Optimal quantitative group testing on cycles and paths

F.K. Hwang\textsuperscript{a} J.S. Lee\textsuperscript{b, *}

\textsuperscript{a}Department of Applied Mathematics, National Chiao-tung University, Hsin-Chu 300, Taiwan, ROC
\textsuperscript{b}Department of Mathematics, National Kaohsiung Normal University, Kaohsiung 802, Taiwan, ROC

Abstract

We determine the minimum number of group tests required to search for a special edge when the graph consists of cycles and paths, generalizing previous results of Aigner on paths and on a simple cycle. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Group testing; Search; Graph

1. Introduction

Suppose that we have a set of items containing exactly two defective ones. The problem is to identify them through quantitative group testing [2]. Any subset $S$ of items can be tested, and the feedback $f(s)$ reveals the number of defectives in $S$, i.e. $f(S) = 0, 1$ or 2. There are constraints on which pairs of items can be the defective pair, and the constraints can be represented by a graph where the vertex-set is the set of items, and the edge-set is the set of allowed pairs. Thus, the problem can also be viewed as searching for a special edge on a graph $G(V,E)$. Suppose $|E| = n$. Since each test has three possible feedbacks, $\lceil \log_3 n \rceil$ is the information lower bound on the number of tests required. Aigner [1] proved

\textit{Theorem 1.} If $G$ consists of paths, then $\lceil \log_3 n \rceil$ tests suffice.

\textit{Theorem 2.} If $G$ is a cycle and $n < 3'$, then $t$ tests suffice. If $n = 3'$, then $t + 1$ tests suffice.

\textsuperscript{*} Corresponding author.

\textsuperscript{☆} This research is supported by NSC grant 88-2115-M-009-016.
In this paper we consider the case that \( G \) consists of any number of cycles and paths. We give the minimum number of tests required for such \( G \).

2. Optimal testing

We first prove an upper bound.

**Theorem 3.** Suppose \( G \) consists of cycles and paths. Then \( 1 + \lceil \log_3 n \rceil \) tests suffice.

**Proof.** If \( G \) contain no cycles, then Theorem 3 follows from Theorem 1. If \( G \) has \( m \) cycles \( C_1, C_2, \ldots, C_m \), test \( S_i = \{ v_1, v_2, \ldots, v_m \} \), where \( v_i \) is an arbitrary vertex on \( C_i \). Suppose \( f(S) = 0 \). Then the two edges incident to \( v_i \) on \( C_i \) cannot be special for each \( i = 1, 2, \ldots, m \). Therefore \( C_i \) is reduced to a path. By Theorem 1, \( \lceil \log_3 n \rceil \) more tests suffice. Suppose \( f(S) = 1 \), then the special edge must be an edge incident to one of the \( v_i \). Again, each \( C_i \) is reduced to a path of two edges and Theorem 1 applies. The proof is completed by noting that \( f(S) \) cannot be 2 since no edge can be incident to two vertices in \( S \).

Consider a test \( S \) on a graph \( G \). An edge \((u, v)\) will be called an \( S_i \)-edge, \( i = 0, 1, 2 \) if \( |\{u, v\} \cap S| = 0, 1, 2 \), respectively. Let \( G_0, G_1, G_2 \) be the partition of \( G \) according to the three feedbacks of \( S \). Then \( G_i = \{ S_i \text{-edge} \} \) for \( i = 0, 1, 2 \). A cycle (path) will be called a mixed cycle (path) if it contain an \( S_1 \)-edge. Otherwise it is called a pure cycle (path), or an \( S_0 \) (\( S_2 \))-cycle if we want to be more specific. We also refer to an edge as pure if it is either \( S_0 \) or \( S_2 \).

**Lemma 4.** Let \( i \) and \( j \) satisfy the conditions \( i \geq 0, j \geq 0 \) and \( i + 2j \leq k \), except when \( j = 0 \), then \( i \) is 0 or \( k \). Then there exists a test \( S \) on a \( k \)-cycle \( C \) such that \( |S_0| = i, |S_1| = 2j \) and \( |S_2| = k - i - 2j \).

**Proof.** If \( j = 0 \), then either \( S \cap C = \emptyset \) or \( S \cap C = \emptyset \). Otherwise, assign arbitrary \( k - i - 2j + 1 \) consecutive vertices to \( S \), and assign the next \( i + 1 \) consecutive vertices to \( \tilde{S} \) (not in \( S \)). The remaining vertices are assigned \( S \) or \( \tilde{S} \) such that \( S \) and \( \tilde{S} \) alternate.

**Lemma 5.** Consider a set \( P \) of paths with \( k \) total edges. Let \( i \) and \( j \) satisfy the conditions \( i \geq 0, j \geq 1 \) and \( i + j \leq k \). Then there exists a test \( S \) on \( P \) such that \( |S_0| = i, |S_1| = j \) and \( |S_2| = k - i - j \).

**Proof.** We order the paths such that the \( k \) edges (hence all vertices) are linearly ordered. Assign the first \( k - i - j \) edges to \( S_2 \), meaning their vertices are all in \( S \). Assign the next \( j \) edges to \( S_1 \), if \( j \) is odd or \( i = 0 \). If \( j \) is even and \( i > 0 \), assign the next \( j - 1 \) edges to \( S_1 \). Furthermore, if there is a change of path during this process, then the vertex starting the new path is in the same set, \( S \) or \( \tilde{S} \), as its preceding vertex. These rules assure that this process ends in an \( \tilde{S} \)-vertex which will start the
final assignment of \( i \) edges in \( S_0 \), meaning all their vertices are in \( S \). For \( j \) even and \( i > 0 \), there is one edge left which will be assigned to \( S_1 \), meaning the last vertex is in \( S \). \( \square \)

**Corollary 6.** A partition \((i, 0, k - i)\) is possible if and only if there exists a subset of paths with a total of \( i \) edges.

Let \( M(G) \) denote the minimum number of tests required for \( G \).

**Theorem 7.** Let \( G \) consist only of cycles and paths with \( n \) edges in total, where \( 3^{t-1} < n \leq 3^t \). Then \( M(G) = t \) except

(i) \( G \) consists of cycles only and \( n = 3^t \),

(ii) \( t = 2 \) and \( G \) contains two cycles,

(iii) \( t = 3 \) and \( G \) contains seven cycles,

(iv) \( t = 4 \) and \( G \) contains 26 cycles,

and \( M(G) = t + 1 \) in the four exception cases.

**Proof.** *Sufficiency:* The \( t \leq 2 \) case is easily verified. We prove the general \( t \geq 3 \) by induction. It suffices to prove that if \( G \) is not one of the exception cases, then there exists a test \( S \) where the three feedbacks partition \( G \) into \( G_0, G_1, G_2 \) with \( n_0, n_1, n_2 \) edges, where \( n \leq 3^{t-1} \) and \( G_i \) is not an exception case for \( i = 0, 1, 2 \).

Suppose \( G \) contains \( c \) cycles where \( c \leq 3^{t-1} - 1 \). We consider two cases:

1. \( c < (3^{t-1} - 1)/2 \). Assign \( S_1 \)-edges such that the \( c \) cycles are all mixed. Suppose the \( c \) cycles contain \( n' \) edges. By Lemma 4 we can obtain at least \( 2\lceil (n' - c)/2 \rceil \) \( S_1 \)-edges. Assign \( \min\{2\lceil (n' - c)/2 \rceil, 3^{t-1} - 1\} = 3^{t-1} - j \) edges to \( S_1 \), where \( j \geq 1 \) is odd . Again by Lemma 4, the pure edges in the \( c \) cycles can be divided evenly into \( S_0 \) and \( S_2 \). Since \( 3^{t-1} - j \geq \lceil n'/3 \rceil \), so \( 3^{t-1} - j < \lfloor n/3 \rfloor \) implies the existence of paths with a total of more than \( j \) edges. By Lemma 5, we can obtain \( j \) \( S_1 \)-edges and divide the other edges evenly into \( S_0 \) and \( S_2 \). Note that in the case \( 3^{t-1} - j \geq \lceil n/3 \rceil \), even though no \( S_1 \)-edge is needed on the paths, some \( S_1 \)-edges may be forced in the process of dividing the path edges evenly into \( S_0 \) and \( S_2 \). By Lemma 5, at most one \( S_1 \)-edge needs to be forced. This is alright since \( 3^{t-1} - j + 1 \leq 3^{t-1} \).

2. \( c \geq (3^{t-1} - 1)/2 \). We will assign the \((3^{t-1} - 1)/2\) largest cycles to be mixed each with two \( S_1 \)-edges. Let \( p \) denote the largest size of the pure cycles. Then \( p \leq 5 \) for otherwise the mixed cycles would have consumed \( 3(3^{t-1} - 1) = 3^t - 3 \) edges and there are not enough edges left for a pure \( p \)-cycle with \( p \geq 6 \). Let \((e_0, e_2)\) be a division of edges into the \( S_0 \) and \( S_2 \) type through assigning the pure cycles into \( G_0 \) or \( G_2 \). Then there is a division with \( |e_0 - e_2| \leq 5 \). For \( t \geq 3 \), there are at least four mixed cycles with 12 pure edges on them. By Lemmas 4 and 5, we can divide these pure edges as well as the pure edges on paths (if any) arbitrarily, i.e. the \( n - 3^{t-1} (n - (3^{t-1} - 1) \) if no paths exist) pure edges can be divided evenly into \( G_0 \) and \( G_2 \). Therefore \( n_i \leq 3^{t-1} \) for \( i = 0, 1, 2 \). Furthermore, the number of cycles in \( G_0 \) or \( G_2 \) is at most


\[ \left\lceil \frac{3^{t-1} - 1 - (3^{t-1} - 1)/2}{2} \right\rceil < 3^{t-2} - 1 \quad \text{for} \ t \geq 5, \]
\[ \frac{25 - (3^3 - 1)/2}{2} = 6 \quad \text{for} \ t = 4, \]
\[ \frac{6 - (3^2 - 1)/2}{2} = 1 \quad \text{for} \ t = 3. \]

Hence they are not exception cases.

That \( t + 1 \) tests suffice for the exception cases follow from Theorem 3.

**Necessity:** That \( t \) tests are necessary for the nonexception case follows from the information lower bound. We now prove that the exception cases cannot be done in \( t \) tests.

(i) Since the number of \( S_1 \)-edges on a cycle must be even, there is no way to partition \( 3^t \) edges on cycles into \( 3^{t-1}, 3^{t-1} \) and \( 3^{t-1} \).
(ii) Suppose \( G \) contains two cycles. Then the number of \( S_1 \)-edges on these two cycles must be 2 (it must be even). That means one of the two cycles, of size \( k \), is pure. If \( k > 3 \), then one more test cannot do it by information lower bound. If \( k = 3 \), then again one more test cannot do it since it is the exception case (i).
(iii) Suppose \( G \) contains seven cycles. Since at most \( (3^3 - 1)/2 = 4 \) cycles can be mixed, there are at least three pure cycles. Without loss of generality, assume there are two \( S_0 \)-cycles. Then \( G_0 \) contains two cycles and is the exception case (ii), hence it cannot be done in two more tests.
(iv) Suppose \( G \) contains 26 cycles. Since at most \( (3^4 - 1)/2 = 13 \) cycles can be mixed, there are at least thirteen pure cycles. Without loss of generality, assume there are seven \( S_0 \)-cycles. Then \( G_0 \) contains seven cycles and is the exception case (iii), hence it cannot be done in three more tests. \( \square \)

**References**