International Journal on Advanced Science Engineering Information Technology

Vol.7 (2017) No. 1
ISSN: 2088-5334

A Novel DNA Sequence Approach for Network Intrusion Detection System Based on Cryptography Encoding Method

Omar Fitian Rashid, Zulaiha Ali Othman, Suhaila Zainudin

Faculty of Information Science and Technology, University Kebangsaan Malaysia, Bangi, Malaysia
Email: Omaralrawi08@yahoo.com, zao@ukm.edu.my, and suhaila.zainudin@ukm.edu.my

***Abstract*— A novel method for Network Intrusion Detection System (NIDS) has been proposed, based on the concept of how DNA sequence detects disease as both domains have similar conceptual method of detection. Three important steps have been proposed to apply DNA sequence for NIDS: convert the network traffic data into a form of DNA sequence using Cryptography encoding method; discover patterns of Short Tandem Repeats (STR) sequence for each network traffic attack using Teiresias algorithm; and conduct classification process depends upon STR sequence based on Horspool algorithm. 10% KDD Cup 1999 data set is used for training phase. Correct KDD Cup 1999 data set is used for testing phase to evaluate the proposed method. The current experiment results show that the proposed system has obtained good results and these results are equal to 86.36%, 49.69%, and 77.65% for detection rate, false alarm rate, and accuracy respectively. These results are considered as a better result when it is compared with the other previous basic algorithms. It is possible to conclude that DNA sequence has the potential for NIDS solution and it has potential improvement using a better encoding method.**

***Keywords*— DNA, Horspool algorithm, network intrusion detection system, Teiresas algorithm**

I. INTRODUCTION

Intrusion detection is used to detect and prevent attacks [1] and secure the network as the Deoxyribonucleic Acid (DNA) used to detect disease in the human body. The concept of detection of abnormal functioning of tissues by DNA sequence can be applied to computing systems where the normal functioning of the system can be determined by DNA sequence that differs from DNA sequence of attack. Various DNA encoding methods are used in many reliable computer system techniques such as cryptography, steganography, and a digital signature that have been developed through using various DNA encoding methods.

A system to encrypt and generate digital signature has been built that is based on DNA Cryptography to handles the combination of all characters with superb accuracy [2]. To hide secret data in an image, [3] established a system which is depending on the transfer the image into two security layers, one is DNA sequence and the second is a covered layer. The DNA steganography and RGB colors have been used to create a cryptographic system [4]. This system is depending on the conversion of cipher text to color by supplying cipher DNA sequence.

The DNA sequence can be converted into fingerprint image through using wavelets transform function [5]. The DNA decoding is applied to extract the signature from the watermark image, which is become invisible. Any type of data such as text, image, audio or video can be encrypted via DNA cryptography [6]. Image encryption depending on DNA cryptography and hill cipher is established [7]. In addition, the encryption system is provided that depends upon the using of DNA and a key length of 256 bits [8].

Detection of intrusion can be performed by a system provided by [9] where two grains detection level is used. The coarse grained is detected the intrusion, and the details are done by the fine-grained system. The IDS system for IEEE 802.11 wireless network is suggested by [10] which are established upon behavioural analysis and through applying sequential machine learning techniques. The pattern in the protocol is modelled and characterized the probabilities. The MCLP classifier is improved by the proposed IDS that is carried out by multiple criteria of linear programming and swam optimization [11].

The K-means method, cuttlefish algorithm and five rules algorithms are applied to various numbers of clusters, in order to lower the features number and implement high detection rate and to decrease the false detecting attack [12], [13]. To improve the efficiency of the IDS which is depending on extreme learning machine, a framework is established by [14] that is combined the outputs of simple learners. The hybrid IDS proposed by [15] for a sensor network is used to reduce the communication costs is depending upon the support vector machine algorithm and signature rules. Good performance and maximize the

production of the IDS is achieved by [16] through establishing a system that applies both negative and exclusive pattern matching techniques. The fast heuristic clustering method has been applied to establish a novel intrusion detection system that is based on data mining technique [17].

To discover unknown attacks, [18] proposed an adaptive method depend on ant colony clustering. The method is focused on the clustering process of an ant colony movement. The structure of the intrusion detection system is designed, based on ant colony clustering. This can not only improve the detection rate but also reduce false positive rate significantly, and can automatically detect various kinds of attacks. Is important to prevent sensitive data from attack, like prevent privacy in personal communication. Therefore this system implemented to enables the data owner to detect and prevent data leak. The results show that approach can provide accurate detection with a small number of false alarm [19]. Suggest a step to reshape the policy in order to develop a data protection that leads to creating better confidence for the user, therefore adopted a survey questionnaire methodology by clients [20].

An Extreme Learning Machine-based intrusion detection method for Advanced Metering Infrastructure is presented [21]. Firstly, the method filter and partition the malicious data, and different types of invasion are effectively extracted. Finally, Extreme Learning Machine is used to detect the various attack types of malicious data. ELM tends to have better scalability, and much better generalization performance is achieved at much faster learning speed than traditional SVM. Xing-zhu [22] improved the neural network model for network intrusion detection. The network feature subset and parameters of the Radial Basis Function neural network are regarded as a particle. Then, to establish the optimal network intrusion detection model, collaboration and information exchange between particles and the optimal feature subset and parameters of Radial Basis Function neural network are found. The simulation results showed that this system reduced the feature dimensions, and the best parameters of the Radial Basis Function neural network is obtained which, is a kind of network intrusion detection model with high detection accuracy and high speed.

Promod and Jacob [23] applied the Random Forest to measure the intrusion of unauthorized personnel to certain designated areas of the organization. The system of time attendance acts as a security system as it involves access to doors and barriers through which only authorized personnel should access. The Random Forest classifier is used to build a model for intrusion detection system [24] The Random Forest is an ensemble classifier and performs well compared to other traditional classifiers for effective classification of attacks. The obtained empirical results indicated that the presented model is efficient with low false alarm rate and high detection rate.

The current paper is presenting a new procedure that can be used to detect the intrusion detections based on cryptography DNA encoding approach. Also, the results are compared with previously published results in this field of work.

II. Material and Methods

DNA is the genetic material that existed in most organisms (include human being). It has the advantage of storage of information in the long term. This information is saved as a code made up of four chemical bases, called Adenine, Cytosine, Thymine, and Guanine and they are referred as A, C, T, and G, to form base pairs that attached to a sugar molecule and a phosphate molecule, DNA structure is show in Fig. 1 [25] Nucleotides are the base pair, make up two long spiral strands connected together, based on these base pairs. There are about 3 million bases, 99% of these pairs are similar to all persons, and only 1% is unique. DNA cells include genetic information, shared in human through chromosomes where a total of 46 chromosomes are found, 23 from the father and 23 from the mother. The offspring sharing 99.7% with their parents and only 0.3% is the unique code (repetitive coding) that causes DNA to be as biometrics.

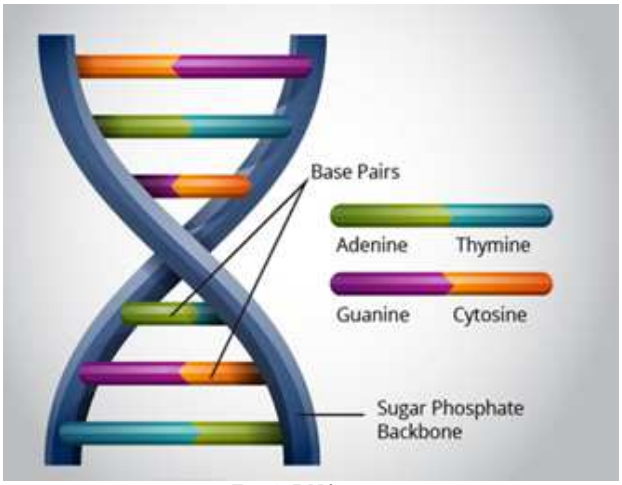

Fig. 1 DNA structure

In DNA, when a pattern of two or more nucleotides is repeated, it is called Short Tandem Repeats (STR), and they are directly adjacent to each other [26]. For example; the sequence ACTT is repeated three times in A-A-A-C-T-T-A-C-T-T-A-C-T-T-A-G. Such repeats are used in the investigation to look for certain particular areas of DNA that make the search much easier than looking at all the DNA sequence.

These special areas of the DNA are believed to be parts that do not code for any genes (non-coding sequences), but they can be changed in various people. Identical repeats of the same pattern exist with the length of 2 to 6 base pairs of DNA, and they can be found anywhere from 1 to 50 times in a row. For example, the sequence "A-C-C-A-C-C-A-C-C-A-C-C" where A-C-C is repeated 4 times as shown in Fig. 2 [27].

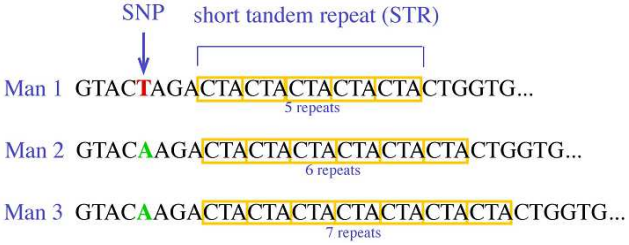

Fig. 2 Exhibited an example of the Short Tandem Repeats

Cryptography is a branch of study within the field of cryptology. The original message is called the plaintext while the coded message is called the cipher text. Encryption and decryption are the processes of converting plaintext to cipher text and vice versa [28]. Encryption algorithms can be grouped into two types which are stream ciphers and block ciphers. In stream ciphers, the image-pixels or text-character are encrypted consecutively, and in block ciphers, blocks of bits or blocks of characters are used [29]. This work is carried out by using Teiresias and Horspool algorithms.

Teiresias algorithm can be used to detect and report all existing patterns in a set of input sequences without using alignment. Let Σ be the alphabet of residues (e.g. the set of the whole DNA sequences), where a regular expression of the form Σ (Σ U {'.'}) Σ is defining a pattern, and the symbol '.' is used to mark a position that can be an arbitrary residue. Every pattern P defines a language G(P) that is consisting of all strings which can be obtained from P by replaced "each don't care" by " an arbitrary residue" from Σ. For example, the pattern "T.AA..C", the following peptides are elements of G ("T.AA..C"): TAAAGCC, TCAAGTC, TTAATGC. For any pattern P, each substring of P that is itself a pattern is called a sub pattern of P. For example, "A..C" is a sub pattern of the pattern "T.AA..C". A pattern P is called a <L.W> pattern (with L <= W) if each sub pattern of P with length W or more contains at least L residues [30].

Horspool matching algorithm is utilized in the present research as an efficient string searching algorithm which has been used to classify the data into attack or normal. The target string (key) that is being searched is pre-processed by the algorithm. Such algorithm does not require checking each character in the searched string, but it skips some of them. However, it becomes faster when the key becomes longer. The efficiency of this algorithm is derived from the fact that each unsuccessful attempt to find a match between the search string and the text used in the searching; since it uses the information gained from that attempt to move as many positions of the text where the string cannot match [31]. Table 1 shows the bad-character table used by Horspool algorithm for the following example (Table 2) that illustrates the application of Horspool algorithm, looking for the key "GCAGAGAG", in the sequence "GCATCGCAGAGAGTATACAGTACG".

TABLE I
BAD-CHARACTER TABLE USED BY HORSPOOL ALGORITHM

A	A	C	G	T
Bc [A]	1	6	2	8

TABLE III
EXAMPLE OF HORSPOOL ALGORITHM

First Attempt

G	C	A	T	C	G	C	A	G	A	G	A	G	T	A	T
							1								
G	C	A	G	A	G	A	G								

Shift by: 1 (**Bc**[A])

Second Attempt

G	C	A	T	C	G	C	A	G	A	G	A	G	T	A	T
	2							1							
	G	C	A	G	A	G	A	G							

Shift by: 2 (**Bc**[G])

Third Attempt

G	C	A	T	C	G	C	A	G	A	G	A	G	T	A	T
			2							1					
			G	C	A	G	A	G	A	G					

Shift by: 2 (**Bc**[G])

Fourth Attempt

G	C	A	T	C	G	C	A	G	A	G	A	G	T	A	T
					2	3	4	5	6	7	8	1			
					G	C	A	G	A	G	A	G			

The current proposed system is done based on three steps, these steps are; DNA sequence for NIDS, STR extraction, and matching process as shown in the following steps:

- Based on Encoding table, the network traffic of 10% KDD Cup dataset is converted to DNA sequence.
- Teiresias algorithm is used to extract STR (key).
- The network traffic of the correct KDD Cup dataset is converted to DNA sequence based on Encoding table.
- Horspool algorithm is used to classify network traffic as normal or attack based on STR extracted in step 2.
- Detection Rate, False Alarm Rate, and Accuracy are calculated to achieve the results.

To perform the first step, the DNA encoding to intrusion detection system based on the DNA encoding table [32] is applied as shown in Table 3. The example below illustrates how the DNA sequence is generated for the following network traffic:

(0,tcp,http,SF,181,5450,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,8,8,0.00,0.00,0.00,0.00,1.00,0.00,0.00,9,9,1.00,0.00,0.11,0.00,0.00,0.00,0.00,0.00)

TTAAATTAATTAATTAATTATTTTGTATTTTTATT
ACTTGCTATCTTGTTAAATTATTAAATTATTAAATT
ATTAAATTATTAAATTATTAAATTATTTTATTATTAA
ATTATTAAATTATTAAATTATTAAATTATTAAATTA
TTAAATTATTAAATTATTAAATTATTAAATTATTAA
ATTAGTATATTAGTATATTATTAAATCCTTAATTAA

ATTATTAAATCCTTAATTAAATTATTAAATCCTTAA
TTAAATTATTAAATCCTTAATTAAATTATTTTATCCT
TAATTAAATTATTAAATCCTTAATTAAATTATTAAA
TCCTTAATTAAATTAGTTGATTAGTTGATTATTTTAT
CCTTAATTAAATTATTAAATCCTTAATTAAATTATT
AAATCCTTTTTTTTATTATTAAATCCTTAATTAAATT
ATTAAATCCTTAATTAAATTATTAAATCCTTAATTA
AATTATTAAATCCTTAATTAAATTATTAAATCCTTA
ATTAA

TABLE IIIII
ILLUSTRATES THE DNA ENCODING METHOD

	C	A	T	G
A	ACAT – a ACTG – b ACCC – c ACGA – d	AAAA – y AATT – z AACC – A AAGG – B	ATAA – W ATTT – X ATCG – Y ATGC – Z	AGAG – { AGTA – [AGCG - } AGGG -]
T	TCAT – e TCTG – f TCCG – g TCGT – h	TAAT – C TATG – D TACC – E TAGA – F	TTAA – 0 TTTT – 1 TTCC – 2 TTGG – 3	TGAA - | TGTT - \ TGCG - + TGGC - =
C	CCAG – i CCTA – j CCCG – k CCGG – l	CAAT – G CATG – H CACG – I CAGT – J	CTAT – 4 CTTG – 5 CTCC – 6 CTGA – 7	CGAA - _ CGTT - - CGCC -) CGGG – (
G	GCAA – m GCTT – n GCCG – o GCGC – p	GAAG – K GATA – L GACG – M GAGG – N	GTAT – 8 GTTG – 9 GTCG - < GTGT - >	GGAT - * GGTG - & GGCC - ^ GGGA - %
A	ACTC – q ACCG – r	AATA – O AACG – P	ATTA - , ATCC - .	AGTT - $ AGCC - #
T	TCTC – s TCCC – t	TATC – Q TACG - R	TTTA - ? TTCG - /	TGTA - @ TGCC - !
C	CCTT – u CCCC – v	CATC – S CACC – T	CTTC - : CTCG - ;	CGTA – ~ CGCG – ‘
G	GCTA – w GCCC – x	GATT – U GACC – V	GTTC – “ GTCC – ‘	GGTC - € GGCG - £

The “KDD Cup 99 is the information source that it is used in the current research, which consists of thousands of records and each record in the dataset has 42 features, 22 of these features explain the connection and 19 features of them describe their connection properties of the same host with the last two seconds [33]. These features are shown in Table 4.

TABLE IV
LIST OF VARIOUS FEATURES OF KDD-CUP99 TASK DESCRIPTION

Number	Feature Name	Type
1	Duration	Continuous
2	protocol type	Discrete
3	Service	Discrete
4	Flag	Discrete
5	source bytes	Continuous
6	destination bytes	Continuous
7	Land	Discrete
8	wrong fragment	Continuous
9	Urgent	Continuous
10	Hot	Continuous
11	failed logins	Continuous
12	logged in	Discrete
13	# compromised	Continuous
14	root shell	Continuous
15	su attempted	Continuous
16	# root	Continuous
17	# file creations	Continuous
18	# shells	Continuous
19	# access files	Continuous
20	# outbound cmds	Continuous
21	is hot login	Discrete
22	is guest login	Discrete
23	Count	Continuous
24	srv count	Continuous
25	serror rate	Continuous
26	srv serror rate	Continuous
27	rerror rate	Continuous
28	srv rerror rate	Continuous
29	same srv rate	Continuous
30	diff srv rate	Continuous
31	srv diff host rate	Continuous
32	dst host count	Continuous
33	dst host srv Count	Continuous
34	dst host same srv rate	Continuous
35	dst host diff srv rate	Continuous
36	dst host same src port rate	Continuous
37	dst host srv diff host rate	Continuous
38	dst host serror Rate	Continuous
39	dst host srv serror rate	Continuous
40	dst host rerror Rate	Continuous
41	dst host srv r error rate	Continuous

The 10% KDD Cup 99 dataset which included 22 types of attacks are used for training as shown in Table 5, and the corrected KDD Cup 99 which included 37 types of attacks are used for testing phases as shown in Table 6.

TABLE V
CLASS LABELS THAT APPEARS IN “10% KDD” DATASET

Attack	Samples	Category
Smurf	280790	DoS
Neptune	107201	DoS
Back	2203	DoS
Teardrop	979	DoS
Pod	264	DoS
Land	21	DoS
Satan	1589	Probe
Ipsweep	1247	Probe
Portsweep	1040	Probe
Nmap	231	Probe
Warezclient	1020	R2L
Guess_passwd	53	R2L
Warezmaster	20	R2L
Imap	12	R2L
ftp_write	8	R2L
Multihop	7	R2L
Phf	4	R2L
Spy	2	R2L
Buffer_overflow	30	U2R
Rootkit	10	U2R
Loadmodula	9	U2R
Perl	3	U2R
Normal	97277	Normal

TABLE VI
CLASS LABELS THAT APPEARS IN "CORRECTED KDD" DATASET

Attack	Samples	Category
Apache2	794	DoS
Back	1098	DoS
Land	9	DoS
Mailbomb	5000	DoS
Neptune	58001	DoS
Pod	87	DoS
Processtable	759	DoS
Smurf	164091	DoS
Teardrop	12	DoS
Udpstorm	2	DoS
Mscan	1053	Probe
Nmap	84	Probe
Portsweep	354	Probe
Ipsweep	306	Probe
Saint	736	Probe
Satan	1633	Probe
Ftp_write	3	R2L
Guess_passwd	4367	R2L
Httptunnel	158	R2L
Imap	1	R2L
Multihop	18	R2L
Named	17	R2L
Phf	2	R2L
Sendmail	17	R2L
Snmpgetattack	7741	R2L
Snmpguess	2406	R2L
Warezmaster	1602	R2L
Worm	2	R2L
Xlock	9	R2L
Xsnoop	4	R2L
Buffer_overflow	22	U2R
loadmodule	2	U2R
Perl	2	U2R
Ps	16	U2R
Rootkit	13	U2R
Sqlattack	2	U2R
Xterm	13	U2R
Normal	60593	Normal

The performance of the present system is determined by three measures called detection rate (DR), false alarm rate (FAR), and accuracy [34]. The DR is the ratio of the number of correctly detected attacks over the total number of attacks. The formulae for calculating DR are shown in equations (1) as follow:

$$DR = \frac{TP}{TP+FN} \qquad (1) \qquad (1)$$

The FAR is the ratio of the number of normal connections that are incorrectly misclassified as attacks to the total number of normal connections. The formula for calculating FAR are shown in equations (2) as follow

$$FAR = \frac{FP}{TN+FP} \qquad (2) \qquad (2)$$

Accuracy is measured by calculating the ratio of the number of truly classified connections over the total number of connections. The formula for calculating Accuracy are shown in equations (3) as follow

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \qquad (3) \qquad (3)$$

III. RESULTS AND DISCUSSIONS

Table 7 exhibited the values obtained in terms of detection rate, false alarm rate and accuracy from the present system and these are compared with two novel intrusion detection systems mentioned by (Duque & Omar [12]; Yu et al., [18]). The published values are obtained from the two systems that have been applied the data mining technique to intrusion detection systems (Duque & Omar [12]; Yu et al., [18]). From the table, it is clear that the detection rate and accuracy obtained by the method of the present system are quite good. This system gives a better detection rate than the previous two systems, and the result is equal to 86.36%. The false alarm rate results for the two systems are not mentioned, and finally, our accuracy result is less than the accuracy of the second system and it equal to 77.65%. The detection rate results for the proposed system and the published one are shown in Fig. 3, The accuracy results for the proposed system and the published one are shown in Fig. 4, and the results of the proposed system are illustrated in Fig. 5.

TABLE VII
COMPARISON BETWEEN THE DR, FAR AND ACCURACY OF THE PROPOSED SYSTEM WITH THE PUBLISHED ONES

	Detection Rate	False Alarm Rate	Accuracy
Duque and Omar [12]	28.78%	-	81.61%
Yu et al., [18]	71.67%	-	-
Proposed System	**86.36%**	**49.69%**	**77.65%**

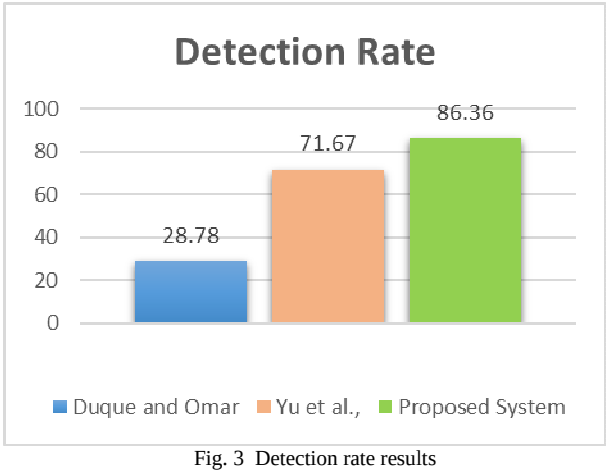

Fig. 3 Detection rate results

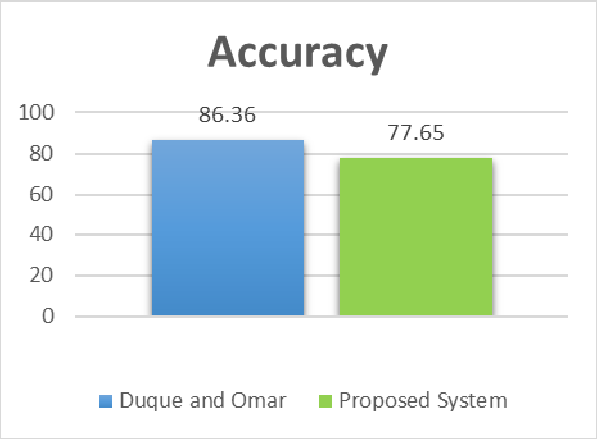

Fig. 4 Accuracy results

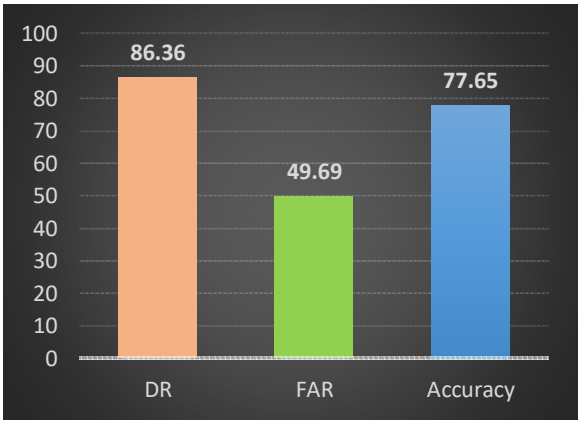

Fig. 5 The results of the present proposed system

IV. Conclusion

This paper has shown how the concepts of DNA sequence disease detection are used for network intrusion detection system. The method looks simple which consist of five steps to conduct detection process. Even though the performance is weak compared with the stated art of IDS, but the proposed method has shown its relatively good result. The performance of applying DNA sequence is very much relying on the DNA encoding techniques. The use of cryptography encoding method may not be suitable for the network. Therefore, the future suggestion is to build a suitable DNA encoding method for network IDS.

References

[1] P. Adlakha, and P. Subramanium, "Various Approaches for Detecting Attacks in Intrusion Detection System (IDS)", *IJCSMC*, vol.2, ISSN 2320-088X, March 2013.

[2] D. S. Chouhan, and R. P. Mahajan, "An Architectural Framework for Encryption & Generation of Digital Signature Using DNA Cryptography", *Computing for Sustainable Global Development (INDIACom)*, ISBN 978-93-80544-10-6, 2014.

[3] P. Das, and N. Kar, "A DNA Based Image Steganography using 2D Chaotic Map", *Electronics and Communication Systems (ICECS)*, 2014 International Conference on, ISBN: 978-1-4799-2321-2, pp. 1-5, 2014.

[4] K. A. Fasila, and D. Antony, "A Multiphase Cryptosystem with Secure Key Encapsulation Scheme Based on Principles of DNA Computing", *Advances in Computing and Communications (ICACC)*, 2014 Fourth International Conference on, ISBN: 978-1-4799-4364-7, pp. 1-4, 2014.

[5] K. K. Ghany, G. Hassan, A. Hassanien, H. Hefny, G. Schaefer, and A. R. Ahad, "A Hybrid Biometric Approach Embedding DNA Data in Fingerprint Images "*Informatics, Electronics & Vision (ICIEV)*, 2014 International Conference on, ISBN: 978-1-4799-5179-6, pp. 1-5, 2014.

[6] S. Jain, and V. Bhatnagar, "A Novel DNA Sequence Dictionary method for Securing Data in DNA using Spiral Approach and Framework of DNA Cryptography", *Advances in Engineering and Technology Research (ICAETR)*, 2014 International Conference on, ISSN: 2347-9337, Pages: 1-5, 2014.

[7] R. Jangid, N. Mohmmad, A. Didel, and S. Taterh, "Hybrid Approach of Image Encryption Using DNA Cryptography and TF Hill Cipher Algorithm", *Communications and Signal Processing (ICCSP)*, 2014 International Conference on, ISBN: 978-1-4799-3357-0, pp. 934 – 938, 2014.

[8] A. Majumder, A. Majumdar, T. Podder, N. Kar, and M. Sharmas, "Secure Data Communication and Cryptography Based on DNA Based Message Encodin", *Advanced Communication Control and Computing Technologies (ICACCCT)*, 2014 International Conference on, ISBN: 978-1-4799-3913-8, pp. 360-363, 2014.

[9] S. O. Al-Mamory, and F. S. Jassim, "On the designing of two grains levels network intrusion detection system", Karbala International Journal of Modern Science, vol. 1, Issue 1, pp. 15-25, September 2015.

[10] H. Alipour, Y. B. Al-Nashif, P. Satan, and S. Hairi, "Wireless Anomaly Detection Based on IEEE 802.11 Behavior Analysis", *IEEE Transactions on information forensics and security*, vol. 10, no. 10, October 2015.

[11] S. M. Bamakan, B. Amiri, M. Mirzabagheri, and Y. Shi, "A New Intrusion Detection Approach using PSO based Multiple Criteria Linear Programming", *3rd International Conference on Information Technology and Quantitative Management*, vol. 55, pp. 231-237, 2015.

[12] S. Duque, and M. N. Omar, "Using Data Mining Algorithms for Developing a Model for Intrusion Detection System (IDS)", *Complex Adaptive Systems San Jose*, CA, vol. 61, 2015, pp. 46-51, 2015.

[13] A. S. Eesa, Z. Orman, and A. M. Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems", *Expert Systems with Applications*, vol. 42, Issue 5, pp. 2670-2679, 1 April 2015.

[14] J. M. Fossaceca, T. A. Mazzuchi, and S. Sarkani, "MARK-ELM: Application of a novel Multiple Kernel Learning framework for improving the robustness of Network Intrusion Detection", *Expert Systems with Applications*, vol. 42, Issue 8, pp. 4062-4080, 15 May 2015.

[15] Y. Maleh, A. Ezzati, Y. Qasmaoui, and M. Mbida, "A Global Hybrid Intrusion Detection System for Wireless Sensor Networks", *The 6th International Conference on Ambient Systems*, Networks and Technologies, the 5th International Conference on Sustainable Energy Information Technology, vol. 52, pp. 1047-1052, 2015.

[16] K. Zheng, Z. Cai, X. Zhang, Z. Wang, and B. Yang, "Algorithms to Speedup Pattern Matching for Network Intrusion Detection Systems", *Computer Communications*, vol. 62, pp. 47-58, 15 May 2015.

[17] Z. Yu, J. Chen, and T. Zhu, "A Novel Adaptive Intrusion Detection System based on Data Mining", *2005 International Conference on Machine Learning and Cybernetics (Volume:4)*, ISSN : 2160-133X, 2005.

[18] Y. Qiang, H. Zhongyu, S. Shikai, and Z. Dawei, "Research of Intrusion Detection Method Based on Ant Colony Clustering", *4th International Conference on Machinery*, Materials and Computing Technology, 2016.

[19] S. Gaikwad, S. Chougule, and S. Charhate, "Detection and Prevention of Sensitive Data from Data Leak Using Shingling and Rabin Filter", *International Journal on Advanced Science, Engineering and Information Technology*, Vol. 6 (2016) No. 5, DOI:10.18517/ijaseit.6.5.997, pp. 663-667, 2016.

[20] M. H. Shaikh and N. A. Ansari, "Examining a Norwegian Client's Response over Information Security and Privacy Policy", *International Journal on Advanced Science, Engineering and Information Technology*, Vol. 5 (2015) No. 3, DOI:10.18517 /ijaseit. 5.3.499, pp. 165-169, 2015.

[21] Y. Li, C. Zhang, and L. Yang, "The Research of AMI Intrusion Detection Method using ELM in Smart Grid", *International Journal of Security and Its Applications*, vol. 10, No. 5, pp. 283-296, 2016.

[22] W. Xing-zhu, "Network Intrusion Prediction Model based on RBF Features Classification", *International Journal of Security and Its*

Applications, vol. 10, no. 4, doi: 10.1109/ICMLC.2002.1174376, pp.241-248, 2016.

[23] N. Farnaaz,, and M. A. Jabbar, "Random Forest Modeling for Network Intrusion Detection System", *Procedia Computer Science*, Volume 89, Pp: 213–217, (2016).

[24] K. V. Promod, and B. Jacob, "Mining a Ubiquitous Time and Attendance Schema using Random Forests for Intrusion detection", *Procedia Technology*, Volume 24, Pp: 1226–1231, (2016).

[25] R. Soram, and M. Khomdram, "Biometric DNA and ECDLP Based Personal Authentication System: A Superior Posse of Security", *IJCSNS International Journal of Computer Science and Network Security*, vol. 10 no.1, January 2010.

[26] E. Oki, S. Oda, Y. Maehara, and K. Sugimachi, "Mutated Gene Specific Phenotypes of Dinucleotide Repeat Instability in Human Colorectal Carcinoma Cell Lines Deficient in DNA Mismatch Repair", Oncogene 18: 2143-3247, 1999.

[27] M. A. Jobling., . [Online]. Available: http://www.le.ac.uk/ge/maj4/NewWebSurnames041008.html, 2009.

[28] B. Purnama, and H. Rohayani, "A New Modified Caesar Cipher Cryptography Method with Legible Ciphertext From A Message To Be Encrypted" *Procedia Computer Science*, vol. 59, pp. 195-204, 2015.

[29] A. Radwan, S. AbdELHaleem, and S. Abd-El-Hafiz, "Symmetric encryption algorithms using chaotic and non-chaotic generators: A review", *Journal of Advanced Research*, Volume 7, Issue 2, pp. 193-208, March 2016.

[30] I. Rigoutsos, and A. Floratos, "Combinatorial Pattern Discovery in Biological Sequences: The TEIRESIAS Algorithm, Bioinformatics", vol. 14 no. 11998, pp. 55-67, 1998.

[31] S. Nimisha, and G. Deepak, "String Matching Algorithms and their Applicability in Various Applications", *International Journal of Soft Computing and Engineering (IJSCE)*, vol. I, Issue 6, January 2012.

[32] N. H. UbaidurRahman, C. Balamurugan, and R. Mariappan, "A Novel DNA Computing based Encryption and Decryption Algorithm", *Procedia Computer Science 46*, pp. 463 – 475, 2015.

[33] W. Chimphlee, M. Mohd Noor, A. Abdul Hanan, and C. Siriporn,,"Anomaly Network Intrusion Detection Method in Network Security Based on Principle Component Analysis", *Journal Teknologi Maklumat*, 18 (2). pp. 114-124, 2006.

[34] S. Wu, and W. Benzhaf, "The Use of Computation Intelligence in Intrusion Detection Systems", *Applied Soft Computing*, pp. 1–35, vol. 10, Issue 1, January 2010.