Combinatorial Group Testing in Bipartite Graphs

C. A. Rodger[1]
and
Meredith Roy[2]
221 Parker Hall
Auburn University, AL
USA 36849-5310

Abstract

The method of large scale group testing has been used in the economical testing of blood samples, and in non-testing situations such as experimental designs and coding theory, for over 50 years. Some very basic questions adressing the minimum number of tests required to identify defective samples still remain unsolved, including the situation where one defective sample in each of two batches are to be found. This gives rise to an intriguing graph theoretical conjecture concerning bipartite graphs, a conjecture which in this paper is proved to be true in the case where vertices in one part of the bipartite graph have low degree.

Keywords: group testing, bipartite, small degree

1 Introduction

Using group testing to economically test blood samples on a large scale began 50 years ago. Since then, the method has been used in industrial settings, as well as in non-testing situations such as experimental designs and coding theory. The mathematical literature contains many results in the area, but an excellent description of the method can be found in [5].

It is very interesting to note that some very basic questions in this area still remain open. For example, despite much attention (see, for example, [5, 6]), it is unknown whether the minimum number of tests needed to identify 2 defective samples among a batch of v samples is $\lceil \log_2(v(v-1)/2) \rceil$ or $\lceil \log_2(v(v-1)/2) \rceil + 1$ [2]. More general results have been obtained in the case where only a subset of all pairs of samples could contain the 2 defective ones [4]. In this paper we address this more general problem in the case where the batch of samples has been collected from two distinct locations, the assumption being that each location will provide at most one defective sample. Clearly a lower bound on the minimum number of tests required to identify the two defective samples in this situation is

[1] rodgec1@auburn.edu
[2] meredity.roy@gmail.com

$\lceil \log_2(x_1x_2) \rceil$, where x_i samples have been collected from the i^{th} location. Determining this minimum $\mu(x_1, x_2)$ is also an open problem; it is conjectured to be $\lceil \log_2(x_1x_2) \rceil$.

It is noted in [5], and reiterated by Frank Hwang at the 2005 conference in Taiwan held in his honor, that this problem can be modeled by a bipartite graph, the vertices in each part representing samples from one of the two locations. The two defective samples are therefore represented by an edge joining the corresponding vertices; so identifying this edge is the objective. Each test involves determining whether or not a selected set of samples contains at least one defective sample; so each test checks whether or not at least one of the edges covering a selected set of vertices (samples) contains a defective sample. Once one notes that the complement of this set of edges is an induced subgraph, and observes that adding "dummy" edges to the bipartite graph until the number of edges is a power of 2 will not change the conjectured value of $\mu(x_1, x_2)$, it can be seen that showing $\mu(x_1, x_2) = \lceil \log_2(x_1x_2) \rceil$ can be settled by proving the following conjecture [1].

Conjecture 1.1 *In any bipartite graph B with 2^m edges, there exists an induced subgraph with exactly 2^{m-1} edges.*

Recently this conjecture has been proved true for all bipartite graphs with at most 32 edges [6], making use of computer searches. In the same paper, a very nice theoretical result was obtained by Juan and Chang, proving a related result that if B is a bipartite graph in which for some $k \geq 6$ the number of edges lies between 2^{k-1} and $2^{k-1} + 2^{k-3} + 2^{k-4} + 2^{k-5} + 2^{k-6} + 2^{k-7} + 27(2^{(k-8)/2}) - 1$, then B contains an induced subgraph S such that both S and $B - S$ contain at most 2^k edges. For more history on using this approach, see [3, 4, 6].

In this paper we provide theoretical evidence supporting Conjecture 1.1, proving in Theorem 2.1 that this conjecture is true in the infinitely many cases where all vertices in one part of B have small degree. Apart from being of interest in its own right, this result is then used in Section 3 to provide an easy proof, without computer aid, that Conjecture 1.1 is true for all bipartite graphs on at most 16 edges.

All set notation used here denotes multisets, and $\{a^b\}$ is used to denote the multiset containing b copies of a.

2 The Main Result

In this section we deal completely with the myriad of cases arising when all vertices in one part of the bipartite graph have very low degree.

Theorem 2.1 *Let B be a bipartite graph with bipartition $\{X, Y\}$ of the vertex set. Suppose $|E(B)| = 2^m$ with $m \geq 2$. If $d(x) \leq 4$ for all $x \in X$ then B contains an induced subgraph containing exactly 2^{m-1} edges.*

Proof: Let $D = D(X)$ be the multiset consisting of the degrees of the vertices in X. Let $P = P(X) = \{S_1, \ldots, S_s, T_1, \ldots, T_t, R\}$ be a partition of D (possibly $R = \emptyset$) for which:

(a) $s_j = \sum_{i \in S_j} i = 4$ for $1 \leq j \leq s$, and s is as large as possible,

and subject to (a), we also have:

(b) $T_j = \{3, 3, 3, 3\}$ for $1 \leq j \leq t$, and t is as large as possible.

Let $t_j = \sum_{i \in T_j} i = 12$ for $1 \leq j \leq t$.

Maximizing s, then t, forces several things to occur, as we now note. Let $d = \sum_{i \in D} i = 2^m$.

(1) $1 \notin R$. To see this, notice that since d, each s_j, and each t_j are all even, $r = \sum_{i \in R} i$ is even. So if $1 \in R$ then either $\{1, 3\} \subseteq R$ or $\{1, 1\} \subseteq R$ (since R contains an even number of odd numbers). But the maximality of s prevents $\{1, 3\} \subseteq R$. Since 4 divides each s_j, each t_j, and $d = 2^m$, 4 divides r. So $\{1\} \subseteq R$ and $3 \notin \{R\}$ implies $\{1, 1, 2\} \subseteq R$ or $\{1, 1, 1, 1\} \subseteq R$, both contradicting the maximality of s.

(2) $4 \notin R$, R contains 2 at most once, and R contains 3 at most three times. These observations follow from the maximality of s, of s, and of t respectively.

(3) $R = \{2, 3, 3\}$ or $R = \emptyset$. This follows from $(1 - 2)$ since 4 divides r.

With the partition P in hand, we now turn to the proof. Notice that since $d = 2^m \geq 4$, $d \equiv 4$ or $8 \pmod{12}$. We consider each case in turn.

Case 1 Suppose $d \equiv 8 \pmod{12}$.

In view of (3), either $s = 0$ and $R = \{2, 3, 3\}$, or $s \geq 2$ (since $t_j = 12$ for $1 \leq j \leq t$).

First suppose $s \geq 2$. Since 12 divides $\sum_{i \in D \setminus (S_1 \cup S_2)} i$, and since 12 divides $2^{m-1} - 4$, there exists a set $D_1 \subseteq D \setminus (S_1 \cup S_2)$ such that $\sum_{i \in D_1} i = 2^{m-1} - 4$ (since each $s_j = 4$ and each $t_j = 12$). Let $V(D_1 \cup S_1)$ be a set of vertices in X whose degrees form the multiset $D_1 \cup S_1$. Then $V(D_1 \cup S_1) \cup N(D_1 \cup S_1)$ induces a subgraph with 2^{m-1} edges.

So now suppose that $s = 0$ and $R = \{2, 3, 3\}$. Then $D = \{2, 3^{4t+2}\}$. Let v be the vertex of degree 2 in X. Let $S \subseteq X$ be a set of $2t + 1$ vertices of degree 3. Let $\bar{S} = X \setminus (S \cup \{v\})$. If $w \in \{v\} \cup \bar{S}$ has exactly 1 or 0 neighbors in $N(S)$ then $S \cup N(S) \cup \{w\}$ together with 0 or 1 neighbors of w respectively induces the required subgraph; so assume this is not the case. If $w \in \bar{S}$ has 2 neighbors in $N(S)$ then $(S \cup \{v, w\} \cup N(S)) \setminus \{s\}$ for any $s \in S$ induces the required subgraph.

So we can assume that $N(v) \subseteq N(S)$, that $N(\bar{S}) \subseteq N(S)$, and by the symmetric argument that $N(S) \subseteq N(\bar{S})$. Therefore we can assume $Y = N(S) = N(\bar{S})$.

Now let $y \in Y$ be such that y is adjacent to v, y has $\alpha \geq 1$ neighbors in S, and y has at least α neighbors in $\bar{S}$ (swapping the roles of S and $\bar{S}$ if necessary). Form a new set T from $S \cup N(S)$ by removing y, adding $\lceil \alpha/2 \rceil$ neighbors of y in $\bar{S}$, and adding v if α is even. Removing y from $S \cup N(S)$ removes α edges from the induced subgraph; each neighbor of y in $\bar{S}$ has exactly 2 edges joining it to vertices in $N(S)\backslash\{y\}$; and, v has exactly one neighbor in $N(S)\backslash\{y\}$. Since $S \cup N(S)$ induces a graph with $2^{m-1} - 1$ edges, T induces a graph in which the number of edges is:

$$(2^{m-1} - 1) - a + \begin{cases} 2\lceil \alpha/2 \rceil + 1 & \text{if } \alpha \text{ is even,} \\ 2\lceil \alpha/2 \rceil & \text{if } \alpha \text{ is odd,} \end{cases}$$
$$= 2^{m-1},$$

thus producing the required subgraph. So is all cases the required subgraph can be found.

Case 2 Suppose $d \equiv 4 \pmod{12}$.

The required subgraph is trivial to find if $d = 4$ (for then vertices in one of X or Y have maximum degree at most 2), so we can assume that $d \geq 16$. In view of (3), in this case if $R \neq \emptyset$ then $s \geq 2$, and if $R = \emptyset$ then $s \equiv 1 \pmod 4$. Also, $2^{m-1} \equiv 8 \pmod{12}$.

Suppose $R \neq \emptyset$ or $s \geq 4$. Let $D_1 = R \cup S_1 \cup S_2$ or $D_1 = \bigcup_{i=1}^{4} S_i$ respectively. Since 12 divides $\sum_{i \in D \backslash D_1} i$, and since 12 divides $2^{m-1} - 8$, there exists a multiset $D_2 \subseteq D \backslash D_1$ such that $\sum_{i \in D_2} i = 2^{m-1} - 8$. Therefore $\sum_{i \in D_2 \cup S_1 \cup S_2} i = 2^{m-1}$, so $V(D_2 \cup S_1 \cup S_2)$ together with its neighborhood in B induces the required subgraph.

Finally, suppose $s = 1$ and $R = \emptyset$; then $D = S_1 \cup \{3^{4t}\}$. If $S_1 = \{2, 2\}, \{1, 1, 2\}$ or $\{1, 1, 1, 1\}$ then let $D_1 = \{2, 3^{2t}\}, \{1, 1, 3^{2t}\}$, or $\{1, 1, 3^{2t}\}$ respectively. In these cases $V(D_1)$ together with its neighborhood induces the required subgraph.

Otherwise, $S_1 = \{1, 3\}$ or $S_1 = \{4\}$, so $D = \{1, 3^{4t+1}\}$ or $D = \{4, 3^{4t}\}$. Let v be the vertex of degree 1 or 4 in X, let S be a set of $2t$ vertices of degree 3 in X, and let $\bar{S}$ be a set of $2t$ vertices of degree 3 in $X \backslash S$. Note that $\sum_{i \in S} i = 6t = 2^{m-1} - 2$.

Let $w \in X \backslash S$ with $d(w) \geq 3$ (possibly $w = v$). If w has 0, 1 or 2 neighbors in $N(S)$ then $S \cup N(S) \cup \{w\}$ together with 2, 1, or 0 neighbors of w that are not in $N(S)$ induces the required subgraph.

Therefore we can assume that $N(\bar{S}) \subseteq N(S)$, and by reversing the roles of S and $\bar{S}$ we can assume that $N(\bar{S}) = N(S)$, regardless of which $2t$ vertices of degree 3 are chosen to form S, and which are chosen to form $\bar{S}$. Also, there is at most one vertex in $Y \backslash N(S)$; if $y_2 \in Y \backslash N(S)$ exists then the only neighbor it can have is v (so $d(y_2) = 1$ if it exists). Let $Y' = Y \backslash \{y_2\}$ if y_2 exists and $Y' = Y$ otherwise. Furthermore, each vertex in Y' must have a neighbor in S and a neighbor in $\bar{S}$.

Since all properties of B described in the last paragraph remain true regardless of how S and $\bar{S}$ are chosen, we now make use of this in each of the following cases.

First suppose $d(v) = 1$. Since $|Y| \geq 3$, let $y_1 \in Y \backslash N(v)$. Choose S so that y_1 has at least as many neighbors in S as in $X \backslash S \cup \{v\}$. Let α be the number of neighbors of y_1 in S. Form a set T from $S \cup N(S)$ by

(a) deleting y_1,

(b) adding $\{v\} \cup N(v)$ if α is odd, and

(c) adding $\lfloor \alpha/2 \rfloor + 1$ neighbors of y_1 in $X \backslash S$.

The result now follows since $S \cup N(S)$ induces a subgraph with $2^{m-1} - 2$ edges, (a) removes α edges, (b) adds 1 edge if α is odd, and (c) adds $2(\lfloor \alpha/2 \rfloor + 1)$ edges.

Now suppose $d(v) = 4$. At least one of the four neighbors of v, say y_1, is non-adjacent to a vertex in X (vertices in $X \backslash \{v\}$ only have degree 3). Therefore we can choose S so that y_1 has as odd number α of neighbors in S, and y_1 has at least as many neighbors in $\bar{S}$ as it does in S (since in this case X has an even number of vertices of degree 3). Form a set T from $S \cup N(S)$ by:

(a) deleting y_1,

(b) adding $\{v\} \cup N(v)$, and

(c) adding $\lfloor \alpha/2 \rfloor$ neighbors of y_1 in $\bar{S}$.

The result now follows since $S \cup N(S)$ induces a subgraph with $2^{m-1} - 2$ edges, (a) removes α edges, (b) adds 3 edges, and (c) adds $2\lfloor \alpha/2 \rfloor = \alpha - 1$ edges. □

3 $\epsilon \leq 16$

We now verify Conjecture 1.1 where $\epsilon \leq 16$, beginning with a simple, but useful, lemma.

Lemma 3.1 *Let $m \geq 2$. Let B be a bipartite graph with bipartition $\{X, Y\}$ of the vertex set. Suppose $|E(B)| = 2^m$. If there exist $u, v \in X$ with $d(u)+d(v) \geq 2^{m-1}$ then B contains an induced subgraph with exactly 2^{m-1} edges.*

Proof: Let $n = |N(u) \cap N(v)|$. If $n \geq 2^{m-2}$ then let $S \subseteq N(u) \cap N(v)$ with $|S| = 2^{m-2}$. If $n < 2^{m-2}$ then let $S = (N(u) \cap N(v)) \cup R$ where $R \subseteq (N(u) \cup N(v)) \backslash (N(u) \cap N(v))$ and $|R| = 2^{m-1} - 2n$; this is possible since $d(u) + d(v) \geq 2^{m-1}$. In either case, $\{u, v\} \cup S$ induces the desired subgraph. □

We are now ready to easily consider all possible bipartite graphs with 2^m edges, $m \leq 4$.

Theorem 3.2 *Let $1 \leq m \leq 4$. Let B be a bipartite graph with 2^m edges. Then B contains an induced subgraph with exactly 2^{m-1} edges.*

Proof: Let $\{X, Y\}$ be the natural bipartition of the vertex set of B. The result is trivial if $m \leq 2$: when $m = 1$ or 2 all vertices in one part must have degree at most 1 or 2 respectively. If $m = 3$ then the result follows from Theorem 2.1 unless each of X and Y contains a vertex of degree at least 5, in which case the result follows from Lemma 3.1.

So suppose B has 16 edges. By Theorem 2.1 we can assume that X (and Y) contains a vertex of degree at least 5. The result follows from Lemma 3.1 if X contains: a vertex of degree at least 7; a vertex of degree 6 and a vertex of degree at least 2; or, a vertex of degree 5 and a vertex of degree at least 3. In each remaining case, X must contain vertices, each having degree 1 or 2, whose degrees sum to at least 8. So a subset S of these vertices can easily be found with degrees adding to exactly 8. Therefore $S \cup N(S)$ induces the desired subgraph in each of the remaining cases. □

References

[1] G. J. Chang and F. W. Hwang, A group testing problem on two disjoint sets, SIAM J. Algebr. Discret. Methods, 2 (1981), 35-38.

[2] G. J. Chang, F. W. Hwang and S. Lin, Group testing with two defectives. Discrete Appl. Math., 4 (1982), 97-102.

[3] J. S. Juan and G. J. Chang, Group testing in bipartite graphs, Taiwan J. Math., 6 (2002), 67-73.

[4] P. Damaschke, A tight upper bound for group testing in graphs, Discrete Appl. Math., 48 (1994), 101109

[5] D. Du and F. W. Hwang, Combinatorial group testing and applications, 2^{nd} edition, World Scientific Publishing Co., River Edge, NJ, 2000.

[6] J. S. Juan and G. J. Chang, Group Testing in Graphs, J. Comb. Optim., 14 (2007), 113-119.