Classification systems based on combinatorial semigroups

J. Abawajy[1], A.V. Kelarev[1,2]

[1]School of Information Technology, Deakin University
221 Burwood Highway, Burwood, Victoria 3125, Australia
jemal.abawajy@deakin.edu.au, kelarev@deakin.edu.au

[2]School of Science, Information Technology and Engineering
University of Ballarat, P.O. Box 663, Ballarat, Victoria 3353, Australia
a.kelarev@ballarat.edu.au

Abstract

The present article continues the investigation of constructions essential for applications of combinatorial semigroups to the design of multiple classification systems in data mining. Our main theorem gives a complete description of all optimal classification systems defined by one-sided ideals in a construction based on combinatorial Rees matrix semigroups. It strengthens and generalizes previous results, which handled the more narrow case of two-sided ideals.

1   Introduction

The present article strengthens and generalizes previous results devoted to the investigation of constructions of multiple classification systems, or classifiers, combining several binary classifiers and based on combinatorial Rees matrix semigroups and two-sided ideals [15]. Our main theorem describes all optimal classification systems defined by one-sided ideals in the corresponding construction. Since the class of one-sided ideals is larger than that of two-sided ideals, considering this more general case may lead to the design of better classification systems. It is also nice that the new formula given by our new main theorem involves fewer terms and may turn out easier to use compared to the previous formula obtained in the more narrow case handled previously.

The design of multiple classifiers combining several binary classifiers is well known in the literature and has valuable applications in data mining (see, for instance, [17, 18]). Recall that a classifier is said to be *binary* if it divides all data into two classes. Every binary classifier operates by producing the labels of its classes as outputs. The set of all outputs of each binary classifier can be regarded as the field $GF(2)$ of order two, which will be denoted by F throughout. This means that the set of the outputs of n binary classifiers can be considered as the vector space F^n.

Rees matrix semigroups can be used in order to generate convenient sets of centroids for centroid-based classifiers and to design combined multiple classifiers capable of correcting the errors of individual binary clusters. These applications are illustrated in Figure 1.
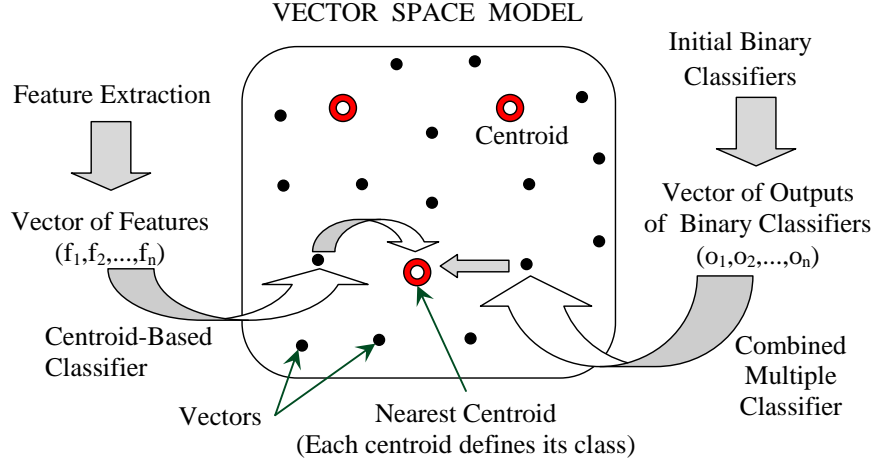
Figure 1: Centroid-based classifiers and combined multiple classifiers.

The classification process using binary classifiers begins with feature extraction and representation of data in a vector space F^n, where n is a positive integer. Every centroid-based classifier selects special elements $c_1, \ldots, c_k$ in F^n, called *centroids*. For $i = 1, \ldots, k$, each centroid c_i defines its class $N(c_i)$ consisting of all vectors v such that c_i is the nearest centroid of v. Every vector is assigned to the class of its nearest centroid.

On the other hand, multiple classifiers and classifiers are often used in analysis of data to combine individual initial classifiers or classifiers. A well-known method for the design of multiple classifiers consists in designing several simpler initial or individual classifiers, and then combining them into one multiple classification scheme with several classes. This method is very effective, and is often recommended for various applications, see [17], Section 7.5. The main advantage of using combined multiple classifiers is in their ability to correct errors of individual classifiers and produce correct classifications despite individual classification errors. Let us refer to [10, 12, 14, 19] for more details.

Denote the number of initial binary classifiers being combined by n. As noted above, the set of the outputs of n binary classifiers can be considered as the vector space F^n. If $o_1, \ldots, o_n$ are the outputs of the initial binary classifiers, then the sequence $(o_1, \ldots, o_n)$ is called a *vector of outputs* of the initial classifiers. In order to define the multiple classifier and enable correction of errors of the initial classifiers, a set of centroids $c_1, \ldots, c_k$ is again selected in F^n. For $i = 1, \ldots, k$, the class $N(c_i)$ of the centroid c_i is again defined as the set of all observations with the vector outputs of the initial classifiers having c_i as its nearest centroid.

In order to introduce convenient sets of generators for classification systems with outputs of initial classifiers represented as vectors in F^n, further we assume that S is a finite semigroup and the number m of the binary classifiers being combined does not exceed the number of nonzero elements of S. Denote by S^* the set of all nonzero elements of S. Since we can extend the vector space F^n to $F^{|S^*|}$ by adding initial binary classifiers, without loss of generality we may assume throughout that the number of nonzero elements of S is equal to n. Further, zeros of semigroups

are denoted by θ, and 0 stands for the zero of a ring.

We refer to the monographs [5, 7, 8] for preliminaries and standard notation used in this research direction. The *semigroup ring* $F[S]$ is the set of all finite sums

$$F[S] = \left\{ \sum_{i=1}^{m} f_i s_i \;\middle|\; f_i \in F, s_i \in S \right\} \tag{1}$$

with addition and multiplication defined by the rules

$$\sum_{s=S} f_s s + \sum_{s \in S} h_s s = \sum_{s \in S} (f_s + h_s)s, \tag{2}$$

$$\left(\sum_{s \in S} f_s s \right) \left(\sum_{t \in S} h_t t \right) = \sum_{s,t \in S} (f_s h_t)st, \tag{3}$$

where $f_s, h_s, h_t \in F$ and $s, t \in S$. It is natural to consider S as being embedded in $F[S]$ by identifying each element s of S with $1s$ in $F[S]$. If S has a zero θ, then the *contracted semigroup ring* $F_0[S]$ is defined as the quotient ring of $F[S]$ modulo its ideal $F\theta$. If S has no zero, then the semigroup ring $F[S]$ is isomorphic to the contracted semigroup ring $F_0[S^0]$, where $S^0 = S \cup \{\theta\}$. In this case we also write $F_0[S]$ for $F[S]$. Thus, we always have

$$F_0[S] = \{0\} \cup \left\{ \sum_{i=1}^{n} f_i s_i \;\middle|\; 0 \neq f_i \in F, \; \theta \neq s_i \in S \right\}. \tag{4}$$

These constructions are standard, well known and have been used in various research directions. Let us refer to the book [7] and survey [11] for a comprehensive bibliography devoted to their applications in coding theory, and to [1], [2], [3], [9] for examples of more recent results. If $\{s_1, \ldots, s_n\}$ is the set of nonzero elements of S, then we can identify the vector space F^n with $F_0[S]$ by identifying $r = (r_1, \ldots, r_n) \in F^n$ with $\sum_{i=1}^{m} r_i s_i \in F_0[S]$. Thereby the set $F^n = F_0[S]$ has been endowed with multiplication.

For the design of efficient multiple classifiers it is essential to find sets with large minimum distance, see [13] and [17], Section 7.5. The *weight* $\mathrm{wt}(r)$ of $r = \sum_{s \in S} r_s s \in F_0[S]$ is the number of nonzero coefficients r_s, for $s \in S$. The weight of a subset $C \subseteq F^n$ is the minimal weight of a nonzero element in C. The *minimum distance* of C is the smallest weight of a nonzero difference $u - v$ for $u, v \in C$. If C is a linear space, then it is well known and easy to verify that the minimum distance of C coincides with its weight.

The concepts of ideals and one-sided ideals of a ring or an algebra over a field are well known. It is also well known and easy to verify that in the case where F is a field of order two, these notions defined for rings coincide with their counterparts defined in F-algebras. For convenience of the readers, let us recall the definitions of ideals and one-sided ideals here. Pick any elements $g_1, \ldots, g_k \in F^n$. The set

$$C(g_1, \ldots, g_k) \quad = \quad F^n g_1 F^n + \cdots + F^n g_k F^n \tag{5}$$

$$= \quad \left\{ \sum_{j=1}^{m_1} \ell_{1,j} g_1 r_{1,j} + \cdots + \sum_{j=1}^{m_k} \ell_{k,j} g_k r_{k,j} \;\middle|\; \ell_{i,j}, r_{i,j} \in F^n \cup F \right\}$$

is called the *ideal*, or *two-sided ideal*, generated by the $g_1, \ldots, g_k$. The sets

$$C_r(g_1, \ldots, g_k) \quad = \quad g_1 F^n + \cdots + g_k F^n \tag{6}$$

$$= \quad \left\{ \sum_{j=1}^{m_1} g_1 r_{1,j} + \cdots + \sum_{j=1}^{m_k} g_k r_{k,j} \;\middle|\; r_{i,j} \in F^n \cup F \right\},$$

$$C_\ell(g_1, \ldots, g_k) \quad = \quad F^n g_1 + \cdots + F^n g_k \tag{7}$$

$$= \quad \left\{ \sum_{j=1}^{m_1} \ell_{1,j} g_1 + \cdots + \sum_{j=1}^{m_k} \ell_{k,j} g_k \;\middle|\; \ell_{i,j} \in F^n \cup F \right\}$$

are called the *right ideal* generated by $g_1, \ldots, g_k$ and, respectively the *left ideal* generated by $g_1, \ldots, g_k$. A *one-sided* ideal is a left or right ideal. The case of left ideals is dual to that of right ideals, which means that every assertion concerning right ideals comes together with an equivalent statement expressed in terms of left ideals with the order of all products interchanged. This is why we consider only right ideals in what follows.

2   Main Theorem

We use standard concepts and refer the readers to the monographs [5, 7, 17] for background information and preliminaries. Let G be a group, I and Λ nonempty sets, and let e be the identity of G. Denote by $G^0 = G \cup \{\theta\}$ the group G with zero θ adjoined. Let $P = [p_{\lambda i}]$ be a $(\Lambda \times I)$-matrix with entries $p_{\lambda i} \in G^0$, for all $\lambda \in \Lambda$, $i \in I$. The *Rees matrix semigroup* $M^0(G; I, \Lambda; P)$ *over* G with *sandwich-matrix* P is the set consisting of zero θ and all triples $(g; i, \lambda)$, for $i \in I$, $\lambda \in \Lambda$, $g \in G^0$, where all triples $(\theta; i, \lambda)$ are identified with θ, and multiplication is defined by the rule

$$(g_1; i_1, \lambda_1)(g_2; i_2, \lambda_2) = (g_1 p_{\lambda_1 i_2} g_2; i_1, \lambda_2), \tag{8}$$

for $g_1, g_2 \in G$, $i_1, i_2 \in I$, $\lambda_1, \lambda_2 \in \Lambda$. Rees matrix semigroups play crucial roles is semigroup theory. Since this article belongs to an intersection of several research directions and is intended also for the readers specializing in data mining, let us refer to [6, 16] for examples of recent results.

A semigroup S is said to be *combinatorial* if every subgroup of S is a singleton. Throughout we assume that $T = M^0(G; I, \Lambda; P)$ is a finite combinatorial Rees matrix semigroup with nonzero sandwich-matrix and S is a subsemigroup of T. Then it is well known and easy to verify that the group G is a singleton, i.e., $G = \{e\}$. Therefore we can use shorthand notation (i, λ) for each nonzero element $(e; i, \lambda)$ in $M^0(G; I, \Lambda; P)$, so that we have $T = (I \times \Lambda) \cup \{\theta\}$ and

$$(i_1, \lambda_1)(i_2, \lambda_2) = \begin{cases} (i_1, \lambda_2) & \text{if } p_{\lambda_1 i_2} \neq \theta \\ \theta & \text{otherwise,} \end{cases} \tag{9}$$

for $g_1, g_2 \in G$, $i_1, i_2 \in I$, $\lambda_1, \lambda_2 \in \Lambda$. Further, let Q be a subset of T. If $j \in I$, $\mu \in \Lambda$, then we use the following notation:

$$Q_{*\mu} = Q \cap \{(i, \mu) \mid i \in I\}, \tag{10}$$
$$Q_{j*} = Q \cap \{(j, \lambda) \mid \lambda \in \Lambda\}, \tag{11}$$
$$Q_{j\mu} = Q \cap \{(j, \mu)\}. \tag{12}$$

For any $X \subseteq I$, $Y \subseteq \Lambda$, let

$$Q_{X*} = \cup_{i \in X} Q_{i*}, \tag{13}$$
$$Q_{*Y} = \cup_{\lambda \in Y} Q_{*\lambda}, \tag{14}$$
$$Q_{XY} = \cup_{i \in X, \lambda \in Y} Q_{i\lambda}. \tag{15}$$

Let us define the subsets

$$I(S) = \{i \in I \mid S_{i*} \neq \emptyset\}, \tag{16}$$
$$\Lambda(S) = \{\lambda \in \Lambda \mid S_{*\lambda} \neq \emptyset\}, \tag{17}$$
$$L(S) = \{\lambda \in \Lambda(S) \mid (\forall i \in I(S)) \, p_{\lambda i} = \theta\}. \tag{18}$$

For any $i \in I$, denote by $E_i(S)$ the set of all subsets Y of $\Lambda(S) \setminus L(S)$ such that $\{i\} \times Y \subseteq S$ and besides, for each $j \in I(S)$, the set Y contains an even number of elements λ with $p_{\lambda j} \neq \theta$. Let $Z_i(S)$ be the set of all the subsets in $E_i(S)$ that have the largest cardinality in $E_i(S)$. Put $N_i(S) = |Y|$ for any $Y \in Z_i(S)$. This makes sense, since the cardinalities of all subsets in each $Z_i(S)$ coincide.

Further, denote by $E(S)$ the set of all unions of the form

$$\bigcup_{i \in I(S)} (\{i\} \times Y_i) \tag{19}$$

where $Y_i \in E_i(S)$ for all i. Likewise, let $Z(S)$ be the sets of all unions of the form (19), where $Y_i \in Z_i(S)$ for all i. We can let $N(S) = |Y|$ for any $Y \in Z(S)$, because the cardinalities of all subsets in $Z(S)$ are equal to each other. Clearly, $E(S)$ and $Z(S)$ are subsets of the power set of S, and $N(S) = \sum_{i \in I(S)} N_i(S)$.

Theorem 1. *Let $T = M^0(G; I, \Lambda; P)$ be a finite Rees matrix semigroup over the group $G = \{e\}$, and let S be a subsemigroup of T. Let $C_r(g_1, \ldots, g_k)$ be a right ideal with the largest weight in $F_0[S]$. Then the weight of $C_r(g_1, \ldots, g_k)$ is equal to*

$$\mathrm{wt}(C_r(g_1, \ldots, g_k)) = \max \left\{ |S_{*L(S)}| + N(S), \max_{\lambda \in \Lambda(S)} |S_{*\lambda}| \right\}. \tag{20}$$

Corollary 2. *Let $T = M^0(G; I, \Lambda; P)$ be a finite Rees matrix semigroup over the group $G = \{e\}$, and let S be a subsemigroup of T. Then the largest number of errors of binary classifiers, which can be corrected by a multiple classifier defined by the set of the form $C_r(g_1, \ldots, g_k)$ in $F_0[S]$ is equal to*

$$\max \left\{ \lfloor (|S_{*L(S)}| + N(S) - 1)/2 \rfloor, \max_{\lambda \in \Lambda(S)} \lfloor (|S_{*\lambda}| - 1)/2 \rfloor \right\}. \tag{21}$$

Remark 1. We can replace $\max_{\lambda \in \Lambda(S)}$ by $\max_{\lambda \in \Lambda(S) \setminus L(S)}$ in Theorem 1 and Corollary 2, because $\lambda \in L(S)$

implies $|S_{*\lambda}| \leq |S_{*L(S)}|$.

Remark 2. Since every finite subsemigroup of an infinite combinatorial Rees matrix semigroup T embeds in a finite Rees matrix subsemigroup of T, it is no increase to generality to shift the word "finite" from T to S in the hypothesis of Theorem 1.

Remark 3. It is well known and easy to verify that a Rees matrix semigroup $M^0(G; I, \Lambda; P)$ over a group and with nonzero sandwich-matrix is combinatorial if and only if the group G is a singleton. Since the case of zero sandwich-matrix is trivial, this means that Theorem 1 handles all combinatorial Rees matrix semigroups.

The following example shows that it is impossible to drop the condition that $G = \{e\}$ from the main theorem and corollary.

Example 1. Let $G = \{e, g\}$ be the group of order two, $I = \{i\}$, $\Lambda = \{\lambda_1, \lambda_2, \lambda_3\}$, $P = \begin{bmatrix} e \\ e \\ e \end{bmatrix}$,

$T = M^0(G; I, \Lambda; P)$, and let $S = T$. Then $L(S) = \emptyset$, $S_{*L(S)} = \emptyset$,

$$E_i(S) = \{\{\lambda_1, \lambda_2\}, \{\lambda_1, \lambda_3\}, \{\lambda_2, \lambda_3\}\},$$

and so $Z_i(S) = E_i(S)$, and $N_i(S) = 2$. Hence every subset in $E(S)$ has two elements, and so the same is true of $Z(S)$; whence $N(S) = 2$. Besides, $|S_{*\lambda}| = 2$ for all $\lambda \in \Lambda$. Therefore the right-hand side of equality (20) is equal to 2. However, it is easily seen that the element

$$g = (e; i, \lambda_1) + (e; i, \lambda_2) + (g; i, \lambda_2) + (g; i, \lambda_3)$$

generates a right ideal of weight 4 in $F_0[S]$, and so equality (21) does not hold in this example, since

$$\mathrm{wt}(C_r(g)) > \max\left\{ |S_{*L(S)}| + N(S), \max_{\lambda \in \Lambda(S)} |S_{*\lambda}| \right\}. \tag{22}$$

3 Proofs

To simplify notation, for any element s in S, we denote by the same letter s the element $1s$ of $F_0[S]$, too. In other words, we identify the elements of S with their images in $F_0[S]$. For any subset X in S, we put

$$g_X = \begin{cases} \sum_{x \in X \setminus \{\theta\}} x & \text{if } X \not\subseteq \{\theta\}, \\ 0 & \text{otherwise.} \end{cases}$$

Evidently, if X and Y are disjoint subsets of S, then $g_{X \cup Y} = g_X + g_Y$. For any

$$r = \sum_{s \in S} r_s s \in F_0[S],$$

the *support* of r is the set

$$\mathrm{supp}(r) = \{t \in T \mid r_t \neq 0\}. \tag{23}$$

Let us define the set

$$\mathrm{Ann}_\ell(F_0[S]) = \{x \in F_0[S] \mid xF_0[S] = 0\}. \tag{24}$$

For $j \in I$, $\mu \in \Lambda$, $X \subseteq S$, and any $r \in F_0[S]$, we put

$$r_j = \sum_{\lambda \in \Lambda} r_{(j,\lambda)} \in F_0[S], \tag{25}$$

$$r_\mu = \sum_{i \in I} r_{(i,\mu)} \in F_0[S], \tag{26}$$

$$r_X = \sum_{s \in X} r_s s \in F_0[S]. \tag{27}$$

Lemma 1. *Let* $T = M^0(G; I, \Lambda; P)$ *be a Rees matrix semigroup over the group* $G = \{e\}$*, and let* S *be a finite subsemigroup of* T*. Then an element*

$$r = \sum_{s \in S} r_s s \in F_0[S]$$

belongs to $\mathrm{Ann}_\ell(F_0[S])$ *if and only if* $\mathrm{supp}(r) \setminus S_{*L(S)} \in E(S)$.

Proof. Consider any element r of $F_0[S]$. Put $X = \mathrm{supp}(r)$, $X_0 = X \cap S_{*L(S)}$ and $X_1 = X \setminus X_0 = X \setminus S_{*L(S)}$, so that $r = g_{X_0} + g_{X_1}$. Since it follows from (18) that $g_{X_0} \in \mathrm{Ann}_\ell(F_0[S])$, our task is to show that $g_{X_1} \in \mathrm{Ann}_\ell(F_0[S])$ if and only if $X_1 \in E(S)$. Therefore, to simplify notation, further we may assume that $r = g_{X_1}$ and $X = X_1$.

The 'if' part. Suppose that $X \in E(S)$. In order to show that r belongs to $\mathrm{Ann}_\ell(F_0[S])$, we choose an arbitrary element $y = (j, \mu) \in S$, where $j \in I(S)$ and $\mu \in \Lambda(S)$, and we are going to verify that $ry = 0$. Since $g_{X_1} = \sum_{i \in I} g_{(X_1 \cap S_{i*})}$, it suffices to demonstrate that $g_{(X_1 \cap S_{i*})}y = 0$, for each $i \in I$.

Now, fix any $i \in I(S)$. It follows from (3) and (9) that

$$g_{(X_1 \cap S_{i*})}y = \sum_{\lambda : (i, \lambda) \in X_1} (i, \lambda) \cdot (j, \mu) = \sum_{\lambda : (i, \lambda) \in X_1, p_{\lambda j} \neq \theta} (i, \mu).$$

Given that $X_1 = \mathrm{supp}(r) \setminus S_{*L(S)} \in E(S)$, we see that $S_{i*} \cap X_1$ belongs to $E_i(S)$; whence the number of elements $(i, \lambda) \in X_1$ satisfying $p_{\lambda j} \neq \theta$ is even. Since $F = GF(2)$, we get $g_{(X_1 \cap S_{i*})}y = 0$. Hence $ry = 0$. It follows that $r \in \mathrm{Ann}_\ell(F_0[S])$, as required.

The 'only if' part. Suppose that $r \in \mathrm{Ann}_\ell(F_0[S])$. Clearly, $X = \cup_{i \in I(S)} X_i$, where $X_i = X \cap S_{i*}$. Since $F = GF(2)$, we get $r = \sum_{i \in I(S)} g_{X_i}$. In order to prove that X belongs to $E(S)$, it suffices

to verify that, for each $i \in I(S)$, the set

$$Y_i = \{\lambda \mid (i, \lambda) \in X_i\}$$

belongs to E_i. To this end, we fix $i \in I(S)$.

Consider an arbitrary element j in $I(S)$. Picking any $\mu \in \Lambda(S)$ and letting $y = (j, \mu)$, we get $ry = 0$ by the choice of r. It follows from (3), (9) and (18) that $g_{X_i}y = 0$.

Therefore the number of elements $(i, \lambda) \in X_i$ satisfying $p_{\lambda j} \neq \theta$ is even. Hence the number of $\lambda \in \Lambda(S)$ such that $p_{\lambda j} \neq \theta$ is even. This is equivalent to saying that the set Y_i belongs to $E_i(S)$. Therefore $\mathrm{supp}(r) \in E(S)$. This completes the proof. $\square$

Lemma 2. *Let $T = M^0(G; I, \Lambda; P)$ be a Rees matrix semigroup over the group $G = \{e\}$, and let S be a finite subsemigroup of T. Then, for any subset $Y \in Z(S)$, the right ideal generated by the element $g_{S_{*L(S)}} + g_Y$ has weight equal to $|S_{*L(S)}| + N(S)$.*

Proof. Since $Z(S) \subseteq E(S)$, Lemma 1 implies that $g_{S_{*L(S)}} + g_Y \in \operatorname{Ann}_\ell(F_0[S])$. Therefore (6) shows that the right ideal $C_r(g_{S_{*L(S)}} + g_Y)$ is equal to the set

$$\{0, g_{S_{*L(S)}} + g_Y\},$$

and so its weight is equal to $\operatorname{wt}(g_{S_{*L(S)}} + g_Y) = |S_{*L(S)}| + N(S)$. $\square$

Lemma 3. *Let $T = M^0(G; I, \Lambda; P)$ be a Rees matrix semigroup over the group $G = \{e\}$, and let S be a finite subsemigroup of T. Let λ be an element of Λ such that the cardinality $|S_{*\lambda}|$ achieves the maximum value. Then the right ideal generated by the element $g_{S_{*\lambda}}$ has weight equal to $|S_{*\lambda}|$.*

Proof. Consider an arbitrary nonzero element x in the right ideal $C_r(g_{S_{*\lambda}})$. Since $F = GF(2)$, it follows from (2), (3), (4), (6) and (9) that there exists a finite subset Y of S and an element $f \in F$ such that

$$x = fg_{S_{*\lambda}} + \sum_{y \in Y} g_{S_{*\lambda}} y.$$

Take any element $y = (j, \mu) \in Y$, where $j \in I(S)$, $\mu \in \Lambda(S)$. If $p_{\lambda j} = \theta$, then $S_{*\lambda} y = \theta$ in S, and so $g_{S_{*\lambda}} y = 0$, which means that we can delete y from the set Y. So we assume that $p_{\lambda j} \neq \theta$. Then (9) implies that $S_{*\lambda} y \subseteq S_{*\mu}$. The maximality of $|S_{*\lambda}|$ shows that $S_{*\lambda} y = S_{*\mu}$. Therefore we get $g_{S_{*\lambda}} y = g_{S_{*\mu}}$. Since $f \in F = GF(2)$, repeating we see that there exist elements $\mu_1, \ldots, \mu_k$ such that

$$x = g_{S_{*\mu_1}} + \cdots + g_{S_{*\mu_k}}$$

and

$$|S_{*\lambda}| = |S_{*\mu_1}| = \cdots = |S_{*\mu_k}|.$$

Since $F = GF(2)$, we see that $\mu_{i_1} = \mu_{i_2}$ implies $g_{S_{*\mu_{i_1}}} + g_{S_{*\mu_{i_2}}} = 0$. It follows that $\operatorname{wt}(x) \geq |S_{*\lambda}|$.

This means that $\operatorname{wt}(C_r(g_{S_{*\lambda}})) \geq |S_{*\lambda}|$. Since $C_r(g_{S_{*\lambda}})$ contains the element $g_{S_{*\lambda}}$ with weight equal to $|S_{*\lambda}|$, we get $\operatorname{wt}(C_r(g_{S_{*\lambda}})) = |S_{*\lambda}|$, as required. $\square$

Proof of Theorem 1. The maximality of $\operatorname{wt}(C_r(g_1, \ldots, g_k))$ and Lemma 2 show that

$$\operatorname{wt}(C_r(g_1, \ldots, g_k)) \geq |S_{*L(S)}| + N(S).$$

Likewise Lemma 3 implies that $\operatorname{wt}(C_r(g_1, \ldots, g_k)) \geq \max_{\lambda \in \Lambda(S)} |S_{*\lambda}|$. Hence we get

$$\operatorname{wt}(C_r(g_1, \ldots, g_k)) \geq \max\{|S_{*L(S)}| + N(S), \max_{\lambda \in \Lambda(S)} |S_{*\lambda}|\}. \tag{28}$$

To prove the reversed inequality, take a nonzero element x with minimal weight in $C_r(g_1, \ldots, g_k)$. Note that $\mathrm{wt}(C_r(g_1, \ldots, g_k)) = \mathrm{wt}(x) = |\mathrm{supp}(x)|$.

If $x \in \mathrm{Ann}_\ell(F_0[S])$, then Lemma 1 tells us that $\mathrm{supp}(x) \setminus S_{*L(S)} \in E(S)$. The cardinality of any set in $E(S)$ does not exceed $N(S)$. Therefore $|\mathrm{supp}(x) \setminus S_{*L(S)}| \leq N(S)$. It follows that $\mathrm{wt}(x) \leq |S_{*L(S)}| + N(S)$. Therefore $\mathrm{wt}(C_r(g_1, \ldots, g_k)) \leq |S_{*L(S)}| + N(S)$ in this case.

On the other hand, if $x \notin \mathrm{Ann}_\ell(F_0[S])$, then there exists an element $y = (j, \mu)$ such that $xy \neq 0$, where $j \in I(S)$, $\mu \in \Lambda(S)$. It follows from (3) and (9) that $\mathrm{supp}(xy) \subseteq S_{*\mu}$, and so $\mathrm{wt}(xy) \leq |S_{*\mu}|$. Since $xy \in C_r(x) \subseteq C_r(g_1, \ldots, g_k)$, we get $\mathrm{wt}(C_r(g_1, \ldots, g_k)) \leq |S_{*\mu}|$.

Therefore, in both of these cases we get

$$\mathrm{wt}(C_r(g_1, \ldots, g_k)) \leq \max\{|S_{*L(S)}| + N(S), \max_{\lambda \in \Lambda(S)} |S_{*\lambda}|\}. \tag{29}$$

Equality (20) follows from (28) and (29). This completes the proof. $\square$

Proof of Corollary 2. It is well known and easy to verify that every multiple classifier, defined by a set that is a linear space of weight W, can correct $\lfloor \frac{W-1}{2} \rfloor$ errors of binary classifiers. Therefore Corollary 2 follows from Theorem 1 immediately. $\square$

4   Acknowledgements

References

[1] Alfaro, R. and A.V. Kelarev, *Recent results on ring constructions for error-correcting codes*, Contemporary Math. 376 (2005), 1–12.

[2] Alfaro, R. and Kelarev, A.V., *On cyclic codes in incidence rings,*, Studia Scientiarum Mathematicarum Hungarica **43** (2006), 69–77.

[3] Cazaran, J., Kelarev, A.V., Quinn, S.J. and Vertigan, D., *An algorithm for computing the minimum distances of extensions of BCH codes embedded in semigroup rings*, Semigroup Forum **73**(2006), 317–329.

[4] Fernandes, V.H. and Volkov, M.V., *On divisors of semigroups of order-preserving mappings of a finite chain*, Semigroup Forum, 81 (2010), 551–554.

[5] Howie, J.M., "Fundamentals of Semigroup Theory", Clarendon Press, Oxford, 1995.

[6] Jackson, M. and Volkov, M., *Undecidable problems for completely 0-simple semigroups*, J. Pure Applied Algebra **213**(10)(2009), 1961-1978.

[7] Kelarev, A.V., "Ring Constructions and Applications", World Scientific, River Edge, 2002.

[8] Kelarev, A.V., "Graph Algebras and Automata", Marcel Dekker, New York, 2003.

[9] Kelarev, A.V. and Passman, D.S., *A description of incidence rings of group automata*, Contemporary Mathematics **456**(2008), 27–33.

[10] Kelarev, A., Ryan, J. and Yearwood, J., *Cayley graphs as classifiers for data mining: The influence of asymmetries*, Discrete Mathematics **309**(2009)(17), 5360–5369.

[11] Kelarev, A.V. and Solé, P., *Error-correcting codes as ideals in group rings*, Contemporary Mathematics **273** (2001), 11-18.

[12] Kelarev, A.V., Yearwood, J.L. and Mammadov, M.A., *A formula for multiple classifiers in data mining based on Brandt semigroups*, Semigroup Forum **78**(2009)(2), 293–309.

[13] Kelarev, A.V., Watters, P. and Yearwood, J.L., *Rees matrix construction for clustering of data*, J. Aust. Math. Soc. **87**(2009), 377–393.

[14] Kelarev, A.V., Yearwood, J.L. and Vamplew, P.W., *A polynomial ring construction for classification of data*, Bull. Aust. Math. Soc. **79**(2009), 213–225.

[15] Kelarev, A.V., Yearwood, J.L. and Watters, P.A., *Optimization of classifiers for data mining based on combinatorial semigroups*, Semigroup Forum 82(2011)(2), 242–251.

[16] Reilly, N.R., *Bipartite graphs and completely 0-simple semigroups*, Semigroup Forum, 84 (2012), 176–199.

[17] Witten, I.H. and Frank, E., "Data Mining: Practical Machine Learning Tools and Techniques", Elsevier/Morgan Kaufman, Amsterdam, 2005.

[18] Yearwood, J., Webb, D., Vamplew, P., Ma, L., Ofoghi, B. and Kelarev, A., *Applying clustering and ensemble clustering approaches to phishing profiling*, Eighth Australasian Data Mining Conference, AusDM 2009, (Melbourne, Australia, December 2009), Conferences in Research and Practice in Information Technology **101**(2009), 25–34.

[19] Yearwood, J.L., Bagirov, A.M. and Kelarev, A.V., *Optimization methods and the k-committees algorithm for clustering of sequence data*, J. Applied & Computational Mathematics 8(2009)(1), 92–101.