UNFORGETTABLE USER DEFINED SEED PHRASE FOR CRYPTOCURRENCY WALLETS

Cheman Shaik

VISH Consulting Services Inc, 6242 N Hoyne Avenue, Chicago IL 60659, USA

cheman_shaik@rediffmail.com

ABSTRACT

In this paper I have discussed a new method of enabling a cryptocurrency wallet user to define his own unforgettable seed phrase. An algorithm named SEEPT (Seed Phrase Transformation) is provided and illustrated with a real user defined seed phrase as input and a BIP39 standard seed phrase generated as output through cryptographic transformations. Discussed in detail is how an unforgettable seed phrase can be generatedfrom a set of names or words that are specific to a user's personal life.

Explained in detail as to how a wallet user can reconstruct his seed phrase on demand without depending on any storage, thereby relieving him from the burden of memorizing it frequently or storing it digitally on a computer hardware or physically on a paper or metal media.

KEYWORDS

Cryptocurrency Wallet, Unforgettable, Seed Phrase, User Defined, Algorithm, XOR operations, Hash Function, BIP39.

1. INTRODUCTION

Cryptocurrency is a digital currency generated through coin mining utilizing the blockchain technology and cryptography techniques[1]. These currencies are regarded as a highly promising asset class for investment and trading purposes. Today there are more than 5000 cryptocurrencies traded around in the financial market out of which Bitcoin and Ethereum are leading with their huge market caps[2].

Cryptocurrency users require a wallet to receive or send coins from or to others as payment for the goods or services bought or sold[3]. Basically, there are three types of crypto wallets available for use –hardware wallets called cold wallets, wallets hosted on the net called hot wallet, and warm wallet that can be installed on users' personal computers, laptops and mobile devices[4].

When a user starts using a cryptocurrency wallet for the first time it generates a long seed phrase of 12 or 24 words from a list of 2048 words specified by the BIP39 standard and instructs the user to carefully write down and store it in a secured place[5][6]. The seed phrase should be kept highly confidential as anyone who knows it can steal all the assets of the user. In case the user's computer, laptop or mobile device is broken, corrupt, stolen or destroyed, the user can recover all his crypto coins by reinstalling the same cryptocurrency wallet on a new system and supplying the seed phrase. If the seed phrase is lost or forgotten, the user losses all his assets with no chance of recovery, whatever be the amount lost[7].

The logic behind generating a seed phrase of 12 words in a crypto wallet involves generating 12 random integers of 11-bit length. Each of these 11-bit numbers falls in the range 0 – 2047 and is used as the index of the word to be picked from the 2048 words of the BIP39 standard list. All the 12 words are displayed on a screen of the crypto wallet and the user is instructed to write down and store them in the same order.

Cryptocurrency wallet users have practical concerns with the random seed phrase generated by the wallet. It is very hard to memorize all the 12 or 24 words. Paper media used to store the seed phrase can be destroyed or damaged in floods and fire accidents or can be lost in shiftingand theft. On the other hand, a seed phrase stored digitally on a computer hardware is vulnerable to hacks and can be stolen, if it is connected to the internet[7].

An unforgettable seed phrase that a user could generate himself, which at the same time translates to some 12 or 24 words from the BIP39 standard list, is highly desirable and useful in resolving the aforementioned practical and security concerns of a random seed phrase generated by the wallet.

2. LITERATURE SURVEY

The BIP39 list of words used to generatewallet seed phrase was published in 2013[8]. Since, then researchers worked on how to secure the seed phrase generated with BIP39 words from hackers and how to recover crypto assets using it in case the device the wallet installed on is lost, stolen, or damaged. However, no one has attempted to devise a method to enable cryptocurrency wallet users to define their own seed phrase which at the same time can be transformedinto a BIP39 seed phrase.

In 2009 Farah MaathJasem conducted some research on enhancing the security of the Bitcoin wallet master seed by introducing an additional source of entropy – non-fixed ASCII encoded characters. She studied on how to increase the resistance of Bitcoin wallet against dictionary attacks by increasing the master seed entropy[9].

In 2015 Vasek*et al*conducted a survey on brain wallets, that is private keys generated from passwords, and discovered that such brain wallets are vulnerable to compromise as passwords are easy to guess[10].Also, Eskandari*et al*, after a study of wallet software, confirmed that brain wallets generated from week passwords erode the security that they offer to crypto assets [11].

In 2019 Volety*et al* revealed in their research report that they could crack the master seeds, through dictionary attacks,of two publicly available bitcoin wallets[12].

No evidences are found in literature on how to derive a seed phrase based on a criterion model that is beyond scope of dictionary attacks by hackers, which can in turn be transformed into a BIP39 seed phrase.

3. GENERATING UNFORGETTABLE USER DEFINED SEED PHRASE

An unforgettable seed phrase can be generated from a list of names or words that are specific to a user's personal life, which the user never forgets, such as names from his family hierarchy. Alternatively, they could be names of his friends from his nursery class to university. Some more suggestions for picking name are rail or bus stations staring from his residence in a particular direction or street names starting from his next street in a direction of his choice. These are only

some typical examples of constructing a list of names for seed phrase generation and such criteria are unlimited and the user can be very innovative in contemplating these criteria.

Seed phrases generated from names or words that a user selects based on his own criteria are unforgettable as he has nothing to memorize and what all he needs to remember is only the criterion model that he used to select the names or words, which the user is very unlikely to forget. It is more advisable to select as many names as the number of seed words that the cryptocurrency wallet will generate for the user.

Fig. 1 below shows a family tree of a crypto wallet user Jose converging upwards with his grandfather at the root. Jose wants to use these names as his seed phrase and will enter them in the same sequence in the wallet screen that computes in background the 12-word seed phrase from words picked from the BIP39 list.

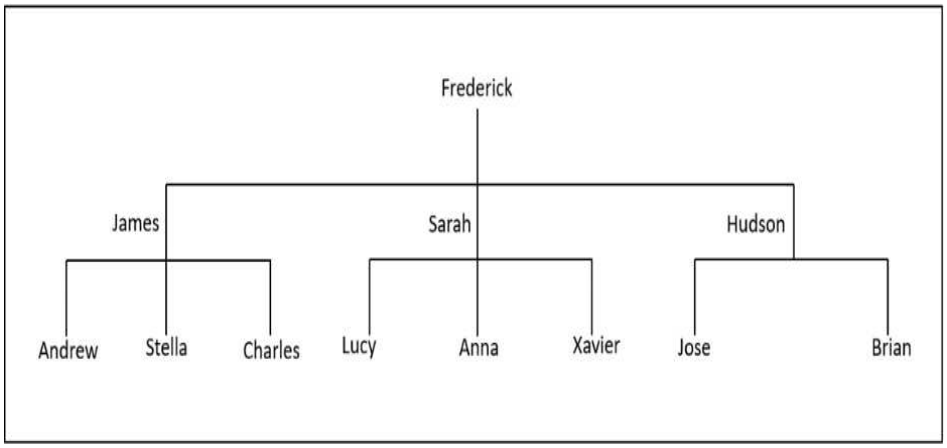

Fig.1 Family hierarchy of a cryptocurrency wallet user

Fig. 2 below shows the crypto wallet screen where the user will be asked to enter his own words for seed phrase generation.

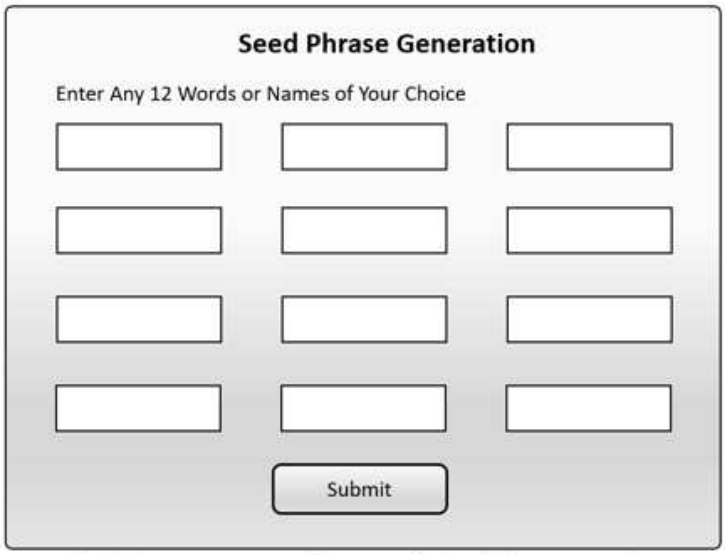

Fig.2 A cryptocurrency wallet screen for Seed Phrase generation

Fig. 3 below shows the seed words entered by the user where after he will click the submit button. Subsequently, the screen will execute logic to generate a seed phrase using some 12 words selected from the BIP39 standard list of 2048 words.

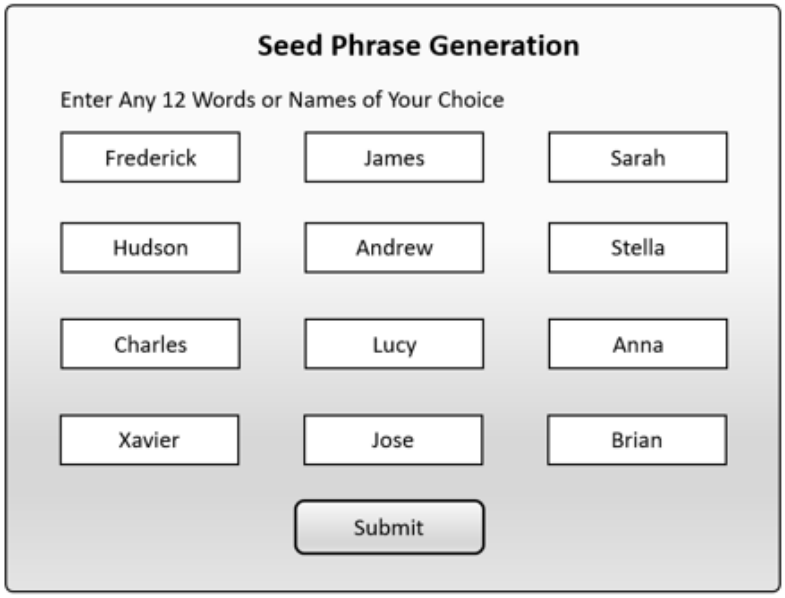

Fig. 3 A cryptocurrency wallet screen filled-in for Seed Phrase generation

A BIP39 compliant seed phrase is generated using the following algorithmic steps:

1. Select the 12 words entered by user
2. Concatenate all the words entered by the user
3. Convert the resulting string to binary
4. Divide the binary string into 8-bit blocks
5. Starting from the first block, perform XOR operation of each string with its succeeding string and replace the succeeding string with the XOR result. XOR the last block with the first block.If a resulting block is less than 8 bits, pad 0's at the left end for the missing bits
6. Convert the binary to string
7. Run SHA256 hash function on the entire string and get the output in binary form
8. Perform XOR operation on the first 132 bits of SHA256 output with the remaining 124 bits
9. Divide the result in to 12 blocks of 11 bits each
10. Convers each block to a decimal number, which will be between 0 – 2047
11. Use the decimal number as the index to pick a word from BIP39 list.
12. Append the 12 words to form BIP39 Seed Phrase

Fig. 4 below shows a flow chart for BIP39 compliant seed phrase generation from user selected seed phrase.

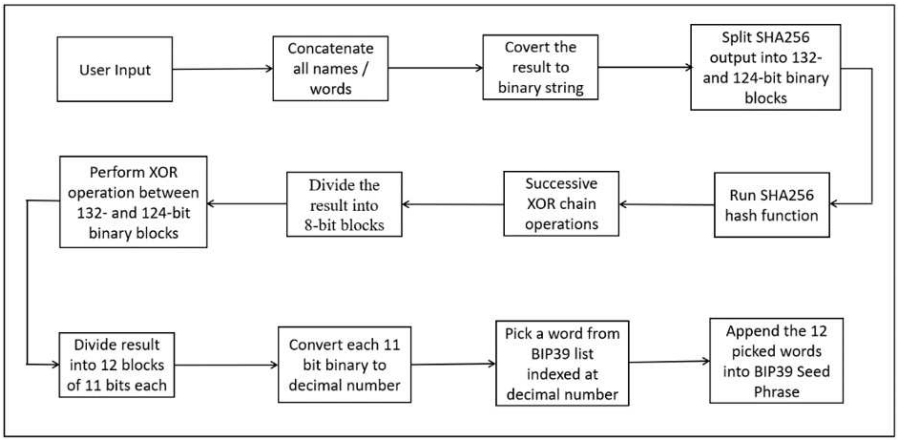

Fig.4 Floe Chart of SPA-11 Algorithm for Seed Phrase Generation

4. USE CASE ILLUSTRATION OF SEED PHRASE COMPUTATION

Starting from the user defined seed phrase words entered in the crypto wallet screen shown in Fig. 3, a step-by-step execution of the SEEPT algorithm is illustrated with the resulting output of each step till the a BIP39 compliant seed phrase is generated.

Step #	Action	Result
1	Concatenate all the words entered by the user	FrederickJamesSarahHudsonAndrewStellaCharlesLucyAnnaXavierJoseBrian
2	Convert the resulting string to binary	01000110011100100110010101100100011001010111001001101001011000110110101101001010011000010110110101100101011100110101001101100001011100100110000101101000010010000111010101100100011100110110111101101110010000010110111001100100011100100110010101110111010100110111010001100101011011000110110001100001010000110110100001100001011100100110110001100101011100110100110001110101011000110111100101000001011011100110111001100001010110000110000101110110011010010110010101110010010010100110111101110011011001010100001001110010011010010110000101101110
3	Divide the binary string into 8-bit blocks	01000110 01110010 01100101 01100100 01100101 01110010 01101001 01100011 01101011 01001010 01100001 01101101 01100101 01110011 01010011 01100001 01110010 01100001 01101000 01001000 01110101 01100100 01110011 01101111 01101110 01000001 01101110 01100100 01110010 01100101 01110111 01010011 01110100 01100101 01101100 01101100 01100001 01000011 01101000 01100001 01110010 01101100 01100101 01110011 01001100 01110101 01100011 01111001

		01000001 01101110 01101110 01100001 01011000 01100001 01110110 01101001 01100101 01110010 01001010 01101111 01110011 01100101 01000010 01110010 01101001 01100001 01101110
4	Starting from the first block, perform XOR operation of each block with its succeeding block and replace the succeeding block with the XOR result. XOR the last block with the first block	01001001 00110100 00010111 00000001 00000001 00010111 00011011 00001010 00001000 00100001 00101011 00001100 00001000 00010110 00100000 00110010 00010011 00010011 00001001 00100000 00111101 00010001 00010111 00011100 00000001 00101111 00101111 00001010 00010110 00010111 00010010 00100100 00100111 00010001 00001001 00000000 00001101 00100010 00101011 00001001 00010011 00011110 00001001 00010110 00111111 00111001 00010110 00011010 00111000 00101111 00000000 00001111 00111001 00111001 00010111 00011111 00001100 00010111 00111000 00100101 00011100 00010110 00100111 00110000 00011011 00001000 00001111
5	Convert the binary to string	I4 !+2 =// $' "+ - ?98/998%'0
6	Run SHA256 hash function on the entire string and get the output	a4c5e547daa8f43eb4a7fbcd69e1dae1c0b2 9ac068f45e15b821ce9b33acea98
7	Convert the Hex to binary	10100100110001011110010101000111 11 01101010101000111101000011111010 11 01001010011111111011110011010110 10 01111000011101101011100001110000 00 10110010100110101100000001101000 11 11010001011110000101011011100000 10 00011100111010011011001100111010 11 001110101010011000
8	Perform XOR operation on the first 132 bits of SHA256 output with the remaining 124 bits	10100100110011101100110011101011 11 01110000100111101100011101111111 10 11110010010111100111001001001101 10 10110110110001010001001000000 100
9	Divide the result in to 12 blocks of 11 bits each	10100100110 01110110011 00111010111 10111000010 01111011000 11101111111 10111100100 10111100111 00100100110 11010110110 11000101000 10010000100

10	Convert each block to a decimal number, which will be between 0 – 2047	1318 947 471 1474 984 1919 1508 1511 294 1718 1576 1156
11	Add 1 to the decimal number	1319 948 472 1475 985 1920 1509 1512 295 1719 1577 1157
12	Use the decimal number as the index to pick the word from BIP39 list.	pill island depend reunion kitchen useful route rude cave strategy shallow mountain

As the user enters the twelve names that he acquired from his family tree based on his criterion, capture the names in the wallet application program and concatenate them in the same order they are entered. Convert the resulting string to binary form and divide it into 8-bit binary blocks. Starting from the first block, perform an XOR operation of each block with its successor block. In each XOR operation, the resulting block of the previous XOR operation is used. XOR the last resulting bloc with the first block. This will make reversal of the computations by attackers very difficult.

Convert the resulting 8-bit binary blocks into a string and run SHA256 Hash Function over it. Covert the hex form output of SHA256 to binary form. Split the 256 bits of the SHA256 Hash output into two blocks – a 132-bit first block and a 124-bit second block. Perform XOR operation between the first and second blocks. If the resulting block has binary output less than 132 bits, pad it with 0's at the left end to make it exactly 132 bits long.

Split the resulting 132 bit binary into 12 blocks of 11 bits each. Convert each of the 12 binary numbers to decimal form. As each of the resulting decimal number falls in the range 0 – 2047 and BIP39 words are numbered from 1 – 2048, add 1 to each of them and use the resulting decimal number as the index of the word to be picked from the BIP39 list and form the BIP39 compliant seed phrase by concatenation.

The cryptocurrency wallet will store the BIP39 seed phrase and use the same for computing public and private keys for different crypto coins. In case the user's computer, laptop or mobile phone where he downloaded his crypto wallet is lost, broken, stolen, or destroyed, he can download the wallet on a new device and enter his own defined seed phrase names or words, which will in turn be translated to the BIP39 seed phrase using the same logic. The user is not required to memorize or retrieve any BIP39 seed phrase from his offline digital or paper storage media.

5. MODIFICATIONS FOR 24 WORD SEED PHRASE

Though the algorithm is illustrated for a 12-word seed phrase generation, it can be implemented for generating seed phrase of any length. Even for generating a 24-word user defined seed phrase, the same steps need to be followed. In case the user has to generate a 24-word seed phrase, he needs to widen his criterion model in order to incorporate more name or words, typically around 24 words, in his input to the algorithm. Each name in the user input may be associated with a word representing name holder'sother attributes.

SHA256 produces an output of 256 bits which will be shorter than 264 which is the total bit length of 24 blocks of 11 bits. In order to overcome this issue SHA512 may use to crate the hash and split it in to two blocks – one with 264 bits and another with 248 bits. Perform XOR operation between the two and divide the result in to 24 blocks of 11 bits and accordingly pick 24 words from the BIP39 list of words. For any seed phrase less than 24 words SHA256 will be sufficient. When a seed phrase of N words needs to be generated, the SHA256 hash should be split at 11N bits into two binary blocks and an XOR operation should be performed between the leading and trailing blocks. Rest all steps discussed above will remain intact.

6. TECHNICAL IMPLEMENTATION

Cryptocurrency wallets can implement the algorithm in their application logic and use it as an interface to the BIP39 seed generation instead of randomly picking the seed words from the standard list of 2048 words. The algorithm can be implemented as a separate loosely coupled procedure in any standard software language such as Java, Python, C++ and C# using any standards open source cryptography libraries.

7. BRUTE FORCE ATTACKS: USER DEFINED SEEDS VS BIP39 SEEDS

A BIP39 seed phrase offers 128-bit security against brute force attacks. As there are 2048 words in the BIP39 list, there are as many options for each seed word which offers 11-bit security. As the seed phrase is constituted by 12 such words, the entire seed phrase offers 11x12, that is 132-bit security. However, as some data in BIP39 seed phrase is not random, a 4-bit deduction applies to the security resulting into a128-bit security[13].

On the other hand, the number of names or words that a user can use in a seed phrase that he defines himself are unlimited. The user can be very innovative in framing the criterion model to select the names or words constituting the seed phrase. Unless a hacker conducts a thorough study of the user's personal life and his life history, it is not possible for him to guess his seed words.

Even after such a time-consuming study of a user, the attacker is not aware of the conceptual model the user devised to select his seed words. When the user's device where he downloaded the cryptocurrency wallet is damaged, lost or stolen, he can derive his seed phrase using the same model without any need to retrieve it from a storage medium or periodically memorizing it once in a while.

8. PRECAUTIONS ON SEED PHRASE SELECTION

Cryptocurrency wallet users should always adopt a criterion model based on which they should select their seed phrase words. They should follow the below guide lines while forming their seed phrase:

- Donot use phrases from songs, poems quotations etc.,
- user names that have at least three characters
- do not use surnames if family tree is used as the criterion model
- never reveal the criterion model used in seed phrase forming

9. CONCLUSION

A Cryptocurrency wallet generates a seed phrase when a user downloads and starts using it. Seed phrase is a list of 12 or 24 words selected from the BIP39 standard list of 2048 words. It is a very confidential piece of string that should be kept secret and will lead to the loss of all crypto assets if lost or stolen. It is very hard to memorize such a long list of words randomly picked from a BIP39 document. A seed phrase written on paper medium is vulnerable to loss or damage by theft, fire, floods, humidity and termites. On the other hand, a seed phrase stored on digital medium connected to the internet is vulnerable to hacks by attackers.

An algorithm called Seed Phrase Transformation (SEEPT) has been developed and discussed with detailed illustration of the steps involved in transforming a user defined seed phrase into a BIP39 seed phrase. Using a family hierarchical chart, a list of names is developed and used as input to the SEEPT algorithm. As per the instructions of the presented algorithm, the input string is processed to perform XOR operations, run SHA256 hash function, further processing the output of the hash function, derive 12 binary blocks of 11 bits, and finally use these blocks to identify the indices of words to be picked from the BIP39 list.

User defined seed phrase is easy to reconstruct without any need to memorize it as it is constructed based on a criterion model devised by the user himself. Unlike the limited number of words in the BIP39 list, the number of names or words that a user can think of are unlimited. Moreover, the user defined seed phrase generation procedure involves a criterion model which is very hard for a hacker to guess because the number of models that the user can contemplate are again unlimited. User defined seed phrase does not yield to brute force attacks unless the attacker conducts a thorough study of the user's personal life.

User defined seed phrase relives cryptocurrency wallet users from the burden of maintaining it safely on paper media or digital media. A user can store his criterion model used to generate the seed phrase only in his brain like a password and reconstruct it any time he needs it.

REFERENCES

1. Chinmay A. Vyas, Munindra Lunagaria, "Security Concerns and Issues for Bitcoin", International Journal of Computer Applications (0975 – 8887) National Conference cum Workshop on Bioinformatics and Computational Biology, NCWBCB- 2014
2. Sam Kopleman, "What are altcoins? Everything you need to know", https://www.techradar.com/news/what-are-altcoins-everything-you-need-to-know
3. Brian Mackay, "Evaluation of Security in Hardware and Software Cryptocurrency Wallets", Research Thesis from School of Computing Edinburgh Napier University Edinburgh, Scotland
4. Crypto Markets Wiki, "Hot vs. Warm vs. Cold Wallets", http://crypto.marketswiki.com/index.php?title=Wallets

5. Bitcoin.com "Bitcoin and Mnemonics: The Art of the Secret Phrase", https://news.bitcoin.com/bitcoin-and-mnemonics-the-art-of-the-secret-phrase/
6. BitcoinSV, "Seed phrase", https://wiki.bitcoinsv.io/index.php/Seed_phrase
7. Exodus, "Everything you need to know about your Secret Recovery phrase", https://support.exodus.io/article/925-everything-you-need-to-know-about-the-secret-recovery-phrase
8. Github.com, "BIP39 English Words List", https://github.com/bitcoin/bips/blob/master/bip-0039/english.txt
9. Farah MaathJasem, Ali MakkiSagheer, Abdullah M. Awad, "Enhancing the Security of the Bitcoin Wallet Master Seed", Conference Paper - University of Anbar, Ramadi, Iraq
10. Marie Vasek, Joseph Bonneau, Ryan Castellucci, Cameron Keith, Tyler Moore1, "The Bitcoin Brain Drain: Examining the Use and Abuse of Bitcoin Brain Wallets", Tandy School of Computer Science, The University of Tulsa
11. ShayanEskandari, David Barrera, Elizabeth Stobert, and Jeremy Clark. "A First Look at the Usability of Bitcoin Key Management". Proceedings of the NDSS Workshop on Usable Security (USEC), 2015
12. TejaswiVolety, Shalabh Saini, Thomas McGhin, Charles Zhechao Liu, Kim-KwangRaymondChoo, "Cracking Bitcoin wallets: I want what you have in the wallets", Future Generation Computer Systems, Volume 91, February 2019 Bitcoin. it, "Seed Phrase", https://en.bitcoin.it/wiki/Seed_phrase

AUTHOR

ChemanShaik is a Research & Development professional in Computer Science and Information Technology for the last twenty years. He has been an inventor in these areas of technology with eight U.S Patents for his inventions in Cryptography, Password Security, Codeless Dynamic Websites, Text Generation in Foreign Languages, Anti-phishing Techniques and 3D Mouse for Computers. He is the pioneer of the Absolute Public Key Cryptography in 1999. He is well known for his Password Self Encryption Method which has earned him three U.S Patents. He has published research papers in the international journals – IJCSEA, IJCIS and the proceedings of EC2ND 2006 and CSC 2008.