

D-030

オントロジーに基づく LBSN 上でのイベント検出

稲葉 鉄平[†] 高橋 正和[†] 築井 美咲[†] 石川 佳治[‡][†]名古屋大学情報科学研究科 [‡]国立情報学研究所

1 はじめに

スマートフォンなどの GPS 機能を持つデバイスの普及により, FourSquare や Facebook のチェックイン機能など, 位置情報を利用したソーシャルネットワークサービス (SNS) が流行している. そのような位置情報を利用した SNS のことを位置に基づくソーシャルネットワーク (Location Based Social Networks, LBSN) [3] という.

本研究は, さまざまな LBSN を構築するための基盤として利用可能な汎用的な LBSN ツールキットの開発を大きな目的としている. LBSN に共通する要素を整理し, 共通基盤となる LBSN システムのシステムアーキテクチャを開発する. また, 特定目的の LBSN を構築するためのシステム拡張機構についての開発も行う.

LBSN システムの実現において重要となる技術の一つとして, 複合イベント処理 (complex event processing, CEP) [4] がある. これは, データストリームからより高次のイベントを検出しようとする処理のことを指す. LBSN における多くのサービスがイベントの組合せに対する処理で実現されることから, 複合イベント処理は LBSN の実現のために必要不可欠な技術である. そこで, 本研究で開発する LBSN ツールキットは, 複合イベント処理機構を基本的な構成要素として有するものとする.

LBSN における複合イベント処理では, アプリケーション (例: ユーザ対応型の観光案内, モバイル環境での友人同士の出会い支援) に応じて発生するイベントやその意味が異なる. この問題に対応するため, 本研究ではセマンティック Web 技術 [1] を活用する. 対象となるアプリケーションにおけるイベントが RDF データとして表現されることを想定し, イベント検出においては RDF に対するマッチングを行う. これにより, 高レベルの「意味的な」複合イベントの検出を図る.

本フレームワークのもう一つの特徴となるのが, Linked Open Data (LOD) [7, 8] の積極的な活用である. LBSN における高レベルなサービスの実現のためには, 知識情報や周辺環境の情報をネットを介して取得し利用することが必要となる. 本フレームワークでは, コアシステムに LOD を利活用するための機構を実現し, LOD と連携した高レベルの LBSN サービスの実現を可能とする.

以下では, まず, 本研究が目指す汎用の LBSN ツールキットについてその概要を述べる. 次に, 高レベルの複合イベント処理のためのオントロジー記述, そして, それを用いた複合イベント検出について, 基本的なアイデアについて述べる.

2 システムのアーキテクチャ

図 1 に, 構想しているシステムのアーキテクチャを示す. 各構成要素について説明する. まず, 右上の大きい四角が, 対象ドメインに特化して構築された LBSN システムを表している.

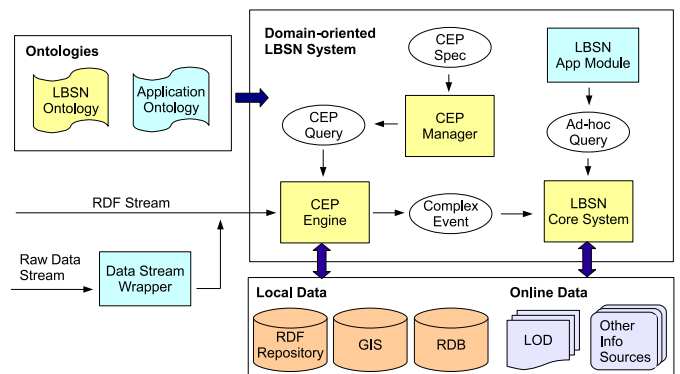


図 1 システムのアーキテクチャ

LBSN コアシステムは, LBSN システムの核となる機能を提供するもので, 各種データへのアクセスと管理, および, 与えられた問合せの処理機能や, LBSN に関するデータの汎用的な分析機能を実現する. アクセス対象となるデータは, ローカルに存在する RDF ストア, GIS, リレーショナルデータベースだけでなく, 外部の Linked Open Data やその他の情報源にも存在する.

CEP エンジン, RDF データストリームを入力とし, 複合イベントの検出を担当する. なお, センサデータなど, RDF データ以外の入力ストリームについては, アプリケーションに応じてデータストリームラッパを構築し, RDF データストリームへと変換する. CEP エンジンにより検出された複合イベントは, LBSN コアシステムに送られ問合せに使用される. 複合イベント処理は, システムに CEP 問合せを登録することで実現される.

CEP 問合せは, システムに対して複合イベントの記述を与えることで生成される. それを担当するのが CEP マネージャである. CEP マネージャはシステムに与えられた複数の複合イベント記述の分析・管理を担当する.

LBSN アプリケーションモジュールは, アプリケーション固有の処理を実現するモジュールであり, 本ツールキットが提供する API を用いて, アプリケーション個別に実装される.

本研究の特徴となるのが, オントロジーの活用である. オントロジーには, LBSN における全般的な概念をカバーする LBSN オントロジーと, 個々のアプリケーションについて対象ドメインの概念を導入するためのアプリケーションオントロジーが存在する.

次節では, 特に複合イベント処理について焦点を当てて説明する.

3 複合イベントの記述と処理

ここでは, 本研究において重要なポイントの一つとなる複合イベント処理の考え方について述べる.

3.1 イベントオントロジーの活用

一般的に LBSN を実現する上で, 共通して現れるイベントがある. たとえば, 「あるオブジェクト o が時刻 t に座標 (x, y) にいた」というイベントは非常に一般的なものである. このような共通なイベントについては, 上述の LBSN オントロジーにおいて,

Event Detection in LBSNs Based on Ontologies

Teppei Inaba[†], Masakazu Takahashi[†], Misaki Yanai[†], Yoshiharu Ishikawa[‡][†]Graduate School of Information Science, Nagoya University[‡]National Institute of Informatics

イベントの知識体系としてモデリングして提供する。

一方で、特定の対象ドメインに固有のイベントも考えられる。たとえば、災害避難に関する LBSN を考えた場合、災害発生に関するイベントや、救助活動に関するイベントなどが考えられる。このようなイベントの共通化は困難であるため、アプリケーション固有のイベントオントロジとして体系化する。対象ドメイン固有の LBSN システムの構築時には、LBSN オントロジとアプリケーションオントロジのイベントの概念を統合して用いることになる。

本ツールキットでは、イベントオントロジによって定義されたイベントを組み合わせ、複合イベントを記述するためのフレームワークを提供する。つまり、あるドメインに対して LBSN システムを構築する際には、ドメイン固有の個々のイベントの概念をオントロジとしてモデル化し、それらの組合せについてはツール提供の機能を用いるものとする。オントロジに関する機能については、今後の研究で詳細化を図る。

3.2 CEP 問合せとその記述

前述のイベントオントロジを用いて記述した、検出対象の複合イベント記述 (CEP description) が与えられると、CEP マネージャはそれをコンパイルして複合イベント問合せ (CEP query) に変換し、連続的問合せ (continual query) としてシステムに登録し実行する。CEP 問合せの記述には、RDF データに対するマッチングを実現するため、SPARQL [9] をベースとした問合せ言語を想定する。ただし、本研究においては、複合イベント記述からコンパイルされる形での問合せの生成を考えているため、SPARQL 言語の機能全体を用いるわけではなく、そのサブセットを想定する。

一方、問合せ言語には RDF データストリームを処理する必要があるため、本研究では C-SPARQL [2] に見られるような連続的問合せの機能を導入する。また、空間データの処理機能も必要となるため、Geo-Store [5] に見られるような空間データおよび空間的問合せ条件の機能も導入する。問合せ言語の仕様についても、今後の研究において明確化していく。

3.3 CEP 問合せの例

LBSN における複合イベント処理の例として、商店のタイムセール情報を考える。食料品店、衣料品店などのさまざまな商店が、安売りなどのタイムセール情報が、あらかじめ決められた RDF スキーマに従って LBSN システムに定期的に配信される。たとえば、以下の図はパン屋のタイムセール情報を示している。店舗名称、店舗の位置情報、タイムセールの情報が含まれている。

```
@prefix ts_event: <http://lbsn.org/ts_event#>.
@prefix ts_ont: <http://lbsn.org/ts_event_ont#>.
ts_event: ev_1 ts_ont:name "Motoyama Bakery" .
               ts_ont:tel "123-4567" .
               rdf:type ts_ont:Bakery .
               ts_ont:coord "(35.163, 136.962)" .
               ts_ont:desc "20% off from 2pm to 3pm!" .
```

図 2 タイムセール情報の例

LBSN システムには、図 2 に示したようなタイムセール情報がイベントとして継続的に送られるとする。また、個々の移動ユーザの位置情報についても、定期的にイベントストリームとしてシステムに送信されているとする。なお、両者のストリームは異なるストリームとして送信されるとする。

一方、タイムセール情報に興味を持っている個々の移動ユーザは、提供されるユーザインタフェースにおいて興味ある商品カテ

ゴリと距離の閾値の情報を記述する。その記述に基づいて、図 3 のような問合せを生成する。基本的には SPARQL 問合せであるが、C-SPARQL 同様、ストリームデータを対象とするための FROM STREAM 句を用いている。RANGE 30min は、過去 30 分の配信データを対象とする時間窓 (window) の指定である。一方、FILTER 句では、ユーザの位置から 1km 以内の食料品店を探している。これは Geo-Store で見られる機能である。

```
SELECT ?name, ?coord, ?desc
FROM STREAM <http://lbsn.org/time_sale.trdf> [RANGE 30min]
FROM STREAM <http://lbsn.org/user_mvmt.trdf> [RANGE 1min]
WHERE { ?shop ts_ont:name ?name .
        ?shop rdf:type ts_ont:Food_Shop .
        ?shop ts_ont:desc ?desc .
        ?shop gml:pos ?coord .
        ?shop ts_ont:time ?time .
        FILTER(glbsn:range(?coord, CUR_LOC, 1, "km"))
```

図 3 連続的問合せの例

RDF データストリームの処理は、C-SPARQL のような連続的問合せ言語だけでなく、RDF データに対する Publish/Subscribe システムに関する研究 [10] (ただし SPARQL は使用されていない) も関連する。ただし、本研究では、複合イベントの記述から CEP 問合せを生成しようという立場であり、考え方は大きく異なる。

4 まとめと今後の課題

本稿では、対象ドメインに特化した LBSN を構築するための LBSN ツールキットについて、そのアイデアを述べた。セマンティック Web 技術の活用と複合イベント処理がそのキーとなる技術である。近年着目されているジオストリーミングでは、たとえば [6] などのツールキットの提案があるが、LBSN を対象とする点が本研究の特徴である。

謝辞

本研究の一部は科学研究費 (22300034, 2580039) による。

参考文献

- [1] 荒木. フリーソフトで学ぶセマンティック Web とインタラクション. 森北出版, 2010.
- [2] D. F. Barbieri, D. Braga, S. Ceri, E. D. Valle, and M. Grossniklaus. Querying RDF streams with C-SPARQL. *ACM SIGMOD Record*, 39(1):20–26, 2010.
- [3] C.-Y. Chow, J. Bao, and M. F. Mokbel. Towards location-based social networking services. In *2nd Intl. Workshop on Location Based Social Networks (LBSN'10)*, pp. 31–38, 2010.
- [4] G. Cugola and A. Margara. Processing flows of information: From data stream to complex event processing. *ACM Computing Surveys*, 44(3), 2012.
- [5] W.-S. Ku, H. Chen, C.-J. Wang, and C.-M. Liu. Geo-Store: A framework for supporting semantics-enabled location-based services. *IEEE Internet Computing*, 17(2):35–43, 2013.
- [6] J. Lee, Y. Liu, and L. Yu. SGST: An open source semantic geostreaming toolkit. In *2nd ACM SIGSPATIAL Intl. Workshop on GeoStreaming (IWGS'11)*, pp. 17–20, 2011.
- [7] 長野, 萩野 (編). 特集: リンクするデータ (Linked Data) - 広がりを始めたデータのクラウド -. 情報処理, 52(3):282–333, 2011.
- [8] 長野, 山口 (編). 特集: Linked Data とオントロジー. 人工知能学会誌, 27(2):162–206, 2012.
- [9] SPARQL 1.1 overview. <http://www.w3.org/TR/sparql11-overview/>, Mar. 2013.
- [10] J. Wang, B. Jin, J. Li, and D. Shao. A semantic-aware publish/subscribe system with RDF patterns. In *COMPSAC*, pp. 141–147, 2004.