

研究室入退室システムによるオンライン時間と在室時間の相関関係

杉田 和敏

Kazutoshi Sugita

東海大学情報メディア学科

Dept. of Informedia Technology

Tokai University

菊池 浩明

Hiroaki Kikuchi

東海大学情報メディア学科

Dept. of Informedia Technology

Tokai University

Abstract: 本研究では、電子的な活動時間と現実的な活動時間の違いと作業効率、および生産性との関係を明らかにする。研究室のメンバー 13 人の入退室時間、メッセージへのログイン時間、ログアウト時間を 2ヶ月に渡って記録し、また同メンバーに研究の状況などのアンケートをあわせておこなった。それらの情報に決定木学習を適用し、各活動時間と卒業研究の関係を調べた。その結果、研究室の在室時間と研究の進捗度には関係があることがわかった。また、個人の嗜好や性格に関するアンケートと組み合わせることによっていくつかの論理的規則が成立することを明らかにした。

1 はじめに

近年、インターネットの普及、高速化に伴い、電子メールやインスタントメッセージ (IM) などといったインターネットを使ったバーチャルなコミュニケーション方法が普及してきた。それに伴い、それらを活用した在宅勤務や e-learning などが増えてきている。ネットワークによる遠隔活動は会社や学校へ行かず自宅で作業を行えるため通勤、通学時間の節約となったり自分の時間に作業が出来るという利点があるが、同時に現実での対面がないため意思の伝達に齟齬がおきる可能性も指摘されている。そこで、本研究では在宅勤務や e-learning が本当に有効な作業、学習環境であるかどうかを知るため、電子的な活動時間と現実での活動時間と作業効率、生産性との関係を明らかにすることを目指す。

ところが、オンラインや実際の作業時間をどのように計測するかがひとつの課題である。そこで、無線 IC タグと IM エージェントを組み合わせることにより、入退室状況およびメッセージログイン状況や、実際に作業している状態をリアルタイムで共有できるシステムを開発した [1]。本稿では、本システムのログデータを利用し、研究室での在室時間を現実での活動時間、メッセージのログイン時間を電子的な活動時間、そして作業効率を研究の進捗度として計測を行った。評価結果が視覚的にわかりやすい決定木評価を行った結果を報告する。

2 要素技術

2.1 RFID

RFID (Radio Frequency Identification) とは、微小な無線チップにより、非接触でモノやヒトを識別・管理することができるシステムである。タグは IC チップとアンテナから構成される。[12] はコイン型の電池を内蔵しないパッシブタグで、複数タグ同時読み書きをサポートしている。

RFID の特徴の中で、非接触で複数同時読み書きが可能であることは最大のメリットであるが、プライバシー保護という観点からすれば、気付かないうちに個人情報を読み取られ、行動を追跡されるなどの問題がある。

2.2 インスタント メッセージ

インスタントメッセージは互いに登録している仲間のオンライン状態の通知と、オンラインであればチャットなどのリアルタイムサービスを提供するシステムである。これを実現するアプリケーションソフトを IM クライアントと呼ぶ。

[13]によると、主要な 4 つのサービス (MSN, Yahoo!, ICQ, AOL) のユーザ数は 250 万人を超え、IETF のプロトコル標準化やオープンソースの動きなどが盛んである。

2.3 入退室管理システム

2.3.1 概要

本システムは在室状況などを管理しているシステムである [1]。本システムによって、次のサービスが提供されている。1. ユーザの入退室情報の管理 (RFID タグ)、2. リアルタイムな伝達、3. 各種メッセージ交換・問い合わせ。

本システムは、主に IM エージェントと入退室管理システムの二つから構成される。図 1 にシステムの構成を示す。本システムは Windows 上の Java アプリケーションで実現されており、RFID リーダとは Java Communications API [15]、IM サーバとは MSN プロトコル [16]、Web サーバとは Samba によるファイル共有で通信している。

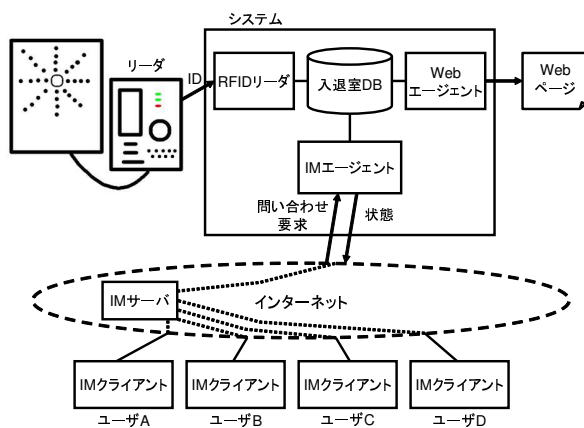


図 1: システム構成

2.3.2 インスタントメッセージングエージェント

IM エージェントは、他のユーザと同じ一般ユーザとして実現している。IM エージェントに話しかけてコマンドを送信することによって、様々なサービスが提供される。

表 1 に主なコマンドとその機能説明を示す。

表 1: IM エージェントのコマンド	
コマンド	機能
ヘルプ	コマンドとその説明
メッセージ	メッセージの代弁
だれ?	現在の在室者と共有備品の貸し出し状態
ゲーム	4 択クイズ出題
退室	遠隔操作の退室

メッセージ代弁機能は複数人にまとめてメッセージを同報することや、オフラインのユーザにメッセージを送ることができる。この場合、メッセージはユーザがオンラインになった時点で送信される。退室機能は退室時にタグの処理を忘れた場合のために用意されている。IM を使用した遠隔操作での退室処理である。

2.3.3 入退室管理システム

本システムは、ユーザが入退室時リーダーに自分の ID が格納されたタグ読み込ませることにより、データベースの情報の退室状態と入室状態を反転させる。その際に、ユーザ情報のデータベースから対応する情報を検索し、入退室時間と氏名、在室者の人数を IM エージェントを通じて、すべてのユーザに送信する。同時に Web ページのユーザー一覧ページへも反映させることで、外出先などからでも携帯端末などへ情報を提供する。オンラインユーザには在室者を問い合わせるコマンドをサービスする。

3 評価実験

3.1 実験目的

研究室での在室時間、メッセージへのログイン時間、研究の進捗状況の関係を調べることにより、電子的な活動時間と現実での活動時間と作業効率、生産性との関係を明らかにする。

3.2 実験方法

東海大学情報メディア第 8 研究室の 4 年生、および院生 13 人を対象として、入退室管理システムを運用し、2006 年 6 月 6 日から 8 月 10 日の約 2ヶ月間、研究室の在室時間、メッセージへのログイン、ログアウト時間を記録する、同メンバーに表 2 にある研究の状況などのアンケートをあわせておこなう。

ただし、在室時間は退室処理をし忘れた場合を考慮し、入室時間と同じ日に退室処理を行ったときの在室時間のみを集計した。

3.3 実験結果

3.3.1 在室時間とログイン時間との比較

全員の合計と代表的なユーザの結果を図 2 から図 4 に示す。また、統計量を表 3 に示す。

表 2: アンケート項目

	質問内容	YES の数
Q1	自分の時間を何より大事にする	3
Q2	学校が遠いと感じる	10
Q3	プログラム作りより発表のほうが得意だ	5
Q4	正直なところ研究が遅れている	9
Q5	自分の研究に自信を持っている	6
Q6	とにかく研究を早く終わらせたい	13
Q7	クーラーの効いているところに行きたくなる	6
Q8	騒音が気になるほうだ	9
Q9	友人は多いほうだと思う	3
Q10	積極的に人の世話をするタイプである	6
Q11	1日の行動予定をあらかじめ決めておく	7
Q12	コソコソやるよりひらめきを重視する	8
Q13	あれこれ考えるよりはまず実行してみる	4
Q14	一人の時間を大切にしよう	12
Q15	人にはいえない大きな目標がある	7
Q16	Web 検索は Google を使ってる	7
Q17	冥王星が惑星から外れるのには反対だ	7
Q18	ベジータよりナッパのほうが好きだ	6
Q19	ツンデレな人と付き合っていく自信がある	8
Q20	フィンガーマウスは使いにくい	10

図 2 より, 研究室へ直接来るよりはメッセージャーでオンラインで活動している時間のほうが一般的であることがわかる。

図 3 と図 4 より, 研究室よりも在宅でオンラインが主な活動時間であるユーザ A と対照的に研究室へ滞在している時間の長いユーザ B の違いが観測できる。B は, メッセージャーにオンラインする時間は短く切り上げている。ただし, 研究室にいてもオンラインになっているケースが多いことにも注意が必要である。

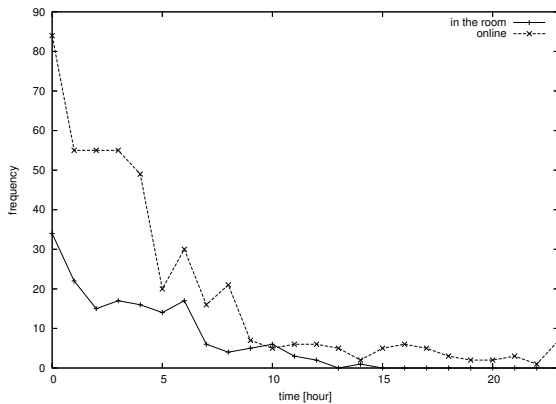


図 2: 在室時間とログイン時間との比較 (合計)

表 3: 在室時間およびログイン時間の統計量

	平均	標準偏差	最大値
在室時間	4.31	3.21	14.1
ログイン時間	5.06	5.15	23.97

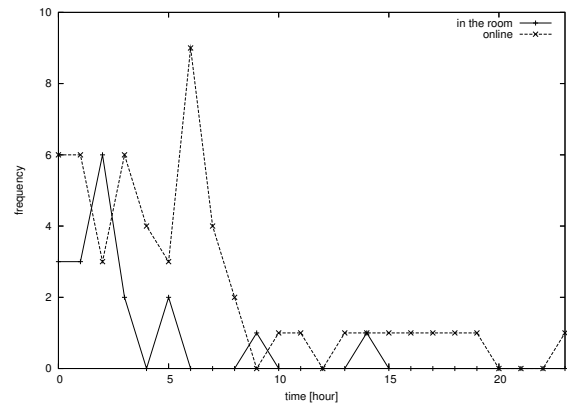


図 3: 在室時間とログイン時間との比較 (ユーザ A)

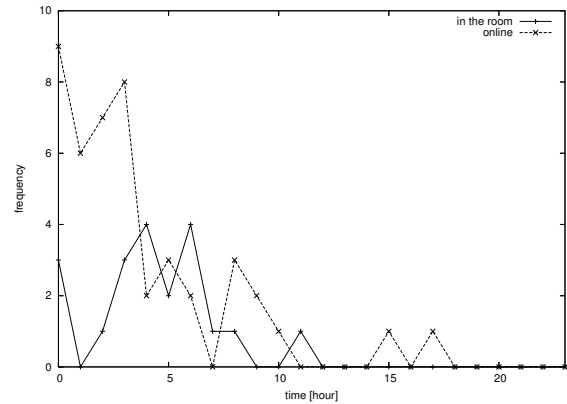


図 4: 在室時間とログイン時間との比較 (ユーザ B)

3.3.2 在室人数とオンライン人数の推移

実験期間中の代表的な日の結果を図 5 と図 6 に示す。

図 6 は典型的な 1 日のパターンであり, 朝 10 ごろから 16 時のピークまで研究室に滞在し, 夜 22 時に全員が帰宅してから深夜までオンラインにて活動している様子がわかる。一方, 図 5 の 6 月 20 日は火曜日なので研究室でゼミがあり, 午後 13 時までにシャープに在室人数が最大化している。(午前中は誰も来ていない)

4 決定木評価

4.1 決定木ツール

今回の評価では決定木を導き出すツールとして並木翼氏の作成した ID3E を使用した [8]。本ツールはユーザビリティ高い決定木ツールを目指したもので, csv ファイルの利用, 視覚的にわかりやすいグラフィカルな出力などの特徴がある。

なお本ツールを利用するに当たって決定木が複雑になるのを防ぐため, 重み付け枝狩りの係数 $M = 2$ とし

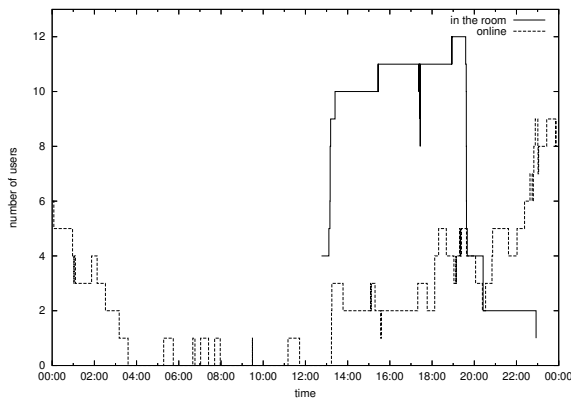


図 5: 在室人数及びオンライン人数 (6/20)

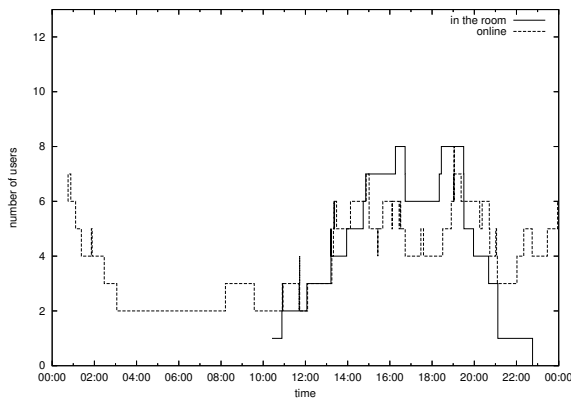


図 6: 在室人数及びオンライン人数 (7/27)

て評価を行った。これにより 2 票以下のノードより下の枝は刈られることとなる。

4.2 決定木アルゴリズム

C4.5 は Quinlan によって提案された決定木学習のアルゴリズムである [3]。C4.5 は、複数の属性値からなる事例の集合と、事例集合の中の意味のある部分集合を表すクラスを入力として取り、クラスを表現する属性から成る論理式を表す論理決定木を出力する。情報量利得を最大化する属性を繰り返し選ぶことで、より「シンプルな」木を出力するように設計されている。「シンプルな」木とは、ノード数が少なく、深さが浅いことを意味する。アルゴリズムは決定的に動き、最適属性の選択も貪欲法で行われるため、大規模なデータに対しても効率的に ($O(n \log n)$) 走る。

ID3 などの決定木アルゴリズムに比べて、次の点が特徴的である。

- 連続値の利用が許されている。
属性に連続値があった時は、全ての閾値について 2

値化を行って情報量利得の意味で最適な閾値を定め、その閾値について、ある条件をノードとした木を生成する。

今回のデータでは在室時間およびオンライン時間を連続値として利用することとした。

- 不明値の扱い。
C4.5 では属性値が不明なとき、属性に重みを与え、該当する属性の生気確率に応じて部分木への分類を行う。例えば、9 割りが Yes の属性値が不明であっても、0.9 の重みで Yes と見なした扱いをする。また、不明値を独立した値と見なして処理を行うため、不明値がある属性はない属性に比べて選ばれにくくなる。
- 利得比による分類。
属性の取り得る値の数が一様でないとき、従来のアルゴリズムでは公平な比較が出来ていなかった点を、分割情報量を導入して改善している。
- 誤差に基づく枝刈り。
決定木学習は事例を誤差なく識別するので、決定木はしばしば複雑になり過ぎてしまい、可読出来ない。そこで、C4.5 では元の決定木に対して誤り率をある定められた上限まで下げることが許すことで、部分木の削除や置き換えを行う。

4.3 評価データ

決定木評価にあたって使用したデータは、利用者に答えてもらったアンケート 20 問に図 7 にある各利用者の合計在室時間、および合計ログイン時間を加えたものである。

しかし、クラスとなる属性に連続値を適用することはできない。そこで、在室時間、ログイン時間をクラスとする場合はそれぞれ時間が長い方から上位 4 名を「長い」、下位 4 名を「短い」、その他 5 名を「普通」とする。

4.4 評価結果

ID3E による評価結果の決定木をを図 8 から図 11 から示す。

4.5 考察

図 8 は Google 利用の有無が入室時間を左右する大きな要因であることを示している。Google は多くの人が使っている一般的な検索エンジンであるが、他の意見

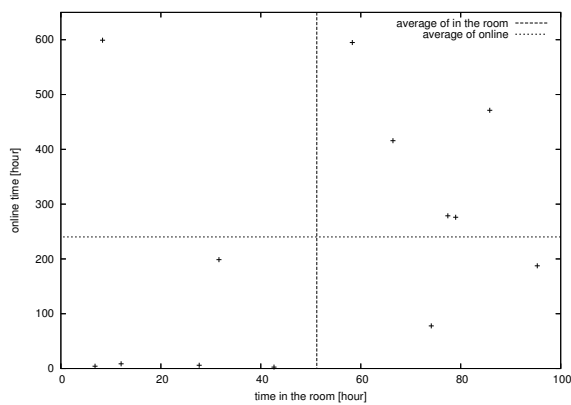


図 7: 在室時間およびログイン時間の散布図

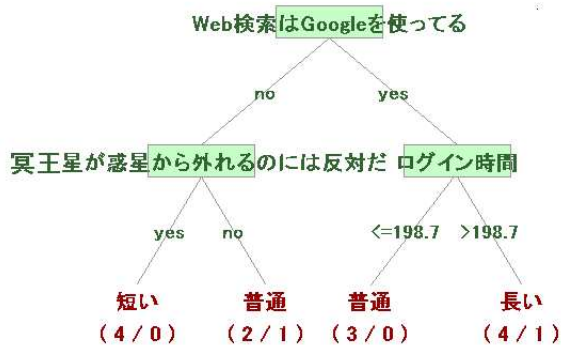


図 8: 在室時間の決定木

に惑わされず自分の使いやすいものを使う。そんなアウトロー的な性格を持つ人が研究室にもあまり来ないのではないかと推察する。

図 9 は発表が得意な人はログイン時間が短いという理論関係を表している。ログイン時間はパソコンを使っている時間と比例するので、ログイン時間が少ない人は一人でパソコンに触れている時間も少なく、従って、社交的な性格で、結局、発表も得意になると推測される。また、逆にログイン時間の長い人は、パソコンに慣れ親しんでいるためプログラム作りが得意と考えられる。

図 10 では入室時間が多く (77.4 時間以上)、または、ログイン時間が短い (4.3 時間未満) 人は卒業研究も進んでいるという結果になり、研究室は卒業研究を進めるのに役に立つ場所であることがわかる。この関係を在室時間とログイン時間の散布図として図 7 に示す。

図 11 では、積極的に人の世話をを行い、なおかつログイン時間が少ない人が友人が多いことが示されており、ここでも図 9 と同様にログイン時間が社交性に結びついている。

以上より、明らかにされた在室時間、オンライン時

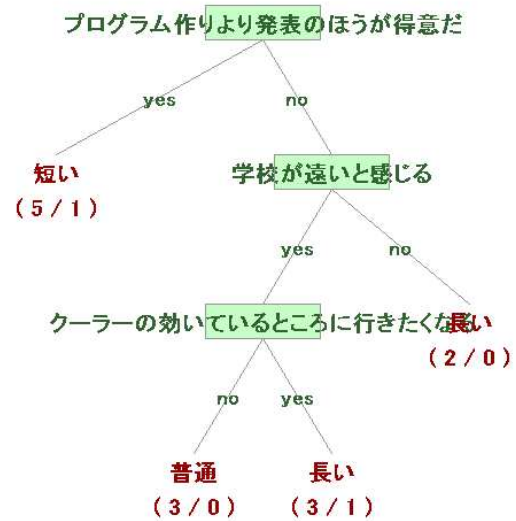


図 9: オンライン時間の決定木

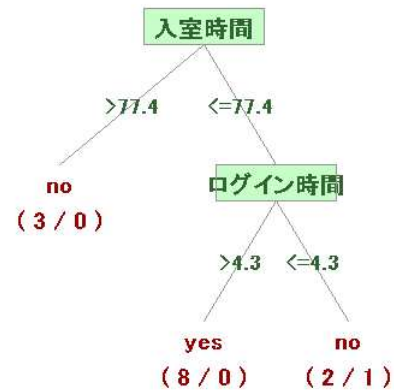


図 10: 正直なところ研究が遅れているの決定木

間と活動成果に関する規則を次に示す。

- web 検索に Google を使用しているかつログイン時間が 198.7 時間より多いならば在室時間が長い。
- プログラム作りより発表のほうに得意ならばログイン時間は短い。
- 在室時間が 77.4 時間以上、もしくは、ログイン時間が 4.3 時間未満の人は卒業研究が進んでいる。
- 積極的に人の世話をを行うかつログイン時間が少ない人は友人が多い。

5 結論

本実験より、作業効率、生産性と現実での活動時間は大きな影響があることが示された。現実で対面しコ

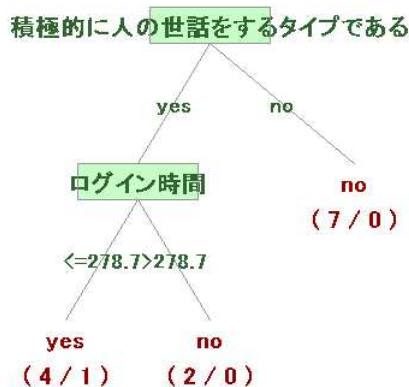


図 11: 友人は多いほうだと思うの決定木

コミュニケーションの取れる環境は作業効率，生産性にとって大切な場所である。また，電子的なコミュニケーション手段に頼ることにより，友人が少ない，発表があまり得意ではないなどの内向的な性格の傾向が見られるようになることもわかった。

謝辞

本研究を遂行するにあたって多大な指導，協力をいただきました中村雄一氏に深く感謝の意を申し上げます。

参考文献

- [1] 中村，菊池，RFID とインスタントメッセージングエージェントによるリアルとバーチャル空間の融合，情報処理学会，CSEC 研究報告，CSEC-032，pp. 25-30，2006.
- [2] 馬野，ID3，日本ファジィ学会誌 vol.6，pp. 502-504，1994.
- [3] J. Ross Quinlan，C4.5: Programs for Machine Learning，Morgan Kaufmann pub.，1993. (日本語訳，古川他，「AI によるデータ解析」，トッパン，1995).
- [4] 高崎，米山，菊池，中西，Web ページの分類におけるキーワードの抽出について，第 7 回曖昧な気持ちに挑むワークショップ，pp.9-14，11 月 2002.
- [5] 米山，高崎，菊池，中西，Web 識別問題における 2 クラス，多クラス決定木の比較，情報処理学会，DPS 研究会，pp.111-114，2 月 2003.
- [6] 松本 他，形態素解析システム「茶筌」version 2.3.3 使用説明書，8 月 2003.
- [7] 徳永，言語と計算 5 情報検索と言語処理，東京大学出版会，1999.
- [8] 並木，菊池，ユーザビリティの高い GUI ベースの決定木学習ツールの開発，情報処理学会全国大会，3 月 2005.

- [9] Tara Calishain, Rael Dornfest, GOOGLE HACKS, O' Reilly Japan, 2003.
- [10] 菅野，岩山，“インスタントメッセージ技術の最新動向とその応用”，情報処理学会研究報告，QAI，pp. 47-54，2002.
- [11] 根本，遠山，“閲覧履歴に基づく情報検索の相互支援”，電子情報通信学会，DEWS2004，2004.
- [12] “電磁誘導方式 RF-ID システム 形 V700 シリーズユーザーズマニュアル”，オムロン株式会社，2001 年 7 月.
- [13] 菊池，多田，中西，“管理者に対して秘匿性を保証したセキュアインスタントメッセージングプロトコル”，情報処理学会論文誌，vol44，pp. 2042-2050，2003.
- [14] MSN Messenger，<http://messenger.msn.co.jp/> (2006 年 10 月参照)
- [15] Java Communications API，<http://java.sun.com/products/javacomm/> (2006 年 10 月参照)
- [16] JMSN，<http://sourceforge.net/projects/jmsn/> (2006 年 10 月参照)

連絡先

杉田 和敏
 東海大学電子情報学部情報メディア学科菊池研究室
 〒 259-1292 神奈川県平塚市北金目 1117
 TEL: 0463-58-1211
 FAX: 0463-59-4014
 E-mail: professor@cs.dm.u-tokai.ac.jp