Zimbabwe Journal of Science & Technology pp 55-59 Vol.16 [2021] e-ISSN 2409-0360

Zimbabwej.sci.technol

A Microsoft Excel based cryptography scheme using matrix operations

Samuel Mukute

Great Zimbabwe University, Munhumutapa School of Commerce, Department of Accounting and Information Systems

E mail: smukute@gzu.ac.zw

ABSTRACT

The Microsoft Office suite which usually is encompassed on the Windows operating system comprises of the Microsoft Excel package which can be of great value if properly utilised. Various daily operations can be modelled as a set of linear equations which can easily be presented in form of a matrix. The limitation on the spectrum of users capable of solving these problems to achieve intended results is mainly based on the complex theoretical mathematical operations and computer programming efficiency usually required. This paper intends to demonstrate the ability to solve modelling problems by the use of Microsoft Excel in the field of Cryptography. The encoding process in the conversion of plaintext to cypher text and the decoding process which is the reverse process are illustrated by utilising in-built formulas associated with the Microsoft Excel package. The thrust of this paper is on the utilisation of the uniqueness of the inverse property of invertible matrices.

Keywords: matrices, modelling, cypher text, plain- text, key, inverse.

Received: 28.02.21 **Accepted**: 01.04.21

1. INTRODUCTION

Model building and simulation usually incorporates mathematical formulations and computer applications for us to produce better results (Velten, 2009). Usually the easiest way of modelling real life situations is to present them as a set of linear equations which can now be transformed into matrix representation (Rashi, 2019)In this era of technological advancement, most if not all modelling and simulation problems are tackled by the use of various software. Generally there are four categories of software platforms to aid in model building and simulation.

a. General purpose computer languages (for example C, C++, FORTRAN, BASIC, Pascal and Java);

b. Specialized Types of modelling software (plat- form) and simulation applications (for example FST, PowerSim, Stella and ExtendSim);

c. Equation solver-based applications (for example Maple, Mathematica, and Mathlab)

d. Spreadsheet based applications (for example Microsoft Excel and OpenOffice Calc1).

Of these four groups of software platforms, spread- sheet applications require the least level of proficiency in computer programming (Teh, 2015).

2. Literature Review

Spreadsheets are immensely popular and widely used because they are easy to

use, versatile and readily available to an extend that as noted in (Baker & Sugden, 2003),the basic idea of the electronic spreadsheet has stood the test of time; it has now become an indispensable item of software, not only in home and in the business , but in academia as well.(Barjis, 2010) summarised their features as:

a. a large number of numerical and non-numerical functions (i.e., dedicated formulas) to do mathematical, statistical, database, date and time, financial, engineering, and other types of calculations;

b. database representations and access;

c. charting and graphing;

d. display and document formatting capabilities such as layout, fonts, and colors to improve presentation;

e. scripting or programming language such as VBA (Visual Basic for Applications) in Excel.

Matrix application can be applied in many everyday operations such as in electrical circuits, in computer graphics conversion of three dimensional images into a two dimensional screen, creating the realistic seeming motions, encryption, page ranking algorithms for search engines and so forth (Anon., 2015). Generally the use of matrices in encryption was developed by Lester Hill , a mathematics professor, as early as 1929 (Barr, 2002). Hill coined what became popularly known as the Hill algorithm by mainly employing a technique based on cipher substitution. During the implementation of the algorithm, both the receiver and the sender have to agree on a key square matrix A which should be invertible mod26 (Tingting, et al., 2018). As an illustration an example shall be used:

Assume A=$\begin{bmatrix} 3 & 25 \\ 24 & 17 \end{bmatrix}$ is key to be used for text encryption. Let us now illustrate encryption of the phrase MISSISSIPPI. Since the key is of size 2x2, this encompasses that the block ciphers should also be size 2x1 , implying that the block will be of size 2 letters. Thus MISSISSIPPI becomes MI SS IS SI PP I. PP I. The Hill algorithm requires the phrase to encrypted to be of size divisible by n , where n , is the size of our square matrix. Hence in this particular example the lonely letter I will be padded by Z. Since Hill encoded the alphabet with A=0,B=1,C=0,...,Z=25 ,our phrase will come in blocks like for illustration sake MI= $\begin{bmatrix} 12 \\ 8 \end{bmatrix}$. Now A$\begin{bmatrix} 12 \\ 8 \end{bmatrix}$=$\begin{bmatrix} 2 \\ 8 \end{bmatrix}$mod26. Thus our blocks repeatedly becomes CI KK UW ER OY. It has to be appreciated that the repeating blocks where masked using the Hill cipher (Sekhon & Bloom, 2019).On deciphering the cipher text, a calculation of the inverse of

$$A= A^{-1}=\det(A)^{-1}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\text{mod26.} \quad (1)$$

Generally the above example summarises the birth of matrice use in cryptography. (Boon, 2006) as a motivation for his students to understand the Hill algorithm then used used the Harry Potter series in encrypting messages using Microsoft Excel. However the following loopholes where noted (Shivdeep, 2017):

a. the issue of the key being invertible by mod26 (this can only be noted if the Greatest Common Divisor(GCD(x,y))is 1)

b. masking of repeated letters using the Hill cipher

c. padding of messages with a selected letter if message is not divisible by n, the size of the square matrix.

As has been alluded to before, these complexities eliminates the target group in this paper. The IRI CellShield Microsoft Excel add-in has also been employed for protection of 1 or more Excel spreadsheets existing in 2 editions namely CellShield Personal Edition (PE) and CellShield Enterprise Edition (EE) (Innovative & Routines., 2021). However it has to be noted that both versions are yet to offer format decryption/encryption but for data masking it uses triple DES(). Currently the CellShield (EE) is compatible only with Excel 2010, 2013, 2016, and 2019 (Innovative & Routines., 2021). As is the case in the previous example, in as much as there is a StartUP menu, it is still on user friendly to the target group of this paper. In networking (specifically routing problems), for instance suppose that we have n cities and a map which shows the roads connecting these cities along with their weights (represented in miles). Then we can represent this information with an *nxn* matrix (Fošner, 2011). Matrix operations can be carried out to determine shortest routes, routes passing through at most two cities, and so forth. Matrix Cramer's Rule and determinants are simple and important tools for solving many problems in business and economics related to maximize profit and minimize loss (Mgr, 2009). Matrices are used to find variance and co-variance. Matrix Cramer's Rule is used to find solutions of linear equations with the help of matrix determinant. The equilibrium of markets in IS-LM model is solved by using determinants and Matrix Cramer's Rule (Shivdeep, 2017)Also matrices can be employed in cryptography which is the core issue of this paper.

2.1 Brief description of encryption

Generally cryptography can be defined as the science of protecting information by transforming it into a secure format. Basically cryptography can be viewed to have basically 5 main functions:

a. Privacy/confidentiality: Only the intended receiver should be able to read the send message

b. Authentication: The identity of the receiver should not be compromised.

c. Integrity: The received message has not been altered along the transmitting channel.

d. Non-repudiation: proving that the sender really sent this message.

e. Key exchange: The method by which cryptokeys are shared between sender and receiver. In cryptography, we start with the unencrypted data, referred to as plaintext. Plaintext is encrypted into cipher text, which will in turn (usually) be decrypted back into usable plaintext. The encryption and decryption is based upon the type of cryptography scheme being employed and some form of key (Gary, 1998). Symbolically it can be denoted as:

$C = E_k (P)$ (2) and

$P = D_k(C)$ (3)

where P = plaintext, C = cipher text, E = the encryption method, D = the decryption method, and k = the key. Usually the handling of the keys can be put across as: Forward Secrecy (Perfect forward secrecy): This feature protects past encrypted sessions from compromise even if the server holding the messages is compromised. This is accomplished by creating a different key for every session so that compromise of a single key does

not threaten the entirely of the communications. Perfect Security: A

sys- tem that is unbreakable and where the cipher text conveys no information about the plaintext or the key. To achieve perfect security, the key has to be at least as long as the plaintext, making analysis and even brute-force attacks impossible. One-time pads are an example of such a system. Deniable Authentication (Message Repudiation): A method whereby participants in an exchange of messages can be assured in the authenticity of the messages but in such a way that senders can later plausibly deny their participation to a third-party (Gary, 1998).

3. METHODS AND RESULTS

Perfect Forward Secrecy key communication is used in this case. The encryption approach used in this paper is mainly based on the universe characteristic of matrices. According to (Wang, 2017) and (He, 2019), "If any given matrix is invertible then the universe of that matrix is unique". To enhance the strength of this approach the universe function was used. A noble idea for representing the idea elaborated before is to use the following example:

Due to some reasons beyond Prof X's control, he has been forced to furnish the department with his Computer Graphics module marks via e-mail. A simple way to ensure security over the unsecure transmission channel has been agreed to use encryption. The recipient is Prof Y with whom earlier on had been agreed that the decryption key is $\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$. The marks are ranging from 0-100%, are going to be represented as an *n*x2 matrix where *n* is the number of candidates. Also each of the scores is prior encoded as follows:

O_{Value}	0	63	47	80	.	.	.	.	100
E_{Value}	100	37	53	20	.	.	.	.	0

Where ***OValue*** is attained mark and ***EValue*** is initial encoded value. Clearly it can be seen that ***EValue*** = **100 – *OValue***.

Consider a set of examination marks for 10 students in the computer graphics module as follows:

63	47	80	37	53	20	91	16	29	89

The policy that should be made is that the marks are presented in ascending order based on the unique student registration number. If there is an absentee then it should be captured as a 0. The matrix for ***OValue*** will be denoted as **A**, for ***EValue*** as **B** and the key as **X** when applying them in Excel.

A was captured in cells A2:B9. N:B: All marks should be captured in row order i.e for this case A2,B2,…….A8,B8,A9,B9.

A	
63	47
80	37
53	20
91	16
29	89

B was captured in cells A9:B13

B	
37	53
20	63
47	80
9	84
71	11

X was captured in cells A16:B17

X	
2	1
3	4

ProfX will compute his **EValue** as illustrated in ***Table B*** on the Excel sheet above.

He then computes the cypher text as illustrated in Table ***ProfX*** **BX** on the Excel sheet below. The formula which was used after highlighting the destination cells D2:E6 is: "=MMULT (A9: B13, A16:B17)" which is the product of the inverse of X

(X') and the marks in Table B using the formulae: "=MMULT(A9:B13,A16:B17)"

Prof X BX	
233	249
229	272
334	367
270	345
175	115

From this argument it means the cypher text message that will be transmitted via the email is the table ***ProfX*** **BX** only.

On receiving ProfY will compute the decrypting key which is the inverse of **X** denoted as **X'**

X'	
0.8	-0.2
-0.6	0.4

in the Excel sheet above using the formulae "=MINVERSE (a16:b17)" after highlighting (A20:B21). To come up with the EValue, ProfY will compute ***ProfX BX × X',*** using the formulae "=MMULT (D2:E6, A20:B21)".

ProfY BX	**InverseX**
37	53
20	63
47	80
9	84
71	11

Finally he will do the basic matrix subtraction as illustrated on Table **A'**.

A'	
63	47
80	37
53	20
91	16
29	89

Please Note: All operations were done on the same spreadsheet for illustration purpose only but in reality the sender and the user will be on different machines.

Conclusion.

This paper has illustrated how encryption can be used by the use of Microsoft Excel. The prerequisite of Computer Programming efficiency and or Mathematical appreciation has been eliminated. Thus all users across various disciplines can confidentially communicate without the need of having scientific background. The researcher intends to further the research by incorporating Linear and Non-linear programming Problem Solving using Microsoft Excel. This is intended to eliminate the need as well of relying on Computer programming efficiency across various fields especially in the area of Optimization.

REFERENCES

Anon., 2015. Real-life-application-of-matrices-some-examples. [Online] Available at: http://mathsguideonline.weebly.com [Accessed 26 April 2021].

Baker, J. & Sugden, S., 2003. Spreadsheets in education: the first 25 years. 1st ed. s.l.:Spreadsheets in Education (eJSiE).

Barjis, J., 2010. Enterprise and Organizational Modelling and Simulation. Tunis, EOMAS2010.

Barr, T. H., 2002. Invitation to cryptology. New Jersey: Prentice Hall.

Boon, L. C., 2006. Harry Potter and the cryptography with matrices. Singapore: National Institute of Edu- cational Technology University.

Christopher, T. B. S., 2015. Building Mathematical Models in Excel: A Guide for Agriculturists. 1st ed. Boca Raton, Florida: Universal-Publishers .

Fošner, D. A., 2011. Matrices and Routing. Primorska: University of Primorska, Slovenia.

Gary, K. C., 1998. Handbook on Local Area Networks. 1st ed. Auerbach: Auerbach.

Gary, K. D., 2020. An Overview of Cryptography. 2nd ed. Auerbach: Auerbach Publishers.

Hanche, S., 2010. Designing a Security Awareness Program. Information Systems Security, 9(6),1-9.

He, J., 2019. Linear Algebra : The Inverse of a Matrix. s.l.:University of Houston.

Innovative, I. & Routines., 2021. What's New in CellShield Version 2.0.. [Online] Available at: https://www.iri.com/blog/ data-protection/cellshield-version-2/ () [Accessed 26 April 2021].

Kaur, S., 2017. Applications of Matrices. International Journal of Engineering Technology Science and Research, 4(11), 2394-3386.

Kessler, G. C., 1998. Handbook on local networks. Auerbach.: J. P. Slone, Ed..

Mgr, E. V., 2009. Determinants and their use in economics. ed. s.l.:Department of Mathematics, University of Economics, Czechia..

Rashi, M., 2019. Application of linear system. [Online] Available at: https://www.cuemath.com [Accessed 2021].

Sekhon, R. & Bloom, R., 2019. Application of Matrices in Cryptography, De Anza College.. [Online] Available at: https://math.libretexts.org/@go/page/37850 [Accessed April 2021].

Shivdeep, K., 2017. Applications of matrices. International Journal of Engineering Technology Science and Research, 11.

Teh, C., 2015 . Building mathematical models in excel: A guide for agriculturists., s.l.: s.n.

Tingting, Y. et al., 2018. The improved hill encryp- tion algorithm towards the unmanned surface vessel video monitoring system based on inter- net of things technology.. [Online] Available at: https://doi.org/10.1155/2018/5183451 [Accessed April 2021].

Velten, K., 2009. Mathematical modelling and simulations.. ed. Wein- heim: WILEY-VCH Verlag GmbH Co.

Wang, Z., 2017. A class of drazin inverses in rings., s.l.: s.n.