Mixed Covering Arrays on Graphs

Karen Meagher
Department of Combinatorics and Optimization
University of Waterloo
Ottawa, Ontario, Canada

Lucia Moura
School of Information Technology and Engineering
University of Ottawa
Ottawa, Ontario, Canada

Latifa Zekaoui
School of Information Technology and Engineering
University of Ottawa
Ottawa, Ontario, Canada

ABSTRACT

Covering arrays have applications in software, network and circuit testing. In this paper, we consider a generalization of covering arrays that allows mixed alphabet sizes as well as a graph structure that specifies the pairwise interactions that need to be tested. Let k and N be positive integers, and let G be a graph with k vertices $v_1, v_2, \dots, v_k$ with respective vertex weights $g_1 \leq g_2 \leq \cdots \leq g_k$. A *mixed covering array* on G, denoted by $CA(N, G, \prod_{i=1}^{k} g_i)$, is a $k \times N$ array such that row i corresponds to v_i, cells in row i are filled with elements from $\mathbb{Z}_{g_i}$ and every pair of rows i, j corresponding to $\{v_i, v_j\} \in E(G)$ contain every possible pair of $\mathbb{Z}_{g_i} \times \mathbb{Z}_{g_j}$ appearing in a column. The number of columns in such array is called its *size*. Given a weighted graph G, a mixed covering array on G with minimum size is called *optimal*. In this paper, we give upper and lower bounds on the size of mixed covering arrays on graphs based on graph homomorphisms. We provide constructions for covering arrays on graphs based on basic graph operations. In particular, we construct optimal mixed covering arrays on trees, cycles and bipartite graphs; the constructed optimal objects

Journal of Combinatorial Designs Vol. **Submitted., ()**

have the additional property of being nearly point balanced.

Keywords: Covering array, Mixed Covering Arrays, Qualitative Independence Graphs, Software testing.

1. INTRODUCTION

Covering arrays are generalizations of orthogonal arrays that are used for testing software [3, 6, 7], networks [14] and circuits [12]. They have also been applied to problems in other domains, such as material science [1] and genomics [13].

Two length-n vectors x, y with entries from $\mathbb{Z}_g$ and $\mathbb{Z}_h$, respectively, are *qualitatively independent* if for any pair $(a, b) \in \mathbb{Z}_g \times \mathbb{Z}_h$, there exists $k \in \{1, 2, \ldots, n\}$ such that $(x_k, y_k) = (a, b)$. Let N, g, k be positive integers. A *covering array*, denoted by $CA(N, k, g)$, is a $k \times N$ array A with entries from a g-ary alphabet such that any two distinct rows of A are qualitatively independent. The parameter N is called the *size* of the array. Given k and g, the covering array number, denoted by $CAN(k, g)$, is the minimum N for which there exists a $CA(N, k, g)$. A $CA(N, k, g)$ of size $N = CAN(k, g)$ is called *optimal*.

The testing application is based on the following translation. The system to be tested has k parameters, each of which can take one out of g possible values. We wish to build a test set that tests all pairwise interactions of parameters with the minimum number of tests N.

Several generalizations of covering arrays have been proposed in order to address different requirements of the testing application (see [4, 8]). A mixed covering array is a generalization of covering arrays that allows for different alphabets in different rows. This meets the requirement that different parameters in the system may take a different number of possible values. Constructions for mixed covering arrays are given in [5, 11]. Another generalization of covering arrays are *covering arrays on graphs*. In these arrays, only specified pairs of rows need to be qualitatively independent and these pairs are recorded in a graph structure [9, 10]. This is usefull in situations in which some pairs of parameters do not interact; in these cases, we do not insist that these interactions be tested, which allows reductions in the number of required tests. In this paper, we consider a structure that generalizes both mixed covering arrays and covering arrays on graphs. This has been considered in the context of software testing where we only test the interactions given by parameters that jointly affect one of the output values [2].

The parameters for mixed covering arrays on graphs are given by a vertex-weighted graph. A *vertex-weighted graph* or simply a *weighted graph* consists of a graph with weight on the vertices. A weighted graph is given by a triple $(V(G), E(G), w_G(\cdot))$ where $V(G)$ is the vertex set, $E(G)$ is the edge set and $w_G : V(G) \rightarrow \mathbb{N}^+$ is the weight function. We assume that the vertices are labelled in a nondecreasing order of weights, that is, $w_G(v_i) \leq w_G(v_j)$, for all $1 \leq i \leq j \leq |V(G)|$. When there is no ambiguity on which graph we are using, we can write $w(\cdot)$ in place of $w_G(\cdot)$.

Definition 1.1. *(Mixed Covering Array on a Graph) Let G be a weighted graph with k vertices and weights $g_1 \leq g_2 \leq \cdots \leq g_k$, and let N be a positive integer. A* mixed covering array *on G, denoted by $CA(N, G, \prod_{i=1}^{k} g_i)$, is a $k \times N$ array with the following properties: 1) the cells in row i are filled with elements from a g_i-ary alphabet, which is usually taken to be $\mathbb{Z}_{g_i}$; 2) row i corresponds to a vertex $v_i \in V(G)$ with $w_G(v_i) = g_i$; and 3) pairs of rows that correspond to adjacent vertices of G are qualitatively independent. Given a weighted graph G with weights $g_1, g_2, \ldots, g_k$, the* mixed covering array number *on G, denoted by $CAN(G, \prod_{i=1}^{k} g_i)$, is the minimum N for which there exists a $CA(N, G, \prod_{i=1}^{k} g_i)$. A $CA(N, G, \prod_{i=1}^{k} g_i)$ of size $N = CAN(G, \prod_{i=1}^{k} g_i)$ is called* optimal.

A mixed covering array, denoted by $CA(N, \prod_{i=1}^{\ell} g_i)$, is a $CA(N, K_\ell, \prod_{i=1}^{\ell} g_i)$, where K_ℓ is the complete graph on ℓ vertices with weights g_i, for $1 \leq i \leq \ell$. A covering array on a graph, denoted by $CA(N, G, g)$, is a $CA(N, G, g^k)$ where $k = |V(G)|$.

In this paper, we generalize some results for covering arrays on graphs [9, 10] to the mixed alphabet case. We also provide constructions for optimal mixed covering arrays on special classes of graphs. In Section 2., we review known results on covering arrays on graphs based on graph homomorphisms. In Section 3., we generalize these results for mixed covering arrays on graphs. In Section 4., we give constructions based on basic graph operations. Using these operations, we build optimal mixed covering arrays on trees and cycles, and using a different method, we build optimal mixed covering arrays on bipartite graphs. We also give upper bounds for all 3-chromatic and for a large number of 4- and 5-chromatic graphs.

2. GRAPH HOMOMORPHISMS AND COVERING ARRAYS ON GRAPHS

In this section, we review basic graph theoretical concepts and results for covering arrays on graphs from [10].

A mapping ϕ from V(G) to V(H) is a *graph homomorphism* from G to H (or simply a *homomorphism*) if for all $v, w \in V(G)$, the vertices $\phi(v)$ and $\phi(w)$ are adjacent in H whenever v and w are adjacent in G. For graphs G and H, if there exists a homomorphism from G to H we write $G \to H$.

The *complete graph* on n vertices, K_n, is the graph with n vertices and with an edge between any two distinct vertices. A *proper colouring* of G with n colours is a map from $V(G)$ to a set of n colours such that no two adjacent vertices are mapped to the same colour. A proper colouring of a graph G with n colours is equivalent to a homomorphism from G to K_n. The *chromatic number* of a graph G, denoted $\chi(G)$, is the smallest n such that $G \to K_n$. A *clique* in a graph G is a set C of vertices from $V(G)$ such that any two distinct vertices in C are adjacent in G. A clique of cardinality n in G is equivalent to a homomorphism from K_n to G. The *clique number* of a graph G, denoted by $\omega(G)$, is the largest n such that $K_n \to G$. For graphs G and H, if there is a homomorphism $G \to H$ then $\chi(G) \leq \chi(H)$ and $\omega(G) \leq \omega(H)$. Also, for all graphs G, $\omega(G) \leq \chi(G)$.

Theorem 2.1 [10]. *Let g be a positive integer and G and H be graphs. If there exists a graph homomorphism $\phi : G \to H$, then*

$$CAN(G, g) \leq CAN(H, g).$$

For any graph G, there are homomorphisms between the following complete graphs

$$K_{\omega(G)} \to G \to K_{\chi(G)}.$$

These homomorphisms can be used to find bounds on $CAN(G, g)$.

Corollary 2.2 [10]. *For all positive integers g and all graphs G,*

$$CAN(K_{\omega(G)}, g) \leq CAN(G, g) \leq CAN(K_{\chi(G)}, g).$$

A k-partition of an n-set is a set of k disjoint non-empty classes whose union is the n-set. Throughout the paper, let P_k^n denote the set of all k-partitions of an n-set. We say that a vector $x \in \mathbb{Z}_g^n$ corresponds to a g-partition $P = \{P_1, P_2, \ldots, P_g\}$ of an n-set if

$$x_i = a \text{ if and only if } i \in P_a, \text{ for all } 1 \leq i \leq n, 1 \leq a \leq g.$$

Throughout this paper, we will denote by x_p the vector corresponding to partition P. We say that partitions P and Q are *qualitatively independent* if their corresponding vectors x_P and x_Q are qualitatively independent.

Definition 2.3. *(Qualitative Independence Graph) Let n and g be positive integers with $n \geq k^2$. Define the* qualitative independence graph*, $QI(n, g)$, to be the graph whose vertex set is the set of all g-partitions of an n-set with the property that every class of the partition has size at least g. Vertices are adjacent if and only if the corresponding partitions are qualitatively independent.*

Theorem 2.4 [10]. *For a graph G and positive integers g and n, there exists a $CA(n, G, k)$ if and only if there exists a graph homomorphism $G \to QI(n, g)$.*

Corollary 2.5 [10]. *For any graph G, and any positive integer g,*

$$CAN(G, g) = \min_{n \in \mathbb{N}} \{n \ : \ G \to QI(n, g)\}.$$

Corollary 2.6 [10]. *Let G be a graph and n, g be positive integers. If there exists a $CA(n, G, g)$, then*

$$\chi(G) \leq \chi(QI(n, g)) \text{ and } \omega(G) \leq \omega(QI(n, g)).$$

3. MIXED COVERING ARRAYS ON GRAPHS

In this section, we generalize results from section 2. to include mixed alphabet sizes.

Let G and H be weighted graphs. A mapping ϕ from $V(G)$ to $V(H)$ is a *weight-restricted graph homomorphism* from G to H if ϕ is a graph homomorphism from G to H such that $w_G(v) \leq w_H(\phi(v))$, for all $v \in V(G)$. For weighted graphs G and H, if there exists a weight-restricted homomorphism from G to H then we write $G \xrightarrow{w} H$.

The *complete weighted graph* on n vertices $K_n(g_1, \dots, g_n)$ is a complete graph on n vertices with weights $g_1, \dots, g_n$ on the vertices.

In the next proof, we will use the concept of *dropping the alphabet size* of a particular row of a mixed covering array (from [11]). Let $h \geq g$. To drop the alphabet size from h to g in a row of a covering array, we replace all symbols from $\mathbb{Z}_h \backslash \mathbb{Z}_g$ in the row by arbitrary symbols from $\mathbb{Z}_g$. Any pair of qualitatively independent rows of the covering array prior to the dropping operation will remain qualitatively independent.

Theorem 3.1. *Let G and H be weighted graphs with weights $g_1, g_2, \dots, g_k$ and $h_1, h_2, \dots, h_\ell$ respectively. If there exists a weight-restricted graph homomorphism $\phi : G \xrightarrow{w} H$ then $CAN(G, \prod_{i=1}^{k} g_i) \leq CAN(H, \prod_{j=1}^{\ell} h_j)$.*

Proof. Let C^H be a $CA(n, H, \prod_{j=1}^{\ell} h_j)$. The covering array C^H will be used to construct C^G, a $CA(n, G, \prod_{i=1}^{k} g_i)$. Let $x \in \{1, 2, \dots, k\}$ be the index of a row of C^G, and let v_x be the corresponding vertex in G. Row x of C^G is constructed from the row corresponding to $\phi(v_x)$ in C^H by dropping its alphabet from $w_H(\phi(v_x))$ to $w_G(v_x)$. Now, for any edge $\{v_x, v_y\}$ in G, the pair $\{\phi(v_x), \phi(v_y)\}$ is an edge in H and rows of C^H corresponding to $\phi(v_x)$ and $\phi(v_y)$ are qualitatively independent. Since dropping the alphabet size preserves qualitative independence, rows x and y of C^G are qualitatively independent. $\square$

The weight-restricted homomorphism defined above gives more than just an upper bound on the size of a mixed covering array on a graph, it also describes a construction for the array.

The next corollary generalizes Corollary 2.2 to the mixed case. For any graph G, there exist homomorphisms between the following graphs $K_{\omega(G)} \rightarrow G \rightarrow K_{\chi(G)}$. These homomorphisms can be extended to weight-restricted homomorphisms and we get the following lower and upper bounds on $CAN(G, \prod_{i=1}^{k} g_i)$.

Corollary 3.2. *Let G be a weighted graph with k vertices and $g_1 \leq g_2 \leq \cdots \leq g_k$ be positive weights. Then,*

$$CAN(K_{\omega(G)}, \prod_{i=1}^{\omega(G)} g_i) \leq CAN(G, \prod_{j=1}^{k} g_j) \leq CAN(K_{\chi(G)}, \prod_{\ell=k-\chi(G)+1}^{k} g_\ell).$$

Proof. Let G be a weighted graph with positive weights $g_1 \leq g_2 \leq \cdots \leq g_k$. From Section 2., we know there exists a graph homomorphism $\phi : K_{\omega(G)} \rightarrow G$. Assign weights $g_1, \dots, g_{\omega(G)}$ to $K_{\omega(G)}$ so that $w_{K_{\omega(G)}}(u) \leq w_{K_{\omega(G)}}(v)$ whenever $w_G(\phi(u)) \leq w_G(\phi(v))$. Since the g_i's are ordered by weight, it is clear that ϕ is a weight-restricted homomorphism. Similarly, from Section 2., there exists a graph homomorphism $\varphi : G \rightarrow K_{\chi(G)}$. Assign weights $g_{k-\chi(G)+1}, \dots, g_k$ to $K_{\chi(G)}$ so that

$w_{K_{\chi(G)}}(\varphi(u)) \leq w_{K_{\chi(G)}}(\varphi(v))$ whenever $w_G(u) \leq w_G(v)$. It is clear that φ is a weight-restricted homomorphism.
$\square$

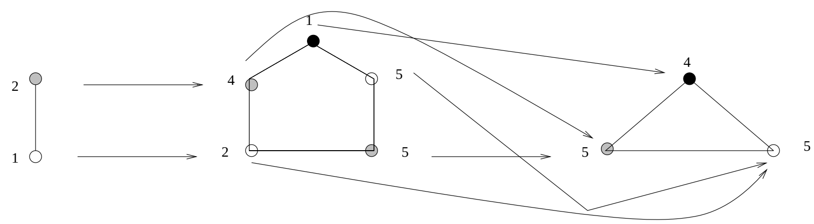

FIG. 1. Illustrating the fact: $K_{\omega(G)} \xrightarrow{w} G \xrightarrow{w} K_{\chi(G)}$

Definition 3.3. *(Mixed Qualitative Independence Graph) Let n and $g_1 < \cdots < g_\ell$ be positive integers. The* mixed qualitative independence graph $QI(n, \prod_{i=1}^{\ell} g_i)$ *is the graph whose vertex set is $P_{g_1}^n \cup P_{g_2}^n \cup \cdots \cup P_{g_\ell}^n$ and such that two vertices are adjacent if and only if their corresponding partitions are qualitatively independent.*

Note that for all $i \in \{1, \ldots, \ell\}$, the graph $QI(n, g_i)$ is an induced subgraph of $QI(n, \prod_{i=1}^{\ell} g_i)$ and if $\ell = 1$ then $QI(n, \prod_{i=1}^{\ell} g_i) = QI(n, g_1)$.

Lemma 3.4. *Let n and $g_1 < g_2 < \cdots < g_\ell$ be positive integers and let $r_i = |P_{g_i}^n|$. Then, $CAN(QI(n, \prod_{i=1}^{\ell} g_i), \prod_{i=1}^{\ell} g_i^{r_i}) \leq n$.*

Proof. Build a $CA(n, QI(n, \prod_{i=1}^{\ell} g_i), \prod_{i=1}^{\ell} g_i^{r_i})$, by assigning to the row corresponding to vertex v (partition P_v), the vector x_{P_v}. $\square$

The next theorem relates the existence of mixed covering arrays on graphs to the mixed qualitative independence graphs via weight-restricted graph homomorphisms.

Theorem 3.5. *For a weighted graph G and positive integers n and $g_1, g_2, \ldots, g_k$, there exists a $CA(n, G, \prod_{i=1}^{k} g_i)$ if and only if there exists a weight-restricted graph homomorphism $G \xrightarrow{w} QI(n, \prod_{i=1}^{k} g_i)$.*

Proof. Assume that there exists a $CA(n, G, \prod_{i=1}^{k} g_i)$; call it C. For any v in $V(G)$, denote by $C_v \in \mathbb{Z}_{g_i}^n$ the row in C corresponding to v and by P_v the partition corresponding to C_v. Consider a mapping $\phi : V(G) \rightarrow V(QI(n, \prod_{i=1}^{k} g_i))$ such that $\phi(v) = P_v$. The map ϕ is a weight-restricted homomorphism. To see this, let $v, w \in V(G)$ be adjacent vertices. Since C is a mixed covering array on G, rows C_v and C_w are qualitatively independent, thus the corresponding partitions P_v and P_w are qualitatively independent and adjacent in $QI(n, \prod_{i=1}^{k} g_i)$.

Conversely, assume there is a weight-restricted graph homomorphism $\phi : G \xrightarrow{w} QI(n, \prod_{i=1}^{k} g_i)$. We build C, a $CA(n, G, \prod_{i=1}^{k} g_i)$, as follows. For each $v \in V(G)$, $\phi(v)$ corresponds to a partition P_v. Take row C_v to be the vector corresponding to P_v. If the vertices $v, w \in V(G)$ are adjacent in G then the partitions $\phi(v), \phi(w)$ are qualitatively independent, thus the corresponding rows C_v, C_w are qualitatively independent. □

Corollary 3.6. *Let G be a weighted graph with distinct weights $g_1, g_2, \dots, g_r$, repeated $s_1, s_2, \dots, s_r$ times, respectively. Then,*

$$CAN(G, \prod_{i=1}^{r} g_i^{s_i}) = \min_{n \in \mathbb{N}} \{n : G \xrightarrow{w} QI(n, \prod_{i=1}^{r} g_i)\}.$$

The next corollary can be used to find a lower bound on the mixed covering array number. Specifically, we conclude that $CAN(G, \prod_{i=1}^{r} g_i^{s_i}) > n$ whenever $\omega(G) > \omega(QI(n, \prod_{i=1}^{r} g_i))$ or $\chi(G) > \chi(QI(n, \prod_{i=1}^{r} g_i))$.

Corollary 3.7. *Let n be a positive integer and let G be a weighted graph with distinct weights $g_1, g_2, \dots, g_r$, repeated $s_1, s_2, \dots, s_r$ times, respectively. Then, If there exists a $CA(n, G, \prod_{i=1}^{r} g_i^{s_i})$ then*

$$\chi(G) \le \chi(QI(n, \prod_{i=1}^{r} g_i)) \text{ and } \omega(G) \le \omega(QI(n, \prod_{i=1}^{r} g_i)).$$

Proof. By the previous Theorem 3.5, there exists a weight-restricted homomorphism from G to $QI(n, \prod_{i=1}^{r} g_i)$. A weight-restricted homomorphism is a homomorphism, therefore we get that $\chi(G) \le \chi(QI(n, \prod_{i=1}^{r} g_i))$ and $\omega(G) \le \omega(QI(n, \prod_{i=1}^{r} g_i))$ by the properties of homomorphisms. □

4. OPTIMAL MIXED COVERING ARRAYS ON GRAPHS

In this section, we give constructions of mixed covering arrays on graphs.

A length-n vector with alphabet size g is *balanced* if each symbol occurs $\left\lfloor \frac{n}{g} \right\rfloor$ or $\left\lceil \frac{n}{g} \right\rceil$ times. A *balanced covering array* is a covering array in which every row is balanced.

Lemma 4.1. *For a balanced length-n vector v on $\mathbb{Z}_g$ and for any $h \le \frac{n}{g}$ there exists a balanced length-n vector w on $\mathbb{Z}_h$ such that v and w are qualitatively independent.*

Proof. Let $n = gk + r$, with $0 \le r < g$. Since v is balanced, we can assume the first $g - r$ symbols each occur k times in v and the last r symbols occur $k+1$ times. Further, we can assume the letters occur in their natural order. That is,

$$v(i) = \begin{cases} \left\lfloor \frac{i}{k} \right\rfloor & \text{for } 0 \le i < (g-r)k, \\ \left\lfloor \frac{i+(g-r)}{k+1} \right\rfloor & \text{for } (g-r)k \le i < n. \end{cases}$$

Build the vector $w \in \mathbb{Z}_g^n$ by $w(i) \equiv i \mod h$. Clearly, w is balanced. We will now show that v and w are qualitatively independent. For $s \in \{0, 1, \dots, g\}$, s

occurs in positions $x, \ldots, x+k$ for some $x \in [0, n]$ in v. The entries of w in these positions are $i \mod h$. Since $h \leq \frac{n}{g}$ then $k \geq h$. Thus, all symbols in the alphabet corresponding to w will occur in these positions. This means we cover all possible pairs with s. Since we can do this for all $s \in \{0, 1, \ldots, g\}$, all pairs are covered between v and w. $\square$

Lemma 4.2. *Let $v_1 \in \mathbb{Z}_{g_1}^n$ and $v_2 \in \mathbb{Z}_{g_2}^n$ be balanced vectors, then for any h with the property that $hg_1 \leq n$ and $hg_2 \leq n$, there exists a balanced vector $w \in \mathbb{Z}_h^n$ such that v_1 and w are qualitatively independent and v_2 and w are qualitatively independent.*

Proof.

Define a bipartite multi-graph H as follows: we have g_1 vertices in the first part, $P \subseteq V(H)$ and g_2 vertices in the second part $Q \subseteq V(H)$. Let $P_a = \{i | v_1(i) = a\}$, for $a = 1, \ldots, g_1$, be the vertices of P, while $Q_b = \{i | v_2(i) = b\}$, for $b = 1, \ldots, g_2$, be the vertices of Q. We also have that $|P_a| \geq \left\lfloor \frac{n}{g_1} \right\rfloor$ and $|Q_b| \geq \left\lfloor \frac{n}{g_2} \right\rfloor$ since the vectors are balanced. For each $i = 1, \ldots, n$ there exists exactly one P_a and one Q_b with $i \in P_a$ and $i \in Q_b$. For each such i, add an edge between vertices corresponding to P_a and Q_b and label it i. Since $|P_a| \geq \left\lfloor \frac{n}{g_1} \right\rfloor$ and $|Q_b| \geq \left\lfloor \frac{n}{g_2} \right\rfloor$, the minimum degree is $\delta = \min\{\left\lfloor \frac{n}{g_1} \right\rfloor, \left\lfloor \frac{n}{g_2} \right\rfloor\}$. By a dual of Konig's theorem, there are $\min\{\left\lfloor \frac{n}{g_1} \right\rfloor, \left\lfloor \frac{n}{g_2} \right\rfloor\}$ classes of edge-disjoint vertex covers. In particular since $h \leq \min\{\left\lfloor \frac{n}{g_1} \right\rfloor, \left\lfloor \frac{n}{g_2} \right\rfloor\}$, we have h edge-disjoint vertex covers. Use these h edge-disjoint vertex covers which form partition R to build our balanced vector $w \in \mathbb{Z}_h^n$. Each edge-disjoint vertex cover corresponds to one of the symbols $0, 1 \ldots, h-1$. Each edge corresponds to a position of a length-n vector. For each edge i in an edge-disjoint vertex cover, assign to $w(i)$, which is the i-th position in vector w, the letter that corresponds to this edge-disjoint vertex cover. Fill the remaining positions in the vector in such a way that the vector remains balanced. Finally, we need to show that the partitions P and R are qualitatively independent. For any $j = 1, \ldots, h$, the class R_j corresponds to an edge-disjoint vertex cover of H, this means that for any $i \in \{1, \ldots, h\}$, there exists an edge a in the edge-disjoint vertex cover incident to P_i. By the definition of R, we have $a \in R_j$. Since the edge labelled a is incident with the vertex corresponding to P_i, we also know that $a \in P_i$. This means that for any $i, j \in \{1, \ldots, n\}, a \in P_i$ and $a \in R_j$ so $P_i \cap R_j \neq \emptyset$. Similarly, we can conclude that Q and R are qualitatively independent. $\square$

4.1 Basic Graph Operations

Let G be a weighted multigraph with k vertices. Label the vertices $v_0, v_1, \ldots, v_{k-1}$ and for each vertex v_i the associated weight is denoted by $w_G(v_i)$. Let the *product weight* of G, denoted $PW(G)$, be $PW(G) = \max\{w_G(v_i)\, w_G(v_j) : \{v_i, v_j\} \in E(G)\}$. All the definitions for graphs extend naturally to multigraphs. Note that $CAN(G, \prod_{i=1}^{k} w_G(v_i)) \geq PW(G)$.

We define three graph operations: *one-vertex edge hooking*, *edge duplication* and the *weight-restricted edge subdivision*. We will see that these operations will have no effect on the covering array number of the modified graphs. A *one-vertex edge hooking* in a multigraph G is the operation that inserts a new edge where one end is in $V(G)$ and the other is a new vertex. An *edge duplication* involves the creation of an edge that is parallel to an existing edge in G, that is, we are creating a multigraph with an existing edge appearing twice. An *edge subdivision* is the operation that replaces an edge by a path with two edges. A *weight-restricted edge subdivision* is an edge subdivision such that if x is the new vertex in G adjacent to vertices y and z then $w_G(x)w_G(y) \leq PW(G)$ and $w_G(x)w_G(z) \leq PW(G)$.

Proposition 4.3. *One-vertex Edge Hooking*
Let G be a weighted multigraph with k vertices. Let G' be the weighted multigraph obtained from G by a one-vertex edge hooking of a new vertex x with a new edge $\{x, y\}$, and $w(x)$ such that $w(x)w(y) \leq PW(G)$. Then, there exists a balanced $CA(n, G, \prod_{i=1}^{k} g_i)$ if and only if there exists a balanced $CA(n, G', w(x)\prod_{i=1}^{k} g_i)$.

Proof. The direct implication is the only non-trivial one. Let C^G be a balanced $CA(n, G, \prod_{i=1}^{k} g_i)$. The balanced covering array C^G will be used to construct a $C^{G'}$, a balanced $CA(n, G', w(x)\prod_{i=1}^{k} g_i)$. Using Lemma 4.1, we know that we can build a balanced length-n vector, call it X, corresponding to vertex x such that X is qualitatively independent with the length-n vector Y corresponding to vertex y in G. $C^{G'}$ is built by appending row X to C^G. Since the only new edge is $\{x, y\}$, and X and Y are qualitatively independent, $C^{G'}$ is a balanced mixed covering array on G'. □

Proposition 4.4. *Edge Duplication*
Let G be a weighted graph with k vertices and let G' be the weighted multigraph obtained from G by duplicating one of its edges. Then, there exists a balanced $CA(n, G, \prod_{i=1}^{k} g_i)$ if and only if there exists a balanced $CA(n, G', \prod_{i=1}^{k} g_i)$.

Proof. The extension of the definition of mixed covering arrays on graphs to mixed covering arrays on multigraphs does not affect the size of a mixed covering array obtained by duplicating edges. The balanced mixed covering array on G is the same balanced mixed covering array on G'. □

Proposition 4.5. *Weight-Restricted Edge Subdivision*
Let G be a weighted multigraph with k vertices. Let G' be the weighted multigraph obtained from G by a weight-restricted edge subdivision creating the new vertex x that is adjacent to vertices y and z in G, with $w(x)$ such that $\max\{w(x)w(y), w(x)w(z)\} \leq PW(G)$. Then, there exists a balanced $CA(n, G, \prod_{i=1}^{k} g_i)$ if and only if there exists a balanced $CA(n, G', w(x)\prod_{i=1}^{k} g_i)$.

Proof. Let C^G be a balanced $CA(n, G, \prod_{i=1}^{k} g_i)$. The balanced covering array C^G will be used to construct a $C^{G'}$, a balanced $CA(n, G', w(x)\prod_{i=1}^{k} g_i)$. By Lemma 4.2, we know that we can build a balanced length-n vector, call it X, corresponding to vertex x such that X is qualitatively independent with the length-n vector Y corresponding to vertex y in G and with the length-n vector Z corresponding to vertex z in G. $C^{G'}$ is built by appending row X to C^G. Since $\{x,y\}$ and $\{x,z\}$ are the only new edges and X and Y are qualitatively independent and X and Z are qualitatively independent, $C^{G'}$ is a balanced mixed covering array on G'. $\square$

The next Theorem can be used to derive an algorithm that reverses the basic operations on a graph, obtaining a simpler graph to be used in a covering array construction.

Theorem 4.6. *Let G be a weighted multigraph and G' a weighted multigraph obtained from G via a sequence of weight-restricted edge subdivisions, one-vertex edge hooking and edge duplication. Let $v_{k+1}, v_{k+2}, \ldots, v_\ell$ be the vertices in $V(G')\backslash V(G)$ with weights $g_{k+1}, g_{k+2}, \ldots, g_\ell$, respectively. Then, there exists a balanced $CA(n, G, \prod_{i=1}^k g_i)$ if and only if there exists a balanced $CA(n, G', \prod_{i=1}^\ell g_i)$.*

Proof. The result is derived by iterating the three previous propositions. $\square$

Example: Let T be a weighted tree with k vertices. We can build T by starting with an edge $\{v_i, v_j\}$ such that $PW(T) = w(v_i) \times w(v_j)$, and by applying successive one-vertex edge hooking in the proper order so as to obtain T. So, $CAN(T, \prod_{i=1}^k g_i) = PW(T)$. $\square$

4.2 n-chromatic graphs for $n = 3, 4, 5$.

Combining Corollary 3.2 with previously known mixed covering array numbers, we get the following theorem.

Theorem 4.7. *Let G be a weighted graph with k vertices with weights $g_1 \leq g_2 \leq \ldots \leq g_k$. If one of the following holds:*
1) $\chi(G) = 3$,
2) $\chi(G) = 4$ *and* $\prod_{i=k-3}^k g_i \notin \{2^4, 6^4\}$, *or*
3) $\chi(G) = 5$ *and* $\prod_{i=k-4}^k g_i \notin \{2^5, 3^5, 23^4\}$ *and* $g_{k-1} \notin \{4, 6, 10\}$,

then $CAN(G, \prod_{i=1}^k g_i) \leq g_{k-1} g_k$.

Proof. Let G be a weighted graph with k vertices. Then by Corollary 3.2, we have $CAN(G, \prod_{i=1}^k g_i) \leq CAN(K_{\chi(G)}, \prod_{l=k-\chi(G)+1}^k g_l)$. Moreover, by Theorems 4.2, 4.3 and 4.4 in [11] for $p = 3$; $p = 4$ and $\prod_{i=k-3}^k g_i \notin \{2^4, 6^4\}$; or $p = 5$ and $\prod_{i=k-4}^k g_i \notin \{2^5, 3^5, 23^4\}$ and $g_{k-1} \notin \{4, 6, 10\}$, we know that $CAN(K_p, \prod_{l=k-p+1}^k g_l) = g_{k-1} g_k$. $\square$

4.3 Cycles

In this section, we solve the problem for cycles. Let C be a cycle with two largest weights g_{k-1} and g_k. Note that if $PW(C) = g_{k-1} g_k$, since cycles are 3-chromatic, Theorem 4.7 solves the problem via the colouring construction (graph homomorphism to K_3). So, the non-trivial case we solve in this section is when $PW(C) < g_{k-1} g_k$.

Theorem 4.8. *Let C be a weighted cycle of length k. There exists a balanced mixed covering array $CA(n, C, \prod_{i=1}^{k} w_C(v_i))$ with $n = PW(C)$. Moreover, this covering array is optimal.*

Proof. Let $\{v_{k-1}, v_k\}$ be an edge in C with $w(v_{k-1})w(v_k) = PW(C)$. Note that in this proof we do not assume that $w(v_{k-1})$ and $w(v_k)$ are the two largest weights. We will build a balanced mixed $CA(n, C, \prod_{i=1}^{k} w_C(v_i))$ with $n = PW(C)$ using the basic graph operations from Section 4.1. Let G_1 be the weighted graph with the single weighted edge $\{v_{k-1}, v_k\}$. From Lemma 4.1, there exists a $CA(n, G, w(v_{k-1})w(v_k))$. Let G_2 be the graph obtained from G_1 by edge duplication. From Proposition 4.4 there exists a $CA(n, G_2, \prod_{i=k-1}^{k} w_C(v_i))$. Let G_3 be the graph obtained from G_2 by applying a weight-restricted edge subdivision with vertex v_1, a minimum-weight vertex among the remaining vertices, and inserting v_1 in the appropriate position based on the original cycle configuration. We construct G_j by applying successive weight-restricted edge subdivisions with the vertex v_{j-2}, one of minimal weight among the remaining vertices, inserted in the appropriate position. Note that $G_k = C$. From Proposition 4.5, for all $j = 3, \ldots, k$ if $PW(G_j) \leq PW(C)$, then there exists a $CA(PW(C), G_j, w_C(v_{k-1})w_C(v_k) \prod_{i=1}^{j-2} w_C(v_i))$. In particular, there exists a $CA(PW(C), C, \prod_{i=1}^{k} w_C(v_i))$. All we need to prove is that $PW(G_j) \leq PW(C)$ for all $j \in \{3, \ldots, k\}$.

To prove this we will assume that for some $j \in \{3, \ldots, k\}$, $PW(G_j) > PW(C)$. Assume at some intermediate step that we have an edge $\{u, v\}$ in the cycle we are building with $w(u) = g$ and the $w(v) = h$ where $gh > PW(C)$. We know from the definition of $PW(C)$ that the vertices u, v are not adjacent in the original cycle. Let $up_1p_2 \ldots p_i v$ be the path in the original cycle that connects u to v and does not contain the edge $\{v_{k-1}, v_k\}$. Then, the vertex p_1 is not in the cycle we are constructing since $gw(p_1) \leq PW(C) < gh$. So, we have that $w(p_1) < h$. This is a contradiction since we have inserted the vertices with the smallest weight first.

This covering array is optimal since $PW(C)$ is a lower bound for the size of a covering array on C. □

4.4 Bipartite Graphs

In this section, we give a construction for optimal mixed covering arrays on bipartite graphs. A *bipartite* graph is one whose vertex set can be partitioned into two sets X and Y, so that each edge has one end in X and one end in Y. Note that if the two largest alphabets in the bipartite graph, denoted by g_k and g_{k-1}, are connected by an edge then we can build an optimal mixed covering array of size $g_k g_{k-1}$ via a weight-restricted homomorphism to K_2. The non-trivial case solved by the next theorem is when $PC(G) < g_k g_{k-1}$.

Theorem 4.9. *Let G be a weighted bipartite graph with k vertices. Then, there exists a balanced $CA(n, G, \prod_{i=1}^{k} w_G(v_i))$ with $n = PW(G)$. Moreover, this covering array is optimal.*

Proof. Let G be a bipartite graph with bipartitions A and B. Denote the vertices of A by $a_1, a_2, \ldots, a_{|A|}$ and their corresponding weights by $w_G(a_i)$ where

$i \in \{1,\ldots,|A|\}$. Similarly, denote the vertices of B by $b_1, b_2, \ldots b_{|B|}$ and their corresponding weights by $w_G(b_i)$ where $i \in \{1,\ldots,|B|\}$. Set $n = PW(G)$. We will give a construction for a $CA(n, G, \prod_{i=1}^{k} w_G(v_i))$ and call this covering array C.

For a vertex $a_i \in A$, where $i \in \{1,\ldots,|A|\}$, let r_{a_i} be the row in C corresponding to a_i. Set the j^{th} entry of r_{a_i} to be equivalent to $j \mod w_G(a_i)$. For a vertex $b_i \in B$, where $i \in \{1,\ldots,|B|\}$, let r_{b_i} be the row in C corresponding to b_i. Set the j^{th} entry of r_{b_i} to be

$$r_{b_i}(j) = \begin{cases} \left\lfloor \frac{j}{\left\lfloor \frac{n}{w_G(b_i)} \right\rfloor} \right\rfloor & \text{if } 0 \leq j < w_G(b_i) \left\lfloor \frac{n}{w_G(b_i)} \right\rfloor, \\ j \mod w_G(b_i) & \text{otherwise.} \end{cases}$$

We will now show that for any two adjacent vertices a and b that the rows r_a and r_b in C corresponding to a and b are qualitatively independent. Since G is bipartite, we can assume that $a \in A$ and $b \in B$. Further, $w_G(a)w_G(b) \leq n$. Thus $\left\lfloor \frac{n}{w_G(b)} \right\rfloor \geq w_g(a)$. For $s \in \{0,1,\ldots,w_G(b)-1\}$, s occurs in positions $x,\ldots,x+\left\lfloor \frac{n}{w_G(b)} \right\rfloor - 1$ of r_b for some $x \in [0, w_G(b_i)\left\lfloor \frac{n}{w_G(b_i)} \right\rfloor]$. The entries of r_a in positions $j \in \{x,\ldots,x+\left\lfloor \frac{n}{w_G(b)} \right\rfloor - 1\}$ are $j \mod w_G(a)$. Since $\left\lfloor \frac{n}{w_G(b)} \right\rfloor \geq w_G(a)$, all symbols in the alphabet corresponding to a will occur in these $\left\lfloor \frac{n}{w_G(b)} \right\rfloor$ positions. This means we cover all possible pairs with s. Since we can do this for all $s \in \{0,1,\ldots,w_G(b)-1\}$, all pairs are covered between r_a and r_b.

This covering array is optimal, since $PW(G)$ is a lower bound on the size of a covering array on G.

□

In conclusion, if a graph G is a tree, a cycle or a bipartite graph, then its mixed covering array number is $PW(G)$.

REFERENCES

[1] J.N. Cawse (ed.). Experimental Design for Combinatorial and High Throughput Materials Development, John Wiley & Sons, New York, 2003.

[2] C. Cheng, A. Dumitrescu and P. Schroeder. "Generating small combinatorial test suites to cover input-output relationships", Proc. of the Third Intern. Conf. on Quality Software (QSIC 03), 2003, p. 76-82.

[3] D.M. Cohen, S.R. Dalal, M.L. Fredman, and G.C. Patton. "The AETG system: an approach to testing based on combinatorial design", IEEE Trans. on Software Engineering, 23 (1997), p. 437-444.

[4] C.J. Colbourn. "Combinatorial aspects of covering arrays", Le Matematiche (Catania), to appear.

[5] C. J. Colbourn, S. S. Martirosyan, G. L. Mullen, D. E. Shasha, G. B. Sherwood, and J. L. Yucas, " Products of mixed covering arrays of strength two", Journal of Combinatorial Designs. to appear.

[6] S.R. Dalal, A. Jain, G.C. Patton, M. Rathi, and P. Seymour. "AETG Web: A web-based service for automatic efficient test generation from functional requirements", Proc. 2nd IEEE Worksh. on Industrial Strength Formal Specif. Techn., IEEE Press, 1998, p. 84-85.

[7] S.R. Dalal and C.L. Mallows. “Factor-covering designs for testing software”, Technometrics 40 (1998), p. 234-243.

[8] A. Hartman and L. Raskin. “Problems and algorithms for covering arrays”, Discrete Math. 284 (2004), p. 149-156.

[9] K. Meagher, “Covering Arrays on Graphs: Qualitative Independence Graphs and Extremal Set Partition Theory”, PhD thesis, University of Ottawa, Ottawa, 2005.

[10] K. Meagher, B. Stevens, “Covering Arrays on Graphs,” J. Combin. Theory. Ser. B 95 (2005), p. 134-151.

[11] L. Moura, J. Stardom, B. Stevens, A. Williams, “Covering arrays with mixed alphabet sizes,” J. Combin. Des. 11(2003), p. 413-432.

[12] G. Seroussi and N.H. Bshouty. “Vector sets for exhaustive testing of logic circuits”, IEEE Trans. on Infor. Theory 34 (1988), p. 513-522.

[13] D.E. Shasha, A.Y. Kouranov, L.V. Lejay, M.F. Chou, and G.M. Coruzzi. “Using combinatorial design to study regulation by multiple input signals: a tool for parsimony in the post-genomics era”, Plant Physiology 127 (2001), p. 1590-1594.

[14] A.W. Williams and R.L. Probert. “A practical strategy for testing pair-wise coverage of network interfaces”, Proc. Seventh Intern. Symp. on Software Reliability Engineering, 1996, p. 246-54.

Received
Accepted