International Journal of Computer Applications Technology and Research
Volume 6–Issue 11, 441-450, 2017, ISSN:-2319–8656

Multi-Level Image Steganography Using Compression Techniques

Mohamed H. Sayed
The National Ribat University
Khartoum, Sudan

Talaat M.Wahby
Sudan University of
Science and Technology
Khartoum, Sudan

ABSTRACT: Steganography is the art and science of writing hidden Messages in such a way that no one, apart from the sender and intended recipient, Suspects the existence of the message. In this research apply Multi-Level Steganography for image steganography was presented. MLS consists of at least two stenographic methods utilized respectively. Two-levels of stenography have been applied; the first level is called (the upper-level), and it has been applied using enhance LSB (secure LSB-L1) image steganography, the secret data in this level is English text, and the cover is Bitmap image, the output is a stego_image called (intermediate image). The second level is called (the lower-level); it has been applied using another enhance LSB (secure LSB-L2) based image steganography. In this level another Bitmap (BMP) image has been used as a cover image and embeds (the BMP image output from level one) as a secure data and generates the new BMP image as stego_image. Lossless data compression technique using Huffman, LZW algorithm and Winrar Application between First and Second level of steganography are applied.

Keyword: steganography, Bitmap, Multi-Level, LSB, Lossless, compression.

1. INTRODUCTION

Security of information is one of the most important factors of information technology and communication. Security of information often lies in the secrecy of its existence and or the secrecy of how to decode it. Cryptography, watermarking and Steganography can be used in information security [1].

Steganography is defined as “the art and science of communicating in a way which hides the existence of the communication”. Methods of steganography have existed for centuries, though with the advent of digital technology, have taken on a new form. Embedding data within the redundancy and noise of media files is among these digital techniques.

Steganography can be classified into image, text, audio and video steganography based on the cover media used to embed secret data. Images are the most popular cover objects used for steganography. In the domain of digital images many different image file formats exist, most of them for specific applications. For these different image file formats, different stenographic algorithms exist.

Steganography (from Greek steganos, or "covered," and graphie, or "writing") is the hiding of a secret message within an ordinary message and the extraction of it at its destination. Steganography takes cryptography a step farther by hiding.

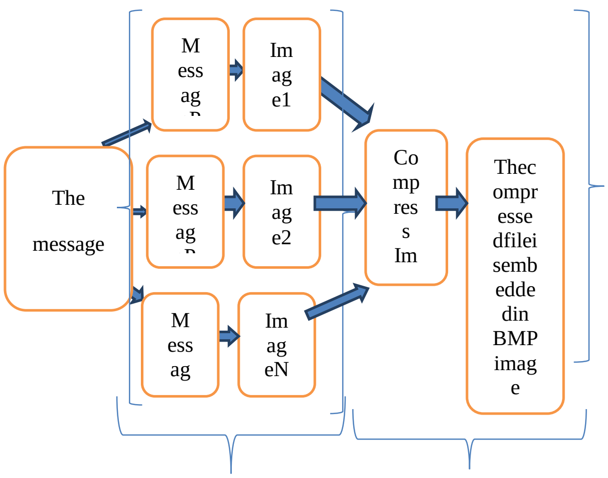


Figure 3.1 shows the proposed algorithm

2. LITERATURE REVIEW

Steganography is the art of hiding and transmitting data through apparently innocuous carriers in an effort to conceal the existence of the data, the word Steganography literally means covered or hiding writing as derived from Greek. Steganography has its place in security .It is not intended to replace cryptography but supplement it. Hiding a message with Steganography methods reduces the chance of a message being detected. If the message is also encrypted then it provides another layer of protection [2].

Data compression techniques, the file could be reduced in size to, say, 15 KB that makes it easier to store on disk and helps faster transmission over an Internet connection.

Data compression is a process by which a file (Text, Audio, and Video) may be transformed to another (compressed) file, such that the original file may be fully recovered from the original file without any loss of actual information [3]. This process may be useful if one wants to save the storage space. For example if one wants to store a 4MB file, it may be preferable to first compress it to a smaller size to save the storage space. Data Compression is possible because most of the real world data is very redundant. Data Compression is basically defined as a technique that reduces the size of data by applying different methods that can either be Lossy or Lossless. A compression program is used to convert data from an easy-to-use format to one optimized for compactness. Likewise, an uncompressing program returns the information to its original form.

3. PROPOSED METHOD

The proposed method is using multilevel image steganography [4], (two levels) level one will be done by embedding the secret message (text) into cover image (cover one) which is a colored image (BMP image) using Least Significant Bit (LSB) image steganography. And then using key to private message, and finally using compression techniques to compress output result that coming from level one. In level two improve the LSB scheme. It overcomes the sequence-mapping problem by embedding the massage into a set of random pixels, which are scattered on the cover-image. Figure 3.1 explain the general overview of the proposed method (embedding process). Steps extracting Process in level one using Modified LSB (secure LSB-L1).

Hide one message in many bitmaps. It is quite similar to writing text across a couple of pages [4]. It means spreading the pixels over multiple images. Figure 3.4 below Shown and more explain:

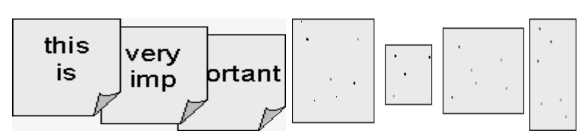


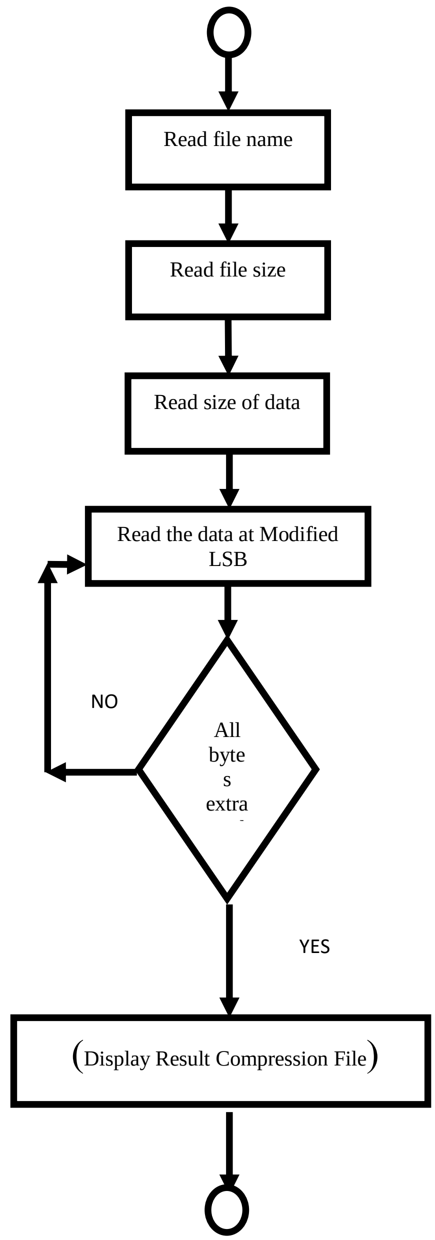
Figure 3.2 spread the information over the images

You can send each image in a separate E-mail, post them in different mailboxes, or store them on different discs. The GUI allows selecting carrier bitmaps the same way as selecting key files. The selection is stored as an array of Carrier Images. Larger images can hide more bytes (more pixels) than smaller images.

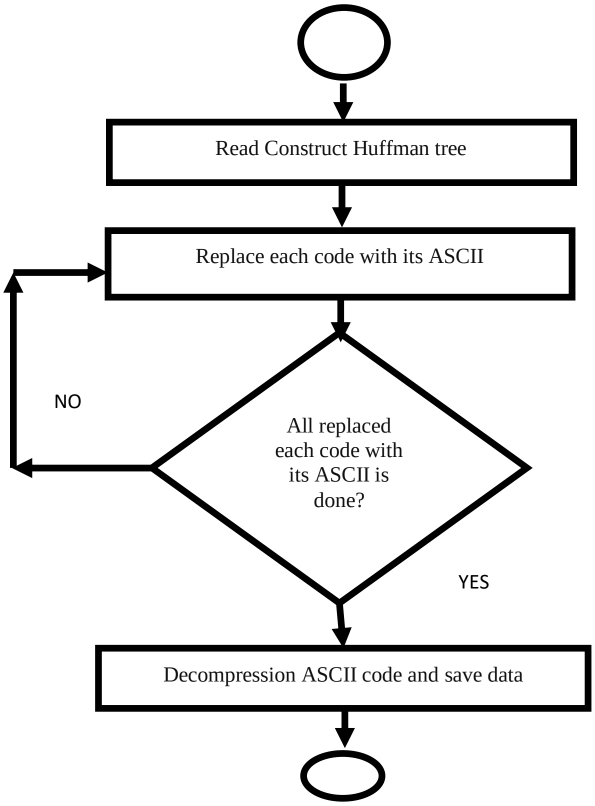
4. PROCESS OF CARRIER UNIT

Now, we start with the first carrier bitmap, loop over the message, hide a number of bytes, and switch to the second carrier bitmap, and so on. Current position in the carrier bitmap Start with 1, because (0,0) contains the message length. At the end, we must save the new images. Each image can be saved using a format (BMP).Steps Embedding Process in level tow using Modified LSB (secure LSB-L2)

Steps extracting Process in level Tow using Modified LSB (secure LSB-L2)



Compression process Using Huffman algorithm



Decompression Process Using Huffman algorithm

5. RESULT

Comparative analysis of multilevel image steganography (secure LSB-L1) and (secure LSB-L2) based image steganography has been done on the basis of parameters like Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR).There is a different message size have been used to embed them in different image size in the upper level of image steganography, the first message (first secret message) will be Use is shown in table 5.1

Size Messages	Images in level one	Images in level tow
270 bytes	black-box	Monaliza
4,650 bytes	Red - box	cyber-security
8,232 bytes	White-box	horse

Table 5.1 different image size in the upper level of image steganography

After the upper level (secure LSB-L1) is applied to the above secret messages the output is more than one image. Figure 5.1 Shown and explain level one applied method

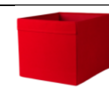
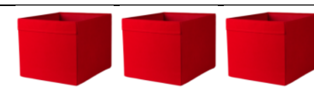
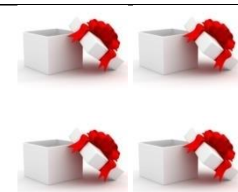
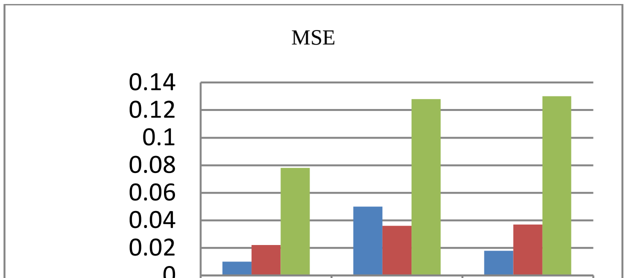
Original Image	Stago-images in level one
	
	
	

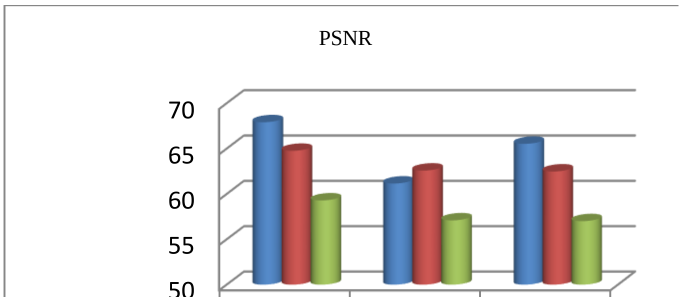
Figure 5.1 explain level one applied method

a below Diagram Shows the experiment results of stego_images and contains the Quality Image parameter values of stego_images above. Figures below is a Diagram showing its Quality Image parameter values (MSE, PSNR) for Monaliza stego-image.

International Journal of Computer Applications Technology and Research
Volume 6–Issue 11, 441-450, 2017, ISSN:-2319–8656

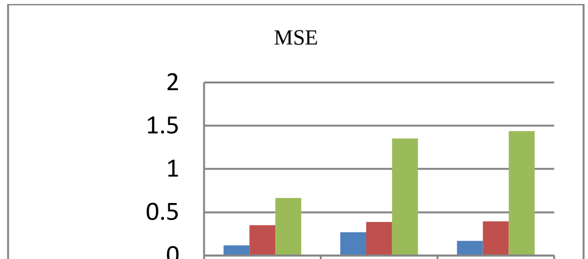


	Winrar	Huffman	LZW
message 1	0.01	0.05	0.018
message 2	0.022	0.036	0.037
message 3	0.078	0.128	0.13

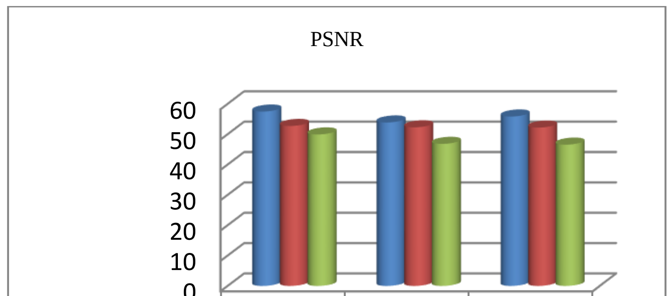


	Winrar	Huffman	LZW
message 1	67.91	61.15	65.53
message 2	64.76	62.58	62.48
message 3	59.29	57.09	57

The Second image is the cyber-security stego- image. Figure 4.12 is a Diagram Showing its Quality Image parameter values (MSE, PSNR) for cyber-security stego-image

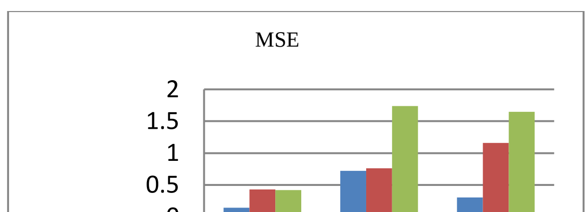


	Winrar	Huffman	LZW
message 1	0.118	0.268	0.17
message 2	0.349	0.387	0.395
message 3	0.663	1.354	1.439

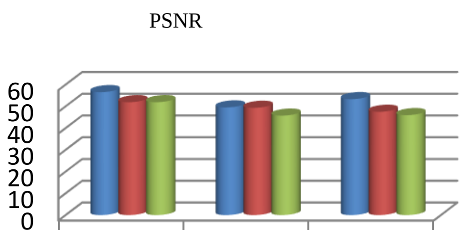


	Winrar	Huffman	LZW
message 1	57.45	53.86	55.82
message 2	52.71	52.26	52.17
message 3	49.91	46.81	46.55

The third image is the horse stego-image. Figures below shoes the Diagram showing its Quality Image parameter values (MSE, PSNR) for horse stego-image



	Winrar	Huffman	LZW
message 1	0.141	0.721	0.304
message 2	0.429	0.759	1.16
message 3	0.419	1.741	1.651



	Winrar	Huffman	LZW
message 1	56.52	49.5	53.31
message 2	51.9	49.33	47.49
message 3	51.91	45.72	45.95

5.1 Conclusion

The proposed model adds a level of security through the main theme of steganography: "hiding information in plain sight". The cover object usually does not invite suspicion, since it looks similar to the original object to the general observer.

The main objective is applying and improves the way to hide the information division the text on more BMP images.

In this thesis, a new concept for performing hidden secret data, called Multilevel Steganography for image steganography, was presented. MLS consists of at least two stenographic methods are utilized respectively, in such a way that one method (called the upper-level) as a carrier for the second one (called the lower-level).

The proposed method is two levels of image steganography, In the level one uses modified least significant bit (secure LSB-L1) image steganography to hide the secret information into more than one image carrier of the text (at least in 2 images). And that improving hide information by being distributed in more than one image carrier. The last step in this level, adding a key string to secure the information.

While level two employs the algorithm use for Encryption and Decryption in this level provides (secure LSB-L2) using several layers lieu of using only LSB layer of the image. Writing data starts from the last layer (8th or LSB layer); because significant of this layer is least and every upper layer has doubled significant from its down layer. So, every step we go to upper layer image quality decreases and image retouching transpires.

Multilevel Steganography has potential benefits, as it may enhance the confidentiality of the secret information by using two level image steganography in one the system and add more complexity to the steganography process through applying it in two levels.

Measuring the performance of proposed algorithm has been applied using many experiments and calculate many values of each experiment, the first value is Peak signal to noise ratio (PSNR), this ratio is used as a quality measurement between two images. If PSNR ratio is high then images are best of quality, the second measurement value is Mean Squared Error is the average squared difference between a reference image and a modified image (stego-image). And other calculates values are Normalized Cross-Correlation, Average Difference, Structural Content, Maximum Difference and Normalized Absolute Error.

There are many experiments have been conducted through the different size of secret messages (secret message one, two and three) utilized as a secret data in level one. And compress in one file, then concealed in one BMP image the output is compressed file or (intermediate object) and it's used as a secret data in level two.

5.2 Recommendations

- The proposed method can be used in military applications for secure communications.

- Try to check the result of proposed algorithm using the grayscale image on both levels to compare the performance results.

- Apply another compression technique.

- Apply compression to a text file.

5.3 Future Work

1- Adding Advance encryption algorithm to in the upper level to encrypt the secure text message to increase the security to proposed method.

2- Adding one more level (level Three)

3- Increase the System functionality to hide all other data types like audio, video not only text data and images.

4- Trying to enhance the performance of algorithms in both levels to increase the system capacity.

References

[1] Al-Dieimy, I.I.U, (2002), "Information Hiding In an Open Environment ", Computer Science & Information System (CSIS), University of Technology Malaysia, Malaysia.

[2] Dorothy, E.R, D.K, (2000), "Cryptography and Data Security", IEEE International Symposium on Canada Electronics (ISKE), University of Canada, Canada, Vol.6, p.p 119-122.

[3] P. Jeyanthi, V. Anuratha, "Analysis of Lossless Reversible Transformation Algorithms to Enhance DATA Compression", Journal of Global Research in Computer Science, Volume 3, No. 8, August 2012, p.p 56-62.

[4] Al-Najjar, Atef Jawad. "The decoy: multi-level digital multimedia steganography model." WSEAS International Conference. Proceedings. Mathematics and Computers in Science and Engineering. No. 12. World Scientific and Engineering Academy and Society, 2008.

[5] C. J. S. B, (2002)," Modulation and Information Hiding in Images", Vol. 1174, of Lecture Notes in Computer Science, University of Technology, Springer, p.p 207-226. Clelland, C.T.R, V.P & Bancroft, (1999), " Hiding Messages in DNA Micro Dots ", Proceedings of IEEE International Symposium on Industrial Electronics (ISIE), University of Indonesia, Indonesia, Vol. 1, p.p 315-327.

International Journal of Computer Applications Technology and Research
Volume 6–Issue 11, 451-460, 2017, ISSN:-2319–8656

Electronic Health Record Implementation Strategies

Richard A. Warren
Walden University
Minneapolis MN, USA

Ify S. Diala
Walden University
Minneapolis MN, USA

Abstract: Adoption of electronic health records (EHR) systems in nonfederal acute care hospitals has increased, with adoption rates across the United States reaching as high as 94%. Of the 330 plus acute care hospital EHR implementations in Texas, only 31% have completed attestation to Stage 2 of the meaningful use (MU) criteria. The purpose of this multiple case study was to explore strategies that hospital chief information officers (CIOs) used for the successful implementation of EHR. The target population consists of 3 hospitals CIOs from a multi-county region in North Central Texas who successfully implemented EHRs meeting Stage 2 MU criteria. The conceptual framework, for this research, was the technology acceptance model theory. The data were collected through semistructured interviews, member checking, review of the literature on the topic, and publicly available documents on the respective hospital websites. Using methodological triangulation of the data, 4 themes emerged from data analysis: EHR implementation strategies, overcoming resistance to technology acceptance, strategic alignment, and patient wellbeing. Participants identified implementation teams and informatics teams as a primary strategy for obtaining user engagement, ownership, and establishing a culture of acceptance to the technological changes. The application of the findings may contribute to social change by identifying the strategies hospital CIOs used for successful implementation of EHRs. Successful EHR implementation might provide positive social change by improving the quality of patient care, patient safety, security of personal health information, lowering health care cost, and improvements in the overall health of the general population.

Keywords: Electronic Health Records; EHR Implementations; Hospitals; Culture Acceptance; Technology.

1. INTRODUCTION

The United States has the largest health care system in the world, representing as much as 17.8% of the total Gross Domestic Production (GDP) in 2015 (Martin, Hartman, Washington, Catlin & The National Health Expenditures Team, 2017). American health care expenditures continue to exceed cost inflation and GDP growth each year (Payne, Pressler, Sarker, & Lussier, 2013). Despite the investment and escalating health care costs, there are shortcomings impacting the quality and efficiency of electronic health care record systems (Zhang et al., 2013). According to Payne et al. (2013), there is a continuing lack of management alignment of information systems (IS) and knowledge management technologies. Cognitive alignment of knowledge management systems with existing infrastructure is paramount in the migration to electronic health record (EHR) use and the articulation of the feasibility of EHR implementation for physicians (Dulipovici & Robey, 2013). Although the United States is a highly industrialized nation, the United States remains behind other countries in developing an interoperable EHR infrastructure (Sao, Gupta, & Gantz, 2013). Among the obstacles to the implementation of EHR systems are underdeveloped infrastructure and widespread concerns of consumers and medical professionals about privacy and security safeguards (Noblin et al., 2013). Technology and information systems abound in the United States, yet standardized, interoperable EHR systems provided by competing proprietary vendors are costly and could undermine patient centeredness (Zhang et al., 2016). Technical experts and technologists are available to help leaders in the United States become the preeminent leaders of EHR implementation. However, hospital leaders in the United States struggle with implementation timelines of EHRs, with some states significantly behind others in the rates of EHR adoption (Sao et al., 2013). The focus of this study was on exploring strategies hospital CIOs in Texas used for the successful implementation of EHR systems.

2. PROBLEM & PURPOSE OF THE STUDY

Adoption of EHR systems in nonfederal acute care hospitals has increased since 2012 across the United States, reaching adoption rates as high as 94% (Henry, Pylypchuk, Searcy, & Patel, 2016). Texas lags behind the rest of the country at 80%, and of the 330 plus acute care hospitals in which EHRs were implemented, only 31% have completed attestation to Stage 2 of the MU criteria (Office of the National Coordinator for Health Information Technology, 2016). The general business problem was that the implementation of EHR systems in Texas is below the national levels, with potential penalties for failure in attestation to MU criteria, leading to lost profits and elevated health care costs. The specific business problem was that some hospital chief information officers (CIOs) in Texas lack information about strategies for successful implementation of EHR systems. The purpose of this qualitative multiple case study was to explore strategies hospital CIOs in Texas used for the successful implementation of EHR systems. The population for this study included 10 acute care hospitals where the successful implementation of EHRs occurred in a multicounty region of North Central Texas. The multiple case study included the investigation of three hospitals' CIOs of the 10 acute care hospitals' CIOs having met Stage 2 attestation of the ONC and CMS specifications of the MU certification standards. Potential benefits of this study to society include the expansion of efficient quality medical practices and reducing medical care costs. The implications for positive social change includes improvements in medical care leading to a healthier society with lower health care costs and higher quality of care (Burns, Dyer, & Bailit, 2014). EHR improvements in medical care include rapid and economical medical diagnoses, less redundancy in diagnostic tests, and the potential reduced medical errors (Bailey et al., 2013). Data mining provides further social benefits through the discovery of new medical treatments, the convergent evolution of health information management, and career opportunities for health informatics specialists and IT professionals in health care settings (Gibson, Dixon, & Abrams, 2015).

3. DISCUSSIONS

3.1 Background of EHRs

Business The history of electronic health records began in the 1960s with the first implementation of computerized patient medical records, which evolved into advanced EHR systems (Murphy-Abdoch & Biedermann, 2014). Over the 50 years that followed the first implementation of computerized patient medical records, technology advances in computer innovations opened the floodgates for advancements in health care (Turk, 2015). Migration from paper documentation of patient data to digital forms of record keeping occurred through the use multiple software applications and stand-alone computer systems (Murphy-Abdoch & Biedermann, 2014). Development of and demand for innovation in health care technology continued as a potential cost-saving and efficient practice that could benefit patients and their health care organizations (Kerwin, Leighton, Buch, Avezbadalov, & Kianfar, 2016).

The demand for more efficient and affordable health care recordkeeping technology created an expanding ecosystem of vendor competition in sales of clinic-specific software (Liu & Zhu, 2013). Liu and Zhu (2013) proposed a unique model of an integrated e-service with the interconnected process and data-oriented grids. The model tied together electronic medical services, records, and application services with implementation architecture, which Liu and Zhu developed and tested through a prototype. Liu and Zhu contributed to the advancement of proposed technology models for health care professionals. However, the authors also highlighted obstacles, such as the need for ongoing updates and economic motivation, inability to interface legacy systems with emerging technologies, lack of interoperability, and cognitive factors involved in adopting new technologies. The explosion of a variety of applications and clinical specific systems exacerbated the problem of non-interoperable data sources.

Arvidsson, Holmstrom, and Lyytinen (2014) argued that strategic blindness becomes detrimental when mistranslating strategic intent, poor communications ensues, and cognitive entrenchment prevails. Accordingly, through the 1980s, desktop computers contributed to further development of non-interoperable, standalone systems and software applications for specific clinical tasks (Murphy-Abdoch & Biedermann, 2014). Legislative actions by Congress, enacting HIPAA in 1996, resulted in extended health insurance coverage and requirements for security of personal health information (Anthony et al., 2014). The governmental action forced health care providers to change the normal operating procedures concerning the protection of personal medical information (Brumen, Heričko, Sevčnikar, Završnik, & Hölbl, 2013). To meet the demands of federally mandated implementation of EHRs, hospital administrators pursued adoption of related technologies without a clear understanding of the total-cost-of-ownership (Legoux, Leger, Robert, & Boyer, 2014).

Implementation of EHRs is not an optional activity for hospital administrators and health care providers because of government legislation (Brumen et al., 2013). Hospital administrators have a stake in fulfilling their responsibility for implementation, patient care, and financial incentives (Boonstra et al., 2014). However, Eastaugh (2013) analyzed data from an independent firm's survey of hospital chief financial officers resulting in evidence of a lack of knowledge and strategy concerning total-cost-of-ownership. Eastaugh also suggested that selection of a vendor for EHR systems involved many variables, such as the number of required EHR support staff and salaries, estimated 10-year costs of operations, anticipated revenues increases or losses, and ongoing upgrade costs.

Developing a strategy that integrates the IT infrastructure with the hospital organization is essential to successful EHR implementation (Silverman, 2013). Both Eastaugh (2013) and Silverman (2013) emphasized the importance of organizational strategy alignment with information systems strategy. Functional structuring of business operations strategies with information systems strategies contributes to successful EHR implementation. Recommendations based on findings reported by Eastaugh and Silverman included ongoing research support to address the need for strategies for successful EHR implementation.

The selection of an IT vendor continues to challenge CIOs. Before the adoption of EHRs, directors and CIOs should consider a myriad of issues (Liebe, Hüsers, & Hübner, 2015). Business leaders should consider several aspects of EHR technology: interoperability, financial requirements, customer accessibility, internal business processes, and the means for learning and training (Loukis & Charalabidis, 2013). Health information technology integration, as described by Silverman (2013), requires careful structuring and thoughtful design to facilitate a variety of uses and to accommodate a variety of users.

Lack of interoperability of ill-conceived information systems undermines the business value of innovations (Hung, Chen, & Wang, 2014). Loukis and Charalabidis (2013) analyzed IS data that indicated establishing interoperability increases the positive impact of communication technology and medical informatics on the financial performance of the organizations. As technologies continue to evolve, assuring interoperability of various departmental and organizational specific computer applications, with security and protection of health information, represent paramount concerns for health care leaders (Meigs & Solomon, 2016; Rodrigues, de la Torre, Fernández, & López-Coronado, 2013; Studeny & Coustasse, 2014).

Anthony et al. (2014) provided a background of the U.S. health care systems, federal and state regulations, and laws while describing how the regulatory system affects health care. The emphasis of Anthony et al. research was in the personal health information regulatory compliance. Findings of the research indicated a variety of ways that hospital leadership implements or circumvents regulatory compliance. Anthony et al. argued that market environment and institutional logics impede standardized compliance. Furthermore, compliance is higher in the case of for-profit hospitals versus not-for-profit institutions (Appari et al., 2013). The conclusion is that organizational differences or changes affect the medical professionals as well as the leadership's strategies.

In their research on requirements set forth for compliance with the HITECH Act and Affordable Care Act, on the implementation of EHR and verification of MU, Appari et al. (2013) described inconsistent results. Specifically, Appari et al. found that implementation of EHRs that met a lower level of MU criteria obtained higher levels of baseline quality of care than those implementing higher levels of MU. The implication is that the acceptance of advanced EHRs requires time for diffusion of technology acceptance.

Enactment of the HITECH Act in 2009 also provided incentive funding for the implementation of EHRs, based on verification of meeting MU criteria (Emani et al., 2014). DesRoches, Aduet, Painter, and Donelan (2013) conducted a national survey of 1820 primary care physicians and specialists in office-based practices to determine the number of physicians who had a basic EHR system and met the MU criteria. The response rate was 60% from which a 43.5% of physicians reported having a basic EHR, and 9.8% met MU criteria (DesRoches et al., 2013). The authors concluded that few physicians could meet the requirements in early 2013. Additionally, physicians varied on their familiarity with MU processes and requirements for meeting MU criteria (Adler-Milstein et al., 2014). In conclusion, the pace of implementation was increasing while there was a continued concern with the ease of use.

A further concern for leadership is that simply replacing paper records with EHRs may fail to produce gains in quality and efficiency or the reduction in costs that EHRs have the potential to achieve (Emani et al., 2014). Ease of use is more likely to contribute to improving the potential effect of EHRs. For example, setting expectations too high is counterproductive and may lead to financial losses due to inadequate research and strategic planning (Appari et al., 2013). Indications are that technology implementation alone is likely, but not sufficient, to produce quality improvements.

Myriad factors contribute to success since each medical facility is unique suggesting that one size might not fit all (Abramson, McGinnis, Moore, & Kaushal, 2014; Meeks et al., 2014). Consequently, specific strategies might have varying success from one facility to another. CIOs and hospital leadership responsible for managing the implementation of EHRs have a stake in successfully guiding the processes and procedures (Gellert et al., 2015). Strategies for implementation involve consideration of domains, human factors, and financial implications, requiring knowledge and understanding of the complexities of the health care industry and specific clinical settings (Wu, Straub, & Liang, 2015). Strategic alignment of multiple domains within the organizational structure may enhance the overall health care system.

3.2 Barriers to EHR Adoption

Acceptance of advanced and innovative technologies is a common phenomenon (King, Patel, Jamoom & Furukawa, 2014). Although implementation of EHRs provides positive performance factor benefits to health care providers, there are obstacles hindering the process (Boonstra et al., 2014). In their systematic review and analysis, Boonstra et al. (2014) identified 19 frameworks for mitigating issues associated with EHR implementation. The three categories of the frameworks are (a) EHR context, (b) EHR content, and (c) EHR implementation process. Boonstra et al. recommended interventions for each in overcoming the obstacles to implementation.

Devkota and Devkota (2013) argued that expanding the use of EHR systems decreases health care costs and improves patient safety, efficiency, and overall organizational outcomes. However, obstacles to implementation, such as lack of funding and interoperability of current systems, retard the adoption of EHRs. Whereas Franzke, Wright, and Hautamaki (2014) argued that usability is a major concern, Devkota and Devkota (2013) noted that patient care and safety are the beneficial outcomes that should be of concern to leaders. Bagyogo, Lapointe, and Bassellier (2014) claimed that the focus of leaders should be on EHR performance, overall technology potential, and user initiative.

User adaptation and ease of use affect efficiency potential of data-intense environments creating opportunities for electronic patient and provider interactions (Ancker et al., 2014). Otto and Nevo (2013) suggested that, along with concern for safety, there are other mitigating factors such as political and economic issues slowing the progress of EHR adoption. Physicians' perceptions and resistance to migrating to EHRs, cited by Otto and Nevo were a loss of control, provider attitude, financial negatives, and continuity of care as an obstacle to adoption. Jamoom, Patel, Furukawa, and King (2014) presented a contrasting view concerning what little knowledge exists about physicians' perspectives on EHR adoption and use; in their research, a comparison of the perspectives of adopters and nonadopters revealed similar results as Otto and Nevo. The greatest obstacles perceived by both adopters and nonadopters included purchase cost and productivity (Jamoom et al., 2014; Otto & Nevo, 2013). Compared to other groups studied, the nonadopters showed considerably more concern with various national health IT policies and financial incentives or penalties for electronic record usage as major factors shaping their EHR adoption potentials.

King, Furukawa, and Buntin (2013) cited lower adoption rates in different geographic locations. King et al. studied EHR adoption rates in a low-income population part of the Midwest, another geographic area with a high population of low-income minorities in the Northeast, and a large metropolitan area in the American West. In contrast to the larger metropolitan area, the two underserved areas in the Midwest and Northeast had lower adoption rates (King et al., 2014). Reasons cited by King et al. for low adoption included limited access to advanced health care technology, organizational complexity, and less favorable business scenarios.

Struik et al. (2014) approached problems related to the slow adoption of EHRs in a discrete choice experiment. The experiment occurred to address the following previously identified barriers in the literature: data entry hardware, technical support, the attitude of the department head, performance feedback, flexibility of interface, and decision support. The perspectives of nurses and physicians were that flexibility of the interface was the factor of highest importance. The results aligned with the TAM, as ease of use, represented an enhancer to the acceptance of technology, described by Davis (1989). Struik et al. demonstrated the internal and external influences on the implementation of health information technologies, and then discussed the social implications affecting organizations.

Cresswell and Sheikh (2013) argued that although much research has covered the health care industry, organizational issues associated with implementation strategies lack adequate research. Zhang et al. (2013) proclaimed that the health care industry is much slower to adopt technology, in comparison to other industries, and there is a larger percentage of adoptions in administrative information technology versus clinical and strategic IT adoptions. Zhang et al. and Cresswell and Sheikh suggested organizational factors dominate as the most influential factors on adoption, requiring research attention.

3.2.1 Regulatory Influences

Concern for the security of patient medical records and the safety of patients prompted legislative action resulting in signing HIPPA into law in 1996 (Anthony et al., 2014). The HITECH Act and American Recovery and Reinvestment Act of 2009 provided incentives to promote the adoption of EHRs and MU of health information technology (Sheikh, Sood, & Bates, 2015). Passage of the Patient Protection and Affordable Care Act of 2010 introduced a far greater emphasis on federal regulations of the American health care industry (Bauer, Thielke, Katon, Unutzer, & Arean 2014). Shaw, Asomugha, Conway, and Rein (2014) proclaimed enactment of the Patient Protection and Affordable Care Act is the greatest change in American health care policy since the 1960s. Legislation restraining discriminatory insurance practices, providing more affordable coverage and methods of reducing costs may lead to considerable benefits and coverage for an additional 25 million American citizens (Shaw et al., 2014).

Due to the pervasiveness of EHRs and health information exchanges, there is increased potential for improved health care. However, Ben-Assuli (2014) argued that serious concerns are legal and privacy issues. Despite these unresolved concerns, incentives provided through the HITECH Act for attesting to MU contributed to the increased adoption of EHRs (Adler-Milstein et al., 2013). Adler-Milstein et al. (2013) emphasized that hospitals ineligible for the federal MU incentives have extremely low adoption rates. MU eligible providers perform quite well with most scoring 90-100 on the 15 measures of MU (Wright, Feblowitz, Samal, McCoy, & Sittig, 2014). The Centers for Medicare and Medicaid Services incentives of $30 billion have been instrumental in the rapid increase of adoption (Mirani & Harpalani, 2014). However, several states including Texas are slow in the adoption of EHRs (Charles et al., 2015).

3.2.2 EHR Benefits

Nationally, the beneficence of EHRs comes in different forms. Physicians' attest to the clinical benefits of providing enhanced patient care overall, ability to access patients' charts remotely, medical medication alerts, and critical lab values (King et al., 2014). In the research, King et al. performed a cross-sectional data examination of the 2011 Physician Workflow study, representative of office-based American physicians. The doctors' perspectives on the benefits of EHRs were that between 30% and 50% of physicians in the study stated clinical benefits were the ability to provide recommended care, appropriate tests, and enhanced patient-provider communications (King et al., 2014). However, Asan, Smith, and Montague (2014) studied 8 family practice physicians and 80 patients, leading to findings that physicians spent more time with the EHR screen than with paper records and less time looking at patients. Asan et al. claimed that their findings could be responsible for negative patient perceptions of physicians who use EHR, with implications for the design and adoption of related technologies.

Effective teamwork directly affects the quality of patient care. Properly aligned and implemented technologies can enhance professional health care teamwork (O'Malley, Draper, Gourevitch, Cross, & Scholle, 2015). Gratez et al. (2014) examined whether primary team cohesion affects outpatient EHRs and clinician-rated care coordinated across delivery sites. Gratez et al. claimed that EHR might not have a positive benefit with less cohesive teams; effectiveness and beneficence depend on the users' proficiencies with the systems. From their study of 63 physicians and health care desk staff, O'Malley et al. (2015) claimed that EHRs could facilitate communication and task delegations of teams but could pose challenges to teamwork if there is a lack of integrated software, poor functionality and interoperability, and inadequate ease of use.

Other noted benefits of EHRs, as indicated by Haegerich, Sugerman, Annest, Klevens, and Baldwin (2014) include injury and error prevention through improved surveillance and monitoring of clinical treatments and outcomes. Hoffman and Podgurski (2013) similarly reported enhanced clinical outcomes from EHRs for the prevention, treatment, and monitoring of infectious diseases, disease outbreaks, and chronic illnesses. EHRs contribute to the rapid analysis of data transmitted electronically to public health authorities. Jaffe, Harold, Frieden, and Thomas (2014) also identified numerous ongoing improvements in health security, enhanced surveillance systems, medical countermeasures, and laboratory networks designed to improve the ability to respond to day-to-day medical issues and emerging health issues.

The ability to store a massive amount of medical data improves continuously yet the accessing data could be challenging depending on the type of database (Wang, Min, Wang, Lu, & Duan, 2015). Communications after health care visits, referrals for specialists or follow-up visits, access to medical records, review of lab results, and maintaining financial records are part of the benefits of EHR systems, which are essentially databases that scholars, such as Wang et al. (2015) continue to try to improve. The benefits of EHRs include the ability to establish a path for accessing relevant data for a variety of medical conditions, transitioning from paper to computers with the potential for reduced health care costs, improved patient care, and safety. However, scholars such as Wang et al. (2015) continue to work toward solutions to the challenges. Noblin et al. (2013) argued that increased numbers of physicians and hospital administrators implementing EHRs assume that the systems will contribute to enhanced safety, efficiency and improved quality of care. As noted, research is replete with analysis indicating that system designs will continue to evolve and continue to garner the beneficial aspects of EHR.

EHRs represented the conversion from paper to digital media to provide physicians, health care staff, and patients the opportunities to store entire medical records and historical data on accessible or mobile storage media (Tansel, 2013). As the patient travels so does the patient's medical record. During medical emergencies, instant access to the patient's medical records might be the difference between life and death. Terry (2013) argued similarly that advances in technology should improve patient health care.

However, Terry (2013), like O'Malley et al. (2015) and Asan et al. (2014), acknowledged drawbacks that accompany the benefits of EHRs. Issues with EHRs such as usability, technological limitations that impede interoperability and safety concerns cast doubt on current EHRs (Terry, 2013). Lee, Kuo, and Goodwin (2013) also highlighted the gap that appeared between expected and actual outcomes of the benefits of EHR implementation. EHRs are inherently expensive because of required infrastructure, electrical power requirements, climatic control, equipment costs, software costs, IT personnel costs, and ongoing updates and maintenance costs. Dey, Sinha, and Thirumalai (2013) suggested that increasing the level of electronic medical records technology might not be beneficial to all providers, based on organizational, environmental, and financial limitations of the providers.

3.2.3 Cost of EHRs

EHR system costs can run into millions of dollars depending on the size and complexity of services provided by hospital organizations (Smith, Bradley, Bichescu, & Tremblay, 2013). Financial decisions made by hospital administrators determine strategies CIO's can pursue in the implementation of EHRs. Investing in information systems is a serious undertaking, but there is a lack of knowledge about how CEOs determine IS funding allocations among other competing expenses for business priorities (Salge et al., 2015). Wang et al. (2015) noted that data conversion and maintenance processes are costly regarding both time and money that escalates with greater numbers of record additions. Therefore, failure to discern the value of EHR implementation might be cause for hesitancy on the part of hospital CEOs' desire to allocate resources.

Determining the value creation by IT investment is difficult due to the differences between health care and other industries. Sherer (2014) argued that there is mixed evidence concerning the value created by health IT systems and implementation costs become difficult to determine because government incentives programs skew results. Adding to the overall cost of IT systems, management must consider the costs of IT employees. Kruse, Mileski, Alaytsev, Carol, and Williams (2015) reported that barriers to EHR adoption include escalating costs, users' negative perceptions, lack of sufficient implementation planning, and lack of proper training requiring potentially expensive support staff or extended education costs. Human resource managers confronted with demands for qualified IT staff work within the confines of budgets for the costs of technical training and IT personnel (Wang & Kaarst-Brown, 2014). As the expansion of technology increases, the need for technologically perceptive human capital also increases (Majumdar, 2014). CIOs obligate funds for EHR infrastructure and the intellectual capital to maintain the increasingly complex innovative technologies; organizations investing in EHR sustainability initiatives must expect increased budgetary expenses (Majumdar, 2014). Terry (2013) estimated that investments in health informatics and technology infrastructure costs approximately $60,000 per bed.

According to the Organization for Economic Co-operation and Development, the United States cost of public health per person exceeds five other high-spending countries examined by Lorenzoni, Belloni, and Sassi (2014). Additionally, the total overall costs of health care in America increased to $2.8 trillion in 2012 (Jaffee & Frieden, 2014). Researching the financial IT investment, Strong et al. (2014) reported falling short of the expected results of lowered costs, higher efficiency, and patient and provider satisfaction from IT medical record advances. In contrast to Strong et al. (2014), Smith et al. (2013) emphasized that sophisticated electronic medical record system investments result in improved financial performance and increased employee productivity. In agreement with Smith et al. 2013, Bardhan and Thouin (2013) reported a positive relationship between reduced costs and improved care with the implementation of financial and clinical information systems. In light of mixed research results, the significance of the EHR problems, and the relative infancy of their applications, a call for ongoing research persists in the literature, concerning the viability of EHRs, cost-effectiveness, and improved health care resulting from EHR investments (Bardhan & Thouin, 2013; Salge et al., 2015; Terry, 2013).

3.3 Barriers to EHRs Implementation

From ongoing research about physicians' perspectives on EHR adoption, the most emphasized barriers to EHR adoption are costs, productivity loss, and decreased interactions with patients due to increased interactions with computers (Bae & Encinosa, 2016; Jamoom et al., 2014; Kruse et al., 2015). Physicians perceived the use of EHRs to be time-consuming activities affecting the amount of time available for patient interaction (Bensefi & Zarrad, 2014). The majority of physicians who participated in EHR studies felt pressure to complete digital forms that detract from time providing health care to the patient (Meigs & Solomon, 2016). Discounting productivity loss, as familiarization increases, proficiency of the user decreases the time associated with the digital input (Bae & Encinosa, 2016). Social-technical acceptance, environmental impact, and organizational factors appear to be among the major factors influencing adoption of information technology (Zhang et al., 2013).

The literature reflects the reality of existing barriers to the adoption of informatics for most industries. However, there is a need to consider factors concerning the protection and security of patient data (Turk, 2015). The U.S. Congress passed ARRA, expanding HIPAA, with specific guideline and safeguards intended to protect the patients' records (Bredfeldt et al., 2013). However, breaches still occur due to the human factor when dealing with large digital data sources, with the potential for hacking and lack of adherence to the guiding principles (Turk, 2014). Across diverse health care settings throughout the world, there are concerns from patients and the public about the security and privacy of their EHR information (Papoutsi et al., 2015). Training increases knowledge and proficiency of system users and can help users understand privacy and security risks and concerns (Kim, 2013). Constant assessments of users' performances and procedures are paramount to mitigating risks, such as commercial exploitation, lack of accountability, data inaccuracies, prejudices, and inequalities in health care provision (Colligan, Potts, Finn, & Sinkin, 2015; Papoutsi et al., 2015).

Colligan et al. (2015) reported that cognitive workload associated with EHR usage increases for nurses. Although Colligan et al. warned against generalizing a one-size-fits-all conclusion about how EHR usage affects every person who uses the technology, the authors did explain that as user experience increases cognitive workload tends to decrease. User workload, cognitive processes required, and investment of time in training and usage may also vary depending on how many systems a user must learn, complicated by a general lack of interoperability among different (Loukis & Charlabidis, 2013). Interoperable information systems in the health care IT industry are uncommon and appears to be one of the most frequently cited problems with health care technology (Slight et al., 2015). There is a growing number of vendors and suppliers of certified health care information systems (Yeung, Jadad, & Shachak, 2013). Consequently, it is beneficial for adopters of IS in establishing an infrastructure focused on interoperability within the organization and with emphasis on collaborators such as customers, the organization supply chain, and business partners (Loukis & Charalabidis, 2013). Interoperability in the case of hospital EHRs is a barrier to implementation (Kruse et al., 2014).

4. METHODOLOGY

This qualitative, multiple case study involved three CIOs from 10 acute care hospitals in a multicounty region of North Central Texas who have implemented EHRs successfully meeting Stage 2 of the ONC and CMS specifications of the MU certification standards. Purposeful sampling for participants from the population who are information-rich subjects heightened the trustworthiness of this study. Purposeful selection of 3 CIOs possessing expertise, qualifications, and experience of having been successful in EHR implementation fulfills the requirement. Data collection processes involved face-to-face semistructured interviews with open-ended questions, outside of the hospital setting, at a private, quiet location mutually accessible to the researcher and participant. Another data collection technique used to gather participant data included publicly available documents, such as the documents and data maintained by the Definitive Healthcare Network, Government agencies, and hospitals to understand the strategies that can be successful for EHR implementation. Corroborating evidence may stem from documents about information system architecture, historical data, illustrative diagrams, detailed specifications, and implementation timelines. Analysis of data included, Yin's (2011) five-phase logical and sequential process: a) compiling, (b) disassembling, (c) reassembling, (d) interpreting, and (e) concluding. NVivo 11 qualitative data analysis software was used because it is used to generate graphs and charts that can add clarity by the graphic representations conducive to a better understanding of findings.

5. FINDINGS AND RESULTS

Findings from this study resulted to four main themes from the data collection and analysis:

5.1.1 Emergent Theme 1: EHR Implementation Strategies

The overall impetus was that governmental directives were not optional and that EHRs were inevitable; without them, the organization would suffer financially. The first decisions that all the participants referred to was that there were joint decisions by the board of directors, CEOs, and CIOs to begin the process of EHR implementation. All participants confirmed that, as a business factor, the board of directors and upper management were supportive of the financial investment required to reach their organizational goals. In alignment with the TAM framework, 100% of the participants admitted expectations of resistance to change and acceptance of the technology changes imposed upon staff, nurses, and doctors. Each of the participants referred to several strategies for implementation of the EHRs, with similar approaches directly supporting the strategies and aligning with the organizational strategy.

5.1.2 Emergent Theme 2: Obstacles to Technology Acceptance

All of the participants referred to the need to obtain user support and acceptance of the forthcoming technology changes associated with EHR implementation. A common practice emerged from the participants' responses: that of developing an implementation team of multilevel employees. Without user acceptance, other implementation strategies might incur opposition at each stage of the process. To that end, each participant developed their implementation team as the primary tool for overcoming the expected resistance to technology changes. All participants referred to the need to incorporate strategies for overcoming the obstacles to the acceptance of new technology. The consensus was that by identifying obstacles to the EHR implementation and ways to overcome obstacles may lead to successfully aligned strategies for improving acceptance.

5.1.3 Emergent Theme 3: Strategic Alignment

All the participants referred to the fact that what works for some organizations may not necessarily work in other organizations. Each hospital organization studied was different in many aspects, size, the number of employees, the number of beds, the number of individual clinics, and geographic location. The common ground for all the participant is that alignment of organizational strategy and EHR implementation is an ongoing challenge. Existing research confirms the participants' assertions in that identifying obstacles to technology acceptance and ways to overcome obstacles might lead to successfully aligned strategies for improving acceptance of the change and new technology.

5.1.4 Emergent Theme 4: Patient Wellbeing

All participants referred to potential benefits because of EHR implementation. The repeated occurrence of the keywords in the participants' responses identified the fourth major theme of improved patient well-being. The theme identified is a primary reason for governmental mandates for EHR implementation. The recurrence of the three terms led to three sub-themes contributing to improved patient well-being. All participants referred to the resulting benefits as improved patient health care, patient safety, and security of medical records. Patient care and safety are the beneficial outcomes for concerns of IT leaders. The participants' hospital websites corroborate responses provided during interviews and member checking procedures.

6. CONCLUSION

Successful implementation of EHRs, which meet the MU criteria, is a very challenging undertaking. The larger the organization, the greater the challenge due to the number of individual clinical applications there are to integrate into the EHR system. Hospital leaders understand the importance that information technology and other technologies contribute to the improvement in patient healthcare while reducing the overall cost of providing that care. Successful EHR system implementation is expensive and carries an ongoing cost with continual updates. Therefore, CIOs must have the full support of hospital leadership in financial matters, administratively, and organizationally for successful implementation. The overall stated goal of the participants was to meet the MU criteria as set forth by the ONC. What followed was numerous additional strategies to comply with the regulatory requirements otherwise face penalties equated in reduced reimbursements for medical services from the Center for Medicare and Medicaid Services (Adler-Milstein et al., 2013). Findings of this study indicate that strategies or methods used for successful EHR implementations are common sense approaches to overcoming resistance to change whether it be technology or changes in procedural practices. All participants confirmed that knowing the employees' perceptions and fears of the technology changes, guides the strategies required to overcoming the obstacles. Users must determine the ease of use and usefulness of the technology in verifying the success of overcoming resistance to technology acceptance (Davis, 1989). The strategies applied by the CIOs in this study such, as the implementation team and informatics team were instrumental in overcoming the resistance to technology acceptance. Identifying the obstacles to resistance enhances the potential for developing solutions to overcome the same.

7. REFERENCES

[1] Martin, A., Hartman, M., Washington, B., Catlin, A., & The National Health Expenditure Accounts Team, 2017). National health spending: Faster growth in 2015 as coverage expands and utilization increases. *Health Affairs, 36,* 1166-1176. doi:1377/hlthaff.2016.1330

[2] Payne, P., Pressler, T. R., Sarkar, I. N., & Lussier, Y. (2013). People, organizational, and leadership factors impacting informatics support for clinical and transitional research. *Biomedical Central Medical Informatics and Decision Making, 13*(20), 1-12. doi:10.1186/1472-6947-13-20

[3] Zhang, N. J., Seblega, B., Wan, T., Unruh, L., Agiro, A., & Miao, L. (2013). Health information technology adoption in U.S. acute care hospitals. *Journal of Medical Systems*, *37*(9907), 1-9. doi:10.1007/s10916-012-9907-2

[4] Dulipovici, A., & Robey, D. (2013). Strategic alignment and misalignment of knowledge management systems: A social representation perspective. *Journal of Management Information Systems*, *29*(4), 103-126. doi:10.2753/MIS0742-1222290404

[5] Sao, D., Gupta, A., & Gantz, D. A. (2013). Interoperable electronic health care record: A case for adoption of a national standard to stem the ongoing health care crisis. *Journal of Legal Medicine*, *35*, 55-90. doi:10.1080/01947648.2013.768153

[6] Noblin, A., Cortelyou-Ward, K., Cantiello, J., Breyer, T., Oliveira, L., Dangiolo, M., … Berman, S. (2013). EHR implementation in a new clinic: A case study of clinician perceptions. *Journal of Medical Systems*, *37*, 1-5. doi:10.1007/s10916-013-9955-2

[7] Henry, J., Pylypchuk, Y., Searcy T. & Patel V. (May 2016). Adoption of Electronic Health Record Systems among U.S. Non-Federal Acute Care Hospitals: 2008-2015. ONC Data Brief, no.35. Office of the National Coordinator for Health Information Technology: Washington DC

[8] Office of the National Coordinator for Health Information Technology. (2016). *Health IT dashboard.* Washington, DC: United States Department of Health and Human Services. Retrieved from https://dashboard.healthit.gov

[9] Burns, M., Dyer, M., & Bailit, M. (2014). Reducing overuse and misuse: State strategies *to improve quality and cost of health care.* Princeton, NJ: Robert Wood Johnson Foundation.

[10] Bailey, L. F. (2014). The origin and success of qualitative research. *International Journal of Market Research, 56*(2), 167-184. doi:10.2501/IJMR-2014-013

[11] Gibson, C. J., Dixon, B. E., & Abrams, K. (2015). Convergent evolution of health information management and health informatics: A perspective on the future of information professionals in health care. *Applied Clinical Informatics*, *6*(1), 163-184. doi:10.4338/ACI-2014-09-RA-0077

[12] Murphy-Abdouch, K., & Biedermann, S. (2014). The electronic health record. In S. H. Fenton, & S. Biedermann, *Introduction to healthcare informatics* (pp. 25-70). Chicago, IL: AHIMA Press

[13] Turk, M. (2015). Electronic health records: How to suture the gap between privacy and efficient delivery of healthcare. *Brooklyn Law Review, 80*, 565-597. Retrieved from https://www.brooklaw.edu

[14] Kerwin, T., Leighton, H., Buch, K., Avezbadalov, A., & Kianfar, H. (2016). The effect of adoption of an electronic health record on duplicate testing. *Cardiology Research and Practice, 2016,* 1-5. doi:10.1155/2016/1950191

[15] Liu, L., & Zhu, D. (2013). An integrated e-service model for electronic medical records. *Information Systems E-Business Management*, *11*, 161-183. doi:10.1007/s10257-012-0188-6

[16] Arvidsson, V., Holmstrom, J., & Lyytinen, K. (2014). Information systems use as strategy practice: A multi-dimensional view of strategic information system implementation and use. *Journal of Strategic Information Systems*, *23*, 45-61. doi:10.1016/j.jsis.2014.004

[17] Anthony, D. L., Appari, A., & Johnson, M. E. (2014). Institutionalizing HIPAA compliance: Organizations and competing logics in U.S. health care. *Health Care Systems*, *55*(1), 108-124. doi:10.1177/0022146513520431

[18] Brumen, B., Heričko, M., Sevčnikar, A., Završnik, J., & Hölbl, M. (2013). Outsourcing medical data analyses: Can technology overcome legal, privacy, and confidentiality issues? *Journal of Medical Internet Research*, *15*, 283-295. doi:10.2196/jmir.2471

[19] Legoux, R., Leger, P. M., Robert, J., & Boyer, M. (2014). Confirmation biases in the financial analysis of IT investment. *Journal of the Association for Information Systems, 15*(1), 33-52. Retrieved from http://aisel.aisnet.org/jais/

[20] Boonstra, A., Versluis, A., & Vos, J. F. J. (2014). Implementing electronic health records in hospitals: A systematic literature review. *Biomedical Central Health Services Research*, *14*, 370-384. doi:10.1186/1472-6963-14-370

[21] Eastaugh, S. R. (2013). The total cost of EHR ownership. *Health care Financial Management*, *67*(2), 66-70. Retrieved from https://www.hfma.org/hfm

[22] Silverman, R. D. (2013). EHRs, EMRs, and health information technology: To meaningful use and beyond. *Journal of Legal Medicine*, *34*(1), 1-6. doi:10.1080/01947648.2013.768134

[23] Liebe, J., Hüsers, J., & Hübner, U. (2015). Investigating the roots of successful IT adoption processes: An empirical study exploring the shared awareness-

knowledge of Directors of Nursing and Chief Information Officers. *BMC Medical Informatics and Decision Making*, *16*(1), 10-24. doi:10.1186/s12911-016-0244-0

[24] Loukis, E. N., & Charalabidis, Y. K. (2013). An empirical investigation of information systems interoperability business value in European firms. *Computers in Industry*, *64*, 412-420. doi:10.1016/j.compind.2013.01.005

[25] Hung, S., Chen, C., & Wang, K. (2014). Critical success factors for the implementation of integrated healthcare information systems projects: An organizational fit perspective. *Communications of the Association for Information Systems*, *34*, 775-796. Retrieved from http://aisel.aisnet.org/cais/

[26] Meigs, S. L., & Solomon, M. (2016). Electronic health record use a bitter pill for many physicians. *Perspectives in Health Information Management*, *13*(1), 1-4. Retrieved from http://perspectives.ahima.org/

[27] Rodrigues, J., de la Torre, I., Fernández, G., & López-Coronado, M. (2013). Analysis of the security and privacy requirements of cloud-based electronic health records systems. *Journal of Medical Internet Research*, *15*(8), 186-191. doi:10.2196/jmir.2494

[28] Studeny, J., & Coustasse, A. (2014). Personal health records: Is rapid adoption hindering interoperability? *Perspectives in Health Information Management*, *11*(2), 1-5. Retrieved from http://perspectives.ahima.org/

[29] Appari, A., Johnson, M. E., & Anthony, D. L. (2013). Meeting meaningful use of electronic health record systems and process quality of care: Evidence from a panel data analysis of U.S. acute-care hospitals. *Health Services Research*, *48*, 354-375. doi:10.1111/j.1475-6773.2012.01448.x

[30] Emani, S., Ting, D. Y., Healey, M., Lipsitz, S. R., Karson, A. S., Einbinder, J. S., ...Bates, D. W. (2014). Physician beliefs about the impact of meaningful use of the EHR: A cross-sectional study. *Applied Clinical Informatics*, *5*, 789-801. doi:10.4338/ACI-2014-05-RA-0050

[31] DesRoches, C. M., Audet, A., Painter, M., & Donelan, K. (2013). Meeting meaningful use criteria and managing patient populations: A national survey of practicing physicians. *Annals of Internal Medicine*, *158*, 791-799. doi:10.7326/0003-4819-158-11-201306040-00003

[32] Adler-Milstein, J., Slazberg, C., Franz, C., Orav, E. J., & Bates, D. W. (2013). The impact of electronic health records on ambulatory costs among Medicaid beneficiaries. *Medicare Medicaid Research Review*, *3*(1), e1-e13. doi:10.5600/mmrr.003.02.a03

[33] Abramson, E. L., McGinnis, S., Moore, J., & Kaushal, R. (2014). A statewide assessment of electronic health record adoption and health information exchange among nursing homes. *Health Services Research*, *49*, 361-372. doi:10.1111/1475-6773.12137

[34] Meeks, D. W., Smith, M. W., Taylor, L., Sittig, D. F., Scott, J. M., & Singh, H. (2014). An analysis of electronic health record-related patient safety concerns. *Journal of the American Medical Informatics Association*, *21*, 1053-1059. doi:10.1136/amiajnl-2013-002578

[35] Gellert, G. A., Hill, V., Bruner, K., Maciaz, G., Saucedo, L., Catzoela, L., … Webster, S. L. (2015). Successful implementation of clinical information technology: Seven key lessons from CPOE. *Applied Clinical Informatics*, *6*, 698-715. doi:10.4338/ACI-2015-06-SOA-0067

[36] Wu, S. P., Straub, D. W., & Liang, T. (2015). How information technology governance mechanisms and strategic alignment influence organizational performance: Insights from a matched survey of business and IT managers. *MIS Quarterly*, *39*, 497-518. Retrieved from http://www.misq.org

[37] King, J., Patel, V., Jamoom, E. W., & Furukawa, M. F. (2014). Clinical benefits of electronic health record use: National findings. *Health Services Research*, *49*, 392-404. doi:10.1111/1475-6773.12135

[38] Devkota, B., & Devkota A. (2013). Electronic health records: Advantages of use and barriers to adoption. *Health Renaissance*, *11*(3), 181-184. doi:10.3126/hren.v11i3.9629

[39] Franzke, M., Wright, S., & Hautamaki, B. (2014). The intersection of IT and human factors: Summative testing in safety-enhanced EHR design. *Biomedical Instrumentation & Technology*, *47*(2), 1-6. doi:10.2345/0899-8205-47.s2.54

[40] Bagayogo, F. F., Lapointe, L., & Bassellier, G. (2014). Enhanced use of IT: A new perspective on post-adoption. *Journal of the Association for Information Systems*, *15*, 361-387. Retrieved from http://aisel.aisnet.org/jais/

[41] Ancker, J. S., Kern, L. M., Edwards, A., Nosal, S., Stein, D. M., & Hauser, D. (2014). How is the electronic health record being used? Use of EHR data to assess physician-level variability in technology use. *Journal of the American Medical Informatics Association*, *21*, 1001-1008. doi:10.1136/amiajnl-2013-002627

[42] Otto, P., & Nevo, D. (2013). Electronic health records: A simulation model to measure the adoption rate from policy interventions. *Journal of Enterprise Information Management*, *26*(1/2), 165-182. doi:10.1108/17410391311289613

[43] Jamoom, E. W., Patel, V., Furukawa, M. F., & King, J. (2014). EHR adopters vs. non-adopters: Impacts of, barriers to, and federal initiatives for EHR adoption. *Healthcare*, *2*(1), 33-39. doi:10.1016/j.hjdsi.2013.12.004

[44] King, J., Furukawa, M. F., & Buntin, M. B. (2013). Geographic variation in ambulatory electronic health record adoption: Implications for underserved communities. *Health Services Research*, *48*, 2037-2059. doi:10.1111/1475-6773.12078

[45] Struik, M. H., Koster, F., Schuit, A. J., Nugteren, R., Veldwijk, J., & Lambooij, M. S. (2014). The preferences of users of electronic medical records in hospitals: Quantifying the relative importance of barriers and facilitators to an innovation. *Implementation Science*, *9*(69), 1-11. doi:10.1186/1748-5908-9-69

[46] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*, 319-340. doi:10.2307/249008

[47] Cresswell, K., & Sheikh, A. (2013). Organizational issues in the implementation and adoption of health information technology innovations: An interpretative

review. *International Journal of Medical Informatics, 82*(5), e73-86. doi:10.1016/j.ijmedinf.2012.10.007

[48] Sheikh, A., Sood, H. S., & Bates, D. W. (2015). Leveraging health information technology to achieve the "triple aim" of health care reform. *Journal of the American Medical Informatics Association, 22*, 849-856. doi:10.1093/jamia/ocv022

[49] Bauer, A. M., Thielke, S. M., Katon, W., Unutzer, J., & Arean, P. (2014). Aligning health information technologies with effective service delivery models to improve chronic disease care. *Preventive Medicine 66*, 167-177. doi:10.1016/j.ypmed.2014.06.017

[50] Shaw, F. E., Asomugha, C. N., Conway, P. H., & Rein, A. S. (2014). The patient protection and affordable care act: Opportunities for prevention and public health. *Lance, 384*, 75-82. doi:10.1016/s0140-6736(14)60259-2

[51] Ben-Assuli, O. (2015). Electronic health records, adoption, quality of care, legal, and privacy issues and their implementation in emergency departments. *Health Policy, 119*(3), 287-297. doi:10.1016/j.healthpol.2014.11.014

[52] Wright, A., Feblowitz, J., Samal, L., McCoy, A. B., & Sittig, D. F. (2014). The Medicare electronic health record incentive program: Provider performance on core and menu measures. *Health Services Research 49,* 325-346. doi:10.1111/1475-6773.12134

[53] Mirani, R., & Harpalani, A. (2014). The Medicare electronic health records (EHR) incentive program: First-year adoption response from inpatient hospitals. *Journal of Organizational Computing and Electronic Commerce, 24,* 388-401. doi:10.1080/10919392.2014.956601

[54] Charles, D., Gabriel, M., & Searcy T. (2015). *Adoption of electronic health record systems among U.S. non-federal acute care hospitals: 2008-2014 ONC Data Brief, No. 23.* Washington, DC: Office of the National Coordinator for Health Information Technology, Government Printing Office.

[55] King, J., Patel, V., Jamoom, E. W., & Furukawa, M. F. (2014). Clinical benefits of electronic health record use: National findings. *Health Services Research, 49,* 392-404. doi:10.1111/1475-6773.12135

[56] Asan, O., Smith, P., & Montague, E. (2014). More screen time, less face time: Implications for EHR design. *Journal of Evaluation in Clinical Practice, 20,* 896-901. doi:10.1111/jep.12182

[57] O'Malley, A. S., Draper, K., Gourevitch, R., Cross, D. A., & Scholle, S. H. (2015). Electronic health records and support for primary care teamwork. *Journal of the American Medical Informatics Association, 22*, 426-434. doi.org/10.1093/jamia/ocu029

[58] Graetz, I., Reed, M., Shortell, S. M., Rundall, T. G., Bellows, J., & Hsu, J. (2014). The association between EHRs and care coordination varies by team cohesion. *Health Services Research, 49,* 438-452. doi:10.1111/1475-6773.12136

[59] Haegerich, T. M., Sugerman, D. D., Annest, J. L., Klevens, J., & Baldwin, G. T. (2015). Improving injury prevention through health information technology. *American Journal of Preventive Medicine, 48*(2), 219-228. doi:10.1016/j.amepre.2014.08.018

[60] Hoffman, J. (2014). Preemption and the MLR provision of the affordable care act. *American Journal of Law & Medicine, 40*, 280-297. Retrieved from http://www.bu.edu/ajlm/

[61] Jaffe, H. W., & Frieden, T. R. (2014). Improving health in the USA: Progress and challenges. *The Lancet, 384*(9937), 3-5.doi:10.1016/S0140-6736(14)61032-1

[62] Wang, L., Min, L., Wang, R., Lu, X., & Duan, H. (2015). Archetype relational mapping: A practical open EHR persistence solution. *Biomedical Central Medical Informatics and Decision Making, 15*, 88-98. doi:10.1186/s12911-015-0212-0

[63] Tansel, A. (2013). Innovation through patient health records. *Procedia Social and Behavioral Sciences, 75*, 183-188. doi:10.1016/j.sbspro.2013.04.021

[64] Terry, N. P. (2013). Meaningful adoption: What we know or think we know about the financing, effectiveness, quality, and safety of electronic medical records. *Journal of Legal Medicine, 34*(1), 7-42. doi:10.1080/01947648.2013.768143

[65] Lee, J., Kuo, Y., & Goodwin, J. S. (2013). The effect of electronic medical record adoption on outcomes in US hospitals. *Bio Medical Central Health Services Research, 13,* 1-7. doi:10.1186/1472-6963-13-39

[66] Dey, A., Sinha, K. K., & Thrumalai, S. (2013). IT capability for health care delivery: Is more better? *Journal of Service Research, 16*, 326-340. doi:10.1177/1094670513478832

[67] Smith, A., Bradley, R. V., Bichescu, B. C., & Tremblay, M. C. (2013). IT governance characteristics, electron medical records sophistication, and financial performance in U.S. hospitals: An empirical investigation. *Decision Sciences Journal, 44,* 483-516. doi:10.1111/deci.1219

[68] Salge, T. O., Kohli, R., & Barrett, M. (2015). Investing in information systems: On the behavioral and institutional search mechanisms underpinning hospitals' investment decisions. *MIS Quarterly, 39*(1), 61-90. Retrieved from www.misq.org

[69] Wang, J., Ho, H., Chen, J., Chai, S., Tai, C., & Chen, Y. (2015). Attitudes toward inter-hospital electronic patient record exchange: Discrepancies among physicians, medical record staff, and patients. *Biomedical Central Health Services Research, 15,* 264-279. doi:10.1186/s12913-015-0896-y

[70] Sherer, S. (2014). Advocating for action design research on IT value creation healthcare. *Journal of the Association for Information Systems, 15,* 860-878. Retrieved from http://aisel.aisnet.org/jais/

[71] Kruse, C. S., Mileski, M., Alaytsev, V., Carol, E., & Williams, A. (2015). Adoption factors associated with electronic health record among long-term care facilities: A systematic review. *BMJ Open, 5*(1), 66-75. doi:10.1136/bmjopen-2014-006615

[72] Wang, C., & Kaarst-Brown, M. (2014). The IT compensation challenges: Theorizing the balance among multi-level internal and external uncertainties. *Journal of the Association for Information Systems, 15*(3), 111-146. Retrieved from http://aisle.aisnet.org/jais

[73] Majumdar, S. K. (2014). Technology and wages: Why firms invest and what happens. *Technology in Society, 39,* 44-54. doi:10.1016/j.techsoc.2014.07.005

[74] Lorenzoni, L., Belloni, A., & Sassi, F. (2014). Health-care expenditure and health policy in the USA versus

other high-spending OECD countries. *Lancet, 384*(9937), 83-92. doi:10.1016/s0140-6736(14)60571-7

[75] Strong, D. M., Johnson, S. A., Tulu, B., Trudel, J., Volkoff, O., Pelletier, L. R., ... Garber, L. (2014). A theory of organization-EHR affordance actualization. *Journal of Computer Information Systems, 15*(2), 55-85. Retrieved from http://aisle.aisnet.org/jais

[76] Smith, A., Bradley, R. V., Bichescu, B. C., & Tremblay, M. C. (2013). IT governance characteristics, electron medical records sophistication, and financial performance in U.S. hospitals: An empirical investigation. *Decision Sciences Journal, 44,* 483-516. doi:10.1111/deci.1219

[77] Bardhan, I., R., & Thouin, M. F. (2013). Health information technology and its impact on the quality and cost of healthcare delivery. *Decision Support Systems, 55,* 438-449. doi:10.1016/j.dss.2012.10.003

[78] Bae, J., & Encinosa, W. E. (2016). National estimates of the impact of electronic health records on the workload of primary care physicians. *Biomedical Central Health Services Research, 16,* 172-194. doi:10.1186/s12913-016-1422-6

[79] Bensefia, A., & Zarrad, A. (2014). A proposed layered architecture to maintain privacy issues in electronic medical records. *E-Health Telecommunication Systems and Networks, 3*(4), 43-49. doi:10.4236/etsn.2014.34006

[80] Bredfeldt, C., Butani, A. L., Pardee, R., Hitz, P., Padmanabhan, S., & Saylor, G. (2013). Managing personal health information in distributed research network environments. *Biomedical Central Medical Informatics and Decision Making, 13*(116), 1-7. doi:10.1186/1472-6947-13-116

[81] Papoutsi, C., Reed, J. E., Marston, C., Lewis, R., Majeed, A., & Bell, D. (2015). Patient and public views about the security and privacy of Electronic Health Records (EHRs) in the UK: Results from a mixed methods study. *Biomedical Central Medical Informatics and Decision Making*, *15*, 86-102. doi:10.1186/s12911-015-0202-2.

[82] Kim M. (2013). Improving electronic health records training through usability evaluation in primary care. Journal of Health and Medical Informatics, *4*(5), 110-115. Retrieved from http://www.ijmijournal.com

[83] Colligan, L., Potts, H. W., Finn, C. T., & Sinkin, R. A. (2015). Cognitive workload changes for nurses transitioning from a legacy system with paper documentation to a commercial electronic health record. *International Journal of Medical Informatics, 84,* 469-476. doi:10.1016/j.ijmedinf.2015.03.003

[84] Loukis, E. N., & Charalabidis, Y. K. (2013). An empirical investigation of information systems interoperability business value in European firms. *Computers in Industry*, *64*, 412-420. doi:10.1016/j.compind.2013.01.005

[85] Slight, S. P., Berner, E. S., Galanter, W., Huff, S., Lambert, B. L., Lannon, C., … Bates, D. W. (2015). Meaningful use of electronic health records: Experiences from the field and future opportunities. *Journal of Medical Internet Research*, *3*(3), 30-44. doi:10.2196/medinform.4457

[86] Yeung, N. K., Jadad, A. R., & Shachak, A. (2013). What do electronic health record vendors reveal About their products: An analysis of vendor websites. *Journal of Medical Internet Research*, *15*(2), 36-58. doi:10.2196/jmir.2312

[87] Kruse, C. S., DeShazo, J., Kim, F., & Fulton, L. (2014). Factors associated with adoption of health information technology: A conceptual model based on a systematic review. *Journal of Medical Internet Research*, *2*(1), 1-10. doi:10.2196/medinform.3106

[88] Yin, R. K. (2011). *Qualitative research from beginning to end.* New York, NY: Guilford Press

International Journal of Computer Applications Technology and Research
Volume 6–Issue 11, 461-466, 2017, ISSN:-2319–8656

Intrusion Detection System Using Genetic Algorithm and Data Mining Techniques Based on the Reduction Features

Mohammad Ghalehgolabi
Electronic and Systems Engineering
Department of Rahjuyan Danesh Borazjan
Bushehr, Iran

Amin Rezaeipanah
Faculty of Computer
Department of Rahjuyan Danesh Borazjan
Bushehr, Iran

Abstract: An intrusion detection system is the process for identifying attacks on network. Choosing effective and key features for intrusion detection is a very important topic in information security. The purpose of this study is to identify important features in building an intrusion detection system such that they are computationally effcient and effective. To improve the performance of intrusion detection system, this paper proposes an intrusion detection system that its features are optimally selected using genetic algorithm optimization. The proposed method is easily implemented and has a low computational complexity due to use of a simplified feature set for the classification. The extensive experimental results on the NSL-KDD intrusion detection benchmark data set demonstrate that the proposed method outperforms previous approaches, providing higher accuracy in detecting intrusion attempts and lower false alarm with reduced number of features.

Keywords: intrusion detection; genetic algorithm; distribution function; NSL-KDD; feature selection.

1. INTRODUCTION

In recent year, due to the growing use of smart devices and the Internet, network traffic is rapidly increasing. A Cisco report found the following : “Global IP traffic in 2012 stands at 43.6 exabytes per month and will grow threefold by 2017, to reach 120.6 exabytes per month” [1]. Intrusions are defined as attempts or action to compromise the confidentiality, integrity or availability of computer or network [2]. Intrusion detection systems (IDSs) are software or hardware systems that automate the process of monitoring the events occurring in a computer system or network, analyzing them for signs of security problems [3]. Feature Selection (FS) is the process of removing features from the original data set that are irrelevant with respect to the task that is to be performed. So not only the execution time of the classifier that processes the data reduces but also accuracy increases because irrelevant or redundant features can include noisy data affecting the classification accuracy negatively [4]. In this paper, we suggest a new feature selection method that uses the features distribution function. The decision tree [5] and k-nearest neighbor [6] classifiers will be evaluated with the NSL-KDD dataset to detect attacks on four attack categories: Dos, Probe, R2L, and U2R. The decision tree classifier’s results are computed for comparison of feature reduction methods to show that our proposed model is more efficient for network intrusion detection.

The remainder of the paper is organized as follows: Section 2 give an overview of feature selection methods and intrusion detection. The basic concept of the proposed method are presented in Sections 3 and the experimental results are presented in Section 4. Finally the paper is concludes with their future work in section 5.

2. RELATED WORKS

Intrusion detection techniques using data mining have attracted more and more interests in recent years. Feature selection is important to improving the efficiency of data mining algorithms [7]. Different researchers propose different algorithms in different categories, from Bayesian approaches [8] to decision trees [9], from rule based models [10] to functions studying [11]. The detection efficiencies therefore are becoming better and better than ever before. In recent years, researchers turn their focus on heuristic and hyper-heuristic methods for features selection. Several examples on these methods including Genetic Algorithm [12], Particle Swarm Optimization [13], and Ant Colony Optimization [14].

Sung and Mukkamala proposed a well-known closedloop FS method for SVM-based IDS, called SVM-RFE, which recursively eliminated one feature at a time and compared the resulting performance in each SVM test [15]. They also ranked six significant features [16]. Intrusion Detection in NEAR System by Anti-denoising Traffic Data Series using Discrete Wavelet Transform was presented by Vancea [17]. In [18] uses NGSA-II for wrapper-based feature selection and GHSOM-pr as the classifier to build efficient IDS. D. Sequeira [19] discussed in their research different types of firewalls. Traditional firewalls cannot detect internal attacks such as flooding attacks, user-to-root attacks, and port scanning because they only sniff out network packets at the network boundaries. Moreover, traditional firewalls cannot differentiate between ordinary traffic and DoS attack traffic, as mentioned by [20]. Warsi et al. [21] present a selective iteration based particle swarm optimization (SIPSO) for intrusion detection system with an upgraded beginning masses and decision director, to capably distinguish diverse sorts of interferences. Aghdam and Kabiri considered the feature selection using ant colony optimization in detecting the attacks [15]. The purpose of this study is to identify important features in building an intrusion detection system such that they are computationally effcient and effective.

3. PROPOSED SYSTEM

Some data sets like NSL-KDD have a lot of features. On the other hand, all of these features do not play a positive role in data categorization. Therefore, you need to select a subset of the best features. In this research, a genetic algorithm is used to select the desired features. This method operates on the basis of the features distribution function analysis. This factor helps to improve the genetic algorithm chromosomes by recognizing the peculiarities. The proposed method can work on a dataset of different dimensions. To evaluate the selected features, two well-known data mining techniques, decision tree (DT) and k-nearest neighbor (KNN) are used. Figure 1 shows the flowchart of the proposed method.

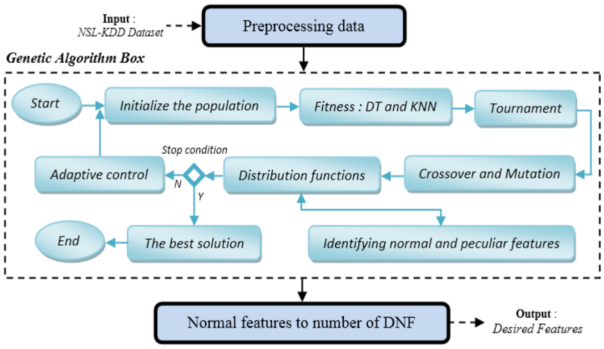


Figure. 1 Overall process of proposed intrusion detection system

3.1 Preprocessing Data

The first step in the creation of any model based on data mining techniques is the preprocessing of data. Pre-processing is done to prepare data for processing as well as improve the quality of real data. This step involves converting string-to-number properties, normalizing and disassembling data.

3.2 Improved Genetic Algorithm for Features Selection

Genetic algorithm was introduced by Holland in 1970, inspired by genetics and Darwinian evolution theory [22]. In this research, the structure of the chromosome is considered with regard to the number of each attribute. Each chromosome is a string of bits with values of 0 and 1 with a length of the total number of features. The genes of a chromosome show the desirable features that will be involved in the classification of the data. In this research, the number of desirable features (DNFs) is fixed in terms of test and error. In the proposed method, the genetic algorithm is implemented sequentially, so each repetition requires the production of a primary population of features. The genetic algorithm begins with an initial population of chromosomes randomly. Then the cluttered and cluttered features of the search space are extracted and used to generate the population in later stages. Compact features are a vector of attributes that are used in the production of the population. Pecularities are vector of features that their use in pre-population generation does not have desirable results and will not be used in the production of new populations. The fitness criterion of chromosomes is the error rate of the classification of data. Because of the expeditious calculation of fitness, two classifiers of KNN and DT have been used.

The chromosome selection operator, the tournament, and the crossover operator was one-point cross over. In the one point cross over operator, single particle genes for parents are exchanged to create new members. After applying this operator to probability Cr, the number of 1 chromosome genes must be constant. A one-point crossover point on both parents' organism strings is selected. All data beyond that point in either organism string is swapped between the two parent organisms. The resulting organisms are the children. An example of this operator is shown in Figure 2.

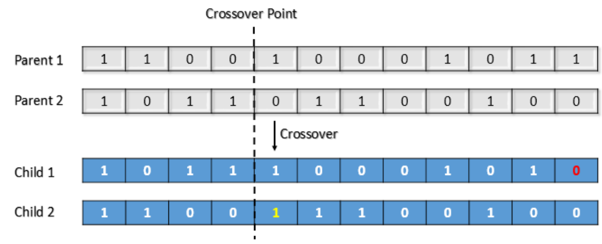


Figure. 2 One-point crossover operator suggested

In this example, features number is 12 and Desired Number of Features is 6, two genes of the children have changed. With applying the one-point crossover, the number of features in the first child is 7 and in the second child is 5. The number of genes in a chromosome should be equal to 6 in the offspring. So, in the first child we will random delete a gene and in the second child, we also random select a gene in unused features.

The mutation change of the bit is applied to one of the produced chromosomes. The role of the mutation in the genetic algorithm is to restore the genetic loss of the population, which provides access to all of the search space. The mutant operator is applied to the probability Mr for each gene. An example is shown in Figure 3.

Figure. 3 Bit change mutation operator suggested

The number of children created is equal to the number of parents. In order to determine the population of the next generation, the chromosomes of the population of the previous and current generations are sorted according to the fitness criterion in descending order. Then, the 25% elemental list (best ones) goes straight to the next generation. The endpoints of 25% chromosomes are removed (the worst ones) and finally the rest of the population are randomly selected from the remaining chromosomes.

One of the interesting phenomena of genetic algorithms is the production of intermediate-generation chromosomes that have a high degree of fitness. These chromosomes may be destroyed due to the application of mutant and crossover operators and no longer be produced.

In this research, elitism is used to preserve these chromosomes. In each generation, a chromosome with the best amount of fitness is transmitted directly to the next generation.

3.3 Using the Features Distribution Function in Identifying Normal and Peculiar Features

At the end of the genetic algorithm, a population of solutions is obtained. In most techniques, the features used in the best solution are considered as desirable features and classify educational data based on these features. The structure of the genetic algorithm is based on random search, which is why it does not always produce the same optimal solution. With these conditions, it will not be possible to find the desirable features that will best serve the classification of data.

Therefore, in this research, an approach has been proposed that largely leads to the selection of the best features. Our goal in this section is to identify the normal and peculiar features due to the outcomes of the genetic algorithm. To realize this goal, features distribution function (FD) has been used in the population. The distribution of any feature in the population indicates the degree of repetition of that feature.

Distribution of the characteristics of the population in the population is the rate of repetition of these features in parts of the population with high fitness. For example, the normal population are solutions that their fitness is greater than the overall fitness of the whole population. Also, the distribution of peculiar features is the frequency of these features in parts of the population with low fitness. For example, a rough population is a solution that is less than the overall fitness of the whole population.

The distribution function of a feature in a normal population is the ratio of its recurrence to the total population and the distribution function of an attribute in the peculiar population is the ratio of its recurrence to the entire peculiar population. Table 1, shows an example of the distribution function of the features.

In this example, according to the average population criteria, 4 solutions for normal population and 6 solutions for pecliar population were selected. The frequency of the first feature (F1) in the normal population is 3 and in the pecliar population is 2.

Table 1. Example of the distribution function

Solution	F1	F2	F3	F4	F5	F6	F7	F8	Fitness
1	0	0	1	1	0	1	0	1	87
2	1	1	0	1	1	0	0	0	85
3	1	0	1	1	0	1	0	0	78
4	1	1	0	1	0	0	0	1	70
5	0	0	1	0	0	1	1	1	50
6	1	0	1	0	1	0	1	0	45
7	0	1	0	1	0	0	1	1	40
8	1	0	1	0	1	1	0	0	39
9	0	0	0	1	1	1	0	1	37
10	0	0	0	1	1	1	1	0	32
normal	$3/4$	$2/4$	$2/4$	$4/4$	$1/4$	$2/4$	$0/4$	$2/4$	-
peculiar	$2/6$	$1/6$	$3/6$	$3/6$	$4/6$	$4/6$	$4/6$	$3/6$	-

Therefore, the distribution function for this property is $FD_1^{normal} = 3/4$ for the normal population and $FD_1^{peculiar} = 2/6$ for the peculiar population. Due to the distribution function of the features, the list of normal and peculiar features are determined. In normal population, features with a distribution function higher than a constant value, such as α, are added to the list of normal properties.

Also, in peculiar population, features with a distribution function less than constant, such as β, are added to the list of pecliar properties. The parameters α and β control the similarity of the solutions (selection pressure) to select a feature in a normal and peculiar population. Given the number of desirable featres, the genetic algorithm is repeatedly repeated to find DNF of normal features. To help the genetic algorithm to find optimal solutions, a list of the normal and peculiar features is used to generate primary population. So that the initial population contains all the normal features and does not include any peculiar features. By fixing a number of features, this strategy significantly reduces the search space. Applying this limitation in the initial population will change the function of the two combinatory and mutation operators. Therefore, these operators should not add or remove features that violate the criterion of building primary population.

3.4 Adaptive Control of Parameters

Adaptive control of the parameters is in fact a method in the control theory in order to adapt the control system to the variable parameters in the system. The basis of comparative control is based on the estimation of the parameters. In this research, the values of the parameters Mr, α and β change during the implementation of the algorithm. The mutation rate parameter at the beginning of work has a relatively high value and decreases sequentially in the process of running the algorithm.The similarity parameter also initially contains a high percentage of the selected space, but it is reduced by repeating the algorithm and because of the difference between the selected features. The α and β parameters decrease by ε in the case of failure to improve the identification of the normal features in a constant number. This method partially solves the problem of the early integration of the genetic algorithm with constant rate operators. Relationships (1) and (2) are used for comparative control of two parameters of mutation rate and similarity.

$$C_r = \begin{cases} k_1 \dfrac{iter}{MaxIter} \times C_r & f' \ge \overline{f} \\ C_r & f' < \overline{f} \end{cases} \qquad (1)$$

$$M_r = \begin{cases} k_2 \dfrac{iter}{MaxIter} \times M_r & f' \ge \overline{f} \\ M_r & f' < \overline{f} \end{cases} \qquad (2)$$

Where k1, k2 < 1 are two constant values that control the deceleration of Cr and Mr. $\overline{f}$ and f' are the population fitness average of the pre-generation and current generation population, respectively.

4. EXPERIMENTAL RESULTS

The NSL-KDD dataset was used to evaluated the performance of the proposed method [23]. This dataset contains 41 features and 5 classes (a normal class and 4 types of attack classes Dos, R2L, U2R and Probing). To implement the proposed method, the Matlab version 2016a software has been used. The results obtained from the experiments were used to increase the accuracy of the evaluation, a mean of 30 repetitions of the test.

In the implementation, the population size of 25, the number of generations 30, the rate of composition is 0.85 and the rate of mutation is 0.15. The pressure rate of the algorithm is considered in selecting the normal features α = 0.95 and the peculiar features β = 0.90. The number of desirable features selected according to the test and error were at best 23. Selected features of the proposed method for the NSL-KDD dataset are shown in Table 2.

Figure 4 and 5 shows the performance of two classifier of KNN and DT in terms of accuracy and Convergence speed on the chromosomes produced, Respectively.

Table 2. Selected features of the proposed method

No.	attribute name	No.	attribute name
1	duration	14	root_shell
2	protocol_type	16	num_root
3	service	17	num_file_creations
4	flag	18	num_shells
5	src_bytes	19	num_access_files
6	dst_bytes	25	serror_rate
7	Land	27	rerror_rate
8	wrong_fragment	29	same_srv_rate
11	num_failed_logins	30	diff_srv_rate
12	logged_in	34	dst_host_same_srv_rate
13	num_compromised	37	dst_host_srv_diff_host_rate
38	dst_host_serror_rate	-	-

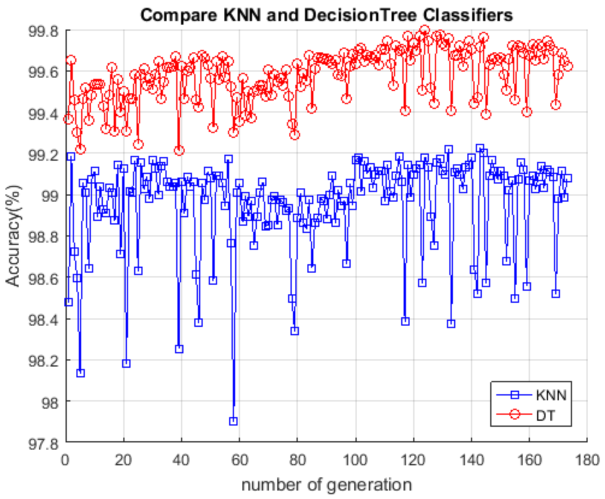


Figure 4. The performance of two classifier of KNN and DT in terms of accuracy on the chromosomes produced

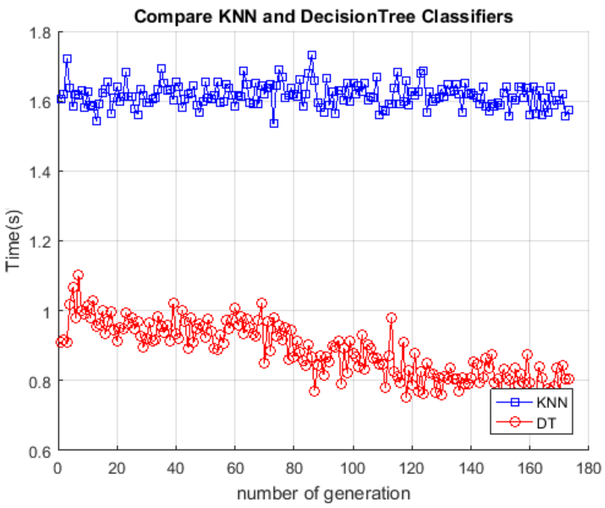


Figure 5. The performance of two classifier of KNN and DT in terms of Convergence speed on the chromosomes produced

The DT classifier method has a better performance than KNN and for this purpose the classification results are for comparison based on DT. In this research, the Accuracy, Precision, Recall and F-measure are used to evaluate the performance of the proposed method. The most important criterion for determining the efficiency of a classification model is Accuracy. This criterion calculates the precision of a single class, defined by relation (3).

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

The FN, TN, FP, and TP parameters represent different states for the classes, which are False Negative, True Negative, False Positive, and True Positive, respectively. The precision criterion shows the precision of the class I classification with respect to all the items that have been proposed for the sample by the classifier. Equation (4) shows how this criterion is calculated.

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \quad (4)$$

The Recall criterion shows the accuracy of the class i classification for all samples with the i label. This criterion is calculated by equation (5).

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \quad (5)$$

The F-measure criterion is calculated from the combination of two precision and recall criteria according to equation (6). This criterion is used in cases where it is not possible to attach special importance to each of the two criteria of Precision and Recall.

$$F - measure = \frac{2 \times Precision_i \times Recall_i}{Precision_i + Recall_i} \quad (6)$$

The effectiveness of intrusion detection systems can be assessed by the proposed criteria. The collision matrix of the Intrusion Detection System data is calculated for each of the four classes of attacks along with the normal class and is shown in Table 3. The table lists the number of records for each attack with the number of predictions.

Table 3. The collision matrix is divided by type of attack

Actual Records			Predicted				
Records Type		Number	Normal	DOS	U2R	R2L	Probe
Normal	Train	67343	67303	6	7	11	16
	Test	9710	9683	3	2	7	1
DOS	Train	45927	8	45909	0	0	10
	Test	7458	1	7454	1	0	0
U2R	Train	52	7	0	44	1	0
	Test	200	2	0	197	3	0
R2L	Train	995	13	0	1	981	0
	Test	2754	9	3	2	2753	2
Probe	Train	11656	22	0	0	1	11633
	Test	2421	7	0	0	0	2413

Table 4, shows the best results classification of the proposed method with different criteria. Results are calculated based on each class against other classes.

In order to further evaluate the above approach, the proposed system performance is compared with other methods of intrusion detection. The methods used to compare the results of their experiments on NSL-KDD data. The results of the proposed method are shown in Table 5 in comparison with the seven methods of intrusion detection.

Table 4. Proposed IDS performance on the NSL-KDD

Records Type		Accuracy	Recall	Precision	F-measure
Normal	Train	99.92	99.89	99.94	99.92
	Test	99.82	99.77	99.87	99.82
DOS	Train	99.93	99.89	99.96	99.93
	Test	99.85	99.73	99.97	99.85
U2R	Train	92.27	99.92	86.66	92.82
	Test	98.68	99.83	97.58	98.69
R2L	Train	99.26	99.93	98.61	99.27
	Test	99.64	99.86	99.42	99.64
Probe	Train	99.87	99.93	99.80	99.87
	Test	99.77	99.82	99.71	99.77

Table 5. Comparison proposed IDS performance with other methods (%)

Methods	Normal	DOS	U2R	R2L	Probe	Accuracy
Fuzzy+ACO [24]	-	-	-	-	-	99.69
ACO+SVM [25]	-	-	-	-	-	98.29
IDS ACO [15]	97.41	99.78	93.51	99.17	74.65	98.9
FARCHD [26]	99.81	98.05	65.38	87.54	95.83	99.00
SIPSO [21]	-	99.80	97.50	82.50	99.70	-
CSM [27]	-	-	-	-	-	99.79
MARS [28]	99.71	99.97	76.00	98.75	**99.85**	92.75
My Method	**99.87**	**99.97**	**97.52**	**99.42**	99.71	**99.81**

As it is known, the proposed method is more accurate than other methods of intrusion detection and for some of the attacks, and in the remaining cases it also provides an accurate precision. In Table 5, the values of each class are based on the values calculated in the relevant research, so some fields may not be presented in the research. The results show that the proposed method works uniformly on all classes and provides the desired accuracy. The reason for this is the selection of features in a hierarchy of high-density populations.

5. Conclusion and Future Work

The accuracy of data mining algorithms depends on the selection of appropriate attributes and the number of records required for learning. The results show that the proposed genetic algorithm chooses appropriate features according to a hierarchical process. The precise and adaptive adjustment of the similarity parameters has led to the identification of more normal and pecliar features, which has led to the effectiveness of the proposed method. The results show that the proposed intrusion detection system has a high accuracy in detecting the intrusion of the DOS type and its underlying attacks. Also, U2R penetration is less accurate than other attacks. The reason for this is the low number of training samples used to test in the dataset. The results of the proposed method showed a precision of %99.81, which is superior to similar algorithms. Another requirement for intrusion detection systems is to find the optimal feature set for each type of attack. Because in this case, the Intrusion Detection System will be able to use only a feature set appropriate to that attack to detect any attack.

As for the future work, intention is to apply the proposed intrusion detection method using complicated classifiers to improve its performance and to combine the proposed method with other population-based algorithms. Analyzing packet payload is recently attracting lots of attention and many researchers report works carried-out in this area. It is notable that feature selection for the payload-based intrusion detection is not mature yet. Intension will be to extract and selection appropriate features from the packet payload to improve the detection rate.

6. REFERENCES

[1] Cisco, I., 2012. Cisco visual networking index: Forecast and methodology, 2011–2016. CISCO White paper, pp.2011-2016.

[2] Lakhina, S., Joseph, S. and Verma, B., 2010. Feature reduction using principal component analysis for effective anomaly–based intrusion detection on NSL-KDD.

[3] Bace, R. and Mell, P., 2001. NIST special publication on intrusion detection systems. BOOZ-ALLEN AND HAMILTON INC MCLEAN VA.

[4] Karabulut, E.M., Özel, S.A. and Ibrikci, T., 2012. A comparative study on the effect of feature selection on classification accuracy. Procedia Technology, 1, pp.323-327.

[5] Safavian, S.R. and Landgrebe, D., 1991. A survey of decision tree classifier methodology. IEEE transactions on systems, man, and cybernetics, 21(3), pp.660-674.

[6] Peterson, L.E., 2009. K-nearest neighbor. Scholarpedia, 4(2), p.1883.

[7] Liu, H., Motoda, H., Setiono, R. and Zhao, Z., 2010, May. Feature selection: An ever evolving frontier in data mining. In Feature Selection in Data Mining (pp. 4-13).

[8] John, G.H. and Langley, P., 1995, August. Estimating continuous distributions in Bayesian classifiers. In Proceedings of the Eleventh conference on Uncertainty in artificial intelligence (pp. 338-345). Morgan Kaufmann Publishers Inc.

[9] Quinlan, J.R., 1993. C4. 5: Programs for Machine Learning Morgan Kaufmann San Mateo. CA Google Scholar.

[10] Witten, I.H., Frank, E., Hall, M.A. and Pal, C.J., 2016. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann.

[11] Werbos, P.J., 1974. Beyond regression: New tools for prediction and analysis in the behavioral sciences. Doctoral Dissertation, Applied Mathematics, Harvard University, MA.

[12] Kratica, J., Kostić, T., Tošić, D., Dugošija, D. and Filipović, V., 2012. A genetic algorithm for the routing and carrier selection problem. Computer Science and Information Systems, (21), pp.49-62.

[13] Xu, Y., Wu, C., Zheng, K., Wang, X., Niu, X. and Lu, T., 2017. Computing Adaptive Feature Weights with PSO to Improve Android Malware Detection. Security and Communication Networks, 2017.

[14] Jovanovic, R. and Tuba, M., 2013. Ant colony optimization algorithm with pheromone correction

strategy for the minimum connected dominating set problem. Computer Science and Information Systems, 10(1), pp.133-149.

[15] Sung, A.H. and Mukkamala, S., 2003, January. Identifying important features for intrusion detection using support vector machines and neural networks. In Applications and the Internet, 2003. Proceedings. 2003 Symposium on (pp. 209-216). IEEE.

[16] Sung, A.H. and Mukkamala, S., 2004, December. The Feature Selection and Intrusion Detection Problems. In ASIAN (pp. 468-482).

[17] Vancea, F., 2014. Intrusion detection in NEAR system by anti-denoising traffic data series using discrete wavelet transform. Advances in Electrical and Computer Engineering, 14(4), pp.43-48.

[18] De la Hoz, E., de la Hoz, E., Ortiz, A., Ortega, J. and Martínez-Álvarez, A., 2014. Feature selection by multi-objective optimisation: Application to network anomaly detection by hierarchical self-organising maps. Knowledge-Based Systems, 71, pp.322-338.

[19] Sequeira, D., 2003. Intrusion prevention systems: security's silver bullet?. Business Communications Review, 33(3), pp.36-41.

[20] Anwar, S., Zain, J.M., Zolkipli, M.F., Inayat, Z., Jabir, A.N. and Odili, J.B., 2015, August. Response option for attacks detected by intrusion detection system. In Software Engineering and Computer Systems (ICSECS), 2015 4th International Conference on (pp. 195-200). IEEE.

[21] Warsi, S., Rai, Y. and Kushwaha, S., 2015. Selective Iteration based Particle Swarm Optimization (SIPSO) for Intrusion Detection System. International Journal of Computer Applications, 124(17).

[22] Gen, M. and Cheng, R., 2000. Genetic algorithms and engineering optimization (Vol. 7). John Wiley & Sons.

[23] Nsl-kdd data set for network based intrusion detection systems.” Available on: http://nsl.cs.unb.ca/KDD/NSL-KDD.html, (March 2009)

[24] Varma, P.R.K., Kumari, V.V. and Kumar, S.S., 2016. Feature Selection Using Relative Fuzzy Entropy and Ant Colony Optimization Applied to Real-time Intrusion Detection System. Procedia Computer Science, 85, pp.503-510.

[25] Mehmod, T. and Rais, H.B.M., 2016. Ant Colony Optimization and Feature Selection for Intrusion Detection. In Advances in Machine Learning and Signal Processing (pp. 305-312). Springer International Publishing.

[26] Elhag, S., Fernández, A., Bawakid, A., Alshomrani, S. and Herrera, F., 2015. On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on Intrusion Detection Systems. Expert Systems with Applications, 42(1), pp.193-202.

[27] Chae, H.S., Jo, B.O., Choi, S.H. and Park, T.K., 2013. Feature selection for intrusion detection using nsl-kdd. Recent Advances in Computer Science, pp.184-187.

[28] Dorigo, M. and Gambardella, L.M., 1997. Ant colonies for the travelling salesman problem. biosystems, 43(2), pp.73-81.