

推薦論文

計画立案の行動に着目したカレンダー情報の ダミーデータ生成手法

吉田 尚史¹ 乃村 能成^{1,a)}

受付日 2017年5月9日, 採録日 2017年11月7日

概要: カレンダー情報を扱う手法の研究やアプリケーションの開発において、個人や組織の特性を反映したカレンダー情報を多数用意したいという要求がある。しかし、カレンダー情報の収集は、プライバシーの保護や守秘義務の観点から困難である。そこで、本稿では、カレンダー情報のダミーデータを生成する手法を提案する。まず、組織や個人の特性を反映したグループカレンダーを表現するモデルとその表現方法を提案した。次に、計画立案の行動に関する時系列を再現することで、現実にも似たダミーカレンダーを生成できることを示した。その結果、複数のカレンダーから構成された、組織の特徴を反映したダミーデータを生成し、組織改変や個人の仕事の変化が既存のカレンダーをどう変化させるかシミュレートできることを示した。

キーワード: カレンダー情報, ダミーデータ

A Method for Generating Dummy Calendar Focusing on Planning Behavior

HISASHI YOSHIDA¹ YOSHINARI NOMURA^{1,a)}

Received: May 9, 2017, Accepted: November 7, 2017

Abstract: In research and development on calendar applications, there is a demand for a large number of calendar data that reflects characteristics of individuals or organizations. In this paper, we propose a method for generating dummy data of calendar information. First, we proposed a model that expresses group calendars reflecting the characteristics of organizations and individuals. Next, by reproducing the time series of planning behavior, we generated dummy calendars with reality. As a result, we showed that we can simulate how organizational changes and personal work changes will change existing calendars.

Keywords: calendar information, dummy data

1. はじめに

カレンダー情報を利用した研究やアプリケーション開発がある。たとえば、複数人の予定から新規予定の登録可能性を可視化する手法 [1] や GPS データと組み合わせた行動予測手法 [2] がある。また、独自の機能を持ったカレンダーシステム [3] を開発する場合もある。これらの手法やシステムを評価したりテストしたりするためには、テストケースとして、もっともらしいカレンダー情報が必要となる。

カレンダー情報に限らず、ソフトウェア開発や研究の目的で、もっともらしい情報、つまりダミーデータを用意することがある。ダミーデータの例として、名前、住所、生年月日、性別、および電話番号といった個人情報やランダムまたは統計情報を基に生成するサービス [4] やクレジットカード番号をランダムに生成するサービス [5] が存在する。また、学生の成績や行動のダミーデータを生成し、教育支援システムのテストケースを用意する手法 [6] が存在する。しかし、これらのデータ生成手法は、まったくのランダムか、実際の利用者からの実データそのものに基づいている

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

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

本論文の内容は 2016 年 11 月の第 168 回マルチメディア通信と分散処理研究会にて報告され、同研究会主査により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である。

ため、カレンダー情報のダミーデータ生成に同様の手法を適用することは難しい。なぜなら、単にランダムに予定を配置したカレンダー情報を作成したのでは、実際のカレンダーの利用状況に則した適切な評価を期待できないという問題があり、実データを利用する場合は、プライバシーや守秘義務に関わる情報を多数含むため、収集できる情報の量が限られるためである。仮にある組織から実データを収集できたとしても、「イベント A は、1年に1回、5月から6月に実施され、土日には絶対発生しない」という組織固有のルールを抽出することは難しい。たとえば、このルールを実データを学習することで再現しようとする、イベント A に関する開催履歴を多数の教師データとして与える必要がある、イベント A が十年単位で安定的に開催されているものでない限り難しい。また、雑多な組織にまたがって多数のデータを収集しても、単一の組織のカレンダーを構成する一貫したルールセットに基づくようなデータとはなりえない。以上のことから、統計情報を利用した手法や機械学習を用いて実データとよく似たデータを生成する手法の適用は困難であると考え。そこで、本稿では、評価やテストの意図に合わせて調整可能なダミーのグループカレンダー情報を複数生成する手法について提案する。

ダミーデータに求める要件「もっともらしさ」は主観的で曖昧である。そこで、ここでは、「もっともらしいデータ」を想定する個人や組織の文化と行動原理をよく反映しているデータであるとした。文化とは、土日や祝日に特別な意味があるか、周期的に発生する行事や会議が多いのか、突発的な行事が多いのかといったことを表し、行動原理とは、日程調整の仕組みや決定権の所在などである。そこで、「もっともらしさ」を反映してカレンダー情報のダミーデータを生成するにあたり、データ構造とアルゴリズムの観点から、2つの考慮すべき点がある。1つは、カレンダーが持つ特徴を表現する要素とは何かを明らかにし、それに対する記法を定義することである。これまでそうした記法が存在しなかったため、ある職業や組織の特徴に則したカレンダーデータが必要な場合には、「同業の多数のユーザ」から「実データ」を提供してもらう必要があった。もう1つは、実際のカレンダー上の予定は、一時にすべての予定が作成されるわけではなく、その成立過程において時系列を持っているため、それを反映した計算手法が求められるという点である。たとえば、早くから計画されカレンダーに書き込まれる予定もあれば、直前になって決まるものもある。あるいは変更されやすいか否か、年間における周期性も関わってくる。これらの要因により日々少しずつ予定が追加されたり変更されたりする。ダミーカレンダー作成の過程にこうした計画立案の過程を反映することが不可欠である。

本手法を用いることで、意図した特徴を持つカレンダー情報を表現可能であり、表現手法が得られることで、客観的で流通可能な情報としてカレンダー情報のダミーデータを扱

える。また、カレンダー情報のパラメータのみを他者から受け取ったり、それらをライブラリ化したりできる。また、本手法により、ある組織構造や職業を反映した架空のカレンダー情報を多数作成することが可能となるため、これらは、単にテストケースとして利用可能なだけでなく、ある組織の行動シミュレーションのためのデータとしても有用である。つまり、組織やプロジェクトの構造改革や事業計画の変更が、個人の忙しさにどう影響を及ぼすかといったシミュレーションを実施するためのデータとしても利用可能であると考え。

提案手法で生成するダミーデータの利用を想定しているアプリケーションは、たとえば、文献 [1], [2], [3] のようなものであり、個人あるいは人間の組織として個々に計画立案の裁量がある共有カレンダーを運用していることを前提にしたアプリケーションであれば、適用可能であると考え。一方、人間の計画立案行動に基づくものではないカレンダー情報を扱うことは難しいといえる。たとえば、航空管制のためのフライトスケジュールや鉄道ダイヤのダミーデータを本手法で生成することは難しく、アルバイトのシフトの再現や病院などの夜勤担当のカレンダーも本手法に馴染まない。

以降では、まず、想定する計画立案の方法について説明し、その中で、組織におけるカレンダーとメンバの関係について述べる。次に、組織におけるカレンダーとメンバの関係をモデル化することを考える。その後、カレンダーの評価値を最大化する計算手法を提案し、実装、機能評価について述べる。

2. 計画立案とは

2.1 想定する計画立案の方法

グループで用いられるカレンダーは複数のカレンダーから構成され、グループのメンバで共有されている。また、共有されているカレンダーのほかにも、グループのメンバごとに個人の予定が登録されるカレンダーや関わっている別グループのカレンダーが存在する。新たな予定を計画する場合には、以下の5つの手順がある。

- (1) 計画立案の時期の決定
- (2) 予定に参加すべきメンバの決定
- (3) 候補日の決定
- (4) カレンダーの評価値の集計
- (5) 予定の調整

まず、(1)について、ある予定を計画する際、予定の内容や周期によって、いつ計画する（日程調整をする）かといったことが、陽にあるいは暗黙に決まっている。たとえば、月例の会議において、次回の会議日程をその会議の場で立案することがある。あるいは、組織の記念行事のような会合であれば、数年前から計画する場合もある。

次に、(2)について、予定に参加するメンバは、グルー

プの全員とは限らない。また、この中には、予定の実施に重要な人物もそうではない人物も含まれており、重要な人物の参加は必須といった条件が存在する。

そして、(3)について、候補日は、いくつかの制約から選ばれることが多い。たとえば、1カ月に1回程度は実施するといった「周期」や、「平日のみ」という条件が考慮される。あるいは、あらかじめ日付が決まっている場合もある。このような条件を満たす1つ以上の日付を候補日とする。

さらに、(4)で、参加者の意向を確認し、各候補日に予定を設定した場合のカレンダーの評価値を集計する。カレンダーの評価値は、カレンダーの個々の予定の満足度の総和によって決まり、満足度とは、参加者が多いかどうかや、重要な人物が参加できるかどうかによって求められる。ここで、候補日によっては、ある参加者の別の予定の日程を変更することが考えられる。この場合、その変更によって失われる別の予定の満足度を加味しなければならない。変更が実施日の直前であるのか、どの程度重要な予定であるのかといったことが考慮されるべきである。また、予定の変更は、推移的に他の参加者にも及ぶため、これらの得失も加味する必要がある。

最後に、(5)で、(4)の集計の結果、最も満足度の高い候補日に予定を作成する。

このような計画立案の行動においては、各メンバが各々関わりのあるカレンダーを複数持ち、それぞれに対してどの程度重要な役割を果たしているかといったことを考慮すべきである。したがって、グループのメンバとカレンダーが相互に接続された重み付きのグラフとしてメンバとカレンダー間の関係を表現することがふさわしい。

また、各カレンダーの個々の予定について満足度を定義し、それらの総和を最大化するような行動をシミュレートすることで、ダミーカレンダーを生成可能である。以降では、まず、グループで利用されるカレンダーを表現するモデルと、カレンダーの満足度についての尺度を定義する。

2.2 グループで利用されるカレンダーを表現するモデル

2.1節での議論をふまえて、グループで利用されるカレンダーを表現するモデルを図1のように設定する。図中の四角はカレンダーを表し、丸がカレンダーの利用者を表す。また、利用者からカレンダーへの矢印はその利用者がカレンダーに与える影響力を表している。影響力が大きい利用者ほど、そのカレンダーの予定に重要な人物であるとする。図1の場合では、5つのカレンダーと4人の利用者の関係を表現している。CalendarAについて、Bobは他の利用者よりも大きな影響力を持っているため、そのカレンダーの予定に重要な人物であるといえる。また、カレンダー利用者の影響力の合計が大きなカレンダーほど、重要なカレンダーであるといえる。

図1において、CalendarAに新たな予定を計画する場

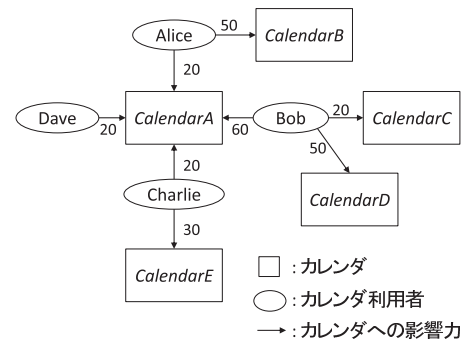


図1 想定するカレンダーのモデルの例
Fig. 1 Example of our calendar model.

合、以下の手順で行う。予定の参加者は Alice 以下 4 人の利用者で、影響力の違いから Bob が重要な人物である。また、計画する予定の候補日は 2 つとし、どちらの日付に実施すると良いか判断するとする。

(手順1) 参加者が参加可能か確認

各候補日に予定を実施する場合、カレンダーの利用者が参加できるか確認する。このとき、各利用者は CalendarA のほかに自身に関わるカレンダーの予定を確認し、参加できるか判断する。たとえば、Bob の場合、CalendarA の予定だけでなく、CalendarC, D の予定も確認する。

(手順2) 予定の実施日を最適化

各候補日に予定を実施する場合のカレンダーについて、それぞれのカレンダーの満足度として、カレンダーの評価値を算出する。評価値の算出法は、次節で説明する。算出された評価値が最も高いカレンダーに基づき、予定を作成する。

2.3 カレンダーの評価値の算出方法

2.3.1 カレンダーの評価値を算出するための要因と計算式

ある予定を計画する際、他の予定と比較し、交渉を行いながら変更する研究が存在する [7], [8]。これらのような手法を用いることで、参加者が最も多くなるカレンダーを生成できる。

しかし、参加者が最も多くなるカレンダーが最も良いカレンダーとはいえない。これは、カレンダーは時系列に沿って予定が登録されたり、変更されたりするため、前の状態のカレンダーからの変化が大きい場合や実施日の直前に変更される場合は良いカレンダーとはいえない。なぜなら、すでに生成されている予定を変更することはコストがかかるためである。同様の考え方で、最適なカレンダーとなるような調整を分散重み付き制約充足問題として考える手法 [9] が存在するが、本提案では、複数のカレンダーを扱うため、カレンダーによって影響を受ける人や影響の大きさが異なる。そこで、予定ごとに変更するコストを算出し、カレンダーの評価値に加味する必要がある。

カレンダーの良し悪しと予定の変更コストを考慮し、以下の3つ要因を用いてカレンダーの評価値を判断する。

(要因1) カレンダー全体の各予定に対する参加可能なメンバーの影響力の合計

カレンダー全体の各予定における、参加可能なメンバーの影響力の合計が大きいほど、予定で重要な人物が参加可能であり、参加人数も多いといえるため、良いカレンダーであると判断できる。

加えて、ある予定を立案した際に他の予定の日程変更を要する場合は、以下の要因を考慮しなければならない。

(要因2) 変更を必要とする予定の参加者の影響力の合計
参加者の影響力の合計が高い予定は、参加者が多い予定や重要な予定である。このような予定を変更する場合、参加者の都合の調整や変更の周知が困難であり、変更コストが高くなると考える。

(要因3) 変更を必要とする予定の実施日までの日数
実施日の前日に変更を行うといったように、急遽予定を変更することは困難であると考えられるため、変更する日付と実施日の日付が近いほど変更コストが高くなると考える。

ここで、各要因を考慮したカレンダーの評価値 U を定義する。まず、(要因1)を反映したカレンダーの良し悪しを $F(Cals_n)$ とする。 $Cals_n$ は、グループで利用される各カレンダーのある時点 n での集合を表す。また、(要因2)と(要因3)を反映した予定の変更コストを $C(e, today)$ とする。 C は、予定の変更が発生するたびに発生し、 F は、予定の数の増加にともない値が増加するため、 C が発生するたびに F を除することで U とする。

以下の関数 $U(Cals_n, today)$ で、予定の変更コストを考慮したカレンダーの評価値を算出する。 $Cals_{n-1}$ は、 $Cals_n$ の前の状態を表す。また、 $Cals_{n-1}$ に登録されていた予定のうち、 $Cals_n$ で変更された予定の集合を $E(Cals_{n-1}, Cals_n)$ と表す。 $today$ は変更を行う日付を表す。

$$U(Cals_n, today) = \begin{cases} \frac{F(Cals_n)}{\prod_{e \in E} C(e, today)} & (E \neq \phi) \\ F(Cals_n) & (E = \phi) \end{cases} \quad (1)$$

ただし、 E は $E(Cals_{n-1}, Cals_n)$ を意味する。 E が空集合の場合、右辺の分母は1と考え、 $U = F$ とする。次項以降で F と C を定義する。

2.3.2 カレンダーの良し悪しを算出する関数

カレンダーの良し悪しは、(要因1)の各予定に対する参加可能なメンバーの影響力の合計で表現できる。各参加者のカレンダーに与える影響力を考慮することによって、重要な人物が参加できる方が良いカレンダーであることを表現できる。以下の関数 $F(Cals_n)$ でカレンダー $Cals_n$ の良し悪しを表現する値を算出する。

$$F(Cals_n) = \sum_{c \in Cals_n} \sum_{e \in c} \sum_{p \in e} I(p, e) \quad (2)$$

ここで、 c は $Cals_n$ に含まれるカレンダーで、 e は c に含まれる予定であり、 p は e に参加できる人物である。また、関数 $I(p, e)$ は p の e が登録されるカレンダーに対する影響力を表す。上記の関数によって、 $Cals_n$ に含まれるすべての予定の参加者の影響力を総和する。

2.3.3 予定の変更コストを算出する関数

予定の変更コストは予定を変更するたびに発生するものとして考える。このため、1つの予定の変更にとともに、他の予定にも変更が発生する場合、それぞれの予定の変更に対して変更コストを算出する。

変更コストは(要因2)と(要因3)の2つの要因を考慮して算出する。(要因2)の参加者の影響力の合計が大きいほど変更コストは大きくなると考えられる。また、(要因3)の実施日までの日数が長い場合の変更は大きなコストではないが、実施日直前には、急激に大きくなると考えられる。このため、実施日までの日数の対数で影響を受ける人数を除いたものを変更コストとする。つまり、ある予定 e 実施日時を変更するコストを以下の $C(e, today)$ で定義する。ここで、変更を行う日付を $today$ と表す。また、 e の実施日を $D(e)$ と表す。

$$C(e, today) = \frac{\sum_{p \in e} I(p, e)}{\log((D(e) - today) + 2)} \quad (3)$$

実施日までの日数は0以上の整数であり、対数をとった後に除算を行うため、右辺の分母が0にならないように2を加えている。また、 $C(e, today)$ の計算結果が1未満となる場合、1として考える。関数 $I(p, e)$ は式(2)の $I(p, e)$ と同様に p の e が登録されるカレンダーに対する影響力を表す。

3. ダミーデータ生成手法

3.1 候補日の選定方式

2.1節で述べたように、提案手法で想定する計画立案では、立案の時期を決定し、その時期になると候補日を設定して調整を繰り返すことでカレンダーを作成していく。つまり、候補日をどの程度の周期でどういう方針に基づいて作成するかがダミーカレンダーの作成にとって重要な要因となる。ここでは、その方式について説明する。

カレンダーに登録される予定は、繰り返し発生する予定と単発で発生する予定の2種類に分けられる。以降では、繰り返し発生する予定の集合をリカーレンスと呼ぶ。たとえば、約1カ月に1回の周期で会議を行う場合、「第1回会議」、「第2回会議」、「第3回会議」という一連の予定の集合を「会議」というリカーレンスとする。多くの予定は、何らかの周期(再び同様の予定がいつかは発生する)がある[10]ため、各カレンダーは複数のリカーレンス系列と単発の予定によって構成されるものとする。

この考えに基づき、候補日作成の方針を以下に述べる。
(方針1) 特徴の異なる複数リカーレンス系列を用意し、立案を開始するタイミングや調整過程を再現

会議や講義、記念日といった、それぞれ立案開始のタイミングや候補日に関する制約条件（たとえば、平日のみに発生など）、および発生周期の異なる複数のリカーレンス系列を用意し、それぞれについて、立案を開始するタイミングで制約条件を満たす候補日を複数設定し、これまでに述べたカレンダーの評価値を最大化する調整を繰り返すことで、カレンダーを逐次的に生成していく。リカーレンスに与える具体的なパラメータについては、4章で述べる。

(方針2) 単発で発生する予定は事前にランダムに生成
単発で発生する予定には、決まった制約や特徴が存在しないため、第2、第3の候補日を設定することができない。つまり調整の対象とできない。そこで、各カレンダーごとに1カ月あたりの単発で発生する予定の件数を設定し、予定をランダムに生成する。また、優先度の高い予定として調整の対象とはしない。

3.2 ダミーデータ生成手順

リカーレンスの多くは1年を周期とした相似形を持つ[11]ため、前年の予定を確認しながら予定を計画することが多い。この過程を再現するために、ダミーデータとして生成したい範囲の1年前から予定を生成する。カレンダー情報のダミーデータの生成手順を以下に示す。

(手順1) 単発で発生する予定を生成

カレンダーごとに単発で発生する予定を生成する。各カレンダーの1カ月あたりの単発で発生する予定の件数に従って、単発で発生する予定をランダムに生成する。

(手順2) 各リカーレンスの1回目の予定を生成

各リカーレンスの特徴を基に1回目の予定を生成する。ここで生成する予定は、生成範囲の開始日の1年前から生成範囲の開始日までの間に生成される。このため、ダミーデータとして出力されるカレンダーには含まれないが、今後の手順において、基準として利用するために生成する。

(手順3) 各リカーレンスの次回生成日を設定

各リカーレンスごとに立案を開始するタイミングが存在する。このタイミングは大きく分けると2種類存在する。1つ目は、タイミングが特定の日に固定されている場合である。たとえば、「忘年会」というリカーレンスは、毎年11月1日に立案するとあらかじめ決めておく。この場合、リカーレンスごとに決まっている日付を設定する。2つ目は、予定の実施時が次の予定の立案開始日の場合である。たとえば、「定例会議」というリカーレンスが該当する。この場合、リカーレンスに属する予定を実施する日付を設定する。ここでは、(手順2)で生成した1回目の予定の実施日を設定する。ここで設定した日付順に予定を生成することで、予定がいつ決まるかという時系列を再現する。

(手順4) 生成日になったリカーレンスの予定を生成

生成範囲の開始日から1年前の日付から1日ずつ順番に確認し、次回生成日であるリカーレンスがあれば予定を生成する。このとき、1年分といった複数の予定を生成する場合もある。たとえば、「大学の講義」という予定では、予定を生成する日付になると1年分の予定を生成する。生成する予定の候補日は、リカーレンスの特徴を考慮して、特徴が反映された日付を順に候補日とする。候補日には順番が設定されており、第1候補日から順に参加者の確認を行う。第1候補日の参加者の都合が良く、予定を実施できる場合、第1候補日に予定を生成する。参加者の都合が悪い場合、すでに決まっている他の予定を別の日付に変更するか第2候補日以降に生成するか判断する。このとき、2.3節の式でそれぞれの変更を行った場合のカレンダーの評価値を算出し、最も良いカレンダーになるように調整する。そして、(手順3)と同様に、各リカーレンスごとの立案を開始するタイミングに従って、次回生成日を更新する。

4. 設定するパラメータ

4.1 概要

候補日生成の基準となるリカーレンスの特徴を記述するためと2.2節で述べたカレンダーのモデルを表現するために、以下の3つについて、それぞれに必要なパラメータ表現を定義する。

(1) リカーレンス

(2) カレンダー

(3) カレンダー利用者

これらのパラメータ表現を定義することで、カレンダーが持つ特徴を表現する記法として利用できる。

4.2 リカーレンスに対して設定するパラメータ

4.2.1 リカーレンスの特徴の調査

リカーレンスの特徴を表す要素とは何かを考えるために、実在のカレンダーを調査する。リカーレンスの特徴を表す要素は、人が予定を立てる際に意識する要因を一般化することで抽出できるものとする。たとえば、「約2週間に1回実施しよう」や「土日祝日、盆正月には行わない」、「ある程度時期が近づいてから実施日を決めよう」といった要因である。これらの要因は、「特定の周期で発生する」や「平日か休日かに影響される」、「どの程度前に予定が確定するか」といった特徴を表現する要素として抽出できる。こうした要素を実在のリカーレンスの調査から抽出する。

4.2.2 調査対象と方法

調査対象は我々の所属する大学の研究室で利用している共有カレンダーに含まれる38個のリカーレンスである。表1に38個のリカーレンスを示す。表1の38個のリカーレン

表 1 調査対象のリカーレンス一覧
Table 1 List of recurrences to survey.

通番	種別	リカーレンス名
1	研究ミーティング (約 2 週間に 1 回)	ミーティング 1
2		ミーティング 2
3		ミーティング 3
4		ミーティング 4
5	進捗報告会 (約 1 カ月に 1 回)	進捗報告会 1
6		進捗報告会 2
7		進捗報告会 3
8	年中行事 (半期に 1 回)	誕生日会
9	年中行事 (約 1 年に 1 回)	研修会
10		忘年会 1
11		忘年会 2
12		送別会 1
13		送別会 2
14		新人歓迎会
15		暑気払い
16		卒業アルバム写真撮影
17		M2 中間発表
18		B4 中間発表
19		防災訓練
20		M1 論文紹介
21		M2 論文紹介
22		卒論発表
23		修論発表
24		オープンキャンパス
25		研究室大掃除
26		大学の編入試験
27		大学の推薦入試
28	大学の卒業式	
29	レクリエーション (1 年に約 3 回)	レクリエーション
30		スポーツ大会
31	講義 (約 1 週間に 1 回 (半期))	講義 1
32		講義 2
33		講義 3
34		講義 4
35		講義 5
36	恒例行事 (決まった日付)	大学の前期入試
37		大学の後期入試
38		大学の入学式

スの中には、さまざまな時間間隔のリカーレンスを含んでおり、一般的な組織で行われる、会議や会合に相当するリカーレンスを含んでいるため、一般的なりカーレンスの特徴の調査に用いるのは有用であると考えられる。各リカーレンスに属する予定を立てる際に意識する固有の要因を列挙する。たとえば、表 1 の通番 1 の「ミーティング 1」というリカーレンスは「2 時間程度」、「約 2 週間間隔」、「平日のみ」、「長期休暇には行わない」、および「特別研究報告書・修士論文の締切前には発生間隔が短くなる」といった固有の要因が存在する。その後、列挙した各リカーレンス固有の要因を一般化する。たとえば、「2 週間間隔」や「1 カ月

間隔」といった、各リカーレンス固有の要因は「ある発生間隔を持つ」と一般化する。一般化したそれぞれの要因がリカーレンスの特徴を表す要素と考える。また、複数の要素を組み合わせることでもリカーレンスの特徴を表現できると考える。

4.2.3 調査結果

カレンダー情報の特徴を表す要素を表 2 に示す。これらの要素は調査対象の各リカーレンスが持つ固有の要因を一般化した要因である。表 2 の通番 10 から 14 は予定を立てる際の行動に関わる要因であり、カレンダーに登録されている予定を後から見ただけでは、判断できない。このような要因を考慮することで、実際の予定を立てる過程を再現できる。

4.2.4 リカーレンスごとに設定するパラメータ

リカーレンスごとに設定するパラメータは表 2 の 14 個の要素の組合せである。ここで、表 2 の通番 14 の TIMING は、3.2 節の (手順 3) で述べたように、2 種類に分けられる。生成する日付が固定されている場合、あらかじめ決まっている日付を設定し、予定の実施時が次の予定の立案開始日である場合、「successively」と設定する。表 2 のパラメータを組み合わせることによって、リカーレンスの特徴を表現する。たとえば、「月例会議」というリカーレンスの場合、1 回あたり 2 時間行う、発生間隔は 30 日、休みの日には行わない、およびお盆や年末年始という休暇期間があるといった特徴を反映させるため、表 2 の通番 1, 2, 3, 4, 6, 10, および 14 のパラメータを図 2 のように設定する。図 2 の上から順に表 2 の通番 1, 2, 3, 4, 6, 10, および 14 のパラメータが設定されており、通番 10 の VACATION_TERM には、お盆の時期と年末年始の時期の 2 つの期間を設定している。他のパラメータについては、それぞれに相当する特徴が存在しないため、設定していない。このような設定によって、特定のリカーレンスを表現する。また、複数のリカーレンスを組み合わせ、カレンダーに登録される予定を表現する。

4.3 カレンダーに対して設定するパラメータ

各カレンダーに対して設定するパラメータとして、3 つのパラメータを考える。表 3 に 3 つのパラメータを示し、以下で各パラメータについて説明する。

(1) カレンダー名

カレンダーの名前を設定する。

(2) 1 カ月あたりの単発で発生する予定の割合

単発で発生する予定が 1 カ月あたりに存在している数を表し、この値に従って、ランダムな予定を生成する。たとえば、月に 2 回程度不定期に出張が発生する場合にはこの値を「2」に設定する。

(3) 予定を実施するための閾値

予定を実施する際、重要な人物の参加は必須といった

表 2 リカーレンスの特徴を表す要素
Table 2 Elements that characterize recurrences.

通番	要素名	パラメータ名	説明
1	予定名	SUMMARY	予定の名前
2	カレンダー名	CALENDAR	登録されているカレンダーの名前
3	実施時間	DURATION	何時間実施するか
4	発生間隔	INTERVAL	周期が存在する
5	曜日	WDAY	特定の曜日に影響される
6	休日	HOLIDAY	土日祝日といった休日に影響される
7	月	MONTH	何月かに影響される
8	週	MONTHWEEK	月の第何週目かに影響される
9	日付	DATE	日付に影響される
10	非発生期間	VACATION_TERM	予定が発生しない期間が存在する
11	順序関係	ORDER	他の予定との間に順序関係が存在する
12	同時発生	SIMULTANEOUS	他の予定と関連して同日に発生する
13	締切	DEADLINE	何らかの締切が存在し、影響される
14	生成タイミング	TIMING	予定がいつ生成されるか

SUMMARY: 月例会議
CALENDAR: 会議
DURATION: 120
INTERVAL: 30
HOLIDAY: false
VACATION_TERM:
- 8/10-8/14
- 12/25-1/3
TIMING: successively

図 2 リカーレンスに対して設定するパラメータの例
Fig. 2 Example of parameters of a recurrence.

表 3 カレンダーごとに設定するパラメータ
Table 3 Parameters of individual calendars.

通番	要素名	パラメータ名
1	カレンダー名	NAME
2	1カ月あたりの単発で発生する予定の割合	RANDOM_EVENTS_PER_MONTH
3	予定を実施するための閾値	PARTICIPANT_THRESHOLD

条件を満たせば、数人の欠席者を許容する場合がある。そこで、予定に参加できる利用者の影響力の合計がどの程度必要か設定する。予定を生成する際、予定に参加できる利用者を確認し、その影響力の合計がこの閾値以上であれば、その日付に予定を生成できるとする。各利用者のカレンダーへの影響力と閾値を設定することで、一定割合の人数の参加できるという条件やある人物の参加は必須という条件を表現できる。たとえば、10人のカレンダー利用者が存在し、1人の重要な人物には影響力として30を設定し、残りの9人にはそれぞれ10を設定する。このとき、閾値として100を設定することで、重要な人物の参加は必須で、残りの9人のうち2人は欠席しても予定を実施できることを表現できる。

カレンダーごとに設定するパラメータの例を図3に示す。図3は図1のモデルの例を基に、CalendarAとCalendarB

NAME: CalendarA
RANDOM_EVENTS_PER_MONTH: 2
PARTICIPANT_THRESHOLD: 100
NAME: CalendarB
RANDOM_EVENTS_PER_MONTH: 1
PARTICIPANT_THRESHOLD: 50

図 3 カレンダーに対して設定するパラメータの例
Fig. 3 Example of parameter of calendars.

Alice:
CalendarA: 20
CalendarB: 50
Bob:
CalendarA: 60
CalendarC: 20
CalendarD: 50
Charlie:
CalendarA: 20
CalendarE: 30
Dave:
CalendarA: 20

図 4 カレンダー利用者に対して設定するパラメータの例
Fig. 4 Example of parameters of users.

の設定を示している。同様の記述を増やすことで、カレンダーの数を増やすことができる。上記3つのパラメータを設定し、カレンダーに含まれるリカーレンスと組み合わせることで、1つのカレンダーを表現する。また、複数のカレンダーを設定することで、グループで利用されるカレンダーを表現する。

4.4 カレンダー利用者に対して設定するパラメータ

カレンダー利用者ごとに各カレンダーに対する影響力を表す数値を設定する。この値が大きいほど、カレンダーの予定に重要な人物であることを表す。パラメータの設定例は図4のようになる。CalendarAは4人全員が関わっており、ここではBobがCalendarAの予定に重要な人物であることを表現している。また、参加者の影響力の合計が高いカレ

ンダほど、変更にかかるコストが大きいことを表す。

5. 提案手法の機能性評価

5.1 参加者を考慮した予定調整機能

5.1.1 評価対象のダミーデータ

提案手法を実装し、ダミーデータの生成を行った。生成したダミーデータを用いて、参加者を考慮した予定調整機能が有効に働くかを評価する。ダミーデータを構成する各カレンダーは個別に利用者が存在する。提案手法では、カレンダーと利用者の関係を重み付きのグラフで表現しており、予定を生成したり、変更したりする際にはカレンダーをまたがった他の予定との日程調整が働く。こうした予定の参加者を考慮した調整が行われていることを示す。

ここで生成するダミーデータは、我々が所属する大学の研究室で利用されるカレンダーを再現したものである。生成したダミーデータは6個のカレンダーで構成され、研究室全体のカレンダーが1個、研究室内のグループのカレンダーが2個、複数の研究室が関係するカレンダーが2個、研究室の教員のカレンダーが1個存在している。これらのカレンダーには表1の38個のリカーレンスがリカーレンスごとに対応するカレンダーに分散されて入っている。また、研究室外の関係するカレンダーとして、研究室の学生が受講する講義のカレンダーを追加する。このカレンダーには講義のリカーレンスが4個含まれている。ここで、4個の講義のリカーレンスは表1の通番31から35の講義のリカーレンスとは異なるものである。表1の通番31から35の講義は研究室の教員が担当する講義である。一方、学生が受講する講義のカレンダーに含まれる4個の講義は研究室外の教員が担当し、研究室の学生が受講する講義である。

上記の7個のカレンダーで構成されるダミーデータを生成する。単発の予定は月に4件の割合で7個のカレンダーに分散して発生するとしている。ダミーデータの生成範囲は2016年4月から2017年3月までとし、予定を生成できる時間帯は10時から17時の7時間とする。

5.2 評価結果

5.1.1項のダミーデータを生成し、各予定を確認した。生成したダミーデータのうち、2016年5月のカレンダーを図5に示す。図5は土日を表示していない。図5には、2週間に1回行われるミーティングが4種類、1カ月に1回行われる進捗報告会が3種類、予定や毎週行われる講義が5種類生成されている。ほかにも、ランダムに発生する予定として、単発の予定が生成されている。また、図5中の講義のうち、「講義1」、「講義2」、および「講義3」は研究室の教員が担当する講義で、「講義A」と「講義B」は他の教員が担当するが、研究室の学生が受講する講義である。

ここで、「講義1」と「講義B」は毎週金曜日に発生している。しかし、27日には研究室の教員が利用するカレンダー



図5 生成したダミーデータ (5月)

Fig. 5 Dummy data generated by our method (May).

で単発の予定が生成されている。これは、研究室の教員が出張することを表しており、出張と講義を同日には行えないため、研究室の教員が担当する「講義1」は前日の26日に変更されている。一方、「講義B」は研究室の教員が関わっていないため、学生の予定としてそのままである。予定の参加者を考慮することにより、時間が重複してしまう予定であっても、参加者が異なる場合は生成できていることが分かる。

5.3 グループの立案行動が変化した場合のシミュレーション

カレンダーとリカーレンスをモデルで表現しているため、パラメータを変化させることで、グループの立案行動が大きく変化する場合に、変化後のカレンダーをシミュレートできる。たとえば、5.1.1項のダミーデータについて、大学でクォータ制が導入され、講義が四半期に集中するといった変化が起こればと考えられる。また、教員の出張が増加する場合もある。このようなグループの立案行動の変化をシミュレーションすることによって、変化後の計画が立てやすくなるを考える。たとえば、予定が多くなりすぎるため、出張を少し減らそうとしたり、ある学期は予定が比較的少ないため、グループ全体で勉強会を開催したりといった計画が立てられる。

図5は大学のクォータ制が導入される前のカレンダーを再現したダミーデータである。このダミーデータのパラメータを変更することで、クォータ制が導入された後のカレンダーを再現できる。具体的には、教員が担当したり、学生が受講したりする「講義」という予定の一部を週2回発生する



図 6 クォータ制導入後のダミーデータ (5月)

Fig. 6 The dummy data after introduction of quarter system (May).

ように変更し、2カ月間に集中するようにする。1学期と2学期を通して開講される講義も存在するため、一部の講義は変更を行っていない。パラメータに変更を加え、クォータ制が導入された後のカレンダーを再現したダミーデータを図6に示す。図6の予定を確認すると講義の予定が増加しており、研究室で行う他の予定の候補日が少なくなっていることが想像できる。一方で、木曜日は比較的予定が少ないことが分かり、突発的な予定が発生した場合でも、予定を変更する余裕は残っていると見える。この状態から、さらに教員の出張が増加した場合を想定し、ダミーデータを生成しようとする、予定が多すぎるため、妥当な日付に「ミーティング1」の予定が生成できない様子を確認した。また、ランダムに発生する予定を生成する際に用いる乱数のシード値を変えると、8回に1回は出張が増加した場合でも「ミーティング1」の予定を設定できた。しかし、これはいくつかの偶然が重なった場合であるため、現実的な計画ではないと判断できる。

6. おわりに

本稿では、想定したグループのカレンダーを再現するダミーデータ生成手法を提案した。まず、ダミーデータに求める要件と考慮すべき点をあげた。次に、グループで利用するカレンダーを表現するモデルと予定の満足度の尺度を設定した。モデルは、カレンダーとカレンダー利用者を接続した重み付きグラフで表現される。その後、ダミーデータの生成手順とパラメータ表現を定義した。最後に、予定の参加者を考慮した予定調整機能を評価し、カレンダーと利用者の関係と行動原理が反映されていることを示した。また、グ

ループの立案行動の変化をシミュレートし、グループの予定が増加した場合を想定したカレンダーを生成でき、さらに予定が増加した場合、予定が妥当な日付に生成できない様子を示した。

残された課題として、本手法の定量的な評価がある。生成したダミーデータと実データとの類似度や生成できるダミーデータの多様性を評価する必要がある。また、パラメータを個別に設定するのではなく、職業、組織の規模、および1カ月や1日あたりの予定数といった情報から、パラメータの組合せを生成する必要がある。これを実現し、ライブラリ化すると、想定するグループカレンダーを表現しやすくなると考える。ほかにも、各カレンダー利用者の影響力を設定する方法を明確化する必要がある。

参考文献

- [1] Faulring, A. and Myers, B.A.: Availability bars for calendar scheduling, *Proc. Conference on Human Factors in Computing Systems Extended Abstracts (CHI)*, pp.760-765, ACM (2006).
- [2] 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ラデータと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).
- [3] 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).
- [4] kazina: なんちゃって個人情報, kazina.com (オンライン), 入手先 (<http://kazina.com/dummy>) (参照 2016-10-12).
- [5] iGoPayGo Soft Gen: Credit Card Generator, iGoPayGo Soft Gen (online), available from (<https://names.igopaygo.com/credit-card>) (accessed 2016-10-12).
- [6] Marciel, M., Michelinakis, F., Fanou, R. and Muñoz-Merino, P.J.: Enhancements to Google Course Builder: Assessments Visualisation, YouTube Events Collector and Dummy Data Generator, *XV Simposio Internacional de Tecnologías de la Información y las Comunicaciones en la Educación (SINTICE 2013)*, pp.6-13 (2013).
- [7] Jennings, N.R. and Jackson, A.: Agent-based meeting scheduling: A design and implementation, *Electronics Letters*, Vol.31, No.5, pp.350-352 (1995).
- [8] 本村真一, 影本憲五, 川村尚生, 菅原一孔ほか: モバイルエージェントに基づく会議日程調整システム, 情報処理学会論文誌, Vol.46, No.12, pp.3123-3126 (2005).
- [9] 磯村厚誌, 大園忠親, 新谷虎松: マルチエージェントカレンダーシステムによるイベント日程スケジューリングについて, 人工知能学会全国大会論文集 2003 年度人工知能学会全国大会 (第17回) 論文集, 社団法人人工知能学会, pp.182-182 (2003).
- [10] 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実現, 情報処理学会研究報告, Vol.2011-DPS-149, No.10, pp.1-6 (2011).
- [11] 吉井英人, 北垣千祐, 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と評価, 情報処理学会論文誌, Vol.56, No.2, pp.543-553 (2015).

推薦文

本論文では、複数および単発で発生する予定を組み合わせ、かつ、計画立案の行動に着目してカレンダー情報のダミーデータを生成する手法を提案しています。システム評価用のデータ生成法としてのほか、機械学習用データ生成法としての利用も期待でき、有用性が高いと思われます。以上の理由により、本論文を推薦いたします。

(マルチメディア通信と分散処理研究会主査 重野 寛)



吉田 尚史 (学生会員)

平成 28 年岡山大学工学部情報系学科卒業。同年同大学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)。本会シニア会員。