

作業発生の規則性に基づく作業予測手法と評価

吉井 英人¹ 北垣 千広¹ 乃村 能成^{1,a)} 谷口 秀夫¹

受付日 2014年5月12日, 採録日 2014年11月10日

概要: 我々は, これまでの研究で作業発生の規則性というモデルを提案した. このモデルを利用することで将来の作業発生を予測できると思われるが, 具体的な予測の手法は, まだ明らかにされていない. 本稿では, まず作業発生の履歴から統計情報や周期に関する情報を抽出し, 繰返情報として定義する. 次に, 繰返情報に基づいて将来の作業発生を予測する手法について述べる. さらに, 周期的に発生する作業についていくつかの実例を示し, 本手法を適用した結果について考察する.

キーワード: 作業予測, カレンダー, 計画立案

A Method for Forecasting the Future Events Based on Task Patterns

HIDETO YOSHII¹ CHIHIRO KITAGAKI¹ YOSHINARI NOMURA^{1,a)} HIDEO TANIGUCHI¹

Received: May 12, 2014, Accepted: November 10, 2014

Abstract: We proposed a model that is capable of handling the patterns of ambiguous recurring tasks. Although, this model is believed useful for prediction of the tasks which might happen in the future, it is not clear how extract the necessary information from the model. This paper proposes a method for forecasting the future tasks by extracting statistical information from the model. First, some additional terms to express the statistical information are introduced. Next, we propose a new method for forecasting. Finally, we show typical examples of recurring tasks, and perform some evaluation of the method on these examples.

Keywords: task forecasting, calendar, planning

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, 多くの作業はある程度決まった周期性に基づいて発生しているからである. たとえば, ミーティングや忘年会は, それぞれ「約2週間に1回」や「毎年12月下旬」といった曖昧な周期を持つ. この周期性に基づき作業の発生を確認できれば, 将来の作業予測や仕事の引き継ぎ時の情報伝搬に有用だと考えられる. 過去の履歴を利用する研究はこれまでにいくつか存在する [1], [2]. また, 仕事引き継ぎの観点から過去の作業履歴の利用法を探る研究も現れてきている [3], [4]. し

かしながら, これらの研究で提案されているモデルは, 一般的なカレンダーシステムが持つデータモデルとは異なるため, データ収集コストが高かったり, 既存のグループウェアとの親和性をとることが難しかったりする.

そこで, 我々は, 曖昧な周期性を扱うモデルとして作業発生の規則性というモデルを提案している [5]. このモデルは, シンプルで既存のカレンダーシステムが扱うモデルと親和性の高いモデルである. このモデルは繰り返し発生する同様の作業を単純な集合として表現する. この集合を利用することで, 将来の作業発生を予測できる.

本稿では, 将来の作業発生を予測するために, 作業発生の履歴から統計情報や周期に関する情報を抽出し, 繰返情報として定義する. また, 繰返情報に基づいてシステムが将来の作業発生を予測する手法について述べる.

関連研究として, ユーザとその周囲の状況をセンサで取

¹ 岡山大学大学院自然科学研究科
Graduate School of Natural Science and Technology,
Okayama University, Okayama 700-8530, Japan

a) nom@cs.okayama-u.ac.jp

得して分析し、将来の行動を予測する研究が行われている [6], [7]. また、次の行動予測することで、行動に合わせたサービスを提供しようという研究もある [8]. これらの研究は、次の人間の「動作の予測」に焦点をあてたものである. ここでは、これらの研究とは異なり、将来しなければならない「作業を提示」することにより、将来の計画を立てる作業（計画立案）の支援を目的としている.

本稿では、まず、繰り返し情報モデルを提案し、作業発生 の規則性モデルとの関係について説明する. 次に、多くの 作業がユーザが感じている作業の周期より長い周期として 発生していることに着目した作業予測手法を提案する. さらに、周期的に発生している作業についていくつかの実例 を示し、本手法を適用した場合の結果について考察する.

2. 作業発生の規則性と繰り返し情報

2.1 作業発生の規則性

周期性と関連性を集合として表現する方法が提案されて いる. この作業発生に関わる周期性と関連性を表現したモ デルを作業発生の規則性と呼ぶ. 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.

タスク (Task) : タスクとは、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持つ.

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 作業発生の規則性に基づく作業予測

タスク発生を予測する単純な方法として、リカーレンス 内のタスク発生日付の間隔の平均（以降、作業発生間隔の 平均日数とする）を求め、平均をそのまま次のタスク発生 の予測に利用する方法が考えられる. しかしながら、この 方法ではタスク発生をうまく予測できない. これは、主に 以下の2つに起因する.

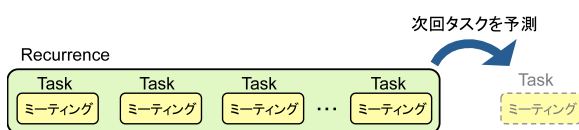


図1 リカーレンスを利用した作業予測

Fig. 1 Task forecasting using recurrence.

(1) 曜日の偏りを考慮していない.

(2) 長期休暇や年末年始といった発生しない期間がある.

作業発生間隔の平均日数のみを用いる予測では、休日には発生しないはずのタスクが日曜日に予測されることがあ る. また、夏期休暇といった長期休暇に容易にぶつかって しまう. さらに、長期休暇が周期の計算に勘定されるので、 平均値が不当に大きくなる. つまり、作業発生間隔の平均 日数のみを用いる予測では、タスクの発生をうまく予測で きない. したがって、タスク発生の予測には作業発生間隔 の平均日数以外の情報が必要である. そこで、単なる履歴 であるリカーレンスからどのような情報を抽出すれば、予 測に有用か考える.

2.3 繰り返し情報

タスク発生の予測に必要な情報を検討するために、実際 の作業発生を分析した. 作業の典型的な例として、著者ら が所属する研究グループでのミーティングについて事例調 査を行った. ミーティングは、約2週間に1回程度の周期 で実施すると決めている作業であるが、実際には、参加者 の都合や休暇の影響があり、その周期はいまいさを含ん でいる. 図2は、2009年度、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. 分析の結果、以下の6つのことが分かる.

(1) 発生する曜日に偏りがあるがいつも同じではない
作業が発生する曜日には偏りがある. これは、他の作業の 曜日が固定されていたり、日付決定の際に2週間という周 期を意識した意思が働いているからだと考えられる. 曜日 ごとの発生頻度を表1に示す. 表中では、祝日も曜日の1 つとしている. 表1から、月曜日が全体の約40%を占めて いることが分かる. また、土日や祝日にはまったく発生し ていないことが分かる.

(2) 発生しない期間が存在する

夏期休暇、年末年始といった作業が発生しない期間が明確

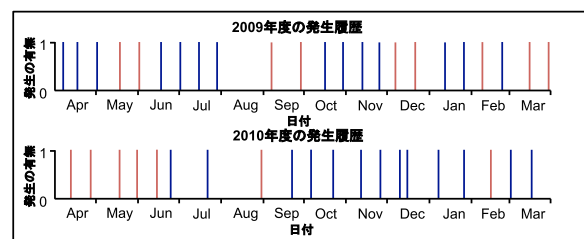


図2 2009年度、2010年度のミーティングの発生履歴

Fig. 2 History of occurrences in 2009 and 2010.

表1 曜日ごとの発生頻度

Table 1 Frequency of task occurrence by day of week.

曜日	日	月	火	水	木	金	土	祝日	合計
発生回数 (回)	0	16	7	9	7	4	0	0	43
発生割合 (%)	0	37	16	21	16	10	0	0	100

表 2 月ごとの発生回数

Table 2 Frequency of task occurrence by month.

月	1	2	3	4	5	6	7	8	9	10	11	12	合計
発生回数 (回)	4	3	4	4	4	4	4	1	3	4	4	4	43

表 3 週ごとの発生頻度

Table 3 Frequency of task occurrence by week number.

週	1	2	3	4	5	合計
発生回数 (回)	9	8	13	8	5	43
発生割合 (%)	21	19	30	19	11	100

表 4 開始時刻の発生頻度

Table 4 Frequency of task occurrence by start time.

通番	開始時刻	発生回数 (回)	累積割合 (%)
1	13:00	12	28
2	10:00	9	49
3	09:30	3	56
4	15:00	3	63
5	09:00	2	67
6	13:10	2	72
7	14:00	2	77
8	10:25	1	79
9	13:30	1	81
10	14:15	1	84

表 5 作業時間の発生頻度

Table 5 Frequency of task occurrence by time-frame length.

通番	作業時間 (分)	発生回数 (回)	累積割合 (%)
1	120 ~ 150	17	40
2	150 ~ 180	12	67
3	90 ~ 120	7	84
4	180 ~ 210	4	93
5	~ 90	2	98
6	210 ~	1	100

に表れている。具体的には、7/30-8/29, また 12/22-1/6 が各年度に共通して作業が発生しない期間と観測される。

(3) 月ごとの発生回数に偏りがある

作業が発生する月にも偏りがある。月ごとの発生回数を表 2 に示す。表 2 から、月ごとの回数はほぼ同じであるが、発生回数が少ない月があることが分かる。具体的には、2月、8月、9月が発生回数が少ない月である。特に8月は、2009年度、2010年度を通して、1回しか発生していない。これは、長期休暇が影響していると考えられる。

(4) 週ごとの発生割合に偏りがある

作業が発生する週にも偏りがある。週ごとの発生頻度を表 3 に示す。表中では、第1週を月の1日目から7日目までとし、7日おきに週が切り替わるものとしている。表 3 から、週ごとの発生割合にもある程度偏りがあることが分かる。第3週目の発生割合は、30%であり、他の週に比べて少し高い値である。

(5) 時間帯に偏りがある

時間帯を開始時刻と作業時間の組みと考えた場合、作業が発生する時間帯はある程度決まっている。開始時刻、作業時間の発生頻度を表 4、表 5 に示す。これによると、開始

時刻は、発生回数の多い上位3つで全体の50%以上を占めており、上位9つまでで全体の80%を超えている。また、作業時間を30分単位で分類した場合、120分から150分が全体の40%を占めていることが分かる。

(6) 年間における差異は大きくない

年単位で発生系列を比較してみると、1年を周期とした相似形を観測できる。つまり、作業自身は、2週間周期で発生しているタスクととらえているが、1年を周期として見直すと、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1つのリカーレンスを形成していると考えられる。この観点から、約1年という周期で各回のミーティングが発生していると考えれば、長期休暇にも同様の周期を見出すことができる。

上記(1),(2),(3),(4),(5)については、我々が普段行っている作業の経験からも分かることである。しかし、具体的な数値についてはあまり意識しないので、これらを数値化し利用することで予測に役立つと考えられる。上記(6)については、今回の事例調査の結果から分かったことである。このことは、昨年のカレンダーを参考にして、今年の前定を決める行動とよく一致しており、いわれれば当然のように映るかもしれない。しかし、我々は約2週間に1回発生すると意識しているミーティングは、その周期にとらわれ、1年前の前定を見返すことはあまりない。つまり、短周期で発生すると意識している作業については、それより長い周期をあまり考慮に入れない傾向にあるといえる。しかし、事例調査の結果からも分かるように、1年単位で作業の発生を眺めた場合、1年ごとに作業の発生が対応していると考えれば、予測に役立つ可能性がある。

そのほか、上記の調査では得られなかったが、誕生日のように毎年決まった日付に発生する作業もある。このような前定の予測も考えて、以下の7つを予測の際に考慮する。

(1) 昨年発生間隔を考慮

分析の結果から、作業発生は年を単位とした相似形を持っていることが分かる。図2から各年度において、作業の発生時期は多少異なるものの、作業発生間隔は類似していることが観測される。このことから、発生間隔として昨年の同時期に発生した作業の発生間隔を用いることで作業発生を予測する。昨年の作業発生間隔は、昨年の同時期に発生した作業とその次の作業との日数差から求められる。

(2) 曜日を考慮

表1から、発生する割合が高い曜日やまったく発生しない曜日があることが分かる。曜日ごとの発生割合を求め、作業が発生する割合の高い曜日や作業が発生しない曜日を考慮して予測する。曜日ごとの発生割合は、リカーレンス内のタスクを解析し、統計情報として得られる。

(3) 作業が発生しない期間を考慮

作業が発生しない期間を求める方法として、ユーザから入力してもらう方法が考えられる。しかし、作業が発生しな

い期間は事前に分らない。また、入力自体が手間になる。したがって、これは有用な方法ではない。そこで、昨年の作業発生間隔を用いる。図 2 から、各年度に共通して、夏期休暇をはさんだ作業の間隔が大きく開いていることが分かる。このことから、昨年の作業発生間隔を用いた場合、夏期休暇を避けて予測できると考えられる。夏期休暇以外の作業が発生しない期間についても同様のことがいえる。

(4) 月を考慮

表 2 から、発生する回数が少ない月があることが分かる。月ごとの発生割合を求め、月を考慮して予測する。月ごとの発生割合は、曜日同様、統計情報として得られる。

(5) 週を考慮

表 3 から、発生する週にも偏りがあることが分かる。週ごとの発生割合を求め、週を考慮して予測する。週ごとの発生割合は、曜日同様、統計情報として得られる。

(6) 日付を考慮

誕生日のように毎年決まった日付に発生する作業もある。このような固定的な日付で発生する作業に関しては、日付をユーザに入力してもらう方法が考えられる。しかし、誕生日のように毎年同じ日付に発生することが明確な作業がある一方、毎年同じ日付に発生しているが、それを認識していない作業もある。また、日付の入力自体も手間となる。このため、日付の発生割合を用いる。日付の発生割合は、曜日同様、統計情報として得られる。

(7) 時間帯を考慮

表 4 から、開始時刻はある程度決まっていることが分かる。作業時間に関しても、表 5 から、発生回数の多い作業時間がほぼ決まっており、その作業時間の近辺でばらつきがあることが分かる。以上から、開始時刻は発生回数の多い上位 3 つ、作業時間は最も発生回数の多いものを用い、この組合せを時間帯とする。開始時刻と作業時間に関しては、曜日の発生割合と同様に統計情報として得られる。

曜日や作業が発生しない期間が発生に影響を及ぼす作業がある一方で、「散髪に行く」といった曜日や発生しない期間とはあまり関係なく、その周期が重要なものもある。このような作業では、その周期を発見して提示する方が有用な場合もある。そこで、作業発生間隔の平均日数も予測に必要な情報の 1 つとする。

作業予測に必要な情報を繰返し情報とする。以上のことから、以下の 8 つを繰返し情報とする。

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 日のまでの 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

2.4 1 年を周期とした周期性の継承

2.4.1 概要

1 年を周期とした作業発生について前節の事例をもとに詳しく述べる。2.3 節で述べたように、作業は 1 年を周期とした相似形を持つ。これが明確に現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ある。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 年度の作業発生から、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予測できる。12 月から 1 月までの作業発生についても同様のことがいえる。

2.4.2 1 年を周期とした周期性の継承

前節の結果から、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、作業発生を予測する。

2.4.3 予測の手順

2.4.2 項の結果をふまえ、1 年を周期とした周期性の継承に基づく作業予測手法を提案する。以下で予測の手順について説明する。

- (1) リカーレンス内の最後に発生したタスクの日付（最終作業日）から 1 年（365 日）前の日付を求める。
- (2) 1 年前の日付を基準日とし、リカーレンス内のタスクの中から、基準日に最も近い作業日を求める。これを対応日とする。
- (3) 対応日とその次の作業日との日数差（昨年の作業発生間隔）を求める。
- (4) 昨年の作業発生間隔を最終作業日に加えた日付を候補日 1 とする。
- (5) 作業発生間隔を用いて、候補日 1 の前後数日を候補日として求め、候補日 1 からの日数差と曜日、月、週、日付の発生割合をパラメータとした評価関数で重み付けをする。
- (6) 候補日の中で評価値が最も高い日付を次回作業日とする。
- (7) 次回作業日に時間帯の情報を加えた日時を次回作業日時として予測する。

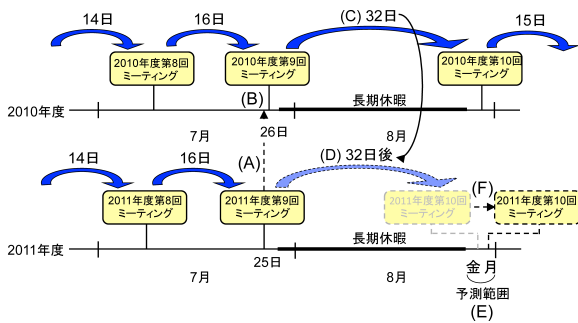


図 3 1年を周期とした周期性の継承に基づく作業予測手法の適用例
Fig. 3 Forecasting operation based on one-year interval.

次に、この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 日の 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 1 とする。
- (E) 候補日 1 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。

2.4.4 予測のアルゴリズム

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予測のアルゴリズムについて述べる。ここでは、予測のアルゴリズムを昨年の発生間隔を用いて候補日を求める部分と統計情報を用いて日付を調整する部分に分けて説明する。

- (1) 昨年の発生間隔を用いて候補日を求める
 - 昨年の発生間隔を用いて候補日を求めるアルゴリズムを図 4 に示す。図 4 の 11 行目から 20 行目までが候補日を求める部分である。具体的には、以下のように、候補日を求める。
 - (a) d_{last} をリカーレンス D の中で最後に発生したタスクの日付とする。
 - (b) d_{last} から、1 年前の日付を求め、 d_{last_year} とする。
 - (c) D の中で d_{last_year} に最も近い日付を求め、

```

01 function next(d:date): date
02 begin
03   Return the next tasking date of d within D
04 end;
05
06 function previous(d:date): date
07 begin
08   Return the previous tasking date of d within D
09 end;
10
11 function pivot(dlast:date): date
12 begin
13   dlast_year := dlast - 365
14   dcorrespondence := if | dlast_year - previous(dlast_year) | < | dlast_year - next(dlast_year) |
15                       then previous(dlast_year)
16                       else next(dlast_year)
17   i := next(dcorrespondence) - dcorrespondence
18   dpivot = dlast + i
19   Return dpivot
20 end
    
```

図 4 昨年の発生間隔を用いて、候補日を求めるアルゴリズム
Fig. 4 Algorithm for calculating candidates using last-year's interval.

```

01 dpivot := pivot(dlast)
02 loop:
03   Dcandidate := (dpivot - rstandard_deviation · · · · · dpivot + rstandard_deviation)
04   for each d in Dcandidate
05     begin
06       E(d) := a1 * day_ratio + a2 * month_ratio +
07              a3 * week_ratio + a4 * date_ratio          ... (A)
08     end
09
10   dmax := Max(E)
11
12   if dmax = 0
13     then dpivot := dpivot + rstandard_deviation
14          Dcandidate := []
15          goto loop
16   else Return dmax
    
```

図 5 統計情報を用いて、日付を調整するアルゴリズム
Fig. 5 Algorithm for adjusting candidates according to statistical information.

- (d) $d_{correspondence}$ と $d_{correspondence}$ の次の作業の日付の間隔 i を求める。
 - (e) $d_{last} + i$ を予測の最初の候補日 d_{pivot} とする。
 - (2) 統計情報を用いて日付を調整する
 - 統計情報を用いて日付を調整するアルゴリズムを図 5 に示す。図 5 では、以下のように日付を調整する。
 - (a) 図 4 のアルゴリズムを用いて、 d_{last} から予測の候補日を求め、 d_{pivot} とする。
 - (b) d_{pivot} を基準として、 D 内の作業発生間隔の標準偏差 $r_{standard_deviation}$ を用いて、候補日の集合 $D_{candidate}$ を求める。
 - (c) $D_{candidate}$ 内の各日付に対して、図 5 内の式 (A) を用いて重み付けを行い、重み付けを行った結果を $E(d)$ とする。
 - (d) 各候補日 $E(d)$ の中で最大の重みを持つ日付を求め、 d_{max} とする。
 - (e) d_{max} が 0 の場合は、 d_{pivot} を求め直し、処理を繰り返す。 d_{max} が 0 でない場合は、 d_{max} を予測日とする。
- 図 5 内の式 (A) は、曜日、月、週、および日付を考慮し、重み付けするものである。式 (A) における $a1, a2, a3, a4$

は、過去の発生割合の情報に基づいて決定する。ここでは、作業発生の有無を 0, 1 の値に置き換えたものを従属変数とし、それらを曜日, 月, 週, および日付のそれぞれからなる独立変数の線形結合で表現することで、線形回帰分析を行う。以下具体的に説明する。

ある日付 d において、作業発生の有無 $Y(d)$ は、 $Y(d) = a_1 \cdot \text{day}(d) + a_2 \cdot \text{month}(d) + a_3 \cdot \text{week}(d) + a_4 \cdot \text{date}(d)$ と表すことができるとする。ここで、 $\text{day}(d)$ は、 d の曜日に対応する要素だけが 1 となる (祝日を曜日として考える) 8 次元のベクトルである。つまり、 $\text{day}(2014/8/18) = (1, 0, 0, 0, 0, 0, 0, 0)$ 、 $\text{day}(2014/8/19) = (0, 1, 0, 0, 0, 0, 0, 0)$ となる。同様に独立変数として扱われる $\text{month}(d)$ 、 $\text{week}(d)$ 、 $\text{date}(d)$ もベクトルであり、それぞれ、12, 5, 365 次元である。したがって、 a_1, a_2, a_3, a_4 も対応する独立変数と同じ次元を持つベクトルとして表現される。

実際の作業発生履歴を $T(d)$ とすると、残差 $|T(d) - Y(d)|^2$ を最小化するように線形回帰計算をすることで、 a_1, a_2, a_3, a_4 を求めた。実際の計算やデータ整形は、R 言語で 500 行程度のプログラムで実現した。

3. 評価

3.1 評価の観点

我々は普段、カレンダーシステムを用いて予定を管理することが多い。この際、将来発生すると推測されるタスクを事前に登録することがある。将来発生すると推測されるタスクをその発生時期を考慮しながらカレンダーに仮登録していく作業をここでは計画立案と呼ぶ。計画立案は、おおよそ以下の手順で行われると想定する。

- (1) 計画立案したい月のカレンダーと前年の同月のカレンダーを並べて開く。
- (2) 前年のカレンダーの各タスクについて
 - (a) そのタスクが、再び発生するかどうか判断する。
 - (b) 発生すると判断したならば、立案したい月のカレンダーに複写登録する。
- (3) 計画について納得できるように調整し確認する。
これを必要な月について繰り返す。

計画立案において、(2-a) のように過去のカレンダー上に登録されている各タスクについて、そのタスクが再び発生するかどうか判断する。また、(2-b) のように発生すると判断した場合、立案したい月のカレンダーに複写登録する作業を行う。これらの作業には、手間がかかることが想定される。なぜならば、過去のカレンダーの中には、繰り返し発生しないタスクも存在するため、繰り返し発生するかどうかを個々のタスクについて吟味する必要がある。また、繰り返し発生すると判断した場合でも、カレンダーに複写する際に、発生時期を考えなければならぬためである。

繰り返し発生するタスクに関して、システムが発生時期を予測して提示してくれる場合、(2-a) や (2-b) にかかる

手間を軽減できると考えられる。

このことから、作業予測は計画立案において有用であるといえる。しかし、作業予測が十分な精度を持ち合わせていない場合、たとえばまったく別の日時に予測されるといった場合に、作業予測の有用性は低下する。なぜならば、誤った日時に予測されたタスクを修正する手間がかかるからである。つまり、作業予測においては、精度が重要であるといえる。

ここで注意すべきは、作業予測における予測の精度は実際のタスク発生日時と一致するかどうかではなく、予測された結果をユーザが納得できるかどうかであるということである。なぜならば、実際のタスクの発生とその日時は直近にならないと決まらないため、計画立案の段階では、正確な日時は分からないからである。また、計画立案の段階においては、タスクの実施日として複数の候補が存在することが考えられる。たとえば、ミーティングの計画を立てる際には、A 日、B 日、C 日といった複数の候補が考えられる。この中で、仮に B 日を選んだとして、その決定を裏付ける明確な根拠はないことがある。この場合、選ばれた B 日だけではなく、A 日、C 日も妥当な日付であり、これらを予測してもシステムとしては正当と考えられる。むしろ、タスクの発生日として妥当でない日付 (休日に発生しないタスクを休日に予測する) を提示しないことが重要だといえる。つまり、ここで問題とすべきは、どれだけユーザが納得できる予定を提示できるかということである。

これをふまえ、作業予測の精度について評価する。具体的には、タスク発生を予測し、予測したタスクにおいて、妥当だと判断したタスクの数を比較する。

これにより、作業予測の精度の評価を行う。

3.2 評価環境の準備

まず、実験に用いた初期データについて述べる。あらかじめ、2011 年度と 2012 年度の過去 2 年間の予定を作業発生の規則性に基づくカレンダーシステムに登録しておく。加えて、2011 年度と 2012 年度分についてリカーレンスの関係を手動で与える。表 6 は、予測に用いたタスクを示したものである。表 6 では、作業ごとの 2011 年度のタスク数、2012 年度のタスク数、および 2011 年度と 2012 年度のタスク数の合計をあげている。また、表 6 を各分類ごとに整理した表を表 7 に示す。

予測に用いるそれぞれの作業について説明する。それぞれの作業を以下の 5 つに分類する。

(1) A (A1~A26)

A1 から A26 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、次の作業日時を決定する際に、周期を意識した意思が働く。

表 6 予測に用いる作業

Table 6 Task list for benchmark.

通番	タスク	作業名	2011 年度	2012 年度	合計
A1	研究ミーティング	ミーティング 1	23	23	46
A2	(約 2 週間に 1 回)	ミーティング 2	15	18	33
A3	進捗報告会	進捗報告会 1	12	13	25
A4	(約 1 カ月に 1 回)	進捗報告会 2	11	12	23
A5		進捗報告会 3	11	12	23
A6	年中行事 (半期に 1 回)	誕生日会	2	1	3
A7	年中行事	研修会	1	1	2
A8	(約 1 年に 1 回)	忘年会 1	1	1	2
A9		忘年会 2	1	1	2
A10		送別会 1	1	1	2
A11		送別会 2	1	1	2
A12		新人歓迎会	1	1	2
A13		暑気払い	1	1	2
A14		卒業アルバム写真撮影	1	1	2
A15		M2 中間発表	1	1	2
A16		B4 中間発表	1	1	2
A17		防災訓練	1	1	2
A18		M1 論文紹介	1	1	2
A19		M2 論文紹介	1	1	2
A20		卒論発表	1	1	2
A21		修論発表	1	1	2
A22		オープンキャンパス	1	1	2
A23		研究室大掃除	1	1	2
A24		大学の卒業式	1	1	2
A25		大学の推薦入試	1	1	2
A26		大学の編入試験	1	1	2
B1	スポーツ大会 (季節ごと)	スポーツ大会	4	4	8
C1	講義	講義 1	14	15	29
C2	(約 1 週間に 1 回 (半期))	講義 2	12	13	25
C3		講義 3	12	11	23
C4		講義 4	15	15	30
C5		講義 5	14	14	28
D1	レクリエーション	レクリエーション	3	3	6
E1	恒例行事	大学の前期入試	1	1	2
E2	(決まった日付)	大学の後期入試	1	1	2
E3		大学の入学式	1	1	2
合計			171	177	348

表 7 予測に用いる作業 (分類別)

Table 7 Task list for benchmark by task type.

分類	2011 年度	2012 年度	合計
A	94	99	193
B	4	4	8
C	67	68	135
D	3	3	6
E	3	3	6
合計	171	177	348

(2) B (B1)

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、A1 から A26 の作業ほど強くない。

(3) C (C1~C5)

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、A1 から A26 の作業とは異なる。

(4) D (D1)

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。毎年 3, 4 回程度発生している。

(5) E (E1~E3)

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。

表 8 タスク C1 における評価関数パラメータ

Table 8 Fitting parameter for task C1.

	要素	パラメータ	p 値
a1	金曜	-0.003692	0.911030
	土曜	-0.004192	0.898907
	日曜	-0.005071	0.877560
	月曜	-0.004070	0.902706
	火曜	-0.004195	0.898694
	水曜	0.284249	< 2e-16 ***
	木曜	-0.002318	0.943999
	祝日	NA	NA
a2	1 週	-0.008614	0.725007
	2 週	0.013551	0.579961
	3 週	0.015268	0.532927
	4 週	0.009292	0.704312
	5 週	NA	NA
a3	4 月	0.077169	0.008443 **
	5 月	0.108063	0.000215 ***
	6 月	0.122532	3.11e-05 ***
	7 月	0.124435	1.99e-05 ***
	8 月	0.002299	0.936811
	9 月	-0.006246	0.830876
	10 月	-0.009258	0.749402
	11 月	-0.006234	0.831248
	12 月	-0.004718	0.870670
	1 月	-0.009436	0.744927
	2 月	-0.013557	0.647697
	3 月	NA	NA

表 6 のデータをもとに、提案手法を含む 4 つの手法で 2013 年度の 1 年分のカレンダーを作成した。

手法 1: 作業発生間隔の平均日数 (平均)

作業発生間隔の平均日数のみを用いる手法である。

手法 2: 作業発生間隔の平均日数と発生確率 (平均+統計)

手法 1 に 2.3 節で述べた曜日, 月, 週, および日付情報と時間帯情報による補正を加えた手法である。

手法 3: 昨年の作業発生間隔 (1 年前)

昨年の作業発生間隔のみを用いる手法である。

手法 4 (提案手法): 昨年の作業発生間隔と発生確率 (1 年前+統計)

手法 3 に曜日, 月, 週, および日付情報と時間帯情報による補正を加えた手法である。

2.4.4 項で述べたパラメータ a1, a2, a3, a4 には、2011 年度, 2012 年度の過去 2 年間の予定から最適な値を求めて設定した。一例として、表 6 中のタスク C1 における各パラメータを表 8 に示す。C1 は、各年度の前期 (4 月から 7 月) の間、毎週水曜日に実施される講義である。a1 においては、水曜日、a2 においては、4 月から 7 月が有意であることを p 値が示している。月の第何日目であるか (a3) にはあまり依存しないで発生していることも分かる。ここで、a4 (特定の日付) については、複数年度で同じ月日に

表 9 予測結果の比較

Table 9 Results of forecasting by method.

通番	手法	妥当	妥当でない			合計
			曜日×	期間×	合計	
手法 1	平均	77	133	140	182	259
手法 2	平均+統計	132	1	138	138	270
手法 3	1 年前	136	27	23	42	178
手法 4	1 年前+統計	163	0	13	13	176

発生したときのみ 1 を入れることとした。このため、C1 においてはすべて 0 となり、候補日の評価には影響を与えないので、表中には記載していない。また、NA と記載のある部分については、他の要素に従属している（つまり 3 月かどうかは、他の 11 カ月の 1/0 から完全に求められる）ため、次元が 1 つ減るだけで、パラメータとしては意味を持たない。

また、今回の評価では、時間帯は最も発生回数の多いものを用いた。

作成したカレンダーをここでは、未来カレンダーと呼ぶ。未来カレンダーには、予測された予定以外は登録されていないこととする。今回の評価では、各手法で予測したタスクが妥当かどうかを人間が 1 つ 1 つ判断し、妥当だとした数、妥当でないとした数を計測する。また、その判断理由を吟味する。なお、判断基準は、評価に先立ってあらかじめ用意しておいた。具体的には、すべてのリカーレンスの系列に対して、2013 年度で発生する妥当な日付の一覧を作成して列挙しておく。たとえば、忘年会 1 (A8) であれば、12 月上旬から中旬の平日を妥当な日として（これは、被験者の経験に基づく主観による）マークし、一覧を作成しておいた。これにより、各手法が生成する予測に対する判断基準に一貫性を持たせている。

平均を用いる手法 1 と 1 年前を用いる手法 3 を比較することで、1 年前を用いた場合の有用性を示せる。平均のみを用いる手法 1 と平均+統計を用いる手法 2、また、1 年前のみを用いる手法 3 と 1 年前+統計を用いる手法 4 を比較することで、統計情報を用いる有用性を示せる。さらに、各分類 (A-E) における差を明確にする。

3.3 実測評価

3.3.1 作業予測の精度の評価

前節で述べた 4 つの手法を用いて作成した未来カレンダーを目視で確認し、予測された個々のタスクについて、日時が妥当かどうかを判断する。妥当でないとしたタスクは、未来カレンダーから除外する。今回の評価における妥当性の判断基準は、予測された日時にタスクが発生しそうでないと実験者が判断するかどうかである。

この作業後に未来カレンダーに残ったタスクは、妥当だと判断されたタスクであるとし、このタスクの数を比較する。

表 9 に評価結果を示す。また、表 9 における作業の分

表 10 予測結果の分類別の比較

Table 10 Results of forecasting by method and task type.

通番	手法	分類	妥当	妥当でない			合計
				曜日×	期間×	合計	
手法 1	平均	A	62	30	19	39	101
		B	1	0	3	3	4
		C	10	103	116	138	148
		D	3	0	0	0	3
		E	1	0	2	2	3
手法 2	平均+曜日	A	90	0	14	14	104
		B	3	0	1	1	4
		C	33	1	123	123	156
		D	3	0	0	0	3
		E	3	0	0	0	3
手法 3	1 年前	A	70	21	17	30	100
		B	3	0	1	1	4
		C	61	6	2	8	69
		D	1	0	1	1	2
		E	1	0	2	2	3
手法 4	1 年前+曜日	A	91	0	9	9	100
		B	3	0	1	0	4
		C	65	0	2	2	67
		D	1	0	1	1	2
		E	3	0	0	3	3

類に従い、比較した結果を表 10 に示す。予測した日時が妥当でないとした理由として、以下の 2 つがある。

(1) 曜日×

発生しない曜日に予測された場合である。つまり、曜日が妥当でないタスクである。この際、祝日も独立した 1 つの曜日として考える。

(2) 期間×

夏期休暇や年末年始といった発生しない期間に予測された場合である。つまり、期間が妥当でないタスクである。別の年度に予測された場合も期間が妥当でないとする。また、決まった日付に発生するタスクの日付が異なる場合も期間が妥当でないとする。

予測されたタスクの中には曜日、期間ともに妥当でないものも存在する。この場合、曜日、期間ともに×としている。このため、妥当でないタスクの数は、単純に曜日×、期間×を合計した値となっていない場合がある。

それぞれ 4 つの手法における評価結果について分析する。

3.3.2 作業発生間隔の平均日数を用いる手法

手法 1 と手法 2 は、どちらも作業発生間隔の平均日数を用いる手法である。表 9 に示すとおり、手法 1 において予測されたタスクの合計は 259 件で、うち、妥当なタスクは 77 件、妥当でないタスクは 182 件であった。また、「曜日×」のタスクは 133 件、「期間×」のタスクは 140 件であった。手法 2 では、合計 270 件で、うち、妥当なタスクは 132 件、妥当でないタスクは 138 件であった。また、「曜日×」のタスクは 1 件、「期間×」のタスクは 138 件であった。

手法 1 では、予測されたタスクの半数以上が妥当でないタスクであった。これは、曜日と期間をまったく考慮して

いないためである。曜日の発生割合を用いることにより、曜日を考慮して予測できることは、表9の手法1と手法2の結果を比較すると分かる。手法1では、「曜日×」のタスクが133件であるのに対して、手法2では1件となっている。このことから曜日を考慮することの有用性が分かる。

表9では、「期間×」のタスクは、手法1で140件、手法2で138件とそれほど差は見られない。分類別に比較した場合、表10から、分類Cのタスクが「期間×」の大半を占めていることが分かる。具体的には、手法1で116件、手法2で123件であり、全「期間×」項目の8割を占める。これは、講義の場合、作業は半期のみであるにもかかわらず、手法1、手法2では、講義が発生しない半期分も周期として計算する。このため、1週間よりも周期が大きくなってしまふ。また、講義が2012年度の前半に発生した場合、発生しないはずの2012年度の後半に予測されるためである。このことから、作業発生間隔の平均日数を用いる手法では、発生しない期間に関しては考慮できないことが分かる。

分類Eの同じ日付に発生する作業に関して、手法1では、3件のうち、2件がうまく予測できていない。これは、2012年度が閏年のため、1年の発生間隔にずれが生じたためである。手法2では、日付を考慮しているため、閏年の影響を受けずに分類Eの3件について、うまく予測できている。

3.3.3 昨年の作業発生間隔を用いる手法

手法3と手法4は、どちらも昨年の作業発生間隔を用いる手法である。手法3において予測されたタスクの合計は、178件で、うち、妥当なタスクは136件、妥当でないタスクは42件であった。また、「曜日×」のタスクは27件、「期間×」のタスクは23件であった。手法4では合計176件で、うち、妥当なタスクは163件、妥当でないタスクは13件であった。また、「曜日×」のタスクは0件、「期間×」のタスクは13件であった。

昨年の作業発生間隔を用いる手法3、手法4では、作業発生間隔の平均日数を用いる手法1、手法2と比べて、「期間×」のタスクが少ない。表10に示すとおり、分類別に見た場合でも、A-Eの5つの分類ほぼすべてにおいて、手法3、手法4の方が「期間×」のタスクが少ないことが分かる。特に、分類Cにおいては、その差が顕著に現れている。これは、昨年の作業発生間隔を用いることで、発生しない期間を考慮できているからである。具体的には、夏期休暇や年末年始などの作業が発生しない期間を考慮できている。また、講義のように半期だけ行われる作業は、発生しない半期を避けて予測できているためである。

表10から、手法1、手法3はどちらも曜日を考慮していないにもかかわらず、「曜日×」は明らかに手法3の方が少ないことが分かる。具体的には、分類Aの作業に関して、「曜日×」は、手法1では30件であるのに対して、手法3

では21件である。また、分類Cの作業に関しては、手法1では103件であるのに対して、手法3では6件である。これは、昨年の作業発生間隔を用いることで曜日をある程度考慮できていることを示す。たとえば、約2週間に1回という周期で発生し、土日には発生しない作業を考える。この作業が月曜日に発生した場合、次回は月曜日から金曜日のいずれかの曜日に発生し、作業日が土日となる間隔では発生しない。この間隔が次年度においても適用されるため、自然に曜日を考慮できている。しかし、表9から分かるように、手法3でも、依然とし「曜日×」のタスクが27件予測された。手法4では、曜日を考慮することにより、この27件に関しても正しく予測できている、「曜日×」のタスクは0件となっている。

分類Eの同じ日付に発生する作業に関して、平均間隔を用いる手法1と同様に、手法3ではうまく予測できていないが、日付を考慮している手法4ではうまく予測できている。このことより、提案手法では、日付を考慮できていることが分かる。

昨年の作業発生間隔と曜日、月、週、および日付を考慮した手法4では、176件中163件と約9割が妥当なタスクである。5つの分類のいずれにおいても、手法4が妥当なタスク数が最も多い。このことから、4つの手法の中では、手法4が最も精度の高い予測であるといえる。

ここで、手法4における妥当でない13件について考察する。13件は、いずれも「期間×」のタスクであり、基本的には、年度間における発生パターンの差異が大きい場合に予測に失敗する。年度間の差異は、大きく分けて4つの理由によることが分かった。

(1) もともと年間の予定が明確でない場合

表6中のタスクD1(レクリエーション)がこれにあたる。D1は、計画について年間3、4回程度発生することしか制約がなく、12月を最後として実施した年度もあれば、2月の年度もあり、もともと妥当な予測をすることが困難なタスクであるといえる。年末年始、夏期長期休暇以外の平日を妥当な日付としていたが、年末に予測された。

(2) 他の作業に依存して実施日が揺れる場合

B1(スポーツ大会)は、年4回のうちの1回は、学園祭と同時期に行うという暗黙の認識があった。しかし、学園祭の実施時期が2012年度に限って例年とは異なる期間に実施されていたため、年度間の差異が大きくなった。

(3) ある年に限って実施されなかった場合

A6(誕生会)は、9月と2月に毎年行われているが、2012年度の2月は、実施されなかった。また、C3(前期講義)、C5(後期講義)においては、2011年度には実施された講義期間の最終日の講義が、2012年度では、それ以前の別の日に2コマまとめて実施されるなどして見掛け上欠落している。このため、予測結果が欠落を補う形で2012年度各期の最終日に現れて問題となった。

(4) 1年前の基準日と対応点との間隔が大き(小)過ぎて、範囲外に予測される場合

たとえば、年に1回発生するタスクが、2011年度には長期休暇の5日前に発生していたとする。同じタスクが2012年度には長期休暇の1日前に発生していた場合、対応点とその次のタスクの間隔は、1年以上になるので、2013年度における予測日は、後ろにずれて、長期休暇内になってしまう。逆に基準日と対応点の間隔が1年未満になり、長期休暇の末尾に予測されるケースもあった。長期休暇に限らず、12月上旬を妥当としていたにもかかわらず11月に予測された忘年会(A9)もこのケースに数えられる。同様の現象は、A4, A5, A9, A10, A17, A18, A19, A25で各1回発生していた。

以上のことから、基準日と対応点との関係をどう扱うかが課題といえる。また、他の作業との関連性を考慮した予測のための尺度が必要といえる。

4. まとめ

作業予測のために繰返情報モデルを提案し、作業発生の規則性モデルとの関係について述べた。また、多くの作業が1年を周期として発生していることに着目した作業予測手法を提案した。さらに、周期的に発生している作業について作業予測手法を適用し、評価を行った。評価の結果、昨年の作業発生間隔と発生確率を考慮することにより、曜日や長期休暇を考慮して予測できることを示した。

なお、評価において、2つ以上のタスクの時間帯が重なることがあった。これは、それぞれの手法において3, 4件確認でき、各タスク間の関連性を考慮していないことが原因であると考えられる。タスク間の関連性を考慮した予測についての検討が望まれる。

また、本稿では、多くの周期が1年とした場合の予測結果について述べたが、大きな波の周期がつねに1年であるかどうかは今後の検討が必要である。今回の実験では、大学の研究室におけるイベントとして1年を意識した運営が行われているために、この仮定は妥当であるといえるが、大きな波の周期は、本来過去の発生履歴から求めるのがよいといえる。たとえば、自己相関関数を用いて表2のミーティングの周期を求めた場合、周期として最初に現れるのは、14日であった。これは、研究ミーティングが約14日周期で発生しているという参加者の感覚とも一致する。順位の2番目には、378日という周期が現れる。我々の主張する365日からは、2週間のずれが生じているが、この関係の意味については今後の検討が必要である。今後、自己相関関数あるいは、その他の手法によって、大きな波の周期の判別を自動化する試みが望まれる。また、提案手法を適用し、さかのぼってどの程度実際に起こったイベントを予測可能かといった評価を試みたり、まったく別の組織形態における年間スケジュールを利用したりすることで、提

案手法の有用性についてより検討したい。

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] 安部田章, 松並勝賢一: スケジュール情報の共有・再利用に着目した協調作業支援システム, 情報処理学会研究報告, Vol.95, No.67, pp.7-12 (1995).
- [2] 山根隼人, 長尾 確: AcTrec: 行動履歴を用いた個人行動支援, 情報処理学会第66回全国大会講演論文集, Vol.66, No.3, pp.115-116 (2004).
- [3] 斉藤典明, 金井 敦: 知識蓄積・継承のためのスケジュールデータ構成法, 情報処理学会研究報告, Vol.2012-GN-82, No.19, pp.1-7 (2012).
- [4] 斉藤典明, 金井 敦, 赤埴淳一: 知識蓄積に適応可能なスケジュールデータ構成方法の評価, 情報処理学会研究報告, Vol.2012-GN-83, No.10, pp.1-8 (2012).
- [5] 三原俊介, 乃村能成, 谷口秀夫: 作業発生の規則性を扱うカレンダーシステムの提案, マルチメディア通信と分散処理ワークショップ(DPSWS2010)論文集, Vol.2010, No.11, pp.215-220 (2010).
- [6] 首藤幸司, 西尾信彦: センサフュージョンを利用した個人行動の未来予測機構, 情報処理学会研究報告, Vol.2006, No.116 (2006).
- [7] 橋本幸二郎, 道木加絵, 道木慎二, 大熊 繁: HMMを用いた人間行動モデルに基づく次行動予測手法の検討, ロボティクス・メカトロニクス講演会講演概要集, pp.1-3 (2009).
- [8] 西野正彬, 中村幸博, 武藤伸洋, 阿部匡伸: あいまいな表現を含むスケジュールデータとGPSデータとの時間的共起関係を利用した行動予測手法の検討, 電子情報通信学会技術研究報告, Vol.109, No.450, pp.73-78 (2010).



吉井 英人

平成24年岡山大学工学部情報工学科卒業。平成26年同大学院電子情報システム工学専攻博士前期課程修了。現在、パナソニックインフォメーションシステムズ(株)勤務。グループウェアによる仕事効率化に興味を持つ。



北垣 千拡 (学生会員)

平成25年岡山大学工学部情報工学科卒業。同年同大学院電子情報システム工学専攻博士前期課程入学。現在、同課程においてグループウェアサービスを研究中。



乃村 能成 (正会員)

平成 5 年九州大学工学部電子工学科卒業。平成 7 年同大学院情報工学専攻修士課程修了。同年九州大学工学部助手を経て、現在、岡山大学工学部准教授、博士 (情報科学)。



谷口 秀夫 (正会員)

昭和 53 年九州大学工学部電子工学科卒業。昭和 55 年同大学院修士課程修了。同年日本電信電話公社電気通信研究所入所。NTT データ通信 (株)、九州大学を経て、現在、岡山大学教授、理事。博士 (工学)。電子情報通信学

会, ACM 各会員。