ホームページ改ざん検知システムにおける コンテンツ更新状況に基づく更新作業効率化

富永 浩之† 高屋敷 光一† 猪股 俊光† 曽我 正和††
†岩手県立大学大学院ソフトウェア情報学研究科 ††岩手県立大学地域連携研究センター

1 はじめに

ホームページに対する改ざん攻撃に対処すべく，筆者らはディジタル署名とパトロールを基にしたホームページ改ざん検知システム iP@TROL[1] の開発を行っている．iP@TROL システムでは，ホームページの作成者であるユーザがサーバ上のコンテンツの更新を行う際，ホームページを構成するコンテンツに対してディジタル署名を作成したのち，アップロードが行われる．その更新作業の効率化を目指し，従来方式における改ざん検知の性能を損なうことなく，必要最小限のコンテンツのみを更新対象とするために，セグメントという概念を導入したセグメント差分更新方式 [2] を提案した．この方式を用いることで，更新時間の短縮がはかられたものの，長期的運用等によってその効果が減少する課題が残された．

本研究では，従来のセグメント差分更新方式におけるこの課題を解決することを目的とし，長期的運用において更新時間の最適化を行う手法を考案した．

2 セグメント差分更新方式

■セグメント 図 1 に示すように，ファイル数の閾値 SC・ファイル容量の閾値 SS に基づき，全コンテンツはセグメントとよぶファイル（ディレクトリを含む）群に分割され（こうして定義したセグメントの情報を，セグメント定義とよぶ），セグメントごとに改ざん検知のための署名データが作成される．あるセグメントの署名データの作成に際し，そのセグメントに含まれる各ファイルから 1 つのダイジェスト値が計算され，秘密鍵媒体を用いてダイジェスト値に対する署名計算が行われる．

■更新リスト・削除リスト・セグメントリスト 更新処理時に各ファイルが生成される．更新リストと削除リストには，それぞれ，更新対象と削除対象のファイルの情報が格納される．セグメントリストには，更新リストか削除リストに含まれるファイルを含むセグメントの情報が格納される．これらを更新情報とよぶ．

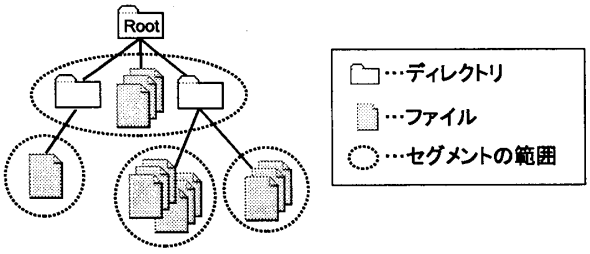

図 1: セグメント定義の例

■セグメント定義ファイル コンテンツのディレクトリ構成，セグメント定義，各ファイルの更新日時情報が記録されたファイル．更新処理の開始時に読み込まれ，終了時に保存される．これを用いて，差分情報（クライアント／FTP サーバ間のコンテンツの差分に関する情報）が算出され，更新情報が生成される．

2.1 問題点

- 一度決定したセグメント定義は（ユーザによって強制リセットされない限り）ホームページの運用の段階で最適化が行われることはないため，「1. 高頻度で同時に更新されるファイル組が複数セグメントに分散し続ける」，「2. ファイルの追加・削除の蓄積によって大規模すぎるセグメントや空のセグメントが発生する」といった現象が起こる．
- 新規に追加されたファイルが複数個あるとき，それらが別々のセグメントに分散される可能性が高くなる．

これらに起因し，不必要なダイジェスト値計算・署名計算が発生する．

3 動的セグメント更新方式

■更新履歴の記録 更新処理の終了時に，その実行日時と，更新されたファイルのリストを，更新スタンプとして更新スタンプファイルに保存する．更新スタンプファイルは図 2 のように高々 n 個が保存され，実行日時が古い更新スタンプファイルはその都度破棄される（n は運用開始前に定める上限値）．

■更新履歴の分析 更新処理において更新情報を生成したのち，更新スタンプファイルの内容を分析し，「ファイル・セグメントごとの更新頻度」，「高頻度で同時に更新されるファイル組」を算出する．

A Method for Efficient Updating Based on State of Update in Tamper Detection System for Web Pages
† Hiroyuki TOMINAGA, Kouichi TAKAYASHIKI, Toshimitsu INOMATA
†† Masakazu SOGA
Graduated School of Software and Information Science, Iwate Prefectural University (†)
Iwate Prefectural University Regional Cooperative Research Center (††)

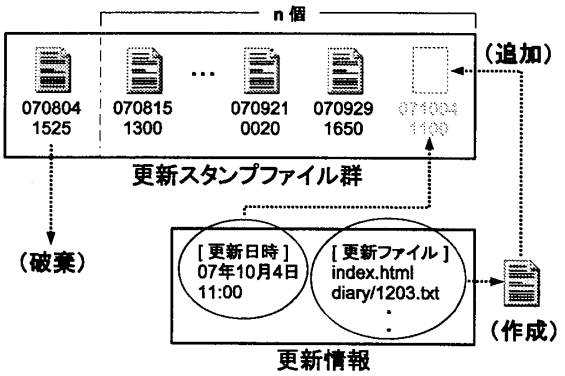

図 2: 更新スタンプの作成と破棄

■セグメント定義の調整　更新履歴の分析結果をもとに,「ファイル組の同一セグメントへの統合」,「空セグメントの削除」,「追加ファイルの特定セグメントへの集積」を行う．このとき，$SC' > SC$ かつ $SS' > SS$ なるファイル数の閾値 SC' とファイル容量の閾値 SS' を用いてセグメントの規模に上限を設け，適宜，新しいセグメントを追加する．

4　評価と考察

クライアントマシン（CPU866MHz・RAM256MB），FTP サーバ（CPU333MHz・RAM256MB），LAN（100Mbps）の環境下で，ファイル数 64・総容量 128MB で構成されるコンテンツを準備し，各方式における更新処理時間を測定した．ここで，F_1〜F_{64} はそれぞれ，2048kB の容量をもつファイルである．

さまざまな閾値設定で F_1〜F_7 を従来方式 [2] で更新した場合の更新処理時間を，図 3 に示す．閾値 "$SC = 1$"，すなわちファイル 1 個ごとにセグメントを定義した場合では，更新するセグメントの数が多いため，他と比較して署名計算の時間が長い．ファイル容量の閾値 "$SS = \infty$" の場合では，セグメントあたりのファイル容量が大きく，他と比較してダイジェスト値計算の時間が長い．これらに対し，閾値 "$SC = 100, SS = 16000$" の場合，それぞれの処理時間がバランスよく短縮され，全体の処理時間も他と比較して短い．これらのことから，関連性の高いファイル組の同一セグメントへの統合，適度なセグメントの規模の維持が，更新時間の短縮効果の向上につながることがわかる．

閾値設定を $SC = 50$，$SS = 20000$，$SC' = 100$，$SS' = 25000$，$n = 16$ とし，1〜8 回目に F_1，F_{11}，F_{21}，... ，F_{61} の計 7 個のファイルを，9〜16 回目に F_1，F_{11}，F_{21}，F_{31} と F_{32}，F_{42}，F_{52}，F_{62} の計 8 個のファイルを更新した場合の，各方式における更新処理時間の推移を，図 4 に示す．全試行を通じてもっとも更新頻度が高いファイル組は，F_1，F_{11}，F_{21}，F_{31} である．ここで，提案補正方式は，提案方式において

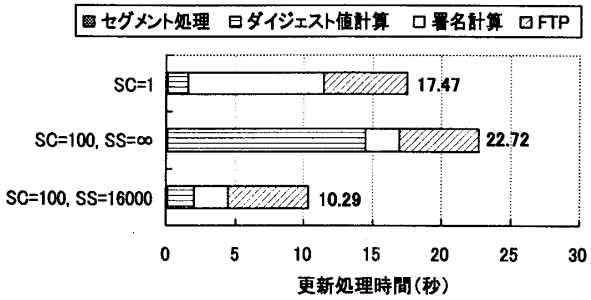

図 3: 従来方式の各閾値における更新処理時間

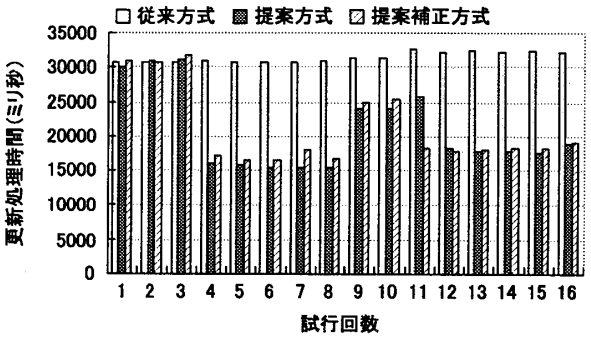

図 4: 各方式の更新処理時間の推移

算出する更新頻度に，実行日時の情報による補正を加える方式である．図 4 において，従来方式では，更新の蓄積によって更新処理時間が短縮されることはなく，ほぼ一定値を示している．それに対し提案方式では，4 回目および 12 回目以降，ファイル組の統合によって更新処理時間が短縮されている．さらに提案補正方式では，日時情報による補正によって 2 度目のファイル組の統合のタイミングが早まっている．これらのことから，提案方式では長期的運用における更新処理時間の向上が実現されているといえる．

5　おわりに

本研究では，ホームページの作成者がコンテンツを更新する際の負荷の軽減をはかるために動的セグメント更新方式を提案し，所期の効果が得られることを確認した．今後の課題として，動的セグメント処理の実行の可否やパトロール対象範囲のセグメント単位での設定機能，ディレクトリ構成に依存しないセグメントの形態の検討などが挙げられる．

参考文献

[1] 猪股俊光，板垣晋，曽我正和，西垣正勝:"ディジタル署名とパトロールを用いた電子情報改ざん検知方式とWWW への応用"，情報処理学会論文誌，Vol.44 No.8, pp.2072–2084(2003).

[2] 富永浩之，高屋敷光一，猪股俊光，曽我正和:"ホームページ改ざん検知システムにおける更新作業効率化のための一手法"，情報処理学会第 69 回全国大会講演論文集 4W-4(2007).