Multicast Session Membership Size Estimation

Timur Friedman and Don Towsley

UMass Amherst Computer Science Department
friedman@cs.umass.edu, towsley@cs.umass.edu

Technical Report TR 98-32
14 August 1998

Abstract

The problem of estimating the number of members in a multicast session through probabilistic polling corresponds to that of estimating the parameter n of the $\text{Binomial}(n, p)$ distribution. This allows an interval estimator for n to be derived. The tradeoff between the relative dispersion of this estimator and the overhead it requires is characterized in a manner that may be mapped to application requirements. Based on the binomial model and its analysis, improvements are suggested to the probabilistic polling mechanisms described in [Bolot, Turletti, and Wakeman 1994], and [Nonnenmacher and Biersack 1998], and [Nonnenmacher 1998]. We also derive the maximum likelihood estimator (MLE) for the last of these.

1 Introduction

There is strong evidence that technologies underlying future wide-scale multicast applications will require estimates of multicast session size. In order to control feedback implosion, for instance, timer delays may be used to regulate feedback suppression. The tuning of these timers typically depends upon some estimate of session size. RTP, the multicast protocol for real-time audio and video transport described in RFC 1899 [Schulzrinne et al 1996], is an example of a protocol that sets timers based upon an explicit session size estimate. It uses the estimate to maintain its control packet bandwidth at a small portion of overall session bandwidth (see section 6.2.1 of the RFC). This same mechanism is used by SRM, the

Scalable Reliable Multicast framework, to restrict its session messages to 5% of overall session bandwidth [Floyd et al 1995].

There are parameters other than timer settings that also depend upon estimates of group size. Rubenstein, Kurose, and Towsley [1998] show that tuning the number of forward error correction (FEC) packets to the total number of data receivers can minimize the number of NAKs needed for reliable multicast.

Two approaches have been taken to group size estimation: direct counting and probabilistic polling. In RTP, each member counts all the other members based upon the receipt of control packets they have sent. A member multicasts a control packet when it joins a session, or remains in a session, and, usually, when it departs from a session. Timers are employed to remove from the count any members that have not been heard from over a long period. The risk of feedback implosion, inherent when all members send packets, is minimized by enforcing a periodic rate above which control packets may not be sent. (See [Rosenberg and Schulzrinne 1998] for a recent proposed extension to this counting and feedback control mechanism.)

RTP's use of this mechanism is tied to a many-to-many communication model that requires individual information from all participants. However, when it is not necessary for all participants to announce themselves, probabilistic polling may be used. The potential savings in either bandwidth or latency are large if only a fraction of the participants are called upon to send packets. Such an approach was first proposed by Bolot, Turletti, and Wakeman [1994] and refined by Nonnenmacher and Biersack [1998]. In this paper, we will refer to the mechanism proposed by Bolot, Turletti, and Wakeman as the BTW mechanism, and to that proposed by Nonnenmacher and Biersack as the NB mechanism. The BTW mechanism requires a succession of one or more polling rounds in order to obtain feedback for an estimate. The NB mechanism obtains feedback in a single polling round.

This paper addresses the problem of multicast session size estimation using end-to-end probabilistic polling by mapping it to the problem of estimating the parameter n of the $\text{Binomial}(n, p)$ distribution. By doing so, we are able to bring statistical tools to bear on the problem to:

- Derive an interval estimator for n that improves with successive polling rounds.
- Derive an upper bound on the minimum number of feedback packets required to generate a desired confidence interval for n.
- Describe the tradeoff between estimation quality and feedback overhead.
- Derive an unbiased point estimator for the NB mechanism.
- Derive the maximum likelihood estimator (MLE) for the NB mechanism.

In addition, we propose and evaluate a solution to a problem that arises with the NB mechanism due to variability in the end-to-end delays between pairs of receivers. The problem is that the NB estimator is constructed under the assumption that these delays are fixed, and is therefore inaccurate when this is not the case. We introduce a mechanism that separates the timer settings in the NB scheme from the end-to-end network delays and we prove that it is accurate in the face of heterogeneous delays.

The outline of this paper is as follows. Section 2 describes the binomial estimation model of probabilistic polling, introducing the notation that will be used throughout. Section 3 is an example of a probabilistic polling session, showing how the analytic tools we develop in this paper may be applied. Sections 4, 5, and 6 give the derivations of those tools: the interval estimator, a point estimator for the maximum number of replies that might result from a poll, and a special interval estimator for the case when there are zero replies to a poll. Sections 7 and 8 examine the BTW and NB mechanisms. We discuss related work in Section 9.

2 The binomial estimation model of probabilistic polling

In this section we describe the binomial estimation model for probabilistic polling and we show how it applies to the BTW and NB models. [Bolot, Turletti, and Wakeman 1994] and [Nonnenmacher 1998] have provided point estimators that have been shown, under a number of different simulation scenarios, to closely track true group size while avoiding feedback implosion. However, no systematic analysis was given to either establish the quality of their estimates or to determine the tradeoffs between their quality and the polling overhead they incur. In this section, we place the problem of probabilistic polling group size estimation in a mathematical context within which we will be able to ask and answer a number of questions. In particular, we will be concerned with the following:

- Can we know, on the basis of polling parameters, what quality of estimate to expect?
- Can anything be said about the amount of polling overhead required to obtain an estimate of a given quality?
- Can polling parameters be matched to an application's requirements, both for a certain quality of estimate and for an estimate that can be achieved within certain bandwidth and latency constraints?

This paper's main contribution is to answer all of these questions in the affirmative.

2.1 Description of the binomial estimation model

Our analytic tool is a model of probabilistic polling group size estimation as estimation of the parameter n in the $\text{Binomial}(n, p)$ distribution. Under this model, the action of receiver j, $j = 1 \ldots n$, in a given polling round i is modeled as an individual sample from identically distributed random variables $X_{i,j} \sim \text{Bernoulli}(p_i)$. (At this level of generality we do not yet specify how the parameter $p_i = \Pr\{X_{i,j} = 1\}$ is defined.) The receiver sends a response if there is a success, defined by $X_{i,j} = 1$, and sends no response if $X_{i,j} = 0$. The number of receiver responses, the result of n independent Bernoulli trials, is binomial. We define $R_i \sim \text{Binomial}(n, p_i)$ to be the random variable that describes the number of receiver responses r_i in round i. The general problem is to construct an estimate $\hat{n}$ that, after k polling rounds, using knowledge of polling probabilities $\boldsymbol{p} = \{p_1, p_2, \ldots, p_k\}$ and the observations $\boldsymbol{r} = \{r_1, r_2, \ldots, r_k\}$, does a good job of estimating n.

For this model, we make two simplifying assumptions: that end-to-end network delays are bounded, and that there is no end-to-end loss. Unbounded delays pose a problem in determining when an observation is complete. End-to-end losses make it difficult to distinguish lack of a reply from loss of a reply. We believe that our model could be extended to deal, at least in part, with a relaxation of these assumptions (for example, delays and losses can be observed when reliability mechanisms are introduced).

2.2 Mapping of the model to the BTW and NB mechanisms

Let us see how our model describes the two existing polling mechanisms.

The BTW mechanism consists of a series of discrete polling rounds. For each round, the reply probability p_i is defined by the sender in the request for feedback that it sends. The number of responses obtained for a round is r_i. This maps directly to our model. The polling probability is set to a low value in the first round, to avoid feedback implosion should n be large. After each round, if no responses are obtained, the reply probability is roughly doubled for the subsequent round. Polling stops when a non-zero response is obtained, and a group size estimate is made based upon the cardinal number of that round.

In the NB mechanism, the situation is a bit more complicated. Instead of sending a reply probability, the sender initiates polling round i by multicasting the parameters of a timer distribution Z_i. Each receiver j sets a timer $z_{i,j}$ based upon its independent sample from Z_i. When a receiver's timer expires, it multicasts a reply packet. Feedback suppression takes place when a receiver receives another's reply packet. If, upon receiving a reply packet, its timer has not yet expired, it cancels its timer. Because at least one receiver is guaranteed to reply, only one polling round is required in order to obtain a reply. If Z_i is tuned properly, feedback implosion is kept to a minimum.

The NB mechanism assumes uniform inter-host delays; the time it takes a message to go from one host to another is a constant c. All receivers thus receive the request for feedback simultaneously, and set their timers at that moment. Without loss of generality, let us order the receivers $j = 1,\ldots,n$ from smallest to largest timer setting. Feedback suppression takes place when the reply from receiver 1 reaches all other receivers. Thus, only those receivers that have timers set in the interval $\left[z_{i,1}, z_{i,1}+c\right]$ will send replies.

Knowing the timer setting $z_{i,1}$ of the first timer to expire, and the distribution, Z_i, from which the timer settings were taken, we can determine the probability that any other receiver j will have a timer set in this interval. Let F_{Zi} represent the cumulative distribution function (CDF) of Z_i. The reply probability, p_i, is a conditional probability, conditioned upon the timer not having expired prior to $z_{i,1}$:

$$\begin{aligned} p_i &= \Pr\left\{z_{i,1} \le z_{i,j} \le z_{i,1}+c \middle| z_{i,j} > z_{i,1}\right\} \\ &= \frac{\Pr\{z_{i,1} \le z_{i,j} \le z_{i,1}+c\}}{\Pr\left\{z_{i,j} > z_{i,1}\right\}} \\ &= \frac{F_{Zi}\left(z_{i,1}+c\right) - F_{Zi}\left(z_{i,1}\right)}{1 - F_{Zi}\left(z_{i,1}\right)} \end{aligned}$$

It would seem that, once the probability p_i is known, the binomial model could be applied directly. However, p_i is not the reply probability for all n receivers. The first receiver to reply sets this probability and is not itself subject to it. Thus, the binomial model may be used to estimate $n-1$ on the basis of $r_i - 1$ of the replies. We will demonstrate, in Section 8 of this paper, how this probability, p_i, can be used to construct an unbiased point estimator for n under the NB mechanism.

3 An example

We have established that probabilistic polling can be modeled, in whole or in part, as the estimation of the parameter n in the $\mathrm{Binomial}(n, p)$ distribution. We now present an example to motivate our derivation from that model of a number of useful interval estimators and other statistical measures.

Suppose a multicast application requires an estimate of the number of participants in its session and it knows that this number is less than one million. It does not require an estimate that is extremely precise, just that the true session size lie within $\pm 90\%$ of the estimated value. Because the application is bandwidth-limited, the number of feedback packets that it receives in making the estimate must not exceed 30. Neither the BTW nor the NB mechanisms permit the application to construct an estimator with these properties in a systematic manner. In this section, we show how an interval estimator and other analytic tools, developed in subsequent sections, can be used to address this problem. Polling in this example will take place over several rounds at sender-defined polling probabilities, much like the BTW mechanism.

The application in our example requires an estimator that gives some guarantee that the true session size lies within $\pm 90\%$ of the value estimated. Thus we need not only a point estimate, $\hat{n}$, but also an interval length no greater than $0.90\hat{n}$ that contains the point estimate. We refer to the ratio of the length of the interval to $\hat{n}$ as β ($\beta = 0.90$ in the example). We will show, in section 4.2, that an upper bound on the amount of overhead, h, required to obtain a given β at the $100(1-\alpha)\%$ confidence level is

$$h \geq \left\lceil \frac{(1+\beta) z_{\alpha/2}^2}{\beta^2} \right\rceil$$

where $z_{\alpha/2}$ is the value for which the CDF of the standard normal distribution $\Phi(z_{\alpha/2}) = (1-\alpha/2)$. To be fairly certain of our estimate we will use a 95% confidence level, so $h = 10$, i.e., an average of ten replies are required. Since the application may safely receive 30 replies, the reply probabilities of the polling mechanism will be tuned to yield an average number of replies in the interval $[10, 30]$.

The reply probability, p, together with the true number of receivers, n, predicts the expected number of replies, $E(R) = np$, R a random variable. When n is unknown and we wish to place an upper bound on $E(R)$ we must either guess or estimate an upper bound for n. In our example, we know a priori that $n \leq 10^6$, and so we must restrict our polling probability to lie in the interval $[10 \times 10^{-6}, 30 \times 10^{-6}]$. For a tight interval it is best to obtain as many replies as possible without exceeding the maximum. If $p = 30 \times 10^{-6}$, $E(R) = 30$ at the upper bound for n, but there would be a good chance that $r > 30$, exceeding the maximum. We will show, in Section 5, that, at the $100(1-\alpha)\%$ confidence level, a maximum number of replies $r_{\max}$ will not be exceeded if

$$p \le \frac{2nr_{\max} + nz_\alpha^2 - z_\alpha \sqrt{n}\sqrt{4nr_{\max} - 4r_{\max}^2 + nz_\alpha^2}}{2n\left(n + z_\alpha^2\right)}$$

where z_α is the value for which the CDF of the standard normal distribution $\Phi(z_\alpha) = (1-\alpha)$. Because we would like to be well assured of not exceeding the maximum number of replies, we will use a 99% confidence level to determine p. At this confidence level, $p = 18.8 \times 10^{-6}$ is the maximum acceptable polling probability.

Now suppose there are no replies to the initial poll, so $r_1 = 0$ is observed. Such an observation would tend to discount the initial guess that n could be as high as 10^6. We will show, in Section 6, that, at the $100(1-\alpha)\%$ confidence level, the maximum possible value for n, following k consecutive polling rounds in which zeroes were observed, is

$$\hat{n}_0 = \left\lceil \frac{\ln \alpha}{\sum_{i=1}^{k} \ln(1 - p_i)} \right\rceil$$

where the values p_i are the polling probabilities for each round $i = 1 \ldots k$. To be conservative in our estimate of the maximum value of n, we will calculate it at the 99% confidence level. At that level, we know that $\hat{n}_0 \le 2.45 \times 10^5$. Even though no replies were received, we have been able to tighten the interval in which n may lie, from its initial $\left[0, 1.00 \times 10^6\right]$ to $\left[0, 2.45 \times 10^5\right]$.

In each subsequent polling round i, if zero replies were received the prior round, a new maximum value may be calculated for n. Based upon that calculation, a new maximum safe p_i can be established. Because $\hat{n}_0$ will decrease with each successive round in which there is a zero observation, p_i will increase. If zeroes continue to be observed until $\hat{n}_0 \le 30$, the maximum acceptable overhead, $p = 1$ may be used in a last polling round to obtain an exact estimate (a zero observation when $p = 1$ means there truly are zero receivers). The following table shows values for $\hat{n}_0$ and p_i for each round, assuming non-zero responses are obtained. The right-hand column shows the expected number of replies should n be the maximum possible value, i.e. if $n = \hat{n}_0$. These figures for $E(R)$ reflect the safety margin built in to the maximum safe value for p. If no such margin were built in, $E(R)$ would be 30, the overhead limit. With an $E(R)$ of 18.8 (somewhat higher for lower $\hat{n}_0$ and higher p), we know, with a confidence level of 99%, that no more than 30 replies will be received.

round i	$\hat{n}_0$	p_i	$E(R; n = \hat{n}_0)$
1	1.00×10^{6}	1.88×10^{-5}	18.8
2	2.45×10^{5}	7.68×10^{-5}	18.8
3	4.81×10^{4}	3.91×10^{-4}	18.8
4	9.47×10^{3}	1.99×10^{-3}	18.8
5	1.86×10^{3}	1.01×10^{-2}	18.9
6	3.64×10^{2}	5.23×10^{-2}	19.0
7	7.0×10^{1}	2.89×10^{-1}	20.2
8	1.2×10^{1}	1.00×10^{0}	12.0

Finally, suppose that 16 replies are obtained on round 4 and polling stops, since $r_4 = 16$ exceeds the ten replies that we require in order to estimate n to ±90%. We will show, in Section 4.1, that the following estimator interval will provide us with a good center point and confidence interval for observations at a fixed probability p:

$$\hat{n} = \frac{\bar{r}}{p} \pm \frac{z_{\alpha/2}}{2kp}\left((1-p)z_{\alpha/2} + \sqrt{(1-p)4k\bar{r} + (1-p)^2 z_{\alpha/2}^2}\right)$$

This is applied to the last observation alone at probability $p_4 = 1.99 \times 10^{-3}$, $k = 1$, and $\bar{r} = 16$. Suppose we use a 95% confidence level. Then $\hat{n} = 8{,}049 \pm 5{,}021$, or $\hat{n} = 8{,}049 \pm 63\%$. Should we wish to refine this estimate, we could continue polling at the same probability and apply this estimator to all subsequent observations.

4 Interval estimator

In this section, we derive an interval estimator for n when one or more replies have been observed at a constant p, and we establish two useful properties of the estimator. More precisely, we compute an interval $[n_l(\alpha), n_u(\alpha)]$ which the parameter n lies within with a $100(1-\alpha)\%$ confidence level. The first property concerns the length of the interval relative to the magnitude of its center, i.e., $2(n_u(\alpha) - n_l(\alpha))/(n_l(\alpha) + n_u(\alpha))$, a measure we call β. We establish an upper bound on β. This upper bound is independent of any prior distribution of n; it is purely a function of the expected estimation overhead and, thus, can be used to predict overhead requirements. Furthermore, if a prior maximum value for n is known, an even tighter upper bound can be established. The second property is that $[n_l(\alpha), n_u(\alpha)]$ is centered on the method of moments estimator, an estimator that we will argue is a reasonable point estimator for n.

4.1 Derivation of the interval estimator

We derive the interval estimator for n in the $\text{Binomial}(n, p)$ distribution in a manner similar to Lloyd's derivation [1984] of an estimator for the parameter p (see pp. 169-170). The estimator is, at the $100(1-\alpha)\%$ confidence level:

$$\hat{n} = \frac{\bar{r}}{p} \pm \frac{z_{\alpha/2}}{2kp}\left((1-p)z_{\alpha/2} + \sqrt{(1-p)4k\bar{r} + (1-p)^2 z_{\alpha/2}^2}\right)$$

where $z_{\alpha/2}$ is the value for which the CDF of the standard normal distribution $\Phi(z_{\alpha/2}) = (1-\alpha/2)$. In other words, at the $100(1-\alpha)\%$ confidence level, $n_l(\alpha) \le n \le n_u(\alpha)$ where

$$n_l(\alpha) = \frac{\bar{r}}{p} - \frac{z_{\alpha/2}}{2kp}\left((1-p)z_{\alpha/2} + \sqrt{(1-p)4k\bar{r} + (1-p)^2 z_{\alpha/2}^2}\right)$$

$$n_u(\alpha) = \frac{\bar{r}}{p} + \frac{z_{\alpha/2}}{2kp}\left((1-p)z_{\alpha/2} + \sqrt{(1-p)4k\bar{r} + (1-p)^2 z_{\alpha/2}^2}\right)$$

This interval estimator is derived from the interval estimator for the mean of a normal distribution with one degree of freedom:

$$\mu = \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{k}}$$

The approximation is justified when n is large and np is not extremely close to either 0 or n. (A rule of thumb given in [Arnold 1990] pp. 143-144 is that the approximation is valid so long as $5 \le np \le n - 5$.) We approximate the number of replies $r_1, r_2, \ldots, r_k$ in each observation as being drawn from IID distributions $R_i \sim \text{Normal}(np, np(1-p))$ so

$$np \doteq \bar{r} \pm z_{\alpha/2} \frac{\sqrt{np(1-p)}}{\sqrt{k}}$$

$$(np - \bar{r})^2 \doteq z_{\alpha/2}^2 \frac{np(1-p)}{k}$$

This is a quadratic for which the solution for n is:

$$\hat{n} = \frac{\bar{r}}{p} + \frac{(1-p)z_{\alpha/2}^2}{2kp} \pm \frac{z_{\alpha/2}}{2kp}\sqrt{(1-p)4k\bar{r} + (1-p)^2 z_{\alpha/2}^2}$$

For reasons that we elaborate on below, we would like the center point of the estimator to be $\hat{n} = \bar{r}/p$. However we observe that the center point of the solution to the quadratic is somewhat higher than this value. The amount is small so long as

$$\bar{r} >> \frac{(1-p)z_{\alpha/2}^2}{2k}$$

which will tend to be the case; for typical values of $z_{\alpha/2}$ ($z_{0.05/2} \doteq 1.96$ or $z_{0.01/2} \doteq 2.58$), $(1-p)z_{\alpha/2}^2 / 2k$ will not exceed 3.33. To center the estimator, we add the second term of the solution to the quadratic into the length of the interval. This has the effect of, lowering $n_l(\alpha)$ by this (typically small) amount.

4.2 Upper bound on β

We define β to be the length of the interval relative to the magnitude of its center, i.e.

$$\beta = 2\frac{n_u(\alpha) - n_l(\alpha)}{n_l(\alpha) + n_u(\alpha)}$$
$$= \frac{z_{\alpha/2}}{2k\bar{r}}\left((1-p)z_{\alpha/2} + \sqrt{(1-p)4k\bar{r} + (1-p)^2 z_{\alpha/2}^2}\right)$$

At the $100(1-\alpha)\%$ confidence level, $\hat{n} = \bar{r}/p \pm 100\beta\%$.

We are interested in the behavior of β as a function of overhead $h = k\bar{r}$ when estimation probabilities are low, as they would typically be in situations in which probabilistic polling is used.

$$\lim_{p\to 0}\beta_\alpha = \frac{z_{\alpha/2}}{2h}\left(z_{\alpha/2} + \sqrt{4h + z_{\alpha/2}^2}\right)$$

This is simplified further for high polling overhead:

$$\lim_{p\to 0}\left(\beta_\alpha \middle| h >> z_{\alpha/2}\right) = \frac{z_{\alpha/2}}{\sqrt{h}}$$

If a maximum value $\hat{n}_0$ can be established for n, and thus a safe minimum $p = h/k\hat{n}_0$, tighter bounds can be drawn as a function of both h and k:

$$\beta_\alpha \le \frac{z_{\alpha/2}}{2h}\left(\left(1-\frac{h}{k\hat{n}_0}\right)z_{\alpha/2} + \sqrt{\left(1-\frac{h}{k\hat{n}_0}\right)4h + \left(1-\frac{h}{k\hat{n}_0}\right)^2 z_{\alpha/2}^2}\right)$$

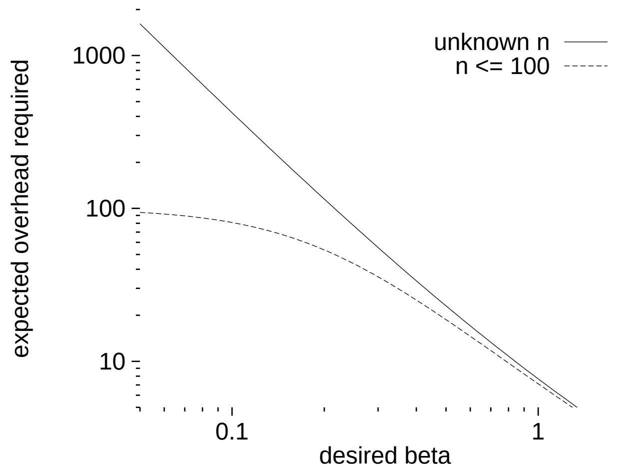
Any of these limiting equations may be solved for h to discover an upper bound on the minimum average overhead required to establish a given β at the confidence level specified by α. The first of the three, for an unknown prior distribution of n, yields

$$h \ge \frac{(1+\beta)z_{\alpha/2}^2}{\beta^2}$$

If there is a known upper bound, $\hat{n}_0$, on n then the required overhead is lower:

$$h \ge \frac{k\hat{n}_0(1+\beta)z_{\alpha/2}^2}{k\hat{n}_0\beta^2 + z_{\alpha/2}^2 + \beta z_{\alpha/2}^2}$$

The following graph displays the minimum expected overhead required to obtain a range of values of β, from ±5% to beyond ±100%, at the 95% confidence level. Two curves are plotted: one for unknown prior distribution of n, the other for when it is known a priori that $n \le 100$ (i.e. $\hat{n}_0 = 100$):



4.3 Center of the interval estimator's interval

A good property for an interval estimator is for it to be centered upon a reasonable point estimator. We will show that this interval estimator is centered upon the method of moments estimator (MME), and we will argue that the MME is a reasonable point estimator.

The MME is found by setting the first sample moment equal to the first population moment (see, e.g., [Casella and Berger 1990] p. 285):

$$m_1 = \mu_1$$
$$\frac{1}{k}\sum_{i=1}^{k} r_i = E(R)$$
$$\bar{r} = np$$
$$\hat{n} = \frac{\bar{r}}{p}$$

We argue that the MME is a reasonable estimator because it possesses two good qualities: it is unbiased, and it is superefficient, as shown in [Blumenthal and Dahiya 1981]. An unbiased estimator is one for which the expected value is equal to the parameter being estimated (see, e.g., [Mood 1950] p.

149). A superefficient estimator is one for which it is certain that the estimate is equal to the parameter being estimated in the limit as the number of the observations increases (see, e.g. [Blumenthal and Dahiya 1981] p. 905).

There is one special case to consider: occasionally, an unusually large number of replies r_i might be observed during the i^{th} polling round. By unusually large we mean that it is larger than the MME for n, an event that will be rare when p is small, as it typically is in situations that lend themselves to probabilistic polling. Since such an observation r_i is necessarily a lower bound on n, $\hat{n} = \max_i r_i$ is a better point estimator than the MME in this (rare) case. We do not know what interval to associate with such a point estimate.

5 Bounding overhead

The preceding section illustrated how to determine the expected number of packets required to obtain a desired β for an estimate of n. However, if a polling mechanism is designed to return an expected number of packets $E(R)$, more than $E(R)$ packets may well be received. If we can place a upper bound on this number of packets, then we can establish the maximum overhead that an estimate for a desired β will incur.

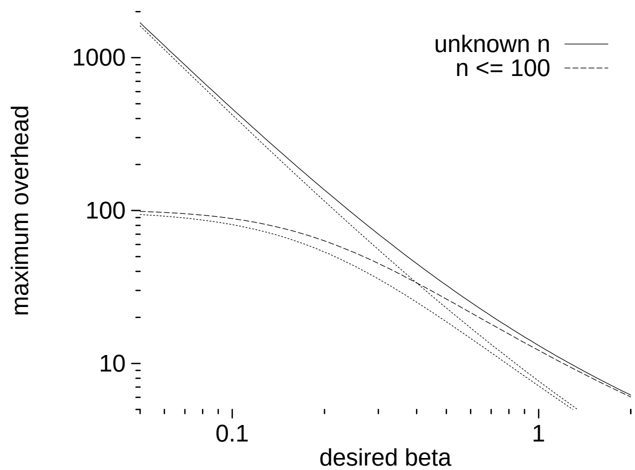
As we did when deriving the interval estimator, we approximate the distribution of the number of replies to be normal: $R \sim \text{Normal}(np, np(1-p))$. For any p, the maximum number of replies that can be expected, at the $100(1-\alpha)\%$ confidence level, is

$$\begin{aligned} r_{\max} &= E(R) + z_\alpha \sigma \\ &= np + z_\alpha \sqrt{np(1-p)} \end{aligned}$$

where z_α is the value for which the CDF of the standard normal distribution $\Phi(z_\alpha) = (1-\alpha)$. Substituting $h = p/kn$ for the expected overhead, we obtain the following expression for $r_{\max}$:

$$r_{\max} = \frac{h}{k} + z_\alpha \sqrt{\frac{h}{k}\left(1 - \frac{h}{nk}\right)}$$

The following graph displays the maximum overhead required to obtain the same range of values of β as shown in Section 4.2, again at the 95% confidence level. As in that graph, two curves are plotted: one for unknown prior distribution of n, the other for when it is known a priori that $n \le 100$ (i.e. $\hat{n}_0 = 100$). In addition, the two curves from the prior graph are traced for comparison:



If the overhead bound is a rigid one, one might wish to use a confidence level higher than 95% in order to establish the maximum overhead.

It is convenient to have an expression for the maximum probability p that, given an upper bound on n, will not cause the number of replies to exceed $r_{\max}$. The equation for $r_{\max}$ is quadratic in p, the solution of which yields

$$p \le \frac{2nr_{\max} + nz_\alpha^2 - z_\alpha \sqrt{n}\sqrt{4nr_{\max} - 4r_{\max}^2 + nz_\alpha^2}}{2n\left(n + z_\alpha^2\right)}$$

at the $100(1-\alpha)\%$ confidence level. The smaller of the two solutions to the quadratic is used, as the larger one yields a p that would cause the number of replies to exceed $r_{\max}$.

6 Special interval estimator for observed zeroes

As we have seen in the example of Section 3, the BTW mechanism is designed to produce zero replies in early polling rounds, so as to ensure minimum feedback. In this section we derive the special interval estimator for the case when all observations are zeroes (and p is known and allowed to vary). At the $100(1-\alpha)\%$ confidence level:

$$\hat{n}_0 = \left\lceil \frac{\ln\alpha}{\sum_{i=1}^{k}\ln(1-p_i)} \right\rceil$$

where $\hat{n}_0$ defines the upper bound of the confidence interval $[0,\hat{n}_0]$. This upper bound is the lowest value of n for which the probability of all observations being zero would be at most α, i.e. $\hat{n}_0 = \min_n\left(n\colon \Pr\{r_1 = r_2 = \ldots = r_k = 0; n, \boldsymbol{p}\} \le \alpha\right)$.

The probability of all observations being zero, given some n, is

$$\Pr\{r_i = 0; n, p_i\} = \frac{n!}{(n-0)!0!} p_i^0 (1-p_i)^{n-0}$$
$$= (1-p_i)^n$$
$$\Pr\{r_1 = r_2 = \ldots = r_k = 0; n, \boldsymbol{p}\} = \prod_{i=1}^{k}(1-p_i)^n$$

We can find a minimum n because this probability is strictly increasing in n. This minimum n is the smallest integer that exceeds the non-integer n for which $\Pr\{r_1 = r_2 = \ldots = r_k = 0; n, \boldsymbol{p}\} = \alpha$:

$$\prod_{i=1}^{k}(1-p_i)^n = \alpha$$
$$n = \frac{\ln\alpha}{\sum_{i=1}^{k}\ln(1-p_i)}$$

7 Application to the BTW mechanism

The example given in Section 3 is essentially a tuned version of the BTW mechanism. We have seen how, with the tools developed in Sections 4, 5, and 6, the following parameters may be freely adjusted to match application requirements:

- β: The size of the interval relative to the magnitude of its center, at a desired confidence level.
- h: The overhead, in number of replies, used for estimation.
- α: The level of confidence that no more than a given number of replies will be received.

We recommend that our interval estimator be used in place of what appears to be the estimator in [Bolot, Turletti, and Wakeman 1994] (equation (1) in that paper). That estimator has three properties our estimator improves upon: it is a point estimator, it depends upon a priori knowledge of n, and it does not make full use of the number of replies in the final polling round (it only uses the fact that the number is non-zero).

8 Analysis of the NB mechanism

In Section 8.1, we apply the binomial model to the NB mechanism and derive an unbiased estimator for n. We show that the estimator in [Nonnenmacher 1998] is biased in a manner that depends upon actual group size. In Section 8.2, we suggest a provably correct revision to the NB mechanism: one that fixes the error that results when the assumption of uniform inter-host delays does not hold. In Section 8.3 we present the maximum likelihood estimator (MLE) for n under the NB mechanism.

8.1 Unbiased estimator for the NB mechanism

Recall, from Section 2.2, that, in each polling round i, each of the r_i replies received under the NB mechanism is accompanied by the timer setting $z_{i,j}$ that triggered receiver j's sending of a polling response. Without loss of generality, we order the receivers $j = 1,\dots,n$ from earliest timer setting to latest. The earliest timer setting, $z_{i,1}$, thus defines an interval $[z_{i,1}, z_{i,1} + c]$ in which all r_i of the unsuppressed replies fall (assuming uniform inter-host delays of magnitude c). The reply probability is the conditional probability that a timer falls within this interval, given that the earliest timer setting is $z_{i,1}$:

$$p_i = \frac{F_{Zi}\left(z_{i,1} + c\right) - F_{Zi}\left(z_{i,1}\right)}{1 - F_{Zi}\left(z_{i,1}\right)}$$

where F_{Zi} is the CDF of the timer distribution. This conditional probability applies to all receivers except for the one with the earliest timer.

We can use the conditional probability to estimate the session size under the binomial model, on the understanding that we use $r_i - 1$ replies to estimate the quantity $n - 1$. This results in the following modification of the interval estimator we derived in Section 4.1:

$$\hat{n} = \frac{r-1}{p} + 1 \pm \frac{z_{\alpha/2}}{2kp}\left((1-p)z_{\alpha/2} + \sqrt{(1-p)4k(r-1) + (1-p)^2 z_{\alpha/2}^2}\right)$$

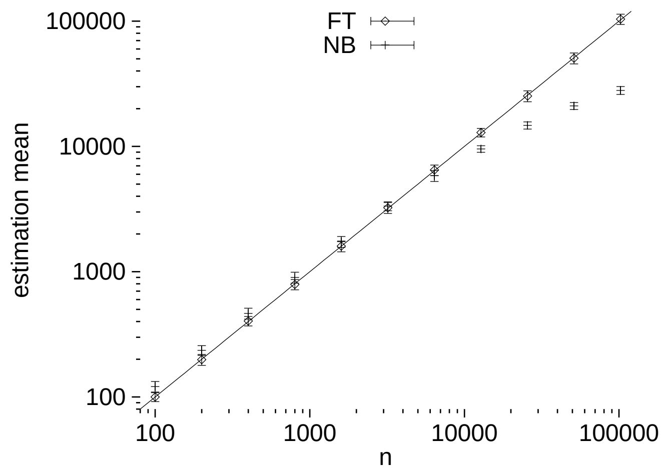
Note that, when applied to the NB mechanism, the estimator is only valid for a single observation because it relies upon a constant probability p across observations, and p varies across observations in the NB mechanism.

We compare the point estimator that is at the center of this interval estimator with the point estimator in [Nonnenmacher 1998]. We will label these estimators $\hat{n}_{\mathrm{FT}}$ and $\hat{n}_{\mathrm{NB}}$:

$$\hat{n}_{\mathrm{FT}} = \frac{r-1}{p} + 1$$

$$\hat{n}_{\mathrm{NB}} = \frac{r}{p}$$

Through simulations of session sizes between 100 and 100,000, we have found $\hat{n}_{\mathrm{FT}}$ to be unbiased, whereas $\hat{n}_{\mathrm{NB}}$ is biased in a manner that depends upon the session size, as seen in the diagram below.



The estimator $\hat{n}_{\mathrm{FT}}$ closely fits the curve $\hat{n} = n$ that has been plotted, whereas the estimator $\hat{n}_{\mathrm{NB}}$ is biased above n for low values of n, and then below n for higher values. For example, at $n = 100$, $\hat{n}_{\mathrm{FT}} = 99.99 \pm 8.17$ whereas $\hat{n}_{\mathrm{NB}} = 121.05 \pm 11.68$, and at $n = 819{,}200$, $\hat{n}_{\mathrm{FT}} = 823{,}479 \pm 82{,}345$, whereas $\hat{n}_{\mathrm{NB}} = 51{,}433 \pm 4{,}759$. Confidence intervals for this simulation are all at the 95% level or better. These simulations have been conducted using the particular timer distribution suggested in [Nonnenmacher and Biersack 1998] and [Nonnenmacher 1998], with the parameters $\lambda = 10.0$, $T = 10.0$, and $c = 1.0$ as described therein.

8.2 Revised NB mechanism

The NB mechanism relies upon one clearly unrealistic assumption that renders it unusable in a real-world internetwork environment. It assumes that the time required for a packet to travel from any host in the network to any other host is the same for all host pairs: a constant c. This assumption is central to the estimator in [Nonnenmacher 1998], which calculates the integral of the timer setting density function $f_Z(t)$ over the interval from the first timer expiry, $z_{i,1}$, to the time when all feedback is presumed to be suppressed, $z_{i,1} + c$. If inter-host delays, instead of being a constant c, are distributed above and below c, then some feedback suppression will take place before $z_{i,1} + c$, and some will take place after. To know the probable effects of this upon the estimator, one would have to know the prior distribution of inter-host delays, something that may be extremely difficult to determine.

We modify the NB mechanism so that the estimator uses the same number of replies that it would receive if inter-host delays were constant, even when they are not. This is achieved by transforming c into a sender-specified parameter, not dependent upon inter-host delays, and modifying the receiver and sender behaviors so as to assure the same number of replies under any delay distribution. While our modification always gives the correct number of replies for estimation purposes, we will show through simulation, using delay assumptions from [Nonnenmacher and Biersack 1998], that the unmodified NB mechanism gives the incorrect number.

Under our modified mechanism, the sender's request for feedback includes the parameter c along with the other parameters that receivers use to set their timer distributions. A receiver j monitors other receivers' feedback messages and keeps track of the minimum timer setting $z_{\min}$ from amongst those messages and its own timer setting $z_{i,j}$. Instead of suppressing its reply upon receipt of the first feedback message, it only suppresses its reply if, at any time, $z_{\min}$ drops to the point where $z_{i,j} > z_{\min} + c$. If its feedback is not suppressed by time $z_{i,j}$, receiver j multicasts a reply. We summarize this receiver algorithm as:

```
request(Z,c) arrives from sender
obtain timer z_{i,j} from random distribution Z
assign z_min = z_{i,j}
while time t < z_{i,j}
   for each reply(z_{i,x}) that arrives
      if z_{i,x} < z_min then
         assign z_min = z_{i,x}
         if z_{i,j} > z_min + c then exit (reply suppressed)
multicast reply(z_{i,j}) and exit (reply sent)
```

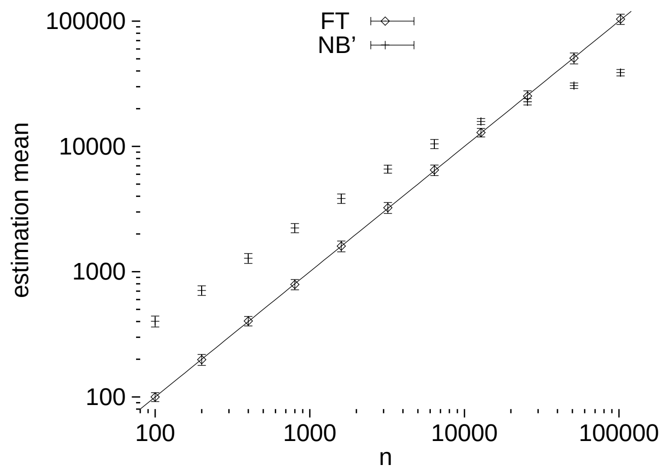
This procedure guarantees that all timers that are set in the interval $[z_{i,1}, z_{i,1} + c]$ will be counted, where $z_{i,1}$ is the lowest timer setting amongst all the receivers in polling round i. To prove this guarantee we suppose j to be a receiver with a timer setting in $[z_{i,1}, z_{i,1} + c]$. If receiver j receives the reply containing

$z_{i,1}$ before sending its own reply, then its $z_{\min} = z_{i,1}$ and the condition $z_{i,j} > z_{\min} + c$ will never obtain, and so its reply will not be suppressed. If receiver j does not receive the reply containing $z_{i,1}$ before sending its own reply, then its $z_{\min} \geq z_{i,1}$, because it is set from a timer value necessarily greater than or equal to $z_{i,1}$, and, again $z_{i,j} > z_{\min} + c$ will never obtain, and its reply will not be suppressed. We assume no loss, and bounded end-to-end delays. Therefore, its reply will be counted.

However, it is possible that a reply will be sent from a receiver j for which the timer setting $z_{i,j}$ does not lie in the interval $[z_{i,1}, z_{i,1} + c]$. Because $z_{i,1}$ is the minimum timer setting this can only occur if, at the time of sending its reply, no reply with timer setting $z_{i,x}$ such that $z_{i,j} > z_{i,x} + c$ has yet arrived at receiver j. The job of filtering out these extra replies before making an estimate falls to the sender. Once it has established the minimum timer setting for round $i, z_{i,1}$, it only counts those replies that fall within the interval $[z_{i,1}, z_{i,1} + c]$. This procedure guarantees that no receiver with a timer value that falls outside of the interval $[z_{i,1}, z_{i,1} + c]$ will be counted.

The two guarantees (that every timer that expires in $[z_{i,1}, z_{i,1} + c]$ will be counted, and that no timer that expires outside $[z_{i,1}, z_{i,1} + c]$ will be counted) combine to guarantee that exactly those timers that expire in $[z_{i,1}, z_{i,1} + c]$ will be counted, regardless of the distribution of inter-host delays. Since the estimator depends upon a correct observation of the number of timers that expire in $[z_{i,1}, z_{i,1} + c]$, our method guarantees that inter-host delays will not affect the correctness of the estimator.

We examine the effects of the modification through simulation. To do so we compare estimators that are both corrected for the bias described in Section 8.1. We label the estimator for the unmodified NB mechanism $\hat{n}_{\mathrm{NB}'}$, and the estimator for the modified mechanism $\hat{n}_{\mathrm{FT}}$. We simulate session sizes ranging between 100 and 100,000, with inter-host delays distributed as $\mathrm{Beta}(2.0, 2.0)$ over the interval $[0, 2c]$, as described in [Nonnenmacher and Biersack 1998], and the other parameters $\lambda = 10.0$, $T = 10.0$, and $c = 1.0$ as described therein:



We find $\hat{n}_{\mathrm{FT}}$ to be unbiased in the presence of these heterogeneous inter-host delays. (By construction, it behaves identically to the situation of homogeneous delays.) However, $\hat{n}_{\mathrm{NB'}}$ is biased: above n for low values of n, and below n for high values of n. For example, at $n = 100$, $\hat{n}_{\mathrm{NB'}} = 402.00 \pm 40.20$, and at $n = 819{,}200$, $\hat{n}_{\mathrm{NB'}} = 63{,}926 \pm 4{,}817$.

Because the modified NB mechanism lessens feedback suppression, we also analyzed the reply overhead incurred by the modified mechanism. It is of the same order of magnitude. There may be means to lower the additional overhead required to obtain a correct estimate. The closer c is to the highest inter-host delay, the smaller the proportion of replies that the sender will need to discard. However, changing c has other effects on the number of replies, and the way in which other polling parameters might adjusted in conjunction with c deserves further study.

8.3 MLE for the NB mechanism

Recall, from the beginning of this paper, that the binomial estimation model did not entirely match the NB mechanism because more is observed than simply a number of polling replies. Along with each of those replies is a number z_j that indicates the timer setting at which the receiver j's reply timer expired. Those settings were independently drawn from identically distributed random variables Z_j. Perhaps the values z_j could be used to provide a more refined estimate than would be possible under the binomial model? Unfortunately we find that the values z_j do not contribute to the MLE, except for the smallest z_j, which is already accounted for in the binomial model by the reply probability, which it determines. Here we derive the maximum likelihood estimator (MLE) for the mechanism proposed by Nonnenmacher and Biersack [1988].

Label the receivers $j = 1,\ldots,n$. Associated with receiver j is a sequence of IID non-negative random variables $\{Z_{i,j}\}_{i=1}^{\infty}$. These sequences are assumed to be mutually independent and $Z_{i,j} =_d Z, \ (1 \le j \le n; i = 1,\ldots)$ where Z has probability density function $f_Z(z)$ and cumulative distribution $F_Z(z) = P(Z \le z)$. Let $Z_{i,[j]}$ denote the j – th smallest value of $\{Z_{i,j}\}$.

The i –th experiment is performed as follows. Receiver j returns a probe to the sender if $Z_{i,j} < Z_{i,[1]} + c$ for some value $c > 0$, $i = 1,\ldots$ Let the results of the i – th experiment be an ordered vector $(z_{i,1}, z_{i,2},\ldots,z_{i,n})$, $z_{i,1} \le z_{i,2} \le \cdots \le z_{i,n}$ where $z_{i,j}$ is an instantiation of the $Z_{i,[j]}$. Define $z_i = (z_{i,1}, z_{i,2},\ldots,z_{i,r_i})$ where r_i denotes the number of receivers such that $z_{i,j} - z_{i,1} < c$.

The MLE for n based on k of the above experiments can be derived in a manner similar to that used to derive the MLE for n for the Binomial(n, p)distribution (see [Feldman and Fox 1968], reproduced in [Casella and Berger 1990] pp. 292-3). Let $z = z_1, z_2,\ldots,z_k$ denote the observations. The likelihood function, $L(n|z)$ for the parameter n after k experiments is

$$L(n|z) = \prod_{i=1}^{k} \binom{n}{r_i} r_i! \prod_{j=1}^{r_i} f_Z(z_{i,j})\left(1 - F_Z(z_{i,1} + c)\right)^{n-r_i}$$
$$= \prod_{i=1}^{k} \frac{n!}{(n-r_i)!} \prod_{j=1}^{r_i} f_Z(z_{i,j})\left(1 - F_Z(z_{i,1} + c)\right)^{n-r_i}$$

The MLE $\hat{n}$ is obtained by recognizing that $L(\hat{n}|z)$ is a unique maximum for $n = \hat{n}$ and so for any value of n $L(\hat{n}|z)/L(n|z) \ge 1$. This ratio is easily simplifiable at the value $\hat{n} - 1$:

$$\frac{L(\hat{n}|z)}{L(\hat{n}-1|z)}=\frac{\prod_{i=1}^{k}\hat{n}\left(1-F_Z(z_{i,1}+c)\right)}{\prod_{i=1}^{k}(\hat{n}-r_i)}$$

The value of n at which this ratio is equal to 1 is found by solving

$$\prod_{i=1}^{k}\left(1-F_Z(z_{i,1}+c)\right)=\prod_{i=1}^{k}\left(1-\frac{r_i}{n}\right)$$

Although this is a k-degree polynomial it can be solved by simple search because the right hand side of the equation is strictly decreasing in n. This search is bounded by the condition that $\hat{n}$ not be less than any of the observations r_i. Rather than search on the open-ended interval $[\max_i r_i,\infty)$ one can perform a binary search for $1/n$ on the closed interval $[0,1/\max_i r_i]$. The n so found is such that $\hat{n}$ lies in the interval $[n-1,n]$ and since $\hat{n}$ must be integer valued it is simply the floor of n. Note that the MLE for one observation turns out to be the heuristic estimator $\hat{n}=r/F_Z(z_{1,1})$. Last, observe that the MLE depends on $\{z_{i,1}\}$ but not on $\{z_{i,j}: j=2,\ldots,r_i\}$, $i=1,\ldots,k$.

We now consider a different approach to estimating the population size n. We instead derive the MLE for $m=n-1$ conditioned on the observed minimum timer values $z_{1,1},z_{2,1},\ldots,z_{k,1}$. As before, let $z=(z_1,z_2,\ldots,z_k)$. The likelihood function, $L_{z_{1,1},z_{2,1},\ldots,z_{k,1}}(m|z)$ for the parameter m after k experiments, where the minimum timer values are $z_{1,1},z_{2,1},\ldots,z_{k,1}$, is

$$\begin{aligned}L_{z_{1,1},\ldots,z_{k,1}}(m|z)&=\prod_{i=1}^{k}\binom{m}{r_i-1}(r_i-1)!\prod_{j=2}^{r_i}f_{Z|Z>z_{i,1}}(z_{i,j})\left(1-F_{Z|Z>z_{i,1}}(z_{i,1}+c)\right)^{m-r_i+1}\\&=\prod_{i=1}^{k}\frac{m!}{(m-r_i+1)!}\prod_{j=2}^{r_i}f_{Z|Z>z_{i,1}}(z_{i,j})\left(1-F_{Z|Z>z_{i,1}}(z_{i,1}+c)\right)^{m-r_i+1}\end{aligned}$$

Using similar arguments as above, the MLE $\hat{m}$ is obtained by solving

$$\prod_{i=1}^{k}\left(1-F_{Z|Z>z_{i,1}}(z_{i,1}+c)\right)=\prod_{i=1}^{k}\left(1-\frac{r_i-1}{m}\right)$$

Although this is a k-degree polynomial it can be solved by simple search because the right hand side of the equation is strictly decreasing in m. This search is bounded by the condition that $\hat{m}$ not be less than any of the observations $r_i - 1$. Rather than search on the open-ended interval $\left[\max_i r_i - 1, \infty\right)$ one can perform a binary search for $1/m$ on the closed interval $\left[0, 1/(\max_i r_i - 1)\right]$. The m so found is such that $\hat{m}$ lies in the interval $\left[m-1, m\right]$ and since $\hat{m}$ must be integer valued it is simply the floor of m. Note that this produces the estimate $\hat{n} = \hat{m} + 1$. For the case of one observation this results in the estimate $\hat{n} = (r_1 - 1)/(F_Z(z_{1,1} + c) - F_Z(z_{1,1})) + 1$. Last, observe that the MLE $\hat{m}$ depends on $\{z_{i,1}\}$ but not on $\{z_{i,j} : j = 2, \ldots, r_i\}$, $i = 1, \ldots, k$.

9 Related work

We have analyzed the problem of group size estimation through probabilistic polling on an end-to-end basis in a multicast-capable internetwork, focussing on mechanisms described in [Bolot, Turletti, and Wakeman 1994], [Nonnenmacher and Biersack 1998], and [Nonnenmacher 1998]. Another end-to-end mechanism for group size estimation is direct counting, as used in RTP [Schulzrinne et al 1996], [Rosenberg and Schulzrinne 1998], which is necessary when all participants must be heard from.

Our analysis has drawn from the statistics literature on estimation of the parameter n in a $\text{Binomial}(n, p)$ distribution. The MLE is given in [Feldman and Fox 1968], and we adapt it for the NB mechanism. Much of the literature focuses on point estimation and the problem of simultaneous estimation of n and p, and so is not directly applicable to interval estimation or probabilistic polling in which p is known. A survey of results in this area is found in [Olkin, Petkau, and Zidek 1981].

Acknowledgments

This research was funded in part by DARPA contract N66001-97-C-8513. We thank Michael Sutherland, John Buonaccorsi, and Lisa Ditto, of the UMass Amherst Statistical Consulting Center, and Roger Pinkham, of the Stevens Tech Mathematics Department for their help with the statistics literature and clarification of statistical issues. However, they bear no responsibility whatsoever for any shortcomings that might appear in the statistical analysis in this paper. We thank our colleagues Supratik Bhattacharyya, Sneha Kasera, Dan Rubenstein, and Maya Yajnik for their thoughts and advice. We thank Henning Schulzrinne, of Columbia University, for initially encouraging the first author to work on the multicast session membership size problem, and our colleague Jim Kurose for continuing that encouragement. Finally, we owe a great debt to Jörg Nonnenmacher for being so forthcoming in providing us with the latest versions of his work.

References

[Arnold 1990] Steven F. Arnold. *Mathematical Statistics*. Englewood Cliffs, NJ: Prentice-Hall (1990).

[Blumenthal and Dahiya 1981] Saul Blumenthal and Ram C. Dahiya. "Estimating the Binomial Parameter *n*". *Journal of the American Statistical Association*, vol. 76, no. 376 (December 1981), pp. 903-9.

[Bolot, Turletti, and Wakeman 1994] Jean-Chrysostome Bolot, Thierry Turletti, and Ian Wakeman. "Scalable Feedback Control for Multicast Video Distribution in the Internet". In *Proceedings of ACM SIGCOMM'94*. *Computer Communication Review*, volume 24, number 4 (October 1994) pages 58-67.

[Casella and Berger 1990] George Casella and Roger L. Berger. *Statistical Inference*. Belmont, California: Duxbury Press (1990).

[Feldman and Fox 1968] Dorian Feldman and Martin Fox. "Estimation of the Parameter *n* in the Binomial Distribution". *Journal of the American Statistical Association*, vol. 63, pp.150-8 (March 1968).

[Floyd et al 1995] Sally Floyd, Van Jacobson, Steven McCanne, Ching-Gung Liu, and Lixia Zhang. "A Reliable Multicast Framework for Light-weight Sessions and Application Level Framing". In *Proceedings of ACM SIGCOMM'95*. *Computer Communication Review*, vol. 25, no. 4 (Oct.1995) pp. 342-56.

[Lloyd 1984] Emlyn Lloyd. “Chapter 4: Interval Estimation”. In *Volume VI: Statistics, Part A* (Emlyn Lloyd, ed.) of *Handbook of Applicable Mathematics* (Walter Ledermann, ed.), pp. 137-207. Chichester, United Kingdom: John Wiley & Sons (1984).

[Mood 1950] Alexander Mood. *Introduction to the Theory of Statistics*. New York: McGraw-Hill (1950).

[Nonnenmacher 1998] Jörg Nonnenmacher. Ph.D. dissertation. Sophia Antipolis, France: Institut Eurecom (1998).

[Nonnenmacher and Biersack 1998] Jörg Nonnenmacher and Ernst W. Biersack. “Optimal Multicast Feedback”. In *Proceedings of IEEE INFOCOM ’98*. Los Alamitos, CA: IEEE Computer Society Press (1998).

[Olkin, Petkau, and Zidek 1981] Ingram Olkin, A. John Petkau, and James V. Zidek. “A Comparison of *n* Estimators for the Binomial Distribution”. *Journal of the American Statistical Association*, volume 76, number 375 (September 1981), pages 637-642.

[Rosenberg and Schulzrinne 1998] Jonathan Rosenberg and Henning Schulzrinne. “Timer Reconsideration for Enhanced Scalability”. In *Proceedings of IEEE INFOCOM ’98*. Los Alamitos, California: IEEE Computer Society Press (1998).

[Rubenstein, Kurose, and Towsley 1998] Dan Rubenstein, Jim Kurose, and Don Towsley. “Real-Time Reliable Multicast Using Proactive Forward Error Correction”. In *Proceedings of NOSSDAV’98*. Berlin: Springer-Verlag (1998).

[Schulzrinne et al 1996] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson. RFC 1899: “RTP: A Transport Protocol for Real-Time Applications”. Internet Engineering Task Force Request For Comments (January 1996).