

3B-6

ユビキタス情報蓄積システムにおける 体験記録データの複製方式

井上 知洋† 中村 隆幸† 中村 元紀† 山口 正泰†

† 日本電信電話(株) NTT 未来ねっと研究所

1 はじめに

写真や電子メールなどの個人の体験に関わるデジタル情報を個人所有の PC に永続的に蓄積していく事は徐々に一般的に成りつつある。最近ではウェアラブルセンサーなどを用いてより大容量で網羅的な体験記録情報を蓄積し活用する取り組みも始まっており、ライフログ、体験記録システムなどの名前で新しい研究分野として広まりつつある。我々も、様々な場所に設置された公共的なセンサーを用いて個人の日常の行動履歴を記録し蓄積するシステムを提案している[1]。

我々は、このように集められた膨大な量の体験記録データが個人所有のストレージに蓄積される事を前提に、蓄積データを活用するシステムについて研究を行っている。本稿では、蓄積されたデータの一部をモバイル端末などに複製して利用する際の、複製対象を選択する方式について提案する。

2 体験記録データの特徴

現在のユーザは電子メールやオフィス文書など、個人の活動履歴に関わる電子データを個人 PC の中に蓄積している。将来のユーザは、様々な場所に設置されたセンサーやユーザ自身が持ち運ぶウェアラブルセンサーによって、行動や体験に関わる様々な種類のデータを収集して、同じように蓄積していくことになる。このように集められる情報を本稿では体験記録データと呼ぶ。

体験記録データは、現在のユーザが収集している一般的なデータとは次の二点で異なる。まず、蓄積されるデータ量が現在と比較して著しく大きい。次に、センサーなどによって収集されたデータは、何らかの後処理的な意味づけが施されない限りデータ単体ではほとんど価値を持たない。

このため、蓄積された履歴データを事後的に解析し何らかの形で意味づけ・関連づけて整理することが一般に行われる[2]。この結果、体験記録データは様々な意味体系によって整理される。メディア種類別や時系列別などの一般的な整理方法の他、各アプリケーションは関連性を抽出した結果を独自の体系で整理する。例えばユーザの行動履歴からコンテキストを抽出し、特定のコンテキストにおかれていた時間帯に収集されたデータは「イベント」という単位に纏められて整理される。体験記録データでは、

これらの意味的な依存関係にあるデータ集合の扱いが特に重要になる。

3 複製対象選択システム

前述のような体験記録データを複数端末にまたがって利用する事を考える。ユーザの全ての履歴データは、ホームサーバのような単一のストレージに蓄積されていることを前提とし、その上の蓄積データをモバイル端末に複製して利用することとする。将来においても蓄積データの全てをモバイル端末に複製することは難しいため、蓄積データの中から適当な部分集合を選択してモバイル端末に複製することになる。この部分集合選択の機能を本稿では複製対象選択システムと呼ぶ。

3.1 課題

複製対象選択システムは、2 節で説明した体験記録データの特徴を踏まえると、以下の条件を満たす必要がある。(a) 膨大な量の蓄積データに対して複製対象選択が効率的に動作する。(b) 単独では価値を持たないセンサーデータを複製する際には「イベント」のような意味的な関係を保存したまま複製対象を選択する。

4 提案方式

本稿では、上記の課題を解決する複製対象選択システムの構成を提案する(図1)。システムは大きく、ストレージ部と複製マネージャ部に分かれる。

4.1 ストレージ部

ストレージ部は内部的にデータストアとメタデータ DB に分けられ、データストアは収集されたコンテンツそのものを格納する一方、メタデータ DB は蓄積コンテンツのメタデータを管理する。

提案システムでは、ストレージ部に蓄積されるコンテンツの内、関連性のある複数のデータを既存のディレクトリ構造とは独立の構造を用いて構造化して格納する。本システムを利用する外部アプリケー

A Selection and Replication Method for the Ubiquitous Logging System of Personal Data
Tomohiro Inoue, Takayuki Nakamura, Motonori Nakamura, Masayasu Yamaguchi

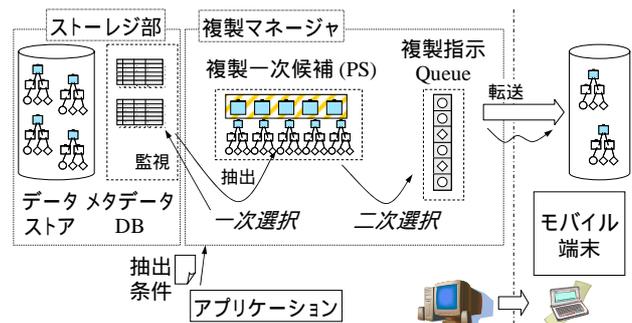


図1 複製対象選択システム

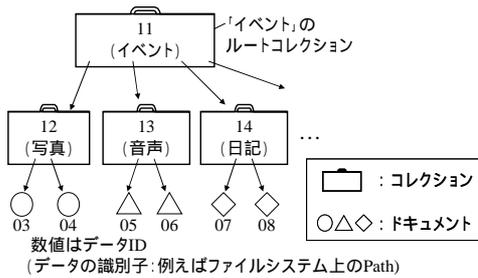


図2 コレクションツリーの構造

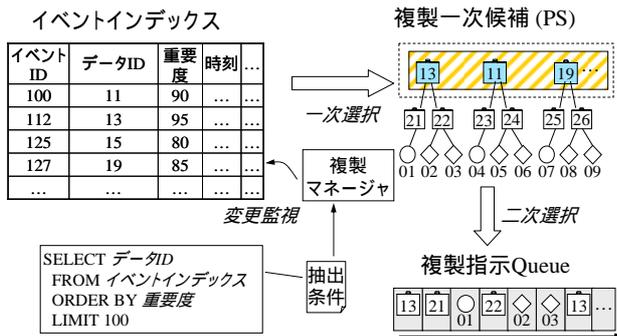


図3 複製マネージャの動作

ションは、例えば「イベント」という単位で関連づけられたデータを図2のように構造化する。この構造の中で、データはコレクションまたはドキュメントに分類される。コレクションはドキュメントあるいは別のコレクションへのリンクを保持する。これにより、蓄積コンテンツを多数のコレクションによって構成される木構造（以下コレクションツリーと呼ぶ）に整理する。

コレクションツリーは一般のファイルシステムにおけるディレクトリ構造と相似だが、任意のデータ（コレクションまたはドキュメント）は複数のコレクションからリンクされ得る点が異なる。さらに、既存のディレクトリがファイルシステム上の単一の構造を与えるために構成されるのに対し、コレクションツリーは各アプリケーションが独自の体系でデータを関連づけるために作られる。このため、同じデータに対して複数のアプリケーションがそれぞれ別の目的でコレクションツリーを構成する事も可能とする。本稿では、あるアプリケーションが蓄積コンテンツを「イベント」という単位でコレクションツリーとして整理した状態を想定しているが、整理の単位は「イベント」のような特定の方法に限らない。

メタデータ DB は、蓄積コンテンツについてアプリケーションが付加する任意のメタデータを管理する RDB システムである。メタデータは例えば「イベント」の日時や重要度、注釈などの情報でありアプリケーションによって変更される。WinFS[3] などの将来の PC ストージでは、このようなメタデータに関する高機能な DB インタフェースを備えることが一般的になると考えられている。

4.2 複製マネージャ部

複製マネージャは、アプリケーションによって指定された抽出条件に基づいてストレージ部に蓄積された膨大な量のコンテンツの中の一部を抽出し、モバイル端末に複製する。抽出条件は、対象とするメタデータ DB の表と選択条件からなり、例えば SQL によって記述される。

図3で示すように、複製選択は二段階に分けて行われる。複製マネージャはまず、与えられた抽出条件をメタデータ DB に対して発行し、得られた抽出結果を「複製一次候補 (Preliminary Selection; 以下 PS とする)」として保持する（一次選択）。また、抽出条件で指定された表の変更を監視し、表に変更があった場合に条件を再適用して PS を更新する。

PS は特定の表から抽出された、コレクションのリストである。典型的には「イベント」のような形で整理されたデータ集合のルートとなるコレクション（例えば図2のコレクション11など）が並べられる。

次に、PS に並ぶコレクションから関連するデータが抽出され、複製指示 Queue と呼ばれるリストに並べられる（二次選択）。関連性の抽出はコレクションツリーのリンクを辿る事によって行われる。Queue はモバイル端末とのネットワーク接続が検出されるたびに複製マネージャによって読み込まれ、個々のデータはリスト順にモバイル端末に複製される。

5 効果

我々の提案する複製候補選択システムは、メタデータ DB の特定の表の変更を契機に PS を更新する。よって大量に蓄積された体験記録データのコンテンツそのものの変更監視を必要としない。また、PS は整理済データのルート要素のみを並べたリストであるため、PS の変更を管理するコストはそれほど大きくない。このため、蓄積データ量へのスケーラビリティを確保できる。

さらに、データストアにコレクションツリーを導入し、木構造を考慮して複製候補を選択する事で、アプリケーションが独自に定義する「イベント」のようなデータ集合を、その構造を保ったままモバイル端末に複製することができる。

6 おわりに

本稿では、将来の個人ストレージが抱える膨大な量の蓄積コンテンツを複数端末間で利用する際に必要となる複製対象選択方式について基本方式を提案した。現在、システムの詳細の設計を進めるとともに、プロトタイプシステムの開発を行っている。

参考文献

- [1] 中村隆幸, 井上知洋, 中村元紀, “コピキタス情報蓄積環境 GreenTown の実装”, DICOMO 2004, pp. 293296
- [1] Nicky Kern, et al., “Context Annotation for a Live Life Recording”, Proc. Pervasive 2004 Workshop on Memory and Sharing of Experiences, April 2004
- [2] <http://www.microsoft.com/japan/msdn/longhorn/introducing/longhornch04.asp>