The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

オフィスワーカーの入退室データを用いた移動時間パターンの分析

Analysis of Moving-Time Patterns from Entry and Exit Data of Office Workers

小島世大 Seidai Kojima　石榑隼人 Hayato Ishigure　坂田美和 Miwa Sakata　武藤敦子 Atsuko Mutoh　森山甲一 Koichi Moriyama　犬塚信博 Nobuhiro Inuzuka

名古屋工業大学 大学院 工学研究科 情報工学専攻
Department of Computer Science, Graduate School of Engineering, Nagoya Institute of Technology

Recently, IC card systems are popular and their log data are used for analyzing human behaviors. In this paper, we propose an analysis method of employee's moving-time patterns using log data in entry and exit management system. We express employees with feature quantities focused on moving time from room to room, extract a set of employees with similar moving-time patterns using clustering, and confirm the relationship between moving-time patterns and employee attributes. We examine our proposed method using actual entry and exit data and confirm the effect.

1. はじめに

近年 IC カードの普及に伴い、IC カード利用履歴を用いた人の行動分析の研究が増えている [1][2]。

さらに、電子錠と IC カードにより、人の入退を制御する入退室管理システムの導入が増えている。入退室管理システムは、電子錠に備えられたカードリーダにかざされた IC カードの情報を読み取り通行履歴として記録する。現在では、企業は単純な入退室制御だけではなく、社員や部門毎の入退室傾向の分析や社員がどの部屋にいるのかを把握したいなどの要望がある [4]。また、近年、センサノード及びネットワークの技術が急速に発展している。環境に設置されたセンサノードと人が身に着けるウェアラブルセンサーの併用は、企業におけるオフィス内のワーカーの行動のモニタリングへ利用されている。センサデータを用いたオフィスワーカーの行動分析 [5] もされているが、これには特別な装置を必要とする。本研究では、多くの組織で普及が進んでいる入退室管理システムから得られる入退室データを用いて、社員の移動時間パターンを抽出し、移動時間パターンと社員属性の関係を分析する。本研究で移動とはある部屋を退室し、次にある部屋に入室するまでのこととする。入退室データから社員の行動分析ができれば、企業の働き方改革等への利用も考えられる。

2. 関連研究

IC カードの利用者数の増加により、IC カード利用履歴データが人の行動分析手段として注目を集めている。 鈴木ら [2] は交通 IC カードの利用履歴を用いて人が駅の改札口を出て、次の改札口に入るまでの間にその人の滞在目的があるとし、生活パターンを定量的に表す生活行動属性を提案した。 カード利用者の生活圏 (駅) によらず似た生活パターンを持つ人を抽出しマーケティング等へ活用し得ると示した。また、嶋本ら [1] は英国・ロンドンで導入されている Oyster Card の 4 週間分の利用履歴データを用いて公共交通の変動を把握した。料金支払い形態に応じて利用者の利用属性を分類することで，1 人あたりの利用回数の変動の 4 割以上を説明できることを示した。さらに，分析対象の 4 週間の期間に学生の秋休み期間が含まれていることに着目した分析を行い、秋休み期間中には直接関連する学生のみならず学生以外の利用者の公共交通の利用が減少することを明らかにした。

水岡ら [3] は 大量の視聴データにクラスタリング手法を適用し、好みのチャンネルの連続視聴や特定日時のチャンネル切り替えなど特徴的な視聴パターンを持つ集合を機械的に抽出できることを示した。視聴パターンとは、時間帯ごとの視聴行動の組み合わせのことであり、特徴的な視聴パターンの抽出や、その視聴パターンを持つ視聴者集合の傾向分析をすることは、視聴者の特性を生かした番組や CM を作成することが容易になり、テレビメディアの価値の向上、視聴者に合わせた魅力的なコンテンツの提供などの抽出、分析することのメリットについても述べている。

本研究では、会社内の入退室データからクラスタリングを用いて、似た移動時間パターンを持つ社員の集合を抽出する。その集合に含まれる社員の属性分布と移動時間パターンの関係について分析し、考察する。

3. 入退室データ

本研究では (打刻日、打刻時刻、場所、操作、社員 ID) の 5 つの属性で構成された入退室データを扱う (表 1)。操作には、部屋を退室もしくは入室したかが記録されている。移動時間に着目するために、退室と入室がセットになっているものを結合し、表 2 のようなデータ構造に変換する。移動時間は入室時刻 − 退室時刻で計算する。

表 1: 入退室データ

打刻日	打刻時刻	打刻場所	操作 (退室 or 入室など)	社員 ID

表 2: 変換後のデータ構造

退室日	退室時刻	移動時間	退室場所	入室場所	社員 ID

4. 事前調査

協力企業の入退室データ (2016 年 6 月 332 名分) から社員の移動時間に着目した調査を行った。図 1 に移動開始時刻と移動時間の散布図を示す。図 1 より、どの時間帯においても 0-5

連絡先: 小島世大, 名古屋工業大学 大学院, 〒 466-8555 愛知県名古屋市昭和区御器所町, s.kojima.571@nitech.jp

分の移動時間である移動が多く、昼休憩の時間帯 (12 時から 12 時 50 分) は移動時間が長い移動が多くみられる。さらに、協力企業の聞き取り調査より昼休憩で移動時間の長い移動は社外にランチを取りに行っている可能性があると考えられる。

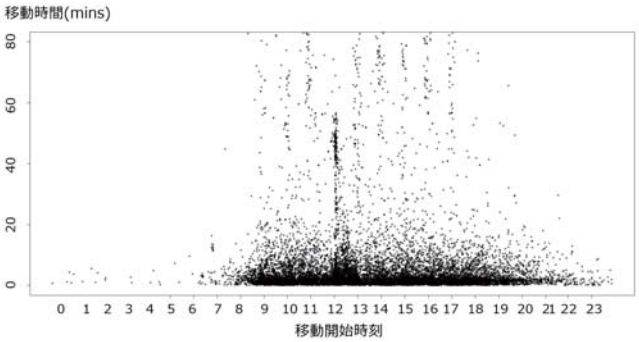

図 1: 移動開始時刻と移動時間の散布図

次に移動開始時刻毎の移動回数の頻度を図 2 に示す。移動開始時刻を 10 分で区切り、6 種類の色分けで示した。図 2 から毎時 50 分から 0 分 (ピンク) 及び 0 分から 10 分 (赤) に移動が多くみられる。また、特に昼 12 時頃の移動が多くみられた。

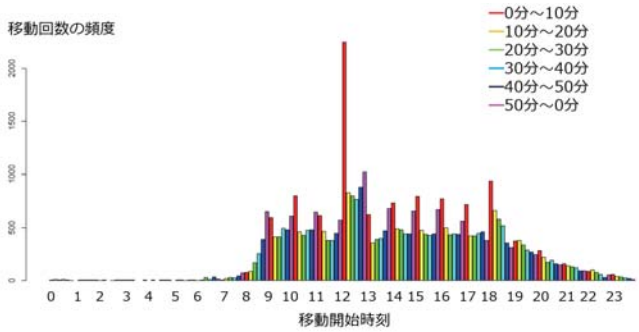

図 2: 移動開始時刻毎の移動回数の頻度

これらの事前調査より、時間帯や移動時間区分毎に移動回数を数える方法で社員の移動時間のパターンを読み取ることが可能だと考える。

5. 提案手法

提案手法の流れを図 3 に示す。まず、全社員に対して時間帯及び移動時間区分毎の移動回数を移動特徴量として算出する。算出した移動特徴量を正規化しクラスタリングを行い、適合度検定を用いてクラスタ毎の社員属性の分布とそのクラスタの移動パターンの特徴について考察をする。

5.1 移動特徴量

社員 p を次の特徴量ベクトル F_p で表す。

$$F_p = (m_{p1}, m_{p2}, ..., m_{pi}, ..., m_{pk})$$

社員 p の移動特徴量 m_{pi} は移動区分 i(時間帯及び移動時間区分) の移動回数とする。例えば、図 4 に示す社員 ID1 の午前の移動時間区分 (0-5 分) の移動回数は、入退室データより 10:00 と 11:00 の 2 回となる。

移動特徴量の算出 (5.1 節)
↓
移動特徴量の正規化 (5.2 節)
↓
移動特徴量によるクラスタリング (5.3 節)
↓
適合度検定 (5.3 節)

図 3: 提案手法の流れ

退室日	打刻時刻	移動時間	退室場所	入室場所	社員ID
6/12	10:00	4分	Room1	Room1	1
6/12	11:00	3分	Room1	Room2	1
6/13	13:00	8分	Room2	Room4	1
6/12	14:00	7分	Room2	Room3	2
6/12	15:00	2分	Room3	Room3	2

入退室データ(変換後)

社員ID	午前 移動時間区分 (0-5分)	午前 移動時間区分 (5-10分)	午後 移動時間区分 (0-5分)	午後 移動時間区分 (5-10分)
1	2回	1回	0回	0回
2	0回	0回	1回	1回

図 4: 社員毎の移動特徴量の算出例

5.2 移動特徴量の正規化

5.1 節で算出した移動特徴量からクラスタリングによって似た移動時間パターンを持つ社員の集合を抽出するために、値の正規化を行う。似た移動時間パターンとは、社員の総移動回数に対する移動区分毎の移動回数の割合が似ている場合と考えられるため、まず、社員に対して総移動回数のみでクラスタリングをする。クラスタリングには k-means 法を用い、クラスタ数は elbow 法で決定する。クラスタリングの結果、クラスタ内人数の少ないクラスタや総移動回数の極めて少ないクラスタは移動時間パターンの抽出が困難であると判断し、取り除く。

次に総移動回数でクラスタリングした結果と次の式を用いて 5.1 節で表した社員 p の移動特徴量 m_{pi} を正規化する。

$$m_{pi} \leftarrow \frac{m_{pi}}{\sum_{j \in c} m_{ji}}$$

c は社員 p を含むクラスタである。移動特徴量毎に社員の移動特徴量の値の割合を計算することで移動特徴量毎の値の大きさを揃える。しかし、全社員に対して移動特徴量毎の計算を行うと、移動時間パターンの似た移動回数の少ない社員と移動回数の多い社員の移動特徴量の割合は異なる値となる。そこで、社員の移動特徴量の割合をクラスタ毎に計算することで、移動回数の少ない社員と移動回数の多い社員を比較したときに、それぞれの移動区分毎の移動回数の割合が同じであった場合に似たパターンを持つ社員であるとみなす。

5.3 移動特徴量によるクラスタリングと適合度検定

5.2 節の方法で正規化した移動特徴量を用いて全社員に対してクラスタリングを行う。クラスタリングには k-means 法を用い、クラスタ数は elbow 法で決定する。次に、クラスタ毎の社員属性の分布と全社員属性の分布に対して適合度検定を行い、検定結果の中で有意差のあったクラスタ毎に移動時間パターンと社員属性の分布の関係について考察する。

6. 実験

6.1 実験方法

実験では協力企業の 2016 年 6 月の 332 名の入退室データを表 2 の形式に変換し、60 分以内の移動時間であった移動のみを用いる。移動特徴量の正規化を行い、社員を 9 次元 (時間帯 (午前・昼休憩・午後の 3 次元) ×移動時間区分 (0-5 分・5-20 分・20-60 分の 3 次元)) の特徴量で表した。適合度検定で用いる社員属性は、(性別、部署、職種、採用種別、年代、社歴) の分布とし、有意水準は 5 %とした。

6.2 結果と考察

移動特徴量の正規化のためのクラスタリングの結果、クラスタ内人数が少ないクラスタ 1 と移動回数の少ないクラスタ 7 の 97 名を取り除いた (表 3)。残りの 235 名の社員を正規化した移動特徴量によるクラスタリングを行った結果、9 つのクラスタに分かれた (図 5)。ここで、クラスタ同士の移動時間パターン (社員の総移動回数に対する時間帯及び移動時間区分毎の割合) がすべて異なる特徴を持っていたことが確認できた。表 4 に検定結果の p 値を示す。赤字で示す 5 つの項目において有意差を確認した。次に、有意差の出た項目について、クラスタの時間帯及び移動時間区分毎の移動回数の特徴と社員属性分布の関係について考察する。

表 3: 移動特徴量の正規化のためのクラスタリング結果

クラスタ番号	クラスタ内人数	移動回数の平均
1	2	581
2	13	308
3	46	214
4	60	150
5	71	71
6	45	58
7	95	8

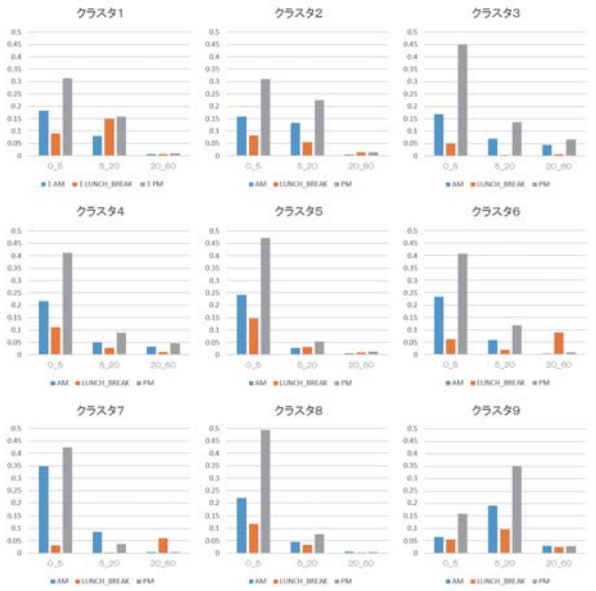

図 5: 全クラスタの移動時間パターンの特徴 (x 軸は移動時間区分,y 軸は移動回数の割合の平均)

表 4: 検定結果 (赤字は p< 0.05,df は自由度)

クラスタ	クラスタ内人数	性別 (df:2)	部署 (df:17)	職種 (df:9)	採用種別 (df:4)	年代 (df:8)	社歴 (df:6)
1	20	0.3	0.3	0.6	0.1	0.2	0.9
2	39	0.08	0.4	0.9	0.1	0.9	0.7
3	4	0.08	0.008	0.09	0.1	0.5	0.3
4	28	0.4	0.1	0.3	0.4	0.4	0.3
5	98	0.6	0.8	0.9	0.6	0.9	0.9
6	26	0.3	0.04	0.4	0.3	0.5	0.3
7	2	0.06	0.4	0.02	0.02	0.5	0.4
8	8	0.7	0.2	0.3	0.4	0.6	0.08
9	10	1	0.9	0.5	0.2	0.8	0.03

クラスタ 6 の人数は 26 人であった。図 6 にクラスタ 6 に含まれる社員、図 7 に全社員の時間帯及び移動時間区分毎の総移動回数に対する移動回数の割合の平均を示す。図 6 と図 7 を比較すると、移動時間が長い移動が昼休憩に多い特徴を持つクラスタであった。次に、図 8 にクラスタ 6 に含まれる社員と全社員の部署の分布を示す。このクラスタ 6 に含まれる社員は全社員の分布に比べて 1800 の部署が特に多く、逆に 1500 の部署は全社員の分布に比べて少ないことが分かる。この結果から 1800 の部署は部署内で社員同士が仲が良く昼休憩も一緒にとりに行く人が多い部署であるのではないかという推測もできる。

次に、図 9 にクラスタ 9 の 10 人の時間帯及び移動時間区分毎の総移動回数に対する移動回数の割合の平均を示す。クラスタ 9 の 10 人と全社員の移動時間区分毎の特徴を比べたところ、クラスタ 9 は午前、昼休憩、午後で少し長めの移動時間区分 (5-20 分) の移動が多い特徴を持つクラスタであった。有意差がでた社歴の分布を確認すると、クラスタ 9 に含まれる社員は全社員の社歴分布に比べて 10 年以上 15 年未満の社歴である社員が特に多く、社歴が 5 年未満の社員が少ないことが分かった (図 10)。この結果から、社歴が少し長くなってきた社員は移動に時間がかかることが多く、たばこ等の小休憩を取りやすくなっている可能性も推測できる。

他に、クラスタ 3 の 4 人と全社員の時間帯及び移動時間区分毎の特徴を比べたところ、クラスタ 3 は昼休憩の移動回数が少ない特徴を持つクラスタであった。有意差がでた部署の分布を確認すると、クラスタ 3 に含まれる 4 人はそれぞれが違う部署に属していた。適合度検定を行って有意差が出たのは期待度数の設定が全データの社員属性の分布の割合とするためであると示唆される。よって、このクラスタの移動時間のパターンが部署に関係しているとは考えにくい。

また、クラスタ 7 の 2 人と全社員の時間帯及び移動時間区分毎の特徴量を比べたところ、午前の短い時間の移動が多く、昼

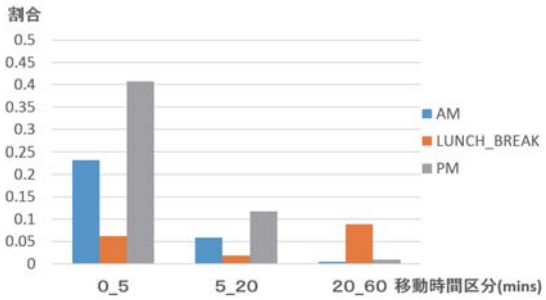

図 6: クラスタ 6 に含まれる社員の移動時間区分毎の移動回数の割合の平均

休憩は長めの時間の移動が多い特徴を持つクラスタであった。有意差がでた職種と採用種別の分布を確認すると、社員 2 人のそれぞれが違う職種と採用種別であった。適合度検定を行って有意差が出たのは期待度数の設定が全データの社員属性の分布の割合とするためであると示唆される。よって、このクラスタの移動時間のパターンが職種と採用種別に関係しているとは考えにくい。

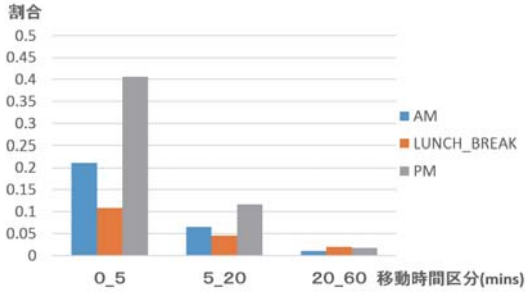

図 7: 全社員の移動時間区分毎の移動回数の割合の平均

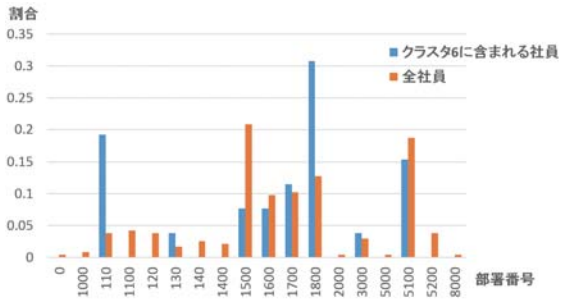

図 8: クラスタ 6 に含まれる社員と全社員の部署分布

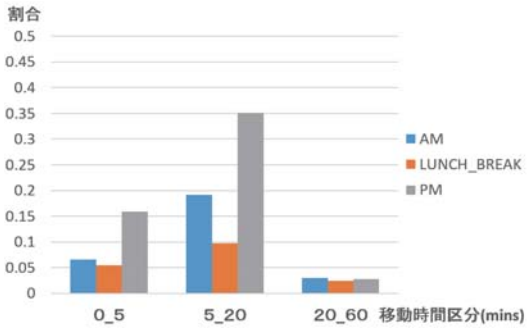

図 9: クラスタ 9 に含まれる社員の移動時間区分毎の移動回数の割合の平均

7. まとめと今後の課題

本研究では、入退室データを用いて社員の移動時間パターンを抽出し、分析をした。入退室データから社員の移動時間に関する事前調査を行った結果、時間帯及び移動時間区分毎の移動回数に社員毎の特徴があることが示唆された。そこで、社員毎の総移動回数に対する時間帯及び移動時間区分毎の割合が似たパターンの社員の集合を抽出するために正規化した移動特徴量を用いて、クラスタリングを行った。クラスタ毎の社員属性の分布と全社員属性の分布に対して適合度検定を行い、有意差がでた結果の考察を行った。クラスタリングでは特徴量を正規化したことで、似た移動時間のパターンを持つ社員の集合を抽出し、クラスタ同士の移動時間パターンがすべて異なる特徴を持っていたことが確認できた。このことより、提案手法を用いて特徴ある似た移動時間パターンを持つ社員の集合を抽出することができたと考えられる。

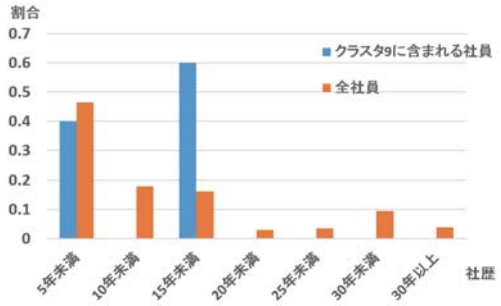

図 10: クラスタ 9 に含まれる社員と全社員の社歴分布

今回社員に対して特徴量を移動時間区分毎の移動回数で表したが、そのほかにも場所に注目したクラスタリングをすることで特徴的な行動の抽出をすることも可能であると考える。その際には、次元数も増えるため、データ期間を増やして実験することや、クラスタリング方法についても再検討する必要があると考える。また今回はクラスタ毎の社員属性について適合度検定を行ったが、標本サンプル数が少ないという欠点もある。そこで、他の検定方法を試すことや、検定対象についても検討し直してみる必要がある。

謝辞

本研究を進めるにあたり、入退室データを提供して頂いた協力企業に感謝の意を表する。

参考文献

[1] 嶋本寛, 北脇徹, 宇野伸宏, 中村俊之, “IC カード利用履歴データを用いた公共交通需要変動分析”, 土木学会論文集 D3(土木計画学),Vol.70,NO.5(土木計画学研究・論文集 第 31 巻) ,pp.605-610,2014.

[2] 鈴木敬, 相薗敏子, “交通 IC カード利用履歴を用いた生活行動属性指標の提案”, 信学技報,IEICE,Technical Report,LOIS2011-84,2012.

[3] 水岡良彰, 陶亜玲, 中田康太, 折原良平, “大規模テレビ視聴データクラスタリングによる視聴パターンの分析”, 特集論文, 理論と方法 (Sociological Theory Methods), Vol.26, No.1:pp.83-97, 2011.

[4] 佐藤雅之, 及川和彦, 永嶋規充,“入退室管理システムにおける通行履歴の応用”, 情報処理学会第 77 回全国大会,2G-05

[5] 岡田将吾, 神谷祐樹, 佐藤祐作, 藤田義弘, 山田敬嗣, 新田克己,“センサ環境を利用したオフィスワーカーの行動パターン分析”,The 27th Annual Conference of the Japanese Society for Artificial Intelligence,2013