On the Performance of Protocols for Collecting Responses over a Multiple-Access Channel *

Mostafa H. Ammar and George N. Rouskas
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332

Abstract

We consider a generalization of the multiple access problem where it is necessary to identify a subset of the ready users, not all. The problem is motivated by several "response collection" applications that arise in distributed computing and database systems. In these applications, a collector is interested in gathering a set of responses from a number of potential respondents. The collector and respondents communicate over a shared channel. We define three collection objectives and investigate a suite of protocols that can be used to achieve these objectives. The protocols are based on the use of polling, TDMA, and group testing. Using a binomial respondent model we analyze and, where applicable, optimize the performance of the protocols. Our concern is with cost measures that reflect the computational load placed on the system, as well as the delay incurred for achieving a particular objective.

1 Introduction

We investigate the problem of how to best collect a specified number of responses from a set of nodes over a multiple access channel. Several situations in distributed systems where such a problem arises are described later. We consider a system where nodes share a common communication channel. One node in the system is interested in collecting responses from the other nodes. Not all nodes can or will respond when requested and the node soliciting responses is interested in achieving a collection objective.

The problem we consider is actually a generalization of the multiple access communication problem where we are concerned with identifying a *subset* of ready users, not all. A response collection process will be aimed at achieving one of a set of *collection objectives* to be described later. We describe and analyze a suite of protocols that can be used for response collection. Our concern is with the cost of the collection process in terms of the amount of computation resources it consumes, as well as the amount of time expended to achieve a certain collection objective. The protocols we use are based on the use of polling, time division multiple access (TDMA) and group testing.

Whereas polling and TDMA are well known multiple access techniques, group testing warrants a short introduction. It is a technique that can be used to efficiently identify "defective" items in a set. It has been studied extensively in different contexts (see for example [1, 2, 3, 4]). The basic idea of the technique is the testing of items being inspected in groups. The composition of the group to be tested at any one point in time being dictated by the history of previous test outcomes. Each test is counted as a single step and the objective is to determine group composition rules to minimize the number of steps. In its original form, the problem assumes the outcome of each test would indicate one of two situations: "all items are not defective" or "there is at least one defective item." We are concerned here with the potential use of group testing as a technique for collision resolution over a multiple access channel. Such use has been described in [5, 6, 7, 8]. The additional feature when using group testing over a multiple-access channel is the ability to differentiate among three possible outcomes when a group is enabled: no transmission, one transmission, and more than one transmission (a collision).

This paper is organized as follows. In section 2 we discuss some applications that motivate our work. Section 3 contains a model of our system. Section 4 presents a description and analysis of the Polling and TDMA protocols. In section 5 we describe and analyze an approach based on the staging of the response collection process where in each stage a TDMA protocol is employed. Sections 6 and 7 investigate the group testing and staged group testing protocols. Some numerical examples are presented in section 8 and section 9 contains some concluding remarks.

2 Some Applications

The following are some applications that make use of response collection.

A database system with multiple query optimization: Here we have a shared channel LAN with the primary purpose of giving a set of attached users access to a database (also connected to the network). The users are moderately active and the database employs sophisticated query processing techniques. These include schemes to speedup query processing through the use of multiple query optimization (see e.g., [9]). Rather than processing each query individually, the database tries to process a number of queries (up to a maximum) at a time. In order to manage memory and processing resources efficiently, the database processor prefers to actively collect responses, rather than receiving responses asynchronously.

*This work is supported in part by NSF grant NCR-8604850.

CH2979-3/91/0000-1490 $1.00 © 1991 IEEE

A broadcast delivery information system: An information system using broadcast as a delivery mechanism has the potential for *shared response*, i.e., responding to several users with one transmission (see e.g., [10, 11, 12, 13]). In order to maximize the benefit of the response sharing feature, the information source needs to be aware of the information needs of a representative set of users at any one time. Thus, before transmitting responses, the information source needs to collect a set of requests from the users.

Quorum collection for synchronization in a distributed system: Quorum consensus is a general class of synchronization protocols for distributed systems [14, 15, 16, 17]. An operation may proceed to completion only if it is granted permission from a number of nodes. If mutually exclusive execution of operations is desired, e.g., as would be required when writing replicated data, then the node executing the operation needs to collect permission from a majority of nodes. Other applications, such as the reading of replicated data, may require permission from a certain number of nodes, not necessarily a majority [18]. When quorum consensus protocols are used, if a quorum cannot be collected, the operation requesting the quorum aborts. Nodes, in some situations, may not be able to grant permission when requested if they have already granted permission to another node. They may also not be able to grant permission because they have failed or are too busy.

Finding multiple instances of a named resource: An application running in a distributed system often requires access to multiple instances of a resource. The application typically knows the name or property of such a resource, and may not be aware of where the resource is physically located in the network. The searching application needs to determine a set of addresses where the resource resides [19]. Some examples of this are: 1) a node is searching for four or more processors that are lightly loaded in order to run a parallel program, 2) a node is searching for three copies of a replicated data item in order to update it, and 3) a node is searching for up to four disks with a given amount of available space to store a certain file.

3 System Model

This section describes a model of a system that captures the salient features of the request collection applications discussed in section 2.

3.1 The Collector and Respondents

The system under consideration has a node (connected to a shared channel) which is attempting to collect responses. We call this node *the collector*. The collector actively solicits responses by transmitting messages on the channel.

All the nodes that can potentially respond to a collector's request are called *respondents*. We assume there are N such respondents. This collector-respondent classification may be permanent (as in the multiple query optimizer and the shared response system discussed above) or it may be temporary (as in the other two applications.) In the latter case, the collector will abandon its role once its response collection objective has been achieved. At that time another node may assume the collector's role. As several nodes may desire to become collectors at the same time, a fair "election" protocol needs to be available for use by the nodes. In this paper we only concern ourselves with the system behaviour from the time a new collector is identified until the collector's objective is achieved.

The collector is the (perhaps temporary) master in the system and actively solicits responses from the respondents. We distinguish between the *soliciting* and *enabling* of a respondent. A respondent is solicited once it receives a message from the collector making it aware that a collection process is underway and indicating the collection objective. A respondent is enabled if the protocol rules allow it to transmit a response on the channel if indeed it can respond. A respondent can be enabled only after or at the same time as it is solicited.

The collector is assumed to operate with some statistical knowledge of the state of the respondents. In our analysis we will assume the *binomial respondent* model. At the instant the collection process begins, each respondent will transmit a response when enabled with a probability q and with probability $p = 1 - q$ a respondent will not transmit a response when enabled.

3.2 Collection Objectives

With respect to a collector's request, a respondent is classified as *active* if it will respond when given a chance (i.e., solicited and enabled). A respondent is said to be *inactive* otherwise. The goal of the collector is to identify and collect "enough" responses from active respondents to satisfy its application. Note that in some instances, the collector's goal may be achieved if it determines (from the *lack* of responses) that the desired number of responses cannot be collected. We consider three distinct collection objectives.

1. *L or Nothing:* Terminate successfully after collecting exactly L responses or abort when a determination is made that the number of active respondents is less than L.

2. *L or Maximum:* Terminate successfully after collecting L responses, or after all respondents have been given a chance to respond, whichever occurs first.

3. *L or More:* Terminate successfully if L or more responses have been collected *and* all respondents have been given a chance to respond. Abort when a determination is made that the number of active respondents is less than L.

For example, assume the number of respondents N is 20 and $L = 6$. A collector with the "6 or Nothing" objective will terminate if 6 responses have been received or it will abort the search if out of the respondents enabled a total of 15 did not transmit responses. With the "6 or More" objective, the collector will abort in the same situation above, it will, however, continue to gather responses after 6 responses have been received. In the case of a "6 or Maximum" objective the collector will terminate (before all respondents have been enabled) only if 6 responses have been collected.

We note the following equivalencies between the various objectives: (recall that N is the total number of respondents)

$$N \text{ or } Maximum \equiv 0 \text{ or } More \equiv Find\ All\ Active \quad (1)$$

We will use the superscript (y) to denote a collection objective. In this paper we present results that pertain to the L or Maximum collection objective and we use the superscript (ℓ) to indicate the ℓ or maximum objective. Results for the other objectives have been obtained and can be found in [20].

3.3 Network Environment

All communication takes place over an error-free shared channel with capabilities for single destination, multicast and broadcast addressing. Simultaneous transmissions over the shared channel result in a collision. All response packets are assumed to be of the same size and the network can operate in a slotted mode where each slot is long enough for the transmission of a response packet. Respondents are constrained to begin transmission at a slot boundary and thus all collisions are the result of the complete overlap of response packets. The channel is assumed to provide the so-called (0, 1, e) feedback where the nodes on the channel are informed whether the previous slot contained no transmissions (0), one transmission (1), or a collision (e).

3.4 Collection Costs

For a particular protocol z and a given collection objective y, we identify three types of costs incurred in the collection process:

1. Delay Cost $D_z^{(y)}$: The average number of response slots needed until the collection objective is achieved or until a determination is made that the collection objective is not attainable.

2. Respondent Solicitation Cost $S_z^{(y)}$: The average number of respondents that are solicited in the collection process. As each solicitation message received requires interpretation and perhaps the generation of a response this measures the computation cost incurred by the respondents.

3. Collector Solicitation Cost $C_z^{(y)}$: The average number of solicitation messages sent by the collector during the collection process. This is a measure of the computation cost incurred by the collector, as well as the delay incurred each time the collector needs to send a solicitation message.

The *total cost* incurred by collection protocol z with objective y is given by:

$$A_z^{(y)} = \alpha D_z^{(y)} + \beta S_z^{(y)} + \gamma C_z^{(y)} \tag{2}$$

where α, β, and γ are weights assigned to the various costs. We also define the *total solicitation cost*

$$B_z^{(y)} = \beta S_z^{(y)} + \gamma C_z^{(y)} \tag{3}$$

We will drop the subscript describing the protocol when it is clear from the context to which protocol the quantity refers.

4 Polling and TDMA

4.1 Description

The collector may employ several techniques to achieve its objective. One possible approach is to poll all respondents individually. Each polling message sent by the collector solicits and enables one respondent. The major disadvantage of the polling approach is that it may require a significant amount of time to complete as each poll requires two messages to be sent if the outcome is positive (i.e., a response is generated by the respondent) or a message followed by a timeout period if the outcome is negative.

Another approach which would require less time is for the collector to declare its objective to the entire network via a broadcast message and have the active respondents send their responses. If the responses are transmitted using a random access scheme, a considerable amount of time and bandwidth may be wasted until the required number of respondents successfully transmit their responses. Alternatively, the protocol may operate by having the respondents ordered in some (perhaps random) way and allocate a time slot to each respondent. Active respondents transmit their responses in the allocated slot. Slots allocated to inactive respondents remain idle. As all respondents can hear channel activity, they all know when the collection objective (declared by the collector in its broadcast message) has been achieved and this collection phase is terminated. We call this technique, *the TDMA collection protocol.* Note that in this scheme the respondents are all solicited by the initial broadcast message. A respondent is enabled at the beginning of its allocated slot.

The TDMA protocol will achieve the collector's objective in less time than a polling procedure. The TDMA protocol, on the other hand, will involve all the respondents (whether or not they are active) as they will receive the initial broadcast message which will have to be interpreted by all the receiving hosts.

4.2 Analysis

For the polling protocol we let $R^{(y)}$ be the average number of respondents that need to be polled before collection objective (y) is achieved. We have that $D_{polling}^{(y)} = C_{polling}^{(y)} = S_{polling}^{(y)} = R^{(y)}$.

For the *L or Maximum* collection objective we get:

$$\begin{aligned} R^{(L)} &= N\sum_{k=0}^{L-1}\binom{N}{k}p^{N-k}q^k \\ &+ \sum_{k=L}^{N}k\binom{k-1}{L-1}p^{k-L}q^L \end{aligned} \tag{4}$$

The expression in (4) is derived by noting that the procedure will require N polls (i.e., poll all respondents) if $L-1$ or less respondents are active. Otherwise, exactly k respondents are polled if the first $k-1$ polls result in discovering $L-1$ active respondents and the last poll discovers an active respondent.

If the TDMA collection protocol is used, a single solicitation message is sent which reaches all N respondents and thus we have for all collection objectives y: $C^{(y)}_{TDMA} = 1$ and $S^{(y)}_{TDMA} = N$. The average number of slots needed to achieve the collection objective will be the same as the number of polls required to reach the same collection objective when polling is used. We thus get that for all collection objectives y, $D^{(y)}_{TDMA} = R^{(y)}$.

5 Staged TDMA

5.1 Description

Staged TDMA is a generalization of both the polling and TDMA protocols described above. The set of potential respondents is subdivided into disjoint groups, say g_i for $i = 1, \ldots, M$. Where M is less than or equal to the total number of respondents. Responses are gathered by having the collector send a multicast message to each group one at a time. All the respondents in a group are ordered and they are allocated slots in which to respond if they are active. If the collection objective is achieved after or during the exploration of the ith group, the procedure is terminated. Otherwise, the collector goes on to explore the $(i+1)$th group and so on. The multicast solicitation messages sent by the collector contain the collection objective. Note that, in general, the objective declared in the $(i+1)$th solicitation message is a "reduced" version of the one declared in the ith message. The amount of reduction is determined by the number of responses collected in the ith stage. All respondents solicited in stage i operate with the knowledge of the collection objective and the number of the not-yet-solicited respondents. Thus during a stage a determination can be made when the collection objective has been achieved or cannot be achieved because the number of unexplored respondents is not sufficient.

Here all respondents in group g_i are solicited once they receive the collector's multicast message. Each respondent is subsequently enabled during its allocated slot. (Similar ideas for the staging of a search can be found in [21].)

We distinguish between *fixed-group* and *adaptive-group* staged TDMA. When fixed groups are used, a set of mutually exclusive and collectively exhaustive groups are determined *a priori*. In an adaptive-group staged procedure, on the other hand, we decide on the constitution of a group after the result of exploring the previous groups is known. The use of optimal adaptive groups will intuitively incur less or equal cost than the use of optimal fixed groups. Adaptive groups, however, may require the use of multiple destination addresses in multicast messages, as single multicast addresses cannot be set up ahead of time.

The staged TDMA protocol has the advantage that it may achieve its objective without involving (i.e., soliciting) all the respondents. It may, however, require somewhat more time to complete when compared to the single-stage TDMA protocol described above because of the delay involved in sending solicitation messages. For a performance measure that incorporates the time to complete, as well as the number of involved respondents, the performance of the staged TDMA protocol can be optimized by selecting the groups appropriately. Observe, however, that if the collection objective is to identify all active respondents, as is the case in [6, 8], then no advantage is gained by staging the TDMA collection procedure. In such situations a single-stage TDMA procedure is always superior to polling or to staged TDMA.

5.2 Analysis and Optimization

We first observe that the average number of respondents that need to be enabled in a staged TDMA protocol will be the same as the average number enabled in a (single-stage) TDMA protocol. Thus we have that for all collection objectives y: $D^{(y)}_{st-TDMA} = D^{(y)}_{TDMA} = R^{(y)}$, where $R^{(y)}$ is given in equation (4), for the L or Maximum collection objective. We emphasize that the above is true regardless of the method of staging (fixed or adaptive) or of the actual grouping used. We next consider the other two performance measures.

5.2.1 Fixed Groups

When fixed-group staged TDMA is used a set of M disjoint groups, g_i, for $i = 1, \ldots, M$ are given. The size of group g_i is given by n_i and $\sum_{i=1}^{M} n_i = N$. (We address the determination of the best such grouping shortly.) We let $\underline{n} = (n_1, n_2, \ldots, n_M)$. Both the collector and respondent solicitation costs will be a function of $\underline{n}$. We also define the integer $1 \le J(k) \le M$ as the smallest integer such that $\sum_{i=1}^{J(k)} n_i \ge k$ for $1 \le k \le N$. This represents the number of groups that need to be solicited if k respondents should be enabled.

Using the same reasoning as that used in the derivation of equation (4), we obtain the following set of expressions:

$$\begin{aligned} C^{(L)}(\underline{n}) &= M \sum_{k=0}^{L-1} \binom{N}{k} p^{N-k} q^k \\ &+ \sum_{k=L}^{N} J(k) \binom{k-1}{L-1} p^{k-L} q^L \end{aligned} \quad (5)$$

$$\begin{aligned} S^{(L)}(\underline{n}) &= N \sum_{k=0}^{L-1} \binom{N}{k} p^{N-k} q^k \\ &+ \sum_{k=L}^{N} \sum_{i=1}^{J(k)} n_i \binom{k-1}{L-1} p^{k-L} q^L \end{aligned} \quad (6)$$

We next investigate how the total cost of the fixed-group, staged TDMA procedure may be optimized by selecting the appropriate fixed group sizes. For any given collection objective y the total cost is given by:

$$\begin{aligned} A^{(y)}(\underline{n}) &= \alpha R^{(y)} + \beta S^{(y)}(\underline{n}) + \gamma C^{(y)}(\underline{n}) \\ &= \alpha R^{(y)} + B^{(y)}(\underline{n}) \end{aligned} \quad (7)$$

Thus minimizing the total cost can be achieved by minimizing $B^{(y)}(\underline{n})$ which satisfies the following recursive equation:

$$\begin{aligned} B^{(\ell)}(\underline{v}) &= \beta v_1 + \gamma \\ &+ \sum_{x=0}^{v_1} B^{(\ell-x)}(\underline{v}^{-1}) \binom{v_1}{x} p^{v_1-x} q^x \end{aligned} \quad (8)$$

where $\underline{v} = (v_1, v_2, \ldots, v_T)$ and $\underline{v}^{-1} = (v_2, v_3, \ldots, v_T)$. The cost in equation (8) is derived as the sum of the solicitation cost for the first group (of size v_1) plus the cost of the collection protocol as it proceeds through the rest of the groups with a diminished collection objective. To complete the expression in (8) we need the following boundary conditions where $\underline{v} = (v_1, v_2, \ldots, v_T)$:

$$B^{(j)}(\underline{v}) = 0 \qquad for \quad j \leq 0 \tag{9}$$

$$B^{(\ell)}(v_1) = \beta v_1 + \gamma \qquad for \quad 0 < \ell \leq v_1 \tag{10}$$

$$B^{(\ell)}(\underline{v}) = \beta \sum_{i=1}^{T} v_i + \gamma T \; for \; \ell > \sum_{i=1}^{T} v_i \tag{11}$$

This last boundary condition warrants some explanation: This objective is equivalent to the objective "find all active" which requires the solicitation and enabling of all the remaining respondents thus the solicitation cost is as given above.

In order to determine the best fixed grouping that will minimize the $B^{(\ell)}(\underline{n})$ for a given q, β and γ, one straightforward method is to enumerate all potential groupings of the N respondents and evaluate the cost of each (using (8) or the appropriate expressions in (5) - (6)). The optimum grouping is the one with the minimum such cost. The approach just described is obviously not feasible as it is prohibitively time consuming even for moderate values of N. We thus adopt a heuristic approach aimed at determining a near-optimal grouping. Our approach is based on the assumption that the optimal solicitation cost for achieving objective y given the set of m respondents are subdivided into T groups, satisfies the following recursive relationship (based on (8)):

$$B_{opt}^{(\ell)}(\underline{v}_{T,m}) = \min_{0<v_1 \leq m-T+1} \{\beta v_1 + \gamma + B_{opt}^{(rnd[\ell - q v_1])}(\underline{v}_{T-1,m-v_1})\} \tag{12}$$

where $rnd[\bullet]$ rounds its argument to the nearest integer. Note that if a group of size v_1 is tested, an average of $q v_1$ responses are anticipated. Thus, the expression in (12) is based on the assumption that the optimal grouping for finding exactly ℓ is approximated by the optimum grouping for finding an average of ℓ.

The optimization procedure is thus as follows:

1. For each possible number of groups $M = 1, \ldots, N$ determine the grouping of the respondents into M groups using (12) and the following boundary conditions

$$B^{(j)}(\underline{v}_{T,m}) = 0 \qquad for \quad j \leq 0 \tag{13}$$

$$B^{(\ell)}(\underline{v}_{1,m}) = \beta m + \gamma \quad for \quad 0 < \ell \leq m \tag{14}$$

$$B^{(\ell)}(\underline{v}_{T,m}) = \beta m + \gamma T \; for \; \ell > m \tag{15}$$

   Note that the values of the boundary conditions above are independent of the grouping used. Thus, whenever, while using (12), the value of B is evaluated using these boundary conditions, we assume that the T remaining groups are such that the last group contains $m - T + 1$ respondents and the other $T - 1$ groups contain one respondent each.

2. Choose the grouping (from among the N different ones produced in Step 1) that yields the lowest cost as evaluated by (8).

We can judge how near-optimal the grouping found using the heuristic above by comparing its cost to the cost of the best adaptive grouping (as determined in the next subsection). This latter cost is a lower bound on the best fixed-group cost.

5.2.2 Adaptive Groups

In contrast to the fixed-groups protocol described above where a given set of M groups are used, the groups used in an adaptive protocol can be described by a tree. The root of the tree represents the group solicited in the first stage. The number of active respondents discovered at the end of the first stage determines the size of the next group to be solicited if the objective has not been reached. The same occurs at the conclusion of the second stage and so on.

Rather than proceeding as before by analyzing the protocol for any given tree of groups, we proceed directly to the discussion of the optimization step. As in the fixed-group staged TDMA case, the only part of the cost that is amenable to optimization is the total solicitation cost. We define $B_{opt}^{(y)}(m)$ as the optimum total solicitation cost when the number of unsolicited respondents is m. Using the same arguments leading to (12) we can write:

$$B_{opt}^{(\ell)}(m) = \min_{0<n\leq m} \left\{\beta n \gamma + \sum_{x=0}^{n} B_{opt}^{(\ell - x)}(m-n) \binom{n}{x} p^{n-x} q^{x}\right\} \tag{16}$$

The above equation with the following boundary conditions can be used to determine the optimum tree of groups that needs to be used.

$$B_{opt}^{(j)}(m) = 0 \qquad for \quad j \leq 0 \tag{17}$$

$$B_{opt}^{(j)}(0) = 0 \quad ; \quad B_{opt}^{(1)}(1) = \beta + \gamma \tag{18}$$

$$B_{opt}^{(\ell)}(m) = \beta m + \gamma \quad for \quad \ell > m \tag{19}$$

This last boundary condition stems from the fact that, as the objective stated is equivalent to finding all active users, it is best to use a single stage (i.e., a single broadcast solicitation message) which will incur the given solicitation cost.

6 Group Testing

6.1 Description

The group testing response collection procedure is initiated by a broadcast solicitation message sent by the collector and received by all respondents. Once this message is received by all respondents, the channel operates in the slotted mode where a group of respondents is enabled at the beginning of each slot. The choice of group to enable is determined entirely by the respondents by observing the channel activity and does not require intervention by the collector. The protocol operates in a similar manner to the one described in [6], with the major difference being that the protocol will terminate whenever the collection objective is achieved.

Each respondent observes the channel activity during each slot and updates its knowledge of the state of the respondents accordingly. The state of the system is described by membership in four sets [6, 1]: a *classified set*, a *binomial set*, a *defective set*, and a *conflicted set*. More details on the operation of this protocol, including the definitions of these sets can be found in [6].

6.2 Analysis

The size of the groups enabled can be found through the solution of a set of recursive equations shown below. In the following $H^{(y)}(n)$ denotes the average number of slots needed to satisfy the collection objective y when the binomial set is of size n, $F^{(y)}(m,n)$ denotes the average number of slots when the defective set is of size m and the binomial set is of size $n-m$ and $G^{(y)}(k,m,n)$ denotes the average number of slots when the defective set is of size k, the conflicted set is of size m and the binomial set is of size $n-m$. Then we can write for $\ell \geq 1$: (These equations are similar to those shown in [6]. There are, however, small but critical differences that have to do with the fact that the collection objective is now an influencing parameter.)

$$\begin{aligned} H^{(\ell)}(n) \; &= \\ 1 + \min_{1\leq x\leq n} &\{P_0 H^{(\ell)}(n-x) + P_1 H^{(\ell-1)}(n-x) \\ &+(1-P_0-P_1)G^{(\ell)}(0,x,n)\}, \quad n \geq 1 \end{aligned} \tag{20}$$

$$F^{(\ell)}(m,n) \;=\; 1+\min\{A_1, A_2\}, \quad n \geq 1 \tag{21}$$

$$\begin{aligned} A_1 \;&=\; \min_{1\leq x\leq m}\{P_2 F^{(\ell)}(m-x,n-x) \\ &+P_3 H^{(\ell-1)}(n-x) + (1-P_2-P_3)G^{(\ell)}(0,x,n)\} \\ A_2 \;&=\; \min_{m<x\leq n}\{P_4 H^{(\ell-1)}(n-x) \\ &+\; (1-P_4)G^{(\ell)}(m,x,n)\} \end{aligned}$$

$$\begin{aligned} G^{(\ell)}(0,m,n) \;&= \\ 1 + \min_{1\leq x<n}&\{P_5 G^{(\ell)}(0,m-x,n-x) + P_6 \\ +F^{(\ell-1)}(m-x,n-x) &+ (1-P_5-P_6)G^{(\ell)}(0,x,n)\}, \\ &n\geq m\geq 2 \end{aligned} \tag{22}$$

$$\begin{aligned} G^{(\ell)}(k,m,n) \;&=\; 1+\min\{B_1,B_2\}, \\ n\geq m\geq 2, &\quad m-1\geq k\geq 1 \end{aligned} \tag{23}$$

$$\begin{aligned} B_1 \;&=\; \min_{1\leq x\leq k}\{P_7 G^{(\ell)}(k-x,m-x,n-x) \\ +P_8 F^{(\ell-1)}(m-x,n-x) &+ (1-P_7-P_8)G^{(\ell)}(0,x,n)\} \\ B_2 \;&=\; \min_{k<x\leq m}\{P_9 F^{(\ell-1)}(m-x,n-x) \\ &+\; (1-P_9)G^{(\ell)}(k,x,n)\} \end{aligned}$$

where P_0 = probability that no transmission occurs = q^x and P_1 = probability that exactly one transmission occurs = $xq^{x-1}p$. The expressions for P_2 through P_9 are identical to those in [6, equations (5.1)-(5.4)] and are not repeated here.

In addition, the following boundary conditions are applicable:

$$H^{(0)}(n) = H^{(\ell)}(0) = 0; \;\; F^{(0)}(m,n) = 0 \tag{24}$$

$$F^{(\ell)}(0,n) = H^{(\ell)}(n) \tag{25}$$

$$F^{(\ell)}(1,n) = 1 + H^{(\ell-1)}(n-1) \tag{26}$$

$$G^{(0)}(k,m,n) = G^{(\ell)}(0,1,n) = 0 \tag{27}$$

$$G^{(\ell)}(1,m,n) = 1 + F^{(\ell-1)}(m-1,n-1) \tag{28}$$

$$G^{(\ell)}(k,2,n) = 2 + H^{(\ell-2)}(n-2) \tag{29}$$

Since all respondents are initially in the binomial set we have that the average number of slots needed to achieve the collection objective is $D^{(y)}_{GT} = H^{(y)}(N)$. If the group testing collection protocol is used, a single solicitation message is sent which reaches all N respondents and thus we have for all collection objectives y: $C^{(y)}_{GT} = 1$ and $S^{(y)}_{GT} = N$.

7 Staged Group Testing

7.1 Description

Staged group testing is a generalization of both the polling and (the single stage) group testing protocols. The protocol operates in a similar manner to the staged TDMA protocol. The set of respondents are solicited in a set of stages, where each stage begins by a multicast solicitation message sent to a subset of the respondents. As before these solicitation messages contain the collection objective and are modified from the initial objective as responses are collected in each stage. The groups of respondents solicited in different stages are disjoint.

Whereas in the staged TDMA protocol solicited respondents are enabled during slots allocated to each individually, in the staged group testing protocol solicited respondents are enabled according to a group testing procedure that involves only the members of the group being explored during the current stage. Within stage i a group testing procedure that involves only the respondents in group g_i is carried out. At the beginning of each slot in stage i the respondents involved know the distribution of the respondents in g_i into each of the classified, binomial, defective and conflicted sets. In addition they know the collection objective (declared by the collector in its ith solicitation message) and the number of remaining and not-yet-solicited respondents. This information allows the procedure to terminate before all the respondents in g_i have been enabled when either the collection objective is achieved or it is determined that the objective is not attainable as the number of remaining respondents in this and further stages is not sufficient.

We again distinguish between fixed and adaptive groups as was done for the staged TDMA protocol. By appropriately selecting the groups used in either a fixed or adaptive group staged group testing procedure, its cost may be minimized.

7.2 Analysis and Optimization

7.2.1 Fixed Groups

As before we are given a set of M disjoint groups defined by the vector $\underline{n}$ whose elements are the sizes of each group. We first observe that for a given value of $\underline{n}$, the respondent solicitation cost (i.e., the number of respondents solicited) and the collector solicitation cost (i.e., the number of solicitation message sent by the collector) will be identical to those obtained for the fixed groups staged TDMA protocol in section 5.2.1. We thus assert that for any collection objective y:

$$C^{(y)}_{FG,st-GT}(\underline{n}) = C^{(y)}_{FG,st-TDMA}(\underline{n}) \quad and \tag{30}$$

$$S^{(y)}_{FG,st-GT}(\underline{n}) = S^{(y)}_{FG,st-TDMA}(\underline{n}) \tag{31}$$

Equations (5) - (6) contain the appropriate expressions for the L or Maximum collection objective.

It remains to determine the average delay cost (or the average number of response slots required for the collection objective to be satisfied). We first note that unlike the staged TDMA protocol where the average delay cost was the same as that incurred by the single stage TDMA protocol, here staging will, in general, affect the delay cost. In order to capture the fact that the group testing procedure within any one stage operates with the knowledge of the number of the not-yet-solicited respondents, we define: $\mathcal{H}^{(y)}(n;t)$ as the average number of slots remaining in a stage when the binomial set is of size n, the number of not-yet-solicited respondents is t and the current collection objective is y. $\mathcal{F}^{(y)}(m,n;t)$ and $\mathcal{G}^{(y)}(k,m,n;t)$ are defined in a similar manner. The above three quantities are related in exactly the same way as the corresponding quantities in section 6.2. The boundary conditions for these quantities are essentially the same as those shown in section 6.2.

We note the following equivalence relations:

$$\mathcal{H}^{(y)}(n;0) = H^{(y)}(n); \quad \mathcal{H}^{(L)}(n;t) = H^{(L)}(n) \tag{32}$$

The average delay cost satisfies the following recursive relationship:

$$D^{(\ell)}(\underline{v}) = \mathcal{H}^{(\ell)}(v_1; \sum_{i=2}^{T} v_i) + \sum_{x=0}^{v_1} D^{(\ell-x)}(\underline{v}^{-1}) \binom{v_1}{x} p^{v_1-x} q^x \tag{33}$$

where $\underline{v}$ and $\underline{v}^{-1}$ are as defined in section 5.2.1.

In addition we use the following boundary conditions

$$D^{(j)}(\underline{v}) = 0 \qquad for \quad j \le 0 \tag{34}$$

$$D^{(\ell)}(v_1) = H^{(\ell)}(v_1) \quad for \quad 0 < \ell \le v_1 \tag{35}$$

$$D^{(\ell)}(\underline{v}) = \sum_{i=1}^{T} H^{(v_i)}(v_i) \; for \; \ell > \sum_{i=1}^{T} v_i \tag{36}$$

For a given set of weights, α, β, and γ, a near-optimum set of fixed groups can be determined by assuming the following recursive equation for the optimal total cost, $A^{(y)}_{opt}(\underline{v}_{T,m})$ for achieving objective y given a set of m respondents are subdivided into T groups. (see equation (12) and arguments leading to it):

$$A^{(\ell)}_{opt}(\underline{v}_{T,m}) = \min_{0<v_1 \le m-T+1} \{ \alpha \mathcal{H}^{(\ell)}(v_1; m-v_1) + \beta v_1 + \gamma + A^{(rnd[\ell - qv_1])}_{opt}(\underline{v}_{T-1,m-v_1}) \} \tag{37}$$

where the following boundary conditions are obeyed:

$$A^{(j)}_{opt}(\underline{v}_{T,m}) = 0 \qquad for \quad j \le 0 \tag{38}$$

$$A^{(\ell)}_{opt}(\underline{v}_{T,m}) = A^{(m)}_{opt}(\underline{v}_{T,m}) \; for \; \ell > m \tag{39}$$

$$A^{(\ell)}_{opt}(\underline{v}_{1,m}) = \alpha H^{(\ell)}(m) + \beta m + \gamma \; for \; 0 < \ell \le m \tag{40}$$

The optimization procedure using (37) is described in section 5.2.1. Also, as for the staged TDMA case, the adaptive groups optimum cost derived in the next subsection can be used as a lower bound by which we can judge the "goodness" of our heuristically obtained fixed grouping.

7.2.2 Adaptive Groups

The optimum groups to use in an adaptive-group, staged group testing procedure can be determined by considering the following recursive equation (see equation (16) and arguments leading to it):

$$A^{(\ell)}_{opt}(m) = \min_{0<n\le m} \{ \alpha \mathcal{H}^{(\ell)}(n; m-n) + \beta n + \gamma + \sum_{x=0}^{n} A^{(\ell-x)}_{opt}(m-n) \binom{n}{x} p^{n-x} q^x \} \tag{41}$$

where the following boundary conditions are satisfied:

$$A^{(j)}_{opt}(m) = 0 \qquad for \quad j \le 0 \tag{42}$$

$$A^{(j)}_{opt}(0) = 0 \quad ; \quad A^{(1)}_{opt}(1) = \alpha + \beta + \gamma \tag{43}$$

$$A^{(\ell)}_{opt}(m) = \alpha H^{(m)}(m) + \beta m + \gamma \; for \; \ell > m \tag{44}$$

8 Numerical Examples

8.1 Polling and TDMA

First we consider the performance of polling and TDMA protocols. The quantity of interest for both protocols is $R^{(\nu)}$ (see section 4.2). For the polling protocol, this quantity indicates the average number of polling messages sent (which is the same as the number of response slots required and the number of respondents enabled). For the TDMA protocol, the quantity indicates the number of response slots required. We refer to this measure generically as the average number of "steps". The variation of the average number of steps is shown as a function of q for the L or Maximum collection objective in Figure 1. Systems where the number of respondents N is 20 are considered. (Recall that for the TDMA protocol, exactly one solicitation message is sent and that all N respondents are solicited regardless of the collection objective and of the value of q.)

For the L or Maximum collection objective (Figure 1) as the likelihood of receiving a response from a respondent increases, the average number of steps required decreases. The average number of steps increases if the value of L required to achieve the collection objective increases.

In general, polling will be preferred over TDMA only if the weight of the respondent solicitation cost, β, is high as polling's only advantage is that less respondents are solicited.

8.2 Staged TDMA

The delay cost of the staged TDMA procedure is the same as that for a TDMA protocol. The variations of this cost with q are shown in Figure 1. As mentioned earlier, the only advantage of a staged TDMA protocol is that it may incur a lower respondent solicitation cost at the expense of a slightly higher collector solicitation cost. As an example, we consider achieving the L or Maximum collection objective using a fixed group staged TDMA procedure in a system where the 20 respondents are subdivided into four groups of sizes (6, 6, 4, 4). The respondent and collector solicitation costs are shown as a function of q in Figures 2 and 3, respectively. Whereas in the single stage TDMA protocol the respondent solicitation cost is always 20, using a staged procedure can provide for a lower cost especially for high values of q. This is achieved at the expense of increasing the collector solicitation cost from the value 1 in the single stage TDMA protocol to a value between 1 and 4 (as shown in Figure 3) in this case.

Table 1 shows near-optimal groupings of 15 respondents in a fixed-group procedure for various values of the cost weights and for the L or Maximum collection objective. The table also shows the cost of using an optimal adaptive-group, staged TDMA procedure. Observe that the heuristically obtained fixed groupings result in the same or slightly higher costs than the optimal adaptive groupings. The equality in cost happens when the optimal adaptive grouping is equivalent to the fixed grouping shown. Note also that the "best" fixed-group staged TDMA procedure is sometimes one where there is one group of size 15 or 15 groups of size one. This matches our intuition that in certain instances, a single stage TDMA procedure or polling will be best.

8.3 Group Testing

Group testing (in a single stage) incurs a respondent and collector solicitation costs of N and 1, respectively regardless of the value of q. Figure 4 shows the average delay cost as a function of q for the L or Maximum collection objective. By comparing to Figure 1, we make the following observations: 1)For the entire range of q the use of group testing provides for a lower or equal delay cost than a single stage TDMA protocol. 2)The group testing collection procedure adapts to the (single stage) TDMA procedure (i.e., in each slot a group of size 1 is enabled) when $q \geq \frac{1}{\sqrt{2}}$. (This was found to be true in all the numerical experiments we conducted. No formal proof is available yet.)

8.4 Staged Group Testing

The respondent and collector solicitation costs of the staged group testing procedure are the same as those for the staged TDMA procedure. For the L or Maximum collection objective the variations of these costs with q are shown in Figures 2 and 3, where the respondents are subdivided into four groups of sizes (6,6,4,4). The delay cost of the staged group testing procedure for the same grouping of the respondents is shown in Figure 5. Comparing this delay with that incurred by a single stage group testing procedure (as shown in Figure 4) we observe that: 1)For values of $q \geq \frac{1}{\sqrt{2}}$ the delay cost for the two approaches is the same since in both the group testing procedure adapts to a TDMA procedure. 2) For values of $q < \frac{1}{\sqrt{2}}$, the delay incurred is lower when single stage group testing is used. Particularly, in the limit as q approaches zero, the single stage group testing procedure needs only one group test to determine that the respondents are not active, whereas the staged group testing procedure needs a number of group tests equal to the number of groups of respondents.

Table 2 shows the grouping of 15 respondents in a fixed-group, staged group .esting procedure for various values of the cost weights and for the L or Maximum collection objective. It also shows the cost when an optimal adaptive-group, staged group testing procedure is used. Observe the following: 1)As in the staged TDMA case, the near-optimal fixed groupings achieve the same or slightly higher cost than the optimal adaptive groupings. 2)The costs of the staged group testing procedure are close to those of the staged TDMA procedure if the parameters are such that the optimal grouping results in small size groups. Otherwise, the staged group testing costs can be lower.

9 Concluding Remarks

In this paper we have considered response collection strategies that can be used over a multiple access channel. We were motivated by some distributed computing applications to define a set of collection objectives. Five protocols that can be used to achieve these collection objectives were investigated: Polling, TDMA, staged TDMA, group testing and staged group testing. In analyzing the performance of these protocols, three cost components were taken into account: the number of steps required to complete the objective, the number of solicitations required by the collector, and the the number of respondents receiving solicitation messages. The idea of staging stems from the inclusion of the latter two cost components and from the fact that the request procedure will terminate once the collection objective has been achieved.

Our findings are summarized in Table 3 where we use the terms low, medium, and high to denote relative values of the costs. Our conclusion is that, in general, a suitably optimized adaptive-group, staged group testing protocol can achieve the best performance. A near optimal fixed-group staged group testing procedure can achieve almost similar performance but can be easier to implement as the groups are determined a priori.

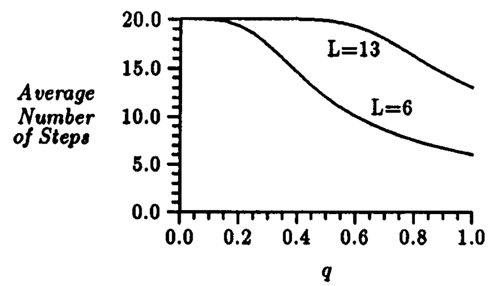

Figure 1: Average Number of Steps for the L or Maximum Collection Objectives (Polling or TDMA)

References

[1] M. Sobel and P. A. Groll, "Group testing to eliminate efficiently all defectives in a binomial sample," *The Bell Systems Technical Journal*, September 1959.

[2] F. K. Hwang, "On finding a single defective uisng binomial group testing," *Journal of the American Statistical Association*, vol. 69, pp. 146–150, MARCH 1974.

[3] M. R. Garey and F. K. Hwang, "Isolating a single defective using group testing," *Journal of the American Statistical Association*, vol. 69, pp. 151–153, March 1974.

[4] M. C. Hu, F. K. Hwang, and J. K. Wang, "A boundary problem for group testing," *SIAM Journal of Algebra and Discrete Mathematics*, vol. 2, pp. 81–87, June 1981.

[5] J. Wolf, "Born again group testing: Multiaccess communications," *IEEE Transactions on Information Theory*, vol. 31, pp. 185–191, March 1985.

[6] T. Berger, N. Mehravari, D. Towsley, and J. Wolf, "Random multiple-access communication and group testing," *IEEE Transactions on Communications*, vol. 32, pp. 769–779, July 1984.

[7] N. K. Garg and S. Mohan, "Group testing with capture for random access communication," *IEEE Transactions on Communications*, vol. 35, pp. 849–854, August 1987.

[8] D. Kurtz and M. Sidi, "Multiple-access algorithms via group testing for heterogeneous population of users," *IEEE Transactions on Communications*, vol. 36, pp. 1316–1323, December 1988.

[9] T. K. Sellis, "Multiple-query optimization," *ACM Transactions on Database Systems*, vol. 13, pp. 23–52, MARCH 1988.

[10] M. H. Ammar, "Teletext-like information delivery using broadcast polling," *Computer Networks and ISDN Systems*, vol. 12, pp. 107–115, MARCH 1987.

[11] M. H. Ammar and H. J. Kim, "Prototyping a broadcast delivery information system," Tech. Rep. GIT-ICS-90/16, Georgia Institute of Technology, Atlanta, GA, 1990.

[12] G. Herman, G. Gopal, K. Lee, and A. Weinrib, "The datacycle architecture for very high throughput databases," in *Proceedings of SIGMOD*, ACM, 1987.

[13] J. W. Wong and M. H. Ammar, "Analysis of broadcast delivery in a videotex system," *IEEE Transactions on Computer*, vol. 34, pp. 863–866, September 1985.

[14] L. Lamport, "The implementation of reliable distributed multiprocess systems," *Computer Networks*, vol. 2, pp. 95–114, 1978.

[15] D. Gifford, "Weighted voting for replicated data," in *Proceedings of 7th Symposium on Operating Systems*, pp. 150–162, ACM, 1979.

[16] M. Ahamad and M. H. Ammar, "Performance characterization of quorum-consensus algorithms for replicated data," *IEEE Transactions on Software Engineering*, vol. 15, pp. 492–496, April 1989.

[17] D. Barbara and H. Garcia-Molina, "Mutual exclusion in partitioned distributed systems," *Distributed Computing*, vol. 1, pp. 119–132, 1986.

[18] S. Y. Cheung, M. Ahamad, and M. H. Ammar, "Optimizing vote and quorum assignments for reading and writing replicated data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 1, pp. 387–397, September 1989.

[19] S. Mullender and P. Vitanyi, "Distributed matchmaking," *Algorithmica*, vol. 3, 1988.

[20] M. H. Ammar and G. N. Rouskas, "On the performance of protocols for collecting responses over a multiple access channel," Tech. Rep. GIT-ICS-90/53, Georgia Institute of Technology, Atlanta, GA, 1990.

[21] J. Bernabeu, M. H. Ammar, and M. Ahamad, "Optimal selection of multcast groups for resource location in a distributed system," in *Proceedings of INFOCOM '89*, IEEE, 1989.

Cost Measure	Polling	TDMA	Staged TDMA	Group Testing	Staged Group Testing
Delay Cost (no. of slots)	High	High (=Polling)	High (=Polling)	Low	Medium
Respondent Solicitation Cost	Low	High (= N)	Medium	High (= N)	Medium (= staged TDMA, for same grouping)
Collector Solicitation Cost	High	Low (=1)	Medium	Low (=1)	Medium (= staged TDMA, for same grouping)

Table 3: Relative Performance of the Various Protocols

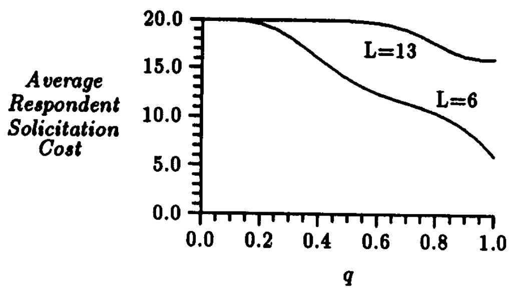

Figure 2: Average Respondent Solicitation Cost for the L or Maximum Collection Objectives (Fixed Group Staged-TDMA)

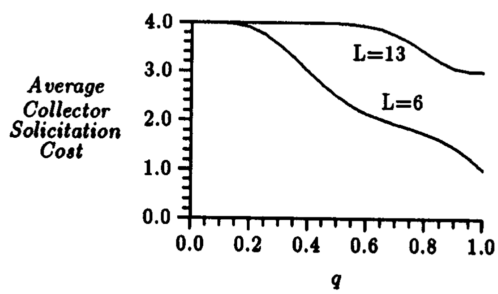

Figure 3: Average Collector Solicitation Cost for the L or Maximum Collection Objectives (Fixed Group Staged-TDMA)

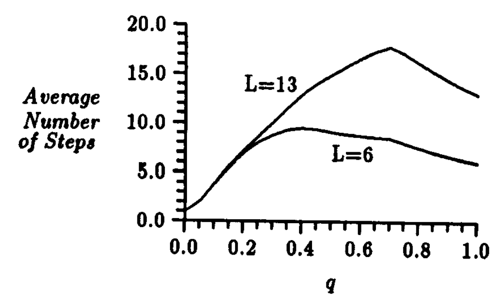

Figure 4: Average Delay Cost for the L or Maximum Collection Objectives (Group Testing)

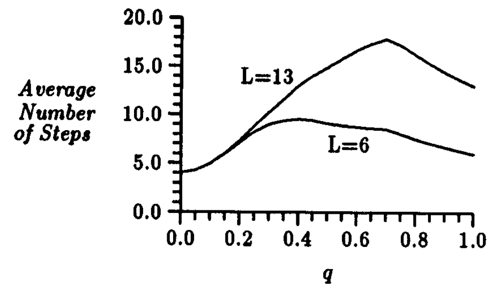

Figure 5: Average Delay Cost for the L or Maximum Collection Objectives (Fixed Group Staged Group Testing)

L	q	α	β	γ	Fixed Groups	FG cost	AG cost
3	.85	1	10	1	3 1 . . . 1	40.35	40.29
1	.9	1	10	1	1 . . . 1	13.33	13.33
1	.1	1	10	1	1 . . . 1	95.29	95.29
1	.1	1	.1	1	15	10.44	10.44
8	.8	1	.1	1	10 3 1 1	12.45	12.34
8	.1	1	10	1	15	166.00	166.00

Table 1: Near-Optimal Fixed Group and Optimal Adaptive Group Staged TDMA (L or Maximum)

L	q	α	β	γ	Fixed Groups	FG cost	AG cost
3	.85	1	10	1	3 1 . . . 1	40.35	40.29
1	.9	1	10	1	1 . . . 1	13.33	13.33
1	.1	1	10	1	2 2 2 1 . . . 1	93.35	92.09
1	.1	1	.1	1	15	4.33	4.33
8	.8	1	.1	0	10 3 1 1	12.45	12.28
8	.1	1	10	1	15	153.86	153.86

Table 2: Near-Optimal Fixed Group and Optimal Adaptive Group Staged Group Testing (L or Maximum)