

# 集成的ライフログを活用したレコメンドシステムの構築 Recommend System using Life Log

茂木 学<sup>†</sup>  
Manabu Motegi

山田 智広<sup>‡</sup>  
Tomohiro Yamada

永徳 真一郎<sup>†</sup>  
Shin-ichiro Eitoku

手塚 博久<sup>†</sup>  
Hirohisa Tezuka

武藤 伸洋<sup>†</sup>  
Shin-yo Muto

阿部 匡伸<sup>†</sup>  
Masanobu Abe

## 1 まえがき

リアル世界およびサイバースペースにおける複数人のライフログを活用してコンシェルジュサービスを実現するシステムを構築した。

近年、携帯電話をはじめとする端末の小型化、および各種センサの搭載が進んでいる。これにより、普段の生活を妨げることなく、GPSや加速度データ等、リアル世界上のライフログ収集が可能になってきた。一方、インターネット内での人の活動に着目すると、Webアクセスログ、検索履歴、ログデータは比較的容易に取得可能である。

図1は、本稿でターゲットとするサービス領域を示している。x軸はライフログのソースを、リアル世界とサイバースペースの2つの方向性で示している。y軸は、ライフログが個人に閉じて利用されているのか、あるいは複数人で協調的に利用されているのかを示している。図1に示すように、Google[1]はサイバースペースから収集した検索語、およびWeb閲覧履歴からなる複数人のライフログを用いたパーソナライズド検索を行っている。同様に、Amazon[2]も購買履歴からなる複数人のライフログを活用した協調フィルタリングを用いてレコメンドーションを実現している。一方、MyLifeBits[3]は、サイバースペースのデータ以外に、ビデオ、音声等、リアル世界のライフログも用いている。そして、これらを主に個人の履歴管理に適用している。また、ロケーションベースドサービスに関しては、ユーザ本人の現在の位置情報に合わせた、様々なサービスが実現化されている。

本稿のターゲットは、AmazonやGoogleに代表されるような、サイバースペースでの複数人のライフログを利用したサービスをリアル世界に拡張することである。そして、リアルとサイバースペースにおける人の活動から、知識抽出や人の関係性抽出を行うことにより、新たなライフログ活用サービスを開拓することである。

## 2 リアルとサイバースペースを融合した複数人のライフログ活用サービス

日常のライフログを基本として、ユーザの行動を予測し、適切なタイミングで適切な情報を提示するコンシェルジュサービスに関して主に検討を行ってきた。具体的

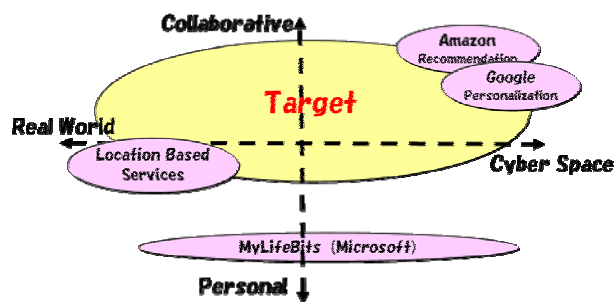


図1 既存のサービス及びターゲット領域

には、商品レコメンドと生活リマインドの2つのサービスをターゲットとしている。

以下では、想定している商品レコメンドサービスの一例に関して説明する。まず、システムはリアル世界におけるユーザのライフログとして日常行動パターンを蓄積する。また、サイバースペースにおける情報として、ユーザのWeb検索および閲覧履歴を蓄積する。そして、ユーザの現在の行動と蓄積した行動パターンとを照合し、そこから次の行動を予測する。その後、予測した行動およびユーザのサイバースペースにおける情報を用いて、適切な商品レコメンドを行う。その際、同様な日常行動パターンを有する他者のサイバースペース上の情報も用いて商品レコメンドを行う。

一方、生活リマインドサービスに関しては、リアル世界のライフログから現在いる滞在地を算出し、その滞在地と時間帯から適切な生活リマインドを行う。例えば、朝、自宅を出ることを検出すると、借りているDVD等を忘れないようにリマインドを行うことが考えられる。

上記2つのサービスに関しては、リアル世界のライフログとして主にGPSデータを用いている。これは、GPSが携帯電話すべてに搭載され、最も利用しやすいセンサデバイスの一つであることから選択した。また、サイバースペースの情報としては、Web検索・閲覧履歴だけでなく、ユーザ自身あるいは家族等他者のスケジュール情報の利用も検討している。これらを含んだ上記サービスの具定例は、第5章で改めて述べる。

## 3 ライフログ活用システムアーキテクチャ

図2にシステムコンセプトを示す。また、図3に全体のシステム構成を示す。ライフログを活用したサービスを提供するためには、図2に示すようにcollection, estimation, extraction, storage&management, search&matching, servicesの各機能が必要である。

<sup>†</sup> 日本電信電話株式会社 NTTサイバースソリューション研究所, NTT Cyber Solutions Laboratories, NTT Corporation

<sup>‡</sup> 日本電信電話株式会社 研究企画部門 NTT Corporation

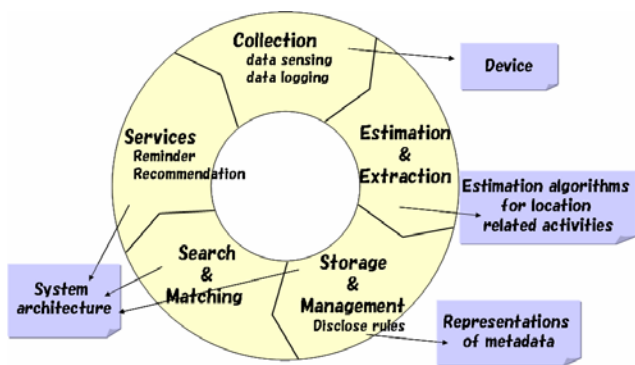


図 2 システムコンセプト

collection 機能は、ユーザの活動状況をセンシングし、収集する機能であり、それを実現するデバイスが必要となる。しかし、collection 機能で収集したセンシングデータは単なる数値データでしかない。そこで、estimation&extraction 機能では、collection 機能により収集したセンシングデータから、より抽象化したライフログデータの抽出および推定を行う。storage&management 機能は、抽出したライフログデータの蓄積および情報の開示、非開示等を管理する機能である。search&matching 機能は、必要な情報の検索およびマッチングにより、ユーザへのサービス提供に必要な情報を抽出する機能である。services 機能は、実際にユーザにサービスを提供する機能である。その後、collection 機能により、システムは提供されたサービスがユーザに効果的であったか否かを判断し、サイクリックにパーソナライズを行い提供するサービスの精度を向上させる。このように、これら機能間をライフログ情報が流通することにより、ユーザに効果的なサービスを提供可能となり、またパーソナライズが促進される。以上のコンセプトを具現化するために、図 3 に示すシステムを構築した。主に個人のユーザのライフログを取得・管理するローカルシステム（ユーザの PC・ホームサーバ等）と、ユーザの許可を得てローカル環境からアップロードされた複数ユーザのライフログを管理するセンタサーバ、ならびにライフログを用いてユーザへサービスを提供するアプリケーションサーバから構成される。また、ローカルシステムから、センタサーバにアップロードするライフログをユーザが制御するための開示支援機能を用意した。これにより、利用者が安心してライフログを提供可能となる。

ローカルシステムへの入力として、様々なセンサからの情報を扱うために、センサからの情報をドライバで FDML フォーマット[4]に変換することとした。このように、ローカルシステムにおける IF を統一することで、新たにセンサあるいは PC を追加する場合でも、ドライバを開発することで容易に拡張可能となる。

収集したデータからは、リアルタイムで個人ライフログ生成エンジンにより特徴量が抽出される。そして、可視化・開示制御を容易にするため 3W（What, Who, Where）に分類された内部流通用の XML フォーマットに変換される。また、個人ライフログとしてローカルデータベースに蓄積される。

ユーザが開示許可と設定したライフログはセンタ側へアップロードされる。つまり、センタ環境には各ユーザが「サービス提供者へ開示を許可したライフログ」のみが複数人分蓄積されることとなる。アップロードされたライフログは、集合ライフログ生成エンジンにより、他ユーザのライフログ等も含めた処理が行われ、該当のアプリケーションへ送信される。センタへは多数のユーザからのライフログが集約され、そこで知識抽出や関係性抽出を行うためのマイニング処理等が実施される。センタでの処理は、多数のユーザのデータを扱うだけではなく、様々なサービスプロバイダからの要件に対応したマイニング処理に対応していく必要がある。さらに、取得されるライフログの種類が増加することにより、扱う情報種別の拡張も必要となってくる。本稿では、スケールアウトの発想に基づき、規模成長と機能成長を実現したデータベースおよび分散処理との連携を行っている。これにより、サービスを停止することなく、柔軟かつ容易にマシンを追加することを可能としている。詳細は、引用文献[5]を参考していただきたい。

サービス利用にあたっては、ユーザはサービス登録などによりサービス提供者からサービス提供に必要なライフログの種類等の情報を事前に取得しておく。

サービス利用時にはローカルシステムで蓄積されているライフログの中から、該当のサービスで要求されている種類のライフログに対して開示/非開示判断を行う。なお、サービス提供者が要求しない種類のライフログについては、該当のサービス提供者に対しては全て「非開示」として扱う。

取得したライフログは、ローカル環境に一旦全て蓄積され、センタ環境側へは、ユーザがサービス提供者毎に開示/非開示の判断を行った上で、開示許可を行った情報だけが送信される。つまり、各ユーザのライフログ管理は基本的にユーザ自らが意識的に行うこととなる。これによりユーザは「自らのデータを手元でコントロールできる」という安心感が得られると考えられる。

ライフログ種別が拡大することにも対応した可視化および開示支援に対応するため、3W によるライフログの分類とそれらの情報の時間軸上で連携させるユーザ IF を提供する。

以降の章では、複数人のライフログを利用する場合に重要となる可視化・開示支援について基本的な考え方の説明を行う。

#### 4 可視化・開示支援の実現方法

ライフログの可視化では、多くのライフログが『時間』ないしは『空間』に紐付けられることから時間軸上での可視化や地図上へデータをマッピングする可視化方式をとっているものが多い。しかしながら、これら研究の可視化は自らの過去の振り返りの観点から行なわれているものがほとんどであり、ライフログを公開する場合を考慮していないようである。

ライフログ開示/非開示判断サポートに注目した可視化を考えた場合、以下の点が必要である。

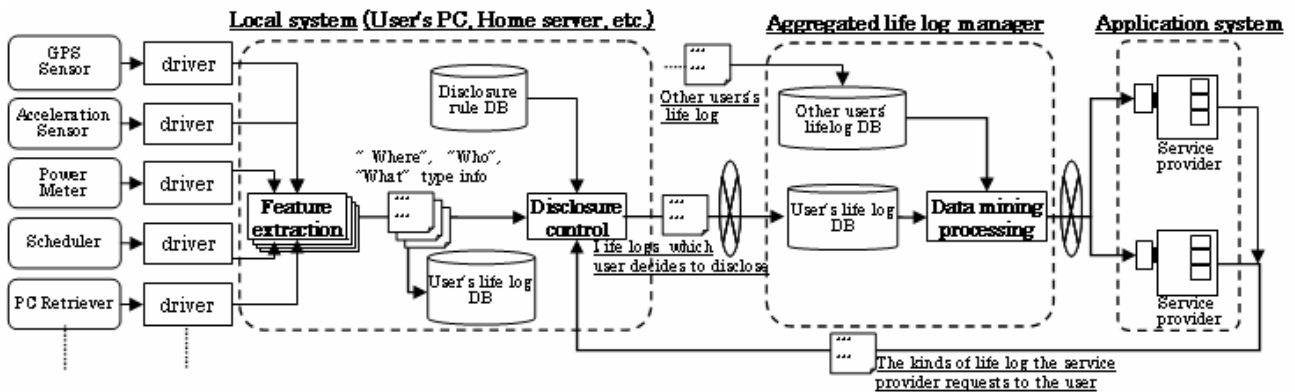


図3 ライフログ活用システム構成

- ・複数のサービス提供者へのライフログ提供
- ・複数種類のライフログの蓄積・利用
- ・時間経過に合わせたライフログの増大

上記の観点から、下記3機能の検討・実装を行なった。

- (1) サービス提供者毎のライフログ表示切替機能
- (2) 組み合わせを考慮した可視化機能
- (3) 様々な軸での開示/非開示の判断可能なサポート機能

上記(1)に関しては、サービス提供者ごとにユーザーが提供すべきライフログが異なると考えられる。そのため、サービスごとにどのライフログを開示/非開示するかをユーザーが決定可能かつくりとする必要がある。上記(2)に関しては、例えば、ユーザーが「ある場所にいた」という情報だけは開示してもかまわないが、「〇〇さんと一緒に、ある場所にいた」となると、その場所情報を非開示

にしたいという要望が考えられる。このように、提供するライフログの組み合わせによってもユーザーの開示/非開示の判断が変わることが考えられる。したがってライフログの開示/非開示を選択する上では、表示すべき複数種類のライフログ間の組み合わせを考慮した表示を行う必要がある。上記(3)に関しては、ユーザーが蓄積しているライフログ全種類の中から、3W (What, Who, Where) 毎にライフログの表示・検索を行い、サービス提供者から提供されるライフログだけを可視化する必要がある。

## 5 プロトタイプ構築

### 5.1 プロトタイプ概観

図5に商品レコメンドの処理フローを、図6に実際に携帯電話で受信したレコメンドメール例を示す。実際のアーキテクチャの実現性を検証するため、図3に示したプロトタイプシステムの実装を行った。センサはGPSを搭載した携帯電話を想定している。加えてWeb閲覧履歴を取得するためのアプリケーション[6]、スケジューラ、携帯電話のピックアップを検出する電力計を接続した。

個人ライフログ生成エンジンとしては、滞在抽出、移動手段判定、外出・帰宅検出等を実装し、入力されるデータに対してリアルタイムで処理結果を出力可能としている。また、開示制御を行うために、図4に示すユーザインタフェースを開発した。

第4章で記載した機能要件を満たすため、画面の右側にサービス毎のタブを用意し、サービス毎に必要なライフログをユーザーが確認可能な構成とした。また、ユーザーが利用するサービスを追加すると、それに合わせてサービスタブが追加されるように設計している。組み合わせを考慮した可視化を実現するために左側の閲覧画面にあるように、What情報をキーにして、関連するWho、Where情報をリンクさせて表示可能としている。また、様々な単位での開示制御が行えるように、開示したくない情報をWhat、Who、Whereの情報をキーにして検索し、個別に開示制御可能としている。

集合ライフログ生成エンジンとしては、行動パターン抽出[7]、同行判定[8]、行動パターンが類似の人の興味語を検出する協調フィルタリング、手すきタイミング検出等を実装している。行動パターンの抽出には、シーケンシャルパターンマイニングを用いている。複数利用者の特徴行動

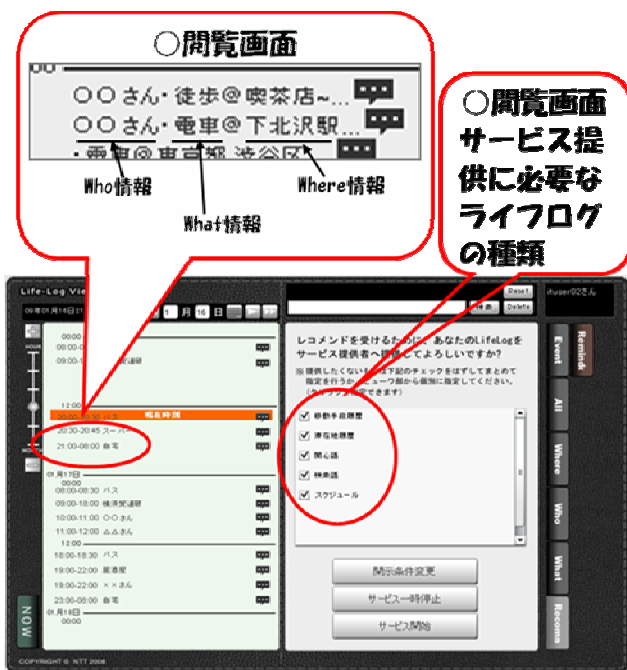


図4 開示制御インタフェース



パターンを抽出することで、行動パターンおよびそれに含まれる頻出の滞在情報から、ユーザ属性の類似性判定を行っている。同行判定では、GPS データに含まれる誤差情報を考慮した判別関数を定義し、SVM により最適な閾値を学習させることで、同行者を自動的に抽出可能としている。

アプリケーションとしては、商品レコメンド (2 種類)、生活リマインド (2 種類) を実装した。図 5 に商品レコメンドの処理フロー例を示す。商品レコメンドにおいては、まず、Web 閲覧履歴から推定したユーザの関心語、および滞在地系列より算出した行動パターンを用いて行動予測を行い、予想滞在地を算出する。次に、店舗情報・商品情報を検索し、同行判定、移動手段判定の結果を用いた手すき判定を行い、手すきのタイミング (例えば、一人で電車に乗った時) にレコメンドメールが送出される。そして、レコメンドメールは図 6 に示すように、ユーザの携帯電話で受信される。ここで、手すき検出を行う理由は、ユーザが手すき状態であれば、レコメンドメールを読みやすい環境にいる可能性が高いと考えたためである。

## 5.2 プロトタイプの評価

システム評価をより実際の使用に近い形態で行うために、以下のようなシーンを想定して検証を行った。あるビジネスマンの 1 日の行動を想定しており、以下の 4 つのシーンを用意した。

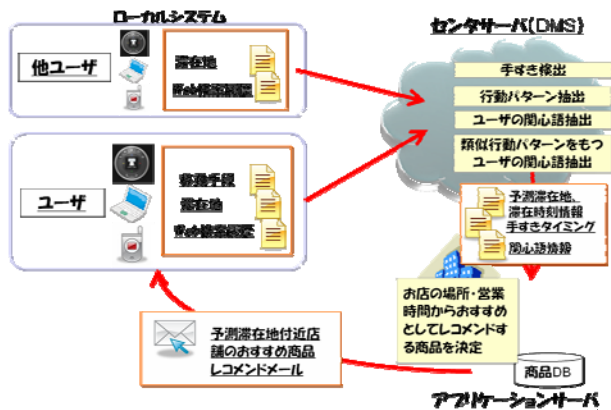


図 5 商品レコメンドの処理フロー例



図 6 受信レコメンドメール例

ここで、ユーザは、DVD をレンタルショップから借りており、返却日を翌日に控えていること、また、家族は妻と娘であり、妻は旅行中であるという想定である。

### ○シーン 1.

ユーザが出勤時に携帯電話をピックアップする

- ・システムは時間帯から出勤と判断して、DVD 返却日が近いことを携帯電話にメールでリマインド

### ○シーン 2.

ユーザが普段通勤している電車に乗車

- ・システムはユーザが電車に乗ったことを検知
- ・ユーザが 1 人であり、同行者がいないことを検知し、ユーザが手すき状態でありレコメンドメールを読む可能性が高いと判断
- ・行動パターンを、時間情報および滞在地系列情報として常に蓄積
- ・過去に蓄積した通勤ルートから目的地を判断
- ・最近のユーザの検索履歴から、興味があるキーワードを抽出
- ・同様な行動パターンの他ユーザが興味を持っているキーワードに関するレコメンド情報を協調フィルタリングにより抽出
- ・ユーザの興味があるキーワードに関する目的地の情報、および他ユーザが興味を持っているキーワードに関する情報を携帯電話にメールでレコメンド

### ○シーン 3.

ユーザが会社で、突然の出張によりスケジュール変更

- ・システムはスケジュール変更により、ユーザが DVD 返却不可能なことを検知
- ・ユーザの妻と娘のスケジュールを検索し、妻は旅行中だが、娘は DVD 返却可能であると判断
- ・ユーザの娘の携帯電話に代理で DVD 返却の依頼をメールすると共に、ユーザの携帯電話に娘に代理依頼をしたことをメール通知

### ○シーン 4.

ユーザが通常行動とは異なり、出張するために電車に乗車

- ・システムは過去に蓄積した時間情報および滞在地系列情報と比較して、目的地を検出しようとするが、この場合は検出不可能なので、スケジュールから目的地を推定
- ・最近のユーザの検索履歴から、興味があるキーワードを抽出
- ・目的地によく行く他ユーザが興味を持っているキーワードに関するレコメンド情報を協調フィルタリングにより抽出
- ・ユーザの興味があるキーワードに関する目的地の情報、および他ユーザが興味を持っているキーワードに関する情報を携帯電話にメールでレコメンド

上記シーン 1, 3 は生活リマインドの例であり, 上記シーン 2, 4 は商品レコメンドの例である. 開発したシステムに対して, 5 人分のライフログを最大 1 ヶ月間蓄積し, システム動作の検証を行った. また本システムを NTT R&D フォーラム 2009 に展示し, 実時間での動作がほぼ安定して行えることを確認した. なお, 今回は屋内でのデモ展示のため, あらかじめ取得しておいた GPS データを 1 秒ごとにローカルサーバに送信するもので, シーンごとに対応する GPS データを送信するつくりとした. これにより, 擬似的にリアルタイム動作を行った. 2 日半のデモ展示を行い, 適切なタイミングで携帯電話にメールが送信されない事象が数回ほどあったが, 全体的には安定動作した. メールが送信されなかった原因は, ネットワークによる遅延と考えられる. デモ展示は盛況であり, 見学した人々からはプライバシーを心配する声が聞かれた一方, 大容量ネットワークを生かすアプリケーションになりうると期待する声も多かった.

簡易的なシステムパフォーマンス評価として, 上記 4 つのシーンごとに以下の時間を各 5 回ずつ計測した.

- ・シーン 1:  
ユーザが携帯をピックアップした時刻と, 携帯電話でリマインドメールを受信した時刻の差
- ・シーン 2:  
ユーザが電車に乗った時刻と, 携帯電話でレコメンドメールを受信した時刻の差
- ・シーン 3:  
スケジュール変更した時刻と, ユーザと娘の各携帯電話でメールを受信した時刻の差
- ・シーン 4:  
ユーザが電車に乗った時刻と, 携帯電話でレコメンドメールを受信した時刻の差

上記計測時間の平均値を表 1 に示す. これより, シーン 1 およびシーン 3 では, ユーザがアクションを起こしてから 30 秒程度でリマインドメールが携帯電話に届く. これは, 例えばシーン 1 では, 一般的に携帯電話をピックアップしてから家を出るまでに 30 秒程度はかかると思われるので, 実用上問題ない処理時間と考えられる.

一方, シーン 2 およびシーン 4 では, ユーザが電車に乗ってから, レコメンドメールが来るまで 60 秒程度かかっている. これは, 移動手段判定を行うのに, 事前に一定量蓄積された GPS データが必要であるためである (現状は 30 秒分の GPS データが必要). このように, 電車に乗ってから 60 秒程度後にレコメンドされたほうが, ユーザがレコメンドメールを見る余裕があると思われる.

## 7. まとめ・今後の展開

本稿では, ライフログ活用システムのローカルサーバ, センタサーバ, アプリケーションサーバから構成されるアーキテクチャを提案した. そして, 提案するアーキテクチャに基づき, リアル世界とサイバー空間のライフログを活用可能なシステムを構築した. 4 つの利用シーンを

表 1 システムターンアラウンドタイム

	Scene1	Scene2	Scene3		Scene4
			Father	Daughter	
Average period[s]	30	65	25	21	61

設定し, 各シーンにおいて簡易的な評価を行い有効性を確認した. 今後は, さらなるシステム評価と, 実サービスにむけたシステムブラッシュアップを行う予定である.

## 文献

- [1] Google, <http://www.google.com>
- [2] Amazon, <http://www.amazon.com>
- [3] J. Gemmell, et al, " MyLifeBits: A Personal Database for Everything", Microsoft Research Technical Report MSR-TR-2006-23, (2006).
- [4] 製造業 XML 推進協議会, <http://www.mstc.or.jp/mfgx/>
- [5] 赤間他, “追記・参照型データ管理システムの設計と評価”, 情処学会論文誌, Vol.49, No2, (2008).
- [6] 森田他, “Memory-Retriever:体験獲得情報を想起させる行動検索手法”, 情報処理学会論文誌, Vol.48 No.3, pp.1197-1208, (2007).
- [7] M. Nishino, et al, “A Place Prediction Algorithm Based on Time-Sensitive Frequent Patterns”, Proc. of Pervasive 2009
- [8] S. Seko, et al, “An Algorithm to Estimate the Level of Friendship Based on the Mode of Transportation and the Time Spent Sharing Movement Tracks”, Proc. Pervasive 2009