mathrm{n}})$using equation (6)
                    if($\eta(\mathfrak{l}_{\mathrm{n}}, \mathfrak{a}_{\mathrm{n}})$ > 0)
                    $\mathfrak{X} = \mathfrak{X}$ + [ $\eta^{\sigma}_{\mathfrak{l}_{\mathrm{n}},\mathfrak{a}_{\mathrm{n}}}\ \tau^{\gamma}_{\mathfrak{l}_{\mathrm{n}},\mathfrak{a}_{\mathrm{n}}}$]
               }
              if($\mathfrak{X}$ != 0)
              {
                      In (7) replace $\mathfrak{X}$ as denominator & select next attribute $\mathfrak{a}_{\mathrm{n}}$
                $\mathfrak{F} = True$
                  $\mathcal{K}_n = \{ \mathcal{K}_n \cup \mathbb{a}_n \}$
                  $\mathbb{L}_n = \mathbb{L}_n + 1$
              }
          }while($\mathfrak{F} = True$);
        Update-Pheromone:
            $\forall\ \mathbb{i}, \mathbb{j} \in \mathcal{K}_n$
                $\tau(\mathbb{i}, \mathbb{j}) = \tau(\mathbb{i}, \mathbb{j}) + (q/|\mathcal{K}_n|)$
        if ($\mathbb{L}_n < \mathbb{l}$)
        {
            $\mathbb{r} = \mathcal{K}_n$;
            $\mathbb{l} = \mathbb{L}_n$;
        }
    }
    $\mathbb{c} = \mathbb{c} + 1$;
    Evaporate-Pheromone:
        $\forall\ \mathbb{i}, \mathbb{j} \in \mathcal{A}_{\mathcal{C}}$
            $\tau(\mathbb{i}, \mathbb{j}) = \tau(\mathbb{i}, \mathbb{j}).\rho$
}
Output $\mathbb{r}$ & $\mathbb{l}$
```

4. Experimental results and discussion

The experiments were conducted on a windows 7 based PC which has 3 GB RAM and i3 processor. Java Programming was used to implement the algorithm. The ACO free parameters selected for the experiments are, $\sigma = 1, \gamma = 0.01, \rho = 0.9, q = 0.1, initial\ pheromone = 0.5, heuristic\ restricting\ constant = 0.001.$ Triangular membership function with shoulders is used in Fuzzification process.

The proposed algorithm was first tested on standard benchmark data set from UCI[20]. The algorithm was best suited for real valued data sets. The algorithm was run for three ants and three iterations for *iris data set* which consists of four attributes. The outputs of second iteration for a sample are shown below. It can be observed that the Ant number 2 has produced the local best or the iteration best solution. Finally the best solution of all the iterations is considered as the global best solution. In the case of *Iris* data the feature numbers 3 and 4 are selected by our algorithm, and the classification accuracy was tested on J48 decision tree algorithm and was found to be 96% with all the attributes and 96% with only two attributes. Similarly, for UCI *Cleveland data set* which contains 13 features our algorithm has identified 6 best features which are 2, 3, 6, 7, 9, and 13. When it comes to classification accuracy of *Cleveland data set* 52.2% was obtained when all the 13 features were present and an accuracy of 53.5% was obtained with the selected 6 features.

```
Data set: Iris, Iteration Number: 2, Total Ants released = 3
Solution Construction of Ant-1:
First Attribute covered by Ant-1:   1; Fuzzy-Entropy: 1.484478213609575
Next Attribute covered by Ant-1:  2; Fuzzy-Entropy: 1.5156788512812662
Next Attribute covered by Ant-1:  3; Fuzzy-Entropy: 1.3766725438292826
Next Attribute covered by Ant-1:  4; Fuzzy- Entropy: 1.3381856873930147
Attribute Chosen by Ant-1: 4
Ant-1 Covered So Far: 1, 4
Ant-1 covered: 1, 4 Fuzzy-Entropy: 1.3381856873930147
Next Attribute covered by Ant-1: 2; Fuzzy-Entropy: 1.385095764066342
Next Attribute covered by Ant-1: 3; Fuzzy-Entropy: 1.3060895061390123
Attribute Chosen by Ant-1: 3
Ant-1 Covered So Far: 1, 4, and 3
Ant-1 covered: 1, 4, 3; Fuzzy-Entropy: 1.3060895061390123
Next Attribute covered by Ant-1: 2; Fuzzy-Entropy: 1.34962844693125
```

Output of Ant-1: 1, 4, and 3.

Solution Construction of Ant-2:
First Attribute covered by Ant-2: 4; Fuzzy-Entropy: 1.271447825647451
Next Attribute covered by Ant-2: 1; Fuzzy-Entropy: 1.3381856873930147
Next Attribute covered by Ant-2: 2; Fuzzy-Entropy: 1.3675786444761084
Next Attribute covered by Ant-2: 3; Fuzzy-Entropy: 1.255311115925029
Attribute Chosen by Ant-2: 3
Ant-2 Covered So Far: 4, 3
Ant-2 covered: 4, 3; Fuzzy-Entropy: 1.255311115925029
Next Attribute covered by Ant-2: 1; Fuzzy-Entropy: 1.3060895061390123
Next Attribute covered by Ant-2: 2; Fuzzy-Entropy: 1.3296198664289236
Attribute Chosen by Ant-2: None
Output of Ant-2: 4 and 3.

Solution Construction of Ant-3:
First Attribute covered by Ant-3: 1; Fuzzy-Entropy: 1.484478213609575
Next Attribute covered by Ant-3: 2; Fuzzy-Entropy: 1.5156788512812662
Next Attribute covered by Ant-3: 3; Fuzzy-Entropy: 1.3766725438292826
Next Attribute covered by Ant-3: 4; Fuzzy-Entropy: 1.3381856873930147
Attribute Chosen by Ant-3: 4
Ant-3 Covered So Far: 1 and 4
Ant-3 covered: 1, 4; Fuzzy-Entropy: 1.3381856873930147
Next Attribute covered by Ant-3: 2; Fuzzy-Entropy: 1.385095764066342
Next Attribute covered by Ant-3: 3; Fuzzy-Entropy: 1.3060895061390123
Attribute chosen by Ant-3: 3
Ant-3 Covered So Far: 1, 4, and 3
Ant-3 covered: 1, 4, 3; Fuzzy-Entropy: 1.3060895061390123
Next Attribute covered by Ant-3: 2; Fuzzy-Entropy: 1.34962844693125
Attribute chosen by Ant-3: None
Output of Ant-3: 1, 4, and 3

Table.3: Evaluation of Feature Selection Algorithm using different classifiers.

Features/Attributes	**J48**		**Random Tree**		**Random Forest**		**JRIP**		**Average**
	Accuracy	Time	Accuracy	Time	Accuracy	Time	Accuracy	Time	Accuracy
All (21)	99.67%	0.34 s	99.48%	**0.08 s**	**99.83%**	1.21 s	99.34%	3.97 s	99.58%
1, 2, 3, 5, 6, 8, 9, 11, 12, 18, 19, 20, and 21 (Total 13 features)	99.59%	0.2 s	99.75%	**0.06 s**	**99.81%**	0.88 s	99.64%	1.77 s	99.69%
How much Faster Model was generated after feature reduction?	41.17% faster		25% faster		27.21% faster		55.41% faster		37.19% faster
Percentage Change in Accuracy	-0.08%		+ 0.27%		-0.02%		+0.30%		+ 0.11%

Table 3 shows the classification accuracies and time taken by the model generation of the respective classifiers for full feature real-time IDS data and also for the reduced feature real-time IDS data. Four classifiers were compared here J48, Random Tree, Random Forest and JRIP rules. 50% split was used for training and testing. Weka 3.6 was used to evaluate the performance of real-time IDS before and after feature reduced. The table shows the list of features that were identified as relevant according to the proposed Fuzzy-Entropy based ACO search algorithm. The percentage change in the accuracy before and after feature reduction is also indicated, it shows that almost all the classifiers produced near full attribute accuracy even after feature reduction. Among the four, Random Forest classifier produced a 0.27% increase of accuracy after reduction. The advantage of feature reduction can be clearly found by looking at the model generation time before and after reduction. The model generation time after feature reduction, is 37.19% faster compared to full feature data set and without losing any bit of classification accuracy.

5. Conclusion

This paper proposed a real-time IDS which can detect most common type of intrusions like DoS, DDoS, probing, and account hijacking. 21 Network traffic attributes were identified and data set for training and testing was constructed. In order to reduce the training, testing and classification times and optimize the data set feature selection using fuzzy-entropy and ACO was proposed. The proposed algorithm was first tested on standard UCI data sets and found to be effective. Top 13 features were identified from the 21 features of real-time IDS data and the same was evaluated using four popular classification algorithms. This research presents a simple and faster way of detecting real-time intrusions on computer networks.

References

1. E. D. Dorothy, An Intrusion-Detection Model, IEEE Transactions on software engineering, 1987:**13**:222-232.

2. M. John, C. Alan and A. Julia, "Defending Yourself: The role of Intrusion Detection Systems," IEEE Software, 2000:**17**(5):42-51.

3. W. Lee and S. J. Stolfo, "A Framework for constructing features and models for intrusion detection systems," ACM Transaction on Information and System Security, 2000:**3**(4): 227-261.

4. K. Labib and R. Vemuri, "NSOM: A Real-Time Network Based Intrusion Detection System Using Self-Organizing Maps," Department of Applied Science, University of California, Davis, 2002.

5. Ming-Yang Su, Gwo-Jong Yub and Chun-Yuen Lina, "A real-time network intrusion detection system for large-scale attacks based on an incremental mining approach," Computers & Security, 2009:**28**:301-309.

6. P. Sangkatsanee, N. Wattanapongsakorn and C. Charnsripinyo, "Practical real-time intrusion detection using machine learning approaches," Computer Communications, 2011:**34**:2227-2235.

7. Ming-Yang Su, "Real-time anomaly detection systems for Denial-of-Service attacks by weighted k-nearest-neighbour classifiers," Expert Systems with Applications, 2011:**38**:3492-3498.

8. Ravi Kiran Varma. P and Valli Kumari. V, "A Security Framework For Ethernet Based Embedded Web Server," International Journal of Embedded Systems and Applications, 2012:**2**(2):17-27.

9. Ravi Kiran Varma. P, Valli Kumari. V and Srinivas Kumar. S, "A Novel Rough Set Attribute Reduction Based on Ant Colony Optimization," International Journal of Intelligent Systems Technologies and Applications, 2015:**14(3/4)**:330-353

10. Ravi Kiran Varma. P and V. Valli Kumari, "Feature Optimization and Performance Improvement of a Multiclass Intrusion Detection System Using PCA and ANN," International Journal of Computer Applications, 2012:**44**(13):4-9.

11. Nandita Sengupta, Jaydeep Sen, Jaya Sil and Moumita Saha, "Designing of online intrusion detection system using rough set theory and Q-learning algorithm," Neuro Computing, 2013:**111**:161-168.

12. Nguyen H S, "Discretization problem for Rough Set Methods," RSCTS98, LNAI, 1998:**1424**:545-552.

13. Pramod Kumar P, Prahlad Vadakkepat and Loh Ai Poh, "Fuzzy-rough discriminative feature selection and classification algorithm, with application to microarray and image datasets," Applied Soft Computing, 2011:**11**:3429-3440.

14. Richard Jensen and Qiang Shen, "Fuzzy-rough data reduction with ant colony optimization," Fuzzy Sets and Systems, 2005:**149**:5-20.

15. Neil Mac Parthal´ain, Richard Jensen and Qiang Shen, "Fuzzy Entropy-Assisted Fuzzy-Rough Feature Selection," in Proc. 15th Int Conf Fuzzy Systems, Vancouver, 2006.

16. Jonathan J Davis and Andrew J Clark, "Data pre-processing for anomaly based network intrusion detection: A review," Computers & Security, 2011:**30**:353-375

17. C. E. Shannon, "A Mathematical Theory of Communication," Bell System Technical Journal, 1948:**27**(3):379-423.

18. M. Dorigo, V. Maniezzo and A. Colorni, "The Ant System, Optimization by a Colony of Cooperating Agents," IEEE Transactions on Systems, Man, and Cybernetics- Part B, 1996:**26**(1):29-41.

19. M. Dorigo and D. G. Caro, "Ant Colony Algorithm for the Travelling Salesman Problem," BioSystems, 1997:73-81.

20. M. Lichman, "{UCI} Machine Learning Repository," University of California, Irvine, School of Information and Computer Sciences, 2013. [Online]. Available: http://archive.ics.uci.edu/ml. [Accessed 2015].