Dutse Journal of Pure and Applied Sciences (DUJOPAS) Vol. 4 No. 1 June 2018

APPLICATION OF TRANSFORMATION-PERMUTATION TO LOSSLESS DATA ENCRYPTION AND COMPRESSION

Zirra P. B. Department of Mathematics and Computer Science,
Faculty of Science,
Federal University Kashere, Gombe State,
Nigeria.

Babayo A.M. Department of Mathematics and Computer Science,
Faculty of Science,
Federal University Kashere, Gombe State,
Nigeria.

Abstract

The more axioms are imposed on an algebraic structure, the more the structure shrinks. Group theory is the generalization of both ring and field. Every group have a transformation representation. That is, both Cayley Theorem of group and the Preston-Wagner Theorem for semigroup theory hold. This research work presents how to compress and encrypt data using the transformation-permutation arc. Set of permutations is a subset of the set of transformations. Data encryption is a flat-line for securing data and compressed data are easier transmitted. All mathematical singled-valued encryption and compression functions in theories and theorems are subsets of transformations and compression equations are mostly functions.

Keywords: Data Compression, Mutiple-Valued Functions, Cayley Theorem, Embedding, Coding Theory.

Introduction

The more axioms are endowed by an algebraic structure, the more the structure shrinks: There are more number of mathematicians compared to modern abstract algebraists and there are more number of the aforesaid algebraists compared with the number of semigroup-group theorists. Transformation-permutation encrypted and compressed data are light secured data transmission from an algebraists to the theorists.

The analogue of the Cayley Theorem of Group which states that every group is embeddable in the Permutation Group (Heinstein, 1964)is the Preston-Wagner Theorem of Semigroup which states that every semigroup is embeddable in the transformation semigroup Higgins (1992). By this, the researchers understand that every compression equation and encryption function are under the custody of transformations.

The number theoretical motivational background behind the study of semigroup theory is the set of natural numbers. It is important recommending that every alphabet has a natural number representative and not integer ring representative or continuum (Evis and Neuson, 1997)numbers ($\mathbb{R}$and $\mathbb{C}$). Since every semigroup is embeddable in the transformations, every word can be represented with a transformation notation. The word partial is too involved now because of the so-called identity of the free semigroup of words with juxtapose of binary operation. This research work presents that a partial transformation is capable of embedding, encrypting and compressing every sentence—including the longest-and every word—including the longest in the dictionary, interdenominationalism.

A subset of cartesian product is a relation. The relation is binary if it is a subset of duplets. The subsets of binary relations, $A \times B$ say, are transformations if all the elements of A are involve in the relationship and no two elements in B have one pre-image. A function is a transformation that is single-valued not double-valued, multiple-valued and down the road. Data Encryption has origin from abstract algebra Gallian(2013).

In the mid-1970s, Ron Rivest, Ad Shamir, and Len Adleman devised ingenious method that permits each person who is to receive a sent message to publicly explicate how to scramble messages sent to him. And even though the method used to scramble (encoding) the message is own publicly, only the person for whom it is intended will be able to unscramble (decoding) the message. One of the most interesting and important applications of finite fields has been the development of algebraic coding theory. This theory originated in the late 1940s.It was created in response to practical communication problems. Algebraic codes now used in compact disc players, fax machines, modems, and bar code scanners and are essential to computer maintenance, Gallian(2013).

Background of the Study

Finite field has application to coding theory and Dihedral Group has application to cryptography Gallian(2013). Both Field and Dihedral Group have permutation representation and permutations are subgroups of the transformation semigroups. Semigroup has applications to coding theory and encryption, see Howie(2003). It is not quite difficult to see that compound words are forming semigroup: A word space a word is equal to a compound word and a word comma a compound word is the same as a compound word comma a word. The comma is standing for bracket and shifting of bracket is the

associativity. The associativity gleaned from the fact that shifting bracket is only possible because of the often overlooked gab filling.

Data compression is defined as the reduction of the volume of data file without loss of information, Franser(2010). Ziviani(2000) and (Kattan, 2006) pointed out that data compression aims to condense the data in order to reduce the size of a data file.

The following is the General Encoding-Decoding Channel:

Figure 1: General Encoding-Decoding Channel

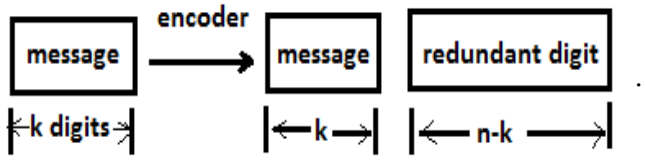

Figure 2: Encoding Process

A check digit scheme (Gallian, 2013) of dihedral group of order 10 can be used to detect all single- digit errors, which is due to Verhoeff(1969) and Winter(1990).

There are "lossless" and "lossy" forms of data compression, Franser(2010). Lossless data compression is used when the data has to be uncompressed exactly as it was before compression. Lossy compression works on the assumption that the data does not have to be stored perfectly. See Lossless Method in Guy(2011).

Methodology

The methodologies of this research work are the blended methodologies of computer science data encryption and compression, and methodologies of universal algebra. Universal algebra (Stanley and Sankappanavar, 1981) is the umbrella of all algebraic structures (transformation semigroup and permutation group inclusive). The methodologies of universal algebra are: Philosophy and Axiomatic Method.

Many a times, we say that mathematics is the father of science possibly because most scientific propositions are expressed using the mathematical model theory, the mathematical

models viz: Boyle Law, Charles Law, Gay-Lussac Law, etc. in Chemistry, Pecto-dectyl Limp System, Genetic Ratios, etc. in Biology, formulas in Physics, etc.

The question that may intrude is who and what is the father of mathematics, ie the grandfather of science? The answer is philosophy. Philosophy gives us the flat-line and avenue to use the question tag, 'Why ?' with no bound. This is in line with Albert Einstein Philosophy on Science, 'The method of science is asking the right question and…', Gallian (2013). An original research thesis used 'why'. That is why it earns a PhD (Doctor of Philosophy not Computer Science). It needs the 'why' part before the mathematics. For example, can 1 and 1 produce 2 comes before $1 + 1 = 2$, the mathematics which is in this sense defined 'and' as addition '+'. In this occasion, conjectures are philosophical and theorems (lemmas, corollaries and propositions) are mathematical.

Axiomatic Methodology is the methodology of Axiomatic System. The word system on this occasion means set. Thus, axiomatic system means set of axioms. Axiomatic Methodology is the set of axioms equipped with the operation of deductions. Mostly, the axioms are self evident and deductions include intermarriage.

Just like in language, we have syntax meaning 'grammar' and semantics meaning 'meaning', axiomatic system are the syntax of mathematics and model theory are the semantics.

Universal Algebra is the umbrella of all algebraic structures, Stanley and Sankappanavar (1981). An algebra is a set equipped with at least unary operation. Unary operation is a map inputting a single element and outputting a single element viz: $\emptyset(a) = a^{-1}$, the inverse element of an element of a set. A set endured with one unary operation is called a mono-unary algebra. A set enclosed with two unary operations is called a bi-unary algebra. A set equipped with a binary operation $\emptyset$ by which it means $\emptyset: S \times S \to S$, defined by $\emptyset(s,t) = s\emptyset t$, where $s, t \in S$ is called a groupoid. If $\emptyset$ is associative, it is called a semigroup. Group is a set endured with a binary operation, a unary operation and a nullary operation. So group has three operations: One binary operation, one unary operation and one nullary operation (identity element operation). Binary relation, the generalization of transformations, is the domain of binary operation.

Transformation semigroup is an algebraic structure under universal algebra which also uses Philosophy and Formal and Material Axiomatic Systems.

Results and Discussion

$$\begin{aligned}\text{Let } \mathrm{A} = 1 = \mathrm{a}, \mathrm{B} = 2 = \mathrm{b}, \mathrm{C} = 3 = \mathrm{c}, \mathrm{D} = 4 = \mathrm{d}, \mathrm{E} = 5 = \mathrm{e}, \mathrm{F} = 6 = \mathrm{f}, \mathrm{G} = 7 = \mathrm{g}, \mathrm{H} = 8 = \mathrm{h}, \mathrm{I} = 9 \\ = \mathrm{i}, \mathrm{J} = 10 = \mathrm{j}, \mathrm{K} = 11 = \mathrm{k}, \mathrm{L} = 12 = \mathrm{l}, \mathrm{M} = 13 = \mathrm{m}, \mathrm{N} = 14 = \mathrm{n}, \mathrm{O} = 15 = \mathrm{o}, \mathrm{P} \\ = 16 = \mathrm{p}, \mathrm{Q} = 17 = \mathrm{q}, \mathrm{R} = 18 = \mathrm{r}, \mathrm{S} = 19 = \mathrm{s}, \mathrm{T} = 20 = \mathrm{t}, \mathrm{U} = 21 = \mathrm{u}, \mathrm{V} = 22 \\ = \mathrm{v}, \mathrm{W} = 23 = \mathrm{w}, \mathrm{X} = 24 = \mathrm{x}, \mathrm{Y} = 25 = \mathrm{y}, \mathrm{Z} = 26 = \mathrm{z}, \quad ", " = 27, \quad "." = 28.\end{aligned}$$

That is, case sensitivity does not matter for and as when it matters, the transformation will be but only longer. Then the sentence: "The quick brown fox jumps over a lazy dog" is encrypted by:

$$\begin{pmatrix} 01\ 2\ 3\ \mathbf{4}05\ 06\ 7\ 8\ 09\ \mathbf{10}11\ 12\ 13\ 14\ 15\ \mathbf{16}17\ 18\ 19\ \mathbf{20} \\ 20\ 8\ 5\ \emptyset 17\ 21\ 9\ 3\ 11\ \emptyset\emptyset\ 02\ 18\ 15\ 23\ 14\ \emptyset\emptyset\ 06\ 15\ 24\ \emptyset\emptyset \\ 21\ 22\ 23\ 24\ 25\ \mathbf{26}27\ 28\ 29\ 30\ \mathbf{31}32\ \mathbf{33}34\ 35\ 36\ 37\ \mathbf{38}39\ 40\ 41 \\ 10\ 21\ 13\ 16\ 19\ \emptyset\emptyset\ 15\ 22\ 05\ 18\ \emptyset\emptyset\ 01\ \emptyset\emptyset\ 12\ 01\ 26\ 25\ \emptyset\emptyset\ 04\ 15\ 07 \end{pmatrix}$$

And it is encrypted as well as compressed by:

$$\begin{pmatrix} 01\ 2\ 3\ 05\ 06\ 7\ 8\ 09\ 11\ 12\ 13\ 14\ 15\ 17\ 18\ 19\ 21\ 22\ 23 \\ 20\ 8\ 5\ 17\ 21\ 9\ 3\ 11\ 02\ 18\ 15\ 23\ 14\ 06\ 15\ 24\ 10\ 21\ 13 \\ 24\ 25\ 27\ 28\ 29\ 30\ 32\ 34\ 35\ 36\ 37\ 39\ 40\ 41 \\ 16\ 19\ 15\ 22\ 05\ 18\ 01\ 12\ 01\ 26\ 25\ 04\ 15\ 07 \end{pmatrix}$$

Alphabets are embedded, encrypted and compressed by:

$A = \begin{pmatrix} 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14\ 15\ 16\ 17\ 18\ 19\ 20\ 21\ 22\ 23\ 24\ 25\ 26 \\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14\ 15\ 16\ 17\ 18\ 19\ 20\ 21\ 22\ 23\ 24\ 25\ 26 \end{pmatrix}$. They are also compressed because A ⊙ B ≠ AB, where ⊙ is a space.

Vowel will be encrypted and embedded by:

$$V = \begin{pmatrix} 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14\ 15\ 16\ 17\ 18\ 19\ 20\ 21\ 22\ 23\ 24\ 25\ 26 \\ 1\ \emptyset\emptyset\ 5\ \emptyset\emptyset\ 9\ \emptyset\emptyset\emptyset\emptyset\emptyset\ 15\ \emptyset\emptyset\emptyset\emptyset\emptyset\ 21\ \emptyset\emptyset\emptyset\emptyset\emptyset \end{pmatrix}$$

$$V = \begin{pmatrix} 1\ 5\ 9\ 15\ 21 \\ 1\ 5\ 9\ 15\ 21 \end{pmatrix}$$

Consonants are $A \backslash V$.

One good point is that the sentence "The quick brown fox jumps over a lazy dog'' has all the alphabets embedded – abuse of term − in it. The cyclic notation of the sentence is the following set.

$$\{(0), (20\ 8\ 5), (17\ 21\ 9\ 3\ 11), (2\ 18\ 15\ 23\ 14), (6\ 15\ 24)\} \cup$$

$$\{, (10\ 21\ 13\ 16\ 19), (15\ 22\ 5\ 18), (1), (12\ 1\ 26\ 25), (4\ 15\ 7)\}$$

Where (0) represents an often overlooked space before the first later of the first word of a sentence and each cycle is representing a word in the sentence.

Since (abc) is a cyclic notation, $(bca) = (cab) = (abc)$. This can be used in encrypting a word or all the words in a sentence.

To encrypt the sentence "The quick brown fox jumps over a lazy dog"using cyclic-permutation notation, we adjoin zero map as follows:

$$\{(0), (20\ 8\ 5), (17\ 21\ 9\ 3\ 11), (2\ 18\ 15\ 23\ 14), (6\ 15\ 24)\} \cup$$
$$\{, (10\ 21\ 13\ 16\ 19), (15\ 22\ 5\ 18), (1), (12\ 1\ 26\ 25), (4\ 15\ 7)\}$$

And one simple way to compress the data as well encrypt it is to work over $A\backslash V$.

Thus, the sentence "the quick brown fox jumps over a lazy dog" may be encrypted as well as encoded using$A\backslash V \wedge \begin{pmatrix}1&2&3&4&5\\1&2&3&4&5\end{pmatrix} \Rightarrow \begin{pmatrix}1&2&3&4&5\\2&3&4&5&1\end{pmatrix}$ as: "HT CKQ RWNB XF MPSJ RV ZYL GD" which only an algebraist cryptanalysts can understand the compression key. See (Winter, 1990) for such an idea. The compression key is:

$$A\backslash V \wedge \begin{pmatrix}1&2&3&4&5\\1&2&3&4&5\end{pmatrix} \Rightarrow \begin{pmatrix}1&2&3&4&5\\2&3&4&5&1\end{pmatrix}$$

Now, the decompression expression will now be $(A\backslash V \vee \quad V) \wedge \begin{pmatrix}1&2&3&4&5\\2&3&4&5&1\end{pmatrix} \Rightarrow \begin{pmatrix}1&2&3&4&5\\1&2&3&4&5\end{pmatrix}$, where both$\begin{pmatrix}1&2&3&4&5\\1&2&3&4&5\end{pmatrix}$ and $\begin{pmatrix}1&2&3&4&5\\1&2&3&4&5\end{pmatrix}$ are elements of the dihedral group of degree 5.

$\begin{pmatrix}1&2&3&4&5\\2&3&4&5&1\end{pmatrix} \Rightarrow \begin{pmatrix}1&2&3&4&5\\1&2&3&4&5\end{pmatrix}$implies "HT CKQ RWNB XF MPSJ RV ZYL GD" $\Rightarrow$ "TH QCK BRWN FX JMPS VR LZY DG".

The probability of each element in V that may come after "TH" is $\frac{1}{5}$. This probability by natural intuition will become $\frac{2}{5}$ since most English sentences starts with "The or They or so". The more the number of combinations, the more this probability is getting closer to 1.

For example, FX can imply FIX, FOX, etc; but very rarely BRWN, JMPS, LZY imply any other set of words beside BROWN, JUMPS and LAZY respectively.

Haven decoded these, the two letter words FX, VR and DG will have probability $\frac{3}{5}$ which is 60% of success. With the Check-Digit Scheme (Gallian, 2013), the probability will be $\frac{4}{5}$ (80% success) if not 100%. But, even without, we have seen how the partial transformation encrypted, compressed and embedded the sentence.

The choice of $\begin{pmatrix}1&2&3&4&5\\2&3&4&5&1\end{pmatrix} \Rightarrow \begin{pmatrix}1&2&3&4&5\\1&2&3&4&5\end{pmatrix}$ is arbitrary. Any arrangement over transformation-permutation is possible.

Conclusion

Transformation is another name for one to one relations between two sets (A and B, say) such that all the elements of A are involve in the relationship and no two elements of B have one pre-image in A. Transformation plays significant role in the theory of semigroup whence the Presto-Wagner Theorem still stands. A bijective transformation of a set on to itself is called permutation. Every group is also embeddable in the permutation group. Partial transformations are also forming semigroup just as compound nouns and adjectives are forming group. The idea of this application is inherited from the fact that finite field

which finite semigroup is its generalization has applications to coding theory Gallian(2013). Compression of data exploits size of memory and decrease the cost of communications Winter (1990). Every Shakespearean play is encyptable, embedded and compressed by a partial transformation element of a semigroup. Semigroup has application to coding theory, see Howie (2003).

Recommendations

Data encryption is not as important as de-encrypting it. Encryption code is useless without the de-encryption code. If a function which is also an element of a subset of set of transformations encrypt a data, its inverse will de-encrypt it. Hence, any data (image, text, numerals, etc.) on a plain or space are encrypted and decoding using transformations and permutations consciously or unconsciously. This work recommend that further researches be carried out on applications to image encryption and compression.

Images are on a plain, the complex plane and transformation equations are conformal. We recommend that further research be carried out on conformal image encryption and compression. Still-life hard wares and objects can be encrypted and decoded with matrices and their transformations which are rotations and reflections. Matrices form an algebraic structure (Ruskuc, 2000), which through the Preston-Wagner Theorem (Higgins, 1992) are embedded in transformation algebraic structure. We finally recommend for encryption and compression of n-dimensional still-life matters using tensors, the partial generalizations of matrices, whence physics is the study of matter with respect to energy.

REFERENCES

Evis, H. and Neuson, V. C. (1997). *An Introduction to the Foundation and Fundamental Concept of Mathematics,* Holt, Reinhart, Winston. Retrieved from http://www.amazon.com

Fraser, A.G. (2010). *Data Compression and Automatic Programming*. Retrieved from www.comjnl.oxfordjournals.org.

Gallian, J.A. (2013). *Contemporary Abstract Algebra*, Cengage Learning, India, 8th Edition.

Guy, E.B. (2011). *Introduction to Data Compression*. Retrieved from www.eecs.harvard.edu/~michaelm/CS222/compression.pdf.

Heinstein, I. N. (1964). *Topics in Algebra*, John Wiley and Sons, 2nd Edition.

Higgins, P. M. (1992). *Techniques of Semigroup Theory*, Oxford University Press, Oxford, United States of America.

Howie, J. H. (2003). *Fundamentals of Semigroup Theory*, Oxford University Press, U.S.

Kattan, A. (2006). Universal Lossless Compression Technique with Built in Encryption, *MSc Thesis*, University of Essex, UK.

Ruškuc, N. (2000). *Semigroup Presentations*. Ph.D. Thesis, University of St Andrews, pp. 1-256.

Stanley, B. and Sankappanavar, H. P. (1981). *A Course in Universal Algebra*, Springer-Varlag, United States of America.

Steven, W.S. (2007). *The Scientist and Engineer's Guide to Digital Signal Processing*, Technical Publishing, California.

Verhoeff, J. (1969). *Error Detecting Decimal Codes*, Mathematisch Centrum,Amsterdam.

Winter, S. (1990). Error-Detecting Schemes Using Dihedral Groups, *UMAP Journal,* 11(4): 299-308.

Zirra, P.B. and Gregory, M.W. (2011). Radical Data Compression Algorithm Using Factorization, *International Journal of Computer Science and Security*, 5(2): 221-226.

Ziviani, N., Moura, E., Navarro, G., and Baeza, Y.R. (2000). Compression: A Key for Next Generation Text Retrieval Systems, *IEEE Computer Society*, 33(11): 37-44.