

# 系列マイニングによるスポットパス抽出に関する研究

渡邊 直人<sup>†</sup> 佐藤 哲司<sup>††</sup>

<sup>†</sup> 筑波大学大学院 図書館情報メディア研究科 〒305-8550 茨城県つくば市春日 1-2

<sup>††</sup> 筑波大学 図書館情報メディア系 〒305-8550 茨城県つくば市春日 1-2

E-mail: †{wnaoto,satoh}@slis.tsukuba.ac.jp

**あらまし** 近年、ユーザが自身の訪れたスポットを他のユーザに知らせるチェックイン機能が様々な SNS で提供されている。ユーザは駅や飲食店など、様々なスポットでチェックインをしている。ユーザがどのようなスポットからスポットへと移動をしているのかを明らかにすることで、企業のマーケティングやユーザ行動のナビゲーションへの応用が可能となる。本論文では、系列マイニングを用いることで、チェックインデータの中から頻出するスポットパスを抽出した。その中で、スポットのカテゴリに応じた特徴が見られたので報告する。

**キーワード** スポットパス, ロケーションベース SNS, 系列マイニング, ユーザ行動

## A Study on Extracting the Spotpath by Sequential Mining

Naoto WATANABE<sup>†</sup> and Tetsuji SATOH<sup>††</sup>

<sup>†</sup> Graduate School of Library, Information and Media Studies, University of Tsukuba

1-2 Kasuga, Tsukuba, Ibaraki, 3058550 Japan

<sup>††</sup> Faculty of Library, Information and Media Science, University of Tsukuba

1-2 Kasuga, Tsukuba, Ibaraki, 3058550 Japan

E-mail: †{wnaoto,satoh}@slis.tsukuba.ac.jp

**Abstract** Latest, Social Networking Services have a function called check-in that users record a spot visited. Users check-in at variety spot. For example, a station and a restaurant on weekdays, a shopping mall and a electronics retail store on holidays. To know the behavior of users is useful for a marketing plan on business or a navigation for user behavior. In this paper, we extracted the SpotPath appeared frequently among check-in data. And we report the special SpotPath depending on category and usage of user.

**Key words** SpotPath, Location-Based Social Network Services, sequential mining, user behavior

### 1. 背景

近年、ユーザが自身の訪れたスポットに対してチェックインをする、ロケーションベース SNS (Location-Based Social Network Services, 以下 LBSNS) が普及してきている。LBSNS の代表的なサービスの 1 つである Foursquare<sup>(注1)</sup> は、2009 年に米国でリリースされ、2012 年 4 月現在で、ユーザ数は全世界で 2,500 万人を突破し、チェックイン数は累計 25 億回を超えている<sup>(注2)</sup>。日本での利用も盛んであり、東京でのチェックイン数が米国ニューヨークよりも多いという調査もある<sup>(注3)</sup>。

ユーザの利用目的は、チェックインをすることで得られる

サービス内での称号獲得、訪れたスポットを記録するライフログとしての利用、友人との情報共有など様々である。また、Foursquare 上でのチェックインは Twitter<sup>(注4)</sup> や Facebook<sup>(注5)</sup>、Mixi<sup>(注6)</sup> など既存の SNS サービスとの連携が可能なことも、Foursquare の利用を促進している要因と考えられる。

チェックインデータは、店舗名や所在地 (緯度・経度情報)、店舗の業種・業態を表すカテゴリ情報など、スポットに関する情報だけでなく、チェックインをしたユーザに関する情報も含まれており、Foursquare のサーバ上に蓄積されている。チェックインデータから観測される、誰がいつ、どこにいるのか、次にどこへ行くのかという情報は、ユーザ行動をモデリングする上で

(注1) : Foursquare, <https://ja.foursquare.com/>

(注2) : <http://markezine.jp/article/detail/16431>

(注3) : <http://scobleizer.com/2010/05/16/the-king-and-queen-of-location-based-services/>

(注4) : Twitter, <https://twitter.com/>

(注5) : Facebook, <https://www.facebook.com/>

(注6) : Mixi, <http://mixi.jp/>

有用である。例えば、雑貨店を訪れたユーザは、次にカフェを訪れる傾向があるということが分かれば、雑貨店にカフェの割引クーポンを置くなどのマーケティングが行える。また、ユーザの移動軌跡を利用した行動ナビゲーションに関する研究も盛んである。

本研究では、LBSNS のチェックインデータを対象として、系列マイニングの手法を適用し、頻出する移動軌跡をスポットパスとして抽出する。これによりユーザがどのようなスポットを訪れ、次にどこへ向かうのか、その特徴を明らかにすることを目的とする。LBSNS の代表的なサービスである Foursquare から実際のチェックインデータを収集し、24 時間を単位としたチェックインのシーケンスをユーザ毎に作成する。得られた個人毎のシーケンスに対し、系列マイニングのアルゴリズムである PrefixSpan を用いることで典型的なスポットパスを抽出する実験を行う。

## 2. 関連研究

本研究に関連する、スポット推薦に関する研究、系列マイニングを利用した研究について概観することで、本研究の位置付けを明確にする。

篠田ら [1] は、ユーザが訪れたスポットと、そこに滞在している時間に着目し、これらを行動特性の素性として、類似するユーザの情報を参照することで、ユーザが興味を示すと思われるスポットを推薦する方法を提案している。zhijun ら [2] は、写真共有サービス Flickr<sup>(注7)</sup> に投稿される写真に付与されたジオタグ (位置情報) を利用し、世界の 12 都市での旅行ルートの作成を試みている。Choudhury ら [3] も Flickr の写真のジオタグを利用し、ユーザ毎のアップロードした写真の撮影時間の間隔から、2 地点のルートの長さを定義して、頻出する旅行ルートの抽出を試みている。一般ユーザによる手作業で作成されたルートと、提案手法により自動抽出されたルートの比較を行い、スポット推薦をするための良好な結果が得られたと報告している。Kurashima ら [4] は、トピックモデルと隠れマルコフモデルを適用したスポット推薦方法を提案し、Flickr のデータを用いて提案法の有効性を検証している。cho ら [5] は、ユーザの行動モデルを明らかにすることを目的とし、LBSNS のデータ分析をしている。その結果、ユーザの行動モデルには個人毎の行動パターンの他、友人関係も影響を及ぼしていることを報告している。

旭ら [5] は、ブログから様々な人間の行動連鎖を系列マイニングによって抽出し、ユーザに抽出したシーケンスを提示する検索システムを作成している。例えば、結婚式に関するクエリを入力すると、結婚式での流れを理解できるとともに、結婚式に関係するクエリが表示される仕組みになっている。郡ら [6] は、ブログ記事からスポット名を抽出してシーケンスを生成し、スポット名とそのルート、スポット名に関連するキーワードをマップ上に表示する。これは旅行プラン作成補助を目指したシステムとなっている。箴島ら [7] は、商品レビューを評価するた

めに、レビュー文章を分単位に分割し、文単位での評価の並びをシーケンスとして、商品レビューに付与されている評価値とシーケンスとの関係を調査している。山田ら [9] は、プログラミングのコードに対して系列マイニングを適用している。オブジェクト、メソッド、引数などをシーケンスとしてマイニングすることで、プログラムの欠陥を発見する手法を提案している。

本研究では、代表的な LBSNS である Foursquare の実データに対して、系列マイニングを適用することで、多くのユーザが訪れる共通のスポットパスを抽出し、行動モデルの分析を試みる。LBSNS は、写真共有サイト Flickr に写真をアップすることと比較して簡単に利用できることから、ユーザは日々の生活での利用が多いと考えられる。そのため、抽出されるスポットパスも日々の生活を直接的に反映しているものと思われる。

## 3. 系列マイニングを用いた分析法

ユーザが日々の生活の中でスポットを訪れるにあたり、スポットを訪問する順序も重要な要素であると考えられる。例えば、居酒屋でお酒を飲んだ後に締めラーメンを食べに行く、ショッピングモールで買い物をして歩き疲れたらカフェで休憩するなどが挙げられる。そこで本研究では、スポットの訪れる順序を考慮してシーケンスを抽出することができる系列マイニング (sequential mining) を行う。本章では、系列マイニングの定義と具体的なアルゴリズムである PrefixSpan について説明する。

### 3.1 系列マイニングの定義

系列マイニングとは、大量の系列データ (sequence data) を蓄積した系列データベース (sequence database) の中から、系列情報を保持したまま頻出する部分シーケンス (sequential pattern) を抽出する手法である。

系列データベースに格納される系列データは、一般に表 1 のような構造を持つ。アルファベットで表記されているアイテムは、本研究のチェックインを表しており、シーケンス中の何番目に出現するかを表す、識別子 (TID) が付与されている。これらアイテムで構成されたシーケンスにも、それぞれ識別子 (SID) が付与されている。

シーケンスの長さは、構成するアイテムの個数によって定義され、シーケンス  $\alpha$  が  $k$  個のアイテムで構成されているとき、シーケンス  $\alpha$  の長さは  $k$  とする。また、同一シーケンス、すなわち、同一のアイテムが同じ順序で出現するシーケンスの出現回数は支持度と定義される。系列マイニングの処理開始時に、抽出するシーケンスの最小支持度 (minimum support) を設定し、系列データベース内で長さ  $k$  のシーケンス  $\alpha$  の支持度を調べ、最小支持度を満たすシーケンスを抽出する。

表 1 系列データベースの例

SID	TID		
	t1	t2	t3
s1	A	C	D
s2	A	B	C
s3	A	A	B

(注7) : Flickr, <http://www.flickr.com/>

表 2 長さ 1 のシーケンスの支持度

A : 3
B : 2
C : 2
D : 1

表 3 アイテム A に対する Prefix database

射影元の SID	部分シーケンス
s1	C D
s2	B C
s3	A B
s3	B

### 3.2 Prexspan の特徴

系列データベースから、与えられた長さ（支持度）以上の部分シーケンスを抽出するためのアルゴリズムを説明する。本研究では、pei [10] らによって提案された PrefixSpan アルゴリズムを用いて部分シーケンスを抽出する。系列データベースに対して Prefix projection と呼ばれる射影を行い、それによって生成される Prefix database を探索することで、候補となるシーケンスを生成することなく高速な処理を実現している。Prefix projection では、射影元のシーケンスから射影対象より後ろに存在するシーケンスのみを抽出する。

具体的な PrefixSpan の処理は、初めに長さ 1 の頻出シーケンスを抽出し、その後、深さ優先探索で Prefix projection と Prefix database のマイニングが繰り返される。ここで、表 1 に示した系列データベースに対して、PrefixSpan を適用する際の処理手順を示す。抽出する条件は長さが 2 以上の部分シーケンスで、最小支持度を 2 とする。まず、各アイテムを長さ 1 のシーケンスとして抽出し、その支持度を数える。表 2 は、その結果であり、アイテム A、アイテム B、アイテム C はそれぞれ支持度が 2 以上であるので適当である。アイテム D は支持度が 1 のため不適となり、条件を満たすアイテム D から始まる部分シーケンスは存在しないことがわかる。次に、射影対象をアイテム A とした場合の処理は、各 SID に対して、表 3 に示すアイテム A に後属する部分シーケンスの抽出である。これに対して、各アイテムの支持度は表 4 であり、この結果から、支持度が 2 以上であるアイテム B、アイテム C は、アイテム A とつなげたシーケンス AB、シーケンス AC として抽出されることになる。続いてシーケンス AB について射影するをすると、抽出されるシーケンスとその支持度は C:1 となり、最小支持度 2 を満たさないのが不適である。続いてシーケンス AC について射影すると、抽出されるシーケンスとその支持度は D:1 となり、これも不適である。これでアイテム A から始まるシーケンスの抽出の処理が終わる。以上の処理を、シーケンス B、シーケンス C に対しても行う。その結果、表 5 の部分シーケンスが抽出される。

## 4. 系列データの分析と評価

本研究では、実際の LBSNS のデータを分析する。4.1 節で分析に用いるチェックインデータについて説明し、4.2 節でス

表 4 アイテム A に続く各アイテムの支持度

A : 1
B : 3
C : 2
D : 1

表 5 PrefixSpan による抽出結果

部分シーケンス	支持度
A B	3
A C	2

ポットパスの抽出結果について分析を行う。

### 4.1 分析に使用するチェックインデータ

本研究ではユーザ数やサービスの普及度合いから、Foursquare のチェックインデータを収集し、分析に使用することとした。収集したデータセットの基本的な情報を表 6 に示す。

Foursquare では、既に登録されているスポットが存在すれば、そのスポットに対してチェックインをして、無ければ新たにスポットを登録してチェックインすることが可能である。スポットには、スポット名、カテゴリ、緯度経度、国名、住所などの情報項目があり、ユーザによって情報が付与されている。カテゴリは現時点では 400 種類以上のカテゴリがあらかじめ決められており、ユーザは選択してスポットに付与する。

### 4.2 ユーザのチェックイン数

ユーザあたりのチェックイン数を示したのが図 1 であり、ユーザによってその利用頻度に差があることが確認できる。その分布は、べき乗分布に近いことがわかる。

### 4.3 チェックインされるスポット

表 7 はチェックインの多い上位 10 位のスポットである。都市部の鉄道駅が 9 割を占めている。11 位以下も鉄道駅が頻出

表 6 分析のデータセット

期間	2012/06/15 - 2012/06/30
チェックイン数	419,349
ユーザ数	45,447
スポット数	135,655
カテゴリ数	394

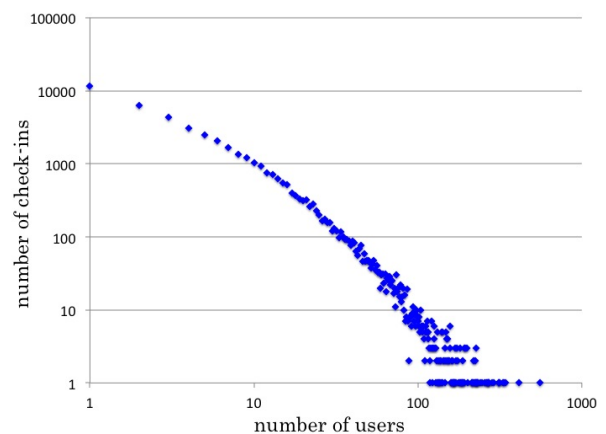


図 1 ユーザのチェックイン数

表 7 チェックイン数上位 10 位のスポット

No	スポット名	件数
1	秋葉原駅	2,583
2	新宿駅	2,163
3	東京駅	1,946
4	渋谷駅	1,347
5	横浜駅	1,177
6	池袋駅	1,166
7	大阪駅	1,026
8	大宮駅	885
9	ヨドバシカメラ マルチメディア Akiba	800
10	川崎駅	799
23	渋谷ヒカリエ	482
36	横浜アリーナ	313

表 8 チェックイン数上位 10 位のカテゴリ

No	カテゴリ名	件数
1	鉄道駅	102,899
2	ラーメン	16,384
3	コンビニエンスストア	16,307
4	和食	13,553
5	地下鉄	12,902
6	ショッピングモール	11,982
7	スーパーマーケット	8,329
8	カフェ	7,936
9	電器店	7,358
10	コーヒーショップ	6,975

表 9 平日と休日のチェックイン数

種類	1日あたりの		
	チェックイン数	日数	平均チェックイン数
平日	245,151	11	22,286.5
休日	174,198	5	34,839.6
全日	419,349	16	26,209.3

していることが確認できた。鉄道駅以外には、都内の大型家電量販店であるヨドバシカメラ マルチメディア Akiba、2012 年春にオープンした複合商業施設の渋谷ヒカリエ、多目的イベントホールである横浜アリーナなどが見られる。

スポットを個別に観測するには、データがスパースであることから、ここから先はカテゴリに着目して分析を進めていく。チェックインの多いカテゴリ上位 10 件を示したのが表 8 である。鉄道駅が非常に多く、全体の約 25%を占めていることがわかる。同様に、交通機関である地下鉄もランクインしている。グルメ系のカテゴリとしてラーメン、和食、カフェ、コーヒーショップなどがランクインをしている。生活用品を買うためのコンビニエンスストアやショッピングモール、スーパーマーケットもランクインしている。

次に、チェックインがされる日を平日と休日に分けて調査した。表 9 から、休日は平日に比べると 1 日あたりの平均チェックイン数が多いことがわかる。しかし、表 10、表 11、表 12、表 13 から、チェックインの多いスポットとカテゴリに、大きな差異は見られなかった。

表 10 チェックイン数上位 10 位のスポット (平日)

No	スポット名	件数
1	秋葉原駅	1,330
2	新宿駅	1,327
3	東京駅	1,138
4	渋谷駅	821
5	横浜駅	725
6	池袋駅	662
7	大阪駅	634
8	大宮駅	568
9	川崎駅	544
10	品川駅	494

表 11 チェックイン数上位 10 位のスポット (休日)

No	スポット名	件数
1	秋葉原駅	1,253
2	新宿駅	836
3	東京駅	808
4	渋谷駅	526
5	池袋	504
6	横浜駅	452
7	ヨドバシカメラ マルチメディア Akiba	443
8	大阪駅	392
9	新横浜駅	373
10	名古屋駅	321

表 12 チェックイン数上位 10 位のカテゴリ (平日)

No	カテゴリ名	件数
1	鉄道駅	66,465
2	コンビニエンスストア	10,955
3	ラーメン	9,857
4	地下鉄	9,063
5	和食	8,202
6	ショッピングモール	5,464
7	スーパーマーケット	4,363
8	カフェ	4,312
9	コーヒーショップ	4,300
10	オフィス	3,561

表 13 チェックイン数上位 10 位のカテゴリ (休日)

No	カテゴリ名	件数
1	鉄道駅	36,434
2	ラーメン	6,527
3	コンビニエンスストア	6,518
4	和食	5,352
5	地下鉄	5,351
6	ショッピングモール	3,966
7	スーパーマーケット	3,928
8	カフェ	3,839
9	電器店	3,835
10	コーヒーショップ	3,624

#### 4.4 チェックインの時間帯

次にチェックインがされる時間帯を調べた。データセットのチェックインデータ数を 1 時間単位として表したのが図 2 であ

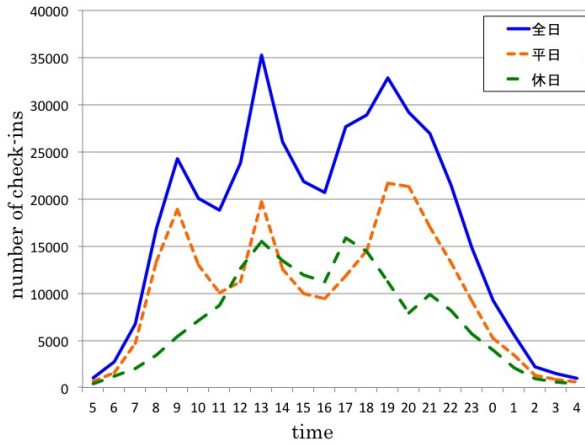


図2 チェックインの時間帯

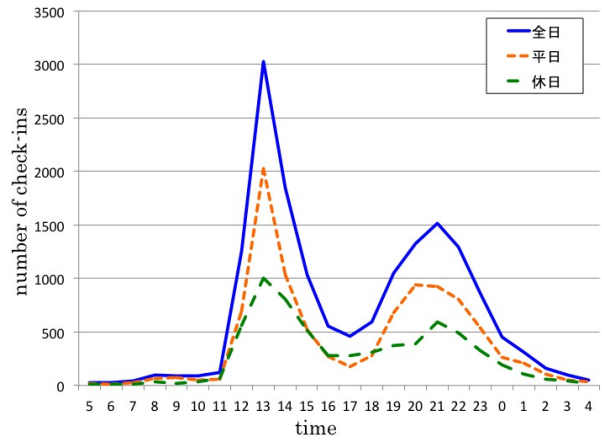


図4 ラーメンのチェックインの時間帯

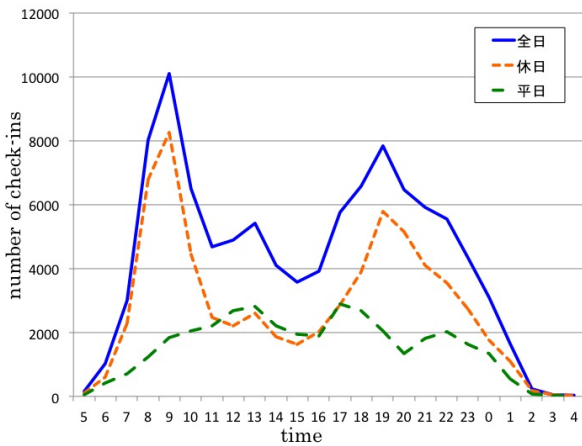


図3 鉄道駅のチェックインの時間帯

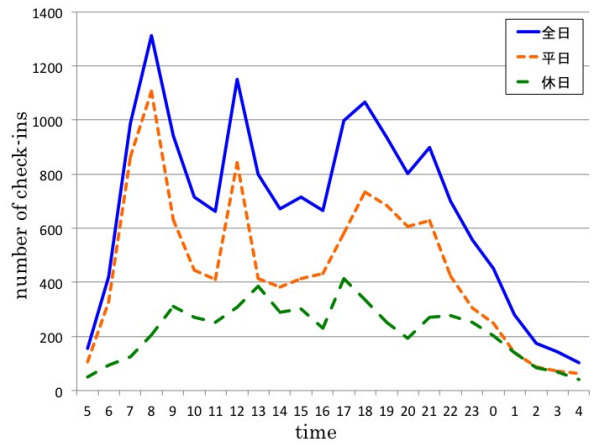


図5 コンビニエンスストアのチェックインの時間帯

る。全日では、9時前後、13時前後、19時前後でチェックインが多くされていることがわかる。平日のチェックインも同様の様子を示している。一方、休日では朝のチェックインは少なく、13時前後、そして17時にピークが存在している。チェックインが少ない時間帯は4時、5時であることがわかる。

また、カテゴリによるチェックインのピークの時間帯を調査した。チェックインの多いカテゴリ上位3つの鉄道駅、ラーメン、コンビニエンスストアを調べた結果が図3、図4、図5である。鉄道駅は平日では9時と19時にピークがあることがわかる。休日では、13時、17時でのチェックインがされてるが、極端に多いというわけではない。ラーメンでは、平日と休日の共に、13時台にチェックインが集中していて、その後21時前後でのチェックインが多いことがわかる。コンビニエンスストアでは、平日の9時前後、12時台、18時前後にそれぞれピークがあることがわかる。休日では17時がピークとなっており、平日とは異なる形となっている。さらに、1日のチェックインに対する時間帯あたり比率を示したのが表6であり、休日は22時以降の比率が、平日に比べ高くなっていることがわかる。このように、チェックインされる時間帯は、平日か休日か、さらにカテゴリによって異なる特徴を持つことが判明した。

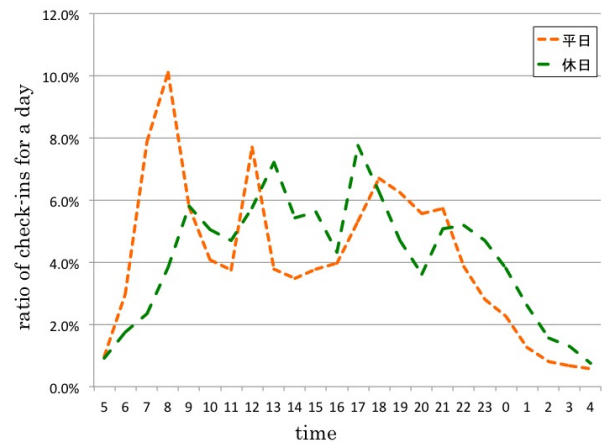


図6 コンビニエンスストアの時間帯あたりのチェックイン比率

#### 4.5 チェックインシーケンス

次にチェックインデータからシーケンスを作成し、その特徴を調べる。チェックインの時間帯は、4時と5時が少ないという傾向から、ユーザ毎に05:00から翌04:59までの時間のチェックインを1シーケンスとして作成する。その結果、シーケンス数と平均シーケンス長は表14の通りであった。この結果から、あるユーザがチェックインを1回以上する日には、1日のうち

シーケンス種類	件数
シーケンス数	182,755
平均シーケンス長	2.29

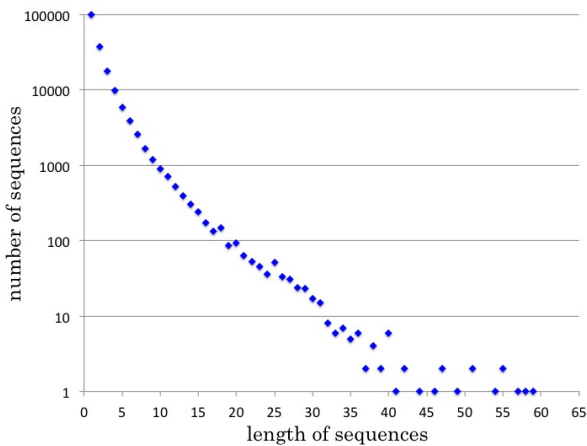


図 7 シーケンス長の乗分布

長さ	件数
2	165
3	80
4	35
5	4
7	1
8	1

平均は 2.29 回チェックインすることが判明した。

表 7 はシーケンス長の分布である。長さ 1 がシーケンスが最も多いが、スポットパス抽出が可能となるシーケンス長 2 以上のシーケンスを合計すると、その割合は半数以上となっている。

## 5. 系列マイニングの結果

4.1 節で作成したシーケンスに対して、PrefixSpan を適用してスポットパスの抽出を行った。最低支持度を 2 と設定し、全シーケンス数の約 0.1% に当たる 180 以上存在するスポットパスを抽出した。その結果、288 種類のスポットパスが得られた。スポットパスの長さの分布を表 15 に示す。スポットパスの平均長は 2.6 であり、多くのスポットパスが長さ 2 であることがわかる。また、長さ 7 および 8 のスポットパスが 1 つずつ存在しているが、これはすべてカテゴリが鉄道駅のスポットによって構成されている。その他に、288 のスポットパスのうち 198 のスポットパスが、鉄道駅を含んでいた。

最後に抽出した 288 種類のスポットパスを、長さ 2 毎のパスに分割し、スポットをノード、次のスポットへの経路をエッジとして有向グラフを生成した。その結果を図 8 に示す。鉄道駅の入出力のエッジ数が共に最も多いことが確認できる。中央の円の外側には、チェックイン数のトップ 50 にもランクインしていないカテゴリが並んでおり、鉄道によって様々な種類のスポットパスが形成されていることがわかる。また、拡大した図

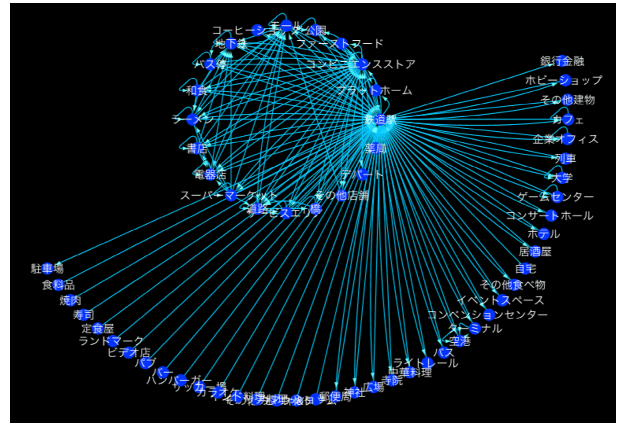


図 8 スポットパスにおける頻出アイテム

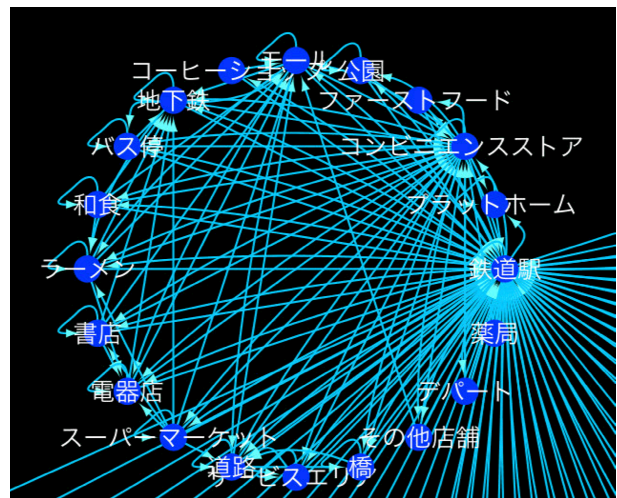


図 9 スポットパスによる頻出アイテム (拡大)

9 から、出現数の多いカテゴリ同士でつながっていることがわかる。その中でもコンビニエンスストア、ショッピングモール、地下鉄、スーパーマーケットなどもエッジ数からハブとなっていることが確認できる。

## 6. 考察

### 6.1 チェックインデータに関する考察

Foursquare のチェックインデータを対象分析をしたところ、図 1 より、ユーザのチェックイン回数はべき乗分布に近い傾向が見られた。その中でも複数回のチェックインをしたユーザは半数程度であったことから、今回抽出されたスポットパスは、Foursquare をある程度利用しているユーザ層の特性を表していることになる。また、チェックインされたスポットのカテゴリを確認したところ、表 8 より鉄道駅が全体の約 25% を占めていた。この要因として、Foursquare が同じ場所でチェックインを重ねることで特典が得られるというサービス設計がされている点、鉄道駅という場所が通勤や通学など普段の生活の中で多く訪れるスポットであるという点、鉄道駅での電車の待ち時間はスマートフォンを操作してチェックインするのに適している点などが考えられる。その他、ラーメンやコンビニエンスストア、和食、地下鉄、ショッピングモールといったカテゴリで

のチェックインが多いのは、これらのスポットを生活の中で利用する機会が多いためと考えられる。チェックインの時間帯に着目すると、図2より、全体の傾向として9時・13時・19時でのチェックインが多いことがわかった。これは、Foursquareでは社会人や学生のユーザが多く、これらの時間が通勤通学時、昼休み、退社、下校などの時間となっていることが考えられる。一方で、カテゴリや、平日か休日かによってピークの分布が異なっていることも確認できた。

チェックインデータからシーケンスを作成した結果、表14よりシーケンスの平均長は2.29であることがわかった。

## 6.2 スポットパスに関する考察

スポットパスを抽出した結果、288パターンのスポットパスを抽出した。113のカテゴリから構成されており、表15より、スポットパスの約70%は長さが2であった。であった。その中でも、鉄道駅が多く出現していることが確認できた。最後に抽出したスポットパスからネットワーク図を作成した。図8、9より、鉄道駅が移動軌跡の主要なハブとなっていること、コンビニエンスストア、ショッピングモール、スーパーマーケットなどのカテゴリも出入次数が多く、生活に密着した近いスポットがやはりハブとなっている。その他、鉄道としか移動軌跡を形成していないカテゴリが存在しており、外出時に移動手段として鉄道駅でチェックインをして、目的地のスポットでチェックインをするといったような利用がされていることがわかる。

## 7. おわりに

本研究では、LBSNSの代表的なサービスであるFoursquareに蓄積されたチェックインデータを対象に、系列マイニングを適用することでユーザが訪問したスポットのシーケンスであるスポットパスの抽出した。

まず、Foursquareのチェックインデータの概要を分析したところ、カテゴリによってチェックイン数や、チェックインの時間帯などが異なる特徴を示すことがわかった。特に鉄道駅のチェックインが頻出しており、その他、ラーメンや和食などの飲食店、コンビニエンスストアやショッピングモールなど日々の生活で利用するカテゴリが多く見られた。

作成したシーケンスに対し、PrefixSpanを適用することで、288種類のスポットパスが抽出された。スポットパスの長さは多くが2および3であり、鉄道駅を含むスポットパスが多く抽出された。次に、抽出した288種類のスポットパスを長さ2毎のパスに分割し、スポットをノード、次のスポットへの経路をエッジとして有向グラフを生成した。その結果、鉄道カテゴリを主要なハブ、その他コンビニやショッピングモールなど生活に近いカテゴリもハブとなるようなネットワークが確認できた。

今後の課題として、Foursquareはユーザによって利用具合が異なるため、ユーザ層によって特徴的なスポットパスを抽出することが挙げられる。そのためには、他の期間のチェックインデータでの分析や、チェックイン数やカテゴリを考慮したユーザのクラスタリングの処理が挙げられる。

## 謝 辞

本研究の一部は筑波大学図書館情報メディア系プロジェクト研究による助成を受けたものである。

## 文 献

- [1] 篠田裕之, 竹内亨, 寺西裕一, 春本要, 下條真司: 行動履歴に基づく協調フィルタリングによる行動ナビゲーション手法. 情報処理学会研究報告, 2007-DPS-132, pp. 87-92, 2007.
- [2] Z. Yin, L. Cao, J. Han, J. Luo, and T. S. Huang: Diversified Trajectory Pattern Ranking in Geo-tagged Social Media. In *SDM*, pp. 980-991, 2011.
- [3] M.D. Choudhury, M. Feldman, S. Amer-Yahia, N. Golbandi, R. Lempel and C. Yu: Constructing Travel Itineraries from Tagged Geo-Temporal Breadcrumbs, in *Proc. Int. Conf. on World Wide Web (WWW)*, pp. 1083-1084, 2010.
- [4] T. Kurashima, T. Iwata, G. Irie, and K. Fujimura: Travel route recommendation using geotags in photo sharing sites. In *CIKM*, pp. 579-588, 2010.
- [5] E. Cho, S. A. Myers, and J. Leskovec: Friendship and mobility: User movement in location-based social networks. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2011.
- [6] 旭直人, 山本岳洋, 中村聡史, 田中克己: 行動連鎖を用いた情報検索支援とWebからの行動連鎖の抽出. 第1回データ工学と情報マネジメントに関するフォーラム (DEIM2009), A7-2, 2009.
- [7] 郡宏志, 服部峻, 手塚太郎, 田島敬史, 田中克己: ブログからのビジターの代表的な行動経路とそのコンテキストの抽出. 信学技報, DE2006-55(2006-7), pp.29-34, 2006.
- [8] 箴島郁子, 嶋田和孝, 遠藤勉: 系列パターンを利用した評価表現の分類. 言語処理学会第11回年次大会, 2009.
- [9] 山田 吾郎, 吉田 則裕, 井上 克郎: シーケンシャルパターンマイニングに基づくオブジェクト指向プログラムのための欠陥検出手法. 情報処理学会研究報告, Vol.2009-SE-164/Vol.2009-EMB-13/Vol.2009-CSEC-45, No.15, pp.1-8, 2009.
- [10] J. Pei, J. Han, B. Mortazavi-Asl, H. Pinto, Q. Chen, U. Dayal, and M.-C. Hsu: PrefixSpan mining sequential patterns Efficiently by prefix projected pattern growth. In *ICDE 2001*, pp. 215-226, Heidelberg, Germany, 2001.