

The Security Algorithm RBB01 used to Secure the Social Media Analyzed Data

Dr. C. Bagath Basha¹, K. Vishal Raj², M. Vivek Vardhan Reddy³, Syed Muzammil⁴, G. Venu Madhav⁵ and T. Vishwith Reddy⁶

^{1,2,4}Department of Artificial Intelligence & Machine Learning, Kommuri Pratap Reddy Institute of Technology, Ghanpur Village, Ghatkesar (M), Malkajigirl Dist., Hyderabad, Telangana, India.

Email: { chan.bagath@gmail.com, vishal0605raj@gmail.com, msyed0287@gmail.com }

^{3,5,6}Department of Computer Science and Engineering, Kommuri Pratap Reddy Institute of Technology, Ghanpur Village, Ghatkesar (M), Malkajigirl Dist., Hyderabad, Telangana, India.

Email: { maddivivekvardhanreddy02@gmail.com, mr.venumadhav124@gmail.com, Vishwithreddy2003@gmail.com }

Abstract—The significance of current global information is best understood in light of the reality that life on Earth is impossible without information. This is a lot of information, and it is growing every day thanks to social media sites like “Twitter, Facebook, YouTube, WhatsApp, Instagram, linkedIn”, and so on, which just provide a rapid general assessment online. Many more people groups used “this media,” notably Twitter. In this manner, “machine learning algorithms” are employed to analyse the public perception expressed in tweets and make predictions about the future. This information under examination has to have a strong extremity score. Since this score may be altered and has the potential to affect outcomes, it is quite vulnerable. Therefore, the current security calculations are “Salsa and ChaCha.” The “Salsa” calculation is slanted in a way that pushes initial attributes. Awkward traits are shifted to the first portion of the “ChaCha” computation. The existing computations don't have very good security because they focus mostly on execution rather than security. RBB01 is in this way the intelligent security calculation. Five phases make up this calculation. The first step is to locate the enigmatic “key N, n, and p values” via the lattice. The secret key is applied in condition in the second stage (1). The third stage of PN(n) upto n=1 activity. All PN(n) actions are made into a single line in the fourth step. The fifth stage involves trading and pairing attributes on a grid. In contrast to current calculations, the suggested calculation offers excellent security and execution.

Index Terms— Twitter, Machine Learning Algorithm, Security, Encryption, and RBB01.

I. INTRODUCTION

In the modern age, social networking sites like Twitter, Facebook, and others are increasing the amount of big data every day. Huge amounts of data are swiftly transported to several fields is what big data refers to. Three characteristics of large data are volume, velocity, and diversity. Data volume refers to data size. The data are being evaluated by velocity. Variety is data that is generated from several sources.

Twitter: Twitter is currently one of the more widely used forms of media in society. Anyone, wherever in the world, may tweet using the internet through this form of communication. Through the application of machine learning algorithms, this tweet was utilised to forecast the future and determine its polarity score.

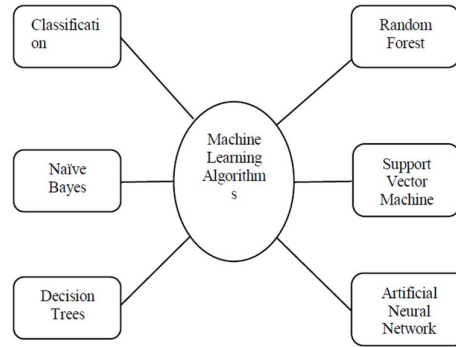


Figure 1. Machine Learning Algorithms [1]

The categorization algorithm's yes or no idea may be shown in Fig. 1. Mean Square Error is the basis of the Random Forest algorithm. Probability of class and condition is the foundation of the Naive Bayes method. Classification and regression are both aspects of the SVM technique, however only classification issues are employed more frequently. The idea of a decision tree is a depiction of a tree. Back propagation is a concept in artificial neural networks.

The fastest way to share peoples' opinions in the public today is via social media. People utilise platforms like "Twitter," "Facebook," "WhatsApp," "YouTube," "Instagram," "LinkedIn," etc. to disseminate information on sports, entertainment, politics, government employment, etc.

Twitter is one of these media that is specifically chosen for study since it is crucial for making predictions. Because this data might alter the future, it is less secure; thus, "apply the existing security algorithm is Salsa and ChaCha". These algorithms run in quarter-round manner and provide quick encryption performance. Because they primarily concentrated on efficiency rather than security, current algorithms do not have acceptable security. RajBagathBasha (RBB) 01 is the innovative security algorithm in this case.

II. RELATED WORK

The performance of more methods using sentiment analysis and machine learning was compared, and it was discovered that the classification technique is extremely simple to comprehend [1]. Salsa's execution time was compared to that of the proposed algorithm CBB21 [2]. They speak about the CBB20 security algorithm and the tweets. Salsa is used to compare the encryption time of the CBB20 proposed method [3]. They researched Twitter's security system and analysis. SRB21, a suggested security method, was used to compare Salsa's running time for encryption [4]. They mostly focus about security and future predictions made on Twitter. Salsa's encryption running time is compared to SRB18's security algorithm [5]. They investigated the RBJ25 cryptography algorithm and contrasted its encryption efficiency with that of "ChaCha and AES" [6]. Also in the recommended "security algorithm RB20", it compares Salsa and AES in terms of speed while calculating the running time of encryption [7]. The CBB22 method is suggested and offers extended data security [8]. The "SRB21 Phase 1 algorithm" is suggested for twitter data confidentiality. Salsa20/4 is used to compare the encryption and decryption running times of this algorithm [9]. Through classification and SVM algorithms, sentiment analysis is mostly used to analyse movie reviews [10]. They outlined 7 steps for ensuring data security [11].

III. ENCRYPTION ALGORITHM

- Obtain the data that have been generalised and convert it into matrix form.
- Determine the matrix's N, n, and p values.

$$P_N(n) = (N! / (n! * (N - n)!)) p^n (1 - p)^{(N - n)}$$

do $P_N(n)$ operations upto $n = 1$ Equation (1)

where N is the size of the matrix
n is upper triangle values, n is upto 1
p is lower triangle values

- Combine all values from PN (n) into a single line of activity.
- In Step 4, two values from each pair upon left.
- In Step 6, the matrix's values via Step 5 must be replaced.

IV. RESULT AND DISCUSSION

$$P_{BD} = \begin{bmatrix} 2101/7 & 2102/7 & 2103/7 & 2104/7 & 2105/7 \\ 2106/7 & 2107/7 & 2108/7 & 2109/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

where P_{BD} is Probability of Bernoulli Distribution

Step 1: n=4

“N= 5, p=7, N!=120, n!=24”

- Using “Equation (1)”

$$P_5(4) = (5! / (4! * (5-4)!)) (7^4 ((1-7)^{(5-4)}))$$

$$P_5(4) = (120 / (24 * (1)!)) (7^4 ((-6)^{(1)}))$$

$$P_5(4) = (120 / (24 * 1)) (2401(-6))$$

$$P_5(4) = (120 / 24) (-14406)$$

$$P_5(4) = (5) (-14406)$$

$$P_5(4) = 72030$$

Step 2: n=3

“N= 5, p=7, N!=120, n!=6”

$$P_5(3) = (5! / (3! * (5-3)!)) (7^3 ((1-7)^{(5-3)}))$$

$$P_5(3) = (120 / (6 * (2)!)) (7^3 ((-6)^{(2)}))$$

$$P_5(3) = (120 / (6 * 2)) (343(36))$$

$$P_5(3) = (120 / 12) (2058)$$

$$P_5(3) = (10) (2058)$$

$$P_5(3) = 20580$$

Step 3: n=2

“N= 5, p=7, N!=120, n!=2”

$$P_5(2) = (5! / (2! * (5-2)!)) (7^2 ((1-7)^{(5-2)}))$$

$$P_5(2) = (120 / (2 * (3)!)) (7^2 ((-6)^{(3)}))$$

$$P_5(2) = (120 / (2 * 6)) (7^2 (-216))$$

$$P_5(2) = (120 / 12) (49(-216))$$

$$P_5(2) = (10) (10584)$$

$$P_5(2) = 105840$$

Step 3: n=1

“N= 5, p=7, N!=120, n!=1”

$$P_5(1) = (5! / (1! * (5-1)!)) (7^1 ((1-7)^{(5-1)}))$$

$$P_5(1) = (120 / (1! * (4)!)) (7^1 ((-6)^{(4)}))$$

$$P_5(1) = (120 / (1! * 24)) (7^1 (1296))$$

$$P_5(1) = (120 / 24) (7(1296))$$

$$P_5(1) = (5) (9072)$$

$$P_5(1) = 45360$$

$$P_5 = 720302058010584045360$$

$$P_5 = (7,2), (0,3), (0,2), (0,5), (8,0), (1,0), (5,8), (4,0), (4,5), (3,6), 0$$

Pair 1: (7,2)

$$P_{BD} = \begin{bmatrix} 2101/7 & 2102/7 & 2108/7 & 2104/7 & 2105/7 \\ 2106/7 & 2107/7 & 2103/7 & 2109/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

Pair 2: (0,3)

$$P_{BD} = \begin{bmatrix} 2104/7 & 2102/7 & 2108/7 & 2101/7 & 2105/7 \\ 2106/7 & 2107/7 & 2103/7 & 2109/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

Pair 3: (0,2)

$$P_{BD} = \begin{bmatrix} 2108/7 & 2102/7 & 2104/7 & 2101/7 & 2105/7 \\ 2106/7 & 2107/7 & 2103/7 & 2109/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

Pair 4: (0,5)

$$P_{BD} = \begin{bmatrix} 2106/7 & 2102/7 & 2104/7 & 2101/7 & 2105/7 \\ 2108/7 & 2107/7 & 2103/7 & 2109/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

Pair 5: (8,0)

$$P_{BD} = \begin{bmatrix} 2109/7 & 2102/7 & 2104/7 & 2101/7 & 2105/7 \\ 2108/7 & 2107/7 & 2103/7 & 2106/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

Pair 6: (1,0)

$$P_{BD} = \begin{bmatrix} 2102/7 & 2109/7 & 2104/7 & 2101/7 & 2105/7 \\ 2108/7 & 2107/7 & 2103/7 & 2106/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

Pair 7: (5,8)

$$P_{BD} = \begin{bmatrix} 2102/7 & 2109/7 & 2104/7 & 2101/7 & 2105/7 \\ 2106/7 & 2107/7 & 2103/7 & 2108/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

Pair 8:(4,0)

$$P_{BD} = \begin{bmatrix} 2105/7 & 2109/7 & 2104/7 & 2101/7 & 2102/7 \\ 2106/7 & 2107/7 & 2103/7 & 2108/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

Pair 9: (4,5)

$$P_{BD} = \begin{bmatrix} 2105/7 & 2109/7 & 2104/7 & 2101/7 & 2106/7 \\ 2102/7 & 2107/7 & 2103/7 & 2108/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

Pair 10: (3,6)

$$P_{BD} = \begin{bmatrix} 2105/7 & 2109/7 & 2104/7 & 2107/7 & 2106/7 \\ 2102/7 & 2101/7 & 2103/7 & 2108/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

Pair 11: (0,0)

$$P_{BD} = \begin{bmatrix} 2105/7 & 2109/7 & 2104/7 & 2107/7 & 2106/7 \\ 2102/7 & 2101/7 & 2103/7 & 2108/7 & 2110/7 \\ 2111/7 & 2112/7 & 2113/7 & 2114/7 & 2115/7 \\ 2116/7 & 2117/7 & 2118/7 & 2119/7 & 2120/7 \\ 2121/7 & 2122/7 & 2123/7 & 2124/7 & 2125/7 \end{bmatrix}$$

The 3x3=>24 bytes , 6x6=>76 bytes, 10x10=> 312 bytes, 15x15=> 812 bytes, 20x20=> 1531 bytes, 40x40=> 6580 bytes as displayed in the Table I.

Table I. RBB01 encryption performance

File Size	ChaCha	Salsa	RBB01
24	1.69	3.4	4
76	1.29	2.2	4.2
312	2.73	2.3	4.5
822	2.64	3.9	4.9
1531	3.4	9.1	9.6
6580	2.27	10.5	12.5

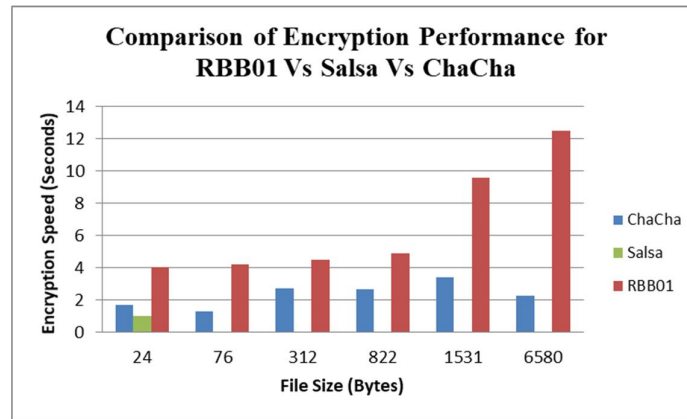


Figure 2. Encryption performance

In Fig.2, the RBB01 methodology and quickly examined the cypher speed. The authentication execution of “3.4 (s), 2.2 (s), 2.3 (s), 3.9 (s), 9.1 (s) and 10.5 (s) for the Salsa” and “4 (s), 4.2 (s), 4.5 (s), 4.9 (s), 9.6 (s) and 12.5 (s)” for the RBB01. When compared to current approaches, the RBB01 offers greater insurance of the information.

IV. CONCLUSION

In this day and age information is most significant in open on the grounds that, without information can't live on the planet. This information is enormous information and that information day to day increment through Social handles like "Twitter, Facebook, Youtube, WhatsApp, Instagram, linkedIn", and so forth... From this media, more number of people groups utilized particularly Twitter. As a necessary consequence, media is used to examine the overall evaluation of "tweets" and envision the "future through Machine Learning Algorithm" and used to make it extremity endlessly score seems to have exceptionally less security because this score can be changed and have an impact on fates, so implement the current security calculations is "Salsa and ChaCha". Since they focused only on execution rather than security, their estimates don't provide very effective security. Furthermore, RBB01 is the suggested security calculation. There are 5 steps in this calculation. Finding the mysterious "key N, n, and p values from" framework is the first step. Employing the mysterious key to the situation is the second step (1). The third stage of PN(n) upto n=1 activity. All PN(n) operations are made into one line in the fourth stage. The fifth stage involves trading and pairing attributes on a grid. In contrast to other computations, the RBB01 calculation has provided excellent security and execution. To boost the number of numerical security models in the future.

REFERENCES

- [1] C. Bagath Basha, and K. Somasundaram K, "A Comparative Study of Twitter Sentiment Analysis Using Machine Learning Algorithms in Big Data", International Journal of Recent Technology and Engineering, pp.591-599, 2019.
- [2] C. Bagath Basha, and S. Rajaprakash, "Applying The CBB21 Phase 2 Method For Securing Twitter Analyzed Data", Advances in Mathematics: Scientific Journal, pp. 1085-1091, 2020.
- [3] C. Bagath Basha, S. Rajaprakash, S. Muthuselvan, P. Saisatishsunder, and SVL Alekhya Rani, "Applying the CBB20 Algorithm for Twitter Analyzed Data", In: Proc. of First International Conference on Advances in Physical Sciences and Materials, Coimbatore, Tamil Nadu, India, 2020.
- [4] C. Bagath Basha, and S. Rajaprakash, "Applying the SRB21 Phase II Methodology for Securing Twitter Analyzed Data", In: Proc. of the International Conference on Mechanical Electronics and Computer Engineering, 2020.
- [5] C. Bagath Basha, and S. Rajaprakash, "Enhancing The Security Using SRB18 Method of Embedding Computing", Microprocessor and Microsystems, 2020.
- [6] S. Rajaprakash, C. Bagath Basha, S. Muthuselvan, N. Jaisankar, and Ravi Pratap Singh, "RBJ25 Cryptography Algorithm For Securing Big Data", In: Proc. of First International Conference on Advances in Physical Sciences and Materials, Coimbatore, Tamil Nadu, India, 2020.
- [7] K. Karthik, C. Bagath Basha, U. Bhaswanth Thilak, T. Sai Kiran, and J. Raj "Securing Social Media Analyzed Data Using RB20 Method", Advances in Mathematics: Scientific Journal, Vol. 3, 2020.
- [8] C. Bagath Basha, S. Rajaprakash, V.V.A. Harish, M.S. Krishna, and K. Prabhas, "Securing Twitter Analysed Data Using CBB22 Algorithm", Advances in Mathematics: Scientific Journal, pp. 1093-1100, 2020.

- [9] C. Bagath Basha, and S. Rajaprakash, "Securing Twitter Data Using SRB21 Phase I Methodology", International Journal of Scientific & Technology Research, pp.1952-1955, 2019.
- [10] R. Jaichandran, C. Bagath Basha, K. L. Shunmuganathan, S. Rajaprakash, and S. Kanagasuba Raja, "Sentiment Analysis of Movies on Social Media using R Studio", International Journal of Engineering and Advanced Technology, Vol. 8, 2019.
- [11] C. Bagath Basha, S. Rajaprakash, R. Meenakumari, M. Suresh, P. Hitesh, T. Kokilavani, and R. Ashok Kumar, " A novel security algorithm RPBB31 for securing the social media analyzed data using machine learning algorithms", Wireless Personal Communications, <https://doi.org/10.21203/rs.3.rs-1860348/v1>.