

A Distance-Based Approach for Binary-Categorical Data Bi-Clustering

Sadikin Mujiono

Abstract—Bi-clustering is the newest approach in dealing with multi-view clustering problems. The most common application of bi-clustering is in biological data set domain, even though there are many areas that can utilize the bi-clustering approach. The paper presents the description and the experimental results of another approach to performing bi-clustering task for binary categorical data. The approach is based on the Hamming Distance (HD) measurement of categorical data to extract a collection of bi-sets. The performance evaluation and analysis of the proposed method confirm that the approach has a wide opportunity to be applied in the area. The feasibility of the method is measured with the Match Score (MS) parameter in a bi-set extraction process for 96 experiments. The average of MS score values for the overlapping bi-set experiments is 0.869, whereas the minimum value is 0.815, and the maximum value is 0.984. In extracting disjoint bi-set, this method achieves a perfect performance with all MS values equal to 1.

Index Terms—bi-clustering, binary categorical data, bi-set, hamming distance.

I. INTRODUCTION

Bi-Clustering is the approach to overcoming a clustering problem by more than one view (variable). The most common application of Bi-Clustering methods and their variations are in the biological domain field to extract the corresponding of gene and its expression condition to some characteristics or a certain function such pathway or other biological function [1]–[8]. Bi-clustering threats data to be clustered in matrix representation and its clustering is done in two ways: rows and column. As a conventional clustering, there are some bi-clustering approaches to measure the similarities between more clusters have been proposed. Some of those common approaches are constant value, constant value on its rows or its column, coherence values, and evolution coherence values. Whereas some of the proposed bi-clustering algorithms are: bimax, CC bi-clustering, ISA, Plaid, Quest, Spectral, sv4d, Xmotifs, and CTWC as described in [2]. The other methods are Bimax proposed in [5] and Bitmap Searching in [9]. Both methods are used to find a collection of Bi-Set.

Categorical attributes are the sample's attributes that its values are specifically based on a certain category. One of the differences with the numerical attribute values is that

those values are not following a certain sequence. So the categorical attributes value cannot be sorted. Some samples of categorical values are the geometric shape such as a circle, rectangular, or diamond; human sex male or female, etc. Due some constraints and those differences, the algorithms that traditionally applied to numerical values cannot be directly applied to categorical attribute values [10]–[14].

The paper presents the results study of the utilization of total categorical attribute values distances to perform bi-clustering. The distance measurement that is used in this study is Hamming Distance, which is also used in [15]. The objective of the proposed bi-clustering methods is to extract bi-set contained in binary categorical matrix data. The detail explanation of this method will be described in section 3.

The rest section of this paper is organized as follows: Section 3 describes the approach and methodology to overcome the challenges in bi-set extraction. Section 4 will describe the experiments scenario and the data set materials. In section 4, it also illustrates the evaluation mechanism. This study uses match score parameter, which is also used in some previous study in bi-clustering. Section 5 contains the explanation of the experiment results and its analysis, whereas section 6 explains the achievement, the prospects, and the shortcoming of this study. Several potential explorations in the future study are also discussed in the last section.

II. RELATED STUDY

There are many proposed methods regarding bi-clustering and categorical data clustering as well. Some of them are presented in this section. Doruk B. et. al. [16] applied CC, HARP, and CPB, to evaluate the data set of yeast, mouse, and Drosophila gene. Among of those algorithms, CPB provides the best result in the capabilities to identify the similarities of local pattern and the finding of pattern shifting and scaling simultaneously. The improvement of the basic bi-clustering algorithm is performed by Wei Shen et. al. [3], and Mihee Lee et. al. [7]. In [2], the residue of CC bi-clustering algorithm is improved with the usage of the GA (Genetic Algorithm) model. In [3], it is performed a dimensionality reduction of the high-dimension matrix with SSVD (Sparse Singular Value Decomposition) method. Bi-clustering applied to Boolean categorical is also proposed in [17].

Limbo, which is proposed by Andritsos et. al. [18], is one of some methods to cluster categorical data set. This algorithm is based on hierarchical clustering algorithm that uses a framework called Information Bottleneck (IB) to define the distance measurement of a categorical tuple.

Manuscript received July 1, 2015.

Sadikin Mujiono is a faculty member of Universitas Mercu Buana, Jakarta, Jl. Meruya Selatan No. 1 Jakarta Barat. Currently he is a doctoral student at Universitas Indonesia. (Corresponding author to provide phone: +62-8111566600; fax: +62-021-5840813; e-mail: mujiono.sadikin@mercubuana.ac.id).

ROCK (A Robust Clustering Algorithm for Categorical Attributes) is the other methods developed by S. Guha, R. Rastogi, and K. Shim [12]. This method is based on agglomerative clustering. ROCK algorithm is started with the allocation of each tuple data set as a separate cluster; then those tuples are joined iteratively based on their closed relative distances. The distance is measured by the predefined between tuple correlations. CACTUS, as is presented in [19], is the summarization based bi-clustering algorithm. The phases of CACTUS consist of three steps: summarization, clustering, and validation. In the summarization step, it is performed a computation to summarize the information contained in the data set. The summarized results then are used to find cluster candidates in the clustering step. In the validation step, finally, it is defined the actual clusters based on previous candidate clusters. The framework of a bi-clustering of categorical data set based on K-Means is proposed in [17]. The other methods for categorical data clustering (not bi-clustering), which is based on Hamming Distance, have also been studied by [15].

Several methods that utilize the Bi-Clustering approach to extract bi-set collection have been proposed. Two of them are *Bit-Pattern Bi-clustering Algorithm (BiBit)* [20] and *Bimax* [5]. BiBit approach works based on the similarities of a binary number of binary matrices. Each pair of rows is used to perform the initial bi-set. If this initial bi-set is not the member of previously found bi-set, then the initial bi-set is included as their member. Afterward, the initial bi-set is used as a comparison of the rest rows of the matrix. Before the comparison is performed, it is needed to encode the binary to a decimal number, So, there are two phases of BiBit approach: encoding and searching. BiBit approach also requests minimum rows and minimum column as input parameters. The Bi-Set extraction algorithm, *Bimax*, is the most famous algorithm in this area and has become a model reference in comparison analysis [21], [9], [22], and [23]. In the beginning, *Bimax* algorithm is proposed to extract Bi-Set contained in gene expression data. A bi-cluster / Bi-Set (G, C) correlation of any subset of gene $G \subseteq \{1, \dots, n\}$ which collectively correspondences to any subset of samples $C \subseteq \{1, \dots, m\}$. The other words, a couple of (G, C) is a sub matrix E with all of those value element is 1 (one), and each e_{ij} element with its value is 1 represents one bi-cluster. *Bimax* algorithm works based on *divide and conquer* mechanism.

III. APPROACH AND METHODOLOGY

A. Objective

The objective of the research is to extract a collection of bi-set contained in binary matrix dataset. The bi-set is as representative of a group of samples which are simultaneously influenced by a group of attributes, or, in other words, a group of sample which correspondence each other based on a group of attributes. Some examples of this case are: a group of samples based on a group of gene expression, a group of drug which interacts each other based on a group of their compound, a group of buyer and seller that interacts each other based on a group of product or services, etc. To achieve this objective, the approach, first of

all, defines a based distance reference. Based on the based distance reference, then it is computed the actual distance of all of the rows in the matrix to this distance reference. The rows member afterward, are grouped based on the actual distance.

B. Framework

a. Categorical Data Sample Space

Referred to Peng Zhang [15], Adendberg stated that there are only two values can be used to specify the differences between two values in categorical sample space, i.e. equal or unequal. Therefore if there are two samples x_1 and x_2 then the difference measurement between both of these samples in j^{th} attribute is $x_{1j} = x_{2j}$ XOR (exclusive or) $x_{1j} \neq x_{2j}$. An instance of such attributes, for example, is a gender which its value can be *male* or *female* only.

In general, if there are a certain p as the number of sample attributes and there is j^{th} attribute with $1 \leq j \leq p$, the sample space is categorized in m_j which is a subset of attribute level $A_j = \{a_{j1}, \dots, a_{jm}\}$, then categorical data sample space is defined as:

$\Omega = A_1 \otimes A_2 \dots \otimes A_p$, with a_{ji} is stated as state or modality.

Categorical data set is:

$$X = (x_1, x_2, \dots, x_n)^T, x_i = (a_{i1}, a_{i2}, \dots, a_{ip}), i = 1, 2, \dots, n;$$

n = number of samples,

p = number of attributes of each sample

The categorical data which each of its attribute value is one of two $\{0, 1\}$ or $\{T, F\}$, is defined as binary categorical data set, and $A_j \in \{0, 1\}$.

b. Hamming Distance

The Hamming Distance (HD) of binary categorical data [15]

$$d(x_{iq}, x_{jq}) = \begin{cases} 0, & x_{iq} = x_{jq} \\ 1, & x_{iq} \neq x_{jq} \end{cases} \quad (1)$$

$$d_{ij} = dx_{ij} = \sum_{q=1}^{q=p} d(x_{iq}, x_{jq}), 0 \leq d_{ij} \leq p \quad (2)$$

Definition (1)

$C = (c_1, c_2, \dots, c_p)$, $c_j \in \{0, 1\}$, and $C \in X$ is the center of the X sample if:

$$\sum d(C, X) = \operatorname{argmin}(s \in X) \sum d(s, X) \quad (3)$$

The center is a certain element which has the most equivalence of its attributes with all of the rest elements in X sample space, therefore c_q will be more frequent in q position for all of x_i , then $d(c_q, x_{iq})$ will be minimum, and the its consequence is $\sum d(C, X)$ will be minimum also.

Definition (2)

- $R \in X$ is a distance based reference of all samples element $x_i \in X$, $i \in \{1, 2, \dots, n\}$
- $D = (0, 1, 2, \dots, p)$ is the possible distance vector of all $s_i \in \Omega$ to an arbitrary R
- $d = (d_0, d_1, d_2, \dots, d_q) \subseteq D$ is the actual distance of all samples $x_i \in X$ to an arbitrary R , and $d_0 < d_1 < d_2 < \dots < d_q$

If it is defined that $R = C$, and

$Vd_i \subseteq X$ is the samples collection $x_i \in X$ which their distance to R is d_i , $i \in \{1, 2, \dots, n\}$, $t \in \{1, 2, \dots, p\}$, then

1. Since all of x_i 's distance to R , $x_i \in Vd_i$, is equal, then all $x_i \in Vd_i$, is more similar each other than the other

samples $x_k \notin Vd_i$, or than the other samples x_k which their distance to R does not equal to x_i 's distance to R.

- Based on no. 1, then the quantity of cluster candidates=the number of $d_i=|d_i|$, where $Vd_i \neq \{\}$ and $\max(|d_i|)=p$

c. *Extreme Condition Weighting*

Let $R = (0,0,0,0,0)$, and $x_i = (1,1,0,0,0)$ and $x_k = (0,0,1,1,0)$, then by applying the rule of Hamming Distance based computation as presented in the previous section, then the computational will provide the same results i.e. $d(R, x_i) = d(R, x_k)$ and the final result is $x_i \sim x_k$ (x_i similar with x_k). This is not true, since both of the samples are different. In this study, this situation is handled with *weighting* adjustment. The weighting adjustment is performed by considering the attributes position. The most left attributes are given the most weight. With the weighting adjustment the formula of Hamming Distance is modified as:

$$d(x_{iq}, x_{jq}) = \begin{cases} 0 + 1/(q + 1), & x_{iq} = x_{jq} \\ 1, & x_{iq} \neq x_{jq} \end{cases} \quad (4)$$

d. *Binary Bi-Clustering Objective*

Definition (3)

$S = \{s_1, s_2, \dots, s_m\}$, sample set

$G = \{g_1, g_2, \dots, g_n\}$, attribute/properties set

The Objective of Bi-Clustering:

To extract a collection (K) Bi-Set $B = \{b_1, b_2, \dots, b_K\}$, $b_i = (S_i, G_i)$, $S_i \subseteq S$, $G_i \subseteq G$ is a set of sample s_i which are influenced by/contain a set of property G_i . b_i is defined as Boolean vector:

$$\langle s_i \rangle, \langle g_i \rangle = \langle s_{i1}, s_{i2}, \dots, s_{im} \rangle, \langle g_{i1}, g_{i2}, \dots, g_{in} \rangle,$$

$$s_{ik} = 1, \text{ if } s_k \in S_i, 0 \text{ else}$$

$$g_{ik} = 1, \text{ if } g_k \in G_i, 0 \text{ else}$$

The Fig. 1 and its description below explain the bi-clustering objective.

	g1	g2	g3	g4	g5
s1	1	1	0	1	0
s2	0	0	1	0	1
s3	0	1	1	1	0
s4	0	1	1	0	1
s5	0	1	1	0	1
s6	1	1	0	1	0

Fig.1. Binary Matrix as representation of samples and their attributes

If the binary matrix input as Fig. 1, then some of the bi-sets as the results of a *Bi-Clustering* process are as follow:

$$b_1 = \langle s_4, s_5 \rangle \langle g_2, g_3, g_5 \rangle = \langle 0, 0, 0, 1, 1, 0 \rangle \langle 0, 1, 1, 0, 1 \rangle$$

$$b_2 = \langle s_1, s_6 \rangle \langle g_1, g_2, g_4 \rangle = \langle 1, 0, 0, 0, 0, 1 \rangle \langle 1, 1, 0, 1, 0 \rangle$$

....

$$b_K = \dots$$

IV. EXPERIMENTS SCENARIO

A. *HD Application for Bi-Clustering*

In the experiment that is performed in this study, we apply our proposed HD based framework to cluster rows and columns simultaneously to perform a bi-clustering. Each segment as the intersection of row clusters and column clusters than is threatened as a bi-set candidates.

B. *Evaluation*

The performance of this proposed method is evaluated by the using of a *Match Score/MS*. The MS parameter was also previously used by Prelic et.al and Domingo [5], [20]. This next section explains the MS parameter.

Definition (4)

Referred from [5], MS is used as performance evaluation parameters of bi-clustering algorithm that is applied to gene representation of binary matrix.

Let, $G_1, G_2 \subseteq \{1, \dots, n\}$ are two sets of gene. The MS of G_1 and G_2 is defined as:

$$SG(G_1, G_2) = |G_1 \cap G_2| / |G_1 \cup G_2| \quad (5)$$

This SG presents the characteristics of the correspondence between two sets of gene. The MS value will be symmetrical, which mean is $SG(G_1, G_2) = SG(G_2, G_1)$, and will have value between zero and one. The MS = 0 if both of gene sets are disjoint, and will be 1 if those gene sets are identical. The MS in this study is modified by to include its attributes (column of the matrix) and the modified formula is:

Let there are two bi-sets $b_1 = (S_1, G_1)$, $S_1 \subseteq S$, $G_1 \subseteq G$ and $b_2 = (S_2, G_2)$, $S_2 \subseteq S$, $G_2 \subseteq G$, The *Match Score / MS* b_1 subject to b_2 is

$$Sc(b_1, b_2) = |b_1 \cap b_2| / |b_1 \cup b_2|$$

$$= \frac{1}{2} \{ (|\langle s_{11}, \dots, s_{1m} \rangle \cap \langle s_{21}, \dots, s_{2m} \rangle| / |\langle s_{11}, \dots, s_{1m} \rangle \cup \langle s_{21}, \dots, s_{2m} \rangle|) + (|\langle g_{11}, \dots, g_{1n} \rangle \cap \langle g_{21}, \dots, g_{2n} \rangle| / |\langle g_{11}, \dots, g_{1n} \rangle \cup \langle g_{21}, \dots, g_{2n} \rangle|) \}$$

The MS of two bi-set sets B_1, B_2 , *MS* B_1 subject to B_2

$$Sc(B_1, B_2) = \frac{\sum_{(b_i \in B_1) \text{ argmax } (b_j \in B_2) (Sc(b_i, b_j))}{|B_1|} \quad (6)$$

C. *Data Set*

It is used synthetic data set to validate the proposed distance based bi-clustering method. The synthetic data set is adapted to and modified from data set was used by [20]. The bi-set contained in the binary matrix as an input, is implanted manually. The characteristics of the data sets are classified into two kinds based on the implanted bi-set, i.e. disjoint and overlapping. The variable to identify those both of the data set characteristics is matrix size, and its bi-set size, whereas the overlapping data set characteristics are matrix size, maximum main bi-set size and one additional variable i.e. the maximum of overlapping size in rows and columns. Fig. 2 and Fig. 3 illustrates those kind of data sets. Fig. 2 is the disjoint bi-set data set with binary matrix size = 20 x 20, bi-set size = 5 x 5, whereas the Fig. 3 is the overlapping bi-set data set with matrix size 26 x 26, maximum bi-set size = 6x6 and maximum overlapping size in rows and columns is 1 respectively. The maximum bi-set is the main bi-set plus the maximum overlapping size. For

example, in Fig. 3, the main bi-set is 5 and the maximum overlapping size is 1.

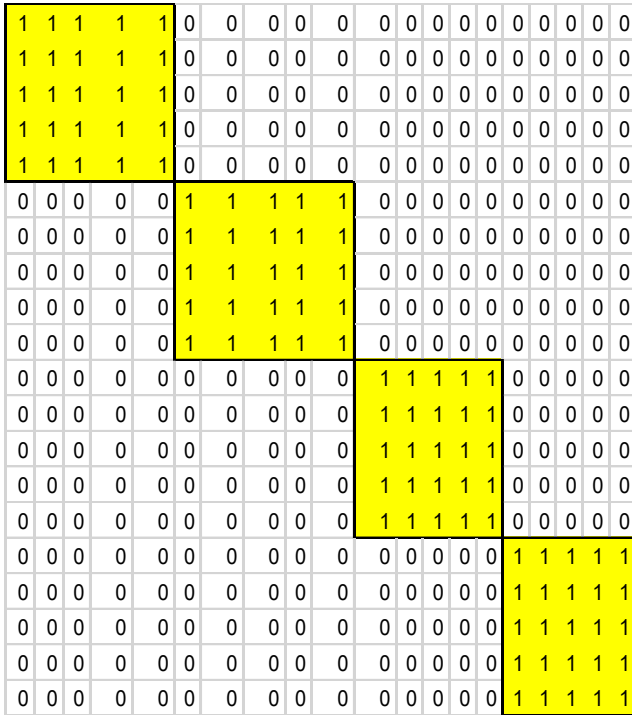


Fig.2. The disjoint bi-set data set

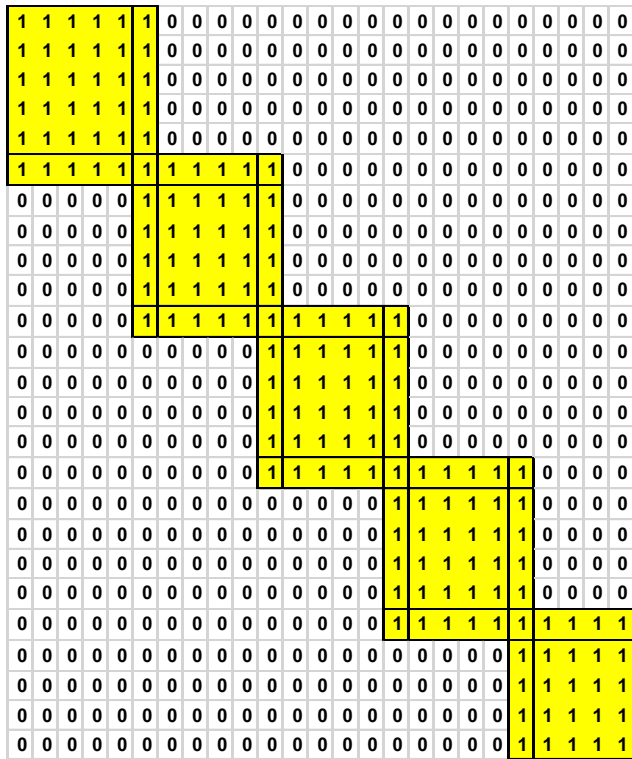


Fig.3. The overlapping bi-set data set

In the experiment execution, we use six disjoint data sets which their matrix size is 50 x 50 and 100 x 100 with the bi-set size ranges from 5 to 25. While for the overlapping data set we use 18 variants based on: matrix size, maximum bi-set size, and maximum overlapping size. Each variant of the data set is generated 5 derivative randomly. Therefore there are 90 overlapping data set members

V. RESULTS AND ANALYSIS

A. Disjoint Data Set

The performance of this proposed method that is applied to the disjoint data set is perfect. For all of the data set member, its MS is one. That means that all of implanted bi-set is extracted with 100 % accuracy. The performance is presented as Fig. 4 as follows.

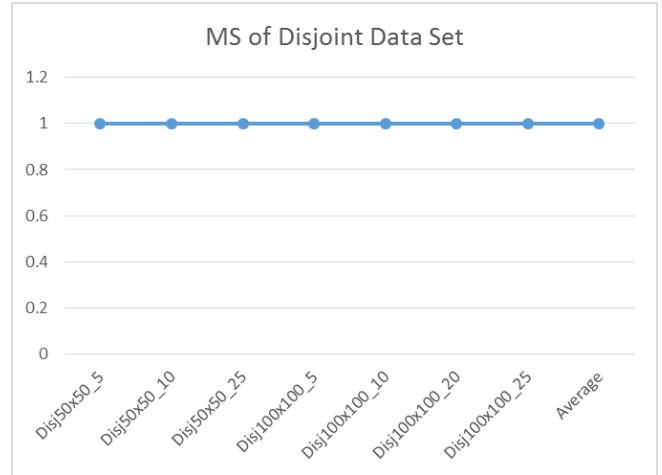


Fig.4. The MS of disjoint data set

B. Overlapping Data Set

The performance of the various overlapping bi-set data set based on its matrix size are presented in figures as follows. For each kind of data set, it is generated 5 data sets randomly and in these graphs we present the MS average of the 5 data sets. The data set id $S_x\text{-}C_m\text{-}O_n$ means that the matrix size is $x * x$, its main bi-set size is $m * m$, and its maximum size of overlapping in rows and in columns is n . Overall, the performance of the proposed method to overlapping bi-set data set is good since the minimum MS is 0.814968116137707 which is achieved for data set ID S100_C20_O18. While of 90 experiments, the achieved maximum MS is 0.983735380116959 for data set ID S60_C20_O18 in experiment number 2. The MS average of all 90 experiments is 0.869736601.

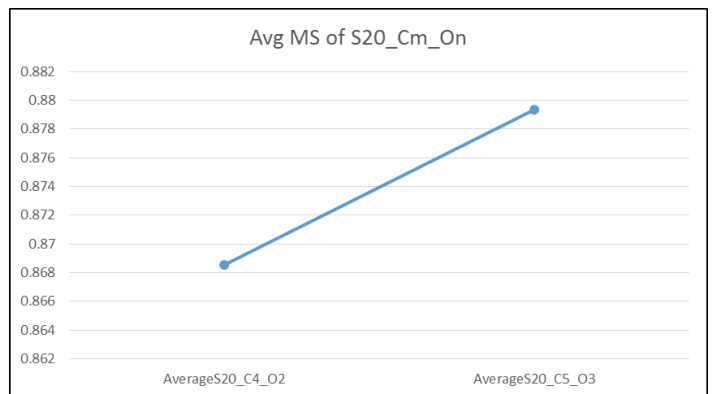


Fig.5. The performance for data set ID S20_Cm_On

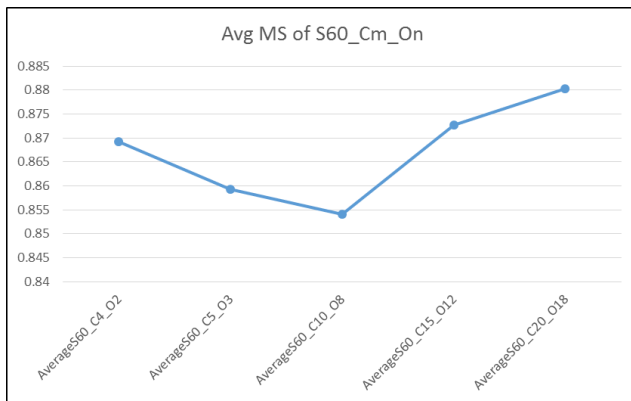


Fig.6. The performance for data set ID S60_Cm_On

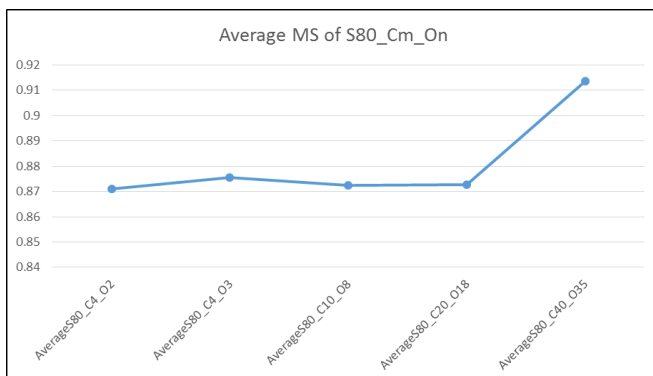


Fig.7. The performance for data set ID S80_Cm_On

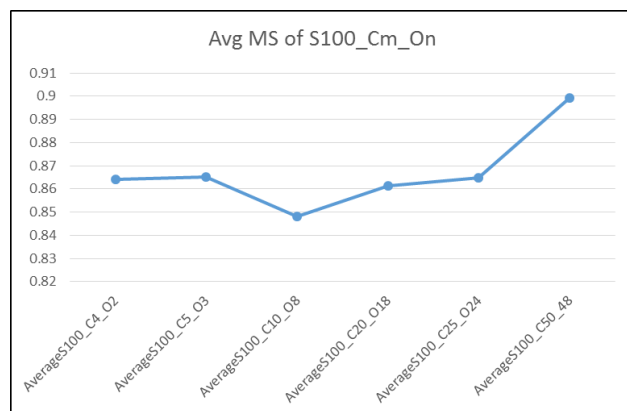


Fig.8. The performance for data set ID S100_Cm_On

As presented in the four figures of performance to overlapping bi-set, there are similar trends in the performance. The MS value tends to better (with higher MS value) for the bigger bi-set size and bigger overlapping size.

VI. CONCLUSION & FUTURE WORKS

The paper presented the results study and the validation of another approach based on HD measurement to extract bi-set contained in a binary matrix. The evaluation of the performances showed that the proposed method is feasible to be explored in the future with some improvements and modifications. The achieved minimum MS of the overlapping bi-set data set was 0.815 which was good enough. Although this proposed approach provided a good MS value, but it also had some drawbacks. One of these drawbacks was the consumed time when this method had to handle an input matrix which its size is more than 100 x

100. The most time consuming was the phase in finding a certain sample or attribute that will be used as a base reference.

In future studies, we will explore several potential improvements and modifications such: the improvement of bi-set finding algorithm in each segment of bi-cluster, the possibilities to create new bigger bi-clusters based on the initial clusters. There are also many challenges related to the comparison with the others approach such Bibit and Bimax. It will also be elaborated the other alternatives in determining of the base reference to overcome the time consuming problem. The other challenges are the application of the proposed method to experimental (real) data set. Some experimental data sets are drug-compound interactions, gene expression or even seller-buyer interactions. To illustrate, in the drug-drug interaction (DDI) applications, the representation of the DDI consists of drug samples represented as rows and compound as columns. The matrix value will be one if a certain drug contains such certain of compounds, and the value will be zero otherwise.

REFERENCES

- [1] J. Gu and J. S. Liu, "Bayesian biclustering of gene expression data.," *BMC Genomics*, vol. 9 Suppl 1, p. S4, Jan. 2008.
- [2] S. Kaiser, "Biclustering: Methods, Software and Application," Ludwig-Maximilians, Munchen, 2011.
- [3] W. Shen, G. Liu, M. Zheng, Z. Li, and Y. Zhong, "A Novel Biclustering Algorithm and Its Application in Gene Expression Profiles *," *J. Inf. Comput. Sci.*, vol. 9, no. 11, pp. 3113–3121, 2012.
- [4] C. Cano, L. Adarve, J. López, and a. Blanco, "Possibilistic approach for biclustering microarray data," *Comput. Biol. Med.*, vol. 37, pp. 1426–1436, 2007.
- [5] B. Peter, A. Prelić, S. Bleuler, P. Zimmermann, A. Wille, P. Bühlmann, W. Gruissem, L. Hennig, L. Thiele, E. Zitzler, A. Prelic, A. Wille, and P. B., "Comparison of Biclustering Methods: A Systematic Comparison and Evaluation of Biclustering Methods for Gene Expression Data," *Bioinformatics*, vol. 22, no. 9, pp. 1122–112930, 2006.
- [6] Y. Cheng and G. M. Church., "Biclustering of expression data," in *Proc. of the International Conference on Intelligent Systems for Molecular Biology*, 2000, pp. 93–103.
- [7] M. Lee, H. Shen, J. Z. Huang, and J. S. Marron, "Biclustering via sparse singular value decomposition.," *Biometrics*, vol. 66, no. 4, pp. 1087–95, Dec. 2010.
- [8] F. Angiulli, E. Cesario, and C. Pizzuti, "Random walk biclustering for microarray data," *Inf. Sci. (Ny)*, vol. 178, pp. 1479–1497, 2008.
- [9] D. S. Rodriguez-baena, A. J. Perez-pulido, and S. Jesus, "A biclustering algorithm for extracting bit – patterns from binary datasets," pp. 1–8, 2011.
- [10] A. Ahmad and L. Dey, "A k-mean clustering algorithm for mixed numeric and categorical data," *Data Knowl. Eng.*, vol. 63, pp. 503–527, 2007.
- [11] H.-J. Do and J. Y. Kim, "Clustering categorical data based on combinations of attribute values," *Int. J. Innov. Comput. Inf. Control*, vol. 5, no. 12, pp. 4393–4405, 2009.
- [12] S. Guha and M. Hill, "ROCK: A Robust Clustering Algorithm for Categorical Attributes," pp. 1–25.
- [13] Z. He, X. Xu, and S. Deng, "Clustering Mixed Numeric and Categorical Data: A Cluster Ensemble Approach," p. 14, 2005.
- [14] S. S. Mesakar and M. S. Chaudhari, "Review Paper On Data Clustering Of Categorical Data," *Int. J. Eng. Res. Technol.*, vol. 1, no. 10, pp. 1–3, 2012.
- [15] P. Zhang, X. Wang, and P. X.-K. Song, "Clustering Categorical Data Based on Distance Vectors," *J. Am. Stat. Assoc.*, vol. 101, no. 473, pp. 355–367, 2006.
- [16] D. Bozdağ, A. S. Kumar, and U. V. Catalyurek, "Comparative analysis of biclustering algorithms," *Proc. First ACM Int. Conf. Bioinforma. Comput. Biol. - BCB '10*, p. 265, 2010.
- [17] R. G. Pensa, "A Bi-clustering Framework for Categorical Data," in *In Proceedings of the 9th European Conference on Principles and*

- Practice of Knowledge Discovery in Databases (PKDD)*, 2005, pp. 643–650.
- [18] P. Andritsos, P. Tsaparas, J. Miller, and K. C. Sevcik, “LIMBO : Scalable Clustering of Categorical Data.”
 - [19] V. Ganti, J. Gehrke, and R. Ramakrishnan, “CACTUS---clustering categorical data using summaries,” *Proc. fifth ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD '99*, pp. 73–83, 1999.
 - [20] D. S. Rodriguez-Baena, A. J. Perez-Pulido, and J. S. Aguilar-Ruiz, “A biclustering algorithm for extracting bit-patterns from binary datasets,” *Bioinformatics*, vol. 27, no. 19, pp. 2738–2745, 2011.
 - [21] J. Gupta, S. Singh, and N. Verma, “MTBA: MATLAB Toolbox for Biclustering Analysis,” in *MTBA: MATLAB Toolbox for Biclustering Analysis*, 2013, vol. 1, no. July, pp. 94–97.
 - [22] H. C. Chen, W. Zou, Y. J. Tien, and J. J. Chen, “Identification of Bicluster Regions in a Binary Matrix and Its Applications,” *PLoS One*, vol. 8, no. 8, 2013.
 - [23] K. Benabdeslem and K. Allab, “Bi-clustering continuous data with self-organizing map,” *Neural Comput. Appl.*, vol. 22, pp. 1551–1562, 2013.

Mujiono Sadikin was born in Magetan, East Java, Indonesia, December 6th 1970. He holds a Bachelor degree in Informatics from Bandung Institute of Technology, Bandung, Indonesia and master degree in Informatics from Bandung Institute of Technology, Bandung, Indonesia. Currently, he is a doctoral student in Computer Science, University of Indonesia. His focus area study covers machine learning, data mining and bioinformatics.