

入退室カードのログとネットワーク接続情報を用いた 部屋利用状況の未来予測

野本 聡史[†] 西脇 雅幸[‡] 竹島 亮[‡] 櫻井 瑛一^{†‡} 本村 陽一^{†‡} 長尾 確[‡]

名古屋大学 工学部電気電子情報工学科情報工学コース[†]

名古屋大学 大学院情報科学研究科[‡]

産業技術総合研究所サービス工学研究センター大規模データモデリングチーム^{†‡}

1. はじめに

人間の行動を予測することには様々なメリットがある。例えば、予測された未来に対する何らかのナビゲートをすることができたり、相手の現在の行動を予測することで適切なコミュニケーション手段を選択することができる。また、そういった「行動を予測する」という研究において、人間のあらゆる行動履歴を利用し予測を行う、という様々な研究がなされている。

名古屋大学には、入退室時に IC カードを用いる部屋がある。その部屋は、Leaders' Saloon と呼ばれ、特定の学生が、研究を進めたり、ミーティングを行ったりする場として活用されている。本研究では、その部屋を利用する学生が入退室する際に残るログや、その部屋の LAN 接続情報などを用いて、特定の部屋を利用する学生の利用状況を予測するシステムを設計する。

2. 部屋利用状況予測システム

2.1. 使用データ

本研究で使用するデータは、IC カードのログ、部屋の LAN への接続情報、スケジュールデ

ータであり、それぞれデータベースに保存されている。

前述の Leaders' Saloon という部屋には特定の学生だけが出入りしており、学生それぞれに対して ID つきの IC カードが配布されている。IC カードのログは、入退室時にカードを用いる際に自動的に取得され、誰が、いつ入室または退室したのかが分かるようになっている。しかし、2 人以上が同時に入退室した場合などは、1 人分の IC カードのログしか残らないため、完全な入退室ログとは言えない。したがって、欠落部分を補完する必要がある。欠落部分のある入退室履歴の補完を行うために、LAN への接続情報を用いる。この補完手法については 3.1 節で述べる。

Leaders' Saloon を利用する学生にはあらかじめ専用のタブレットが配布されており、それらが無線 LAN に接続されているかどうかを逐一検査して、いつ LAN に接続されて、いつ切断されたかをデータベースに自動的に記録するようになっている。

スケジュールデータには、その部屋で行われるミーティング等のイベント情報、また各個人の研究室などでのゼミなどのイベント情報が記録されている。

2.2. ユーザーインタフェース

提案システムのユーザーインタフェースを図 1 に示す。ブラウザ上で見ることができ、予測を行う時間を図 1 a に入力し、さらに予測を行う対象となる人を図 1 b で選択すると、計算された確率が出力される。

さらに、その周辺の時間の部屋に居る確率グラフ図 1 c で表示され、どの時間帯にその部屋に行けば、その人と会いやすいかということを知ることができる。

Future prediction of room usage situation based on card key logs and network connection data

[†]NOMOTO, Satoshi (nomoto@nagao.nuie.nagoya-u.ac.jp)

[‡]NISHIWAKI, Masayuki (nishiwaki@nagao.nuie.nagoya-u.ac.jp)

[‡]TAKESHIMA, Ryo (takeshima@nagao.nuie.nagoya-u.ac.jp)

^{†‡}MOTOYAMA, Yoichi,

^{†‡}SAKURAI, Eiiichi

[‡]NAGAO, Katashi (nagao@nuie.nagoya-u.ac.jp)

Collage of Technology, Nagoya University[†]

Graduate School of Information Science, Nagoya University[‡]

Large Data Based Modeling Team, Center for Service

Research, National Institute of Advanced Industrial Science

and Technology^{†‡}

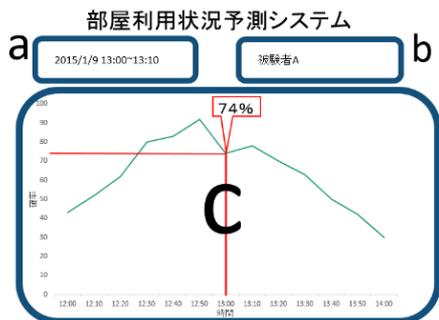


図 1、ユーザーインターフェース

3. PLSA によるデータ分析とベイジアンネットワークの生成

本研究では、PLSA（確率的潜在意味解析）[1]によって、IC カードログを元に作られる入退室履歴に含まれる人のセグメント化を行い、さらにベイジアンネットワークを用いて構造的なモデル化を行うことで、ある人物がある時間帯に部屋に滞在する確率を算出する。

3.1. 入退室履歴の補完

前章で説明した通り、IC カードのログには欠落部分があり、補完する必要がある。この補完には、タブレットの無線 LAN 接続情報を利用する。欠落した箇所の周辺の時間にその人の使用するタブレットの LAN への接続、または切断の時間を見て、近い時間のカードログを代わりに埋めるという作業を行う。

補完後に図 2 のような形でデータベースに保存されている。左から順にカード ID、入室日時、退室日時を表している。

85059116930915422	2014-10-19 16:35:28	2014-10-19 16:52:19
85059116930915422	2014-10-19 16:56:04	2014-10-19 17:49:49
85059116930915422	2014-10-19 17:50:23	2014-10-19 18:34:32
85059116930915422	2014-10-19 18:36:14	2014-10-19 21:19:15

図 2 データベースに保存される入退室履歴

3.2. PLSA（確率的潜在意味解析）

このようなデータによって、各個人の「いつ入室していつ退出した」といった行動履歴が作られる。被験者十数人の 3 か月分の行動履歴に対して、月ごとかつ曜日ごとに分類した個人と、24 時間を 10 分ごとに区切った時間をそれぞれ縦軸と横軸に設定した共起行列を作る。区切られた 10 分の間に少しでも滞在していれば 1 となる。その共起行列は、比較的疎であるため、有用なセグメントを抽出するために PLSA を用いる。これによって、時間と個人を同時にセグメント化することができる。

実験対象の個人 x_i が時間 y_j に部屋に滞在している状態を、個人と時間が所属する潜在クラスを z_k と表す。個人 x_i 、時間 y_j 、潜在クラ

スを z_k の関係を次式

$$P(x_i, y_j) = \sum_k P(x_i|z_k)P(y_j|z_k)P(z_k)$$

のように定義する。

3.3. ベイジアンネットワーク

構造的なモデル化を行うために、ベイジアンネットワークを用いる。ベイジアンネットワークは人間の行動予測など、不確実で絶えず変化するために数式で表現が困難なものを、うまく扱うことができる確率モデルの一種であり、複数の確率変数の確率を条件付確率で表している。

条件付確率は、ある事象が起きる確率変数を $P(a)$ 、求める確率変数を $P(b)$ とすると、事後確率 $P(b|a)$ で求めることができる。

PLSA によって作成されたカテゴリと、スケジュールデータのイベント情報を確率変数にしたものを用いてベイジアンネットワークを構築し、確率を算出した。

4. おわりに

本研究では、PLSA によるベイジアンネットワークを利用したある部屋の未来の状況予測を行うシステムを構築した。今後の課題として、どれくらいの精度で予測が行われているかを求めるための実験と、その評価を行ってきたい。

また、今回使用した入退室履歴は 3 か月分のみであり、PLSA による分析にかけるには比較的少ないデータであるため、あまりいい結果が得られなかったと考えられる。今後は、さらに多くのデータを用いた分析を行い、またそれを評価していきたい。

さらに、今回構築したシステムは単に入力した時間にその人が部屋に居る確率を出力させるだけのものであるのだが、今後の展望として、アンケートなどを実施し、どのような部屋状況、例えば人が多いときまたは少ないときに、その部屋で作業を行いたいかなど、個人の異質性なども考慮していきたいと考えている。

参考文献

[1]村山敬祐・吉田真・山下真一郎・廣川典昭・川島健佑・善甫啓一・桜井瑛一・本村陽一：確率的潜在意味構造モデルを用いたビッグデータ活用事例～ベイジアンネットワークと PLSA 利用によるデータモデリング～，人工知能学会 第 96 回基本問題研究会 SIG FPAI