

# 通信履歴データを用いたスケジュールリング形成のデータマイニング

## Data Mining Analysis on the Relations between Tele-Communication and Activity Scheduling

佐々木邦明<sup>\*1</sup>, 西井和夫<sup>\*2</sup>, 鈴木篤<sup>\*3</sup>

By Kuniaki SASAKI<sup>\*1</sup>, Kazuo NISHII<sup>\*2</sup> and Atsushi SUZUKI<sup>\*3</sup>

### 1. はじめに

Activity Based Approach では, 交通行動はある活動に付随した活動とされているが, その活動の多くはコミュニケーションを第一の目的とすることが多い. このことから情報通信と交通の関係性が古くから指摘されてきた<sup>1)</sup>が, 具体的なデータ解析を通じた説明はそれほど進んでいない. その理由として, 個人の通信に関するデータの収集が容易ではなかったことや, 通話回数などよりその中で取り扱われる情報が問題であるのに対して, 情報内容の計量化とそのデータ収集法が存在しないなどの理由があげられる. また, 活動や交通行動との関係も, その関係が代替だけでなく, 補完関係などもあり交通行動の分析手法としてよく用いられる選択分析では全てを取り扱うことが難しいため, 分析の方法論としても問題があった. その一方で, 固定電話からファクシミリの時代を経て, ページャー, 携帯電話, 携帯電話によるメールサービスと, 中心となる形態を進化させつつ大きく変貌してきた. 現在では携帯電話の契約数が 8000 万を超え, その普及も飽和過程に入り, 通話数(トラフィック数)も固定電話と同程度にまで伸びてきた. 現在では携帯電話は生活のあらゆるところで利用され, 携帯電話の利用を前提とした行動が増加してきている. このような状況下で, 交通への関係が指摘されてきた通信との関係性を携帯電話の利用を中心として分析する必要性が高まっている

2). しかし, 携帯電話や携帯メールの利用と行動との関係は, 通信の位置や時間も自由であるため, 固定電話と比較してより複雑さを増していると予想される. このようなことを考慮した, モバイル通信と行動との関係を模式的に示したのが図 1 である. 左の現実空間での分析は従来型のアクティビティパスだけの分析であり, 図 1 には仮想空間でのコミュニケーションパスが新たに加わっている. 携帯電話を通じて認識される仮想空間は, 距離の概念がなくなるかわりに, コミュニケーションロスが発生する空間である. 仮想空間でのコミュニケーションは現実空間と相互に関係しながらも, 異なった特性を持ち従来型の分析手法では理解が困難なものがあると考えられる. そこで本研究では, 「仮説構築 データによる検証」のプロセスではなく, データマイニング手法を用いてデータ間の関係を帰納的に導き出すアプローチをとることとする.

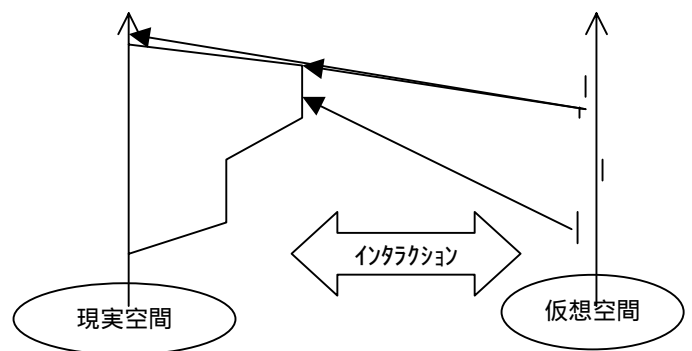


図 1 モバイル通信を考慮した時空間

また, これとは別に, 近年交通データや行動データの収集技術が各段にデータの精度を向上しつつある. 従来型のアンケート調査などの思い出し記述調査に頼らずに, 通信技術を利用した位置測定による詳細な経路情報などが大量に得ることができると

キーワード: モバイル通信, データマイニング, アクティビティ分析

\*1 正員: 山梨大学助教授 大学院医学工学総合研究部

\*2 正員: 山梨大学教授 大学院医学工学総合研究部

\*3 非会員: 三立プレコン

連絡先: 〒400-8511 山梨県甲府市武田 4-3-11

Tel.055-220-8671 sasaki@yamanashi.ac.jp

である．このように現実空間での情報収集精度と量の向上は，より本源的な行動の理解への足がかりとなる可能性を秘めている．つまり現実空間での行動データの精度向上は，仮想空間でのコミュニケーションの効果を無視して分析することの限界を示すことになると考えられる．そこで本研究は従来型の調査だけでなく，モバイル通信の通信履歴を詳細に調査することとする．これは思い出し型調査によるのであるが，モバイル通信履歴はデバイス内に履歴を保存する機能が存在するため，非常に精度高く収集可能であると考えられる．

## 2．利用データ

利用したデータは，2002年秋に行われたSCAT (Survey of Communication, Activity and Travel)である<sup>3)</sup>．その調査目的は，モバイル通信が人々の生活行動のどのような局面に影響を与えているかを把握することである．とくにモバイル通信の詳細に焦点を当て，生活・交通行動の予定取り決め(意思決定)に関する内容や予定変更等の特徴，さらにはその通信が以降の被験者の生活行動に与える影響について調査を行っている．調査の実施方法は，被験者に対してアンケート(記述式)を行ったが，多くの記入が必要であるため，直接配布，直接回収形式で実施した．対象としたのは流通科学，京都，大阪市立，東京，山梨の各大学の学生で，回収総数は152票であった．調査の具体的項目は以下のとおりである．

- ・ 1週間予定調査
- ・ 活動・移動調査(7日間)
- ・ 電話(固定・携帯とも含む)，メール(PC・携帯とも含む)，情報サイト利用など情報収集・コミュニケーション活動調査
- ・ 情報通信に関する意識調査

の4項目である．

## 3．モバイルコミュニケーションと意思決定構造

本研究で目的とするのモバイルコミュニケーションと行動の意思決定構造の関係を分析するにあたって，モバイルによる情報通信の特性を考慮する必要がある．モバイル通信の主な特性とは

- ・ 個人の独立な意思決定でない場合が多い
  - ・ 動的な情報を小さなタイムラグで得られる
  - ・ 空間の認識が実空間とは異なるように感じられる
- などである．つまりこれらのモバイル通信の特性をこうりよするとき，モバイル通信利用を前提として成立する活動の意思決定例としていくつかの例をあげることができる．

時間・空間を明確に特定しない活動の配置が行われ，意思決定者同士の即時的コミュニケーションによって，動的に時空間上のスケジュールを調整する．

常に連絡が可能なため，非計画的な行動パターンが可能になる．

これらの事例から考えると，その意思決定構造は，従来型の個人の効用最大化に基づいた，一日のスケジュールの最適化という考え方を当てはめることは困難に思える．そこで，これらの行動を理解するためには，意思決定がどのように行われているかを，データから帰納的に発見することが必要であると考えられるため，いかにデータマイニング手法を用いた分析を展開する．

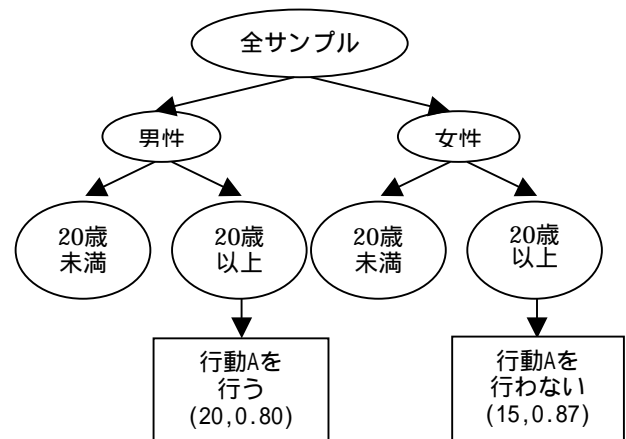


図 1 デシジョンツリーの例

## 4．デシジョンツリーを用いたデータマイニング

データマイニングとは，マーケティング分野で発展してきた手法であり，一般的には大規模なデータの集まりから，より有効なパターンや隠された知識を発見するための手法と定義できる<sup>3)</sup>．この中で，本研究では，ルール帰納と呼ばれるアプローチ<sup>4)</sup>の適用を行う．

このルール帰納においては，大規模データの集合の中に存在する意思決定ルールをデシジョンツリー

の構成によって帰納的に導出できる。また、ディシジョンツリーの構成にあたっては、ある結果に対してそれを説明するために役立つ諸要因がその影響力の強い順に選択される。このとき、諸要因(条件)と結果との関係は、図-1に示すツリー構造を構成する。すなわち、木の根元にあたるルーツ(root)からいくつかの枝が分岐・成長するように、ある結果を支配するいくつかの要因(条件)間の一連の階層的な構造として表現できる。

図-1では、ある結果をツリーの最上段に位置する親ノード(とくに、ルーツノード)とし、それを支配する要因(条件)が子ノードとして連結されている。また、この決定木が下方へ成長するかどうかは利得比基準<sup>5)</sup>に基づき判断される。決定木の最下段では、分岐がこれ以上進まないことからターミナルノードとして定義される。ここで帰納されたルールとは、この決定木における最上段の親ノードから分岐を経てターミナルノードに至る要因(条件)の連結関係を指す。

## 5. ディシジョンツリーの実証分析

ここでは、前述のディシジョンツリーをSCAT調査データに適用することにより、調査日当日に予定された活動の予定変化(具体的には、予定活動に関する変更、確認、追加、休止の4つの結果)の発生に関するルールを分析する。また、求められたルールに基づき、基準変数(ここでは予定変化のタイプ)との関係を明らかにするため、着目する要因による予定変化への影響の程度をみることにする。

表-1は、本分析で用いる変数の一覧であり、個人属性、通信に関する属性、生活行動に関する属性、調査場所の立地条件、の合計4つに分類される。なお、生活行動に関する属性の中で、活動内容は、家事、交際、自由、学習、移動の5つに集約し、また活動時刻についても、3-11時、12-18時、19-21時、22-27時の4時間帯に集約している。

全体のディシジョンツリーは分岐数が大きく広がったものとなったため、一つのストリームを図-2に示した。このルールでは、予定活動が「交際の活動」が親ノードとして登場し、7つの要因(条件)からなる。このルールが満たされる場合において「通信内容が予定の新規であること」の確率は75%となっている。

表 - 1 分析に用いた変数

変数大分類	具体的な変数
個人属性	性別、学年、住居区分、免許保有、車の利用有無
通信状況に関する属性	通信手段、通信相手、通話回数、メール回数、通信時間、通信時の活動、受信/発信、予定活動開始時刻(予定と実際)、通信時刻と予定時刻との差
生活行動に関する属性	学習・仕事の活動時間、交際の活動時間、自由活動時間、家事的活動時間、睡眠時間、移動の時間、自宅関連の移動回数、自宅外関連の移動回数
調査場所の立地条件	調査場所の立地条件が都心型/郊外型
通信時の予定活動に関する内容	予定活動に関する『新規』/『変更』/『確認』/『休止』

このルールを構成する条件には「対象の活動内容に関わるもの」と「個人の属性に関わるもの」および「個人の(通信)活動にかかわるもの」があるため、例えば、個人に関わるものは、「そのような特性を持つ意思決定者が」と解釈するとわかりやすい。つまりこのルールは【車を保有しない個人が、その日自由活動総時間が270分以内で自宅をベースとしない移動が5回以下、メールの回数が13回以下のような状況において、学習・仕事的内容の活動を行っているときに、その時点から30分以内に新規の交際の活動を通信によって取り決める可能性が75%である】と解釈できる。

次に、各ルールにおけるターミナルノードとなる要因(条件)がどのような形で予定変化に関わっているかを明らかにするために、それぞれのルールについて各要因(条件)ごとの通信内容の構成割合の変化を求めるとし、その結果を表-2に示す。

これらの表より、2つのルールで、それぞれ要因(条件)が加えられたときの各条件による予定変化の割合を讀取ることができる。なお、表中では、各条件が加えられたときに最も予定変化の構成割合が大きく増減したケースを識別できるようにセルに網掛けをつけた。その結果、予定の新規の割合が最も増加したのは、条件7(通信時刻と予定活動開始時刻との差が30分前以内)である。

これらの結果より、このツリーを構成する行動に対して、通信による予定活動の変化に最も関連性の強い条件を特定化できた。すなわち、通信時刻が予定活動開始時刻との時間差が30分前以内のときに予定活動の『新規』

表-2 ルールAにおける条件ごとの活動変化の割合

	変更の割合	中止の割合	新規の割合	確認の割合
全サンプル	0.18	0.05	0.32	0.45
条件1の後	0.16	0.06	0.38	0.40
条件2の後	0.18	0.01	0.38	0.43
条件3の後	0.21	0.01	0.30	0.48
条件4の後	0.09	0.01	0.35	0.55
条件5の後	0.04	0.04	0.35	0.57
条件6の後	0.06	0.06	0.41	0.47
条件7の後	0.00	0.00	0.75	0.25

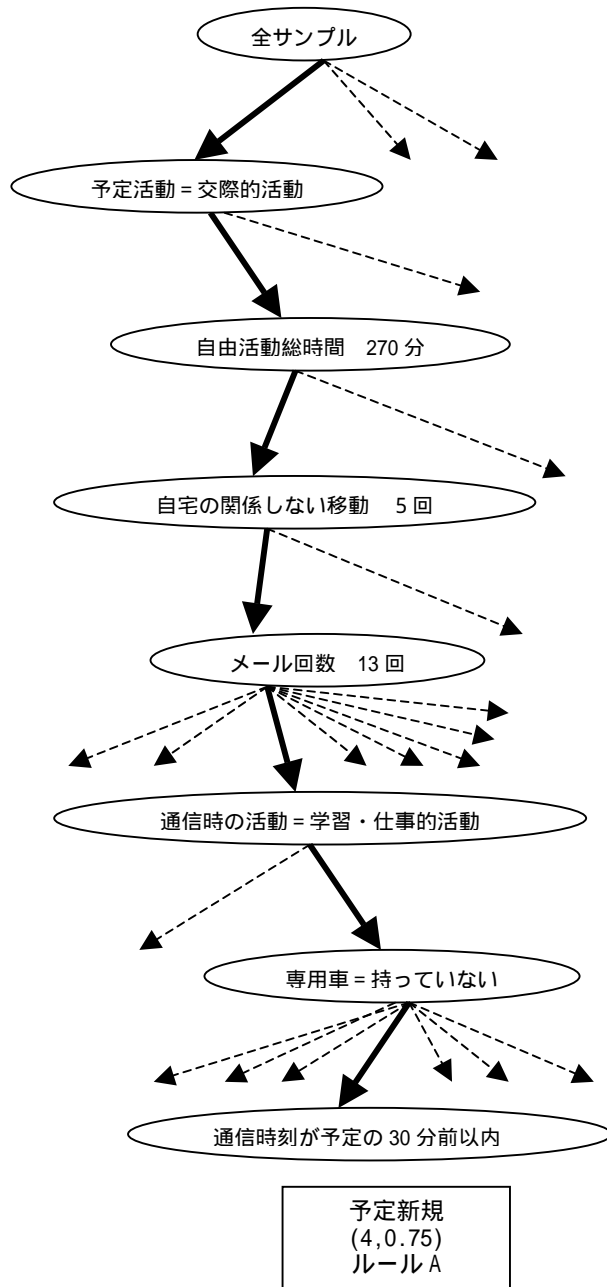


図 2 モバイル行動予定のツリーの例

に関する通信がなされること、例えばルールAにおいて条件7の通信時刻と予定活動開始時刻との時間差が30分前以上のとき、予定活動の変化の割合を求めると、こ

の場合は予定活動の「確認」の構成割合が大きくなる。これは、通信時刻のタイミングが予定活動の新規か確認の判断基準になることを示唆している。

## 6. おわりに

SCAT調査データに関する基礎分析結果より、予定活動については、通信内容が「予定の新規」の場合にはその予定活動が「社交」の場合が多く、また通信内容が「予定の確認」の場合には予定活動が「仕事」の場合が多いことがわかってきた。また、活動の時間帯や場所、通信時点の活動形態などについても予定活動に関する通信内容との関連性があることが確認されていた。そこで本分析では、通信と生活行動上の予定活動の変化との間に何らかのルールが存在するのかを明らかにするため、データマイニングのディジションツリー構成法の適用を試みた。

その結果、予定活動が社交を含む交際的活動の場合、生活行動特性ならびに通信状況に関する7つの条件からなるルールが導出でき、そのルールが満足される場合には通信内容の75%が予定活動の「新規」に関わるものであることがわかった。また、このルールのもとでは、「通信時刻と予定活動開始時刻との時刻差が30分前以内か30分前以上か」によって、予定の「新規」が「確認」が規定されていることがあきらかになった。

本稿では通信によるスケジュール構成を対象として、ディジションツリー分析を行い、その一部だけを取上げて報告を行った。得られたディジションツリーは非常に大きなものであり、報告した以外にも多くのルールが導出されている。それらの詳細についてはまだ分析途中であるが、どのような状況や条件下でどのような行動が行われる傾向があるのかが明らかになってきた。これらの分析を続けることで、通信を利用した意思決定にはルールがあるのかを明らかにできるものと考えられる。

## 参考文献

- 1) IATSS633プロジェクトチーム: 交通と通信の代替・補完関係 - 通信ニューメディアによる交通の代替性 - , 国際交通安全学会誌, Vol.10, No.3, pp.29-37, 1984.
- 2) 近藤勝直: 交通・流通・通信の競合と連携, 土木計画学研究講演集, No.27 (CD-ROM), 2003(予定).
- 3) 西井和夫, 佐々木邦明, 山田宗男: モバイル通信が生活交通行動に与える影響に関する研究, 土木計画学研究講演集, No.27 (CD-ROM), 2003(予定).
- 4) マイケル ベリー, ゴードン リノフ: データマイニング手法, SASインスティテュートジャパン, 江原淳, 佐藤栄作共訳, 海文堂, 1999.
- 5) 豊田秀樹: 金鉢を振り当てる統計学, 講談社, 2001.