Detecting Corrupted Pages in *M* Replicated Large Files

F.K. Hwang and Wenan Zang

Abstract—A file in a distributed database system is replicated on M sites and may contain corrupted pages. Abdel-Ghaffar and El Abbadi gave a detection scheme assuming that the number of corrupted pages $f < M/2$. We replace this assumption by a much weaker one, that, for each page, the majority of copies are correct. Our schemes are based on the structure of the Reed-Solomon code, as proposed by Abdel-Ghaffar and El Abbadi for $M = 2$.

Index Terms—Data consistency, data corruption, fault detection, file comparison, Reed-Solomon code.

1 Introduction

It is often desirable to keep replicated copies of large files at remote sites to prevent loss of information and to permit easy access. However, data in a file can be corrupted due to various reasons. Thus, it is necessary to compare the files from time to time so that data faults can be detected. Metzner and Abidi [10] first proposed the method of *combined signatures*. Divide a file into pages of standard size. A binary parity sequence, called a *signature*, is derived for each page. It is assumed that two copies of a page agree if and only if their signatures agree. However, due to the typically large number of pages in a file, it is still too much work to exchange signatures on all pages for comparison. Instead, combined signatures are exchanged, where a combined signature is a weighted sum of a subset of signatures; when the subsets are properly chosen, all disagreeing pages can be detected.

Many schemes have been proposed (cf. [1], [2], [3], [6], [8], [9], [10], [11]) for detecting disagreeing pages between two replicated files. Let N denote the number of pages and let f denote the number of disagreeing pages, which is assumed to be known. In particular, Abdel-Ghaffar and El Abbadi [2], following an idea of Metzner and Kapturowski [11] of using linear block codes, showed that a one-round exchange of $min\{N, 2f\}$ combined signatures suffice. Abdel-Ghaffar and El Abbadi [1] also proposed a detection scheme for M replicated files. With $M \geq 3$, it is assumed that, for each page, the majority of copies agree (and are presumed to be correct). A *corrupted page*, or a *fault*, is now defined to be a page whose signature disagrees with the majority version. In one model, one site is designated as the *coordinator*, say site 1, while the other sites 2, …, M, are referred to as the *participants*; the coordinator exchanges messages with the participants to detect all corrupted pages in the replicated file. Communication is between the coordinator and the participants, i.e., participants do not exchange any information. Abdel-Ghaffar and El Abbadi gave a detection scheme which transmits at most $(M-2)min\{N, f\} + min\{N, 2f\}$ combined signatures (in possibly two rounds) under the constraint $M \geq 2f + 1$.

In many applications, f is relatively small compared to N, but M can also be a small number. Thus, the constraint $M \geq 2f + 1$ imposes a severe limit on the number of detectable faults. For example, for three sites, only one fault is allowed, even though there are thousands of pages. In this paper, we replace this constraint by the assumption that, for each page, the majority of copies are correct. Note that any model which does not presume an incorruptible copy and allows identical errors must make this assumption, lest the correctness of a page become indeterminable. Thus, our assumption is "minimal" in that sense. We shall also provide insights in the last section on how likely it is that this assumption will be violated and how our algorithms can be modified to detect the violation.

Let us introduce some notions before presenting our schemes. Throughout, f_m stands for the number of corrupted pages at site m, $P_{n,m}$, with $1 \leq n \leq N$ and $1 \leq m \leq M$, stands for the nth page of the copy residing in site m, and $p_{n,m}$ stands for the signature of $P_{n,m}$. Let α be a primitive element in the finite field $GF(2^b)$. Each signature is composed of b bits and, thus, can be considered as an element in $GF(2^b)$. Since, usually, b is much larger than $log_2 N$, we may assume that different pages have distinct signatures. For each site m and for each positive integer j, we define the combined signature $sig_{j,m} = \sum_{n=1}^{N} p_{n,m}\alpha^{jn}$. These signatures are the syndromes of a Reed-Solomon code corresponding to the vector $(p_{1,m}, \ldots, p_{N,m})$ [4]. Finally, given a vector $(e_1, \ldots, e_N)$, its *weight* is defined to be the number of nonzero entries. Our schemes will heavily rely on the following three facts:

FACT 1. *For any given* $E_1, \ldots, E_N \in GF(2^b)$, *the system of equations* $\sum_{n=1}^{N} e_n\alpha^{jn} = E_j$, *where* $j = 1, 2, \ldots, N$, *has a unique solution.*

- *F.K. Hwang is with the Department of Applied Mathematics, Chiao Tung University, Hsinchu 30050, Taiwan, Republic Of China. E-mail: fhwang@math.nctu.edu.tw.*
- *W. Zang is with the Department of Mathematics, University of Hong Kong, Hong Kong. E-mail: wzang@maths.hku.hk.*

Manuscript received 25 Sept. 1996; revised 22 May 1997
For information on obtaining reprints of this article, please send e-mail to: tpds@computer.org, and reference IEEECS Log Number 100309.

FACT 2. *If* $(e_1, \ldots, e_N)$ *has weight, at most,* f *and if* $\sum_{n=1}^{N} e_n \alpha^{jn} = 0$, *where* $j = 1, 2, \ldots, f$, *then* $e_n = 0$ *for each* $n = 1, 2, \ldots, N$.

FACT 3. *For any given* $E_1, \ldots, E_J \in GF(2^b)$, *where* $1 \le J \le N$, *the system of equations* $\sum_{n=1}^{N} e_n \alpha^{jn} = E_j$ *for* $j = 1, \ldots, J$, *has, at most, one solution of* $(e_1, \ldots, e_N)$ *with weight less than or equal to* $\lfloor J/2 \rfloor$. *This solution can be obtained using a Reed-Solomon decoder* [4].

It follows from Fact 3 that, if the number of disagreeing pages between a pair of sites i and j is, at most, $\lfloor f/2 \rfloor$, then $(e_1, e_2, \ldots, e_N) = (p_{1,i} - p_{1,j}, p_{2,i} - p_{2,j}, \ldots, p_{N,i} - p_{N,j})$ is the unique solution with weight at most $\lfloor f / 2 \rfloor$ of the system of equations $\sum_{n=1}^{N} e_n \alpha^{qn} = sig_{q,i} - sig_{q,j}$ for $q = 1, 2, \ldots, f$, and the set of disagreeing pages between i and j is precisely the set of nonzero entries in $(e_1, \ldots, e_N)$.

2 A One-Round Scheme

The algorithm at the coordinator s_1 is shown below.

ALGORITHM (One-Round)
SEND (REQUEST $sig_{1,m}, \ldots, sig_{\min\{N,2f\},m}$) to all participants m
RECEIVE ($sig_{1,m}, \ldots, sig_{\min\{N,2f\},m}$) from all participants m
FOR each pair of sites i and j, where $1 \le i, j \le M$
compute a vector $(e_{1;i,j}, \ldots, e_{N;i,j})$ of weight, at most, f as a solution of the system of equations

$$\sum_{n=1}^{N} e_{n;i,j} \alpha^{qn} = sig_{q,i} - sig_{q,j}, \text{ where}$$

$$q = 1, 2, \ldots, min\{N, 2f\} \quad (1)$$

END(*FOR*)
FOR each page p, where $1 \le p \le N$
CONSTRUCT a graph G_p with vertex-set $\{1, 2, \ldots, M\}$ such that $[i, j]$ isan edge in G_p iff $e_{p;i,j} = 0$
FIND a maximum clique C_p in G_p
END(*FOR*)
FOR each site m, where $1 \le m \le M$
RETURN $F_m = \{p : m \notin C_p\}$ (*F_m is the set of corrupted pages at site m*)
END(*FOR*)

Our algorithm is based on the scheme of Abdel-Ghaffar and El Abbadi [1]. Each participant sends $min\{N, 2f\}$ combined signatures to the coordinator. It follows from Fact 1 and Fact 3 that, by comparing the $min\{N, 2f\}$ combined signatures for any pair of sites i and j, the coordinator can obtain the unique solution $(e_{1;i,j}, \ldots, e_{N;i,j})$ of (1) with weight, at most, f. Since, for each page p, $e_{p;i,j} = 0$ if and only if site i and site j agree on page p, the coordinator can thus identify all the disagreeing pages between any pair of sites i and j. Let G_p be the graph constructed in the algorithm. Then, each connected component of G_p is a clique. Note that the maximum clique problem can be solved in linear time by depth-first search (see, for instance, [12]). By assumption, the majority of sites agree on each page p. Therefore, G_p contains a (unique) maximum clique C_p as $|C_p| \ge \lceil (M+1)/2 \rceil$, where $\lceil x \rceil$ is the smallest integer not less than x. Let $F_m = \{p : m \notin C_p\}$. Clearly, F_m is the set of faults at site m.

THEOREM 1. *The above one-round scheme requires the transmission of* $(M-1)min\{N, 2f\}$ *combined signatures to identify* f *corrupted pages. Moreover,* $(M-1)min\{N, 2f\}$ *is the minimum number of combined signatures needed to be transmitted for any one-round scheme if* $N \ge f$.

PROOF. Abdel-Ghaffar and El Abbadi [2] proved that a site with f faults must transmit $min\{N, 2f\}$ combined signatures to identify all f faults. Suppose that the $M-1$ participants transmit a total of fewer than $(M-1)min\{N, 2f\}$ combined signatures. Then, there exists a participant transmitting fewer than $min\{N, 2f\}$ combined signatures. Consider the case in which this participant contains f faults (we need the assumption $N \ge f$ here). There is no guarantee that the f faults will be identified. □

3 A Two-Round Scheme

Suppose that faults are randomly distributed in the M replicated files. Then, as M grows, the probability that any two replicated files contain more than half of the faults becomes small. Abdel-Ghaffar and El Abbadi [1] took advantage of this observation to reduce the number of combined signatures needed to be transmitted in the first round, with the possibility of transmitting a few more combined signature in the second round in case the small-probability event occurs. We use the same idea here for $M \ge 3$.

In the case $N \le f$, our algorithm is in the same spirit as the one stated in the preceding section: Each participant m sends all its N pages signatures $p_{1,m}, \ldots, p_{N,m}$ to the coordinator. For each page n and each pair of sites i and j, set $e_{n;i,j} = p_{n,i} - p_{n,j}$, where $1 \le n \le N$ and $1 \le i, j \le M$. Let G_p be a graph with vertex-set $\{1, 2, \ldots, M\}$ such that $[i, j]$ is an edge in G_p iff $e_{p;i,j} = 0$. Since the majority of sites agree on page p, G_p contains a (unique) maximum clique C_p as $|C_p| \ge \lceil (M+1)/2 \rceil$. Let $F_m = \{p : m \notin C_p\}$. Clearly, F_m is the set of faults at site m. The total number of signatures sent is $(M-1)N$.

Let us proceed to the case $N \ge f$.

ALGORITHM (Two-Round with $N \ge f$)
SEND (REQUEST $sig_{1,m}, \ldots, sig_{f,m}$) to all participants m
RECEIVE ($sig_{1,m}, \ldots, sig_{f,m}$) from all participants m
FOR each pair of sites i and j, where $1 \le i, j \le M$
Try to compute a vector $e_{i,j} = (e_{1;i,j}, \ldots, e_{N;i,j})$ of weight at most $\lfloor f/2 \rfloor$ as a solution of the system of equations

$$\sum_{n=1}^{N} e_{n;i,j} \alpha^{qn} = sig_{q,i} - sig_{q,j}, \text{ where}$$

$q = 1, 2, \ldots, f$ (2)

Let $w_{i,j}$ denote the weight of the vector $(e_{1;i,j}, \ldots, e_{N;i,j})$ if such a solution exists, and let W be the set of all pairs $\{i, j\}$ of sites such that either (a) no solution of (2) with weight at most $\lfloor f/2 \rfloor$ exists *or*

(b) there is a solution of (2) with weight $w_{i,j}$ at most $\lfloor f/2 \rfloor$ satisfying $w_{i,j} > w_{s,t}$ for any pair $\{s, t\}$ of sites disjoint from $\{i, j\}$

END(*FOR*)
IF $\Omega = \emptyset$ **THEN**
  FOR each page p, where $1 \le p \le N$
    CONSTRUCT a graph G_p with vertex-set $\{1, 2, \ldots, M\}$ such that $[s, t]$ is an edge in G_p iff $e_{p;s,t} = 0$
    FIND a maximum clique C_p in G_p
  END(*FOR*)
ELSE $\Omega \neq \emptyset$
  IF $\Omega = \{\{i, j\}, \{j, k\}, \{k, i\}\}$ for some three sites i, j, and k **THEN**
    SEND (REQUEST $sig_{f+1,m}, \ldots, sig_{min\{N,2f\},m}$) to all sites i, j, and k
    RECEIVE ($sig_{f+1,m}, \ldots, sig_{min\{N,2f\},m}$) from all sites i, j, and k
    FOR each pair $\{s, t\} \in \Omega$
      replace $e_{s,t}$ by the solution $(e_{1;s,t}, \ldots, e_{N;s,t})$ with weight at most f of the system of equations

$$\sum_{n=1}^{N} e_{n;s,t}\alpha^{qn} = sig_{q,s} - sig_{q,t},$$

      where $q = 1, 2, \ldots, min\{N, 2f\}$
    END(*FOR*)
    FOR each page p, where $1 \le p \le N$
      CONSTRUCT a graph G_p with vertex-set $\{1, 2, \ldots, M\}$ such that $[s, t]$ is an edge in G_p iff $e_{p;s,t} = 0$
      FIND a maximum clique C_p in G_p
    END(*FOR*)
  ELSE some site i is contained in each pair in Ω
    FOR each page p, where $1 \le p \le N$
      CONSTRUCT a graph $G_p - \{i\}$ with vertex-set $\{1, 2, \ldots, M\} - \{i\}$ such that $[s, t]$ is an edge in $G_p - \{i\}$ iff $e_{p;s,t} = 0$
      FIND a maximum clique C_p in $G_p - \{i\}$
    END(*FOR*)
    Let S be a minimal subset of $\{1, 2, \ldots, M\} - \{i\}$ which intersects every clique C_p with $|C_p| > (M-1)/2$, where $1 \le p \le N$
    SEND (REQUEST $sig_{f+1,m}, \ldots, sig_{min\{N,2f\},m}$) to all sites m in $\{i\} \cup S$
    RECEIVE ($sig_{f+1,m}, \ldots, sig_{min\{N,2f\},m}$) from all sites m in $\{i\} \cup S$
    FOR each pair $\{i, t\}$ with $t \in S$
      replace $e_{i,t}$ by the solution $(e_{1;i,t}, \ldots, e_{N;i,t})$ with weight at most f of the system of equations

$$\sum_{n=1}^{N} e_{n;i,t}\alpha^{qn} = sig_{q,i} - sig_{q,t},$$

      where $q = 1, 2, \ldots, min\{N, 2f\}$
    END(*FOR*)
    FOR each page p with $|C_p| > (M-1)/2$ and $e_{p;s,t} = 0$ for some $t \in C_p > S$, where $1 \le p \le N$
      SET $C_p = C_p \cup \{i\}$
    END(*FOR*)
    FOR each page p with $|C_p| = (M-1)/2$, where $1 \le p \le N$
      IF there exists $t \in S$ such that $(t \in C_p$ and $e_{p;s,t} = 0)$ or $(t \notin C_p$ and $e_{p;s,t} \neq 0)$ **THEN**
        SET $C_p = C_p \cup \{i\}$
      ELSE
        SET $C_p = \{1, 2, \ldots, M\} - C_p$
      END(*IF*)
    END(*FOR*)
  END(*IF*)
END(*IF*)
FOR each site m, where $1 \le m \le M$
  RETURN $F_m = \{p: m \notin C_p\}$ (*F_m is the set of corrupted pages at site m*)
END(*FOR*)

In the present case, instead of $min\{N, 2f\}$ combined signatures, only f combined signatures $sig_{q,m} = \sum_{n=1}^{N} p_{n,m}\alpha^{qn}$, are transmitted from each participant m to the coordinator, where $1 \le q \le f$. If a pair of sites contains no more than $f/2$ disagreeing pages, then the comparison of f combined signatures between the pair is sufficient to detect all the disagreeing pages. Any pair violating this condition is called abnormal (other pairs are called normal).

For each pair of sites i and j, the coordinator tries to compute a vector $(e_{1;i,j}, \ldots, e_{N;i,j})$ of weight at most $\lfloor f/2 \rfloor$ as a solution of (2).

FACT 4. *If $\{i, j\}$ is an abnormal pair, then $\{i, j\}$ is in Ω (defined in the algorithm).*

PROOF. We aim to prove that either (a) or (b) stated in the algorithm holds for $\{i, j\}$.

Let A denote the nonempty set of all abnormal pairs. Then A can not contain two disjoint pairs or the total number of faults would be greater than f. If $\{i, j\}$ is an abnormal pair, then each pair $\{s, t\}$ disjoint from $\{i, j\}$ is normal, hence, there is a solution with weight $w_{s,t}$ at most $\lfloor f/2 \rfloor$ of (2) with $\{s, t\}$ in place of $\{i, j\}$. Assume that $(e_{1;i,j}, \ldots, e_{N;i,j})$ is a solution of (2) with weight no more than $\lfloor f/2 \rfloor$. Since $(p_{1,i} - p_{1,j}, \ldots, p_{N,i} - p_{n,j})$ is a solution of (2) with weight greater than $\lfloor f/2 \rfloor$, these two solutions are different. Note that $\sum_{n=1}^{N} [e_{n;i,j} - (p_{n,i} - p_{n,j})]\alpha^{qn} = 0$, where $q = 1, 2, \ldots, f$. By Fact 2, there are at least $f + 1$ values of n for which $e_{n;i,j} \neq p_{n,i} - p_{n,j}$. For each two

sites k and l, let $d_{k,l}$ denote the number of disagreeing pages between them. Then, $w_{i,j} + d_{i,j} \geq f + 1$, whence $w_{i,j} + f_i + f_j \geq f + 1$. So, for each pair $\{s, t\}$ disjoint from $\{i, j\}$, we have $w_{i,j} \geq f - f_i - f_j + 1 > f_s + f_t \geq d_{s,t} = w_{s,t}$, the last equality holds because $\{i, j\}$ is an abnormal pair, $f_i + f_j > \lfloor f/2 \rfloor$, hence, $d_{s,t} \leq f_s + f_t \leq \lfloor f/2 \rfloor$. □

Using the fact that Ω can not contain two disjoint pairs, we immediately have the following statement.

FACT 5. *Let $\{i, j\}$ be an arbitrary pair of sites in Ω. Then one of the following three cases occurs.*

Case 1. *There is a site k outside $\{i, j\}$ such that $\Omega = \{\{i, j\}, \{j, k\}, \{k, i\}\}$;*
Case 2. *Each pair in Ω contains site i;*
Case 3. *Each pair in Ω contains site j.*

Let us consider Case 1, now. By Fact 4, any pair of sites outside Ω is normal. By Fact 3, all the disagreeing pages between this pair have already been detected. Now, transmit $min\{N - f, f\}$ additional combined signatures from each site in $\{i, j, k\}$ to the coordinator. By Fact 3, all the disagreeing pages between any two sites in $\{i, j, k\}$ are identified. Thus, G_p is determined for each page p. Let C_p be the maximum clique in G_p. Then, $F_m = \{p: m \notin C_p\}$ is the set of corrupted pages at site m. The total number of combined signatures transmitted is $(M - 1)f + 3min\{N - f, f\}$.

Consider Case 2. (Case 3 is a mirror image of Case 2, we can handle it similarly.) First, note that $G_p - \{i\}$ is already known for each page p. For even M, $G_p - \{i\}$ contains a unique clique C_p of size greater than $(M - 1)/2$, and every vertex not in C_p is a fault. Thus, F_m is known, except for $m = i$. For odd M, $Gp - \{i\}$ may contain two disjoint cliques, both of size $(M - 1)/2$. Call such a page ambiguous. Note that all pages in one of the two cliques induce faults, but site i must be corrected. Let S denote a set of sites excluding i, such that $S \cap C_p \neq \emptyset$ for each nonambiguous page p. Observe that any set of $\lfloor (M + 1)/2 \rfloor$ sites excluding i will do.

Request each site in S as well as site i to transmit $min\{N - f, f\}$ more combined signatures to the coordinator. Then, by Fact 3, all disagreements between i and sites in S are detected. For each ambiguous page p, compare site i with an arbitrary site k in S. Let C_p and C'_p denote the two disjoint maximum cliques in $G_p - \{i\}$ with $k \in C_p$. Then, i agreeing with k on page p implies that every site in C'_p is a fault; i disagreeing with k implies that every site in C_p is a fault. For each nonambiguous page p, compare i with $k \in S \cap C_p$. Then, i agreeing with k on page p implies that i is correct on page p; i disagreeing with k implies that i is a fault on page p. Hence, when a second round is needed, the total number of combined signatures transmitted is $(M - 1)f + (|S| + 1)$ $min\{N - f, f\}$.

For efficiency, it is desirable to have a small S. We shall prove that, if S is an arbitrary minimal set (with respect to set inclusion) of sites excluding i, such that $S > C_p \neq \emptyset$ for each nonambiguous page p, then $|S| \leq min\{\lceil \sqrt{f} \rceil, \lfloor (M + 1)/2 \rfloor\}$.

For each site $s \in S$, there must exist a nonambiguous page p such that s is the only site in $S \cap C_p$; otherwise, we can remove s from S, contradicting the minimality of S. For page p, each site in $S - \{s\}$ corresponds to a fault. Hence, $|S|(|S| - 1) \leq f$, and $|S| \leq \lceil \sqrt{f} \rceil$. On the other hand, since any set of $\lfloor (M + 1)/2 \rfloor$ sites excluding i intersects C_p for any nonambiguous page p, we have $|S| \leq \lfloor (M + 1)/2 \rfloor$. Therefore, $|S| \leq min\{\lceil \sqrt{f} \rceil, \lfloor (M + 1)/2 \rfloor\}$.

To get a minimal set S of sites with the desired property in polynomial time, we first take an arbitrary set S of $\lfloor (M + 1)/2 \rfloor$ sites excluding i. A site m in S is called redundant if $S - \{m\}$ still intersects every clique C_p for each nonambiguous page p. Let m be a redundant site in S, if any, replace S by $S - \{m\}$. The deletion procedure proceeds until S contains no redundant site.

It is worthwhile pointing out that, in the case $f \leq M - 2$, there exists a site $s \neq i$ that contains no corrupted page. Clearly, site s intersects every C_p for each nonambiguous page p. Hence, $|S| = 1$.

Combining all the above statements, we get the following theorem.

THEOREM 2. *The above two-round scheme requires the transmission of at most*

$$(M - 1)min\{N, f\} + min\{1 + \lceil \sqrt{f} \rceil, 1 + \lfloor (M + 1)/2 \rfloor\} min\{N - f, f\}$$

combined signatures to identify f corrupted pages.

In Case 2 and Case 3, we proposed to find a *minimal* (with respect to set inclusion rather than size) set S of sites excluding i such that $S \cap C_p \neq \emptyset$ for each nonambiguous page p. Clearly, the smaller the size of S, the more efficient the algorithm. The reason why we did not try to find a minimum S (with respect to size) with the desired property is as follows: Let T denote the set of all nonambiguous pages. For each site $m \neq i$, let T_m denote the set of all nonambiguous pages p such that $m \in C_p$. To obtain S, we have to solve the following problem:

Input: A finite set T and a family $\{T_m : m \neq i\}$ of subsets of T such that $T = \cup_{m \neq i} T_m$.
Output: A minimum set S such that $T = \cup_{m \in S} T_m$.

which is the well-known set-covering problem and, therefore, NP-hard. Johnson [7] showed that a greedy heuristic algorithm returns a set cover with a ratio bound of $ln|T| + 1$. Feige [5] proved that the set-covering problem can not be approximated within $(1 - \epsilon)ln|T|$ for any $\epsilon > 0$, unless $NP \subseteq DTIME(n^{loglog\, n})$.

4 REMARKS

Our algorithms are based on the assumption that for each page, the majority of copies are correct. Without this assumption, we are unable to distinguish between the correct pages and the corrupted pages, and thus, detection failure occurs. Let us give a rough estimate of the probability that

the assumption fails. For $f \geq \lceil M/2 \rceil$, the probability that a given page has at least $\lceil M/2 \rceil$ corrupted pages is

$$\sum_{i=\lceil M/2 \rceil}^{\min\{M,f\}} \frac{\binom{M}{i}\binom{(N-1)M}{f-i}}{\binom{NM}{f}},$$

which is dominated by the term $i = \lceil M/2 \rceil$; this term is roughly $\binom{f}{\lceil M/2 \rceil} \Big/ N^{\lceil M/2 \rceil}$. Multiplying this by N, we get an estimate of the probability that there exists a page with at least $\lceil M/2 \rceil$ corrupted copies. For N = 100, $f = M$ = 3, this estimate is 0.03; when $f = M$ = 5, it is 0.001. For N = 1,000, $f = M^2$ = 9, this estimate is 0.036; when $f = M^2$ = 25, it is 0.0023. For N = 1,000, $f = 0.01NM$ = 30, this estimate is 0.45; when $f = 0.01NM$ = 50, it is 0.021.

Observe that our algorithms can be modified to detect the case when a majority of copies of some page are corrupted.

In the one-round scheme and in the cases $\Omega = \emptyset$ and $\Omega = \{\{i, j\}, \{j, k\}, \{k, i\}\}$ in the two-round scheme, after obtaining C_p for each page p, let us check the size of C_p. If $|C_p| < \lceil (M+1)/2 \rceil$, we declare detection failure; otherwise, output F_m for each site m.

In the two-round scheme, in case some site i is contained in each pair in Ω, let us modify S to be a minimal subset of $\{1, 2, \ldots, M\} - \{i\}$ that intersects *every* maximum clique C with $|C| \geq (M-1)/2$ in each $G_p - \{i\}$. Possibly some $G_p - \{i\}$ contains two maximum cliques of size $(M-1)/2$. Since each maximum clique is a connected component in $G_p - \{i\}$, we can find these cliques by depth-first search. Then, we modify the loop "for each page p with $|C_p| = (M-1)/2$, where $1 \leq p \leq N$, …" (right above the loop for returning F_m) as follows: For each page p with $C_p = (M-1)/2$, let $t \in S \cap C_p$. If $e_{p;i,t} = 0$, then set $C_p = C_p \cup \{i\}$; if $e_{p;i,t} \neq 0$, then check if $\{1, \ldots, M\} - (C_p \cup \{i\})$ is a clique. If yes, denote this clique by C and take an arbitrary t' in $S \cap C$. If $e_{p;i,t'} = 0$, then set $C_p = C \cup \{i\}$. After obtaining C_p for each page p, let us check the size of C_p. If $|C_p| < \lceil (M+1)/2 \rceil$ for some page p, we declare detection failure; otherwise, output F_m for each site m.

In this paper, we proposed two schemes for efficient detection of corrupted pages in a large replicated file. Our approach is based on the previous schemes of Abdel-Ghaffar and El Abbadi [1], [2]. In these two papers, the lower bounds of the communication complexity of the coordinator-based model were derived and proved to be tight. However, the communication complexity of our two-round scheme is bigger than the lower bound presented in the previous papers. The tight lower bound for the present model remains unknown. We close the paper with this open problem.

ACKNOWLEDGMENTS

Wenan Zang's research was supported in part by CRCG grant 33/024/0010.

REFERENCES

[1] K.A.S. Abdel-Ghaffar and A. El Abbadi, "Efficient Detection of Corrupted Pages in a Replicated File," *Proc. 12th ACM Symp. Principles on Distributed Computing*, pp. 219-229, 1993.
[2] K.A.S. Abdel-Ghaffar and A. El Abbadi, "An Optimal Strategy for Comparing File Copies," *IEEE Trans. Parallel and Distributed Systems*, vol. 5, no. 1, pp. 87-93, Jan. 1994.
[3] D. Barbará, H. Garcia-Molina, and B. Feijoo, "Exploiting Symmetries for Low-Cost Comparison of File Copies," *Proc. Eighth Int'l Conf. Distributed Computer Systems*, June 1988.
[4] R.E. Blahut, *Theory and Practice of Error Control Codes*. Reading, Mass.: Addison-Wesley, 1984.
[5] U. Feige, "A Threshold of $\ln(n)$ for Approximating Set Cover," *Proc. 28th ACM Symp. Theory of Computing*, pp. 314-318, 1996.
[6] W. Fuchs, K.L. Wu, and J.A. Abraham, "Low-Cost Comparison and Diagnosis of Large Remotely Located Files," *Proc. Symp. Reliability Distributed Software and Database Systems*, pp. 67-73, Los Angeles, 1986.
[7] D.S. Johnson, "Approximation Algorithms for Combinatorial Problems," *J. Computer System Science*, vol. 9, pp. 256-278, 1974.
[8] J.J. Metzner, "A Parity Structure for Large Remotely Located Data Files," *IEEE Trans. Computers*, vol. 32, no. 8, pp. 727-730, Aug. 1983.
[9] J.J. Metzner, "Efficient Replicated Remote File Comparison," *IEEE Trans. Computers*, vol. 40, no. 5, pp. 651-660, May 1991.
[10] J.J. Metzner and M.A. Abidi, "Remote Comparison and Correction of Duplicated Data Files," *Proc. Nat'l Telecomm. Conf.*, pp. 59.4.1-59.4.4, Nov. 1979.
[11] J.J. Metzner and E.J. Kapturowski, "A General Decoding Technique Applicable to Replicated File Disagreement Location and Concatenated Code Decoding," *IEEE Trans. Information Theory*, vol. 36, pp. 911-917, July 1990.
[12] R.E. Tarjan, *Data Structures and Network Algorithms*. SIAM, 1988.

F.K. Hwang received his BA from National Taiwan University in 1960 and his PhD from North Carolina State University in 1968. He worked at Bell Labs from 1967 to 1996 as a member of the research staff, and is now a professor at the National Chiaotung University. He has published six books, 300 papers, and has been granted 12 patents.

Wenan Zang obtained his PhD degree in operations research from Rutgers University in 1995. He has been an assistant professor at the University of Hong Kong since 1996. His main research interests are in combinatorics and optimization.