

作業発生の規則性に基づく作業予測の信頼性評価手法

北垣 千拡¹ 乃村 能成¹

概要: 我々は、これまでの研究で過去のカレンダー情報に基づいて周期的に発生する作業から将来の作業の発生を予測する手法を提案した。しかし、この手法による予測結果には予測の妥当性を示す指標が含まれておらず、予測の妥当性を利用者が吟味し、判断する必要がある。そこで、予測結果に対する信頼性の指標を導入することで、利用者の予測結果の妥当性に対する判断を支援する。本稿では、作業予測の結果に対する信頼性評価手法を提案し、評価した。

キーワード: 作業予測, カレンダー, 計画立案

1. はじめに

我々が将来の予定について計画するとき、過去の作業履歴を参考にすることが多い。なぜならば、作業は繰り返し発生している場合が多く、過去の作業の発生から将来の作業の発生をある程度推測できるからである。たとえば、約1カ月に1度発生している定例会議は最後に発生した日付から約1カ月後に発生することが推測できる。このように、作業の発生には何らかの周期性が存在し、この周期を確認できれば作業の発生を予測できると考えられる。将来の作業の発生を予測できれば、作業の発生が集中する繁忙期を推定できるため将来の計画を立てる作業(計画立案)の際に有用である。IT分野専門のリサーチ企業の調査によると、労働者の8割がスケジューラを利用しており[1]、このことから、スケジューラを用いた計画立案の場面は多く、作業の発生の予測は有用であると言える。

過去の履歴を利用し、予測する研究はこれまでもいくつか存在する[2][3][4]。しかし、これらの研究の多くは機械学習を用いたものである。一般に、機械学習によってある程度の予測精度を実現するためには非常に多くのデータが必要となる。しかし、個人が利用するカレンダーの予定数は機械学習に必要なデータ数より少ない場合が多い。このため、機械学習を用いた予測手法では、許容されるレベルの予測精度に達するまでに準備期間が必要となる。また、人間の行動は数年といった単位で変わりうるため、非常に長い期間の行動履歴を行動予測に用いることは用途によっては適切ではない。対して、我々が提案している予測手法

では、これらの手法に比べ少ないデータ数で高い予測精度を実現している[5]。

一方で、作業予測を用いた計画立案を考えた場合、高い予測精度を実現する以外にも重要な要素がある。その1つとして、提案に対して利用者が納得できる理由を提示することが挙げられる。理由の提示方法として、たとえば、述語論理を用いたスケジューリングを行うことによって、結果がどういった論理で導かれたかを利用者に提示する研究がある[6]。しかし、このような方法では、利用者が提案に対して吟味する手間が大きい。また、提案に対する定量的な評価ができないため、機械的な処理に適さない。

そこで、予測の妥当性の指標を数値的に提示する方法を考える。この方法では、利用者は提示された数値の大小を見ることで妥当性判断の確からしさについて知ることができ、計画立案の補助として利用できる。そこで本稿では、予測結果に対する信頼性の指標を導入することで、利用者の予測結果の妥当性に対する判断を支援する手法を提案する。

まず、はじめに、我々の提案している予測手法について述べる。次に、予測手法の妥当性を判断するための指標になり得る考え方を示す。さらに、作業予測の結果に対する信頼度を定義し、信頼度に基づいた信頼性評価手法を提案する。最後に、手法に対する評価を行う。

2. 作業発生の規則性に基づく作業予測

2.1 作業発生の規則性

作業発生に関わる周期性と関連性を集合として表現したモデルを**作業発生の規則性**と呼ぶ[7]。作業発生の規則性には、作業とその周期性を表現する以下の2つの概念がある。

¹ 岡山大学大学院自然科学研究科
Graduate School of Natural Science and Technology,
Okayama University

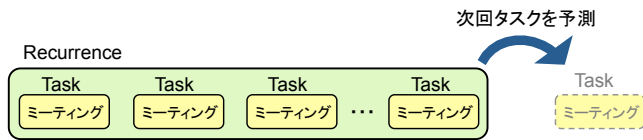


図 1 リカーレンスを利用した作業予測

タスク (Task): タスクは、作業を扱う最小の単位である。タスクは開始時刻と終了時刻を持ち、この間で連続的に行われる作業を表現する。また、個別のタスクは開始時刻による順序関係をもつ。

リカーレンス (Recurrence): リカーレンスはタスクを要素とする集合である。リカーレンスは繰り返し発生している同様のタスクを1つの集合とする。

リカーレンスは繰り返し発生する同様のタスクの集合であり、リカーレンスをうまく利用することで、将来のタスク発生を予測できる。リカーレンスを用いた作業予測の例を図1に示す。図1のように、リカーレンス内に複数のタスクが存在する場合、リカーレンス内のタスクの周期性を解析することにより、次回タスクの発生を予測できる。

2.2 繰返情報

作業発生規則性に基づく作業予測では、作業予測に必要な情報として繰返情報を定義し、これを用いて予測を行う。文献[8]での調査結果と同様に、リカーレンス内のタスクに関する以下の8つの統計情報を繰返情報とする。

作業発生間隔の平均: 作業発生間隔の平均日数と標準偏差
昨年の作業発生間隔: 昨年の同時期に発生した作業の作業発生間隔

曜日情報: 日月火水木金土の7つに祝日を加えた8つの曜日の発生割合

月情報: 1月から12月までの12の月の発生割合

週情報: 月の第1週から第5週までの5つの週の発生割合

日付情報: 1月1日から12月31日までの1年の各日付の発生割合

時間帯情報: 開始時刻の上位3つと発生回数の最も多い作業時間を組み合わせた時間帯の3つのパターン

最終作業情報: リカーレンス内の直近の作業に関する情報

2.3 作業発生規則性に基づく作業予測

2.3.1 1年を周期とした周期性の継承

文献[5]の調査で、多くの作業は「作業発生が年を単位とした相似形をもつ」という性質をもつことが確認されている。この性質について、図2を用いて概略を説明する。図2は、あるミーティングにおける2009年度と2010年度の2年間の発生履歴を示している。横軸は日付で、縦軸は発生の有無を1/0で表現している。「作業発生が年を単位とした相似形をもつ」という性質が明確にあらわれている部分として、7月下旬から9月上旬にかけての作業発生が

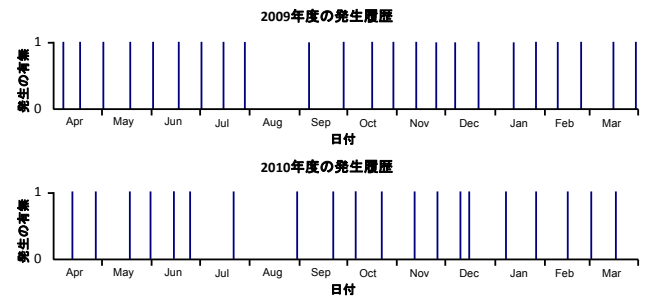


図 2 2009年度、2010年度のミーティングの発生履歴

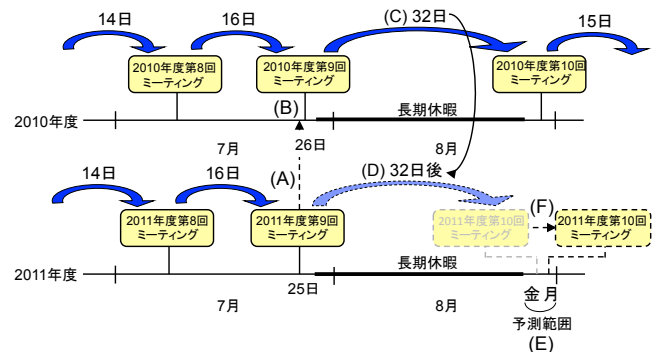


図 3 1年を周期とした周期性の継承に基づく作業予測手法の適用例

ある。この部分は他の作業発生と比べ、発生間隔の幅が広がっている。2009年度においては、7月下旬から9月上旬まで作業が発生していない。また、2010年度においては、7月下旬から8月下旬まで作業が発生していない。各年度において7月下旬から8月下旬が共通して作業が発生していない期間である。2009年度と2010年度の作業の発生履歴から、2011年度も7月下旬に作業が発生し、その次は8月下旬に発生すると推測できる。12月から1月までの作業発生についても同様のことが言える。

以上のことから、各年度において、作業発生が1年ごとに対応している（たとえば、2009年度の第1回ミーティングと2010年度の第1回ミーティングが対応している）と考えた場合、2010年度の作業発生間隔を用いることで2011年度の作業発生を予測できる。つまり、昨年の作業発生間隔を用いて、作業発生を予測できる。

2.3.2 作業予測

本項では、これまでに我々が提案している予測手法[5]について説明する。予測は、昨年の発生間隔を用いておおよその候補日を決定した後、候補日の中からより良い日付を選択するという流れで行う。まず、リカーレンス内の最後に発生した日付に昨年の同時期の発生間隔を適用することで候補日を求める。次に、候補日の前後数日を予測範囲とする。さらに、繰返情報から得られる曜日、月、週、および日付の発生割合をパラメータとして評価関数で重み付けする。最後に、評価関数の評価値に基づいて、予測範囲の中から最も発生確率が高い日を選択する。

次に、この予測手法を適用した例について説明する。

図 3 はミーティングに作業予測手法を適用した例であり、「2011 年度第 10 回ミーティング」の予測の様子を示したものである。以下に手順を示す。

- (A) 「2011 年度第 9 回ミーティング」の発生日である 7 月 25 日から 1 年 (365 日) 前の日付を求め、2010 年 7 月 25 日を得る。この日付を基準日とする。
- (B) 基準日をもとに対応点を求める。具体的には、2010 年 7 月 25 日から最も近いミーティングとその発生日を求め、「2010 年度第 9 回ミーティング」(7 月 26 日) を得る。
- (C) 「2010 年度第 9 回ミーティング」と「2010 年度第 10 回ミーティング」の間隔を求め、32 日を得る。
- (D) 間隔 32 日を 2011 年度にも適用し、「2011 年度第 9 回ミーティング」から 32 日後を候補日とする。
- (E) 候補日の前後 3 日間 (候補日 1 を含めて 1 週間) を予測範囲とする。
- (F) 予測範囲内の日付で曜日、月、週、日付の発生割合を考慮して、発生割合の最も高い日付を発生日として予測する。図 3 では、発生割合の高い月曜日に修正し、その日付を「2011 年度第 10 回ミーティング」の発生日と予測する。また、上記の方法により求めた発生日に時間帯の情報を加えて、発生日時を予測する。

2.3.3 予測精度

実データに対してこの予測手法を適用した場合の予測精度について簡単に述べる。文献 [5] の評価結果によると、予測した 176 件のタスクに対して、全体の約 9 割である 163 件が計画立案における提案として妥当な予測であったことが確認されている。このことから、この予測手法は、ある程度の周期性をもつ作業であれば、高い予測精度をもつと言える。

3. 課題と対処

3.1 課題

2 章で述べた予測手法によって得られた予測結果には実用上の問題がある。それは、予測結果を利用者に提示する場合、利用者が予測結果の妥当性を吟味する手間が大きいと、利用者の負担がかえって大きくなってしまうことである。予測結果を参考にカレンダーに新しい予定を登録するケースを考えた場合、以下の問題がある。

まず、予測結果に基づいて、作業名と予測された日付のみが利用者に提示される。このとき提示される予測結果は、必ずしも利用者にとって妥当ではない。妥当でない予測を多く提示してしまうと、かえって利用者の負担が増してしまう。妥当とはいえない予測結果を利用者に提示してしまう理由は、作業の特徴ごとに予測精度が異なるためである。2.3.1 項で述べたように、この予測手法は、多くの作業は「作業発生が年を単位とした相似形をもつ」という性質に着目した手法である。このため、この性質を満たすこ

とが作業予測に必要な条件である。この性質をどの程度満たすかは、作業の特徴ごとに異なる。作業の特徴として、規則的な周期をもつ作業は、この性質を満たす度合いが大きく、予測精度が高くなる。一方、曖昧な周期をもつ作業は、この性質を満たす度合いが小さく、予測精度が低くなる。たとえば、規則的な周期をもつ作業として「入学式」について考える。「入学式」は「毎年同じ日付に行く」といった規則的な周期をもつ。このため、「作業発生が年を単位とした相似形をもつ」という性質を満たし、予測精度が高くなる。次に、曖昧な周期をもつ作業として「忘年会」について考える。「忘年会」は「毎年 12 月下旬に行く」といった曖昧な周期をもつ。この作業は毎年同じ時期に発生するため、「作業発生が年を単位とした相似形をもつ」という性質を満たすように思える。しかし、12 月下旬であればどこでも発生し得るという曖昧さを含むため、規則的な周期をもつ作業と比べ、年を単位とした相似形にぶれがある。このため、性質を満たす度合いが小さく、予測精度が低くなる。このように、予測結果は作業の特徴ごとに精度が異なる。予測精度が低い場合、妥当でない予測結果が提示されてしまうこともある。

しかし、予測結果のみから予測精度の違いを知るすべが利用者にはないため、予測結果の妥当性が判断できない。利用者が予測結果の妥当性を知るためには、対象の作業の特徴を把握したうえで、提示された日付が妥当であるかを吟味しなければならない。この行為は利用者自身が作業予測を行うことに等しく、機械的な作業予測の利点が損なわれてしまう。

3.2 対処

予測結果を有効に利用するためには、利用者が予測結果の妥当性を吟味する負担を抑える必要がある。ここで、この手間をどの程度軽減できるかについて考える。

例として、利用者が予測結果を参考にカレンダーに新しい予定を登録する状況を考える。理想的には、予測結果を機械的に振り分け、妥当な予測結果だけが自動でカレンダーに登録できれば便利である。しかし、妥当な予測結果だけを自動でカレンダーに登録することは現実的ではない。なぜならば、たとえ予測結果が妥当だとしても、実際に利用者がカレンダーに予定を登録するとは限らないためである。このように、予測結果をどう扱うかの判断は最終的には利用者にあるため、現実的には、利用者による予測結果の取捨選択は必須となる。このため、取捨選択の手間をどの程度軽減できるかが重要となる。

そこで、予測結果に対する妥当性を表す指標として**信頼性**を導入する。これにより、利用者の予測結果に対する妥当性の判断が支援でき、信頼性の低い予測結果をあらかじめ除いたり、取捨選択の手間が軽減できたりする。

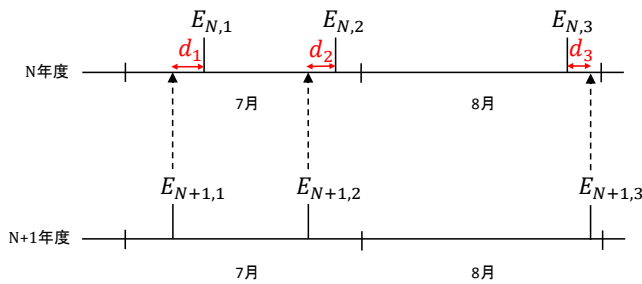


図 4 作業の発生履歴の系列における自己相関

4. 信頼性評価手法

4.1 信頼性

4.1.1 考え方

信頼性について検討するために、予測手法の特徴について考える。この予測手法は、多くの作業は「作業発生が年単位とした相似形をもつ」という性質に着目した手法である。3.1節でも述べたように、予測対象の作業がこの性質を満たすか否かが予測精度に大きく影響する。このため、予測対象となる作業が前述の性質を満たすか否かを判定することによって、予測結果に対する信頼性が表現できると考えられる。

4.1.2 表現方法

前節で述べたように、予測対象となる作業が「作業発生が年単位とした相似形をもつ」という性質を満たすか否かを判定することで信頼性を表現する。本節では、この表現方法について述べる。

2.3.1項で述べたように、作業の発生履歴は1/0の系列と考えられる。ここで、問題を「系列が年単位とした相似形をもつかを判定すること」に置き換えて考える。ある系列が系列中に相似形をもつとは、言い換えれば、系列中に繰返しのパターンが存在することと言える。系列中に繰返しパターンが存在するか求めることは、系列の自己相関を求めることに等しい。このことから、信頼性の判定は、作業の発生履歴の系列に対する自己相関を算出する問題に帰着する。

次に、作業の発生履歴の系列に対する自己相関について考える。自己相関関数は、対象とする現象によって定義が異なる。たとえば、連続信号における自己相関は、ある信号の自分自身との畳込みとして定義される。しかし、この定義は作業の発生履歴の系列に対する定義としては適さない。なぜならば、作業の発生履歴の系列は1/0のインパルス列であるため、信号の畳込みの際、時間軸方向にほんの少しずれただけで全く相関がないと扱われてしまうためである。このため、作業の発生履歴の系列に適した自己相関の定義を与える必要がある。

4.1.3 作業の発生履歴の系列における自己相関

作業の発生履歴の系列は、1/0のインパルス列であるた

めインパルス列の間隔のみで系列の性質が決まる。このため、作業の発生履歴の系列における自己相関は、時間軸方向のずれのみから評価されるべきである。図4を用いて、時間軸方向のずれのみの評価に着目した自己相関の考え方について説明する。図4はN年度とN+1年度の7、8月付近の作業の発生履歴である。E_{N,i}はN年度のi回目の発生日を表す。同様に、E_{N+1,i}はN+1年度のi回目の発生日を表す。まず、E_{N,1}に注目すると、E_{N,1}はE_{N+1,1}の日付の1年前付近に発生していることが分かる。このとき、E_{N+1,1}には対応する発生日が1年前に存在すると言える。同様に、E_{N+1,2}の日付の1年前付近にE_{N,2}が対応する。このように、1年前の系列と対応する組が作れるとき、系列は年単位とした相似形であると言える。次に、系列の自己相関の強さについて説明する。ここで、時間軸方向のずれに着目した評価を行う。E_{N+1,1}の1年前の日付とE_{N,1}の日付の距離をd_1とする。同様に、E_{N+1,2}の1年前の日付とE_{N,2}の日付の距離をd_2とする。このとき、iは組の数だけ存在し、d_iの総和が小さいほど系列の自己相関が高くなる。系列の自己相関が高いほど相関が強く、系列が相似形に近いことを表す。

作業の発生履歴の系列における自己相関関数の定義を式1に示す。なお、τはラグであり、Dは作業発生日の日付の集合、f(t)は日付tに最も近いDの要素を返す関数を表す。R(τ)の値が小さい、つまり、自己相関が高いことは、系列がτ日を単位とした相似形に近いことを意味する。

$$R(\tau) = \sum_{t \in D} (f(t) - f(t - \tau))^2 \quad (1)$$

4.2 信頼度の定義

4.1節で述べたように、自己相関を用いて信頼性を表現する。また、この信頼性の指標を**信頼度**として定義する。作業における信頼度の定義を式2に示す。

$$Reliability = \frac{1}{R(365) + 1} \quad (2)$$

式2にて、τ=365としているのは、作業が「作業発生が年単位とした相似形をもつ」という性質を満たすかを評価するためである。信頼度は最大で1、最小は0に限りなく近い値になる。信頼度を用いることで作業に対する信頼性の程度を表現できる。

5. 評価

5.1 概要

本手法における信頼度は、予測結果に対する妥当性を判断するための指標として提示される。このため、妥当性の高い予測結果の信頼度の値は大きく、妥当性の低い予測結果の信頼度の値は小さくなることは予想できる。しかし、4章では、信頼度がいくつ以上であればその予測結果が信頼足りうるのかといった基準は示されていない。そこで、

まず、人間が予測可能な特徴をもつと判断した作業と予測困難な特徴をもつと判断した実際の作業のデータを用意し、それぞれに対して信頼度を算出する。そして、得られた信頼度の値と作業の特徴との相関関係を調べる。結果から、予測可能な特徴をもつ作業の信頼度と予測困難な特徴をもつ作業の信頼度を分離できる閾値を定める。これにより、信頼度の値のみから予測結果の妥当性を判断可能にする。次に、閾値を定めるのに用いたデータとは別のデータを用意し、予測結果の妥当性を正しく分類できるかを評価する。これにより、本手法の有用性を示す。

5.2 信頼性判定のための閾値の調査

5.2.1 準備

調査に用いる作業のデータについて述べる。作業のデータには、著者らの所属する研究グループの2011年度と2012年度の過去2年間の予定を用いる。予定は36種のリカーレンスに分類され、タスクの合計数は348個である。これらの作業は、特徴ごとに以下の7つに分類できる。

(分類 A) 周期を強く意識する作業

この作業は、それぞれ約2週間、約1カ月、約半年、および約1年という周期を陽に意識している作業である。これらの作業については、次回の作業日時を決定する際に、周期を意識した意思が働く。

(分類 B) 季節ごとに発生する作業

周期を意識しないが、季節ごとに発生する作業である。この作業は、季節ごとというおおまかな周期は決まっているものの、周期が作業日時の決定の際に及ぼす影響は、Aの作業ほど強くない。

(分類 C) 月ごとの頻度に偏りがある作業

半期のみ約1週間に1回という周期で発生する作業である。これらの作業に関しては、我々は約1週間に1回という周期でとらえているが、半期のみしか発生しないという点で、Aの作業とは異なる。

(分類 D) 固定的に発生する作業

固定的な日付に発生する作業である。この作業は、毎回決まった日付に発生する。

(分類 E) 予定変更の多い周期的作業

周期は意識するが、周期が乱れやすい作業である。これらの作業については、次回の作業日時を決定する際に、周期を意識した意志が働くが、周期を維持するという制約は弱く、参加者の都合等の別の要因によって普段意識している周期が無視されることが多い。

(分類 F) 不定期に発生する作業

不定期に発生する作業である。この作業は、発生に関して特に強い制約はない。

(分類 G) 情報の欠落を含む周期的作業

カレンダーへの登録し忘れによる周期の乱れをもつ作業である。もしカレンダーへ正しく登録されていたとすれ

表 1 信頼度と作業の予測可能さとの関係

分類	作業名	信頼度 × 1000000	予測可能か
A	忘年会 1	1000000	○
D	大学の後期入試	1000000	○
A	研究室大掃除	500000	○
A	卒論発表	500000	○
A	大学の推薦入試	500000	○
A	大学の編入試験	500000	○
A	修論発表	500000	○
A	送別会 1	500000	○
A	オープンキャンパス	500000	○
D	大学の前期入試	500000	○
D	大学の入学式	500000	○
A	大学の卒業式	200000	○
A	卒業アルバム写真撮影	200000	○
A	新人歓迎会	200000	○
A	研修会	200000	○
C	講義 4	62500	○
A	B4 中間発表	38461	○
A	M2 中間発表	38461	○
A	M2 論文紹介	27027	○
A	暑気払い	27027	○
A	送別会 2	27027	○
C	講義 3	13157	○
C	講義 2	11627	○
C	講義 5	6802	○
A	忘年会 2	3891	○
B	スポーツ大会	3802	○
A	M1 論文紹介	2762	○
C	講義 1	2695	○
A	防災訓練	2061	○
A	ミーティング 1	1960	○
A	進捗報告会 3	1432	○
A	進捗報告会 1	1197	○
A	進捗報告会 2	1053	○
E	ミーティング 2	523	△
F	レクリエーション	140	×
G	誕生日会	35	×

ば、A から F のいずれかに分類される。

上記の分類のうち、利用者から見て予測可能な特徴をもつと思える作業は分類 A, B, C, D である。一方、予測困難な特徴をもつ作業は分類 F, G である。分類 E は周期の乱れ具合によってどちらかに分類される。

5.2.2 調査

予測可能な特徴をもつ作業と予測困難な特徴をもつ作業に対して信頼度を算出し、得られた信頼度の値と作業の特徴との相関関係を調べる。表 1 に信頼度と作業が予測可能かを人間が判断した結果の関係を示す。

表 1 は、各作業について式 2 で定義した信頼度が高い順に並べている。また、各作業が予測可能かどうかについて、予測可能な特徴をもつ作業は○、予測困難な特徴をもつ作業は×、およびどちらにも分類されうる作業は△を与えている。なお、予測可能かどうかは 5.2.1 項で述べた作業の分類に基づいて与えた。表から、予測困難な特徴をもつ作業は予測可能な特徴をもつ作業と比べて信頼度が低いことが分かる。予測困難な特徴をもつ作業からは妥当でない予測結果が得られる場合が多いことをふまえると、信頼度の値と予測結果の妥当性には相関関係があると言える。

また、予測可能な特徴をもつ作業と予測困難な特徴をもつ作業を「進捗報告会 2」の信頼度と「ミーティング 2」の信頼度の間で分離できることが分かる。具体的には、信頼度がおおよそ 0.001 以下の作業については、予測困難な特徴をもつ作業に分類できる。そこで、閾値を 0.001 に設定し、これより信頼度が小さい作業は信頼性が低いと判定することにする。これにより、各作業の予測結果が妥当か否かを判定できる。無論、この値は著者らの所属する研究グループのカレンダの場合に妥当だと考えられる閾値である。しかし、一般的に考えて、表 1 の分類 A-G のような予定から構成されるカレンダは多く、それらの比率の差によって組織が特徴付けられると考えられる。分類 A-G の比率そのものは今回の閾値には影響しないと考えられるため、他の多くの組織においても同様の効果が得られると考えられる。

5.3 信頼性判定の精度の評価

5.3.1 評価観点

信頼性判定の精度を評価するために、予測困難な特徴をもつ作業を判定できるか評価する。本稿における予測困難な特徴をもつ作業は、(分類 E) 予定変更の多い周期的作業、(分類 F) 不定期に発生する作業、および(分類 G) 情報の欠落を含む周期的作業のいずれかに分類される作業である。評価のためには、これらの予測困難な特徴をもつ作業を表現するデータを用意する必要がある。そこで、各分類の作業をどう表現するか考える。

(分類 E) 予定変更の多い周期的作業は、周期は意識するが周期が乱れやすい作業である。(分類 E) の作業は、周期の乱れが大きいと不規則な作業発生となり、予測が困難になる。ここで、周期の乱れの表現方法を考える。人間が作業の系列を見たときに周期が乱れていると認識するのは、その作業の特徴を把握しているためである。たとえば、おおよそ 2 週間周期で行われるが参加者の都合で周期が乱れる「ミーティング」を例に考える。この作業の発生間隔が参加者の都合で 3 週間になったとする。これに対して、周期が乱れたと感じるのは「ミーティングは通常 2 週間周期で行われる」という作業の特徴を知っているからである。作業の特徴を知らない人間がこの系列を見ても、予定が不規則に発生しているようにしか見え、周期に乱れがあるとは感じない。つまり、作業における周期の乱れとは、系列だけを見ると不規則に発生する作業と変わらない。ここで、系列が不規則であると感じるかどうかは、周期の乱れ具合の問題である。これらのことから、周期の不規則さと周期の乱れ具合には相関があると言える。このため、不規則に発生する作業を用意することによって、周期の乱れが大きい作業、つまり(分類 E) の作業を表現できる。(分類 G) 情報の欠落を含む周期的作業についても同様のことが言える。人間が情報の欠落があると感じるのは作

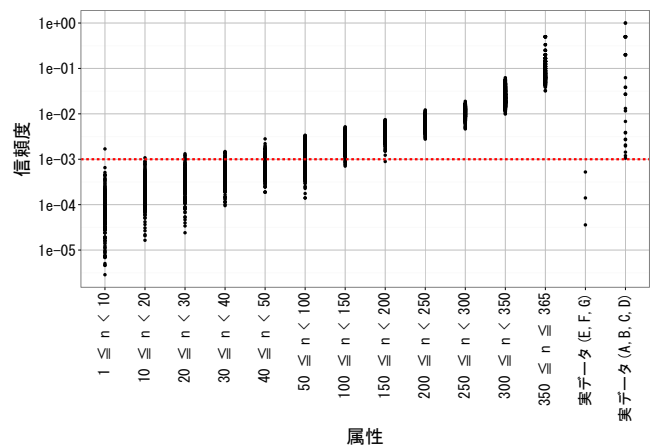


図 5 各 n におけるランダムなデータの信頼度の分布

業の特徴を把握しているためである。このため、不規則に発生する作業を用意することによって、情報の欠落がある作業、つまり(分類 G) の作業も表現できる。最後に、(分類 F) 不定期に発生する作業については、言うまでもなく不規則に発生する作業である。以上から、いずれも不規則に発生する作業を用意することで、予測困難な特徴をもつ作業を用意できる。

用意した作業のデータに対して、5.2.2 項で定めた閾値を境に予測困難な特徴をもつ作業として正しく分類できるか評価する。正しく分類できれば、予測困難な特徴をもつ作業を判定でき、信頼性判定の精度が高いと言える。しかし、ランダムなデータとはいえ、毎日発生するような場合は、擬似的な周期性をもってしまうことは自明である。このため、擬似的な周期性が生まれにくい、年間の発生数の少ない範囲で正確に作業を分類できるかが重要である。

5.3.2 評価環境の準備

予測困難な特徴をもつ作業として、機械的に作成したランダムなデータを評価に用いる。ランダムなデータは、2 年間に不規則な発生間隔で年間平均 n 回発生する予定として作成する。なお、 n の範囲は $1 \leq n \leq 365$ とする。これは年に 1 回の発生頻度から毎日の発生頻度に相当する。このランダムなデータを各 n につき 10 パターンずつ、合計 7300 種類の系列を用意する。なお、5.2.2 項で述べたように、信頼性判定のための閾値は 0.001 とする。

5.3.3 実測

図 5 に各 n におけるランダムなデータの系列とそれに対する信頼度の分布を示す。横軸はデータの属性、縦軸は信頼度を対数軸としてとる。なお、属性は n の範囲別に分類している。また、参考のため、データの属性として 5.2.1 項の実データも載せている。ランダムなデータは、たとえば、 $1 \leq n < 20$ は年間に 1 から 19 回ランダムに発生する作業であり、信頼度は $1.6e-05$ から $1.1e-03$ の間に分布している。 n の小さい範囲においては、閾値である 0.001 以下に信頼度が分布していることが分かる。一方、 n の大きい

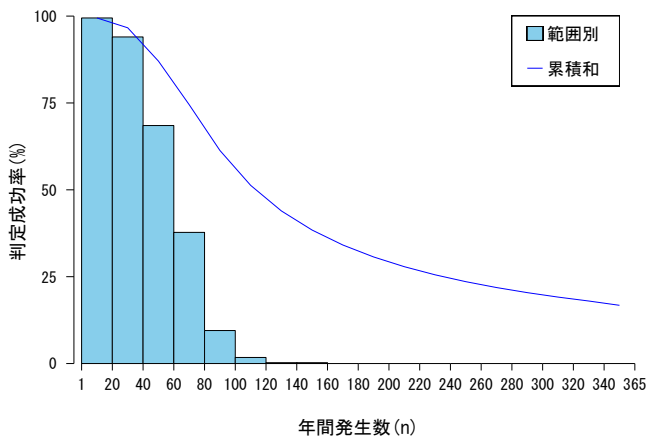


図 6 ランダムデータにおける範囲別の判定成功率とその累積和

範囲においては、閾値である 0.001 以上に信頼度が分布していることが分かる。

また、図 6 にランダムなデータにおける範囲別の判定成功率とその累積和を示す。判定成功率とは、ランダムなデータを予測困難な特徴をもつ作業として正しく分類できた割合を表す指標で、各 n の範囲内のデータ数に対して正しく分類できたデータ数の割合である。棒グラフは各 n の範囲における判定成功率を示し、折れ線グラフは n 以下の範囲における判定成功率を示す。

評価結果から、 $1 \leq n < 20$ の判定成功率は 99% 以上であった。このことから、 $1 \leq n < 20$ の範囲のデータのほとんどは信頼性が低いと判定されたことが分かる。このことから、 $1 \leq n < 20$ 、つまり、不規則に発生する作業で年間の発生回数が 20 回未満の作業においては、予測困難な特徴をもつ作業を正しく分類できていると言える。次に、 $20 \leq n$ の範囲について判定成功率を見ると、 $20 \leq n < 40$ で 94%、 $40 \leq n < 60$ で 68%、 $60 \leq n < 80$ で 38% である。また、 $n \leq 50$ での判定成功率 87% であり、不規則に発生する作業で年間の発生回数が 50 回未満の作業に対する判定成功率は高い水準を示していることが分かる。しかし、個別の範囲で見ると判定成功率は高いとは言えない。たとえば、 $60 \leq n < 80$ の個別の範囲としての判定成功率は 38% であるため、この範囲での判定成功率は $1 \leq n < 20$ の判定成功率と比べて低い。以上から、ほぼ確実に信頼性の判定に成功するのは $1 \leq n < 20$ の範囲であることが分かる。また、 $20 \leq n$ の範囲、つまり、不規則に発生する作業で年間の発生回数が 20 回以上の作業は判定に失敗することがあると言える。判定に失敗する範囲に関して、個別の範囲での判定成功率は n が大きくなるほど減少しており、 $120 \leq n < 140$ では 1% 未満である。これは 5.3.2 項で述べたように、 n の大きな作業については擬似的な周期性をもってしまいうため、信頼性が高くなってしまいうことが原因である。これらの結果から、予測困難な特徴をもつ作業を正しく分類できる n の範囲は $1 \leq n < 20$ であることが分かる。 $20 \leq n \leq 365$ の作業に対しては正しい分類が

できない。しかし、これは実用上問題にはならないと考えられる。なぜならば、現実において、不規則な間隔で年間 20 回以上も発生する作業は多くないと考えられるためである。また、そうした作業が仮にあったとしても、そこからつねに誤った予測が生じる場合は、当該の作業の予測を抑制するような操作を利用者に提供することでこの問題は解決すると考えられる。

6. おわりに

作業予測における予測結果に予測の妥当性を示す指標が含まれないため、利用者が予測結果の妥当性を吟味することが難しい問題を示した。対処として、予測結果の妥当性を示す指標である信頼性の概念を導入した。さらに、作業予測における信頼性の表現方法に基づいて信頼度を定義し、作業予測の信頼性評価手法を提案した。また、信頼性判定のための閾値を実データから調査し、得られた閾値に基づいて提案手法の評価を行った。評価結果から、不規則な間隔で発生する作業のうち年間発生数が 20 回未満の作業について低い信頼性を与えることを示した。

今後の課題として、この調査から得られた閾値が一般的な作業に対しても適用可能かという議論がある。なぜならば、この閾値はカレンダーの性質にも影響するためである。たとえば、本稿で利用したカレンダーは著者らの所属する研究室の共有カレンダーの予定であり、予定が複数の人間によって嚴重に管理されるため、(分類 G) 情報の欠落を含む周期的作業は比較的少ない。一方、個人用のカレンダーは共有カレンダーほど厳密に管理されておらず、前者とは違った性質をもつと考えられる。

謝辞 本研究の一部は、科学研究費補助金・基盤研究(C)(課題番号: 26330224) による研究費を得て実施した。ここに記して謝意を示す。

参考文献

- [1] 株式会社アイ・ティ・アール: クラウド時代のコラボレーション・ツールの方向性—グループウェアの再考, ITR White Paper (オンライン), 入手先 (http://www.itr.co.jp/library/public/ITR_WhitePaper/ITR_WP_C10090023.pdf) (参照 2015-01-14).
- [2] Zunino, A. and Campo, M.: Chronos: A multi-agent system for distributed automatic meeting scheduling, *Expert Systems with Applications*, Vol. 36, No. 3, pp. 7011–7018 (2009).
- [3] Mitchell, T. M., Caruana, R., Freitag, D., McDermott, J., Zabowski, D. et al.: Experience with a learning personal assistant, *Communications of the ACM*, Vol. 37, No. 7, pp. 80–91 (1994).
- [4] 田村洋人, 小川恭平, 竹内俊貴, 鳴海拓志, 谷川智洋, 廣瀬通孝: ライフログを用いた未来予測によるタスク進行管理手法の検討 (サービスや生活と VR/AR, 人工現実感), 電子情報通信学会技術研究報告. MVE, マルチメディア・仮想環境基礎, Vol. 112, No. 106, pp. 43–48 (2012).
- [5] 吉井英人, 北垣千広, 乃村能成, 谷口秀夫: 作業発生規則性に基づく作業予測手法と評価, 情報処理学会論文誌,

Vol. 56, No. 2, pp. 543–552 (2015).

- [6] 福本太郎, 澤村 一: 議論に基づくスケジュール調整エージェント, 電子情報通信学会論文誌, Vol. J91-D, No. 6, pp. 1506–1514 (2008).
- [7] 三原俊介, 乃村能成, 谷口秀夫, 南 裕也: 作業発生の規則性を扱うカレンダーシステムの評価, 情報処理学会論文誌, Vol. 54, No. 2, pp. 630–638 (2013).
- [8] 吉井英人, 乃村能成, 谷口秀夫: 作業発生の規則性に基づく作業予測手法, マルチメディア通信と分散処理ワークショップ (DPSWS2012) 論文集, Vol. 4, pp. 58–64 (2012).