ち sciendo

Commun. Appl. Ind. Math. 10 (1), 2019, 1-18

brought to you by 🔏 CORE

ISSN 2038-0909

DOI: 10.1515/caim-2019-0XXX

Research Article

New Distance Concept and Graph Theory Approach for Certain Coding Techniques Design and Analysis

Khmaies Ouahada^{*}and Hendrik C. Ferreira

Department of Electrical and Electronic Engineering Science, University of Johannesburg, South Africa

*Email address for correspondence: kouahada@uj.ac.za

Communicated by Associate Editor Received on XXXX XX, XXXX. Accepted on XXXX XX, XXXX.

Abstract

A New graph distance concept introduced for certain coding techniques helped in their design and analysis as in the case of distance-preserving mappings and spectral shaping codes. A graph theoretic construction, mapping binary sequences to permutation sequences and inspired from the k-cube graph has reached the upper bound on the sum of the distances for certain values of the length of the permutation sequence. The new introduced distance concept in the k-cube graph helped better understanding and analyzing for the first time the concept of distance-reducing mappings. A combination of distance and the index-permutation graph concepts helped uncover and verify certain properties of spectral null codes, which were previously difficult to analyze.

Keywords: Graph theory, Distance mappings, Spectral codes, Cube graph, index permutation graph. AMS subject classification: 94B05, 94B10, 94C15, 94A24

1. Introduction

Graph theory plays an important and central role in many scientific and engineering fields in connection with the design and the analysis of algorithms such as for computer science, communication networks, coding theory and cryptography [1,2]. A distance in graph theory that is defined as the number of edges between two vertices in the shortest path, plays a central role in the fundamental understanding of the graph itself [3,4]. Therefore many engineering fields that make use of graph theory, take into consideration the distance concept to help understand and analyze better many engineering designs.

For coding technique, the definition of distance may vary from a code to another. In Hamming codes, which are used to correct reversal errors, the distance is defined as the number of different bit between two binary sequences with the same length [5]. In the other hand, Levenshtein codes designed to correct insertions, deletions or substitutions errors, the distance between two words as an example is defined as the minimum number of single character edits, as a result of insertions, deletions or substitutions, required to change one word into the other [6,7]. In general, distance in coding techniques plays a big role in emphasizing the robustness and the capability of the designed codes to correct errors. In the case of convolutional code, which is used in this paper, this distance is called minimum distance, which is defined as the Hamming distance between a codeword and the all-zero path corresponded codeword in the trellis presentation [8].

Distance-preserving mapping codes have been introduced first by Ferreira and Vinck [9] have as anew coding technique for power-line communications [10]. These new codes are based on the convolutional codes structures making use of a mapping techniques [11]-[12] based on algorithms generated from certain designed construction. One of the optimum constructions is called k-cube graph construction [13]. This is actually based on a very popular interconnection topology for parallel computer networks due to its ability to exploit the communication locality found in many parallel applications to reduce message latency [14].

A new way of calculating distance between nodes of a unit cube by making all nodes represented by permutation symbols and the value of the edges is simply a decimal subtraction of the nodes values. A

K. Ouahada, H. Ferreira

geometric sequence was generated as a result of this new introduced distance and played a big role in the generation if the algorithm to design distance preserving mapping codes.

Spectral null codes, are codes with nulls in the power spectral density function and have great importance in certain applications such as transmission systems employing pilot tones for synchronization and track-following servos in digital recording [15]. These codes have many interesting properties. The analysis and presentation of these properties where secured by using graph theory properties and new introduced distance calculation concepts.

The paper is organized as follows: Section 2 introduces definitions and notations to be used for distance-preserving mappings, the cube graph and spectral null codes. Section 3 presents the cube graph construction, where the mapping procedure and algorithm based on the newly introduced distance concept in the k-cube graph. Section 4 explains distance-reducing mappings algorithm and the lower bound on the number of permutations in this special case of mappings. Section 5 presents the index-graphic presentation of spectral null codes. Section 6 makes an approach between distance in graph theory and certain properties in spectral null codes. We conclude with some final remarks and future work in Section 7.

2. Preliminaries

We present in this section a brief overview of related definitions to describe distance-preserving mappings (DPMs), spectral null codes (SNC) and certain graph theory fundamentals.

2.1. Distance-preserving mappings

Distance-preserving mappings is a coding technique, which maps the outputs of a convolutional code to other codewords, e.g. codewords from a permutation codebook with less error-correction capabilities.

Let $\mathcal{P}_M = \{1, 2, \dots, M\}$, an integer sequence of length M and let \mathcal{S}_M denote the set of all permutations of symbols in \mathcal{P}_M . \mathbb{Z}_2^n is the set of all binary vectors of length n.

The mappings considered, take all possible binary sequences of length n, from the binary set denoted by \mathcal{C}_b , with $|\mathcal{C}_b| = 2^n$, to a subset of \mathcal{S}_M , denoted by \mathcal{C}_p where we have $|\mathcal{C}_p| = |\mathcal{C}_b|$.

Applying the definition of the Hamming distance to binary and non-binary sequences, we can denote by $\mathbf{D} = [d_{ij}]$ the distance matrix whose entries are the Hamming distances between any two binary sequences \mathbf{X}^i and \mathbf{X}^j defined as follows:

(1)
$$\mathbf{D} = [d_{ij}] \text{ with } d_{ij} = d_H(\mathbf{X}^i, \mathbf{X}^j).$$

Similarly for permutation sequences we denote by $\mathbf{E} = [e_{ij}]$ the distance matrix whose entries are the Hamming distances between any two permutation sequences \mathbf{Y}^i and \mathbf{Y}^j defined as follows:

(2)
$$\mathbf{E} = [e_{ij}] \text{ with } e_{ij} = d_H(\mathbf{Y}^i, \mathbf{Y}^j).$$

The sum of all the distances in \mathbf{D} and \mathbf{E} are defined as follow:

(3)
$$\|\mathbf{D}\| = \sum_{i=1}^{|\mathcal{C}_b|} \sum_{j=1}^{|\mathcal{C}_b|} d_{ij}.$$

(4)
$$\|\mathbf{E}\| = \sum_{i=1}^{|\mathcal{C}_p|} \sum_{j=1}^{|\mathcal{C}_p|} e_{ij}.$$

The total distances, $\|\mathbf{D}\|$ and $\|\mathbf{D}\|$ play a role in the permutation code's error correcting capabilities, as was shown in [12].

In our case, since only the mappings themselves are considered, we will redefine it purely in terms of the Hamming distance between the binary and permutation sequences on how it is preserved. In general, three types of DPMs can be defined as following,

Definition 2.1. [16] A distance-conserving mapping (DCM) is a distance-preserving mapping that only guarantees conservation of the base code's minimum distance. Hence at least one $e_{ij} = d_{ij}$ while for all other i and j, $e_{ij} \ge d_{ij}$, $i \ne j$.

Definition 2.2. [16] A distance-increasing mapping (DIM) is a distance-preserving mapping that guarantees that the resulting code's distance will have some increase above the base code's minimum distance. Hence, $e_{ij} \ge d_{ij} + \delta$ for $i \ne j$, where $\delta \in \{1, 2, ...\}$, depending on the increase obtained.

Definition 2.3. [16] A distance-reducing mapping (DRM) is a distance-preserving mapping that yields a mapping of which the distance is lower than that of the base code, although this decrease is controlled. The decrease is guaranteed to be not more than a fixed amount, δ , hence with $e_{ij} \ge d_{ij} + \delta$, $\delta \in \{-1, -2, \ldots\}$ for $i \neq j$.

In general and for the sake of simple calculations, we assume that $\delta = (M - n)$.

An upper bound on the sum of the Hamming distances was introduced by Swart and Ferreira [12] as a means to compare between different constructions. The upper bound can be stated as follows:

(5)
$$\|\mathbf{E}_{\max}\| = M\left(|2^n|^2 - (2\alpha\beta + \beta + \alpha^2 M)\right),$$

where $\alpha = \lfloor 2^n/M \rfloor$ and $\beta \equiv 2^n \mod M$, with $\lfloor . \rfloor$, the floor function, producing the integer part after division and mod producing the remainder after division [12].

To compare any mapping to the optimal mapping, a distance optimality factor is defined as follows:

Definition 2.4. [12] The *distance optimality factor* of a mapping is given by

(6)
$$\eta = \frac{\|\mathbf{E}\|}{\|\mathbf{E}_{\max}\|},$$

where $\|\mathbf{E}\|$ is the sum of the Hamming distances for the mapping and $\|\mathbf{E}_{\max}\|$ is the upper bound on the sum of the Hamming distances for the mapping.

In the case of the mapping using the permutation set $\{1243, 1342, 4213, 4312\}$, we have $\|\mathbf{E}\| = 48$ and $\eta = 1$ and the mapping is called to have attained the upper bound.

2.2. Cube Graph Theory

Definition 2.5. The *k*-cube is the graph with vertex set $\{0,1\}^k$ and with an edge between each pair of sequences that differ in exactly one position (i.e. have Hamming distance 1).

The k-cube graph is a bipartite graph since there is a partition of its vertex sets into two parts, A and B, with every vertex in A is adjacent only to vertices in B and vice versa. In the k-cube graph we can find:

- 2^k vertices, each of which has the degree k.
- that the number of edges in a k-cube is equal to $k \times 2^{k-1}$.

Definition 2.6. The *cube distance*, denoted by $D_{\mathcal{Q}}$, is the absolute value of the difference between two integer values, In_1 and In_2 corresponding to two vertices v_1 and v_2 respectively in a cube. It is defined as: $D_{\mathcal{Q}} \{v_1, v_2\} = |In_1 - In_2|, 0 \leq In_1, In_2 \leq 2^k - 1.$

Definition 2.7. A *distance grouping*, denoted by DG_i with $1 \le i \le k$, is a set of all pairs of equidistant vertices.

We have k distance groupings for each k-cube graph defined by $DG_i = \{(v_j, v_t) | 1 \le j, t \le M, D_Q(v_j, v_t) = 2^{i-1}\}.$

2.3. Spectral Null Codes

The technique of designing codes to have a spectrum with nulls occurring at certain frequencies, i.e. having the power spectral density (PSD) function equal to zero at these frequencies, started with Gorog [17,18], when he considered the vector $\mathbf{X} = (x_1, x_2, ..., x_M)$, $x_i \in \{-1, +1\}$ with $1 \le i \le M$, to be an element of a set S, which is called a codebook of codewords with elements in $\{-1, +1\}$. We investigate codewords of length, M, as an integer multiple of N, thus let

M = Nz,

where f = r/N represents the spectral nulls at rational sub multiples r/N [19]. Following [19], we have to satisfy,

$$(7) A_1 = A_2 = \dots = A_N,$$

where

(8)
$$A_{i} = \sum_{\lambda=0}^{z-1} x_{i+\lambda N}, \quad i = 1, 2, \dots, N.$$

If all the codewords in a codebook satisfy these equations, the codebook will exhibit nulls at the required frequencies. henceforth we present the channel symbol -1 with binary symbol 0.

Definition 2.8. A spectral null binary block code of length M is a subset $\mathcal{C}_b(M, N) \subseteq \{0, 1\}^M$ of all binary M-tuples of length M which have spectral nulls at the rational submultiples of the symbol frequency 1/N.

Definition 2.9. The spectral null binary codebook $C_b(M, N)$ is a subset of the M dimensional vector space $(\mathbb{F}_2)^M$ of all binary M-tuples, where \mathbb{F}_2 is the finite field with two elements, whose arithmetic rules are those of mod-2 arithmetic.

For codewords of length M consisting of N interleaved subwords of length z, the cardinality of the codebook $C_b(M, N)$ for the case where N is a prime number is presented by the following formula [20],

(9)
$$|\mathcal{C}_b(M,N)| = \sum_{i=0}^{M/N} \binom{M/N}{i}^N$$

where $\binom{M/N}{i}$ denotes the combinatorial coefficient $\frac{(M/N)!}{i!(M/N-i)!}$.

3. k-Cube Graph Construction

New terminology will be introduced in this section since the k-cube graph is now used in distancepreserving mappings. The vertices will be called the symbols of a permutation sequence and the edges as the transpositions or swaps between them.

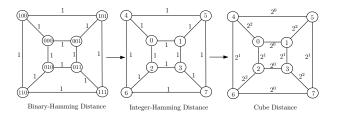


Figure 1. Distance presentation of a 3-cube

The value M representing the length of the permutation sequence is the same as the number of the vertices in a cube graph. In the cube graph the value M is always equal to 2^k . Since we are using permutation sequences with lengths not always equal to 2^k , and thus we are dealing not with a perfect cube, we have to define in general the value of k as:

(10)
$$k = \lceil log_2 M \rceil,$$

where $\lceil . \rceil$, the ceiling function, producing the highest integer close to $log_2 M$. The idea of the k-cube graph construction is based on the concept of the cube distance. Fig. 1 shows different ways of presenting distances in a cube.

It is clear that the Hamming distance, whether for binary sequences or integer sequences, does not help in differentiating the edges since they all have the value 1. On the other hand, the cube distance gives a geometric sequence with a common ratio of 2 since the distances are from 2^0 till 2^{k-1} . Fig. 1 shows that when using the cube distance approach, all parallel edges are equidistant.

We can summarize the properties of a k-cube with cube distances on the edges as follows:

- 1. Any vertex in a k-cube graph has degree k and all the edges incident to a vertex have distances that are in a geometric sequence from 2^0 till 2^{k-1} ;
- 2. Parallel edges are equidistant. Diagonal edges are also considered as parallel to one another since they are equidistant;
- 3. Distances in a k-cube increase when k increases;
- 4. The (k-1)-cube graph is a subgraph of the k-cube graph.

Proposition 3.1. For a k-cube graph, the cube distances for each vertex with other vertices connected to it, form a geometric sequence with a common ratio of 2, from 2^0 to 2^{k-1} .

Proof. We denote by W^i a binary codeword of k bits, where $1 \leq i \leq 2^k$. We denote by W^i_d the corresponding decimal value of the binary codeword W^i .

In general we have $W^i = \left(W^i_{(k)}, W^i_{(k-1)}, \dots, W^i_{(1)}\right)$, where $1 \le i \le 2^k$ and $W^i_{(t)} \in \{0, 1\}$ with $1 \le t \le k$. The corresponding decimal value is then $W^i_d = W^i_{(k)} \times 2^{k-1} + W^i_{(k-1)} \times 2^{k-2} + \dots + W^i_{(2)} \times 2^1 + W^i_{(1)} \times 2^0$.

From the definition of the k-cube graph we have Hamming distance $d_H(W^i, W^j) = 1$ if W^i is adjacent to W^j ; this means that there is a unique $t, 1 \le t \le k$, with $W_t^i \ne W_t^j$.

If we take the following:

$$W^{i} \to W^{i}_{d} = W^{i}_{(k)} \times 2^{k-1} + W^{i}_{(k-1)} \times 2^{k-2} + \cdots + W^{i}_{(2)} \times 2^{1} + W^{i}_{(1)} \times 2^{0}$$
$$W^{j} \to W^{j}_{d} = W^{j}_{(k)} \times 2^{k-1} + W^{j}_{(k-1)} \times 2^{k-2} + \cdots + W^{j}_{(2)} \times 2^{1} + W^{j}_{(1)} \times 2^{0}$$

it then follows, since

$$d_{H}(W^{i}, W^{j}) = 1, \text{ that}$$

$$D_{\mathcal{Q}}\left\{W_{d}^{i}, W_{d}^{j}\right\} = \left|W_{d}^{i} - W_{d}^{j}\right|$$

$$= \left|W_{(t)}^{i} \times 2^{t-1} - W_{(t)}^{j} \times 2^{t-1}\right|$$

$$= \left|1 \times 2^{t-1} - 0 \times 2^{t-1}\right|$$

$$= 2^{t-1}, \text{ where } 1 < t < k.$$

From the definition of the degree of the k-cube, each vertex has degree k and thus the k vertices adjacent to it have each Hamming distance 1. Thus we have k decimal values for the edges as follows:

(11)
$$D_{\mathcal{Q}}\left\{W_d^i, W_d^j\right\} \in \left\{2^{t-1} | 1 \le t \le k\right\}.$$

which can also be written as

(12)
$$D_{\mathcal{Q}}\left\{W_d^i, W_d^j\right\} \in \left\{2^0, 2^1, \dots, 2^{k-1}\right\}.$$

Consequently, this proves that we have a geometric sequence with a common ratio of 2 from 2^0 to 2^{k-1}

3.1. Mapping Procedure

From the properties of the k-cube graph previously presented we can write the steps of our mapping algorithm as follows:

- 1. Consider a random permutation sequence of length M, $(y_1, y_2, \ldots, y_i, \ldots, y_j, \ldots, y_M)$ as a start sequence, where $y_i \in \mathcal{P}_M$,
- 2. Calculate the cube distances on each edge of the cube,
- 3. Calculate the degree of the cube $k = \lceil \log_2 M \rceil$,
- 4. Group the symbols from the start sequence, that have the same cube distance and related to the value of k in the same distance grouping as follows:

(13) Distance grouping 1:
$$(1,2), (3,4), (5,6), \dots$$

Distance grouping 2: $(1,3), (2,4), (5,7), \dots$
Distance grouping 3: $(1,5), (2,6), (3,7), \dots$
Distance grouping 4: $(1,9), (2,10), (3,11), \dots$
 \vdots \vdots

5. Swap the elements in each distance grouping separately. When we have any of the convolutional outputs x_i equal to one then we "swap" the permutation symbols.

We use $\mathcal{Q}(M, n, \delta)$ to denote the DPMs from binary sequences of length n to permutation sequences of length M making use of the k-cube graph construction. The type of mapping is indicated by δ , with $\delta > 0$ for DIMs, $\delta = 0$ for DCMs and $\delta < 0$ for DRMs. Following are a few examples describing the mapping algorithm for the k-cube graph construction taking into consideration the value of M. In both cases where $M = 2^k$ or $M \neq 2^k$, the mapping procedure is the same just the second step changes where the value of k is calculated differently.

3.2. Mapping with $M = 2^k$

Example 3.1. If we take the case of M = 4, which corresponds to a 2-cube graph, we can break down our symbols into two distance groupings as presented in (14). We denote by $D_Q(DG_i)$, with $1 \le i \le k$, the corresponding cube-distance value of DG_i .

(14)
$$D_{\mathcal{Q}} = 2^{0} \rightarrow DG_{1} = \{(1,2); (3,4)\} \\ D_{\mathcal{Q}} = 2^{1} \rightarrow DG_{2} = \{(1,3); (2,4)\}$$

$$\begin{array}{c} 1 & 2^{0} & 2 \\ 2^{1} & 2^{1} \\ 3 & 2^{0} & 4 \end{array} \qquad \left\{ \begin{array}{c} D_{\mathcal{Q}}(DG_{1}) = 2^{0} & \longrightarrow \begin{cases} (1,2) \longrightarrow x_{1} \\ (3,4) \longrightarrow x_{2} \\ \\ D_{\mathcal{Q}}(DG_{2}) = 2^{1} & \longrightarrow \end{cases} \begin{array}{c} (1,3) \longrightarrow x_{3} \\ (2,4) \longrightarrow x_{4} \end{array} \right.$$

Figure 2. Distance groupings of a 2-cube

Graph Theory Approach for Certain Coding Techniques

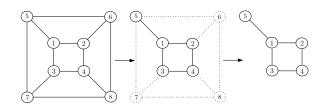


Figure 3. Subgraph presentation of 3-cube

Fig. 2 shows all distance groupings in a 2-cube graph. We associate an input x_i , $1 \le i \le k \times 2^{k-1}$, to each pair having the same distance grouping. The total number of inputs is n = 4, which is the output of the base code.

Starting with any permutation sequence, the mapping algorithm based on the swapping of the symbols of the equidistant edges is presented as follow:

```
Mapping algorithm for Q(4, 4, 0)

Input: (x_1, x_2, x_3, x_4) \in \mathbb{Z}_2^4

Output: (y_1, y_2, y_3, y_4) \in S_4

Start Sequence: (1, 2, 3, 4) \rightarrow (y_1, y_2, y_3, y_4)

begin

if x_1 = 1 then swap(y_1, y_2)

if x_2 = 1 then swap(y_3, y_4)

if x_3 = 1 then swap(y_1, y_3)

if x_4 = 1 then swap(y_2, y_4)

end.
```

From Definition 2.1, we can say that in this example the mapping is considered as conserving since $\delta = 0$.

3.3. Mapping with $M \neq 2^k$

Example 3.2. Here we take the case of $M \neq 2^k$ (e.g. M = 5, where k = 3 from the second step in the mapping procedure), which could be identified by using some graph theoretic definitions or the concept of the cube distance that we have introduced.

3.3.1. Using the Subgraph Theory

We consider the 3-cube graph, which has 8 vertices representing the length of the permutation sequence. We can reduce the length of the permutation sequence by eliminating the vertices that we do not need. In this case, the vertices are 6, 7 and 8. Then we eliminate all the edges incident to those vertices, which means we are not interested in their corresponding swaps. Fig. 3 shows that the vertex 5 is adjacent only to the vertex 1 and the final number of edges, which corresponds to the number of swaps is five. In this case we are dealing with a distance-conserving mapping since $\delta = 0$.

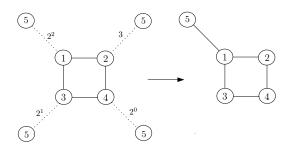


Figure 4. Distance based construction of a cube

The corresponding algorithm for this DCM is,

```
Mapping algorithm for Q(5,5,0)

Input: (x_1, x_2, x_3, x_4, x_5) \in \mathbb{Z}_2^5

Output: (y_1, y_2, y_3, y_4, y_5) \in S_5

Start Sequence: (1, 2, 3, 4, 5) \rightarrow (y_1, y_2, y_3, y_4, y_5)

begin

if x_1 = 1 then swap(y_1, y_2)

if x_2 = 1 then swap(y_3, y_4)

if x_3 = 1 then swap(y_1, y_3)

if x_4 = 1 then swap(y_2, y_4)

if x_5 = 1 then swap(y_1, y_5)

end.
```

3.3.2. Using the cube distance

The 5th vertex could be connected to any vertex of the 2-cube. Fig. 4 shows the possible distances when the symbol 5 is connected to others. It is clear that the distance in the case of (2, 5) is not acceptable since it is equal to 3, which is not in the form of a geometric sequence with a common ratio of 2. So the symbol 5 could be linked to 1 with a distance of 2^2 or to 3 with a distance of 2^1 or to 4 with a distance of 2^0 . Since k is defined in (10), the symbol 5 belongs to the 3-cube and not to a 2-cube and here the geometric sequence ends at the value of $2^{(3-1)}$, which is 2^2 . Thus the symbol 5 should be connected to the symbol 1 as shown in Fig. 4.

It can be seen that we could reach the same results by either using graph theoretic or cube distance approach. Our results prove the reliability of our cube-distance theory without changing the properties of the cube graph.

4. Distance-Reducing Mappings: A Graphical Analysis

The simplicity of the k-cube graph as mentioned previously could help understanding one of the difficult mappings mechanism which is the reducing mappings. To our knowledge previous published constructions did not discuss or touch in details the case of the reducing mappings although some have calculated some cases of reducing mappings without any emphasis of the mechanism of this mapping.

4.1. The lower value of M for a reducing mapping

From Fig. 5 we can see that for the values M = 4 and M = 5, the maximum number of edges corresponding to the number of convolutional outputs n that we can have are only 4 or 5 respectively, which means that we can have in these cases only conserving mappings since n = M. Whereas, in the case of M = 6, as it can be seen from Fig. 5, the number of edges increases to n = 7 and therefore $\delta = -1$; leading to a reducing mapping.

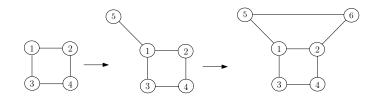


Figure 5. Subgraph presentation of 3-cube

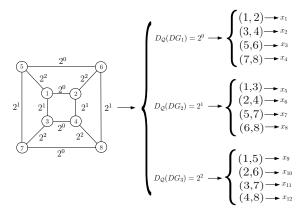


Figure 6. Distance groupings of a 3-cube

It can be clearly seen that the cube graph helped us understanding why we do not have reducing mapping for M = 4 and M = 5 and that the only values to guarantee reducing mappings are for M larger than 6.

Proposition 4.1. Distance-reducing mappings is achievable only with $M \ge 6$.

Proof. To prove this proposition we can suppose the opposite, that we can obtain reducing mappings with values of M smaller than 6. Let us take the example of M = 5, where the symbol 5 in the cube graph is connected to all other symbols, 1, 2, 3, and 4 and the number of edges is 8. As was previously proved, the geometric sequences which is a fundamental property of the cube graph and thus for our optimum construction, makes this assumption not true and therefore the symbol 5 cannot be connected only to the symbol 1. This means that the number of edges for the case of M = 5 is only 5 and in this case we can have only a conserving mapping. A similar analysis holds for the case of $M \leq 4$. The condition in the proposition is then proved.

Example 4.1. In the case of M = 8, it is clear from Fig. 6 that we have three distance groupings presented as follows:

$$D_{\mathcal{Q}} = 2^{0} \rightarrow DG_{1} = \{(1,2); (3,4); (5,6); (7,8)\}$$
$$D_{\mathcal{Q}} = 2^{1} \rightarrow DG_{2} = \{(1,3); (2,4); (5,7); (6,8)\}$$
$$D_{\mathcal{Q}} = 2^{2} \rightarrow DG_{3} = \{(1,5); (2,6); (3,7); (4,8)\}$$

K. Ouahada, H. Ferreira

k-Cube	M	Number of Edges	δ	DPM
2	4	4	0	DCM
3	8	12	-4	DRM
4	16	32	-16	DRM
5	32	80	-48	DRM
6	64	192	-128	DRM
7	128	448	-320	DRM

The corresponding algorithm for this case is,

Mapping algorithm for $\mathcal{Q}(8, 12, -4)$ Input: $(x_1, x_2, \ldots, x_{12}) \in \mathbb{Z}_2^{12}$ Output: $(y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8) \in S_8$ Start Sequence: $(1, 2, 3, 4, 5, 6, 7, 8) \rightarrow (y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8)$ begin if $x_1 = 1$ then swap (y_1, y_2) if $x_2 = 1$ then swap (y_3, y_4) if $x_3 = 1$ then swap (y_5, y_6) if $x_4 = 1$ then swap (y_7, y_8) if $x_5 = 1$ then swap (y_1, y_3) if $x_6 = 1$ then swap (y_2, y_4) if $x_7 = 1$ then swap (y_5, y_7) if $x_8 = 1$ then swap (y_6, y_8) if $x_9 = 1$ then swap (y_1, y_5) if $x_{10} = 1$ then swap (y_2, y_6) if $x_{11} = 1$ then swap (y_3, y_7) if $x_{12} = 1$ then swap (y_4, y_8) end.

We can see that n = 12, which is larger than M = 8. In this case we have $\delta = -4$ and the mapping is considered as reducing.

4.2. The lower bound of δ

We now investigate the issue of distance reducing mappings, on which no results have previously been represented.

Taking into account the number of vertices in a cube as the number of permutation symbols, we can consider the total number of edges as the total number of transpositions or swaps that can be achieved with our distance-preserving mappings, in fact distance-reducing mappings.

Proposition 4.2. The lower bound of δ in distance-preserving mappings for a cube construction is $\delta_{min} = 2^{k-1}(2-k)$.

Proof. From previous definitions and results, the maximum number of swaps for an optimum mapping with cube construction is in fact the number of edges. Thus the maximum number of output bits of a convolutional code should be equal to the total number of edges in the cube, which means that $n_{max} = k \times 2^{k-1}$.

From previous definitions of the distance-preserving mappings we have defined δ as $\delta = M - n$ and thus $\delta_{min} = M - n_{max}$. So the lower bound is $\delta_{min} = 2^k - k \times 2^{k-1}$, which leads to $\delta_{min} = 2^{k-1}(2-k)\Box$

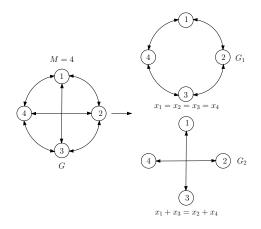


Figure 7. Equation representation for Graph M = 4

Although the value of δ is easy to calculate, the corresponding value for an optimum mapping stays a problem especially when it comes to a distance-reducing mapping. The simplicity of the cube graph construction has solved this problem as presented in the proposition. Table 1 shows the values of δ for different values of M and their corresponding types of mappings.

5. Index-Graphic Presentation of Spectral Null Codes

The idea of the index-graphic [21] presentation of the spectral null codes is actually based on the presentation of the indices of the variables in each grouping of the spectral null equation (7).

Definition 5.1. We denote by $I_p(i, \lambda)$ the permutation symbol of the corresponding index of the variable $x_{i+\lambda N}$ in (8).

(15)
$$I_p(i,\lambda) = i + \lambda N \quad \text{where} \quad \begin{cases} i = 1, 2, \dots, N, \\ \lambda = 0, 1, \dots, z - 1. \end{cases}$$

Definition 5.2. We denote by $\mathcal{P}_{I_p}(M, N)$ the index-permutation sequence from a spectral null equation for variables of length M = Nz as presented.

(16)
$$\mathcal{P}_{I_p}(M,N) = \prod_{i=1}^N \prod_{\lambda=0}^{z-1} I_p(i,\lambda)$$

The production sign in (16) is not used in its traditional way, but just to give an idea about the sequence and the order of the permutation symbols.

Example 5.1. For the case of M = 4 with N = 2, we have

(17)
$$A_1 = A_2 \to x_1 + x_3 = x_2 + x_4$$

We can see from (17), that the indices of the variables x_i , using (16), are represented by the symbols $I_p(1,0) = 1$, $I_p(1,1) = 3$, $I_p(2,0) = 2$ and $I_p(2,1) = 4$. The index-permutation sequence is then $\mathcal{P}_{I_p}(4,2) = (13)(24)$.

An index-permutation symbol is presented graphically by just being lying on a circle, which it is called a state. The state design follow the order of appearance of the indices in (17). The symbols are connected in respect of the addition property of their corresponding variables in (17).

Spectral null codebooks have the all-zeros and all-ones codewords [19], where all the variables x_i are equal. This is depicted in Fig. 7 by the graph G_1 . The obtained relationship between the variables $x_1 = x_2 = x_3 = x_4$ is actually a special case of the equation representing the graph G_2 in Fig. 7. So

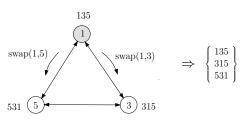


Figure 8. Index-permutation sequences

we limit our studies to the equations of spectral null codes and their corresponding graphs to study the cardinalities and other properties of these codes.

Fig. 7 shows that the graph G, which is the general form of all possible permutations is actually the combinations or the union, $G = G_1 \cup G_2$, of other subgraphs related to the spectral null equations.

6. Graph Theory and Spectral Null Codes

6.1. Cardinalities Approach

6.1.1. Hamming Distance Approach

We make use of the Hamming distance between two permutation sequences representing the indexpermutation symbols of each grouping A_i of the spectral null equation.

To generate the permutation sequences, we start with any state representing an index-permutation symbol in each grouping as appearing in (7). A permutation sequence used as a starting point, contains the symbol from the start state followed by the rest of symbols from the other states taking into consideration the order of the symbols as appearing in (7). Fig. 8 shows the starting permutation sequence as 135. We swap the state-symbol with the following state-symbol in the permutation sequence based on the k-cube construction introduced in the previous sections. We end the swapping process at the last state in the graph. We do not swap symbols between the last state and the starting state for the reason to not disturb the obtained sequences at each state. As an example, for M = 6, Fig. 8 depicts the swaps and shows the resultant index-permutation codebooks for one grouping.

Definition 6.1. The Hamming distance $d_H(\mathbf{Y}^i, \mathbf{Y}^j)$ is defined as the number of positions in which the two sequences \mathbf{Y}^i and \mathbf{Y}^j differ. We denote by $\mathcal{H}_d(M, N)$ the distance matrix, which entries are the distances between index-permutation sequences from a spectral null code of length M = Nz defined as follows:

(18)
$$\mathcal{H}_d(M,N) = [h_{i,j}] \quad \text{with} \quad h_{i,j} = d_H(\mathbf{Y}^i, \mathbf{Y}^j).$$

Definition 6.2. The sum on the Hamming distances in the $\mathcal{H}_d(M, N)$ distance matrix is

(19)
$$\|\mathcal{H}_d(M,N)\| = \sum_{i=1}^M \sum_{j=1}^M h_{i,j}.$$

Example 6.1. For M = 6 with N = 2 and z = 3, the corresponding spectral null equation is

(20)
$$\overbrace{x_1 + x_3 + x_5}^{A_1} = \overbrace{x_2 + x_4 + x_6}^{A_2}$$

The equation (20) is presented by the graph in Fig. 9, where the index-permutation symbols are presented

Graph Theory Approach for Certain Coding Techniques

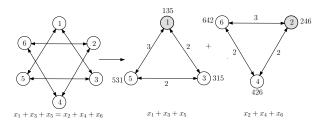


Figure 9. Distances for Graph M = 6 with N = 2

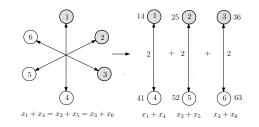


Figure 10. Distances for Graph M = 6 for N = 3

with their corresponding Hamming distances.

(21)
$$\mathcal{H}_{d}(6,2) = \begin{bmatrix} 135 & 315 & 513 & 246 & 426 & 624 \\ 135 & 0 & 2 & 3 & 0 & 0 & 0 \\ 315 & 2 & 0 & 2 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 3 & 0 \\ 426 & 0 & 0 & 0 & 3 & 2 & 0 \end{bmatrix}$$

Each grouping in (20) is represented by a subgraph as depicted in Fig. 9. The Hamming distance matrix for all possible index-permutation sequences is presented in (21), where "0" represents the Hamming distance between same sequences or sequences with non connected symbols.

Example 6.2. For M = 6 with N = 3 and z = 2, the corresponding spectral null equation is

(22)
$$\overbrace{x_1 + x_4}^{A_1} = \overbrace{x_2 + x_5}^{A_2} = \overbrace{x_3 + x_6}^{A_3}.$$

The equation (22) is presented by the graph in Fig. 10. Using the concept of graph distance and the permutation sequences, we can have the distance values as depicted in Fig. 10. The corresponding subgraphs for A_1 , A_2 and A_3 are presented in Fig. 10.

(23)
$$\mathcal{H}_{d}(6,3) = \begin{array}{c} 14 & 41 & 25 & 52 & 36 & 63 \\ 14 & 0 & 2 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 25 & 0 & 0 & 0 & 2 & 0 & 0 \\ 36 & 0 & 0 & 0 & 0 & 2 & 0 \\ 36 & 0 & 0 & 0 & 0 & 2 & 0 \\ 63 & 0 & 0 & 0 & 0 & 2 & 0 \end{array}$$

The Hamming distance matrix for all possible index-permutation sequences is presented in (23).

$$\|\mathcal{H}_d(6,2)\| = 28 > \|\mathcal{H}_d(6,3)\| = 12.$$

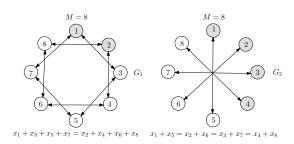


Figure 11. Equation representation for Graph M = 8

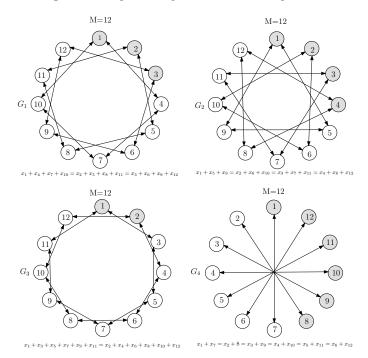


Figure 12. Equation representation for Graph M = 12

Example 6.3. In the case of M = 8, with whether N = 4 or N = 2, we have the corresponding graph of each case as depicted in Fig. 11.

$$\|\mathcal{H}_d(8,2)\| = 40 > \|\mathcal{H}_d(8,4)\| = 16.$$

Example 6.4. In the case of M = 12, we have four combinations where the value of N could be N = 4, N = 3, N = 2 or N = 6. In each case we have a graph representing the spectral null equation as depicted in Fig. 12.

$$\|\mathcal{H}_d(12,2)\| = 64 > \|\mathcal{H}_d(12,6)\| = 24,$$

and $\|\mathcal{H}_d(12,3)\| = 60 > \|\mathcal{H}_d(12,4)\| = 56.$

Theorem 6.1. The sum on the Hamming distances for all index-permutation sequences is

$$\|\mathcal{H}_d(M,N)\| = \begin{cases} 4N, & \text{for } z = 2, \\ 2N(3z-2), & \text{for } z \ge 3. \end{cases}$$

Proof. It is clear that the matrix $\mathcal{H}_d(M, N)$ is symmetric, so we just look for the half of the results of the theorem. For the case of z = 2 the proof is trivial since we swap only two symbols in each indexpermutation sequence. Thus the sum on the distances is $4 \times N$. For the case of $z \ge 3$ we have a cycle graph, where the number of edges is equal to the number of vertices. Since we swap two symbols each

time we move from one state to another, the distance at each edge is equal to two, except for the last edge connecting the first state to the last state where all symbols are swapped and the distance is equal to the length of the index-permutation sequences, which is z. The sum on the Hamming distances for a cycle graph for each grouping is $2 \times (z - 1) + z = 3 \times z - 2$. Thus the result on the sum of the Hamming distances in the matrix is $2 \times N \times (3 \times z - 2)$.

6.1.2. Graph-swap Distance Approach

In coding theory, for the case of convolutional codes, the constraint length K plays an important role in the value of the minimum Hamming distance [8,22]. In our graph approach the length of each grouping A_i , which is equal to the value of z plays an important role in the value of the graph distance.

Definition 6.3. We denote by $\mathcal{M}_{\mathcal{G}_d}(M, N)$ the graph-swap distance matrix, whose entries are the graph distances between two index-permutation symbols from a spectral null code of length M = Nz.

Example 6.5. For the case of M = 8 with N = 2 or N = 4, the corresponding graph-swap distance matrices are

and

$$\mathcal{M}_{\mathcal{G}_d}(8,4) = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 6 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 1 & 0 & 0 & 0 & 0 \\ 8 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

It is clear that we have $\|\mathcal{M}_{\mathcal{G}_d}(8,2)\| = 32 > \|\mathcal{M}_{\mathcal{G}_d}(8,4)\| = 8.$

Theorem 6.2. The sum on the graph distances for all index-permutation symbols is

$$\|\mathcal{M}_{\mathcal{G}_d}(M,N)\| = \begin{cases} \left(\frac{z}{2}\right)^2 M, & \text{for } z \text{ even,} \\ \frac{z^2 - 1}{4}M, & \text{for } z \text{ odd.} \end{cases}$$

Proof. The graphs that we are using are cycle graphs. As long as we go through the edges of a graph the graph distance is incremented by one. When z is even, the first state has the farthest state to it located at $\frac{z}{2}$. So the graph distances from the first state to the $\frac{z}{2}$ state are in a numerical series of ratio one from one to $\frac{z}{2}$. From the state at the position $\frac{z}{2} - 1$ till the first state, the graph distances are in a numerical series of ratio one from one to $\frac{z}{2} - 1$. Adding the two series we get the final sum equal to $\left(\frac{z}{2}\right)^2 M$. Same analogy for the case of z as odd with a numerical series from one till $\frac{z-1}{2}$.

K. Ouahada, H. Ferreira

M N z	$ \mathcal{C}_b(M,N) $	$ \mathcal{H}_d(M,N) $	$ \mathcal{G}_d(M,N) $	$ \mathcal{N}_{A_d}(M,N) $
6 3 2	10	12	4	6
$6\ 2\ 3$	20	28	12	12
8 4 2	36	16	8	8
8 2 4	70	40	32	24
10 5 2	34	20	10	10
$10\ 2\ 5$	252	52	60	40
12 6 2	250	24	12	12
$12\ 4\ 3$	300	56	24	24
$12 \ 3 \ 4$	346	60	48	36
$12\ 2\ 6$	924	64	108	60
$15 \ 5 \ 3$	488	70	30	30
$15 \ 3 \ 5$	2252	78	90	60

6.1.3. Adjacency-swap Matrix Approach

We introduce the adjacency-swap matrix inspired by graph theory as follows.

Definition 6.4. The adjacency-swap matrix of index-permutation symbols is an $M \times M$ matrix $\mathcal{N}_{A_d}(M, N) = (n_{i,j})$ in which the entry $n_{i,j} = 1$ if there is a swap between an index symbol i and an index symbol j and is 0 if there is no swap between index symbol i and index symbol j as presented in each grouping of a spectral null equation.

Example 6.6. For the case of M = 6 with N = 2 or N = 3, the corresponding adjacency-swap matrices are

$$\mathcal{N}_{A_d}(6,2) = \begin{array}{c} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 & 0 \\ 4 & 1 & 0 & 0 & 0 & 0 & 1 \\ 5 & 0 & 1 & 0 & 0 & 0 & 0 \\ 5 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 5 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \end{array} \right],$$

and

We can see that $\|\mathcal{N}_{A_d}(6,2)\| = 12 > \|\mathcal{N}_{A_d}(6,3)\| = 6.$

Theorem 6.3. The total number of swaps in an adjacency-swap matrix is

$$\|\mathcal{N}_{A_d}(M,N)\| = (z-1)M$$

Proof. The proof is trivial as per grouping we have z index-permutation symbols. Thus we have z - 1 ones in each row of the matrix $\mathcal{N}_{A_d}(M, N)$ which refer to the possible swaps of each symbol with others in the same grouping. The total number of swaps is $(z - 1) \times M$.

Graph Theory Approach for Certain Coding Techniques

Table 2 presents few examples of the relationship between the cardinalities of spectral null codes denoted by $C_b(M, N)$ and their correspondences of graph distances. It is clear from Table 2 that the cardinalities of different codebooks with the same length of codewords, increase when the number of swaps increases. This results is also verified in Table 2 based on the concept of distances from graph theory perspective.

7. Conclusion

We have introduced new distances metrics and presented in this paper a graph theoretic approach to certain topics in the field of coding techniques.

We have designed very simple and reliable distance-preserving mappings construction based on graph theory. The new concept of the cube distance introduced in this paper has played a major role in algorithm of this construction without changing the theoretical properties of the cube graph.

The upper bound on the sum of the distances for all values of $M = 2^k$ has been reached with a very simple construction algorithm. The cube graph construction is considered the simplest and easiest construction to be used and to help understand the concept of the mapping technique. To our knowledge this also is the only construction that has emphasized and explained in a very simple manner the mechanism of a distance-reducing mapping.

The use of the selected graph theory properties helped in understanding certain properties of spectral null codes. The introduction of the index-permutation sequences and the use of the concept of distances gave us an idea about the structure and the design conditions of spectral null codes.

The new concepts of distances combined with graph theory properties will open up new research area as in computer since and networking where we can make use of the cube distance to change the complexity of the k-cube from spacial presentation to plan presentation using the same concept of Tanner graph as both share the bipartite property.

References

- 1. N. Deo, Graph Theory with Applications to Engineering and Computer Science. Dover Publications, 2016.
- J. Susymary and R. Lawrance, Graph theory analysis of protein-protein interaction graphs through clustering method, in 2017 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing, pp. 1–5, Srivilliputhur, India, March 23–25, 2017.
- 3. K. Ouahada and H. Ferreira, A similation and graph theoretical analysis of certain porperties of spectrall null codebooks, *SAIEE Africa Research Journal*, vol. 103, no. 3, pp. 106–115, 2012.
- K. Ouahada and H. C. Ferreira, A graph theoretic approach for spectral null codes, in *Proceedings of the Information Theory Workshop*, pp. 369–373, Taormina, Sicily, Italy, October 11-16, 2009.
- A. K. Singh, Error detection and correction by hamming codel, in Proceedings of the 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication, pp. 35–37, Bambhori, Jalgoan, India, December 22- 24, 2016.
- A. Wachter-Zeh, List decoding of insertions and deletions, *IEEE Transactions on Information Theory*, vol. PP, no. 99, p. 1, 2017.
- K. Ouahada, T. G. Swart, and H. C. Ferreira, Binary permutation sequences as subsets of levenshtein codes and higher order spectral nulls codes, in 2006 IEEE Information Theory Workshop - ITW '06 Chengdu, pp. 535–539, Chengdu, China, Oct. 2006.
- 8. A. Viterbi and J. Omura, *Principles of Digital Communication and Coding*. McGraw-Hill Kogakusha LTD, Tokyo Japan, 1979.
- A. J. H. Vinck and H. C. Ferreira, Permutation trellis codes, in in Proceedings of the International Symposium on Information Theory (ISIT 2001), p. 279, Washington, DC, USA, June 24-29, 2001.
- A. J. H. Vinck, Coded modulation for powerline communications, AEU International Journal of Electronics and Communications, vol. 54, no. 1, pp. 45–49, 2000.

- H. C. Ferreira, A. J. H. Vinck, T. G. Swart, and I. de Beer, Permutation trellis codes, *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 1782–1789, 2005.
- 12. T. G. Swart and H. C. Ferreira, A generalized upper bound and a multilevel construction for distancepreserving mappings, *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3685–3695, 2006.
- K. Ouahada and H. C. Ferreira, k-cube construction mappings from binary vectors to permutation sequences, in *in proceedings of IEEE International Symposium on Information Theory*, p. 279, Seoul, South Korea, June 28–July 3, 2009.
- W. J. Dally, Express cubes: Improving the performance of k-ary n-cube interconnection networks, IEEE Trans. On Computers, vol. 40, no. 9, pp. 1016–1023, Sept. 1991.
- M. Haynes, Magnetic recording techniques for buried servos, *IEEE Transactions on Magnetics*, vol. 17, no. 6, pp. 2730–2734, Nov. 198.
- H. C. Ferreira, I. de Beer, and A. J. H. Vinck, Distance preserving mappings onto convolutional codes revisited, in *Proceedings of the IEEE Information Theory Workshop*, pp. 23–26, Breisach, Germany, Apr. 26–29, 2002.
- 17. E. Gorog, Alphabets with desirable frequency spectrum properties, *IBM J. Res. Develop*, vol. 12, pp. 234–241, May 1968.
- K. Ouahada, T. G. Swart, and H. C. Ferreira, Spectral shaping permutation distance-preserving mappings codes, in *IEEE Proc. ITW'07 Conf. California*, USA, pp. 36–41, California, USA, Sept. 2Ö6, 2007.
- 19. K. A. S. Immink, *Codes for mass data storage systems*. Shannon Foundation Publishers, The Netherlands, 1999.
- K. A. S. Immink, Spectral null codes, *IEEE Transactions on Magnetics*, vol. 26, no. 2, pp. 1130–1135, ., Mar. 1990.
- C. Yeh and B. Parhami, 'parallel algorithms for index-permutation graphs. an extension of cayley graphs for multiple chip-multiprocessors (mcmp), in *International Conference on Parallel Processing*, pp. 3–12, California, USA, Sept. 2001.
- 22. L.-H. Chang, C. Wang, P.-N. Chen, Y. S. Han, and V. Y. F. Tan, Distance spectrum formula for the largest minimum hamming distance of finite-length binary block codes, in 2017 IEEE Information Theory Workshop ITW 2017 Kaohsiung, pp. 419–423, Kaohsiung, Taiwan, Nov. 6–10, 2017.