TAIWANESE JOURNAL OF MATHEMATICS
Vol. 6, No. 1, pp. 67-73, March 2002
This paper is availiable online at http://www.math.nthu.edu.tw/tjm/

# GROUP TESTING IN BIPARTITE GRAPHS*

Su-Tzu Juan and Gerard J. Chang†

**Abstract.** This paper investigates the group testing problem in graphs as follows. Given a graph $G = (V, E)$, determine the minimum number $t(G)$ such that $t(G)$ tests are sufficient to identify an unknown edge $e$ with each test specifies a subset $X \subseteq V$ and answers whether the unknown edge $e$ is in $G[X]$ or not. Damaschke proved that $\lceil \log_2 e(G) \rceil \leq t(G) \leq \lceil \log_2 e(G) \rceil + 1$ for any graph $G$, where $e(G)$ is the number of edges of $G$. While there are infinitely many complete graphs that attain the upper bound, it was conjectured by Chang and Hwang that the lower bound is attained by all bipartite graphs. This paper verifies the conjecture for bipartite graphs $G$ with $e(G) \leq 2^4$ or $2^{k-1} < e(G) \leq 2^{k-1} + 2^{k-3} + 2^{k-6} + 19 \cdot 2^{\frac{k-7}{2}} - 1$ for $k \geq 5$.

## 1. Introduction

The idea of group testing originated from the blood testing in 1942 by Dorfman, who published the first paper [8] on this topic. While traditional group testing literature employs probabilistic models, Li [12] was the first to study combinatorial group testing as follows. Consider a population $V$ of $n$ items consisting of an unknown subset $D \subseteq V$ of $d$ defectives. The problem is to identify the set $D$ by a sequence of group tests. Each test is on a subset $X$ of $V$ with two possible outcomes: a *pure* outcome indicates that $X \cap D = \emptyset$, and a *contaminated* outcome indicates that $X \cap D \neq \emptyset$. The goal is to minimize the number $M(d, n)$ of tests under the worst scenario. A best algorithm under this goal is called a *minmax algorithm*. For a good reference, see the book by Du and Hwang [9].

---

Received January 15, 2000, revised March 15, 2000.
Communicated by F. K. Hwang.
2001 *Mathematics Subject Classification*: 05C99.
*Key words and phrases*: Group testing, algorithm, complete graph, bipartite graph, induced subgraph.
*Supported in part by the National Science Council under grant NSC89-2115-M009-007.
†E-mail: gjchang@math.nctu.edu.tw.

As the sample space of the problem consists of $\binom{n}{d}$ samples, we have the following information-theoretic lower bound

$$M(d,n) \geq \lceil \log_2 \binom{n}{d} \rceil,$$

where $\lceil x \rceil$ ($\lfloor x \rfloor$) denotes the smallest (largest) integer not less (greater) than $x$. Using a bisection method, it is easy to get

$$M(1,n) = \lceil \log_2 n \rceil.$$

On the other hand, it is hard to determine $M(d,n)$ for $d \geq 2$. Even for the case of $d = 2$, we only know that

$$\lceil \log_2 \binom{n}{2} \rceil \leq M(2,n) \leq \lceil \log_2 \binom{n}{2} \rceil + 1.$$

Toward the study of $M(2,n)$, Chang and Hwang [4, 5] considered the problem of identifying two defectives in two disjoint sets $A$ and $B$, each containing exactly one defective. At first, it seems that one cannot do better than working on the two disjoint sets separately. Surprisingly, a small example with $|A| = 3$ and $|B| = 5$ shows that $4 = \lceil \log_2(3 \cdot 5) \rceil$ tests is enough rather than identifying the defectives in $A$ and $B$ separately, which takes $\lceil \log_2 3 \rceil + \lceil \log_2 5 \rceil = 5$ tests. In general, they [5] proved that the minmax number to identify the only defective in $A$ and the only defective in $B$ is

$$\lceil \log_2(mn) \rceil,$$

where $m = |A|$ and $n = |B|$. By associating each item to a vertex, Spencer [4] observed that the sample space of this problem can be represented by a bipartite graph where each edge represents a sample in $A \times B$. (Throughout this paper we presume that the reader is familiar with the basic-theoretic notations. See [3, 13] if necessary.) Chang and Hwang [4] conjectured that a bipartite graph with $2^k$ ($k \geq 1$) edges always has an induced subgraph with $2^{k-1}$ edges, or equivalently, $t(G) = \lceil \log_2 e(G) \rceil$ for any bipartite graph $G$. While the conjecture remains open, it has stimulated forthcoming research casting group testing on graphs.

Aigner [1] proposed the following problem: Given a graph $G = (V, E)$, determine the minimum number $t(G)$ such that $t(G)$ tests are sufficient in the worst case to identify an unknown edge $e$ when each test specify a subset $X \subseteq V$ and answers whether the unknown edge $e$ is in $G[X]$ or not, where $G[X]$ is the subgraph of $G$ induced by the vertex set $X$. It is then clear that $t(G) = 0$ if $G$ has exactly one edge, and otherwise

$$t(G) = 1 + \min_{X \subseteq V} \max\{t(G[X]), t(G - E(X))\}.$$

The information-theoretic lower bound for this parameter is

$$\lceil \log_2 e(G) \rceil \le t(G),$$

where $e(G)$ denotes the number of edges in $G$. Chang and Hwang's result [5] becomes that

$$t(K_{m,n}) = \lceil \log_2 e(K_{m,n}) \rceil = \lceil \log_2 (mn) \rceil$$

for complete bipartite graphs $K_{m,n}$, and their conjecture is

**Conjecture 1 [4].** *For any bipartite graph* $G$, *we have* $t(G) = \lceil \log_2 e(G) \rceil$.

From the result in [6], it follows that $t(K_n) \le \lceil \log_2 e(K_n) \rceil + 1$, and there are infinitely many complete graphs attaining the upper bound. Althöfer and Triesch [2] showed that $t(G) \le \lceil \log_2 e(G) \rceil + 1$ for bipartite graphs, and $t(G) \le \lceil \log_2 e(G) \rceil + 3$ for arbitrary graphs. Damaschke [7] proved that $t(G) \le \lceil \log_2 e(G) \rceil + 1$ for arbitrary graphs. In fact, he proved a more general result that $t(G) = \lceil \log_2 e(G) \rceil$ for a graph $G$ with $2^{k-1} < e(G) \le 2^{k-1} + \frac{171}{64} \cdot 2^{\frac{k-1}{2}}$ when $k \ge 13$ and $e(G) \in [1, 14] \cup [17, 25] \cup [33, 45] \cup [65, 83] \cup [129, 155] \cup [257, 295] \cup [513, 568] \cup [1025, 1105] \cup [2049, 2165]$.

The attempt of this paper is to determine the largest number $f(k)$ such that $t(G) = \lceil \log_2 e(G) \rceil$ for any bipartite graph $G$ with $2^{k-1} < e(G) \le f(k)$. Note that Conjecture 1 says $f(k) = 2^k$ for $k \ge 0$. In this paper, we verify the conjecture for $k \le 4$, and show that $f(k) \ge 2^{k-1} + 2^{k-3} + 2^{k-6} + 19 \cdot 2^{\frac{k-7}{2}} - 1$ for $k \ge 5$.

## 2. Graphs $G$ with $t(G) = \lceil \log_2 e(G) \rceil$

It is of our interest to study which graphs $G$ satisfy $t(G) = \lceil \log_2 e(G) \rceil$. The first well-known result of this kind is

**Theorem 2 [5].** *For any complete bipartite graph* $K_{m,n}$, *we have*

$$t(K_{m,n}) = \lceil \log_2 e(K_{m,n}) \rceil = \lceil \log_2 (mn) \rceil.$$

It is not hard to see that acyclic graphs also have this property.

**Theorem 3.** *For any acyclic graph* $G$, *we have* $t(G) = \lceil \log_2 e(G) \rceil$.

*Proof.* Removing successively vertices of degree one, we can get induced subgraphs of $G$ whose numbers of edges range from 1 to $e(G)$. This together with the information-theoretic lower bound gives the theorem. ■

Damaschke's result [7] is for general graphs.

**Theorem 4 [7].** *For any graph $G$ with $2^{k-1} < e(G) \leq 2^{k-1} + \frac{171}{64} \cdot 2^{\frac{k-1}{2}}$ and $k \geq 13$, and $e(G) \in [1, 14] \cup [17, 25] \cup [33, 45] \cup [65, 83] \cup [129, 155] \cup [257, 295] \cup [513, 568] \cup [1025, 1105] \cup [2049, 2165]$, we have $t(G) = \lceil \log_2 e(G) \rceil$.*

In the remaining part of this paper, we employ Damaschke's techniques towards Conjecture 1. For a graph $G$, denote by $\delta(G)$ the minimum degree of a vertex in $G$.

**Lemma 5.** *If $G$ is a bipartite graph with $\delta(G) \geq n$, then $e(G) \geq n^2$.*

*Proof.* The lemma follows from the fact that any part of the vertex set of $G$ has at least $n$ vertices and any vertex is of degree at least $n$. ∎

**Lemma 6.** *If $n^2 - 1 \leq b < (n+1)^2 - 1$ and $a = b - n + 1$, then any bipartite graph $G$ with $e(G) \geq a$ has an induced subgraph $H$ with $a \leq e(H) \leq b$.*

*Proof.* Choose an induced subgraph $H$ of $G$ with as few vertices as possible such that $a \leq e(H)$. Assume $e(H) \geq b + 1$. By the choice of $H$, for any vertex $x$ of degree $\delta(H)$ in $V(H)$, we have $e(H - x) \leq a - 1 < b + 1 \leq e(H)$, which implies that

$$\delta(H) = \deg_H(x) = e(H) - e(H - x) \geq b - a + 2 = n + 1.$$

Assume that $\delta(H) = n + i$, where $i \geq 1$. Then, according to Lemma 5, $e(H) \geq (n + i)^2$. Therefore,

$$\begin{aligned} e(H - x) &= e(H) - \delta(x) \geq (n+i)^2 - (n+i) \geq n^2 + n \\ &= (n+1)^2 - 1 - n > b - n = a - 1, \end{aligned}$$

a contradiction. Hence $a \leq e(H) \leq b$ as desired. ∎

**Lemma 7.** *Suppose vertices $x$ and $y$ are in the same part of a bipartite graph $G$. If $\deg_G(x) + \deg_G(y) \geq 2m$, then $G$ has an induced subgraph with exactly $2m$ edges.*

*Proof.* Suppose $H$ is the subgraph of $G$ induced by $C \cup \{x, y\}$, where $C$ is the set of all neighbors of $x$ and $y$. As $\sum_{v \in C} \deg_H(v) = \deg_G(x) + \deg_G(y) \geq 2m$ and $\deg_H(v)$ is 1 or 2 for any vertex $v$ in $C$, there is a subset $D \subseteq C$ such that $\sum_{v \in D} \deg_H(v) = 2m$. Hence $D \cup \{x, y\}$ induces a subgraph with exactly $2m$ edges. ∎

**Theorem 8.** *Conjecture 1 is valid for $k \leq 4$.*

*Proof.* The theorem is clearly true for $k \le 1$. Now consider a bipartite graph $G$ of $2^k$ vertices for $2 \le k \le 4$. It is sufficient to prove that $G$ has an induced subgraph with $2^{k-1}$ edges. By Lemma 7, we may assume

$(*)$ $\deg_G(x) + \deg_G(y) < 2^{k-1}$
for any two vertices $x$ and $y$ in the same part of $G$.

This in turn implies that for $2 \le k \le 3$ every vertex of $G$ has degree at most two, which allows the existence of an induced subgraph of $2^{k-1}$ edges. So now consider the case of $k = 4$.

According to $(*)$, any part of $G$ has at most one vertex of degree at least $4$. Furthermore, either there is some part in which there are some vertices whose degree sum is $8$, or else the degree sequence of each part is $(4, 3, 3, 3, 3)$ or $(3, 3, 3, 3, 3, 1)$. For the former case, those vertices of degree sum $8$ together with their neighbors induce a subgraph of $8$ edges. For the later case, choose a vertex $x$ in part $A$ with exactly $3$ neighbors $y_1, y_2, y_3$ in $B$. Then, choose a vertex $z$ in $B - \{y_1, y_2, y_3\}$ with exactly $3$ neighbors $w_1, w_2, w_3$ in $A - \{x\}$. At least one of $w_1, w_2, w_3$, say $w_1$, is of degree $3$. Then $G - \{x, z, w_1\}$ is an induced subgraph of $G$ with exactly $8$ edges. ∎

**Theorem 9.** $f(k+1) \ge 2f(k) + 1 - \lfloor\sqrt{f(k)+1}\rfloor$.

*Proof.* Suppose $n^2 - 1 \le f(k) < (n+1)^2 - 1$, i.e., $n = \lfloor\sqrt{f(k)+1}\rfloor$. We only need to show that for any bipartite graph $G$ with $2f(k)+1-n$ edges, $t(G) \le k+1$. Choosing $b = f(k)$ and applying Lemma 6, we infer that $G$ has an induced subgraph $H$ with

$$f(k) + 1 - n \le e(H) \le f(k).$$

And hence

$$f(k) + 1 - n \le e(G - E(H)) \le f(k).$$

Therefore, $t(H) \le k$ and $t(G - E(H)) \le k$, which imply

$$t(G) \le 1 + \max\{t(H), t(G - E(H))\} \le k + 1.$$

This completes the proof of the theorem. ∎

To estimate a good lower bound for $f(k)$ by using the above theorem, we consider the sequence $\{b_k : k \ge 4\}$ defined by $b_4 = 16$ and

$$b_k = 2b_{k-1} + 1 - \lfloor\sqrt{b_{k-1}+1}\rfloor$$

for $k \ge 5$. It is clear that $f(k) \ge b_k$ for $k \ge 4$. Note that

$$b_5 = 2 \cdot 16 + 1 - \lfloor\sqrt{16+1}\rfloor = 29.$$

**Lemma 10.** *For any* $k \geq 5$, *we have* $b_k + 1 \leq 15 \cdot 2^{k-4}$.

*Proof.* First, $b_5 + 1 = 29 + 1 = 15 \cdot 2^{5-4}$. Suppose $k \geq 5$ and $b_k + 1 \leq 15 \cdot 2^{k-4}$ holds. Then

$$b_{k+1} + 1 = 2b_k + 1 - \lfloor \sqrt{b_k + 1} \rfloor + 1 \leq 2b_k + 2 \leq 2(15 \cdot 2^{k-4}) = 15 \cdot 2^{(k+1)-4}$$

and so the lemma follows from induction. ■

**Theorem 11.** *For* $k \geq 5$, *we have* $b_k \geq 2^{k-1} + 2^{k-3} + 2^{k-6} + 19 \cdot 2^{\frac{k-7}{2}} - 1$.

*Proof.* The theorem is true for $k = 5$ as $b_5 = 29 = 2^{5-1} + 2^{5-3} + 2^{5-6} + 19 \cdot 2^{\frac{5-7}{2}} - 1$. Suppose $k \geq 6$ and the theorem is true for $k - 1$. Then

$$\begin{aligned}
b_k &= 2b_{k-1} + 1 - \lfloor \sqrt{b_{k-1} + 1} \rfloor \quad \text{(by the definition of } b_k\text{)} \\
&\geq 2b_{k-1} + 1 - \sqrt{15} \cdot 2^{\frac{k-5}{2}} \quad \text{(by Lemma 10)} \\
&\geq 2(2^{k-2} + 2^{k-4} + 2^{k-7} + 19 \cdot 2^{\frac{k-8}{2}} - 1) + 1 - \sqrt{15} \cdot 2^{\frac{k-5}{2}} \\
&\quad \text{(by the induction hypothesis)} \\
&= 2^{k-1} + 2^{k-3} + 2^{k-6} + (19\sqrt{2} - 2\sqrt{15}) 2^{\frac{k-7}{2}} - 1 \\
&\geq 2^{k-1} + 2^{k-3} + 2^{k-6} + 19 \cdot 2^{\frac{k-7}{2}} - 1 \quad \text{(since } 19\sqrt{2} - 2\sqrt{15} > 19\text{)}.
\end{aligned}$$

The theorem then follows ■

**Corollary 12.** *If* $G$ *is a bipartite graph with* $2^{k-1} < e(G) \leq 2^{k-1} + 2^{k-3} + 2^{k-6} + 19 \cdot 2^{\frac{k-7}{2}} - 1$ *and* $k \geq 5$, *then* $t(G) = \lceil \log_2 e(G) \rceil$.

## References


1. M. Aigner, Search problem on graphs, *Discrete Appl. Math.* **14** (1986), 215-230.
2. I. Althöfer and E. Triesch, Edge search in graphs and hypergraphs of bounded rank, *Discrete Math.* **115** (1993), 1-9.
3. J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*, Macmillan, New York, 1976.
4. G. J. Chang and F. K. Hwang, A group testing problem, *SIAM J. Algebraic Discrete Methods* **1** (1980), 21-24.
5. G. J. Chang and F. K. Hwang, A group testing problem on two disjoint sets, *SIAM J. Algebraic Discrete Methods* **2** (1981), 35-38.
6. G. J. Chang, F. K. Hwang and S. Lin, Group testing with two defectives, *Discrete Appl. Math.* **4** (1982), 97-102.

7. P. Damaschke, A tight upper bound for group testing in graphs, *Discrete Appl. Math.* **48** (1994), 101-109.

8. R. Dorfman, The detection of defective members of large populations, *Ann. Math. Statist.* **14** (1943), 436-440.

9. D. Z. Du and F. K. Hwang, *Combinatorial Group Testing and Its Applications*, World Scientific, Singapore, 1993.

10. F. K. Hwang, Hypergeometric group testing procedures and merging procedures, *Bull. Inst. Math. Acad. Sinica* **5** (1977), 335-343.

11. G. O. H. Katona, Combinatorial search problem, in: *A Survey of Combinatorial Theory*, J. N. Srivastava et al., ed., North-Holland, Amsterdam, 1973.

12. C. H. Li, A sequential method for screening experimental variables, *J. Amer. Statist. Assoc.* **57** (1962), 455-477.

13. D. B. West, *Introduction to Graph Theory*, Prentice Hall, Inc., Upper Saddle River, NJ, 1996.

Su-Tzu Juan and Gerard J. Chang
Department of Applied Mathematics
National Chiao Tung University
Hsinchu, Taiwan 300, R.O.C.
E-mail: gjchang@math.ntu.edu.tw