

# 複数人の経験共有に基づく情報検索システムの提案

板野 愛<sup>†</sup> 小林 一郎<sup>‡</sup>

<sup>†</sup> お茶の水女子大学大学院 人間文化創成科学研究科 理学専攻 小林研究室

〒112-8610 東京都文京区大塚 2-1-1

<sup>‡</sup> お茶の水女子大学大学院 人間文化創成科学研究科 理学専攻 〒112-8610 東京都文京区大塚 2-1-1

E-mail: <sup>†</sup> <sup>‡</sup> {ai\_i, koba}@koba.is.ocha.ac.jp

**あらまし** 近年、パーソナルコンピュータ（PC）の高性能化やインターネットの普及により、PC 内において管理するデータ量が膨大になっている。これに伴い、膨大な量のデータを効率よく管理するためのツールの必要性が高まっている。Google デスクトップのようなインターネットの情報を検索するのと同様な検索システムの利用が可能になったが、検索は依然としてキーワードに依存している。このことから、人の経験に基づき記憶をきっかけにした検索システムの開発などがなされてきている。本研究では、そのような研究をさらに拡張し、将来、家族などの小さな共同体において複数人によって情報が共有され管理される場合に、彼らの経験（イベント）を共有することにより、共有されたイベントを介在して情報を効率的に検索する手法を提案し、個人の情報検索を支援するシステムの開発を目的とする。

**キーワード** 経験共有, 記憶, 情報検索

## An Information Retrieval System Based on Sharing Experiences of Plural People

Ai ITANO<sup>†</sup> and Ichiro KOBAYASHI<sup>‡</sup>

<sup>†</sup> Graduate School of Humanities and Sciences, Advanced Sciences, Kobayashi Laboratory Ochanomizu University  
Otsuka 2-1-1, Bunkyo-ku, Tokyo, 112-8610 Japan

<sup>‡</sup> Graduate School of Humanities and Sciences, Advanced Sciences, Ochanomizu University  
Otsuka 2-1-1, Bunkyo-ku, Tokyo, 112-8610 Japan

E-mail: <sup>†</sup> <sup>‡</sup> {ai\_i, koba}@koba.is.ocha.ac.jp

**Abstract** Recently, as high performance personal computers (PCs) and the Internet have been widely common, the amount of data managed in a PC has also been getting increased. With this, the necessity for the tools to efficiently deal with huge amount of data has increased. We can nowadays use the same kind of information retrieval systems in a desktop environment such as Google Desktop, etc., as the widely used retrieval systems on the Internet. However, the information retrieval system still relies on using keywords when retrieving information. From this fact, many human experienced-based information retrieval systems using their memories as clues for retrieving information have so far studied. In this study, we aim to further develop such an information retrieval method and propose a method so as to efficiently retrieve information by sharing plural peoples' experiences, assuming that it would be common that information will be shared and managed by plural people in a small community like a family in the future, and then develop an information retrieval system with the proposed method.

**Keyword** Experience Sharing, Memory, Information Retrieval for a Small Community

### 1. はじめに

近年、パーソナルコンピュータ（PC）の高性能化やインターネットの普及により、PC 内において管理するデータ量が膨大になっている。これに伴い、膨大な量のデータを効率よく管理する必要性が高まってきた。

現在、検索を中心とするデータ管理の方法としては、

キーワードを利用する方法、ファイル名から検索する方法や拡張子から検索する方法が存在するが、適切なキーワードの選定が必要、ファイルの内容やファイルの場所を把握していなければならない、特定の種類のファイルしか検索対象にならないなど検索を行うためにさまざまな制約が存在し、これらの検索方法には限

界がある．このため，これらのような文字列の一致を基本的とした情報検索システムとは異なり，人の経験に基づき記憶をきっかけにして，欲する情報を検索する検索システムの開発などがなされてきている．

また，PC 内のデータ管理の方法として，現在は一般的に個人のデータを対象とした管理が主流となっている．しかし，将来は今よりも家庭の日常生活の中において家族が電子データを共同で管理する機会が増えると考えられる．

このことから，本研究では，カレンダーに登録された個々の予定を「イベント」として捉え，家族のような小さなコミュニティにおいて複数人によって情報が共有され管理される場合に，お互いの経験（個々の予定であるイベントに相当）を共有することにより共有されたイベントを介在して情報を効率的に検索する手法を提案し，個人の情報検索の精度向上を支援するシステムの開発を目的とする．

## 2. 関連研究

時系列に沿ってデータを管理し検索するシステムとして，暦本[1]の“Time Machine Computing”や松原ら[4]によるデスクトップ検索システムなどが挙げられる．また，イベントや記憶（印象に残ったことがら）に沿って検索するシステムとして，Dumais ら[2]による SIS（Stuff I’ve Seen）などがある．SIS は，ファイルを保存すると自動的に情報が検索に使うメタデータとなり，関連するどのような単語からも情報を検索できるシステムである．本研究では，これらの時系列に沿ったデータ管理や記憶に基づいた情報検索手法などを踏まえ，従来，検索システムの利用として個人を対象としていたものを，小さなコミュニティを対象とし，複数人の経験共有に基づく新しい検索手法を提案する．

コミュニティ内における情報（経験）共有の方法として，小さなコミュニティの中であいまいな知識を共有する方法[5, 6]や組織内のインフォーマル情報を共有する手法[9]などが提案されている．本研究では，情報を共有する媒体を構成員がスケジュールに登録するイベントとし，イベントを機軸にした情報（経験）の共有を実現する．通常，管理対象となる PC 内のデータは，イベントの発生と共に生成される．そのため，イベントがカレンダー上に登録されている場合は，そのイベントを通じて特定のファイルを発見できる場合が多い．このことを利用して，時間的一致により情報を検索するシステムが多く開発されている．しかし，この場合，PC 内に保存されているデータを関連付ける対象となる自分の経験（イベント）が多数存在しないと検索の再現率を上げることはできないことは容易に想像される．そのため，ニュースなど社会で起こった

イベントに管理対象となるファイルを関連づける情報検索システムも開発されている[2,3]．これらの研究においてファイルと結びつけるイベント数の不足に対する対処は可能となったが，関心のないイベントを介在してファイルを検索することは容易でない，などの問題点が存在する．

一方，家族など小さな共同体においては，お互いの予定を把握し日常生活を過ごしている．つまり，個人の予定だけではなく共同体構成員の予定も自分の生活の一部として深く認識していると考えられる．そこで，本研究においては，家族（または小さな共同体）において使用される情報検索システムの開発を主眼に置き，個人の予定だけではなく家族の予定（イベント）を介在することにより，個人のイベントに関連づけられなかったデータを検索可能とする情報検索システムの開発を行う．

## 3. 経験共有に基づく検索

### 3.1 経験共有に基づくイベント増加

複数人で経験を共有することにより以下の利点が挙げられる．

- ① イベントが増えることによりファイルの検索対象となるきっかけが増える
- ② 自分個人のイベントだけでなく，自分が知っているコミュニティ内のイベントを通して自分のファイルを検索することが出来る．  
複数人のイベントを登録することで，自分だけのイベントを検索する場合に比べ，検索の再現率を上げることができる．

### 3.2 イベントとデータの関連性

イベントと取得されたデータには関連性が存在し，特定のイベントごとに特定のファイルが関連付けられるという，イベントのタイプとファイルの種類にある程度の共起関係が存在する．例えば，学校や仕事に関するイベントにおいては，Microsoft Office のファイルや PDF ファイルなどが作成されることが多く，旅行に行った際には画像や動画のファイルが取得されることが多い．また，外出する際には外出先の情報などを Web で調べてから外出するなど，外出と Web 利用の相関性の高さが考えられる．一方，ウェブ利用時に一時ファイルとして自動的に保存されてしまう画像や時間や場所を問わず利用する音楽など，そのファイルの種類からイベントを特定しにくい物もある．以上を踏まえ，イベントとファイルの対応関係を明確にすることによりイベントを機軸にした特定ファイルの検索精度向上を検討する．

表1 イベントとデータの対応関係の例

| イベント | 学校・仕事            | 旅行    |
|------|------------------|-------|
| ファイル | Microsoft Office | 画像・動画 |

### 3.3 イベントを機軸にした情報検索

イベントを機軸とした情報検索は、イベントを考慮して、検索対象ファイルの絞込みを行う操作に相当する。以下に、イベントに対してファイルを特定する方策を示す。

#### (i) 時間的一致性

ファイルの生起がイベントの生起と同時刻の場合、時間的一致性を捉えることにより対象ファイルの絞込みを行う。

#### (ii) イベントの補間

イベントとイベントの間はイベントが存在せず、取得されたファイルを結びつけるイベントが存在しない。これに対処するため、イベント間に架空のイベント（イベントとイベントの「間」というイベント）を設定し、登録されたイベント以外のイベントに取得されたファイルを結びつけ、対象ファイル発見の可能性を増やす。

#### (iii) イベントの前後における時間的一致性

ファイルの生起がイベントの生起の前後の時間帯である場合、その時間帯における時間的一致性を捉えることにより対象ファイルの絞込みを行う。

#### (iv) 自分のイベントと他人のイベントとの相対関係

上記 (i) ～ (iii) におけるイベントを利用したファイル特定の手段は、個人のイベントだけでなく他人のイベントに関しても適用可能である。これらに加え、個人のイベントと他人のイベントの生起の相対関係により対象とするファイルを絞り込むことができる。例えば、「自分が〇〇していて、他人が〇〇していたとき」のファイルなどのような絞込みが相当する。

## 4. 事例による検証

上記に説明した情報検索手法に基づき、個人のイベントのみが登録されている際の情報検索と共同体のイベントが登録されている際の情報検索を比較してみる。

### 4.1 個人のイベントのみ登録されている場合

個人のイベントを機軸とした検索例を以下に示す。

#### (i) 時間的一致性

旅行に出かけたときに撮影した写真を検索する場合、旅行というイベントを登録することによりある程度の時間を特定し、ファイルを検索することができる。



図1 イベントとファイル作成時刻

#### (ii) イベントの補間

結婚式と2次会の間の移動中に撮影した写真を検索する場合、「結婚式」と「2次会」というイベントの間に「結婚式と2次会の間」というイベントを生成し、ファイルを検索しやすくする。



図2 イベントとイベントの「間」というイベント

#### (iii) イベントの前後における時間的一致性

授業の前後に作成した Microsoft Office ファイルを発見する。

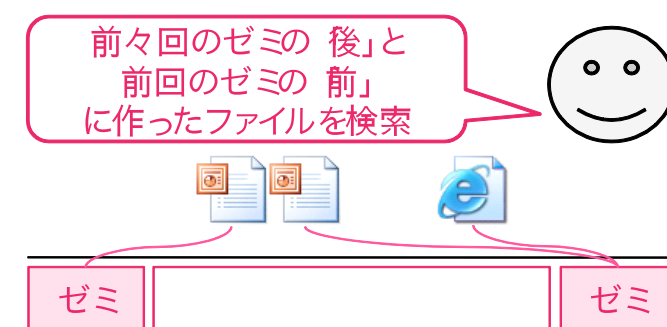


図3 ファイルとイベントの関連

### 4.2 共同体のイベントが登録されている場合

小さな共同体を家族と想定し、家族がイベント登録を行った状況下において、個人のファイルを検索する例を示す。

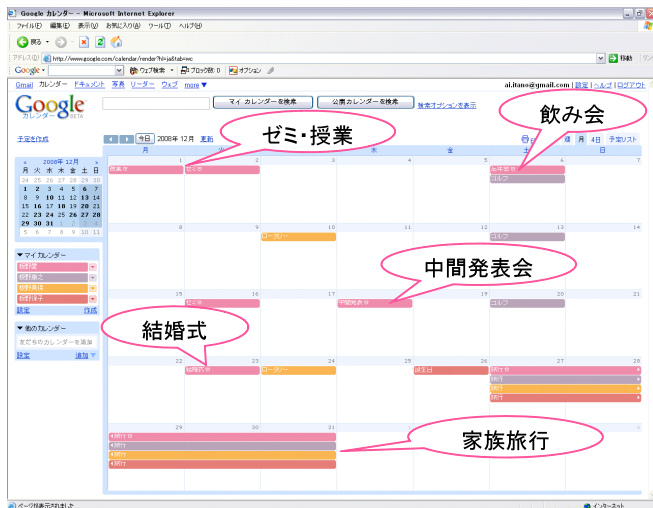


図4 一ヶ月の家族スケジュール

図4に示す一ヶ月の家族スケジュールからは、日課となっているイベント（学校や仕事）から日課ではないイベント（外出）までを知ることができ、特に日課ではないイベントは、スケジュールに登録される可能性が高く、日課となっているものよりもイベントとして取り出しやすい。

#### (iv) 他人のイベントに基づくファイルの特定

自分ひとりのスケジュールに何もイベントが入っていないとしても、同じコミュニティ内の別のイベントからファイルを特定できる。例えば、自分のイベントがない時でも、コミュニティ内のイベントがあれば、ファイルを関連付けることができる（図5参照）。

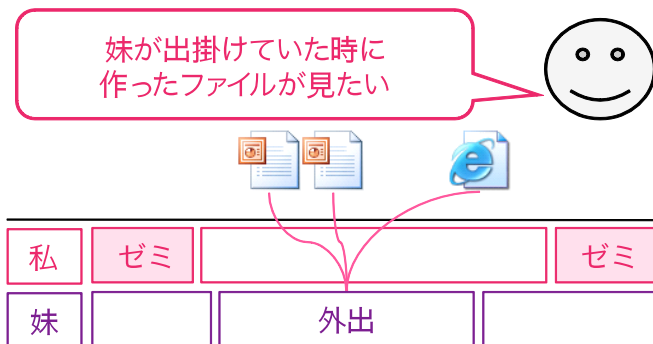


図5 自分以外のイベントとの時間的一致

検索対象候補となっているファイルが自分のイベントに多数関連づけられている際には、対象となるファイルを特定する必要がある。この際、コミュニティ内のそれぞれの人の予定を把握（または記憶）していることで、検索対象となっているファイルの絞込みを行うことが可能となり、検索効率を向上させることが可能となる。



図6 コミュニティ内のイベント

### 4.3 事例の考察

個人のイベントを中心に情報検索を行う際に、検索対象となるファイルの発生と個人のイベントが時間的に一致している際には、一般的に対象ファイルを特定しやすい。

また、イベントの「前」「後」において発生したファイルを特定する際には、前後となる時間幅の特定の仕方が問題になり、イベントのタイプや作成されるファイルの種類などの関係から時間幅を設定する必要があると考えられる。同様に、イベントの「間」というイベントにファイルを関連づける際にも「間」という時間幅の有効性を考慮する必要があると考えられる。4.2の(ii)で挙げた例においては、イベントとイベント間の時間幅が大きくないことからファイルを関連づける「間」のイベントとして有効となるが、イベント間の時間幅が大きい場合には、他のイベントと関連づけて検索を行う前段階の大きな視点からの絞込みをする際に利用されると考えられる。

一方、共同体のイベントが登録されている場合には、個人のイベントが登録されていた際の検索手法を含み、さらに、コミュニティ構成員のイベントを通じて対象となるファイルの検索が行えるようになることから検索の再現率が向上し、さらに構成員の複数のイベントに基づくファイルの絞込みが行えるようになることから適合率も上げることが可能となる。

### 5 おわりに

本研究では、人が記憶しているイベントを通して、そのイベントに関連するファイルを検索する手法の拡張を行った。従来の研究においては、個人のデスクトップ環境を対象に個人のイベントに関連するファイルを検索する手法が開発されてきたが、本研究においては、家族のような生活を共にする小さな共同体において、お互いのスケジュールを把握している環境下で、複数人による経験(イベント)を共有することにより、個人のイベントだけではなく、他人のイベントにも検索対象ファイルを関連づけ、再現率および適合率を向上させる情報検索手法を提案した。

また、提案手法により可能となるイベントに関連づけたさまざまなファイル検索手法を事例を想定し検証を行い、その有用性を考察した。

提案した手法を応用することにより、自分のファイルを検索するだけでなく自分が他人のファイルを検索する、または他人が自分のファイルを検索するなどを行う情報共有のための検索システムの構築も可能になると考えられる。今後の予定として、提案した手法の実装と検証を行うつもりである。

## 文 献

- [1] Jun Rekimoto, "Time-Machine Computing: A Time-centric Approach for the Information Environment", ACM UIST'99, 1999.
- [2] S. Dumais, E. Cutrell, J. Cadiz, G. Janck, R. Strin and D. C. Robbins: "Stuff I've Seen: A Aystem for Personal Information Retrieval and Re-Use", Proc of SIGIR'08, pp. 72-79, 2003.
- [3] Edward Cutrell, Susan T. Dumais, and Jaime Teevan: " Searching to Eliminate Personal Information Management", Communications of the acm, 2006
- [4] 松原靖子, 小林一郎 "ユーザのスケジュール情報と PC 内データ間の相関関係を利用したデスクトップ検索システムの開発" 電子情報通信学会技術研究報告. DE, データ工学, Vol.107, No.131 (20070625) pp. 31-36, 2007.
- [5] 前田晴美, 梶原史雄, 足立秀和, 沢田篤史, 武田英明, 西田豊明 "弱い情報構造を用いたコミュニティの情報共有支援", 情報処理学会研究報告. MBL, [モバイルコンピューティング], Vol.97, No.54 (19970530) pp. 25-30, 1997.
- [6] 福原知宏, 近間正樹, 西田豊明 "コミュニティの知識共有を目的とした話の共有システムの提案"
- [7] 石川雄一, 中内清秀, 森川博之, 青山友紀 "キーワード間の関連性を利用したピアツーピアファイル検索", 電子情報通信学会技術研究報告. NS, ネットワークシステム, Vol.101, No.715 (20020308) pp. 143-150, 2002.
- [8] 岡本雄太, 蝦田智則, 山名早人 "P 2 P ファイル共有ネットワーク上で作動するメタファイルシステム", 日本ソフトウェア科学会インターネットテクノロジーワークショップ 2005 (WIT2005), 2005.
- [9] 敷田幹文, 門脇千恵, 國藤進 "フローに連携した組織内インフォーマル情報共有手法の提案", 情報処理学会論文誌, Vol.41, No.10 (20001015) pp. 2731-2741, 2000.