NEW ZEALAND JOURNAL OF MATHEMATICS
Volume 41 (2011), 65-73

OPTIMIZATION AND MATRIX CONSTRUCTIONS FOR CLASSIFICATION OF DATA

A.V. Kelarev, J.L. Yearwood, P.W. Vamplew, J. Abawajy,
M. Chowdhury

(Received November 2011)

Abstract. Max-plus algebras and more general semirings have many useful applications and have been actively investigated. On the other hand, structural matrix rings are also well known and have been considered by many authors. The main theorem of this article completely describes all optimal ideals in the more general structural matrix semirings. Originally, our investigation of these ideals was motivated by applications in data mining for the design of centroid-based classification systems, as well as for the design of multiple classification systems combining several individual classifiers.

1. Introduction

Semirings have been actively investigated, because they are important in mathematics and have many useful applications, see [**10, 11**]. As a special example of a semiring, let us only mention the max-plus algebra, which plays crucial roles in the study of discrete event systems, see [**3, 10**]. On the other hand, structural matrix rings have also been considered in the literature and many interesting results have been obtained (see, for example, [**5, 8, 12, 22, 23**]).

The present article is devoted to the investigation of the more general structural matrix semirings. Our main theorem gives a complete description of all ideals with largest weights in structural matrix semirings. Originally our investigation of these ideals was motivated by their applications to the design of classification systems, or classifiers, considered in data mining. We refer to the monograph [**24**] for more information on the design of classifiers and their roles in data mining. More detailed explanations are also given in Section 2 below. In particular, special sets satisfying certain optimal properties are required for the design of centroid-based classifiers, as well as for the design of multiple classifiers combining several individual or initial classifiers, see [**19, 20**].

The paper is organised as follows. An overview of applications of matrix constructions for classification of data is given in Section 2 as motivation for this research. The main result of this paper is Theorem 1 in Section 3, which completely describes all ideals with largest weights in structural matrix semirings. A complete proof is given in Section 4.

2010 *Mathematics Subject Classification* : 16S50, 40C05, 65F30.
Key words and phrases: matrix constructions, ideals, semirings, optimization.
The first author was supported by Discovery grant DP0449469 from Australian Research Council. The second author was supported by Queen Elizabeth II Fellowship, Discovery grant DP0211866 and Linkage grant LP0990908 from Australian Research Council. All authors were supported by a UB-Deakin collaboration grant.

2. Motivation and Preliminaries

This section contains a concise review of the main definitions required for our new theorem. We use standard notions and terminology and refer the readers to [**5, 6, 7, 9, 10, 11, 13, 14, 15, 16, 17, 18, 21, 24, 26**] for more detailed discussions of these concepts and examples of recent results.

The design of efficient classifiers is very important in data mining, see [**24**]. Max-plus algebras, more general semirings, and matrix constructions over them 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 initial classifiers.

Classification deals with known classes of data. These classes are represented by given samples of data. The samples are used for supervised training of the classifier to enable it to recognize new elements of the same known classes. The classification process begins with feature extraction and representation of data in a standard vector space F^n, where F can be regarded as a semifield.

A *semifield* is a semiring, where the set of nonzero elements forms a group with respect to multiplication. Recall that a *semiring* is a set F with two binary operations, addition $+$ and multiplication $\cdot$, such that the following conditions are satisfied:

(S1) $(F, +)$ is a commutative semigroup with zero 0,
(S2) $(F, \cdot)$ is a semigroup,
(S3) multiplication distributes over addition,
(S4) zero 0 annihilates F, i.e., $0 \cdot F = F \cdot 0 = 0$.

It is also often assumed that every semiring satisfies an additional property
(S5) $(F, \cdot)$ has an identity element 1.

Our results remain valid without assuming (S5), and so we consider more general semirings, which do not have to satisfy (S5). In analogy with a similar situation in ring theory, we then call every semiring satisfying (S5) a *semiring with identity element*. As usual, such more general terminology adds the convenience of allowing us to consider more general subsets as subsemirings without assuming that all subsemirings contain the identity element. Both terminologies are essentially equivalent, since it is always easy to adjoin an identity element in a standard fashion to every semiring that does not have one.

Every centroid-based classifier selects special elements $c_1, \dots, c_k$ in F^n, called *centroids* (see, for example, [**4**]). For $i = 1, \dots, k$, 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 are often used in analysis of data to combine individual initial classifiers (see, for example, [**27**]). 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 [**24**], Section 7.5 and [**19**]. 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.

Denote the number of initial classifiers being combined by n. If $o_1, \dots, o_n$ are the outputs of the initial classifiers, then the sequence $(o_1, \dots, o_n)$ 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.

The design of multiple classifiers by combining individual classifiers is quite common in the literature. We refer to [**19, 20**] and [**24**] for a list of properties required of the sets of centroids. In particular, it is essential to find sets of centroids with large weights and small numbers of generators. The *weight* $\mathrm{wt}(v)$ of $v \in F^n$ is the number of nonzero components or coordinates in v. The *weight* of a set $C \subseteq F^n$ is the minimum weight of a nonzero element in C. For additional references and discussion of experimental research related to these properties we refer the readers to [**19, 20**].

The *information rate* of a class set C in F^m can be defined as $\log_{|F|}(|C|)/m$. It reflects the proportion of output of the individual initial classifiers used to produce the outcomes of the multiple classification, as opposed to additional efforts spent on increasing reliability and correcting classification errors.

All sequences of the centroid set C can be written down in a matrix M to discuss their properties. If M has two identical columns, this means that two initial classifiers produce identical outputs. This duplication is very inefficient, even though it could help to correct classification errors. Therefore, in a situation like this, one of these classifiers can be removed and a better scheme can be devised. Likewise, it is undesirable to have strong correlation or functional dependencies between very small sets of columns in M or between the initial classifiers.

According to [**24**], Section 7.5, for a classifier with a class set C to be efficient, the class C must satisfy the following most essential basic properties:

(1) The set C must have a large weight.
(2) The information rate of C must be large.
(3) A small set of generators for the set C is essential in order to simplify computer storage and manipulation of the set.
(4) If all vectors of C are recorded in a matrix M, then there should not be strong correlation or functional dependencies between small sets of columns of M. In particular, the matrix M should not have duplicate columns.

Thus, in particular, it is essential to find sets of centroids with large weights and small numbers of generators. For additional references and discussion of experimental research related to these properties we refer the readers to [**19, 20, 26, 27**].

The *max-plus algebra* is the set $\mathbb{R} \cup \{-\infty\}$ with two binary operations, max and $+$. It is very important in the investigation of discrete event systems, see [**3**]. The max-plus algebra is also sometimes called the *schedule algebra*, see [**10**]. Our main results remain valid in the more general case of all semifields, and so we record them in this setting.

Let F be a semiring. Consider the semiring $M_m(F)$ of all $m \times m$ matrices over F. Let ϱ be a binary relation on the set $[1 : m] = \{1, \ldots, m\}$. For $i, j \in [1 : m]$ denote by $e_{i,j}$ the standard elementary matrix in $M_m(F)$ with 1 in the intersection of i-th row and j-th column and zeros in all other entries. It is well known and easy to verify that the set $M_\varrho(F) = \bigoplus_{(i,j)\in\varrho} F e_{i,j}$ is a subsemiring of $M_m(F)$ if and only if the relation ϱ is transitive. In this case $M_\varrho(F)$ is called a *structural matrix* *semiring*. Many interesting results on structural matrix rings have been obtained in the literature (see, for example, [**5, 8, 12, 22, 23**]). Known facts and references concerning structural matrix rings can be also found in [**14**].

If $|\varrho| = n$, then the additive semigroup of $M_\varrho(F)$ is isomorphic to F^n and we can introduce multiplication in F^n by identifying it with $M_\varrho(F)$. Further we consider sets of centroids as subsets generated in $M_\varrho(F)$. Every set of elements $g_1, \dots, g_k \in M_\varrho(F)$ generates the set $C(g_1, \dots, g_k)$ of all sums of these elements and their multiples:

$$C(g_1, \dots, g_k) = $$
$$= \left\{ \sum_{j=1}^{m_1} \ell_{1,j} g_1 r_{1,j} + \cdots + \sum_{j=1}^{m_k} \ell_{k,j} g_k r_{k,j} \;\middle|\; \ell_{i,j}, r_{i,j} \in M_\varrho(F) \cup \{1\} \right\}. \quad (1)$$

The set $C(g_1, \dots, g_k)$ is called an *ideal* generated by $g_1, \dots, g_k$. The concept of an ideal is very important and has been actively investigated in several branches of modern mathematics. In particular, it is used in the investigation of modules over rings (see, for example, [**1, 2**]) and ring constructions (see, for example, [**13, 14**]).

3. Main Results

Let ϱ be a binary relation on the set $[1:m]$. We introduce the following binary relations

$$\varrho_\ell = \{(i,j) \in \varrho \mid \exists k \in [1:m] : (k,i) \in \varrho\}, \quad (2)$$
$$\varrho_r = \{(i,j) \in \varrho \mid \exists k \in [1:m] : (j,k) \in \varrho\}. \quad (3)$$

and put

$$M_Z = |\varrho \setminus (\varrho_r \cup \varrho_\ell)|. \quad (4)$$

For any $i \in [1:m]$, let us define the sets

$$\varrho(i) = \{j \mid (i,j) \in \varrho\}, \quad (5)$$
$$\varrho^{-1}(i) = \{j \mid (j,i) \in \varrho\}, \quad (6)$$
$$R(i) = \{j \mid (i,j) \in \varrho \setminus \varrho_r\}, \quad (7)$$
$$L(i) = \{j \mid (j,i) \in \varrho \setminus \varrho_\ell\}. \quad (8)$$

We introduce the following nonnegative integers

$$M_L = \max\{|L(i)| \,:\, i = 1, \dots, m\}, \quad (9)$$
$$M_R = \max\{|R(i)| \,:\, i = 1, \dots, m\}. \quad (10)$$

Denote by $\mathcal{G}_Z$ the set of all elements $g = \sum_{(i,j)\in\varrho\setminus(\varrho_r\cup\varrho_\ell)} f_{i,j} e_{i,j} \in M_\varrho(F)$, where $0 \neq f_{i,j} \in F$. Let $\mathcal{G}_L$ be the set of all elements $g = \sum_{j\in L(i)} f_j e_{j,i} \in M_\varrho(F)$, where i runs over the set of all integers i such that $|L(i)| = M_L$, and where $0 \neq f_j \in F$. Denote by $\mathcal{G}_R$ the set of all elements $g = \sum_{j\in R(i)} f_j e_{i,j} \in M_\varrho(F)$, where i runs over the set of all integers i such that $|R(i)| = M_R$, and where $0 \neq f_j \in F$. Our main theorem describes all sets $C(g_1, \dots, g_k)$ with the largest weight in $M_\varrho(F)$.

Theorem 1. *Let $M_\varrho(F)$ be a structural matrix semiring over a semifield F. Suppose that $C = C(g_1, \dots, g_k)$ is an ideal with the largest weight in $M_\varrho(F)$. Then the following conditions are satisfied:*

(i) $\mathrm{wt}(C) = \max\{1, M_Z, M_L, M_R\}$;
(ii) *if* $\mathrm{wt}(C) > 1$, *then* $C \cap (\mathcal{G}_Z \cup \mathcal{G}_L \cup \mathcal{G}_R)$ *contains an element of weight* $\mathrm{wt}(C)$;
(iii) $\mathrm{wt}(C(g)) = \mathrm{wt}(g) = M_Z$, *for all* $g \in \mathcal{G}_Z$;
(iv) $\mathrm{wt}(C(g)) = \mathrm{wt}(g) = M_L$, *for all* $g \in \mathcal{G}_L$;
(v) $\mathrm{wt}(C(g)) = \mathrm{wt}(g) = M_R$, *for all* $g \in \mathcal{G}_R$.

4. Proofs

For any semiring F, the *left annihilator* of F is the set

$$\mathrm{Ann}_\ell(F) = \{x \in F \mid xF = 0\}, \tag{11}$$

and the *right annihilator* of F is the set

$$\mathrm{Ann}_r(F) = \{x \in F \mid Fx = 0\}. \tag{12}$$

Lemma 2. *For any structural matrix semiring $M_\varrho(F)$ over a semifield F, the following equalities are satisfied:*

$$\mathrm{Ann}_r(M_\varrho(F)) = M_{\varrho\setminus\varrho_\ell}(F), \tag{13}$$

$$\mathrm{Ann}_\ell(M_\varrho(F)) = M_{\varrho\setminus\varrho_r}(F). \tag{14}$$

Proof. Take any element r in $\mathrm{Ann}_r(M_\varrho(F))$. It can be recorded as

$$r = \sum_{(i,j)\in\varrho} f_{i,j} e_{i,j},$$

where $f_{i,j} \in F$. Consider any pair (i,j) in ϱ_ℓ. There exists $k \in [1:m]$ such that $(k,i), (k,j) \in \varrho$. Hence $e_{k,i} \in M_\varrho(F)$ and $f_{i,j}e_{k,i}e_{i,j}$ is a summand of the product $e_{k,i}r$ in $M_\varrho(F)$. Since $r \in \mathrm{Ann}_r(M_\varrho(F))$, we get $f_{i,j} = 0$. It follows that r belongs to $M_{\varrho\setminus\varrho_\ell}(F)$, and so $\mathrm{Ann}_r(M_\varrho(F)) \subseteq M_{\varrho\setminus\varrho_\ell}(F)$.

To prove the reversed inclusion, let us pick any element r in $M_{\varrho\setminus\varrho_\ell}(F)$. It can be written down as $r = \sum_{(i,j)\in\varrho\setminus\varrho_\ell} f_{i,j}e_{i,j}$, where $f_{i,j} \in F$. In order to verify that $M_\varrho(F)r = 0$, it suffices to show that $e_{a,b}r = 0$ for all $(a,b) \in \varrho$. Suppose to the contrary that $e_{a,b}r \neq 0$ for some $(a,b) \in \varrho$. Then it is clear that at least one of the summands $e_{a,b}f_{i,j}e_{i,j}$ is nonzero for some $(i,j) \in \varrho \setminus \varrho_\ell$. The definition of a structural matrix semiring implies that $f_{i,j} \neq 0$, $b = i$, $(a,i) = (a,b) \in \varrho$ and $(a,j) \in \varrho$. Hence $(i,j) \in \varrho_\ell$. This contradicts the choice of (i,j) in $\varrho \setminus \varrho_\ell$ and shows that $e_{a,b}r = 0$ for all $(a,b) \in \varrho$. Therefore $M_\varrho(F)r = 0$, which means that $r \in \mathrm{Ann}_r(M_\varrho(F))$. Thus $\mathrm{Ann}_r(M_\varrho(F)) \supseteq M_{\varrho\setminus\varrho_\ell}(F)$.

These two inclusions show that equality (13) always holds. The proof of equality (14) is dual and we omit it. □

Proof of Theorem 1. **(iii):** Consider any element $g \in \mathcal{G}_Z$. By definition, we know that

$$g = \sum_{(i,j)\in\varrho\setminus(\varrho_r\cup\varrho_\ell)} f_{i,j}e_{i,j} \in M_\varrho(F),$$

where $0 \neq f_{i,j} \in F$. Clearly, $\mathrm{wt}(g) = |\varrho \setminus (\varrho_r \cup \varrho_\ell)| = M_Z$. Since $\varrho \setminus (\varrho_r \cup \varrho_\ell) = (\varrho \setminus \varrho_r) \cap (\varrho \setminus \varrho_\ell)$, Lemma 2 and (1) show that $C(g)$ coincides with the linear space Fg spanned by g. Since F is a semifield, it follows that all nonzero elements of $C(g)$ have weights equal to the weight of g. Hence $\mathrm{wt}(C(g)) = \mathrm{wt}(g)$ in this case, and so condition (iii) holds.

(iv): Choose any element $g \in \mathcal{G}_L$. It can be represented in the form

$$g = \sum_{j \in L(i)} f_j e_{j,i},$$

where $1 \le i \le m$, $|L(i)| = M_L$ and $0 \ne f_j \in F$. Hence we get $\mathrm{wt}(g) = |L(i)| = M_L$. It remains to verify that $\mathrm{wt}(C(g)) = \mathrm{wt}(g)$. To this end, we choose any nonzero element x in $C(g)$ and are going to verify that $\mathrm{wt}(x) \ge \mathrm{wt}(g)$. It follows from (1) that

$$x = \sum_{j=1}^{k} \ell_j g r_j, \tag{15}$$

for some $\ell_j, r_j \in M_\varrho(F) \cup \{F\}$, where we may assume that only nonzero summands $\ell_j r \ell_j$ are included in the sum. Since $(j,i) \notin \varrho_\ell$ for all $j \in L(i)$, it follows from Lemma 2 that $\ell_j g = 0$ for every $\ell_j \in M_\varrho(F)$. Therefore we may assume that all the ℓ_j are equal to 1 in the expression (15) for x above.

Keeping in mind that $M_\varrho(F) = \bigoplus_{(a,b)\in\varrho} F e_{a,b}$, the distributive law allows us to assume without loss of generality that every element $r_j \in M_\varrho(F)$ in the expression (15) for x belongs to the union $\cup_{a,b} F e_{a,b}$. Since $g r_j \ne 0$, it follows that then all the r_j belong to $\cup_{(i,b)\in\varrho} F e_{i,b}$. The transitivity of ϱ shows that $\varrho(i) \subseteq \varrho(j)$ for all j in $L(i)$. Since $g r_j \ne 0$, we see that all the r_j belong to $\cup_{b \in \varrho(i)} F e_{i,b}$. Since F is a semifield, it follows that $\mathrm{wt}(g r_j) = \mathrm{wt}(g)$ for all $j \in L(i)$. Therefore $\mathrm{wt}(x) \ge \mathrm{wt}(g)$, as required. Thus, condition (iv) holds.

(v): The proof of condition (v) is dual to that of (iv), and so we omit it.

(ii): Suppose that $\mathrm{wt}(C) > 1$. Choose a nonzero element g of minimal weight in C and consider several cases.

Case 1. $g \notin \mathrm{Ann}_r(M_\varrho(F)) \cup \mathrm{Ann}_\ell(M_\varrho(F))$. By Lemma 2, we get

$$g \notin M_{\varrho \setminus \varrho_\ell}(F) \cup M_{\varrho \setminus \varrho_r}(F).$$

Therefore there exist $(a,b), (c,d) \in \varrho$ such that $e_{a,b} g e_{c,d} \ne 0$. However, $e_{a,b} g e_{c,d} \in C$ and $\mathrm{wt}(e_{a,b} g e_{c,d}) = 1$. Hence $\mathrm{wt}(C) = 1$. This contradicts the assumption that $\mathrm{wt}(C) > 1$ and shows that Case 1 is impossible.

Case 2. $g \in \mathrm{Ann}_\ell(M_\varrho(F)) \cap \mathrm{Ann}_r(M_\varrho(F))$. Lemma 2 implies that

$$r \in M_{\varrho \setminus \varrho_\ell}(F) \cap M_{\varrho \setminus \varrho_\ell}(F).$$

It follows from the maximality of $\mathrm{wt}(C)$ and condition (iii), which we have already proved above, that $\mathrm{wt}(g) = M_Z$. Therefore $g \in \mathcal{G}_Z$. Since $\mathrm{wt}(g) = \mathrm{wt}(C)$, this means that condition (ii) holds in this case.

Case 3. $g \in \mathrm{Ann}_r(M_\varrho(F)) \setminus \mathrm{Ann}_\ell(M_\varrho(F))$. Then $g e_{i,t} \ne 0$ for some $(i,t) \in \varrho$. Obviously, $\mathrm{wt}(g e_{i,t}) \le \mathrm{wt}(g)$. By the minimality of the weight $\mathrm{wt}(g)$ in C, we get $\mathrm{wt}(g e_{i,t}) = \mathrm{wt}(g)$, because $g e_{i,t} \in C$. Therefore there exists a subset $S \subseteq \varrho^{-1}(i)$ such that $g = \sum_{j \in S} f_j e_{j,i}$, where $0 \ne f_j \in F$. Clearly, $|S| = \mathrm{wt}(g)$. Since $g \in M_\varrho(F)$, we get $S \subseteq \varrho^{-1}(i)$. Lemma 2 and $g \in \mathrm{Ann}_r(M_\varrho(F))$ show that $(j,i) \in \varrho \setminus \varrho_\ell$ for all $j \in S$. Therefore $S \subseteq L(i)$.

The maximality of $\mathrm{wt}(C) = \mathrm{wt}(g) = |S|$ and condition (iv) proved above imply that $\mathrm{wt}(g) \ge M_L$. By the definition of M_L, we get $M_L \ge |S| = \mathrm{wt}(g)$. Therefore $|S| = M_L$ and $S = L(i)$. It follows that $g \in \mathcal{G}_L$. This means that condition (ii) holds true in this case, too.

Case 4. $g \in \mathrm{Ann}_\ell(M_\varrho(F)) \setminus \mathrm{Ann}_r(M_\varrho(F))$. In this case a dual proof to the proof of Case 3 demonstrates that $g \in \mathcal{G}_R$. Therefore condition (ii) always holds true.

(i): Clearly, condition (ii) implies that $\mathrm{wt}(C) \leq \max\{1, M_Z, M_L, M_R\}$. On the one hand, the maximality of $\mathrm{wt}(C)$ and conditions (iii), (iv), (v), (vi) show that

$$\mathrm{wt}(C) \geq \max\{1, M_Z, M_L, M_R\}.$$

Therefore condition (i) is satisfied. This completes our proof. □

References

[1] M.M. Ali, *Invertiblility of multiplication modules II*, New Zealand J. Math., **39**, (2009), 45–64.

[2] M.M. Ali and D.J. Smith, *Projective, flat and multiplication modules*, New Zealand J. Math, **31** (2002), 115–129. (Corrigendum in New Zealand J. Math. **39** (2009), 241–243.)

[3] F. Baccelli, G. Cohen, G.J. Olsder and J.-E. Quadrat, *Synchronization and Linearity: An Algebra for Discrete Event System*, Wiley Interscience, New York, 1992.

[4] A.M. Bagirov and J.L. Yearwood, *A new nonsmooth optimization algorithm for minimum sum-of-squares clustering problems*, European J. Operational Research, **170** (2006), 578–596.

[5] S. Bereg, A. Kelarev and A. Salagean, *Directed graphs and minimum distances of error-correcting codes in matrix rings*, New Zealand J. Math., **33** (2004)(2), 113–120.

[6] H.Y. Chen and M.S. Chen, *PA note on rings of weakly stable range one*, New Zealand J. Math., **35** (2006), 137–143.

[7] D.M. Clark and B.A. Davey, *Natural Dualities for the Working Algebraist*, Cambridge University Press, Cambridge, 1998.

[8] S. Dăscălescu and L. van Wyk, *Do isomorphic structural matrix rings have isomorphic graphs?*, Proc. Amer. Math. Soc. **124** (1996), 1385–1391.

[9] D. Easdown, J. East, D.G. FitzGerald, *A presentation of the dual symmetric inverse monoid*, Internat. J. Algebra Comput. **18** (2008)(2), 357–374.

[10] J.S. Golan, *Semirings and Their Applications*, Kluwer Academic Publishers, Dordrecht, 1999.

[11] J.S. Golan, *Semirings and Affine Equations over them: Theory and Applications*, Kluwer Academic Publishers, Dordrecht, 2003.

[12] B.W. Green and L. van Wyk, *On the small and essential ideals in certain classes of rings*, J. Austral. Math. Soc. Ser. A, **46** (1989), 262–271.

[13] A.V. Kelarev, *Two ring constructions and sums of fields*, New Zealand J. Math., **28** (1999), 43–46.

[14] A.V. Kelarev, *Ring Constructions and Applications*, World Scientific, River Edge, NJ, 2002.

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

[16] A.V. Kelarev, R. Göbel, K.M. Rangaswamy, P. Schultz and C. Vinsonhaler, *Abelian Groups, Rings and Modules*, American Mathematical Society, Contemporary Mathematics, 273, New York, 2001.

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

[18] A.V. Kelarev, C.E. Praeger, *On transitive Cayley graphs of groups and semigroups*, European J. Combinatorics, **24** (2003)(1), 59–72.

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

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

[21] G.J. Tee, *Eigenvectors of block circulant and alternating circulant matrices*, New Zealand J. Math., **36** (2007), 195–211.

[22] L. van Wyk, *Matrix rings satisfying column sum conditions versus structural matrix rings*, Linear Algebra Appl., **249** (1996), 15–28.

[23] L. van Wyk, *A link between a natural centralizer and the smallest essential ideal in structural matrix rings*, Comm. Algebra, **27** (1999), 3675–3683.

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

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

[26] J.L. Yearwood and M.A. Mammadov, *Classification Technologies: Optimization Approaches to Short Text Categorization*, Idea Group Inc., 2007.

[27] J. Yearwood, D. Webb, L. Ma, P. Vamplew, B. Ofoghi and A. Kelarev, *Applying clustering and ensemble clustering approaches to phishing profiling*, Data Mining and Analytics 2009, Proc. 8th Australasian Data Mining Conference: AusDM 2009, (1-4 December 2009, Melbourne, Australia) CRPIT, Vol.101, pp. 25–34.

A.V. Kelarev, J.L. Yearwood, P.W. Vamplew,
School of Science, Information Technology
and Engineering
University of Ballarat, P.O. Box 663,
Ballarat, Victoria 3353, Australia
{a.kelarev,j.yearwood,p.vamplew}@ballarat.edu.au

J. Abawajy, M. Chowdhury
School of Information Technology, Deakin
University
221 Burwood Highway, Burwood, Victoria
3125, Australia
{jemal.abawajy,morshed.chowdhury}@deakin.edu.au