

# ウィンドウ遷移ログを用いた 従業員のユーザー行動モデルの時系列変化分析

齋藤 良平<sup>1,a)</sup>

**概要：**ユーザー行動モデルは単位時間内に使用されたアプリケーションの組み合わせとその遷移で PC 操作を表現する確率的モデルである。このモデルを企業内 PC から取得されたウィンドウ遷移ログを用いて構築することで業務上の操作の特徴を抽出することが可能である。本稿では、ユーザー行動モデルがどの様に変化していくかを可視化する手法を述べ、実際の企業から取得した約 4 年間分のウィンドウ遷移ログを用い、ユーザー行動モデルの変化を分析することで、従業員の業務と PC 操作の関連を検証した。

## 1. はじめに

近年の IT の発展により、企業における PC の使用はごくありふれた光景となっている。PC の使用はメールや資料作成に始まり、ブラウザを用いたイントラネットシステムの利用や業務アプリケーションの利用等多岐に渡り、業務の多くを PC を用いて行う従業員も少なくない。しかしながら、PC を用いた業務の評価は成果物をもって判断されることが多く、その過程が評価されることは少ないと言える。これは、PC 上の操作は他人が把握することが難しく、個人作業であることが多いためである。同じ業務であっても、それに関わる PC 上の操作は人により異なり、また、同一人物であっても習熟度や環境によって大きく異なることが考えられ、この違いを客観的に判断できるならば業務の生産性や効率を推定する手掛かりとなる。

一方で、これら企業内の PC に記録されている重要情報を狙う外部からの不正アクセスや従業員による内部情報漏えい対策のために、多くの企業でセキュリティ対策ソフトウェアの導入が行われている。これらのソフトウェアではインシデント発生時の状況のトレースや従業員の不正な行動の抑止のために、PC 上の操作ログを記録する機能を持っている。この操作ログの一つとして操作したウィンドウをトレースするためのウィンドウ遷移ログがある。

一般的な企業では従業員は担当する業務が決まっており、その担当業務において使用するアプリケーションは限定的である。このため家庭における PC の使用に比べて PC の使用方法が限定的であると言える。

本研究では、企業内で取得したウィンドウ遷移ログに着目し、筆者らが提案したユーザー行動モデル [1] を約 4 年間分のログを用いて構築、分析する手法を示し、ユーザー行動モデルの変化と業務の関連を検証した。ユーザー行動モデルは単位時間内に使用されたアプリケーションとその遷移を用いて業務の作業を表現することから、同一の従業員が同一の業務を行っていても異なるモデルが構築される。このような違いが発生する要因を分析できるならば、生産性や作業効率の推定に有益な手掛かりとなる。これを行うためには、ログと業務の間に強い関連性が必要であり、提案する手法でこの関連性が導けることを示す。

また、ウィンドウ遷移ログの様に低レベルなログは短期間においても膨大に記録され、その傾向やパターンを直感的に人が判断するのは困難である。ユーザー行動モデルを用いてログの情報を抽象化、可視化することで長期間の大量のログに対する分析を容易にすることが可能であることを示す。

本稿では次節で関連研究について述べ、第 2 章で分析対象とするウィンドウ遷移ログとユーザー行動モデルの概要を述べる。第 3 章ではユーザー行動モデルの時系列変化を可視化する手法と実際に企業から取得したログを用いた分析結果について述べる。

### 1.1 関連研究

ログとして代表される Web のアクセスログやモバイルデバイスの加速度センサー等のライフログからユーザーの行動を抽出や分類する研究は数多く行われている [2]。

近年では、PC 操作ログから人の状態や行動パターンを抽出する研究が多くなされている。ウィンドウ遷移に着目

<sup>1</sup> ハミングヘッドズ株式会社  
Humming Heads, Inc., Chuo, Tokyo 104-0052, Japan  
<sup>a)</sup> ryouhei-s@hummingheads.co.jp

した研究としては、Suzuki らによってウィンドウの切り替え頻度から従業員が休憩中かどうかを推定する手法が提案されている [3].

企業内の操作ログに着目した研究としては鳥羽らによってキーボードのログから得られる特徴量と従業員のストレス量の相関関係が報告されている [4]. また、平山らは PC 操作ログの時系列データから特徴的なパターンを検出する手法の提案を行っている [5].

ログ分析を行う際の、膨大なログデータを可視化、要約する研究は非常に盛んに行われている。本研究で使用するユーザー行動モデルも強く依存する要約手法として Wang らの提案した手法が挙げられる [6].

これらの様に、ログの分析に関する研究は非常に注目されているが、PC の操作ログとユーザー行動を関連付ける研究はまだまだ少ない。それ故、本研究で扱う様な数年に及ぶ実社会での PC 操作ログを対象とした行動分析手法の提案や分析結果は筆者の知る限りでは存在しない。

## 2. ウィンドウ遷移ログとユーザー行動モデル

### 2.1 ウィンドウ遷移ログ

本節では分析対象となるウィンドウ遷移ログについて説明する。ここでは Microsoft 社製 OS Windows® を対象として説明する。

ウィンドウ遷移ログはウィンドウシステムを搭載した OS における、フォアグラウンドなウィンドウの遷移を記録したログである。フォアグラウンドなウィンドウとは通常キーボード入力を受け付けるウィンドウであり、多くの場合でディスプレイ上の最前面に表示される。

ウィンドウ遷移ログはフォアグラウンドであったウィンドウを保持するプロセス名とタイムスタンプ及びフォアグラウンドであった時間 (秒) が時系列的に記録される。表 1 はウィンドウ遷移ログの例である。この例では図 1 の様に朝 8:00:30 から 403 秒間 Web ブラウザを使用し、その後メールを 165 秒間使用していたということが推定できる。

ウィンドウ遷移ログは市販のログ取得ソフトウェアで取得可能である。本研究ではログ取得ソフトウェアを導入している企業から取得したログを用いて分析を行う。

表 1 ウィンドウ遷移ログ

Start Time	Process Name	Duration (sec)
2010-04-01 8:00:30	Web	403
2010-04-01 8:07:13	Mail	165
2010-04-01 8:10:15	Editor	503
⋮	⋮	⋮

### 2.2 ユーザー行動モデル

本節ではウィンドウ遷移ログから作成されるユーザー行

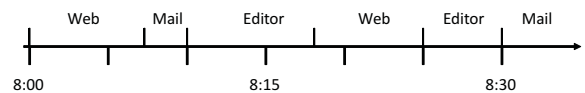


図 1 ウィンドウ遷移

動モデルの概要とその構築手法及び分析手法について記述する。

筆者らが提案したユーザー行動モデルは隠れマルコフモデル (Hidden Markov Model:HMM) を基とした確率的モデルであり、ウィンドウ遷移ログから得られる単位時間内に使用されたアプリケーションの組み合わせとその遷移を表現する [1].

従業員が業務を行う際には、各業務に応じた PC 操作の作業手順がある。例えば、メールで依頼された資料を作成し、送り返す場合にはメールと文章作成ソフトウェアが使用されることが予想できる。これら 2 つのアプリケーションが近傍の時間内で利用される場合は、上記業務の作業を行っているとして、一つのタスクとして表現する。

ユーザー行動モデルではこのタスクとその遷移を用いて従業員の行動と業務を表現する。この様に作成されたユーザー行動モデルは膨大な量のウィンドウ遷移ログを抽象化したモデルであり、このモデルに対して分析を行うことで、ログから特徴やパターンを検出する。

ユーザー行動モデルは図 2 の様にウィンドウ遷移ログから HMM のパラメーター推定を用いて抽象度の高いモデルを作成することで、可視化、比較を行うことが可能である。

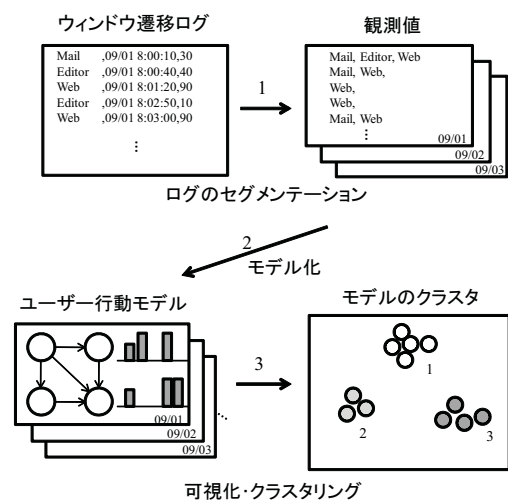


図 2 分析の流れ

### 2.3 ログデータからの観測値

本節ではウィンドウ遷移ログからユーザー行動モデルの観測値の作成方法について説明する。

前述の通り、ユーザー行動モデルでは単位時間内に使用

されたアプリケーションの組み合わせを表現するため、観測値としてアプリケーションの組み合わせを用いる。ウィンドウ遷移をタイムテーブルで表わした図とその時刻におけるイベント集合を図3に示す。このようなイベント集合  $O_t$  を観測値として用いる。

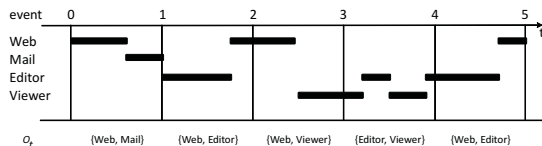


図3 ログから作成される観測値

### 2.3.1 タスク要約

業務において従業員は  $K$  個のタスク状態を持つとして、それぞれに固有の状態  $s_i \in \{s_1, \dots, s_K\}$  を与えることとする。従って、タスク状態は状態の集合として  $S = \{s_1, \dots, s_K\}$  と表記される。

ユーザー行動モデルでは、タスク状態は各タスク状態において観測されたアプリケーションの出現確率によって表現される。これは直感的な観点から自然なモデル化と言える。例えば、ウェブページを作成するという業務を行う場合、ウェブブラウザとHTMLエディターを主に使用し、それ以外のアプリケーションの出現頻度に比べ圧倒的に高くなるのが容易に予想出来る。それ故、タスク  $s \in S$  でユーザーが使用するアプリケーションウィンドウの集合  $o \subseteq \mathcal{E}$  は出力確率  $b_s(o)$  に従うと想定する。ここで、全ての状態  $s \in S$  に対する出力確率  $b_s(o)$  の集合を  $B$  とし、 $o \subseteq \mathcal{E}$  とする。

与えられた集合  $B$  に対して状態  $s$  でアプリケーションウィンドウ  $e$  がアクティブになる確率は  $P_s(e)$  で与えられ、以下のように推定される。

$$P_s(e) = \sum_{o \subseteq \mathcal{E} \text{ such that } e \in o} b_s(o)$$

従って、アプリケーションの数を  $m$  とした時、タスク  $s$  におけるユーザーはそれぞれのアプリケーションウィンドウ  $e_1, \dots, e_m$  に対して確率  $P_s(e_1), \dots, P_s(e_m)$  で出現する。タスク  $s$  に対して、これら出現確率を  $m$  次元のベクトルとして

$$M_s = (P_s(e_1), \dots, P_s(e_m))$$

と定義し、 $s$  のタスク要約と呼ぶことにする。

### 2.4 ユーザー行動モデルの推定

HMMはそのパラメーター  $\lambda = (\Pi, A, B)$  で定義される。HMMの各隠れ状態はそれぞれタスク状態と対応し、その状態がアプリケーションの組み合わせを生成する確率は  $B$  によって決定され、初期確率  $\Pi$ 、遷移確率  $A$  によって系列

が生成される。

ユーザー行動モデルはウィンドウ遷移ログから得られた観測値の系列を用いてHMMのパラメーター  $\lambda$  の推定を行うことで得ることができる。パラメーター推定は非常によく知られた Baum-Welch アルゴリズムを用いて行う [7]。

### 2.5 ユーザー行動モデルの比較

前節で述べた方法で得られるユーザー行動モデルはエルゴード的マルコフモデルである。

ユーザー行動モデルを各頂点がタスク要約でラベル付けされ、各辺がタスク間の遷移確率でラベル付けされた有向グラフであるとし、このグラフをタスク遷移グラフと呼ぶ。2つのタスク遷移グラフの類似性測度として、Kashimaらが提案したラベル付きグラフ間の周辺化カーネル [8] を用いる。この手法を用いて、HMMより得られたユーザー行動モデルの類似度を計算しカーネルPCA [9] を行い、各モデルの主成分得点をプロットすることで可視化およびクラスタリングを行う。

## 3. 実データを用いた分析

本章では長期間に渡るウィンドウ遷移ログからユーザー行動モデルを構築し、このモデルの変化を分析する手法と実際に日本のIT企業から取得したウィンドウ遷移ログを用いて分析した結果について述べる。

### 3.1 対象ログ

ログを収集した企業はソフトウェアの設計、開発、製造、販売を行う日本のソフトウェアメーカーである。各従業員は各自専用のデスクトップPCとユーザーアカウントを使用している。分析においては平日の8:00から21:00までのログを対象に定常的な業務を行う同一部署に所属する4人の2010年9月から2014年6月まで間に渡る長期間のログを用いて行った。

表2 ログ取得対象者

ユーザー名	入社月	有効日数	ログ数	平均ログ数
USER A	2010/09	824日	727426行	882行/日
USER B	2010/09	684日	784328行	1146行/日
USER C	2011/04	616日	828426行	1344行/日
USER D	2011/04	717日	722240行	1007行/日

表2は取得したログにおける各ユーザーの内訳である。有効日数は取得対象時刻内に2時間以上のPCを使用していた日数であり、2時間未満の日のログはノイズとなるため分析対象から除外した。この要因としては休暇、外出、自席以外での作業、PCの故障によるログ欠損等がある。

平均ログ数は一日当たりのウィンドウ遷移ログの記録数を表わし、これは一日当たりのウィンドウ切り替え回数に相当する。この値はUSER AとUSER Cでは1.5倍以上



の差があり、切り替え頻度に個人差があることが分かる。

本研究では分析作業の効率化のため、ログ全体のアプリケーション使用率を計算することにより上位 8 種類のアプリケーションを抜粋し、これら以外のアプリケーションは Etc として下記 9 つのカテゴリを用いてプロセス名を置換し分析に用いた。この 8 種類でアプリケーション使用時間の約 90% を占める。

{Mail, Web, Document, Spreadsheet, Presentation,  
Editor, Viewer, Explore, Etc}

これらの内 Mail, Web, Editor, Viewer はアプリケーショングループであり、それぞれ業務上果たす役割が同一であると考えられるアプリケーションをまとめたものである。Web はウェブブラウザの集まりであり、Mail はメールクライアントソフトウェア、Editor は各種エディタソフトウェア、Viewer は Adobe Reader® 及び画像表示ソフトウェアを含む。

Explore は Windows® の標準のシェルソフトウェアであり、デスクトップや標準のスクリーンセーバー、ファイラーとしての機能を持つ。Document, Spreadsheet, Presentation は、それぞれ文章作成、表計算、プレゼンテーション用ソフトウェアである。

### 3.2 分析手法

本節では、前章で述べたユーザー行動モデルを用いて同一の従業員の時系列的なモデルの変化を分析する手法について説明する。

ユーザー行動モデルは HMM を基としているため、パラメーター推定に使用する観測値に大きく依存する。このため日別や週別の様な少量の観測値を用いた場合、大きな偏りが発生する場合がある。特に日々同じ業務に当たる従業員の場合はこれらの偏りが分析の際に大きなノイズとして発生する。

この問題を解決するため、ウィンドウ幅  $W$  を用いて前後の観測値を合わせてパラメーター推定を行った。この方法を用いることでモデルを時系列的にゆっくりと変化させることが可能となり、時系列変化が視覚的に判断しやすくなる。

ウィンドウ幅を導入する際に、注意すべき点としてログの欠損が挙げられる。単純にカレンダー通りウィンドウ幅を適用すると前後  $W$  日以内にログの欠損がある場合に観測値が減ってしまい前述の効果が期待できなくなる。このため本研究では次の様な方法を用いた。

まず、ログから日別に観測値系列  $I_n$  を作成し日付順に並べると、ログの有効日数を  $T$  とすると系列  $I = (I_1, \dots, I_T)$  を得ることができる。 $n$  番目のユーザー行動モデルを構築する際には  $(I_{n-W/2+1}, \dots, I_n, \dots, I_{n+W/2})$  の観測値系列を用いてパラメーター推定を行い、 $I_n$  に対応する日付をユーザー行動モデルにラベル付けする。この様にすること

で、例えば長期間のログ欠損があった場合でも、その前後で大きく傾向が変わっているかを判断すれば良く、可視化の際に日付からログの欠損を判断可能となる。

分析としては、まず、対象とするユーザーのウィンドウ遷移ログから日別の観測値系列を作成する。次にスライド幅を 1 としてそれぞれ  $W$  日分の観測値系列を用いて、ユーザー行動モデルを作成する。この時、パラメーター推定の初期モデルとして全ての観測値から作成したユーザー行動モデルを用いる。これはエルゴード性マルコフモデルに対して Baum-Welch アルゴリズムを適用する際には初期モデルが大きく影響するためであり、共通のモデルを用いることでモデルのばらつきを小さくする効果がある。

また、パラメーター推定の際に出力確率  $B$  を変化させないで行うことが可能であり、この場合においては全てのモデルでタスク要約が共有され、モデル間の違いは遷移確率のみの違いとなるため直感的な比較が容易となる。

最後に作成された  $T - W$  個のユーザー行動モデルに対して、それぞれの類似度を計算しカーネル PCA を行い、可視化及びクラスタリングを行う。

### 3.3 分析結果

ヒアリング調査の結果 USER A と USER B, USER C と USER D はそれぞれ同一の業務を行うことが多いということが分かった。入社時期の異なるそれぞれのペアについて USER A と USER C の時系列変化についての分析結果を示す。

#### 3.3.1 時系列変化分析

前述した手法を用いて、USER A と USER C それぞれのウィンドウ遷移ログを用いてユーザー行動モデルを作成し、カーネル PCA 及びクラスタリングを行った結果を図 4 に示す。ウィンドウ幅  $W$  は実験的に 20 を使用した。左図が USER A, 右図が USER C のプロットである。各点の色は日付を示しており、図中右下のカラーバーの様の日付と色が対応している。各マーカーはクラスタリングの結果を示しており、どちらの図でも 5 つのクラスタに分割している。

USER A は綺麗に色ごとに固まっており、この塊ごとに別々のクラスタが割り当たっている。一方、USER C はある程度色ごとに固まってはいるが離れた日付の点が混在している。

これは、USER A に関してはカーネル PCA で得られる主成分が時系列と関連性があることを意味する。即ち、モデルの分散は時系列的に大きく時系列に沿ってモデルが変化していることを示している。一方 USER C は時間の経過とモデルの違いにあまり関連性がない。USER A のモデル間の平均類似度は 0.4596, 標準偏差 0.2381 であるのに対して、USER C では平均類似度 0.6668, 標準偏差 0.1794 と全体的に類似性が高いことから分かる様に USER A

に比べてモデルの変化が小さいということを示している。

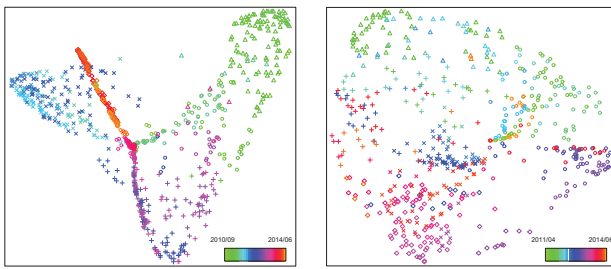


図 4 カーネル PCA の主成分得点のプロット  
 (左:USER A, 右:USER C)

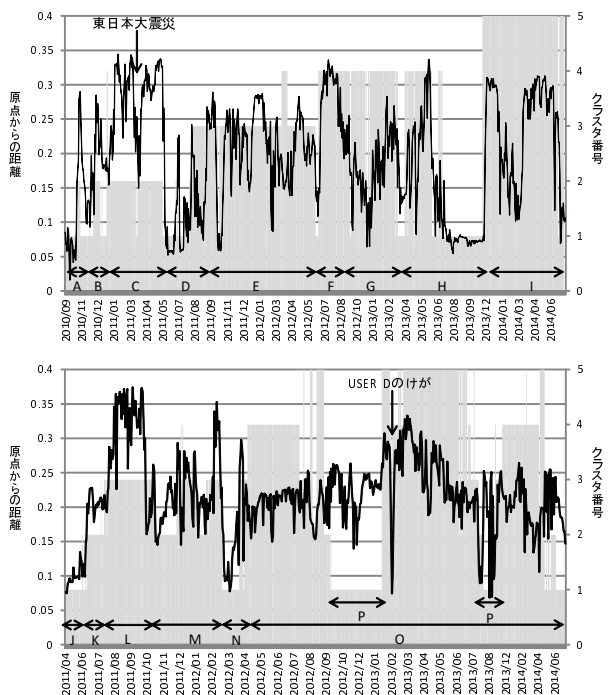


図 5 原点からの距離とクラスター (上:USER A, 下:USER C)

次に各モデルの第一、第二、第三主成分からなる部分空間上の原点からの距離とクラスターの関連を図 5 に示す。x 軸には日付を取り、実線が原点からの距離(左の軸)、灰色の網がけはクラスターのインデックス(右の軸)を示す。

図中の実線は原点からの距離であるため、同一の値であってもモデル間の距離が近いということの意味しない。しかし、近傍でこの値が大きく異なるということは部分空間内での位置が大きく異なることを意味し、ユーザー行動モデルが変化していることを示す。クラスタリングは主成分得点に対して階層化クラスタリングを用いているため、時系列的に見てクラスターの切り替えが起こる点の近傍では原点からの距離も大きく変動している。

また、USER A の 2011/03 付近では近傍に比べ大きく減少している。これは東日本大震災の影響で、企業内での業務体制が変わったことが関連している。USER C では

2013/02 付近にペアとして勤務している USER D のけがの影響で業務パターンが変わった形跡が見られる。

図中 A から O で示されている区間はヒアリングの結果から得られた、同一の業務を行っていた期間である。該当する業務は USER A の A から順に研修、引き継ぎ、事務、引き継ぎ、経理業務、調査業務 1、調査業務 2、資料作成 1、資料作成 2 であり、USER C では J から順に研修、引き継ぎ、事務、事務+自席外作業、資料整理、事務となっている。

USER A に関しては各期間とクラスターの切り替わる地点が概ね一致していることが分かる。ヒアリング結果からも該当業務内容が大きく変わっており、PC の使用方法も変化していることが予想出来る。一方 USER C は入社以来、主に事務業務を行っており、特に期間 O では事務業務のみを行っている。これは USER C の全体的なモデル間の類似性が高いことと合致する。

同一の業務を行っている中でのパターンの違いを分析するため、次項では期間 O 内でもクラスターが異なる期間 P に着目し他の部分との違いを分析する。

### 3.3.2 モデル間比較

モデル間の比較を行うためには、タスク要約を全モデルで共有している方が分析が容易であるため、パラメータ推定の際に出力確率  $B$  を更新せずに同一の観測値系列からモデルを再度作成した。

まず、USER C の全ログから作成した全体的なモデルの定常状態分布を図 6 に示す。この図から USER C の潜在率が高い状態が状態 2、状態 5、状態 8、状態 9 であることが分かる。次に作成した各モデルのから定常状態分布を求め

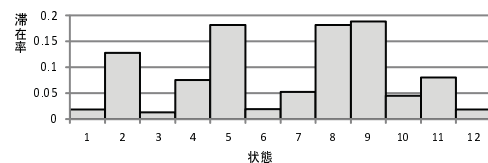


図 6 USER C の定常状態分布

上記 4 つの状態に対しての変化と各状態のタスク要約を図 7 に示す。タスク要約の各グラフ上の値はそのタスク状態でのアプリケーションの使用頻度を示しており、値が 1 の場合には必ずそのアプリケーションが使用されることを示す。状態 2 は全てのアプリケーションでほぼ 0 であり、離席や PC を使用しない業務を意味する。図中の期間 P に該当する箇所では状態 5 と状態 9 の値が下がり、状態 2 と状態 8 の値が上がっていることが分かる。これは期間 P においては Mail と Viewer の使用傾向が変わり、離席または PC を使用しない作業が増えたことを意味する。これは明らかに他の期間とは異なるパターンである。ヒアリング調査時にこの様な話は一切出てこなかったことから、環境や忙しさ等で働き方が変わる等、無自覚な変化であると言える。

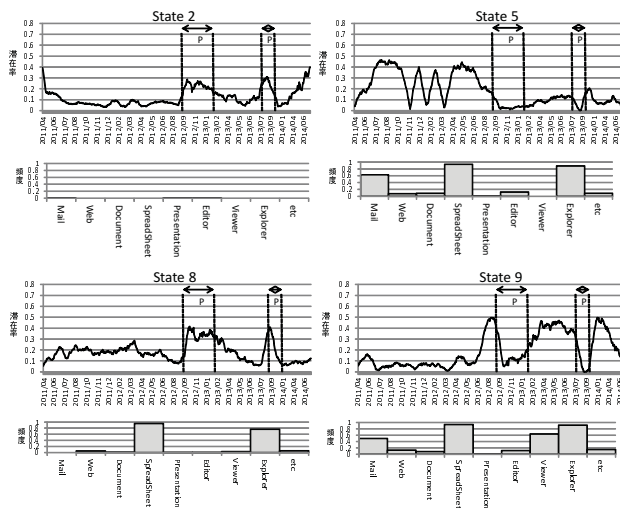


図 7 定常状態分布の変化とタスク要約

## 4. おわりに

本稿では、長期間に渡るウィンドウ遷移ログから作成したユーザー行動モデルの時系列変化分析を行うことで、ユーザー行動モデルと実際の業務との関連性を検証した。ユーザー行動モデルを用いて、4年間・約300万行に及ぶ実際のログを抽象化することで得られた情報は、ヒアリングの結果と合致する点が多く、ログと業務の間の関連性は強いことを示した。また、同一業務期間内における異なった作業パターンを検出することができた。

最後に、本手法の長所と短所についての考察と今後の応用分野について述べる。

### 4.1 考察

前章の分析結果より、ユーザー行動モデルの時系列変化は実際の業務と関連性が強いと言える。しかしながら、複雑なユーザー行動モデルを作成しなくても、もっと簡単な手法でも可能では無かったのかという疑問が残る。

単純な方法としては、ウィンドウ遷移ログからアプリケーションの使用時間を計算しその比率に対してPCAを用いることが考えられる。実際に分析を行ってみたところ、ある程度のクラスタに分割は可能であった。しかし、前章のUSER Cにおける期間Pに対する違いは検出できなかった。これはUSER Cは期間Oにおいてはほとんどアプリケーションの使用頻度が変わらないためである。使用時間による手法では短時間の使用を高頻度で行う場合と長時間低頻度で使用する場合を区別できないため、上記の違いを検出することができない。一方、ユーザー行動モデルを用いる場合には使用頻度と遷移の情報を保持するためこの様な違いが検出可能である。

また、今回用いたウィンドウ幅を用いた分析手法は比較的長い期間同じ業務を担当する従業員にのみ適用可能であ

る。前述の通りウィンドウ幅が小さいとサンプル数が少ないため多くのノイズが混入してしまう。一方、業務が1週間や数日で変わる様な従業員に対しては、一つのユーザー行動モデルの中に複数の業務が含まれてしまうため時系列的に近傍のモデルに対してばらつきが生まれてしまう。この様な業務を行う従業員に対しては短期間でモデルを作成し、業務別にクラスタリングした後、同一クラスタ内の時系列変化をみる等の別の手法を考える必要がある。

### 4.2 今後の展望

今後の展望としては、ユーザー行動モデルを用いた分析を実習や研修中のログを対象として行うことを考えている。企業で行われる実習や研修は、固定されたPC上で操作を行い、様々なバックグラウンドの人が同一の成果物を作成することを目標とすることが多い。より限定されたPC使用環境のログを用いて分析を行うことで、個人の特徴やその変化が抽出できる可能性があり、パターンと研修終了時の客観的評価の関連性を見つめることができると考えている。

### 参考文献

- [1] Saito, R., Kuboyama, T., Yamakawa, Y. and Yasuda, H.: Understanding User Behavior through Summarization of Window Transition Logs, *Databases in Networked Information Systems*, Lecture Notes in Computer Science, Vol. 7108, Springer Berlin Heidelberg, pp. 162–178 (2011).
- [2] Rana, C.: A Study of Web Usage Mining Research Tools, *Information Retrieval*, Vol. 1429, No. 06, pp. 1422–1429 (2012).
- [3] Suzuki, K., Yasuda, H., Shin, K. and Kuboyama, T.: Discriminating User Behavior through PC Operation Logs by PageRank Convergence Patterns, *International Journal*, Vol. 3, No. 1, pp. 37–40 (2014).
- [4] 鳥羽美奈子, 櫻井隆雄, 森 靖英: PC 操作ログの特徴量とオフィスワーカーのストレス量の相関分析 (データマイニング, 特集: ライフログ処理技術とその活用システム論文), 電子情報通信学会論文誌. D, 情報・システム, Vol. 95, No. 4, pp. 747–757 (2012).
- [5] 平山明彦, 原 直, 阿部匡伸: 非負値行列因子分解によるPC操作ログからの勤務パターン抽出 (ライフインテリジェンスとオフィス情報システム), 電子情報通信学会技術研究報告= IEICE technical report: 信学技報, Vol. 114, No. 32, pp. 33–38 (2014).
- [6] Wang, P., Wang, H., Liu, M. and Wang, W.: An algorithmic approach to event summarization, *Proceedings of the 2010 international conference on Management of data*, ACM, pp. 183–194 (2010).
- [7] Rabiner, L.: A tutorial on hidden Markov models and selected applications in speech recognition, *Proceedings of the IEEE*, Vol. 77, No. 2, pp. 257–286 (1989).
- [8] Kashima, H., Tsuda, K. and Inokuchi, A.: Marginalized kernels between labeled graphs, *Proceedings of 20th International Conference on Machine Learning(ICML)*, pp. 321–328 (2003).
- [9] Schölkopf, B. and Smola, A.: *Learning with kernels*, “The” MIT Press (2002).