Secure EVMs against Tampering with AES Encryption and Hashing

A V R M Koushik, A Chandrahas Reddy, G Avinash Reddy, C Fancy

Abstract*: Conducting elections in a country where population is increasing rapidly like wildfire, is the most challenging aspect for the government. Some people take advantage of this aspect and create disturbances by involving in tampering the Electronic Voting Machines (EVMs). This is a situation the authorities face in conducting elections, which is emerging as a problem to be taken care with an utmost priority. This paper presents a method to know if the EVMs have tampered or not and when the EVM is tampered if they are found to be tampered. This method uses the powerful AES Algorithm, an advanced symmetric key cryptographic technique that provides better security and efficiency. It is paired with various hashing algorithms (MD5, SHA-1, SHA-256, and Whirlpool) for validation. Hence, the proposed model defines a unique way to secure the EVMs from tampering. In addition, the method is experimented with various key sizes and hashing algorithms.

Keywords: Cryptography, Hashing, Encryption, EVM

I. INTRODUCTION

In the modern era, everyone wants everything to be secure. This paper deals with the problem of tampering the electronic voting machines (EVMs) using the combination of cryptographic techniques and the hash functions to preserve data integrity, authentication, and confidentiality. Cryptography is the technique to convert the plain text into cipher text and vice-versa. Based on the level of the problem its corresponding cryptographic technique is selected. This problem should be taken care with utmost priority. This paper gives a brief idea about how to end this problem with the implementation of AES, a symmetric key cryptographic technique along with various hashing algorithms. Hashing is a function that converts the arbitrary length into a data with fixed length called hash value.

A. Advanced Encryption Standard

The Advanced Encryption Standard (AES), also known by its original name Rijndael is a specification for the encryption of the electronic data. AES is the advanced version of 3-DES. AES uses a private key symmetric block cipher algorithm providing a 128-bit data having a 128-bit/ 192-bit/ 256-bit as key lengths, which are iterative rather than Feistel cipher. It is much stronger and more faster than Triple-DES. The Rijndael algorithm is clean, fast and good security margin [1].

Revised Manuscript Received on April 25, 2019.

A V R M Koushik, Department of Information Technology, SRM Institute of Science and Technology, Tamil Nadu, India.

A Chandrahas Reddy, Department of Information Technology, SRM Institute of Science and Technology, Tamil Nadu, India.

G Avinash Reddy, Department of Information Technology, SRM Institute of Science and Technology, Tamil Nadu, India.

Ms.C Fancy, Assistant Professor, Department of Information Technology, SRM Institute of Science and Technology, Tamil Nadu, India.

B. Hashing

Hashing is known as a form of cryptographic security that converts any sized input into an irreversibly fixed length. Means it reduces the arbitrary input to a fixed length. Usually, the hash function is public and no key usage. The main purpose of the hash is used to detect the changes to message and most often to provide a digital signature. One block of data is taken as input to produce a hash value of n bits repeatedly.

II. RELATED WORK

In the existing model, the EVMs are designed such that, they produce only the result of the election. They only give the total number of voted casted and the number of votes each candidate is casted. These machines were easy to use when they were introduced back in 1980's. But as the science advanced remarkably, This has led to decrease in the degree of transparency to voters. To overcome this VVPAT was introduced which will show the voter to whom the vote is casted. Based on the conclusion of the paper "Security Analysis of India's Electronic Voting Machines" [2], it is proven that they lack the security features for EVMs and they are vulnerable to serious attacks, it is potentially difficult to trace the attacks and it is a serious issue to deal with [2]. The main drawback of this model is it does not provide when the tampering is done or if the EVM has tampered or not and if the EVM is tampered we cannot have a genuine result [3].

III. BASIC CRYPTOGRAPHY TECHNIQUES

A. Advanced Encryption Standard

The size of the block that is taken as input in AES is 128-bit, with a variable key of length 128-bit, 192-bit and 256-bit (the AES technique used is given by the length of the key size used) [4]. AES encryption and AES decryption take a single block of 128-bit as input. AES has three steps Key Expansion, Encryption, Decryption in it.

B. Key Expansion

In the key expansion step, a set of round keys are produced in the form of a matrix, where each matrix is distinct. In each round, a round key (matrix) is used as an input for AddRoundKey transformation.

C. Encryption

The process of AES encryption takes place in the form rounds and the number of rounds that should be performed in encryption is decided by the size of the key used.

In AES-128, AES-192 and AES-256 there are 10, 12 and 14 rounds respectively that iterate in each technique. The first N-1 rounds in AES has four distinct transformation functions [5] (a single permutation function and three substitution function):

i. SubBytes (The byte by byte substitution in a block is performed using S-box)
ii. ShiftRows (It involves simple permutation function)
iii. MixColumns (Arithmetic is used for Substitution over GF (2^8))
iv. AddRoundKey (A portion of the expanded key and the current block are applied against a bitwise XOR).

All the first n-1 rounds in encryption have 4 rounds but there are only three transformations (SubBytes, ShiftRows, AddRoundKey) present in the last round and an AddRoundKey transformation before the SubBytes transformation in the last round. The input and output produced in each transformation is a matrix. The output of the final round is ciphertext.

- **Decryption**

AES decryption is not as similar as AES encryption. The way in which the transformations in decryption are implemented is different from the way the transformations are implemented in encryption but key scheduling in encryption and key scheduling in decryption is the same. But, the main disadvantage is, software must run in two different modules to perform encryption and decryption. Thus, the first two transformations implemented in encryption should be interchanged and then inversed and the next two transformations in encryption need to be interchanged and inversed. The decryption round has InvShiftRows, InvSubBytes, AddRoundKey, InvMixColumns as the four functions [5].

Only the bytes sequence in the state get affected by InvShiftRows transformation but does not alter or depend on byte contents to perform the transformation. The byte contents in the state get affected by InvSubBytes transformation but do not alter or depend on byte sequence to perform its transformation. As one affects the byte sequence but not affect byte contents and the other affects byte contents but not byte sequence, these both steps can be interchanged and inversed.

The transformations InvMixColumns and AddRoundKey do not alter or depend on byte sequence in state and if the key is viewed as a sequence of words, both the AddRoundKey and InvMixColumns operate one column at a time on the state. With respect to column input, both the operations are linear.

D. Hashing

Hashing converts any length of data to a fixed length. There are many hash functions available that we can use to get a fixed size arbitrary output. Examples of the hash functions are MD2, MD4, and MD5, to a message-digest. Secure Hash Algorithms (SHA-1, SHA-2, and SHA-256) output a large size of message digest and its working is similar to MD4. Whirlpool is a 512-bit hash function. The main application of the hash function is to secure the password and for storing the data (which is hashed). Hash functions have the properties of pre-image resistance (computationally hard to reverse a hash function), second pre-image resistance (should be highly complex to find a hash value of the input) and collision resistance (should be difficult to find an input, which results in same hash as that of the original input).

- **MD5**

The MD5, message-digest algorithm is the most frequently used hash function. It has block and digest of size 512-bit, 128-bit respectively with 4 rounds.

- **SHA-1**

Secure Hash Algorithm-1 has block and digest of size 512-bit, 160-bit respectively rendered as a hexadecimal number with 80 rounds. SHA-1 is developed from MD4. SHA-1 is prone to length extension attacks.

- **SHA-256**

SHA-256 has block and digest of size 512-bit and 256-bit respectively with 64 rounds. The hash length produced is of 256-bit length. This makes it one of the strongest hash functions available. Coding SHA-256 is not that complex to SHA-1 and has not compromised yet. The 256-bit key makes it a good partner for AES in implementing Digital Signatures [6].

- **Whirlpool**

Whirlpool is a block cipher functions developed after the square block cipher is designed. It is a substantially modified AES. It has a block size of 512 bit and 10 rounds in it [7].

IV. PROPOSED MODEL

In this model, we can secure the EVMs from tampering using the methods of encryption and hashing using a secondary partition. In the existing model, it provides information only about the total number of votes casted for each candidate. But, this proposed model gives the number of times the button is accessed right away the polling is started till it ends. This model also contains the counting model that is present in the existing model. Each button is provided with a port id in hexadecimal bits, and once a button is accessed, it stores the id of the button in the secondary partition.

Once the election end, the operator initiates the process to seal the EVM. First, a duplicate dataset is created and stored within a partition in the secondary partition. These datasets contain the order in which the buttons are accesses throughout the process of polling. A hash function is employed to the duplicate dataset of port ids and the dataset is replaced by the hash value, which is obtained from the hash function. Then the original dataset is taken as input for the AES Encryption technique, which provides a ciphertext using the pre-defined key. This ciphertext replaces the existing dataset. This completes the process of sealing the EVMs and they are transferred to a secure place. Here, the Phase-I ends.

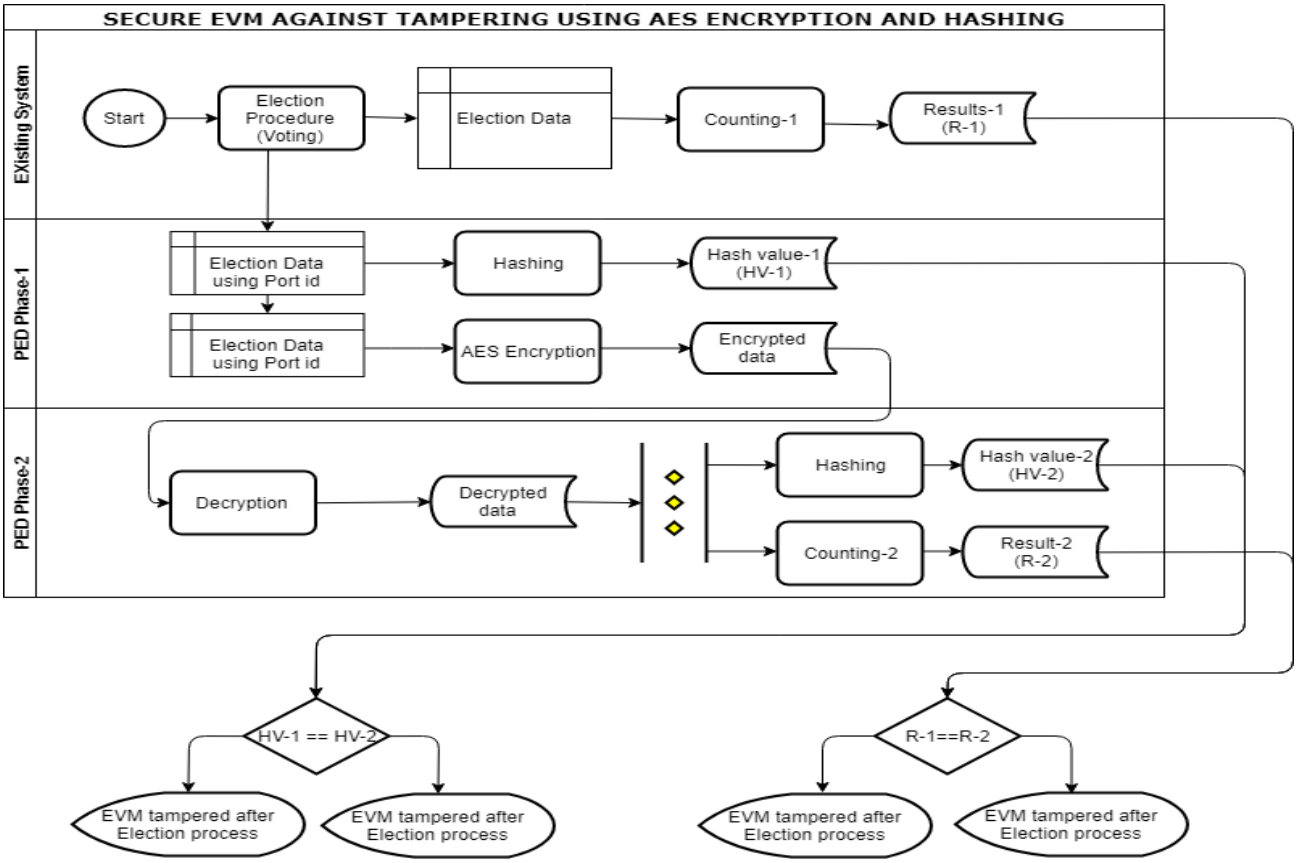

Fig 1. PED Methodology

When the EVMs are used for counting the result, first the ciphertext is decrypted using AES Decryption process with the same pre-defined key that is used during the process of Encryption. And this decryption gives the original plain text which was created initially after the elections.

The same hash function used before AES Encryption is used again to get the hash value. In the obtained plain text, the hexadecimal bits are divided into blocks of bits with each block being the size of default port id. Now segregating the dataset based on the port id, we can find the number of votes casted for each candidate. This way gives the genuine result free from tampering. This is Phase-II. This model is referred as Port Encryption Decryption Model (PED Model).

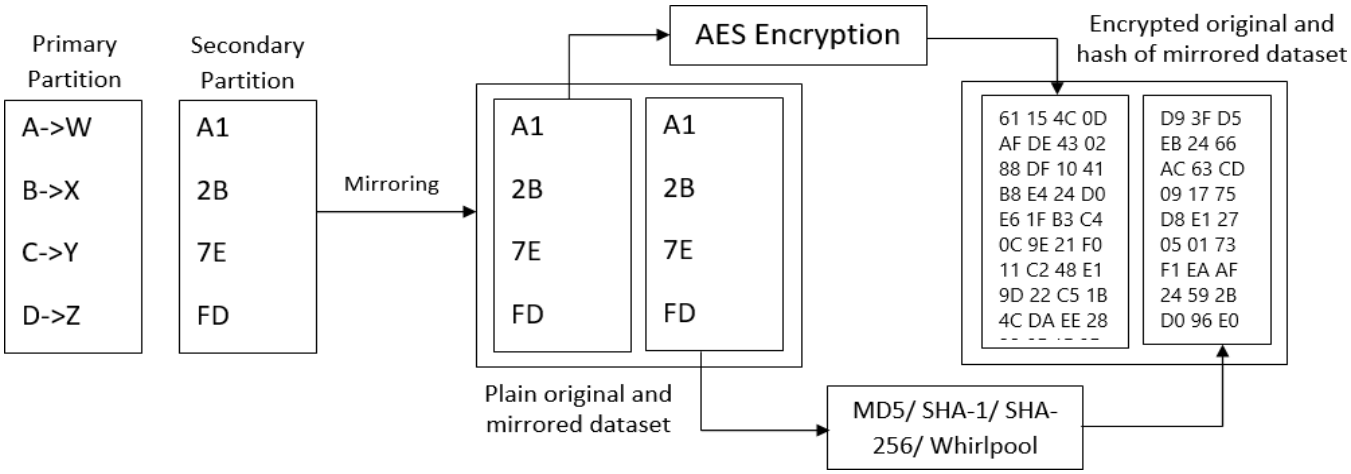

Fig 2. Phase I

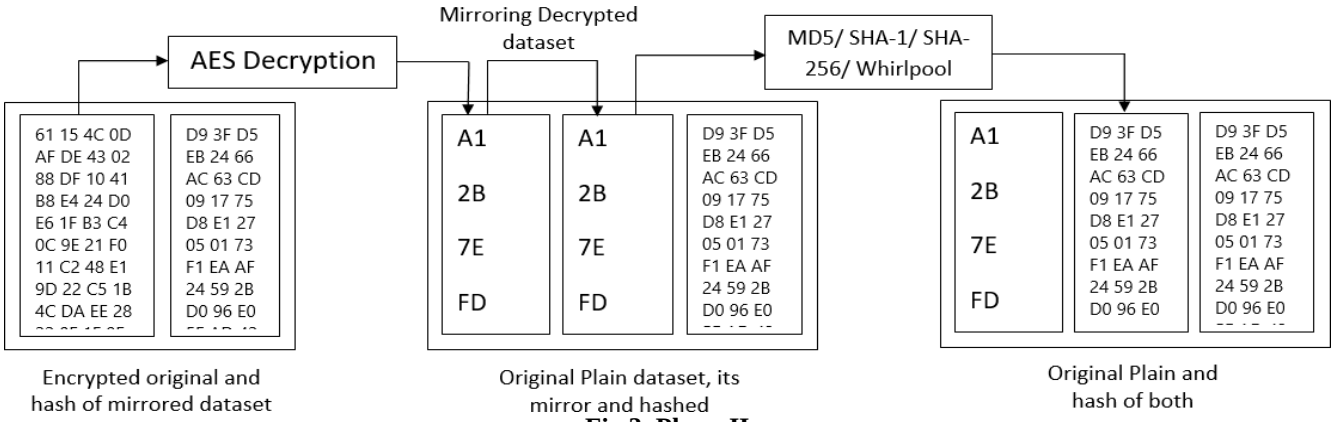

Fig 3. Phase II

A. Ways to check if the EVM has Tampered

i. The hash value obtained after decryption is compared with the hash value obtained before the encryption, stored in the secondary partition, which replaced the duplicate dataset.

ii. The result obtained from the PED Model and the result from the existing model.

B. When is the EVM Tampered?

i. If the values of the hash value obtained after decryption are not the same as the hash value obtained earlier in PED Model, then it means there was an act of tampering. If the hash value obtained before encryption and after decryption is same then it means there was no act of tampering after the election.

ii. If the election result obtained from the existing model is not the same as the result obtained in the PED Model, then it means there was an act of tampering before the elections. If the election result obtained from the existing model is the same as the result obtained in the PED Model, then it means there was no act of tampering done before the elections.

V. ANALYSIS

The analysis of encryption time and decryption time is based on the input of 4096 bits in AES-128, AES-192 and AES-256 Algorithms.

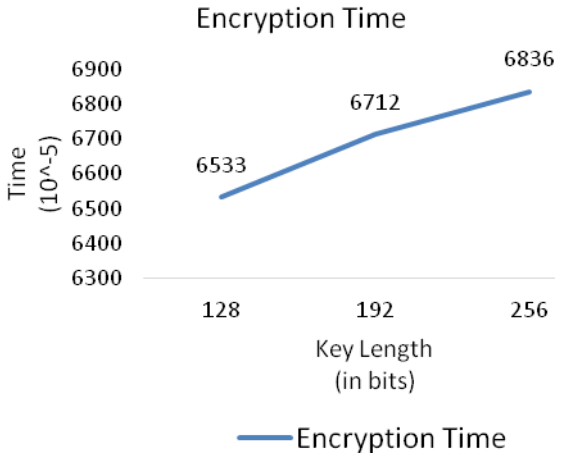

Fig 4. Encryption Analysis

On a detailed study of the encryption analysis graph, it is clear that the encryption time (the unit of time is seconds) is directly proportional to the size of the key used.

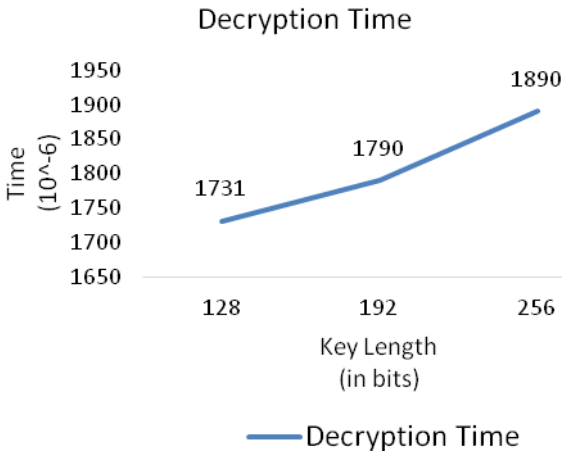

Fig 5. Decryption Analysis

On a detailed study of the decryption analysis graph, it is clear that the decryption time (the unit of time is seconds) is directly proportional to the size of the key used.

On a deep analysis of both the graphs we can also depict that the decryption time taken to convert a cipher text that was generated from a plain text back into its original text is always less than the encryption time taken for a plain text to convert into cipher text for a key length (128/192/256 bit key length).

The memory usage graph signifies how the AES Algorithm for different key lengths (4096 and 16384) and different input sizes uses the memory. The memory consumed by the model is directly proportional to the size of the key and size of the input

Memory Usage Analysis

Table I Memory usage for diferent text-length (in bits) and key-length (in bits)

Plain Text	128	192	256
16384	2150728	2194640	2234640
4096	1804168	1804576	1804785

The output size of the hash functions MD5, SHA-1, SHA-256, and Whirlpool are 128-bit, 160-bit, 256-bit, and 512-bit respectively. The length of hash value obtained attacked by malware that can lead to non-functioning of the EVM. This issue can be taken to work on it as a further study for the model from a hash function is always the same and is not dependent on length of the input. Coming to their cryptanalysis, MD5 can be

cracked within seconds because of its shorter output length. But, for SHA-1, it is theoretically breakable but has never been cracked [8]. SHA-256 and Whirlpool are the hash functions, these hash functions are cracked neither theoretically nor practically. These cannot be broken because of their computational power and their output size (256 and 512) which is not even possible to find the possible ways (2256 and 2512) for applying brute-force [9].

VI. CONCLUSION

The proposed PED model is mainly based on the EVMs in the developing countries or any institutions where the authorities face many claims about the EVMs being tampered. Hence we proposed a methodology to find if the EVMs are tampered or not and if it is tampered, it helps to find out if it is tampered before the elections or between the election and the counting date. We used a high level encryption algorithm that is AES Algorithm and Whirlpool hashing technique to preserve the integrity. We performed the analysis on the time taken to run the encryption process and hashing, time taken to run the decryption process and hashing and memory consumption based on the usage of different key length. From the analysis we pointed that this is one of the best hybrid cryptographic analytics to preserve the data integrity and authentication.

FUTURE ENHANCEMENTS

As the model involves the cryptographic techniques and hash techniques, there can be a chance of EVM being

REFERENCES

1. Herman Isa, Iskandar Bahari, Hasibah Sufian, Muhammad Reza Z'aba "AES: Current Security and Efficiency Analysis of its Alternatives", IEEE, 2011
2. Scott Wolchok, Eric Wustrow, J. Alex Halderman, The University of Michigan, Hari k. Prasad, Arun Kankipati, Sai Krishna Sakhamuri, Vasavya Yagati, Netindia, (P)Ltd., Hyderabad "Security Analysis of India's Electronic Voting Machines", ACM, 2010
3. Tushar Puri, Jaspreet Singh, Hemant Kaushal, "Prototyping of Indian Electronic Voting Machine", IJERD, 2017
4. Mahmoud Alfadel, El-Sayed M. El-Alfy, Khaleque Md Aashiq Kamal, "Evaluating Time and Throughput at Different Modes of Operation in AES Algorithm", IEEE, 2017
5. Luanlan, "The AES Encryption And Decryption Realization Based On FPGA", IEEE, 2011
6. Kadek Dwi Budi Utama, M. Rizqia Al-Ghazali Q, Leonardus Irfan Bayu Mahendra, Guruh Fajar Shidik, "Digital Signature using MAC Address based AES128 and SHA-2 256-bit", IEEE, 2017
7. P. Kitsos, O. Koufopavlou, "Whirlpool Hash Function: Architecture and VLSI Implementation", IEEE©, 2004
8. Zhenqi Wang, Lisha Cao, "Implementation and &omparison of Two Hash Algorithms", IEEE©, 2011
9. Jia Yu, Bo Ren, "Compression Function Design of Secure Hash Algorithm Based on Block Cipher", IEEE©, 2012

AUTHORS PROFILE

A V R M Koushik currently pursuing his B.Tech, 3rd year in Department of Information Technology at SRM Institute of Science and Technology. He is keen on the works which involve implementation of cryptography and security in the areas of IoT and web development. He has zeal to learn new things and interested in managing things.

A Chandrahas Reddy currently pursuing his B.Tech, 3rd year in Department of Information Technology at SRM Institute of Science and Technology. His area of interest is Cryptography, Networking, Cyber Forensics and Data Analytics. Implemented projects in IoT and Networking. He is an avid reader of scientific and Computer related journals. His aim is to contribute himself to the scientific world.

G Avinash Reddy currently pursuing his B.Tech 3rd year in Department of Information Technology from SRM Institute of Science and Technology. His area of interest is Cryptography, IOT, Security and Software Development. He implemented projects like Remote Water quality analysis using IOT, Aadhar based online voting machine.

Ms.C Fancy completed her B. Tech in IT and ME,in Computer Science Engineering. She is working in the department of Information Technology, SRM Institute of Science and Technology as an Assistant Professor since 2013. She is currently pursuing Phd in the SDN Domain. Her research interests include cryptography, medical imaging and cyber security.