Improving Data Security System through Gaussian Windowing / Inverse Gaussian Windowing in PTS

[1]Mamta pant, [2] Arundhati Waliya

[1] *HR Institute of Technology, Uttar Pradesh, India*
[2]*Associate Professor, Dept. of Computer science Engineering, HR Institute of Technology, Uttar Pradesh, India*

***Abstract*— Data security is the process of protecting digital data from unwanted action of unauthorized access and data corruption throughout its life. Through such conditions as cyber-attacks or data breaches on any server or computer system slightly increase the rate of cybercrime. As per the McAfee Corp. report released on Dec. 7, 2020,the estimated annual loss in the global economy exceeded $1 Trillion which is 50 % more from 2018. And if we specifically talk about India, the extent of loss due to cybercrime has increased up to 63 crore in 2020 as compared to 2021 through ATM cards, credit cards, E-commerce, and internet banking.**
Data security should have the ability to protect the confidential and personal data of the customer. In the current scenario, there are so many technologies available that provide the best data security. In this research paper we are working on Gaussian and Inverse Gaussian Windowing methods in the Partially Transmitted Sequence (PTS) technique.

***Index Terms*— Crest Factor, Fast Fourier Transform, Partially Transmitted Sequence Partially Transmitted Sequences, Gaussian Windowing Method, Inverse Gaussian Windowing Method, Crest Factor.**

I. INTRODUCTION

In the way to protect your data from cyber attackers, firstly, you need to know your data, you need to know where the data is located, how much data is sensitive information and what techniques you used in the industries to secure your data. And most important thing you need to know is Laws and Regulations to protect that valued data.
Data Security Techniques
Data Anonymization
In this method, we removed Personally Identifiable Information (PII). We can secure our sensitive data either way by erasing or by encrypting an identifier
As an example: A bank stores all of its customer's data along with their transaction history, so by using that method they remove all the sensitive information from the transaction summary like account number, ATM card number and Debit/credit card number. They only store sender and receiver name, date, and amount. So that information cannot be traced by an attacker.
Data Anonymization Techniques

- Pseudonymization
- Generalization
- c)Data Swapping (Shuffling and permutation)
- d)Data Perturbation
- e) Synthetic Data

Data masking
This method is used to mask your information by replacing a value character with any symbol
Example: You got a payment receipt after payment to any retailer store using credit or debit card so in that method we hide all of your sensitive information like cardholder name or card number. In the same way, that method is used in cyber security to hide or mask your critical information.
Tokenization
Instead of sharing your information like credit card details to a third party while doing an online payment, this method creates one token. Only that token is used to make a payment, which means third party receiver does not know your actual details.
Data Loss Prevention (DLP) - This method contains 2 parts:

- Detection
- Prevention

This method works for both data loss and data leakage prevention .This type of protection is generally used when an employee leaves the organization. Digital Route Management This method is used to protect intellectual property such as prevent piracy or copyright items so that a third party cannot store your data.Data classification based on their sensitivity Sensitive data is private information that should protect from others. There are some types of data classification based on their sensitivity. Regulated Data: Sensitive data comes under this category e.g. credit card details, debit card details, bank account details, etc.Unregulated Data: Data in the form of an Unknowing files or publically known data comes under this category e.g. advertisement details of any recruitment, applications for a vacant post, etc.

II. CREST FACTOR

Crest factor is the ratio of the peak amplitude of a waveform to its mean square value. It is a parameter of the waveform which indicates how extreme the peak in the waveform. There are so many reduction factors that exist such as peak windowing, noise shaping, and peak cancellation. Application of crest factors is in electrical engineering for describing the quality of an AC power waveform and vibration analysis for predicting the amount of impact.

$$CrestFactor(CF) = \frac{\text{maximum } \{Z(t)^2\}}{mean\ \{Z(t)^2\}}$$

Z (t) denotes the transmitted signal Maximum represents the peak of the signal Mean represents the average value of the signal the importance of the CF is attributed to the fact that the CF signifies the amount of signal deviation from the mean power which makes it more susceptible to third party attacks.

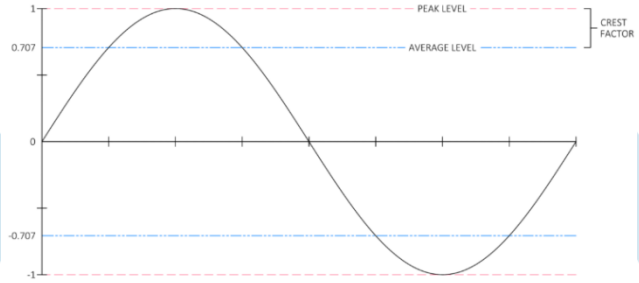

Fig -1: Crest factor of a sine wave

III. PROBLEM FORMULATION

The main objective of this proposed work is to lessen the Crest Factor of the system. This would infer minimum deviation of the signal from the average power. This ensures that the abrupt signal peaks would not be there in the system and hence it would not be recognizable. This would reduce the perceptibility factor and the data would be difficult to be identified by the intruders. As the data protection and confidentiality is a major area of concern, it is essential to use minimum crest factor. But selection of crest factor is to be done prudently such that it doesn't make the system complex because data integrity must be maintained at all measures.

Mathematical Modeling of Partially Transmitted Sequences

Consider the sample space of the shift vectors are given by:

$$S\_B = \begin{matrix} b1 \\ b2 \\ \vdots \\ \vdots \\ bn \end{matrix}$$

Simultaneous versions of the data are generated given by:

$$Y = \begin{matrix} Y1 \\ Y2 \\ \vdots \\ \vdots \\ Yn \end{matrix}$$

The exhaustive search tries to find out:

$Vector$ = $((CF1, F2, CF3, \dots \dots . CFn))$

The shift that results in the crest factor to fall to the minimum value is used for the final data transmission. The PSD of such a shifted version of the data is given mathematically as:

$$\sum_{n=0}^{K-1} \frac{\lambda_n(\hat{H}(f) - \hat{H}_n(f))}{(\lambda_n \hat{H}(f) - \hat{G}_n(f))^2} = 0$$

G represents the shift vector sample space, $\hat{H}(f)$ is the PSD And $\hat{H}n\ (f)$ is the statistical average of (f).

The variation in the maxima of the modulated envelope of H (f) is given by:

$\hat{H}i\ (f) = 1(1 -). d. 2, n = 0, 1, \dots \dots \dots \dots N - 1$

The PSD after vector shifts is computed as:

$$\hat{H}^{k+1}(f) = \left[\sum_{n=0}^{N-1} \frac{\mu_n \hat{H}_n^{(k)}(f)}{\left(\mu_n \hat{H}_n^{(k)}(f) + \mu_n \hat{H}_n(f)\right)^2}\right] + \left[\sum_{n=0}^{N-1} \frac{\mu_n \hat{H}_n^{(k)}(f)}{\left(\mu_n \hat{H}^{(k)}(f) + \hat{H}_n(f)\right)^2}\right]^{-1}$$

Here, f represents the frequency domain dependence, B stands for the shift vector's sample space in the magnitude-square form, we get

$$Z(f) = 2\sum_{n=0}^{N-1}|h_n(f)|^2$$

Z (f) denotes the f-domain total magnitude; hn denotes the samples in discrete frequency
The subsequent process is the rigorous search for the shift causing the signal to attain the minimal values of CF:

$$Sh(f) = \begin{bmatrix} s_0^{(0)}Y_0^{(0)} & s_1^{(0)}Y_1^{(0)} & \cdots & s_{N-1}^{(0)}Y_{N-1}^{(0)} \\ s_0^{(1)}Y & s_1^{(1)}Y_1^{(1)} & \cdots & s_{N-1}^{(1)}Y_{N-1}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ s_0^{(k-1)}Y_0^{(k-1)} & s_1^{(k-1)}Y_1^{(k-1)} & \cdots & s_{N-1}^{(k-1)}Y_{N-1}^{(k-1)} \end{bmatrix}$$

The possibility of finding a vector leading to decreased CF is given by
b€min (CF)
The Hermitan transpose is given by b, the kth value order is denoted by $s.\,d.\,(n)$
The change or variation in the instantaneous value from the mean is given by

$$V(f) = \sum_{k=0}^{k-1} s.d._k\, u_n b_n^{\dagger}$$

s.d. is the standard deviation , (f) is the probability distribution vector, $up\ (f)$ is a shift combination vector, Computing the Eigen values, we get

$$A(f)A^{\dagger}(f) = \sum_{p=0}^{p-1}\sigma_p^2(f)u_p(f)v_n^{\dagger}(f)$$

Here, $\sigma p\ 2\ (f)$ is the mth Eigen value of the Eigen decomposition.
The frequency dependence is therefore given by:

$$H(f)H^{\dagger}(f) = \sum_{k=0}^{k-1} s.d._k^2\,(f)u_n(f)u_k^{\dagger}f)$$

The Eigen values extend to… N-1 is all zero.
The CF exhibits a probabilistic swing of:

$$s.d._n^2\,(f) = \sum_{k=0}^{k-1}|u_{k=1}^n(f)|^2\left|Y_{k=1}^{k=n}(f)\right|^2$$

$s.\,d.n\ 2\ (f)$ depicts the variance or swing in terms of probability
Putting $|bk{=}1\ nk{=}n\ (f)|\ 2 = \mu(f)$ & operating $s.\,d.n\ 2\ (f)\ /\ \sum \mu p\ (n)\ (f)\ k{=}n{-}1\ k{=}0$, the following is obtained:

$$\widehat{H}^{(n)}(f) = \frac{\sum_{k=n=0}^{P-1}\mu_k^{(n)}(f)\left|Y_{k=0}^{k=n-1}(f)\right|^2}{\sum_{k=0}^{k=n-1}\mu_p^{(n)}(f)}$$

Here, $H(n)\ (f)$ represent the discrete f-based samples of H
n denotes the sample number
n=0,1,2……,k-1
Considering a total sample space of B vectors,,

$$\int_{-\infty}^{\infty} B(t,f)df = |y(t)|^2$$

$(t,\ f)$ represents the dependence of B on (t,f), Hence we get

$$|y(t)| = \int_{-1/2}^{1/2}\exp(j2\pi ut)\,dY_x(u)$$

(u) is the equi-probable df between +1/2 & -1/2

IV. RESULTS

The system is designed on the programming and simulation tool MATLAB (Matrix Laboratory) to facilitate the mathematical operations performed on data streams. The PTS is applied and hence after, the windowing functions are applied in the peak window. The results obtained with the different peak windows are shown in the subsequent section.
1. Gaussian Windowing –
The Gaussian Function, is mathematically defined as-

$$w(x) = e^{\frac{-1}{2}\left(\mu\frac{x}{\frac{N}{2}}\right)^2} \qquad 0 \le |x| \le \frac{N}{2}$$

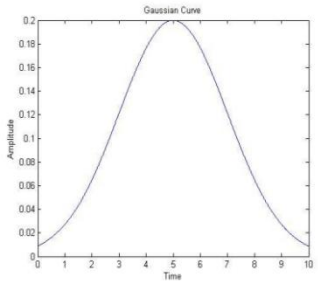

Fig -2: Gaussian Curve

It is important to note that the windowing function rises in such a manner that the maximum value is less than unity and does not have a constant increasing or decreasing gradient. Hence it reduces the peaks more than the adjacent values thereby reducing the CF.

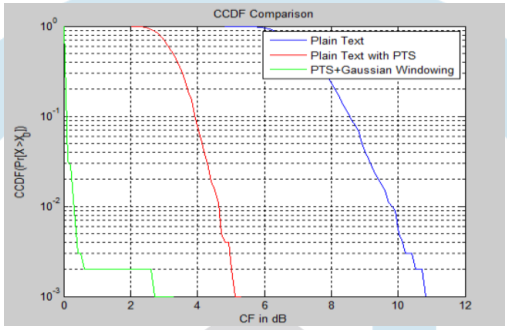

Fig -3: Proposed Techniques with Gaussian Peak Windowing

The CCDF above exhibits the variation of the CF with respect to the threshold in CF. The comparison has been made among 3 systems viz. Plain Text, Plain Text with PTS, and Plain Text with PTS + Gaussian Windowing. It can be clearly seen that the proposed technique of PTS + Gaussian Windowing outperforms the other two techniques thereby clearly indicating that the proposed system achieves better CF reduction.

V. CONCLUSION

In today's world everything trends toward digital. Most of the personal and company data are migrating to the cloud on the individual data center. This business model gives more flexibility and enhances productivity than ever before but it also gives challenges to the data security because it is also very easy for cyber attackers to steal your data and accomplish cybercrime. As data protection is a major area of concern. Nowadays on the top security needs, businesses and services must prefer data security techniques so the objective of this research paper is to minimize the crest factor of the system which is the encryption algorithms work in the form of OSI model layers, this is called bit-level security. This will reduce the perceptibility factor so that data is difficult to identify by the cyber attackers.

REFERENCES

[1] Zhiyi Wang ,Jun Cao ,Rui Deng ,Yi Liu,Jing " Time-frequency domain encryption with SLM scheme for physical-layer security in an USER DATA system", IEEE 2018
[2] Amber Sultan ,Xuelin Yang , Adnan A. E. Hajomer ,Weisheng Hu," Chaotic Constellation Mapping for Physical-Layer Data Encryption in USER DATA", IEEE 2018
[3] Adnan A. E. Hajomer , Xuelin Yang ,Weisheng Hu, "Chaotic Walsh–Hadamard Transform for Physical Layer Security in USER DATA", IEEE 2017
[4] HM Furqan, M Hamamreh , "Enhancing physical layer security of OFDM systems using channel shortening", IEEE 2017
[5] W Liu, M Li, G Ti, X Tian, Q Liu "Transmit filter and artificial noise aided physical layer security for OFDM systems",IEEE 2016
[6] G Shiqi, X Chengwen, F Zesong," Resource allocation for physical layer security in heterogeneous network with hidden eavesdropper", IEEE 2016
[7] Y Zou, J Zhu, X Wang, VCM Leung," Improving physical-layer security in wireless communications using diversity techniques",IEEE 2015
[8] E Jorswieck, S Tomasin, A Sezgin, "Broadcasting into the uncertainty: Authentication and confidentiality by physical-layer processing",IEEE 2015