International Journal of Research in Advent Technology, Vol.7, No.5S, May 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

Noise Concerned Accurate Voice Authentication System For Defense Sector

M.Kathiresh[1] , Dr.R.Shankarasubramanian[2]
Research Scholar[1] ,Associate Professor[2]
Department of Computer Science,
Erode Arts and Science college, Erode.
Email :kathireshcs83@gmail.com ,rshankarprofessor@gmail.com

Abstract: Voice authentication in defense sector plays a more important role in predicting the intruders to avoid the security threats. It is more difficult with voices which can be made from multiple persons with the same frequency. Accurate prediction and differentiation of voice signals were requiredto ensure the security of defense sector. Various research methods have been introduced earlier for the accurate prediction of voice signals. Inaccuracy and lack intraining made the existing system might lead to the wrong recognition.This paper presents a proposed Noise concerned Accurate Voice Authentication System (NAVAS)to provide a flexible and convenient environment to recognize the speech in an understandable manner.Noises presents in the speech signal would leads to minimize the accuracy rate. In this research, adaptive noise cancelling approach has been used for the better noise removal for better accuracy. The accuracy of speech recognition has been improved by introducing the SVM based learning methodology which can classify the speech signalsbased onmore relevantmatching.Security of the proposed work is enhanced by encrypting the voice signals using Hybridized AES+SHA2 method. The performance is estimated with experimentation with the matlab simulation environment which ensuresproposed methodperforms better than the existing techniques.
Keywords: NAVAS, noise signals, irrelevancy, Learning signal features, security enhancement

1. INTRODUCTION

The work of commercial and non-profit organizations, financial institutions and enterprises isassociated with the widespread use of information resources and services, access to which is carried out by using modern telecommunication systems [5]. In present economy, profitability and mobility of enterprises and organizations are the most important factors for their existence and development in a competitive environment [7]. To keep the activity of enterprises efficient and profitable, they have to be flexible with respect to changes in the business climate, and their production operations must be carried out rapidly in order to quickly react to these changes [17].

This dynamic environment requires a flexible network organization that supports the rapidly changing conditions and needs of an enterprise during the process of their functioning and development [3]. Under these circumstances, the telecommunications networks of enterprises are becoming an essential part of informational infrastructure that provides the conditions for their successful operation. Network infrastructure must be flexible and scalable, expendable, productive enough, secure and well-managed, to ensure the functioning of the information environment of the enterprise without significant expenditures [14]. Given that access to information and financial resources is performed via open communication channels, special attention should be given to methods and means of protection.

The main global trend is to focus on the construction of telecommunication networks based on TCP/IP architecture, within which secure data links are created and implemented on the basis of the methods and means of encryption and authentication [19]. In addition, this authentication method is the first barrier, which is intended to combat intruders and determines the rights and opportunities of the authorized user. In recent years, biometric features (images) of a user, primarily face features, fingerprints and iris, have been used to improve the reliability of authentication [13].

Biometric authentication has some key advantages over knowledge and token - based authentication techniques [9]. Biometric characteristics are not easily forgotten, like a password, or lost like a key [12]. One can hardly lend someone your finger nor can someone easily steal your eye. That makes them fairly secure, and convenient. Unfortunately, they've had to wait for technology to catch up to the level where it can support their effective use. Only recently has technology provided the statistical, analytical and data processing techniques to support it properly.For the majority of biometric authentication techniques, sophisticated equipment and the physical presence of the person being authenticated is required [6]. For example, fingerprint scanning, pen signatures and retinal scans – not so with voice authentication, where authentication may be give n remotely via a device commonly known as the telephone [4]. Given the use of the correct analytical techniques, a person's voiceprint can

be as unique as any other biometric characteristic, but yet can be used for authentication remotely and has the added benefit of being less personally intrusive than say, subjecting the person to a retinal or fingerprint scan.

Voice authentication is a fairly simple process. To register, a user records sample(s) of their voice which are stored in the authenticating system and become known as their 'voiceprint' [18]. Then, to access this resource subsequently, they supply a sample of their voice to the system, and it decides if it matches their voiceprint before allowing them access.When deciding whether or not to employ a voice authentication system it is important to consider the application [11]. If it is to be used to authenticate a user to administer their bank accounts for example, this is a completely different risk than say accessing their voicemail on their mobile phone. Should a false acceptance result in the banking application, the consequences would be considered much more severe.

The main goal of the research method is to introduce the method that can perform the authentication using voice biometric. This is done by introducing the method namely Noise concerned Accurate Voice Authentication System (NAVAS). The more noises present in the speech signal would lead to failure of accurate recognition of speech. This can be avoided by removing the noises. In this research, adaptive noise cancelling approach can be used for the better noise removal, so that accurate recognition can be attained. The accuracy of speech recognition can be improved by introducing the SVM based learning methodology which can classify the speech signals based on which more relevant matching can be done. Security of the proposed work is enhanced by encrypting the voice signals using Hybridized AES+SHA 2 method.

2. RELATED WORKS

Monrose et al [10] suggested a method to extract repeatable key using user input voice to perform it in cryptographic process of the personal digital assistants (PDA) which resist storage full access attackers. The approach using to captured 12-ceptral coefficient voice vector frames using autocorrelation analysis to produce robust representing user's voice vector, they used (spectral mean subtraction) method. authors also depend on two database sets, the first one is used as phrases training set to produce biometric voice features descriptor and the second use each phrases in first database as a passphrases to evaluate the proposed method by taking any 5 recorded phrase to generate the feature descriptor key and take two recorded trying to regenerate the same key.

Carrara and Adams [2] introduced modified Randomized Biometric Templates RBTs algorithm that's proposed in [1] was performed to extracting reliable features from voice biometrics and generates a cryptographic key. Modified RBTs used the differentiate between individuals in the population called: inter-user variation and similarities of samples of the same individual called: low intra-user variation, with quantization process to correct differentiate readings from the same individual.The algorithm has two main steps: Enroll step, which contain: user features assign, features error correction, a user cryptographic key production, and secure model encoding. The second step is KeyGen step while template will be decrypted by password entered from the user, then biometric sample used to reproduce key of the user. Modified RBTs according to the TI46 database speech samples, generate cryptographic keys with varying Entropy ,FAR and FRR (as an evaluated measure) according to a sensitive parameter k that's indicate the number of biometric voice features distinguished.

Inthavisas and Lopresti [8] proposed speech biometric feature vector as a binary vector which combined with a pseudo-random key for a cryptographic matter, and using Dynamic Time Warping (DTW) to secure template. The first word speech signal considered as keying signal and by performing DTW to the rest words to get averaged result, and then the Euclidean distance between it and the input is determined. A decision will make to determine acceptance or not to the user depending on the comparison between Euclidean distance and a specific threshold. Basically; the system contains two stages training and verification, and it used 13 orders Mel-Frequency Ceptral Coefficient (MFCC) for training and verification. These biometric voice features are used with the data in the template to generate a key binding as a binary vector varying in size according to the number of features that extracted and according to this work; they generate a 128-bit Advanced Encryption Standard (AES) with a 0% EER against attackers in both databases.

Sin et al [16] introducing a system for template update that used for verification by fingerprint. Authors exchange templates with target structure matched inputs. The ERR dropped to 2% after evaluation updating. Their system was relayed on actual mobile phones trading usage from 2009.

Paulini et al [15] introduced a new concept of the human voice called :binarization where the authors through using voice recognition system based on GMM-UBM to draw distinctive binary vectors to encrypt the biometric templates system, however, the authors acknowledge the existence of a type of degrading in biometric achievement resulted from binary representation.

3. NOISE CONCERNED VOICE AUTHENTICATION SYSTEM

The main goal of this research method is to introduce the method that can perform the authentication using voice biometric. This is done by introducing the method namely Noise concerned Accurate Voice Authentication System (NAVAS). The more noises present in the speech signal would lead to failure of accurate recognition of speech. This can be avoided by removing the noises. In this research, adaptive noise cancelling approach can be used

for the better noise removal, so that accurate recognition can be attained. The accuracy of speech recognition has been improved by introducing the SVM based learning methodology which can classify the speech signals based on which can classify the speech signalsbased onmore relevantmatching. Security of the proposed work is enhanced by encrypting the voice signals using Hybridized AES+SHA 2 method.

3.1. Noise Reduction of Voice Signals

The voice communication between humans and machines is the idea people have been thinking about a long time. For high level of the voice communication with the control system it is important to ensure good quality of the speech signal processing with additive noise in real environments. This work describes a proposed method for optimal adjustment parameters of the adaptive filter with an LMS algorithm in the practical application of suppression of additive noise in a speech signal for voice communication with the control system. With the proposed method, the optimal values of parameters of an adaptive filter are calculated which guarantees the stability and convergence of the LMS algorithm.The least mean – square (LMS) algorithm was developed by Widrow and Hoff in 1960. This algorithm is a member of the stochastic gradient algorithms. The LMS algorithm is a linear adaptive filtering algorithm, which, in general, consists of two basic processes:

- The filtering process,
  - which involves computing the output y(n) of the adaptive filter in response to the vector input signal x(n),
  - generating an estimation the error e(n) by comparing this output y(n) with the desired response d(n),
- An adaptive process,
  - which involves the automatic adjustment of the parameters w(n+1) of the filter in accordance with the estimation error e(n),

$y(n) = w^T(n)x(n)$

w(n) the tap – weight vector,

$e(n) = d(n) - y(n)$

$w(n+1) = w(n) + 2\,\mu\, e(n)\, x(n)$

$w(n) = [w_0(n)\ w_1(n) \ldots w_{M-1}(n)]^T$ tap weights,

w(n+1) the tap – weight vector update,

μ step size parameter.

The LMS algorithm adapts the filter tap weights so that e(n) is minimized in the mean-square sense. When the processes x(n) and d(n) are jointly stationary, this algorithm converges to a set of tap weights which, on average, are equal to the WienerHopf solution

$$w = w_0 = R^{-1}p$$

p the cross – correlation vector M x 1 of the input signal x(n) and the desired signal d(n)

$$p = E\,[x(n)\, d(n)] = [p_0 p_1 \ldots p_{M-1}]^T$$

R the Toeplitz autocorrelation matrix M x M of the input signal.

$$R = E\,[x(n)x^T(n)]$$

If prior knowledge of the tap-weight vector w(n) is available, will use it to select an appropriate value for w0. Otherwise, set w0 = 0. The LMS algorithm can be used to solve the Wiener-Hopf equation without finding matrix inversion R-1. It does not require the availability of the autocorrelation matrix of the filter input and the cross correlation between the filter input and its desired signal.

3.2. Learning Voice Signal Features Using Support Vector Machine

An important objective of extracting the features is to compress the speech signal to a vector that is representative of the meaningful information it is trying to characterize. In these works, acoustic features namely MFCC features are extracted.Mel Frequency CepstralCoefficients (MFCCs) are short-term spectral based and dominant features and are widely used in the area of audio and speech processing. The mel frequency cepstrum has proven to be highly effective in recognizing the structure of music signals and in modeling the subjective pitch and frequency content of audio signals. The MFCCs have been applied in a range of audio mining tasks, and have shown good performance compared to other features. MFCCs are computed by various authors in different methods. It computes the cepstral coefficients along with delta cepstral energy and power spectrum deviation which results in 26 dimensional features. The low order MFCCs contains information of the slowly changing spectral envelope while the higher order MFCCs explains the fast variations of the envelope.

A machine learning technique which is based on the principle of structure risk minimization is support vector machines. It has numerous applications in the area of pattern recognition. SVM constructs linear model based upon support vectors in order to estimate decision function. If the training data are linearly separable, then SVM finds the optimal hyper plane that separates the data without error. Fig. 1 shows an example of a non-linear mapping of SVM to construct an optimal hyper plane of separation. SVM maps the input patterns through a non-linear mapping into higher dimension feature space. For linearly separable data, a linear SVM is used to classify the data sets. The patterns lying on the margins which are maximized are the support vectors.

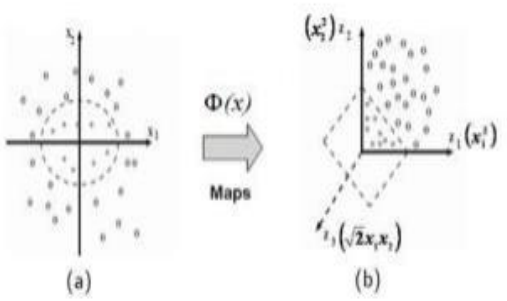

Fig -1: Example for SVM Kernel Function Φ(x) Maps 2-Dimensional Input Space to Higher 3-Dimensional Feature Space. (a) Nonlinear Problem. (b) Linear Problem.

The support vectors are the (transformed) training patterns and are equally close to hyperplane of separation. The support vectors are the training samples that define the optimal hyperplane and are the most difficult patterns to classify. Informally speaking, they are the patterns most informative of the classification task. The kernel function generates the inner products to construct machines with different types of non-linear decision surfaces in the input space.

3.3. Signal Encryption Before Voice Communication

The AES key expansion algorithm takes the input which is a 4-word key and produces a linear array of 44 words. Each round uses 4 of these words. Each word contains 32 bytes which means each sub-key is 128 bits long. The key is copied into the first four words of the expanded key. The remainder of the expanded key is filled in four words at a time. Each added word w[i] depends on the immediately preceding word, w[i−1], and the word four positions back w[i−4]. In three out of four cases, a simple XOR is used. For a word whose position in the w array is a multiple of 4, a more complex function is used. SHA-2 is a type of cryptographic hash functions that is designed by the NSA (U.S. National Security Agency). SHA stands for Secure Hash Algorithm. Cryptographic hash functions are the mathematical operations that run on the digital data. A person can determine the data's integrity, by comparing the computed "hash" to a known and expected hash value. SHA-256 can accept messages with arbitrary lengths up to 264-bit. The Hash computation produces a final digest message of 256 bits that depends upon the input message, composed by multiple blocks of 512-bit each. This input block is expanded and it is fed to the 64 cycles of the SHA-256 function in words of 32-bit each.

The proposed architecture i.e., Hybrid cryptosystem is represented in the following fig. It depicts the integration of AES algorithm with the SHA-2 hash function. The AES and SHA-2 algorithms are explained in previous sections. Accordingly the two algorithms are designed and integrated as shown in the Fig 2.

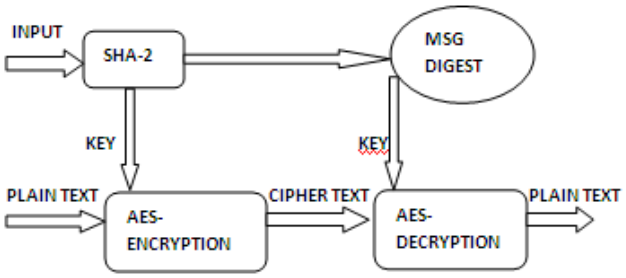

Fig.2 Hybrid Cryptosystem

An input of arbitrary length is given to the SHA-2 module. A message digest of fixed length is generated which is 256 bits. This message digest is used in the encryption and decryption process as key. As shown in the fig.8, after generating the message digest, it is given as a key for the encryption of the plain text which in turn generates a cipher text. Later by making use of same key, decryption is performed to retain back the original plain text. AES itself is a strong security mechanism. Since SHA-2 is being used here along with the AES, this design ensures higher security since complexity of the design increases. Here security is given in terms of complexity.

4. RESULTS AND DISCUSSION

The evaluation of the proposed Noise concerned Accurate Voice Authentication System (NAVAS) developed using matlab simulation environment. The voiceprint is evaluated against each other to establish which performs better than the other.

Sensitivity

Sensitivity evaluates the percentage of actual positive which classifies actual voice as actual. The sensitivity is defined as below:

$$Sensitivity = \frac{T_p}{T_p + F_n} \quad \ldots(1)$$

where T_p defines the voice correctly as authenticated voice. F_p defines the intruder voice incorrectly as the intruder voice. F_n defines the intruder voice incorrectly as the authenticated voice. T_n defines the intruder voice correctly as intruder voice.

Precision

Precision is defined as the proportion of the true positives against both true positives and false positives results for voice signals. It is defined as follows

$$Precision = \frac{T_p}{T_p + F_p} \quad \ldots(2)$$

Accuracy

Accuracy is defined as the overall correctness of the model and is calculated as the sum of actual classification parameters $(T_p + T_n)$ separated by the total number of classification parameters $(T_p + T_n + F_p + F_n)$

$$\text{Accuracy} = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad \ldots(3)$$

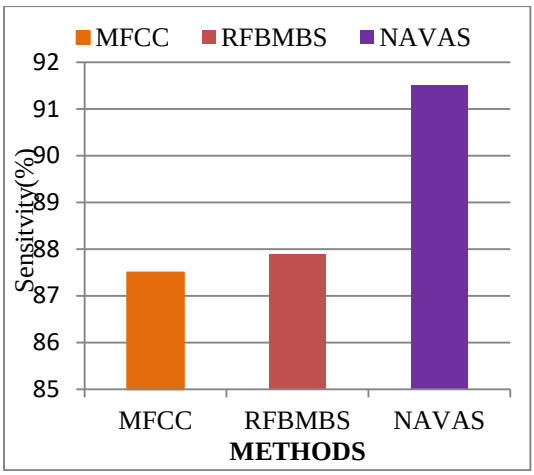

Fig 3. Performance Evaluation of Sensitivity with Different Methods

In this sensitivity evaluation the proposed method is better than the existing methods. The MFCC is 87.5 %, RFBMBS (Retina and Fingerprint Based Multi-Biometric Systems) method metric is 87.9% and the NAVAS is 91.5%. The figure 3 shows the better sensitivity analysis compare than the other methods.

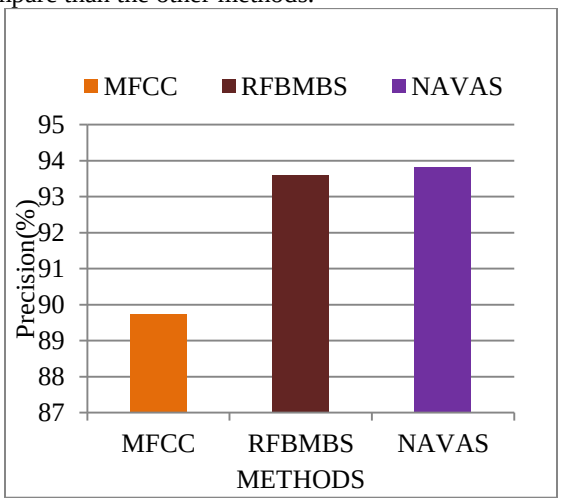

Figure 4 Performance Evaluation of Precision with Different Methods

In this precision evaluation the proposed method is better than the existing methods. The MFCC method metric is 89.7 %, RFBMBS method metric is 93.5% and the NAVAS is 93.8%. In the figure 4 shows the better sensitivity analysis compare than the other methods.

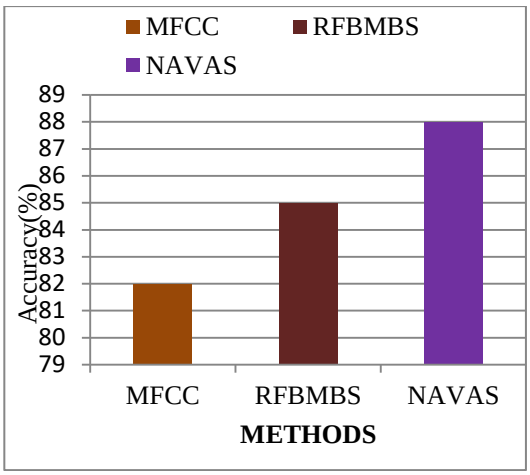

Fig5: Performance Evaluation of Accuracy with Different Methods

In this accuracy evaluation the proposed method is better than the existing methods. The MFCC method metric is 82.00 %, RFBMBS method metric is 85.00% and the NAVAS is 88.00%. In the figure 5 shows the better sensitivity analysis compare than the other methods.

5. CONCLUSION

In this research, Noise concerned Accurate Voice Authentication System (NAVAS) is proposed for accurate voice authentication. Noise elimination is important in the speech signal for better accurate recognition of speech. In this research, adaptive noise cancelling approach which provides better noise removal and the accurate recognition is attained. The accuracy of speech recognition has been improved by introducing the SVM based learning methodology which can classify the speech signalsbased onmore relevantmatching.Security of the proposed work is enhanced by encrypting the voice signals using Hybridized AES+SHA 2 method. The experimentation is accomplished with the matlab simulation environment and the performance is evaluated by compared with the existing methods. Finally,the proposed method ensures better performance than the existing techniques.

REFERENCE

[1] Ballard, L., Kamara, S., Monrose, F., & Reiter, M. K. (2008, October). Towards practical biometric key generation with randomized biometric templates. In *Proceedings of the 15th ACM conference on Computer and communications security*(pp. 235-244). ACM.

[2] Carrara, B., & Adams, C. (2010, August). You are the key: generating cryptographic keys from voice biometrics. In *2010 Eighth International Conference on Privacy, Security and Trust*(pp. 213-222). IEEE.

[3] Davies, A., & Brady, T. (2016). Explicating the dynamics of project capabilities. *International Journal of Project Management*, *34*(2), 314-327.

[4] Dewangan, S. K. (2015). Human Authentication Using Biometric Recognition. *International Journal of*

Computer Science & Engineering Technology (IJCSET), *6*(04), 240-245.
[5] Dunlop,J.(2017). Telecommunications engineering. Routledge.
[6] Durst, D. I., Kaish, N., & Fraser, J. (2018). *U.S. Patent Application No. 15/804,923*.
[7] Epstein, M. J. (2018). *Making sustainability work: Best practices in managing and measuring corporate social, environmental and economic impacts*. Routledge.
[8] Inthavisas, K., &Lopresti, D. (2011, May). Speech biometric mapping for key binding cryptosystem. In *Sensing Technologies for Global Health, Military Medicine, Disaster Response, and Environmental Monitoring; and Biometric Technology for Human Identification VIII* (Vol. 8029, p. 80291P). International Society for Optics and Photonics.
[9] Katsini, C., Belk, M., Fidas, C., Avouris, N., & Samaras, G. (2016, November). Security and usability in knowledge-based user authentication: a review. In *Proceedings of the 20th Pan-Hellenic Conference on Informatics* (p. 63). ACM.
[10] Monrose, F., Reiter, M. K., Li, Q., Lopresti, D. P., & Shih, C. (2002, August). Toward Speech-Generated Cryptographic Keys on Resource-Constrained Devices. In *USENIX Security Symposium* (pp. 283-296).
[11] North, R., Norris, J., & Chu, F. (2017). *U.S. Patent No. 9,639,682*. Washington, DC: U.S. Patent and Trademark Office.
[12] Ogbanufe, O., & Kim, D. J. (2018). Comparing fingerprint-based biometrics authentication versus traditional authentication methods for e-payment. *Decision Support Systems*, *106*, 1-14.
[13] Olade, I., Liang, H. N., & Fleming, C. (2018, October). A Review of Multimodal Facial Biometric Authentication Methods in Mobile Devices and Their Application in Head Mounted Displays. In *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)* (pp. 1997-2004). IEEE.
[14] Pastushenko, M., &Faizulaieva, O. (2016, October). Employment of phase characteristics of user voice signal in authentication systems. In *2016 Third International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T)* (pp. 205-206). IEEE.
[15] Paulini, M., Rathgeb, C., Nautsch, A., Reichau, H., Reininger, H., & Busch, C. (2016, June). Multi-bit allocation: preparing voice biometrics for template protection. In *Odyssey 2016* (pp. 291-296).
[16] Sin, S. W., Zhou, R., Li, D., Isshiki, T., &Kunieda, H. (2012). Narrow fingerprint sensor verification with template updating technique. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, *95*(1), 346-353.
[17] Teece, D., Peteraf, M., &Leih, S. (2016). Dynamic capabilities and organizational agility: Risk, uncertainty, and strategy in the innovation economy. *California Management Review*, *58*(4), 13-35.
[18] Vegh, L. (2018, February). Cyber-physical systems security through multi-factor authentication and data analytics. In *2018 IEEE International Conference on Industrial Technology (ICIT)*(pp. 1369-1374). IEEE.
[19] Wollschlaeger, M., Sauter, T., &Jasperneite, J. (2017). The future of industrial communication: Automation networks in the era of the internet of things and industry 4.0. *IEEE Industrial Electronics Magazine*, *11*(1), 17-27.