

都市内回遊行動評価のための空間データマイニング*

Spatial data mining for city-round trip analysis*

羽藤 英二**，中西雅一***，寺谷寛紀***，柏谷増男****

By Eiji HATO, Masakazu NAKANISHI, Hiroki TERATANI and Masuo KASHIWADANI

1. はじめに

従来の都市圏レベルでの交通需要予測では、個人の一日の活動のうち目的のはっきりした活動に派生するトリップを一旦分解し、分析の単位として独立に取り扱う。大括りなゾーン単位での OD 表やそれに基づく総移動量の予測及び経路への配分では、こうしたアプローチが適している。一方本研究では、高度に都市機能が集積したある限られた狭い地区における移動-活動分析手法として空間データマイニングを考える。

都市再生事業では、狭域エリアにおける人の移動-活動パターンそのものに着目した上で、滞在する複数の施設やその回遊経路、滞在時間などについて詳細に分析する必要がある。この際、高度な都市機能が集積する地区では多様な活動パターンが複合的に存在する。分析の基となる回遊行動データを定型のトリップに分解して取り扱うよりも、そのまま一つのトランザクションとして取り扱うことが重要となる。通勤のような定型のトリップが存在しない場合、OD 表に交通需要を集約化し、なんらかの意思決定モデルにデータを当てはめて活動空間を物理的に結びつけるネットワーク整備や施設配置の効果を予測することは難しい。寧ろ都市空間に実装された Bluetooth や無線タグ、擬似 GPS 衛星、移動体通信を基にして多量かつ正確な位置データを収集し、このデータを基にして、都市空間において頻出する重要な移動-活動パターンを抽出することが重要となる。パターンの抽出のためには某大で常に化するデータを「ざっと見渡す」ことが必要であり、抽出*

キーワード：データベース、データマイニング

**正員，工博，愛媛大学工学部環境建設工学科

(松山市文京町3, hatoen2@ehime-u.ac.jp)

***学生員，愛媛大学工学部環境建設工学科

****フェロー，工博，愛媛大学工学部環境建設工学科

するパターンはむやみに数が多くなく、理解できることが求められる。

本研究では、松山市の中心な商業地域を対象に得られた回遊行動データを基に空間データマイニングを行ない、頻出する移動-活動パターンを抽出する方法を検討する。

2. データマイニングについて

データマイニングとは、データベースから潜在する構造や規則性を抽出し、モデル化を行ない、仮説検証や予測を行なう手法である。データに空間情報、位置情報を用いると、頻出する都市の空間パターンを抽出することができる。この空間パターンを用いることで、特定の商業施設と近隣の他の施設との相互関係などが分析できる。規則性の抽出方法として、決定木、相関ルールなどが挙げられる。本研究では相関ルールを用いる。

相関ルールとは、スーパーマーケットでの顧客の購買パターンを分析する方法として、Agrawalらにより1993年に提起されたものである。POSシステムで得られる顧客毎の大量の情報(客層、購入商品、購入金額、購入時間など)の中から「目玉商品 A を購入した顧客は高い確度で日用品 B を購入する」というような規則を導く。このような規則が得られたとすれば、目玉商品 A が他のどの商品の売上に貢献するかが明瞭になる。さらに目玉商品 A を購入した顧客の何パーセントが日用品 B を購入するかが分かれば、売上の予測にもある程度つながる。一般に、 X, Y を商品の集合として(この例では $X = \{ \text{目玉商品 A} \}$, $Y = \{ \text{日用品 B} \}$)

$X \rightarrow Y$

という形をした規則を相関ルールと呼ぶ。また、相関ルール $X \rightarrow Y$ の左辺 X を条件部、右辺 Y を結論部と呼ぶ。

相関ルールの価値基準には、確信度とサポートが

表 2 1 観測度数（上段）と期待度数（下段）

	H	-H	合計
X	$N S_{XY}$ $N S_X S_Y$	$N(S_X - S_X S_Y)$ $N(S_X - S_X S_Y)$	$N S_X$
-X	$N(S_Y - S_X S_Y)$ $N(S_Y - S_X S_Y)$	$N(1 - S_X - S_Y + S_X S_Y)$ $N(1 - S_X - S_Y + S_X S_Y)$	$N(1 - S_Y)$
合計	$N S_Y$	$N(1 - S_Y)$	N

ある．確信度とは，条件部を満たすデータが結論部も満たす割合のことである．サポートとは，条件部と結論部を同時に満たすデータの全データに対する割合のことである．相関ルールの価値を測る上で，確信度が高いことが重要であることは当然である．また，サポートも，高いことが重要である．

相関ルールの価値基準には条件部と結論部の独立性に基づく考え方もある．相関ルール $X \rightarrow Y$ を考える． $X, Y, X \rightarrow Y$ のサポートをそれぞれ S_X, S_Y, S_{XY} とし，トランザクションの総数を N とする． X と Y が独立としており，同じトランザクション内に含まれるのが単なる偶然であると仮定したときに期待される期待度数と，実際に観測されたトランザクション数（観測度数）は，表 2-1 のようになる．分割表に対して，次の検定量

$$T_{dep} = \sum_{\text{各セル}} \frac{(\text{観測度数} - \text{期待度数})^2}{\text{期待度数}}$$

は自由度 1 のカイ 2 乗分布に従うことが知られている．これに表 2-1 の各値を代入して簡略化すると，

$$T_{dep} = N \frac{(S_{XY} - S_X S_Y)^2}{S_X S_Y (1 - S_Y)(1 - S_X)}$$

を得る． T_{dep} が 0 に近づけば X と Y は互いに独立であり，大きければ正負を問わず互いの相関が強いといえる．そこで，ある有意水準 α を定めて， $T_{dep} < c^2$ () であれば， X と Y が独立であると見なし， $X \rightarrow Y$ が発見されたのは単なる偶然であるから，価値がないとして捨てる．

本研究では，商品を回遊行動中に立ち寄った店舗に置き換え，店舗間の関係性を調べた．また，中心市街地の空間上での回遊行動の頻出パターンを得るために，対象地域を図 2 のような 8 個のエリアに分類して，そのエリア間の関係性を調べた．

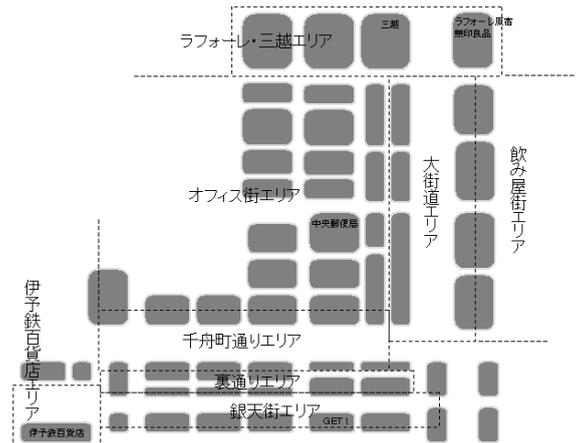


図 2 エリア分類図

3. データについて

中心市街地の回遊行動の分析フレームワークを検討するために，松山市の中心市街地でパイロット調査を実施した．被験者は大学生で，1人または2人のグループで自由に行動してもらう方法を用いた．同時に被験者に対して調査前及び調査後に，調査対象地域の地図を描いてもらうアンケートも実施した．調査後のアンケートに関しては，実際に通った経路も記入してもらった．

調査は5月と12月に行ない，被験者は5月の調査では10人（1人×2組，2人×4組），12月の調査では，27人（1人×11組，2人×4組）で行った．

分析データは，回遊行動追跡調査のアンケートで得られた地図から得られる 25 グループの行動データを用いる．相関ルールの価値の測定については，サンプルデータ数（25）を全データ数，立ち寄り店舗もしくはエリアを条件部及び結論部として確信度とサポートを計算した．

本研究での相関ルールの算出方法は以下に述べる手順で行った．

- ・店舗間の関係性については，被験者が立ち寄った店舗を表 3-1 で示すカテゴリー（31 種類）に分類し，被選択店舗の組み合わせを算出する．

- ・エリア間の関係性については，被験者が移動した経路を図 2 に基づいてエリア間の移動に変換し，エリアの組み合わせを算出する．

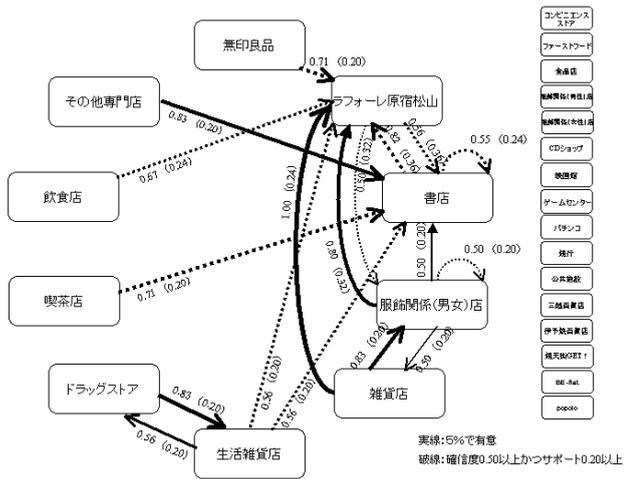


図4 1 4 店舗間相関ルール結果

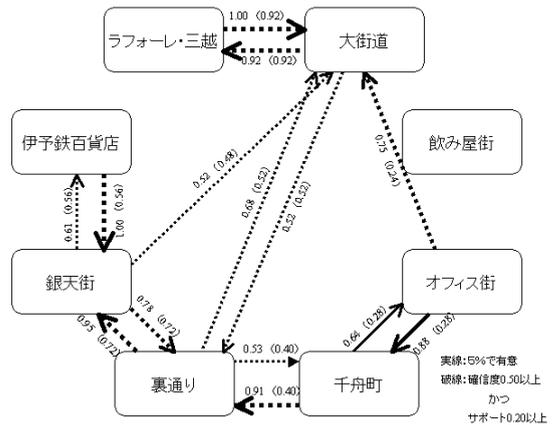


図4 2 3 エリア間相関ルール結果

遊行動が多く見られる。銀天街エリア訪れると、再び銀天街を訪れる確率は、大街道と比較すると小さい。大街道エリアから伊予鉄百貨店への回遊行動は少ないと考えられる。

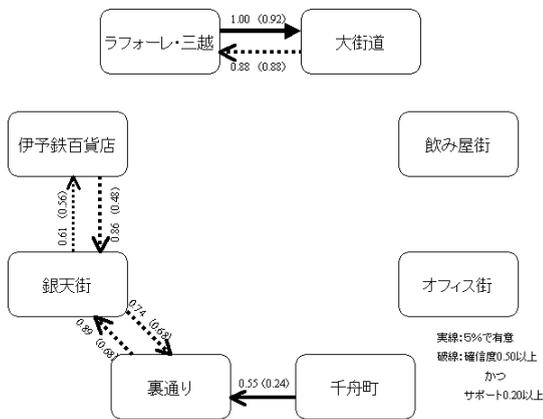


図4 2 1 エリア間相関ルール結果

5. まとめ

膨大な数のデータから知識としての都市回遊行動の特徴を抽出することで、都市空間の再評価を行なうことが可能となる。分析を行なった結果以下のことが明らかとなった。

店舗属性的には同じ属性の店舗を選択する可能性が高い。

回遊行動の多くは大街道エリアに集中する傾向がある。

銀天街エリアでは、南北方向の回遊行動が多く見られる。

今後の課題としては、移動体通信から得られる膨大な位置データから自動的に移動-活動パターンを判別し相関ルールなどの抽出を行なうツールの実装、組み合わせ数を増やしたときの効率的なアルゴリズムの開発が考えられる。

参考文献

- 1) 福田剛志, 森本康彦, 徳山豪; データマイニング, 共立出版株式会社, 2001.

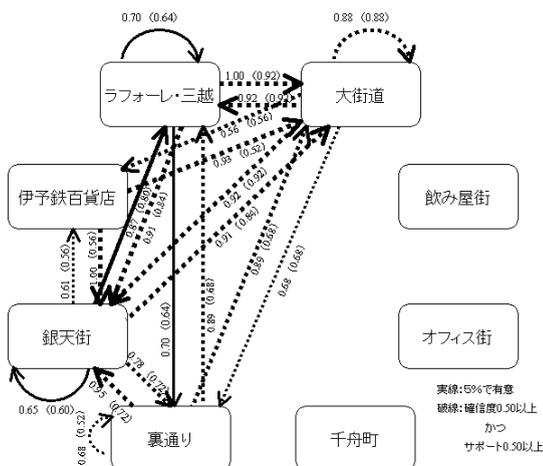


図4 2 2 エリア間相関ルール結果

以上の結果より、銀天街エリアでは、東西方向の回遊行動以外に、裏通りエリアとの南北方向の回