An Urn Model with Applications to Database Performance Evaluation

Rahul Simha

Department of Computer Science
College of William and Mary
Williamsburg, VA 23185
simha@cs.wm.edu

Amitava Majumdar

Sunrise Test Systems Inc.
47211 Lakeview Blvd.
Fremont, CA 94538-6530
amit@srtest.viewlogic.com

Statement of scope and purpose: Urn models are commonly used to motivate and formulate problems in discrete probability. Typically, an urn contains some number of colored balls from which the balls are drawn according to rules particular to an application. The problem is usually to characterize the probability distribution of the balls drawn or those remaining in the urn. Numerous textbooks on the subject point to the wide applicability of urn models. This paper presents an urn model that generalizes several previous urn models and shows how the extensions can be applied to certain performance problems related to distributed databases.

Abstract: We consider two problems in distributed databases that have identical probabilistic structure, both of which have received significant attention in the literature. One is the problem of characterizing the number of distinct sites accessed by transactions in a distributed database and the other is the problem of determining the number of block accesses in a relation. We focus in particular on obtaining the distribution of this number when accesses are generated randomly. Previously published research has derived the mean number of sites or blocks accessed under some assumptions about the system parameters. The results presented in this paper generalize this work in several ways. First, we weaken the standard uniformity assumption to allow for a transaction accessing a random number of distinct sites or blocks and also consider a non-uniform access pattern in which one site or block (a 'hotspot') is accessed more frequently than others. Second, we compute not only the mean and variance but also the entire distribution of the number of sites (blocks) accessed, a measure found useful in the performance analysis of distributed databases. Additional generalizations are discussed in the paper.

1 Introduction

We consider two problems of identical probabilistic structure in the area of databases. One is the problem of determining the number of distinct sites accessed by transactions in a distributed database [18] – the *site access* problem. The other is the problem of determining the number of distinct blocks of data accessed by the average transaction [1] – the *block access* problem. Both problems have received significant attention in the literature [1, 4, 2, 3, 7, 10, 11, 14, 16, 17, 18, 19, 20] and both can be formulated using the same urn model (described below).

In the site access problem, a typical database transaction (via the operations that compose the transaction) generates accesses to data at different sites in a distributed database. Before the transaction can successfully complete, all sites must 'commit' [13] and thus, the time to complete the commit operation for the whole transaction depends on the number of *distinct* sites accessed – in general, the more the number of sites accessed, the longer the commit time. Since the actual time for each site to commit is usually modeled as a random variable, the overall time to commit is then a random number (the number of sites) of random variables, the statistics of which are known to be useful in performance evaluation [18]. In the block access problem, a database relation is horizontally partitioned into blocks, which are then used as units in storage management [13]. Consequently, if a relation is accessed a certain number of times, the true cost of access is the number of *distinct* blocks accessed. Knowledge about the statistics of block accesses can then be used to influence storage management policies. Since both problems have the same probabilistic structure, we will focus on only one of them, the site access problem, with the understanding that most of the distributions derived in the paper also apply to the other problem.

The site access problem we examine is simply stated as the following question: how many distinct sites are accessed by a transaction, when each of the transaction's t operations accesses a random number of different sites among a total of n_s sites? If random variable X_t denotes the number of sites accessed, then the mean and variance of X_t are important performance measures [18]. The distribution of X_t is also useful, as outlined in [18](Section III, pp. 100): the time the commit coordinator waits to obtain responses from the different sites can be expressed as a maximum of X_t random variables. That is, if R_i is a random variable denoting the time the coordinator waits for one site to send its response to a precommit message, then $R(t) = \max\{R_1, \ldots, R_{X_t}\}$ is the time waited before the coordinator commits or aborts the transaction. The distribution of X_t is needed in order to compute $E[R(t)]$. For example, it is known that when the R_i's are independent and identically distributed exponential random variables with parameter μ then $E[R(t)|X_t = n] = \frac{H_n}{\mu}$ where H_n is the n-th harmonic number [6]. Then, knowing the distribution of X_t, we can compute the mean response time as $E[R(t)] = \sum_n \frac{H_n}{\mu} P[X_t = n]$.

The above cited papers have used an expression (derived in [1] and expounded in [18, 20]) for the expected value of X_t, the number of sites accessed. While the distribution of X_t has been computed for a restricted case [7], we show how to compute both the distribution as well as the mean and variance of X_t for several useful generalizations. Our contributions may be summarized as follows: We generalize the transaction model so that each operation accesses multiple *distinct* sites, a generalization useful when the transaction's operations are known apriori. Distributions are then computed *conditioned* on the knowledge about the operations. Previous work has only studied single accesses to distinct sites. Additionally, we consider the case in which the number of distinct sites accessed is itself random. Our most important contribution is to incorporate the effect of a 'hotspot' (one site accessed more frequently than others). Finally, we provide general expressions when the number of operations is itself random. These expressions are useful when a transaction's operations depend in large part on the data; then, if the statistics of this number can be obtained, they can be used to accurately compute the distribution of X_t via compounding.

The research presented in this paper, to the best of our knowledge, also extends existing results in the general area of urn models [9]. In the language of urn models, our model can be described as follows. Suppose we are sampling with replacement from an urn initially containing n_s white balls. With each drawing a random number of them (black or white) are painted black and the entire sample is returned back to the urn. For a fixed (or random) number of drawings, what is the distribution of the number of black balls in the urn? Urn models are an important framework for the discussion of several discrete distributions of interest. Indeed, our solution technique, the method of *factorial moments*, is commonly used in characterizing several such discrete distributions [8].

A comment may be made about the general form of our expressions. We note that the expression for the mean [1] is intuitively useful because of its simple form. Although our generalizations are much more powerful, they surprisingly maintain this simplicity of form, which is both analytically tractable as well as intuitively useful. A more precise statement of this fact is provided in the next section.

As mentioned earlier, the expression for the mean number of sites accessed appears in several places in the literature, starting with [1], either in the context of the block access problem [2, 3, 4, 7, 10, 11, 14, 17, 19, 20] or the site access problem [18]. A comprehensive discussion of the general block access problem appears in [7], in which assumptions of independence in several versions [12, 15, 17] of the same problem are discussed. Additional related work on this problem can be found in [4, 10, 14, 17, 21, 2, 19]. Note that [10] examines the case in which a single transaction accesses a random number of distinct sites, one of the many results we generalize in this paper. A different type of analysis is found in [3], in which the functional form of the mean number accessed is studied using the theory of majorization. In this context, the recent work of [18] uses Jensen's inequality and results in convex stochastic orderings to compare expected values for fixed and random numbers of transactions.

In the next section we give a precise formulation of our problem. Section 3 contains the main results. Following that, we consider special cases in Section 4 and non-uniform access patterns in Section 5, before making concluding remarks in Section 6.

2 Problem Formulation

We follow the formulation adopted in [7], albeit with slightly different notation. A database is partitioned across n_s sites. We define the following terms (the reader is referred to [13] for general terminology in distributed databases).

- W_t – the random number of *distinct* sites accessed by operation t, $t \in \mathbb{N}$, the set of natural numbers. For example, $W_3 = 7$ implies that the third operation makes accesses to seven *different* sites. We refer to the seven sites as the 'group' of sites accessed by the transaction. We will assume that every group of seven sites (out of $\binom{n_s}{7}$ sites) is equally likely to be the group of sites accessed by this operation. While this uniformity assumption is weaker than the standard uniformity assumption, we also extend our results to the case in which one particular site is more frequently accessed than other sites. Also, the distribution of W_t can be arbitrary; in the literature it is most often assumed that $W_t = 1$ or $W_t = k$ (fixed-size).
- Z_t – for convenience we define $Z_t = n_s - W_t$. When $W_t, t \geq 1$ is an *i.i.d.* sequence, we will, for ease of notation, drop the subscripts from W_t and Z_t.
- X_t – the random number of sites accessed by the first t operations. It is the analysis of this rv that forms the core of our paper.
- Y_t – let $Y_t = n_s - X_t$.

When W has the degenerate distribution $W = 1$ (hence, $Z = n_s - 1$), the expected value of X_t is straightforward to compute:

$$E[X_t] = n_s \left(1 - \left(1 - \frac{1}{n_s}\right)^t\right) \tag{1}$$

To paraphrase the explanation in [7, 18], note that the probability of a transaction *not* accessing a particular site i is $\left(1 - \frac{1}{n_s}\right)$ and therefore, the probability this site is not accessed by t transactions is $\left(1 - \frac{1}{n_s}\right)^t$. Hence, the probability that it is accessed at least once is $\left(1 - \left(1 - \frac{1}{n_s}\right)^t\right)$. The result follows by summing expectations for each site (even though there is dependence). Next, observe that we may write (1) in terms of Y_t:

$$\begin{aligned} E[Y_t] &= E[n_s - X_t] \\ &= n_s \left(\frac{n_s - 1}{n_s}\right)^t \\ &= n_s \left(\frac{E[Z]}{n_s}\right)^t \end{aligned}$$

In this paper, we generalize the above expression for Y_t to

$$E\left[Y_t^{\underline{m}}\right] = n_s^{\underline{m}} \left(\frac{E\left[Z^{\underline{m}}\right]}{n_s^{\underline{m}}} \right)^t \tag{2}$$

where $x^{\underline{m}}$, for a non-negative integer x, is defined as $x(x-1)\dots(x-m+1)$. Observe that for a rv X, $E\left[X^{\underline{m}}\right] = E\left[X(X-1)\dots(X-m+1)\right]$ is the m-th descending factorial moment of X [8]. It is well-known [5] that the mean and variance, as well as the distribution of X, can be expressed in terms of its factorial moments. Note also that, while $W = 1$ is typically assumed [1, 3, 7, 18, 20], our generalization extends to arbitrary W. In this paper we show how to compute expressions for $E\left[X_t^{\underline{m}}\right]$ (or equivalently, $E\left[Y_t^{\underline{m}}\right]$, since they are directly related), a matter to which we next turn.

3 Distribution of the Number of Sites Accessed

We will now solve for $E\left[Y_t^{\underline{m}}\right]$ and use our solution to obtain expressions for the mean, variance and distribution of X_t, our rv of interest. The following lemma will prove useful. We do not lay any claim to originality for equation (4); it is possible to derive it from known combinatorial identities. However, we believe a direct proof is simpler and worth presenting here because the proof reveals ideas useful in later derivations.

LEMMA : *Let V_i, $i = 1, \dots, n$ be a collection of n indicator rv's and let $S = \sum_{i=1}^{n} V_i$. Next let A be the collection of all $(\alpha_1, \dots, \alpha_m)$, $1 \le \alpha_1 < \dots < \alpha_m \le n$, such that $V_{\alpha_1}^{j_1} V_{\alpha_2}^{j_2} \dots V_{\alpha_m}^{j_m}$, $j_i > 0$, is a term in the expansion of*

$$S^{\underline{m}} = \left(\sum_{i=1}^{n_s} V_i \right) \left(\sum_{i=1}^{n_s} V_i - 1 \right) \dots \left(\sum_{i=1}^{n_s} V_i - m + 1 \right). \tag{3}$$

Then,

$$E\left[S^{\underline{m}}\right] = \sum_{(\alpha_1, \dots, \alpha_m) \in A} E\left[V_{\alpha_1} V_{\alpha_2} \dots V_{\alpha_m}\right]. \tag{4}$$

PROOF : First, observe that the right hand side of (3) will contain terms of the type $\gamma E\left[V_1^{j_1} V_2^{j_2} \dots V_n^{j_n}\right]$ when the expectation is distributed over the terms in the product $(\sum_{i=1}^{n} V_i)^{\underline{m}}$. In the full expansion, we will obtain terms in which $\gamma \ne 1$ and $j_i \ne 0$ for only some i's. We will show that such terms cancel out in a proof by induction. Clearly, the result is true for $m = 1$. Assume it is true for m.

First note that, since the V_i's are indicator rv's, $E\left[V_{\beta_1}^{j_1} V_{\beta_2}^{j_2} \dots V_{\beta_k}^{j_k}\right] = E\left[V_{\beta_1} V_{\beta_2} \dots V_{\beta_k}\right]$ for $j_i \ge 1$, i.e. exponents of the V_i's are suppressed. When expectations are simplified in this manner in the expansion of (3), we will refer to the expansion as a full expansion.

Next, note that

$$\begin{aligned} E\left[S^{\underline{m+1}}\right] &= E\left[S^{\underline{m}}\left(\sum_{i=1}^{n} V_i - m\right)\right] \\ &= E\left[S^{\underline{m}}\left(V_1 + V_2 + \ldots + V_n - m\right)\right]. \end{aligned} \tag{5}$$

Thus, terms in the full expansion of $E\left[S^{\underline{m+1}}\right]$ are obtained by taking expectation of terms in $S^{\underline{m}}$ multiplied by the different V_i's in (5), as well as '$-m$'. Consider multiplication by V_1 (in general by V_i). By hypothesis, terms in the full expansion of $E\left[S^{\underline{m}}\right]$ consist of expected values of product terms of the following two types:

1. $E\left[V_{\beta_1} V_{\beta_2} \ldots V_{\beta_m}\right]$, where $V_{\beta_1} = V_1$ and
2. $E\left[V_{\beta_1} V_{\beta_2} \ldots V_{\beta_m}\right]$, where $V_{\beta_1} \neq V_1$

Suppose there are k_A terms of the first type in the full expansion of $S^{\underline{m}}$. Then, $E\left[S^{\underline{m+1}}\right]$ will have k_A terms of the type $E\left[V_1 V_{\beta_2} \ldots V_{\beta_m}\right]$ (after suppressing the exponent of V_1) with total contribution $k_A E\left[V_1 V_{\beta_2} \ldots V_{\beta_m}\right]$. Likewise, there is a similar contribution made by multiplication of these k_A terms in the full expansion of $E\left[S^{\underline{m}}\right]$, by the other $m-1$ V_{β_i}'s $(2 \leq i \leq m)$. Thus the total positive contribution of these terms in $E\left[S^{\underline{m+1}}\right]$ is

$$m k_A E\left[V_{\beta_1} V_{\beta_2} \ldots V_{\beta_m}\right] \tag{6}$$

for each distinct group of m V_i's. In essence, due to exponent suppression, V_1 and the V_{β_i}s are absorbed in the expected value terms of the first type. Now, the $k_A E\left[V_1 V_{\beta_2} \ldots V_{\beta_m}\right]$ in $E\left[S^{\underline{m}}\right]$ will multiply with negative m in (5) and cancel with the term in (6). Thus, the only positive contribution is made by terms of the second type, which will result in terms with $m+1$ distinct V_i's. Furthermore, since these terms are obtained by taking products of the $m+1$ terms in all possible distinct orderings we get our result. ∎

The above lemma will prove useful in arriving at the factorial moment of Y_t. Define the following indicator rv's:

$$U_{it} = \begin{cases} 1 & \text{if site } i \text{ is } \textit{not} \text{ accessed by transactions } 1, \ldots, t \\ 0 & \text{otherwise.} \end{cases}$$

Note that $Y_t = \sum_{i=1}^{n_s} U_{it}$. We will first condition on values for each W_i, $i = 1, \ldots, t$ and compute $E\left[Y_t^{\underline{m}} \middle| W_1 = k_1, \ldots, W_t = k_t\right]$. From the above lemma,

$$\begin{aligned} E\left[Y_t^{\underline{m}} \middle| W_1 = k_1, \ldots, W_t = k_t\right] &= \sum_{(\alpha_1, \ldots, \alpha_m) \in A} E\left[U_{\alpha_1 t} U_{\alpha_2 t} \ldots U_{\alpha_m t} | W_1 = k_1, \ldots, W_t = k_t\right] \\ &= \sum_{(\alpha_1, \ldots, \alpha_m) \in A} P\left[\text{all these } U_{\alpha_i t}\text{'s are equal to } 1 | W_1 = k_1, \ldots, W_t = k_t\right] \end{aligned}$$

$$
\begin{aligned}
&= |A| \prod_{i=1}^{t} P\left[U_1 = U_2 = \ldots = U_m = 1 | W_i = k_i\right] \\
&= n_s^{\underline{m}} \prod_{i=1}^{t} \frac{\begin{pmatrix} n_s - m \\ k_i \end{pmatrix}}{\begin{pmatrix} n_s \\ k_i \end{pmatrix}}.
\end{aligned}
$$

To understand the above derivation observe that the set A contains all m-sized subsets of sites and therefore, our uniformity assumption allows us to consider any one such subset (sites 1 through m) and multiply the expression by the size of A. The probability of leaving sites 1 through m untouched by the accesses is the number of ways of selecting accesses to other sites ($n_s - m$ of them) divided by the total number of ways of selecting the accesses. Finally, note that there are $n_s^{\underline{m}} = n_s(n_s-1)\ldots(n_s-m+1)$ such m-sized product terms in the expansion of equation (4). The resulting expression is easily simplified to[1]

$$
E\left[Y_t^{\underline{m}} \middle| W_1 = k_1, \ldots, W_t = k_t\right] = n_s^{\underline{m}} \prod_{i=1}^{t} \frac{(n_s - k_i)^{\underline{m}}}{n_s^{\underline{m}}}. \tag{7}
$$

While the above conditional expectation is itself useful (for example, when given information about particular operations), we will now remove the conditioning in order to complete our goal of obtaining $E\left[Y_t^{\underline{m}}\right]$. To achieve this, first note that $W_i = k_i$ implies $Z_i = n_s - k_i$. Then, in order to remove the conditioning we sum the above product expressions together with $P\left[W_i = k_i\right] = P\left[Z_i = n_s - k_i\right]$ over all possible k_i's (assuming the W_i's are independent):

$$
\begin{aligned}
E\left[Y_t^{\underline{m}}\right] &= \sum_{k_1,\ldots,k_t} n_s^{\underline{m}} \prod_{i=1}^{t} \frac{(n_s - k_i)^{\underline{m}}}{n_s^{\underline{m}}} P\left[W_1 = k_1\right] \ldots P\left[W_t = k_t\right] \\
&= \sum_{k_1,\ldots,k_t} n_s^{\underline{m}} \prod_{i=1}^{t} \frac{(n_s - k_i)^{\underline{m}}}{n_s^{\underline{m}}} P\left[Z_1 = n_s - k_1\right] \ldots P\left[Z_t = n_s - k_t\right] \\
&= n_s^{\underline{m}} \sum_{k_2,\ldots,k_t} \prod_{i=2}^{t} \frac{(n_s - k_i)^{\underline{m}}}{n_s^{\underline{m}}} P\left[Z_2 = n_s - k_1\right] \ldots P\left[Z_t = n_s - k_t\right] \sum_{k_1} \frac{(n_s - k_1)^{\underline{m}}}{n_s^{\underline{m}}} P\left[Z_1 = n_s - k_1\right] \\
&= n_s^{\underline{m}} \sum_{k_2,\ldots,k_t} \prod_{i=2}^{t} \frac{(n_s - k_i)^{\underline{m}}}{n_s^{\underline{m}}} P\left[Z_2 = n_s - k_1\right] \ldots P\left[Z_t = n_s - k_t\right] \frac{E\left[Z^{\underline{m}}\right]}{n_s^{\underline{m}}}
\end{aligned}
$$

Continuing this manipulation, we get our key result, equation (2):

$$
E\left[Y_t^{\underline{m}}\right] = n_s^{\underline{m}} \left(\frac{E\left[Z^{\underline{m}}\right]}{n_s^{\underline{m}}}\right)^t.
$$

[1] For $m > k_i$ multiply numerator and denominator of the combination term by $(n_s - k_i)(n_s - k_i - 1)\cdots(n_s - m + 1)$. For $m \le k_i$ cancel these terms.

The pmf's of Y_t and X_t, can be expressed in terms of their factorial moments using standard methods (see page 73 in [5]). As mentioned earlier, equation (2) can be used to compute the mean and variance as well as the generating function of Y_t. Due to the simplicity of expressions for statistics of Y_t (as compared with those for X_t), we use them more often in deriving further results. In particular, the following general expressions are useful:

$$\begin{aligned}
E[Y_t] &= n_s \left(\frac{E[Z]}{n_s}\right)^t \\
E[X_t] &= n_s \left(1 - \left(\frac{E[Z]}{n_s}\right)^t\right) \\
\mathrm{Var}[Y_t] &= n(n_s-1)\left(\frac{E[Z^{\underline{2}}]}{n_s(n_s-1)}\right)^t + n_s\left(\frac{E[Z]}{n_s}\right)^t - n_s^2\left(\frac{E[Z]}{n_s}\right)^{2t} \\
\mathrm{Var}[X_t] &= \mathrm{Var}[Y_t]
\end{aligned}$$

It is also easy to obtain (from equation (7)) expectations conditioned on $k_1, \ldots, k_t$. For example,

$$E[Y_t | W_1 = k_1, \ldots, W_t = k_t] = n_s \prod_{i=1}^{t} \left(1 - \frac{k_i}{n_s}\right).$$

Thus, we have achieved the first few goals described in the introduction: we have generalized earlier results (equation (1)) to the cases in which a transaction's operations access a fixed or random number of distinct sites and obtained conditioned as well as unconditioned (factorial) moments of the number of accessed sites. In the following sections, we consider specific distributions of interest and non-uniform access patterns. But first, we present an illustrative example that shows how equation (2) may be used in an application.

Example: Consider a transaction with $t = 5$ operations in a system with $n_s = 8$ remote sites. Suppose that each operation accesses one of the remote sites (hence, $W_t = 1$ and $Z_t = n_s - 1 = 7$). Next, suppose the time required for a commit is exponentially distributed with an average commit time of 10 milliseconds (Thus, $\mu = 0.1$). Now, if n distinct sites are accesssed and $R_1, \ldots, R_n$ are exponentially distributed rv's then the time to commit the transaction, as explained earlier, is $R(t) = \max(R_1, \ldots, R_n)$ with expectation

$$\begin{aligned}
E[R(t)] &= E[\max(R_1, \ldots, R_n)] \\
&= \frac{H_n}{\mu}
\end{aligned}$$

where $H_n = \sum_{i=1}^{n} \frac{1}{i}$ [6]. However, the number of sites accessed is itself random. Hence, we need to use the distribution of this number and condition on the number of sites:

$$\begin{aligned} E\left[R(t)\right] &= \sum_n E\left[E\left[R(t)\right] | X_t = n\right] P\left[X_t = n\right] \\ &= \sum_n \frac{H_n}{\mu} P\left[X_t = n\right]. \end{aligned} \tag{8}$$

Thus, what remains to be computed is the distribution of X_t. In our example, note that $X_t \in \{1, \ldots, 5\}$ and thus, we need to compute $P\left[X_t = n\right], n = 1, \ldots, 5$. To do so, we compute the distribution of Y_t and use the relation between X_t and Y_t. Substituting the parameters $t = 5, n_s = 8$ and $Z_t = 7$ in equation (2), we get:

$$E\left[Y_t^{\underline{m}}\right] = 8^{\underline{m}} \left(\frac{7^{\underline{m}}}{8^{\underline{m}}}\right)$$

for $m = 3, \ldots, 7$, the values of Y_t. Now recall that the z-transform of Y_t, $\sum_i z^i P\left[Y_t = i\right]$, can be successively differentiated at unity to give factorial moments. Thus, to obtain probabilities from factorial moments, it is easy to work backwards and obtain these probabilities recursively from the highest factorial moments. In our example, we obtain

$$\begin{aligned} P\left[Y_t = 7\right] &= \frac{1}{7!} E\left[Y_t^{\underline{7}}\right] \\ P\left[Y_t = 6\right] &= \frac{1}{6!} \left(E\left[Y_t^{\underline{6}}\right] - 7^{\underline{6}} P\left[Y_t = 7\right]\right) \\ &\vdots \\ P\left[Y_t = 3\right] &= \frac{1}{3!} \left(E\left[Y_t^{\underline{3}}\right] - \sum_{j=4}^{7} j^{\underline{3}} P\left[Y_t = j\right]\right). \end{aligned}$$

Next, we use $P\left[X_t = i\right] = P\left[Y_t = 8 - i\right]$ and compute the distribution of X_t: $P\left[X_t = 1\right] = 0.0002, P\left[X_t = 2\right] = 0.0256, P\left[X_t = 3\right] = 0.2563, P\left[X_t = 4\right] = 0.5127, P\left[X_t = 5\right] = 0.2051$. Substituting these values in equation (9), we obtain the mean commit time as $E\left[R(t)\right] = 20.45$ milliseconds.

4 Special Cases of Interest

We now focus on special instances of the random variable W. We consider the following three distributions for W. The first is the most immediate generalization of $W = 1$, namely $W = k$. The second is the uniform distribution over $\{1, \ldots, n_s\}$, which we write as $W \sim U\{1, \ldots, n_s\}$. This case is useful when no particular size dominates. Finally, we consider $W \sim \text{Mix}(k_1, k_2, \beta)$, where

W takes two values, k_1 and k_2 with probabilities β and $1-\beta$ respectively. This latter distribution allows analysis with a mix of operations — small ones, with k_1 accesses and large ones with k_2 accesses ($k_1 < k_2$). By varying β, we can control the transaction mix. Note that this two-valued distribution can easily be generalized to an arbitrary but finite discrete distribution.

For the first case, the degenerate distribution at $W = k$, note that $E[Z] = n_s - k$ and hence [10, 14],

$$E[X_t] = n_s \left(1 - \left(1 - \frac{k}{n_s}\right)^t\right).$$

This is a simple and straightforward generalization of (1). Next, when $W \sim U\{1, \ldots, n_s\}$, we have $E[Z] = \frac{n_s - 1}{2}$; this results in

$$E[X_t] = n_s \left(1 - \left(1 - \frac{n_s + 1}{2n_s}\right)^t\right).$$

Third, if $W \sim \mathrm{Mix}(k_1, k_2, \beta)$, it is easily seen that $E[Z] = \beta(n_s - k_1) + (1-\beta)(n_s - k_2)$, from which we get

$$E[X_t] = n_s \left(1 - \left(\frac{\beta(n_s - k_1) + (1-\beta)(n_s - k_2)}{n_s}\right)^t\right).$$

Next, we turn to the computation of rate information. For example, if transactions arrive to the system at a Poisson rate of λ, we can use the above formulas to obtain an expression for the effective rate at which sites are accessed. Equivalently, we may consider an interval of length and count the number of sites accessed during the interval. Rate information, we note, is essential for use in any queueing-theoretic analysis of performance.

For a given interval length d let T_d denote the random number of operations (from randomly arriving transactions) arriving in that interval. Suppose that T_d has a probability mass function (pmf) $f_{T_d}(t)$ and let the associated z-transform be denoted by $\phi_{T_d}(z)$. Then

$$\begin{aligned} E\left[(Y_{T_d})^{\underline{m}}\right] &= E_{T_d}\left[E\left[Y_t^{\underline{m}}\right] \mid T_d = t\right] \\ &= \sum_t E\left[Y_t^{\underline{m}}\right] P\left[T_d = t\right] \end{aligned}$$

Substituting from equation (2) we get

$$\begin{aligned} E\left[(Y_{T_d})^{\underline{m}}\right] &= \sum_t n_s^{\underline{m}} \left(\frac{E\left[Z^{\underline{m}}\right]}{n_s^{\underline{m}}}\right)^t P\left[T_d = t\right] \\ &= n_s^{\underline{m}} \phi_{T_d}\left(\frac{E\left[Z^{\underline{m}}\right]}{n_s^{\underline{m}}}\right). \end{aligned} \tag{9}$$

Equation (9) provides a characterization for the pmf of Y_{T_d}. A similar characterization can be derived for the pmf of X_{T_d}. In particular the expected value of X_{T_d} is given by

$$E\left[X_{T_d}\right] = n_s\left(1 - \phi_{T_d}\left(\frac{E[Z]}{n_s}\right)\right), \tag{10}$$

a simple and direct generalization of equation (1).

As an example, suppose $T_d \sim \text{Poisson}(\lambda d)$ implying that (since $\exp(\lambda(z-1))$ is the z-transform of a Poisson rv with rate λ),

$$E\left[X_{T_d}\right] = n_s\left(1 - \exp\left\{\lambda d\left[\frac{E[Z]}{n_s} - 1\right]\right\}\right) = n_s\left(1 - \exp\left\{\frac{-\lambda d E[W]}{n_s}\right\}\right).$$

We substitute expressions for $E[Z]$ for the three different cases under consideration, namely (1) the degenerate distribution at $W = k$, (2) $W \sim U\{1, \ldots, n_s\}$ and (3) $W \sim \text{Mix}(k_1, k_2, \beta)$. This yields

$$E\left[X_{T_d}\right] = \begin{cases} n_s\left(1 - \exp\left\{\frac{-\lambda d k}{n_s}\right\}\right) & \text{for degenerate } W = k \\ n_s\left(1 - \exp\left\{\frac{-\lambda d(1+n_s)}{2n_s}\right\}\right) & \text{for } W \sim U\{1, \ldots, n_s\} \\ n_s\left(1 - \exp\left\{\frac{-\lambda d(\beta k_1 + (1-\beta)k_2)}{n_s}\right\}\right) & \text{for } W \sim \text{Mix}(k_1, k_2, \beta). \end{cases}$$

Thus, our results in compounding allow for simple derivations of the compounded number of sites accessed by a random number of transactions.

Example: As an example, consider the case when transactions arrive at the rate of 1 per second (thus $\lambda = 1$ and $d = 1$) to a system with 50 sites ($n_s = 50$). Suppose the transaction mix is such that small operations access only one site ($k_1 = 1$). Then, if large operations access, say, 30 sites and 10 percent of the operations are small operations, the third expression above yields an average of 20.92 distinct sites accessed, or about 40 percent of the sites. In Figures 1 and 2, we plot the number of sites accessed when using the $\text{Mix}(k_1, k_2, \beta)$ distribution with $n_s = 50$ and $k_1 = 1$. Clearly, as the fraction of large operations increases, the average number of sites accessed decreases.

5 Non-Uniform Access Patterns

We now consider the case in which one site is accessed more frequently than others. We present a model for this non-uniform access pattern and generalize previous results to incorporate this type of non-uniformity. Without loss of generality we will assume that site 1 is the special 'hotspot' site. Let p be the probability that an operation accesses site 1; consequently, $1-p$ is the probability that an access does not access site 1. We assume accesses to other sites are uniform among the other

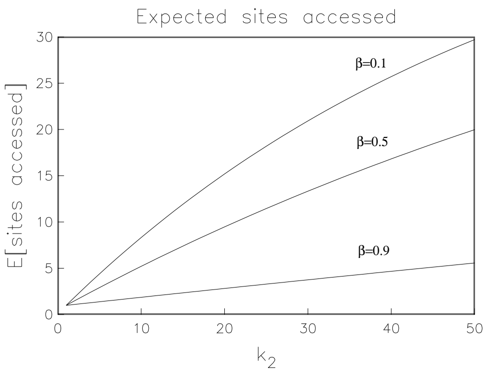

Figure 1: Varying the size of operations

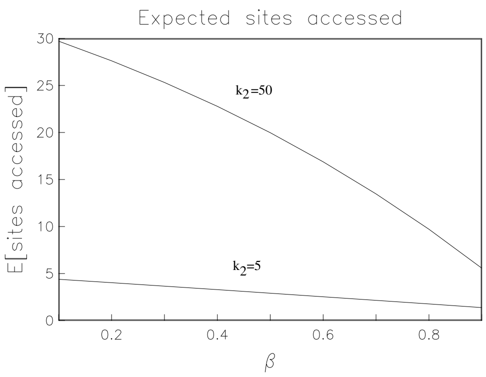

Figure 2: Varying the fraction of small operations

sites. Note that in the case where no hotspots are present, if operation i accesses k_i different sites then $p = \frac{k_i}{n_s}$. We will find the following definitions convenient (the set A is defined in Section 3):

$$\begin{aligned}
A_1 &= \{(\alpha_1, \ldots, \alpha_m) \in A : \alpha_1 = 1\} \\
\bar{A}_1 &= A - A_1 \\
\nu_1(k_1, \ldots, k_t) &= \sum_{(\alpha_1,\ldots,\alpha_m)\in A_1} E\left[U_{\alpha_1 t} U_{\alpha_2 t} \ldots U_{\alpha_m t} | W_1 = k_1, \ldots, W_t = k_t\right] \\
\nu_2(k_1, \ldots, k_t) &= \sum_{(\alpha_1,\ldots,\alpha_m)\in \bar{A}_1} E\left[U_{\alpha_1 t} U_{\alpha_2 t} \ldots U_{\alpha_m t} | W_1 = k_1, \ldots, W_t = k_t\right]
\end{aligned}$$

From the lemma in Section 3,

$$E\left[Y_t^{\underline{m}} |\, W_1 = k_1, \ldots, W_t = k_t\right] = \nu_1(k_1, \ldots, k_t) + \nu_2(k_1, \ldots, k_t).$$

We will first analyze ν_1. Since A_1 contains those m-sized subsets of sites containing site 1, the size of A_1 is $|A_1| = m(n_s - 1)^{\underline{m}}$. Hence, following the analysis of the uniform case presented earlier,

$$\begin{aligned}
\nu_1(k_1, \ldots, k_t) &= \sum_{(\alpha_1,\ldots,\alpha_m)\in A_1} E\left[U_{\alpha_1 t} U_{\alpha_2 t} \ldots U_{\alpha_m t} | W_1 = k_1, \ldots, W_t = k_t\right] \\
&\overset{\text{WLOG}}{=} |A_1| \prod_{i=1}^{t} P\left[U_{1t} = U_{2t} = \ldots = U_{mt} = 1\right] \\
&= m(n_s - 1)^{\underline{m}} \prod_{i=1}^{t} (1-p) \frac{\binom{n_s - m}{k_i}}{\binom{n_s - 1}{k_i}} \\
&= m(n_s - 1)^{\underline{m}} \prod_{i=1}^{t} (1-p) \frac{(n_s - 1 - k_i)^{\underline{m-1}}}{(n_s - 1)^{\underline{m-1}}}.
\end{aligned}$$

Observe that $U_{1t} = 1$ occurs with probability $1-p$ and that the event $\{U_{2t} = \ldots = U_{mt} = 1\}$ implies k_i sites are accessed among $n_s - m$ sites in the numerator of the third step. For the denominator, note that since the terms in the set A_1 have non-zero expectation only when $U_{1t} = 1$, k_i sites are selected from $n_s - 1$ sites.

The calculation of ν_2 is slightly more involved. Define the following events:

$$\begin{aligned}
E_u &= \{U_{2t} = \ldots = U_{(m+1)t} = 1\} \\
E_1 &= \{U_{1t} = 1\}
\end{aligned}$$

Observing that $P[E_1] = (1-p) = 1 - P[E_1^c]$ and $|\bar{A}_1| = (n_s-1)^{\underline{m}}$, we can derive an expression for ν_2 as follows:

$$\begin{aligned}
\nu_2(k_1,\ldots,k_t) &= \sum_{(\alpha_1,\ldots,\alpha_m)\in\bar{A}_1} E\left[U_{\alpha_1 t}U_{\alpha_2 t}\ldots U_{\alpha_m t}|W_1=k_1,\ldots,W_t=k_t\right] \\
&\overset{\text{WLOG}}{=} |\bar{A}_1|\prod_{i=1}^{t} P\left[U_{2t}=\ldots=U_{(m+1)t}=1\right] \\
&= (n_s-1)^{\underline{m}}\prod_{i=1}^{t}\left(P[E_u|E_1]P[E_1]+P[E_u|E_1^c]P[E_1^c]\right) \\
&= (n_s-1)^{\underline{m}}\prod_{i=1}^{t}\left((1-p)\frac{\binom{n_s-m-1}{k_i}}{\binom{n_s-1}{k_i}} + p\frac{\binom{n_s-m-1}{k_i-1}}{\binom{n_s-1}{k_i-1}}\right) \\
&= (n_s-1)^{\underline{m}}\prod_{i=1}^{t}\left((1-p)\frac{(n_s-1-k_i)^{\underline{m}}}{(n_s-1)^{\underline{m}}} + p\frac{(n_s-k_i)^{\underline{m}}}{(n_s-1)^{\underline{m}}}\right)
\end{aligned}$$

Since site 1 can be accessed in $\bar{A}_1$, we condition on the events E_1 (site 1 is not accessed) and E_1^c (site 1 is accessed). When E_1^c occurs, k_i sites are accessed out of the remaining n_s-1 and therefore, for the event $\{U_{2t}=\ldots=U_{(m+1)t}=1\}$ to occur, the k_i accesses must be selected among n_s-m-1 sites. The explanation of the second combination term (when E_1^c occurs) is similar.

Next, for $i=1,2$ let $\bar{\nu}_i$ be the function $\bar{\nu}_i(n_s-k_1,\ldots,n_s-k_t) = \nu_i(k_1,\ldots,k_t)$. Then,

$$\begin{aligned}
E\left[Y_t^{\underline{m}}\right] &= \sum_{k_1,\ldots,k_t}\nu_1(k_1,\ldots,k_t)P[W_1=k_1]\ldots P[W_t=k_t] \\
&\quad + \sum_{k_1,\ldots,k_t}\nu_2(k_1,\ldots,k_t)P[W_1=k_1]\ldots P[W_t=k_t] \\
&= \sum_{k_1,\ldots,k_t}\bar{\nu}_1(n_s-k_1,\ldots,n_s-k_t)P[Z_1=n_s-k_1]\ldots P[Z_t=n_s-k_t] \\
&\quad + \sum_{k_1,\ldots,k_t}\bar{\nu}_2(n_s-k_1,\ldots,n_s-k_t)P[Z_1=n_s-k_1]\ldots P[Z_t=n_s-k_t] \\
&= m(n_s-1)^{\underline{m-1}}\left((1-p)\frac{E\left[(Z-1)^{\underline{m-1}}\right]}{(n_s-1)^{\underline{m-1}}}\right)^t \\
&\quad + (n_s-1)^{\underline{m}}\left((1-p)\frac{E\left[(Z-1)^{\underline{m}}\right]}{(n_s-1)^{\underline{m}}} + p\frac{E\left[Z^{\underline{m}}\right]}{(n_s-1)^{\underline{m}}}\right)^t
\end{aligned}$$

As in the earlier uniform case, the distributions of Y_t and X_t can be obtained from the above expression for the factorial moments of Y_t. The mean and variance are similarly derived. For example, in the special case when $W=1$, the mean value of X_t is found to be:

$$E\left[X_t\right] = n_s - (1-p)^t - (n_s - 1)\left(\frac{n_s - 2 + p}{n_s - 1}\right)^t.$$

Note that the above expression is minimized at $p = \frac{1}{n_s}$, the uniform case, as we might intuitively expect.

6 Summary

In this paper we have considered the problem of computing the distribution of the number of sites accessed by transaction, a distribution that is useful in calculating the time required for a transaction to commit (or abort). We have generalized known results about this problem in several useful ways. First, the standard uniformity assumption has been weakened by considering operations that access a given fixed number of different sites and by allowing one site to be more frequently accessed than others. Second, we provide distributions and higher moments in addition to the mean. Third, we study the case in which the number of operations that make accesses is itself random and provide expressions for some commonly used distributions of this random number. Finally, our results are applicable to the problem of estimating block accesses as well as to problems that fit our class of urn models.

References

[1] A.F. Cardenas. Analysis and Performance of Inverted Database Structures. *Comm. ACM.*, Vol. 18, No. 5, May 1975, pp. 253-263.

[2] T-Y. Cheung. Estimating Block Accesses and the Number of Records in File Management. *Comm. ACM*, Vol. 25, No. 7, July 1982, pp. 484-487.

[3] S. Christodoulakis. Implications of Certain Assumptions in Database Performance Evaluation. *ACM Trans. Database Systems*, Vol. 9, No. 2, June 1984, pp. 163-186.

[4] P. Ciacca, D. Maio and P. Tiberio. A Unifying Approach to Evaluating Block Accesses in Database Organizations. *Info. Proc. Lett.*, Vol. 28, 1988, pp. 253-257.

[5] F.N. David and D.E. Barton. Combinatorial Chance. *Hafner Publishing Company*, New York, NY, 1962.

[6] J. Galambos. The Asymptotic Theory of Extreme Order Statistics. *R.E.Krieger Pub. Co.*, Florida, 1987.

[7] A.M. Langer and A.W. Shum. The Distribution of Granule Accesses Made by Database Transactions. *Comm. CACM*, Vol. 25, No. 11, Nov 1982

[8] N.L. Johnson and S. Kotz. *Discrete Distributions. Houghton Mifflin Co.* Boston, MA. 1969.

[9] N.L. Johnson and S. Kotz. Urn Models and Their Applications. *John Wiley & Sons*, 1977.

[10] W.S. Luk. On Estimating Block Accesses in Database Organizations. *Comm. ACM*, Vol. 26, No. 11, Nov 1983, pp. 945-947.

[11] D. Maio, M.R. Scalas and P. Tiberio. On Estimating Access Costs in Relational Databases. *Info. Proc. Lett.*, Vol. 19, 1984, pp. 157-161.

[12] T.H. Merett and E. Otoo. Distribution Models of Relations. *Proc. 5th Conf. Very Large Databases*, Rio de Janeiro, 1979.

[13] M.T. Ozsu and P. Valduriez. Principles of Distributed Database Systems. *Prentice Hall*, 1991.

[14] P. Palvia and S. March. Approximation Block Accesses in Database Organizations. *Info. Proc. Lett.*, Vol. 19, 1984, pp. 75-79.

[15] D. Potier and P. Leblanc. Analysis of Locking Policies in Database Management Systems. *Comm. ACM*, Vol. 10, Oct 1980, pp. 584-593.

[16] J.B. Rothnie and T. Lozano. Attribute Based File Organization in a Paged Memory Environment. *Comm. ACM*, Vol. 17, No. 2, Feb 1974, pp. 63-69.

[17] K.F. Siler. A Stochastic Evaluation Model for Database Organizations in Data Retrieval Systems. *Comm. ACM*, Vol. 19, No. 2, Feb 1976, pp. 84-95.

[18] A. Thomasian. Determining the Number of Remote Sites Accessed in Distributed Transaction Processing. *IEEE Transactions on Parallel and Distributed Systems*, Vol. 4, No. 1, Jan 1993, pp. 99-103.

[19] K-Y. Whang, G. Wiederhold and D. Sagalowicz. Estimating Block in Database Organizations: A Closed Noniterative Formula. *Comm. ACM*, Vol. 26, No. 11, November 1983, pp. 940-944.

[20] S.B. Yao. Approximating Block Accesses in Database Organization. *Comm. ACM*, Vol. 20, No. 4, April 1977, pp. 260-261.

[21] P.C. Yue and C.K. Wong. Storage Cost Considerations in Secondary Index Selection. *Int. J. Comp. Info. Sci.*, Vol. 4, No. 4, 1975.