e-Fusion Security Using Face Recognition, Voice Recognition & Code Generator

Md. Ashraful Islam, Md. Shamim Reza Sajib, Md. Ariful Islam Malik

Department of Computer Science and Engineering
Bangladesh University of Business and Technology (BUBT)

Abstract—Security nowadays is a major issue. Day by day, privacy breaches are becoming more prevalent and easy to cause. To reduce and even prevent unauthorized cyber security attacks, organizations are taking various measures in handling these. With state-of-the-art technology flourishing exponentially, we're here to put forth ideas for reducing cyber-attacks to a great extent. The idea is to create a fusion of image processing, voice recognition, and password protection to build a robust and effective anti-security attack system. After face recognition, security steps have been provided through voice recognition and then with the use of keyword generators, the entire process would protect the targeted items in a dynamic way. To ensure the system works at a substantially fast speed, the space and time complexity of the algorithm has been properly pointed. The complexity of data finding has been minimized by using hashing functions.

Index terms- Image Processing, Voice Recognition, and Code Generator

1. Introduction:

The internet is indeed a blessing – however this immense web of connectivity permits vulnerability of private information, and ease of access of this information in unauthorized manners.

Face recognition algorithms are becoming a dominant source of information pool nowadays. The most common technology used behind this is image processing. And ultimately, the use of image processing can be implemented as a very powerful tool for access control. As every human being has a uniquely identifiable facial structure, it is proper to use this feature to the advantage of cyber security. Another very popular tool for securing systems is speech recognition. From the technology perspective, speech recognition has a long history with several waves of major innovations. Most recently, the field has benefited from advances in deep learning and big data. The advances are evidenced not only by the surge of academic papers published in the field, but more importantly by the worldwide industry adoption of a variety of deep learning methods in designing and deploying speech recognition systems. [15] Similar to how each person has unique and specific facial features, they have unique voices as well.

Combining these two advanced technologies with a third but very common and popular tool, it can be determined that the future of cyber security attacks will be reduced to great percentage. The third and final tool in question is the code generator. Automatic Code Generation refers to using programs to generate code that the user would otherwise have to write themselves.

2. Aims and Objectives:

i) To prevent unauthorized access
ii) To determine the advent of cyber-crime.
iii) To determine how bio security reduces the treat of cyber-crimes.

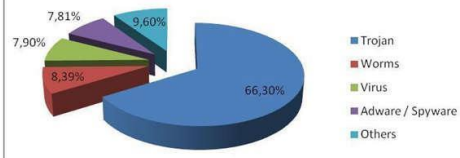

Fig.-1: Various Types of Virus

3. Data Analysis:

Understanding the nature and function of cyber-crimes and network security; the qualitative descriptive mechanism is the most ideal means of collecting and analyzing data due to the flexibility, adaptiveness, and immediacy of the topic. This brings inherent biases, but another characteristic of such research is to identify and monitor these biases, thus including their influence on data collection and analysis rather than trying to eliminate them. Finally, data analysis in an interpretive qualitative research design is an inductive process. Data are richly descriptive and contribute significantly to this research. [17]

4. Code Generator:

Automatic code generation has certain advantages over traditional coding.

i) High Quality
ii) Consistent
iii) Productive
iv) Coding Abstractly

Passive- Passive code generators creates code, then has nothing more to do with the project.
Active- Active code generators create code then keep track of that code during its lifecycle.

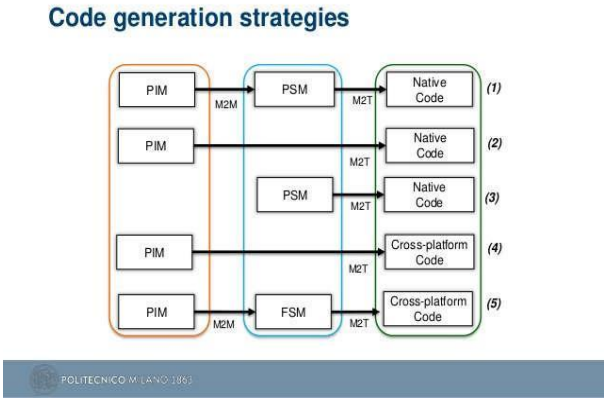

Fig-2: Code Generation [18]

5. Facial Recognition:

Facial expression detection is not a very modern concept. In old days, facial expression detection was in practice in the field of science and technology. Scientists in ancient times used face detection in their researches. In the book "Aristotle's De Motu Anamaliam", [8] written by Martha Carven Nussbaum mentioned the facial symptoms of animals. Facial Expression Recognition Using Back Propagation algorithm is a very common Facial Expression Recognition technique. It is done by Indra

Adji Sulistijono, Zaqiatud Darojah Naoyuki Kubota Abdurahman Dwijotomo, Dadet Pramadihanto from Indonesia and Japan University. At first, it is important that the machine can detect and track the face. This procedure used face capturing to capture moving faces. The method used to recognize the facial expression in this work is back propagation neural network. Ekman and Friesen [8] developed the facial action coding system (FACS) to measure the facial behavior. The FACS codes different facial movements into Action Units (AU) based on the underlying muscular activity that produces momentary changes in the facial expression. An expression is further recognized by correctly identifying the action unit or combination of action units related to a particular expression.[9].

Tells the identification of the nature of the noise is an important part in determining the type of filtering that is needed for rectifying the noisy image. Mehrabian [10] pointed out that 7% of human communication information is communicated by linguistic language (verbal part), 38% by paralanguage (vocal part) and 55% by facial expression. Therefore, facial expressions are the most important information for emotions perception in face to face communication. For classifying facial expressions into different categories, it is necessary to extract important facial features which contribute in identifying proper and particular expressions. Pentland uses a nearest neighbor classifier while feature-line-based methods explained by Li and Lu in [11], replace the point-to-point distance with the distance between a point and the feature line linking two stored sample points. The face regions are located by matching the window patterns at different image locations and scales against the distribution-based face model. Yang et al. [12] proposed a hierarchical knowledge-based method consisting of three levels for detecting the face region and then locating facial components in an unknown picture. Mosaic images of different resolutions are used in the two higher levels. In order to improve the level of detection reliability, the lighting effect is also considered and alleviated for the possible face regions. Two sets of rules based on the characteristics of a human face region are applied to the mosaic images. At third level, the edges of facial components are extracted for the verification of face candidates. However, the computational requirements of these methods may be too high for some applications, which may be unable to detect and locate a tilted human face reliably. Sung et al. [13] proposed an example-based learning approach for locating vertical frontal views of human faces in complex scenes. A decision-making procedure is trained based on a sequence of

Face and non-face examples. Six facial clusters and six non-facial clusters are obtained according to normalized frontal face patterns [14].

6. Proposed Methodology

To identify a person from several persons using Facial Expression Detection is a very efficient modern technology. The basic idea of Facial Expression Detection consists of the following steps of capturing image.

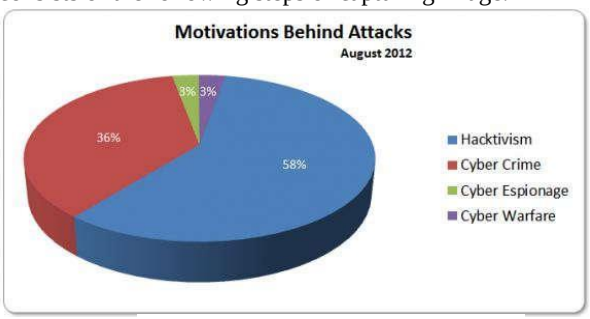

Fig-3: Code Generation

Then feed the feature into the machine i.e. into the neural network. Then another unknown image sequence of unknown person was taken as input. From that, input images are also generated. The similarity of the input image with the trained data

Are calculated and result was shown to the user.

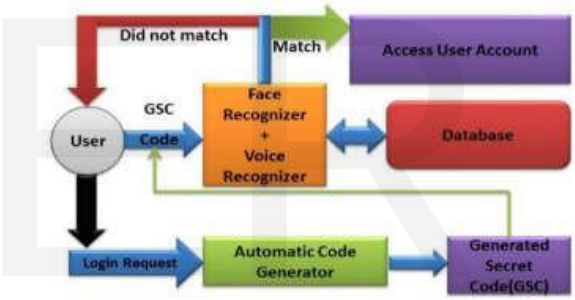

Fig-4: System Architecture

When the user initially registers, and creates an account, his/her image of their face and a voice clip is recorded and taken as input and stored into the system's database. The next time the user attempts to log in, and automatic code is generated which then causes the webcam and speaker to be activated on the user's device. The person then has to speak aloud the generated code. The voice and face recognition software will then match the voice and image input with the ones stored in the database for that specific user. If they match, the user will then be allowed access, otherwise not.

6.1 Capturing Frames

A camera is used to capture people from a specific distance in front of his face. The camera stores data into hard disk of computer. For the identification purpose the captured images are stored in computer and then compared with the storage image to find the matched image. To verify a person if he is an authenticated or not in the area the small image of possibly one gait cycle are storage in a very fast and near storage. The most relative image is considered to be the output.

6.2 Edge Detection

By subtracting the background of the three frames as shown below we get the following temporary image. The background subtracted images are then used to get the binary image. Binary images contain only the black (0) or white (1). From these binary images GEI is extracted. The Canny edge detector uses a filter based on the first derivative of a Gaussian, because it is susceptible to noise present on raw unprocessed image data, so to begin with, the raw image is convolved with a Gaussian filter. The result is a slightly blurred version of the original which is not affected by a single noisy pixel to any significant degree. The ne next is finding the intensity gradient of the image. An edge of an image may point in a variety of directions, so the Canny algorithm uses four filters to detect horizontal, vertical and diagonal edges in the blurred image. The edge detection operator Roberts, Prewitt, Sobel for example returns a value for the first derivative in the horizontal direction (Gy) and the vertical direction (Gx). From this the edge gradient and direction can be determined as following.

$|G| = |Gx| + |Gy|$…….. ….. … . …. …. ….. … (i)

Next is finding the edge direction. The formula for finding the edge direction is given below

$\theta = \tan^{-1}(Gy/Gx)$…… ….. …… ….. …. …. .(ii)

The edge direction angle (θ) is rounded to one of four angles representing vertical, horizontal and the two diagonals (0). Once the edge direction is obtained, the next step is related the edge direction to a direction that can be traced in an image. Finally, hysteresis is used as a means of eliminating streaking. Hysteresis uses two thresholds, a high and a low. Any pixel in the image that has a value greater than T1 is presumed to be an edge pixel, and is marked as such immediately. Then, any pixels that are connected to this edge pixel and that have a value greater than T2 are also selected as edge pixels. We have implemented BPN algorithm to train data that has the following basic steps-

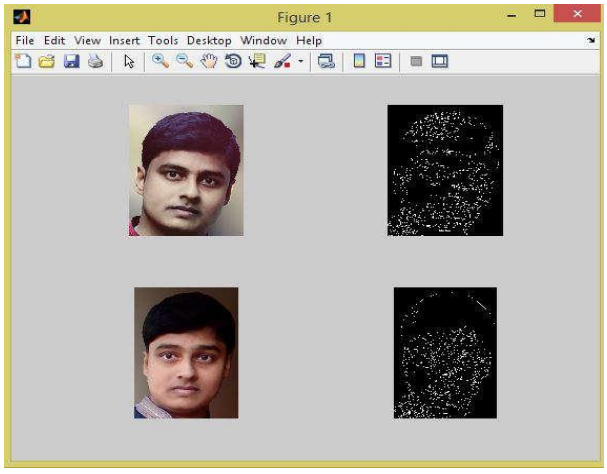

Figure 5: Canny edge detection from captured frame

6.3 Voice Recognition:

To convert speech to on-screen text or a computer command, a computer has to go through several complex steps. When you speak, you create vibrations in the air. The analog-to-digital converter (ADC) translates this analog wave into digital data that the computer can understand. To do this, it samples, or digitizes, the sound by taking precise measurements of the wave at frequent intervals. The system filters the digitized sound to remove unwanted noise, and sometimes to separate it into different bands of frequency (frequency is the wavelength of the sound waves, heard by humans as differences in pitch). It also normalizes the sound, or adjusts it to a constant volume level. It may also have to be temporally aligned. People don't always speak at the same speed, so the sound must be adjusted to match the speed of the template sound samples already stored in the system's memory.[16]

Next the signal is divided into small segments as short as a few hundredths of a second, or even thousandths in the case of plosive consonant sounds -- consonant stops produced by obstructing airflow in the vocal tract -- like "p" or "t." The program then matches these segments to known phonemes in the appropriate language. A phoneme is the smallest element of a language -- a representation of the sounds we make and put together to form meaningful expressions. There are roughly 40 phonemes in the English language (different linguists have different opinions on the exact number), while other languages have more or fewer phonemes.

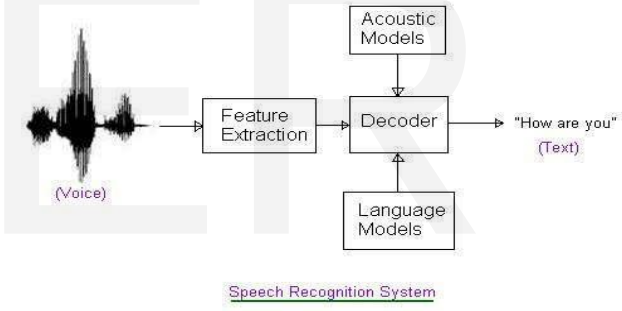

Fig-6: Speech Recognition

7. Attacks on Information: What are the Threats?

Not forgetting that the latter are always a combination of tools that have to do with technology and human resources (policies, training). Attacks can serve several purposes including fraud, extortion, data theft, revenge or simply the challenge of penetrating a system. This can be done by internal employees who abuse their access permissions, or by external attackers to remotely access or intercept network traffic. At this stage of development of the "information-society" and computer technology, hacker's is no longer new.

Some date back to emergence of digital networks, a good few years ago, no doubt as access to electronic communication networks became more widespread, also went-by multiplying the number of those entering "illegally" to them, for different purposes. Another common attack on a computer system is the creation and distribution of malicious computer code, called "viruses".[17]

8. Benefits of network Security

1. Prevents unauthorized users from accessing your network/account. [17]
2. Provides transparent access to Internet-enabled users.
3. Ensures that sensitive data is transferred safely by the public network.
4. Help your managers to find and fix security problems.
5. Provides a comprehensive system of warning alarms attempt to access your network/account.

9. Conclusion:

In conclusion, it can be said that attacks on machines connected to the Internet have increased by a very large extent. Considering customer lists and records of shareholders, trading and marketing materials, marketing strategies and product design, a simple "hack" in the system can mean complete downfall of an entire organization. With advances in technology, no one is safe from an attack by hackers. It is quite effortless to gain control of a machine on the Internet that has not been adequately protected. Companies invest a significant portion of their money in protecting their information, since the loss of irreplaceable data is a real threat to their business. The technology boom in the development of networks, digital communications and advances in software technology allowed the birth of a virtual world whose ultimate expression is the Internet. In order to be safe from these in the modern day, it is necessary to use multiple methods of security simultaneously for maximum effectiveness and protection. Therefore, combining several technologies to build an even more robust security system is the ultimate target.

REFERENCES:

[1] G.Donato, M.S.Bartlett, J.C.Hager, P.Ekman, T.J.Sejnowski, "Classifying Facial Actions", IEEE Trans.Pattern Analysis and Machine Intelligence, Vol. 21, No.10, pp. 974-989, 2009.

[2] R. Chellappa, C.L.Wilson, S.Sirohey, "Human and Machine recognition of Faces: a Survey", Proc. IEEE,Vol. 83, No. 5, pp. 705-741, 2015.

[3] V. Bruce, "What the Human Face Tells the HumanMind: Some Challenges for the Robot-HumanInterface", Proc. IEEE Int.Workshop Robot and Human Communication, pp. 44-51, 2012.

[4]I.A. Essa, A.P. Pentland, "Coding, Analysis, Interpretation, and Recognition of Facial Expressions", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 757-763, 2007.

[5] K.-M. Lam, H. Yan, An analytic-to-holistic approach for face recognition based on a single frontal view, IEEE Trans. Pattern Anal. Mach. Intell. 20 673-686, 2008.

[6]T. Kanade, J.F. Cohn, and Y. Tian, "Comprehensive Database for Facial Expression Analysis", Proc. 4th IEEE Int.Conf. on Automatic Face and Gesture Recognition, p. 46–53, 2000

[7] T. Kanade, J.F. Cohn, and Y. Tian, "Comprehensive Database for Facial Expression Analysis", Proc. 4th IEEE Int.Conf. on Automatic Face and Gesture Recognition , p. 46–53, 2010

[8]Ashish Kumar Dogra, Nikesh Baja, Harish Kumar Dogra, "Facial Expression Recognition using Neural Network withRegularized Back-propagation Algorithm",International Journal of Computer Applications (0975 – 8887) Volume 77 – No.5, September 2013.

[9] Pushpaja V. Saudagare, D.S. Chaudhari, "Facial Expression Recognition using Neural Network –An Overview",International Journal of Soft Computing and Engineering (IJSCE)ISSN: 2231-2307, Volume-2, Issue-1, March 2012.

[10] Surbhi, Mr. Vishal Arora, "The Facial expression detection from Human Facial Image by using neural network", International Journal of Application or Innovation in Engineering & Management (IJAIEM) Volume 2, Issue 6, ISSN 2319 – 4847, 2013.

[11] L. Zhang, Automatic adaptation of a face model using action units for semantic coding of videophone sequences, IEEE Trans. Circuits Systems Video Technol. 8 (6) ,781-795, 1998.

[12] Jagdish Lal Raheja, Umesh Kumar, "Human facial expression detection", International Journal of Application or Innovation in Engineering & Management, 2014.

[13] C. Sanchez Avila, R. Sanchez Reillo, D. de Martin Roche, "Facial Expression based biometrics recognition", IEEE AESS System, October 2002.

[14] G. Yang, T.S. Huang, Human face detection in a complex background, Pattern Recognition 27 (1), 53-63, 1994.

[15]https://en.wikipedia.org/wiki/Automatic_programming

[16]http://electronics.howstuffworks.com/gadgets/high-tech-gadgets/speech-recognition1.htm

[17] Ammar Yassir and Smitha Nayak, Cybercrime: A threat to Network Security, IJCSNS International Journal of Computer Science and Network Security, VOL.12 No.2, February 2012

[18] https://www.slideshare.net/mbrambil/automatic-code-generation-for-cross-platform-multidevice-mobile-apps-an-industrial-experience

Author Details:

Md. Ashraful Islam
Lecturer,
Department of CSE
Bangladesh University of Business and Technology (BUBT)
Email: ashacse42@gmail.com
Phone: +8801723777711

Md. Shamim Reza Sajib
Lecturer,
Department of CSE
Bangladesh University of Business and Technology (BUBT)
Email: sajib1717@gmail.com
Phone: +8801857415874

Md. Ariful Islam Malik
Lecturer,
Department of CSE
Bangladesh University of Business and Technology (BUBT)
Email: malikariful@gmail.com
Phone: +8801722540484