

Title	情報共有促進支援に向けた情報利用推移モデルの一提案
Author(s)	門脇, 千恵; 爰川, 知宏; 杉田, 恵三; 國藤, 進
Citation	情報処理学会論文誌, 40(11): 3856-3867
Issue Date	1999-11-15
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/4059
Rights	<p>社団法人 情報処理学会, 門脇千恵/爰川知宏/杉田恵三/國藤進,情報処理学会論文誌, 40(11), 1999, 3856-3867. ここに掲載した著作物の利用に関する注意: 本著作物の著作権は(社)情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 The copyright of this material is retained by the Information Processing Society of Japan (IP SJ). This material is published on this web site with the agreement of the author (s) and the IP SJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IP SJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.</p>
Description	

情報共有促進支援に向けた情報利用推移モデルの一提案

門脇 千恵^{†1} 爰川 知宏^{†2,☆}
杉田 恵三^{†3,☆☆} 國藤 進^{†4}

計算機を用いた情報共有が社会の基盤となりつつある現在、情報共有促進のためのキーファクターの1つは、共有物となるコンテンツである。本研究では、オフィスで共有されている組織情報に着目し、組織情報の利用推移を、コンテンツの再利用過程を示すものという観点から分析する。そして、利用推移に応じた蓄積情報の再利用支援という見地から、組織情報共有促進支援を目指す。情報利用者の立場から関心の動向やコンテンツ全体の流行傾向を見るという従来の視点に対して、本論文では個々のコンテンツの再利用のされかたに焦点を当てて分析を行い、情報利用推移モデルを導く。分析では、組織情報の1つであるノウハウが実組織で長期間利用された履歴を用いて、特徴的な再利用の傾向を抽出し、モデル化を行った。本モデルを組織情報共有支援システムに対して応用することにより、組織情報の中でも特に組織内で世代を超えて利用される蓄積情報の再利用支援が可能であり、この支援方法についても示した。

A Proposal of an Information Life Cycle Model to Support of Promoting Information Sharing

CHIE KADOWAKI,^{†1} TOMOHIRO KOKOGAWA,^{†2,☆} KEIZO SUGITA^{†3,☆☆}
and SUSUMU KUNIFUJI^{†4}

Information sharing using computers is becoming an important infrastructure of society, and a key factor in the promotion of information sharing is sharable contents. This research focuses on the information possessed by an organization that comes to be shared within an office environment, and analyzes the life cycle of such organization information to illustrate the process of reusing contents. Moreover, from the view of supporting the reuse of accumulated information within a life cycle, this research aims to promote and support the sharing of organization. In contrast to the conventional approach that examines the overall attractiveness and interests in contents from the viewpoint of users, this paper analyzes contents with a focus on the way that individual units of contents may be reused in order to propose an information life cycle model. To this end, we use the utilization logs about the long-term use of know-how in an actual organization, we extract and model characteristic reuse tendencies. By applying this model to an organizational information sharing support system, it becomes possible to support the reuse of accumulated information particularly suitable for reuse from one generation to the next in the organization. This support method is also presented here.

1. ま え が き

近年、様々な領域において、計算機を介した情報共有が重要視されている。さらに、情報共有の方法だけでなく、共有物であるコンテンツにまで、研究やビジネスの焦点が広がってきた。最近になって使用されるコンテンツ戦略という言葉の意味からもその一端がうかがえる。このような中、コンテンツがどのくらいの期間どの程度利用されるのか、といった利用推移を把握することは重要である。たとえば、利用推移の調査結果がビジネスに反映されている。米国では、World-Wide Web (WWW) 上のコンテンツの利用推移が、

†1 北陸先端科学技術大学院大学情報科学研究科
School of Information Science, Japan Advanced Institute of Science and Technology

†2 NTT マルチメディアネットワーク研究所
NTT Multimedia Networks Laboratories

☆ 現在、西日本電信電話株式会社
Presently with NTT West

†3 NTT 研究開発推進部
NTT R&D Management Department

☆☆ 現在、NTT ソフトウェア株式会社
Presently with NTT Software Corporation

†4 北陸先端科学技術大学院大学知識科学研究科
School of Knowledge Science, Japan Advanced Institute of Science and Technology

それらコンテンツに付随して提供されている広告料金に影響することから、コンテンツの利用推移の調査方法には非常な関心が寄せられている。

WWW上のコンテンツに限らず、組織情報共有支援システムによってオフィスで共有される組織情報のコンテンツも、CSCW (Computer Supported Cooperative Work) や経営戦略論の研究領域から重要視されている^{1),2)}。そして、情報共有の枠組みに主眼を置いた研究が進められ^{3),4)}、多くのコンテンツが獲得可能となった反面、有用なコンテンツが多くのコンテンツに埋もれてしまうという情報埋没化の問題も起き始めている⁵⁾。このような背景を受けて、組織情報共有を促進するためには、新しいコンテンツの獲得を図ることのみならず、蓄積されている有用なコンテンツの再利用を促すことも重要である。再利用を促すためには、不用になったコンテンツを特定しそれらを淘汰してゆくことによって有用なコンテンツの埋没化を防ぐという考え方と⁶⁾、有用なコンテンツを抽出し他より表出化させるという考え方がある⁵⁾。そして、どちらの場合の実現にも、“不用なコンテンツ”、“有用なコンテンツ”を取捨選択するための判断基準が必要となる。しかし、各コンテンツに対して有用/不用などの情報生命の判断や、なぜそのコンテンツを再利用するのかといった再利用目的を、利用者から毎回聞き出すことは不可能である。そこで、利用推移の分析によって、たとえば、コンテンツが情報生命を持つ期間や再利用の目的が発生する可能性の有無を推測できれば、そのコンテンツが組織にとって有用か不用か取捨選択の判断材料の1つとなる。また、利用者に依存せず、コンテンツの観点から利用推移が分析可能となれば、分析の大量処理や自動化が可能となる。利用推移の分析によって、組織で繰り返し利用されるコンテンツを抽出し、コンテンツの取捨選択を行うことは、蓄積された組織情報を洗練し、組織知²⁾の再構築へとつながる。しかし、このようなコンテンツの利用推移に着目した蓄積情報の再利用支援の研究は確立されていない。

組織情報と呼ばれる情報には、対象とする協調作業の種類により様々であるが、その内容はおおむね、仕様書のような作業の対象情報、途中成果物に関する議論のような会話情報、仕事に関するノウハウや本のような参照情報に大別される⁷⁾。これらの組織情報を、利用推移という観点から見た場合、流行情報のような一過性の情報から、ノウハウのように組織内で繰り返し参照される再利用性の高い情報もある。ノウハウは再利用価値が高い反面、支援システムの検索機能などだけでは、効率の良い再利用は難しいといわれる⁸⁾。

本研究のねらいは、ノウハウのように組織内でメンバを越えて繰り返し利用される情報を対象に、それらコンテンツの再利用過程に着目することで、利用推移の特徴を分析によって抽出し、その結果に基づいて、組織情報の再利用支援の土台となるモデルの提案を行うことである。

以下、2章ではまず、情報の利用や利用動向に関する従来の研究を紹介しつつ、本研究におけるモデル提案のためのアプローチと分析手法を述べる。3章では実組織を対象に行った分析のケーススタディについて概略を述べ、4章では分析の結果を受けて情報利用推移モデルを提案する。5章では蓄積情報の再利用支援への提案モデルの適用について検討する。

2. 従来研究とアプローチ

2.1 従来研究

コンテンツの利用結果を評価し次の新たな利用へとつなげる研究には、推薦システムや協調フィルタリングの研究があり^{9),10)}、WWWやネットニュースをコンテンツ源に研究が行われている^{11)~13)}。これらの研究の主眼は、利用者の関心に応じた情報の提供であり、またキーワードの利用履歴の分析をもとにした Mailing List における興味の変遷についての研究¹⁴⁾はあるが、情報の利用推移を表現するモデルに関する研究は行われていない。

一般社会における情報の利用や利用動向に関するモデルには、社会科学や普及学の分野において、注射針モデル、二段階流れモデルなどが提唱されている¹⁵⁾。これらのモデルは、情報の再利用の過程ではなく、情報や知識などが伝達され普及してゆく過程をモデル化している。

計算機を介した電子メディア社会において、情報の利用や利用動向に関するモデルは少ない。ネットワーク利用のための情報入手先の調査結果に基づく情報流通の三段構造¹⁶⁾、集団の性質や導入される電子メディアの性質に基づいて提案された定着段階モデル¹⁷⁾や、グループウェアからの脱落を考察したサービス間利用遷移概念モデル¹⁸⁾がある。これらのモデルは、利用者/組織、利用プロセスといった概念レベルをモデル化の範疇としている。そのため、本論文で取り扱う対象、つまり、再利用の過程の中で具体的にやりとりされるコンテンツに比べて情報の粒度が大きい。また、塚田らは、オフィスワーク支援ツールを分析し、様々なツールが扱う情報の形態を包括的に表現可能な情報のアクセスモデルを提案している¹⁹⁾。しかし、情報共有のシーム (アクセス形態が異なるツール間で情報を

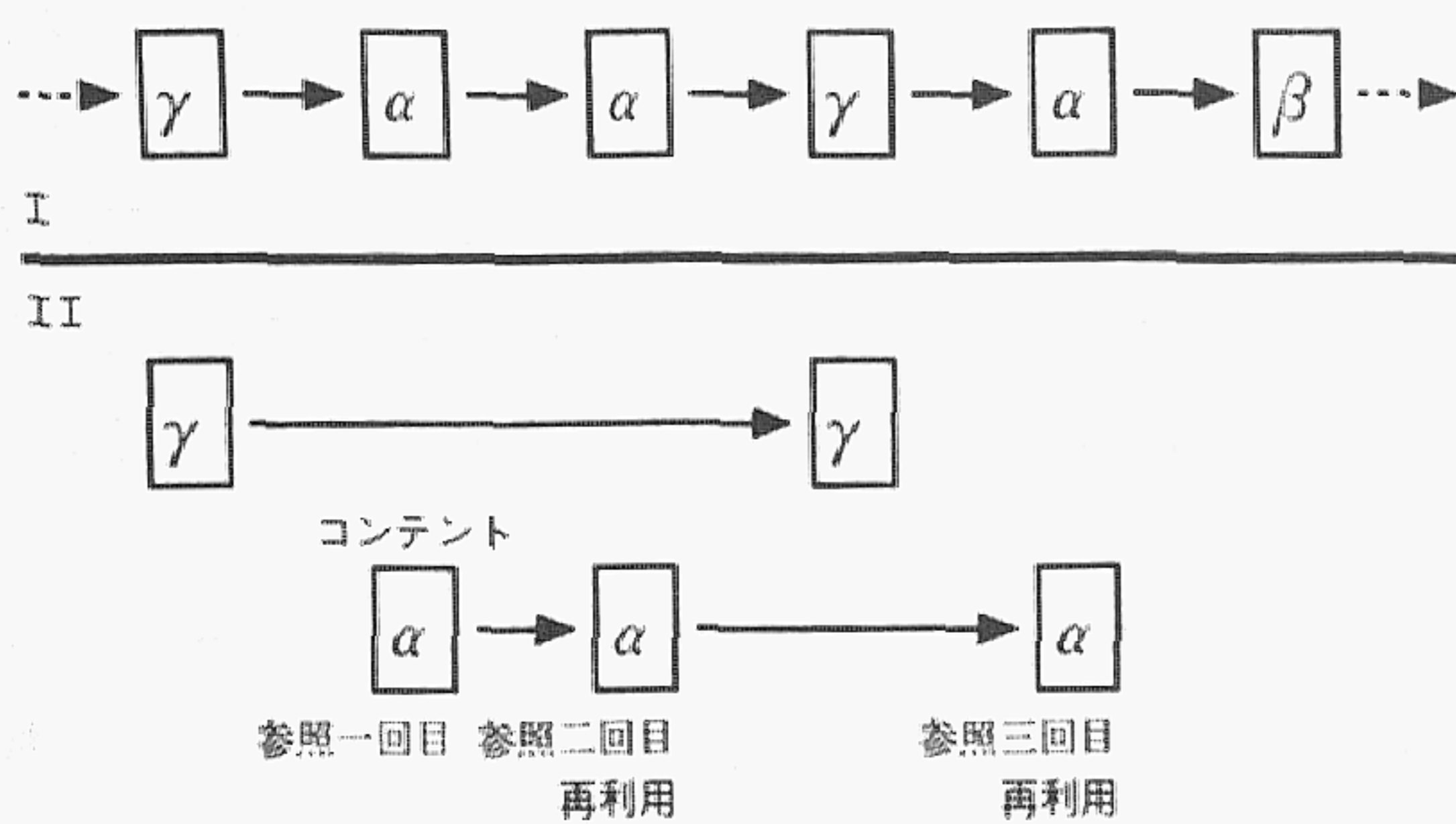


図1 利用推移のとらえ方

Fig. 1 Standpoints of the information life cycle.

共有する際のオーバーヘッド、たとえば、情報の交換操作など)の解消を目指しているために、このモデルが対象とする情報の粒度も、コンテンツというレベルよりも大きい。

2.2 アプローチ

上記のように、従来提案されたモデルでは、コンテンツという粒度の細かい対象を取り扱うには十分でなく、本論文で取り扱う利用推移の観点とは以下のように異なる。利用推移の分析を行う際、次の2つの観点が考えられる(図1)。(1)利用者に利用された情報の全体像を分析する(図1(I))、(2)利用者に利用された情報を各コンテンツごとに着目して分析する(図1(II))、の2つの場合である。前者は従来研究で用いられてきた観点であり、取り扱われる情報の粒度は大きい。この観点からの分析では、利用された一連のコンテンツの変遷が明らかになり、利用者の情報への関心の動向を推測しうる。これは、その利用者の中での情報の流行傾向とも見ることができる。後者の場合、あるコンテンツ、あるいはある種のコンテンツ^{*}が最初に利用されてから、そのコンテンツの利用が終結するまでの経過が明らかになり、コンテンツが情報生命を持つ期間が推測可能となる。これは、本分析の目的とするところである。よって、後者の観点から利用推移の分析を行う。以降、本論文では、後者の観点における利用の繰返し、すなわち、参照の繰返しを“再利用”と呼ぶ。参照が1回のみの場合には再利用とは呼ばない。1回目の利用から、2回目以降再利用されなくなるまでが、再利用の過程である。“利用推移”とは、初回の利用に続く再利用の過程を指し、あるコンテンツの1利用推移とはそのコンテンツの1再利用

過程という位置付けである。たとえば、図1(II)のコンテンツαの場合、参照の2回目以降がコンテンツαの再利用となる。そして、参照1回目から参照3回目までの過程がコンテンツαの利用推移である。

ここで、前述の塚田ら¹⁹⁾の情報アクセスモデルにおける「人」、「物」、「時」、および「目的」の4つのクラスは、種々のグループウェアで生じる情報空間を構成する基本的要素であり、本研究の対象である組織情報共有支援システムにも適用可能である。そこで、上記4つのクラスを本研究に導入し、以下のようにこれらクラスを具体的に設定して分析を行うことで、利用推移を細かく分析する。

- 「人」というクラスは、利用者各個人という視点からとらえる。
- 「時」というクラスは、ある情報の利用経過過程を表すものとして、コンテンツの再利用間隔として表す。たとえば、図1(II)のコンテンツαにおける参照1回目と参照2回目の時間間隔である。
- 「物」というクラスは、分析のターゲットとなるコンテンツである。たとえば、図1(II)のコンテンツαにあたる。具体的には、3章で述べるカード1枚もしくは、ある視点で包括されるカード群である。
- 「目的」というクラスは、情報共有システムの場合、情報の利用目的となる。特に、本分析では情報の再利用目的となる。この再利用目的に起因して、コンテンツの再利用のほとんどが生じ、それら再利用の積み重ねが利用推移となる。

今回の分析によって最終的に求めようとしている利用推移は、再利用目的の推移と表裏一体の関係にある。つまり、分析で求めようとしている利用推移の意義を換言すれば、コンテンツの有用/不用などの再利用価値の判断材料として各コンテンツに今後再利用目的が生ずるか否かを推測することである。再利用目的は利用者の内的要因であるため、利用者によって再利用目的についての入力処理が行われない限り、分析可能な計算機上の記録としては残らない。しかし、逆に再利用の過程の記録を分析することによって、再利用の目的がいつの時点で発生し、いつの時点以降には発生していないかを推測することは可能である。つまり、計算機を用いて、再利用目的の詳細まで分析することはできないが、再利用目的の生じた時期や頻度は推測可能である。よって、情報空間における「目的」の推移は、残りの3つのクラスの記録から間接的に推測可能である。情報空間の3要素の記録を分析することで、利用推移の特徴を抽出するアプローチをとる。

^{*}ある種のコンテンツとは、ある視点によって同じ範囲に入るコンテンツを指す。ケーススタディの場合では、各コンテンツにキーワードが付随するが、たとえば、“freshman”というキーワードを持つという視点からは、コンテンツの区別は関係なく“freshman”キーワードを持つコンテンツの集合を指す。

2.3 分析の手法

情報空間の3要素の記録を分析する手法としては、アンケートやインタビューなどによって利用者の再利用の記憶を調査する方法と、システムの利用の際に生じたログなどの利用履歴を解析する方法が考えられる。前者はコンテンツの再利用状況を細かく聞き出すことはできるが、主観的な回答になりやすく、また、再利用に関する記憶をすべて正確に拾い出すことは困難である。一方、後者における統計的処理は機械的な処理を主とするため、利用履歴さえ残っていれば、多くの利用者の長年にわたる再利用の傾向をより簡易に割り出すことができる。また、客観的評価を行ううえで有用である。そこで、本研究では利用履歴解析によって、利用推移を分析する。

3. ケーススタディ

オフィス業務には棚卸しなど季節的な周期性を持つものもあり、このような業務に係る組織情報の利用推移を分析するためには、再利用の周期性も考慮して同一組織の同一メンバーを短期でなく観察できることが望ましい。

よって、本論文の目的とする利用推移の分析には、同一利用者の再利用の記録、つまり、「人」、「時」、「物」という情報空間の要素が非短期間途切れることなく記録されており、かつ、それらの記録を各コンテンツ、すなわち、図1(II)の視点から分析できるという条件を備えた組織情報共有支援システムの利用履歴を解析する必要がある。また、再利用の対象となった組織情報の種類に偏りがないことが必要である。本論文ではケーススタディとして、組織内で実運用されていたノウハウ蓄積システムの利用履歴を取り上げた。さらに、組織情報の投稿にあたって入力制限がないため、豊富なジャンルの情報が蓄積されている。また、ほぼ3年間にわたり同一グループに在籍した利用者の利用履歴が残っている。これらの利用者には、システムを介した組織情報の利用は義務付けられていないため、自然な再利用の傾向が観察可能である。

3.1 分析対象

システム：ノウハウ蓄積システム FISH²⁰⁾であり、組織におけるインフォーマルな情報の蓄積と共有を目的に、1991年1月に観察グループに導入された。その後、1994年4月に分散環境への適応性を考慮した GoldFISH²¹⁾に移行された。(1)分散環境という要因の影響を受けていないサービスを分析することにより分析の精度を上げる^{*}、(2) FISHでは運用開始時から同種の代替サービスのない状況において安定して

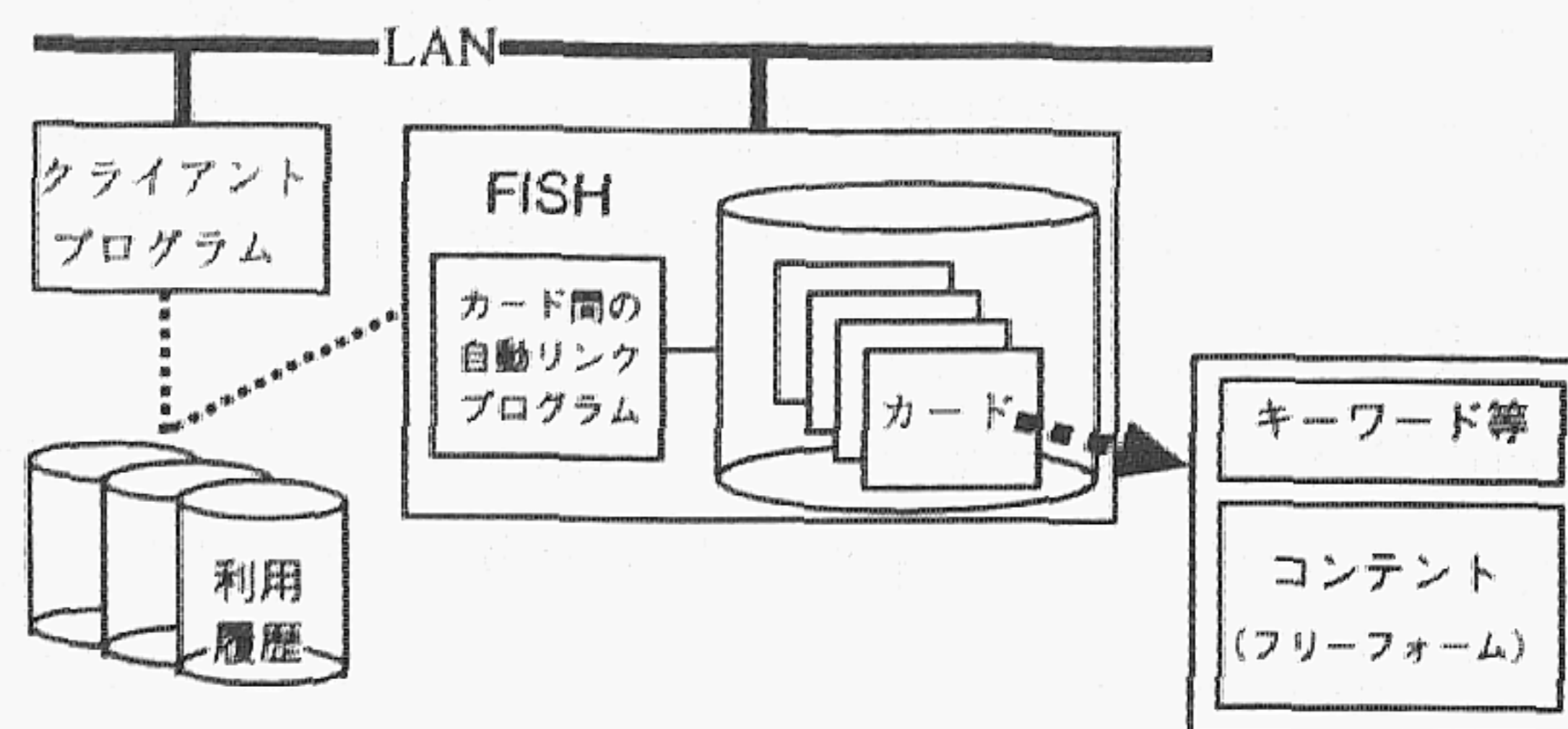


図2 FISH概要

Fig. 2 Outline of FISH.

% kpick perl			
1	05Jun93	作成者名 A	perl のループ文
2	23May93	作成者名 B	perl の日本語入力
3	17Feb93	作成者名 C	jperl

図3 参照可能候補の表示例

Fig. 3 Display example of candidates for reference.

記録された利用履歴が残っている、という理由から、今回は FISH を対象とした。

観察グループ：研究開発を主たる業務としているグループであり、分析対象となったメンバーは、FISH 運営期間の9割以上の期間をこのグループに在籍していた10人である。

組織情報：FISHでは、蓄積情報は「カード」というメタファとして蓄えられ、複数キーワードによる検索が可能である(図2)。カードは1枚ずつカードIDで管理され、属性部とテキスト部より構成される。属性部は、カード作成者や検索用キーワードなどの項目から構成されている。テキスト部は、自由文入力のテキストとして組織情報を蓄えている。カード間にはフルテキストサーチによる自動リンクが張られ、関連情報の参照が容易である。蓄積カード数は約1200件で、カード内容は、たとえば、業務関連規則、伝票処理手順、トラブル処理手順、計算機環境の説明、社内用語解説、FISHに関連した情報、時刻表、連絡先である。

利用方法：カードの利用は、UNIX上のコマンドインタフェースにより行われる。特定のカードを利用するためには、まずキーワード入力による検索を行う。検索が成功すると、入力したキーワードに適合するカードの見出しが適合した数だけ提示される(これを参照可能候補と呼ぶ)。図3は参照可能候補の表示例である(たとえば、入力キーワード「perl」で検索が成功し、参照可能候補が3件表示された場合である)。これら参照可能候補から番号を選択すると、カードの

^{*} モデル化にあたり分散環境の要因を全面的に排除するのではなく、よりシンプルなシステムでモデル化を行った後に、複雑なシステムへモデルを適用する方針である。

属性部とテキスト部が表示される。以降、本論文における“コンテンツの利用1回”とは、参照可能候補から番号を1つ選択し、カード1枚の属性部とテキスト部を表示することを指す。

利用履歴：利用履歴は、分析の解析プログラムにおいて、インプットファイルとなる直接の分析対象である。利用履歴の1レコードは、「コンテンツの利用1回」を記録しており、利用の年月日時刻、カードID、利用者IDの項目から成り立っている。観察グループでは、FISHの導入後から運用終了までの3年余りの期間に、カードの参照によって生じた約1万4千件の利用履歴が記録されている。このうち、対象利用者10人分の約9800件を今回の分析に用いた。

3.2 分析結果

再利用の過程から再利用の特徴を抽出するために、以下の手順で分析を行った。まず、利用履歴をもとに、(1) 利用履歴全体に対して再利用の起こった割合と利用履歴に記録された再利用過程の数、(2) 1つの再利用過程における再利用の度合を計るために、再利用回数、再利用間隔、再利用期間を、cshとperlで作成した解析プログラムを用いて各々調べた。次に、(3) (2)で得られた実測数値から抽出可能な再利用傾向の表現方法を考察した。(4) (2)で得た実測数値に(3)の表現方法を適用し、特徴的な再利用の傾向を整理した。なお、手順の(2)~(4)における「1つの再利用過程」とは、「ある利用者があるコンテンツ、すなわち、特定のカード1枚を再利用した過程」を指す(観察条件a)。以下に、(1)~(4)の手順に沿って分析結果を示す。

(1) 再利用の発生割合：分析対象の利用履歴全体に対して、再利用されなかった割合は16.7%であった。残り83.3%は、少なくとも2回以上は再利用されており、それら履歴における利用者とコンテンツの組合せ数、つまり、利用推移の観察対象となる再利用過程は1217である。

(2) 再利用の度合：(1)で得た再利用過程の各々について、再利用の度合を調べた(図4)。図4は、図1(II)における1利用推移(たとえば、コンテンツ α の利用推移)を詳細化したものである。図4中の“ n ”とは、あるコンテンツがある利用者に利用された回数を表している。また、 x を2以上 n 以下とすると、図4中の“ t ”は、 $(x-1)$ 回目の利用時から x 回目の利用時までの時間間隔(再利用間隔)である。そして、 n 回目の時刻から1回目の時刻の差が再利用期間である。ここでは、すべての再利用過程における再利用回数の分布を図5に示す。X軸が再利用回数で、Y軸はその再利用回数の過程が占める数の割合である。最大再

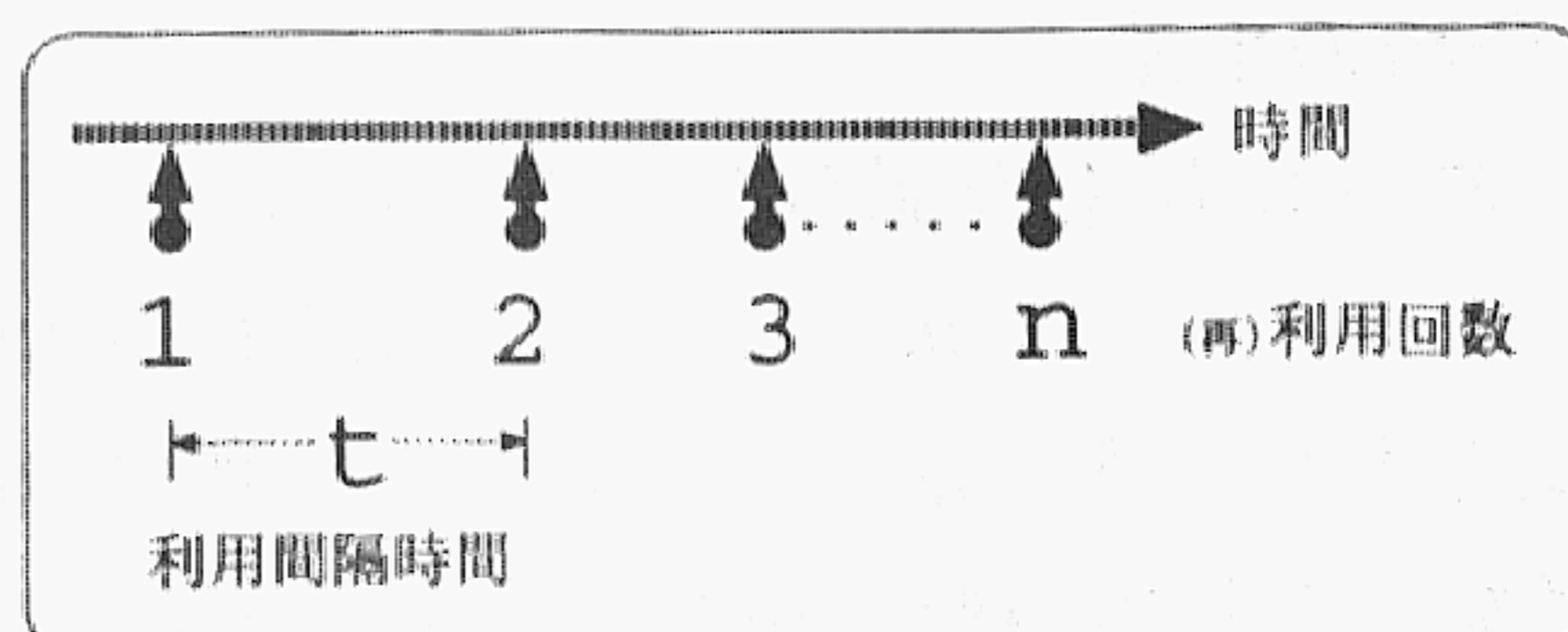


図4 1 再利用過程における分析点
Fig. 4 The point of an analysis.

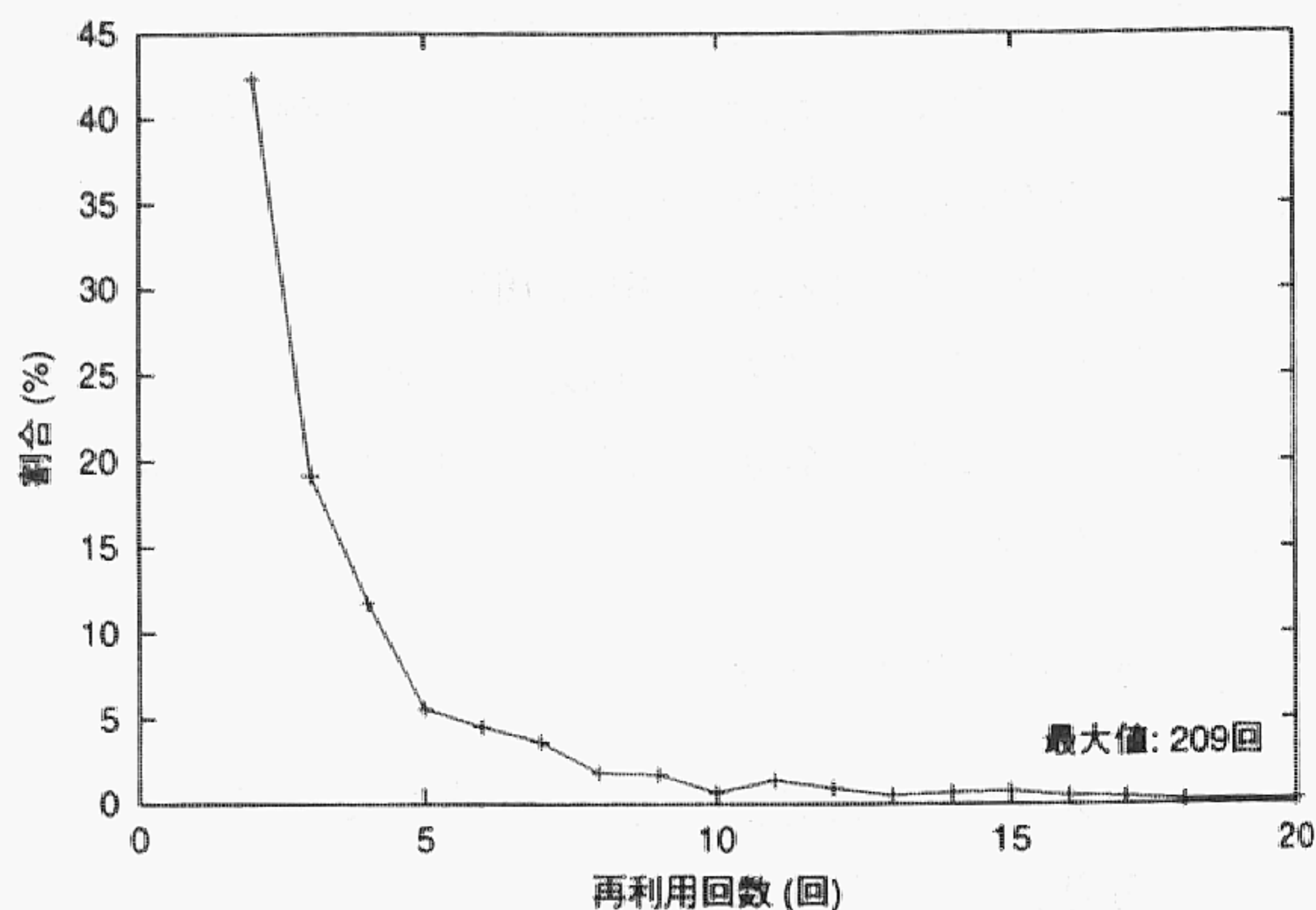


図5 再利用回数の分布
Fig. 5 Distribution of number of reusing.

利用は209回であったが、回数が多い場合は割合がきわめて小さいため、再利用回数20回以上の部分は省略した。また、すべての再利用過程における再利用回数の平均値は5.50回である。

(3) 特徴の表現方法：(2)で得た結果から、再利用過程にはいくつかの傾向があることが分かった。しかし、分析結果が数値の羅列では、再利用の特徴を明確なイメージとしてとらえ難い。そこで、再利用の特徴を視覚化するために2次元のグラフとして表現する。グラフ化では、図4における $(n-1)$ をX軸に t をY軸にとり、Y軸の単位は時間とした。たとえば、X軸上における最初のプロット点は、コンテンツがある利用者に、最初に利用されてから次に(2回目)に再利用されるまでの時間の間隔を表している。

(4) 特徴的な再利用傾向の抽出：本分析の目的は、コンテンツの利用に関するすべての傾向を求めることではなく、再利用過程における特徴的な傾向を抽出することにあるため、(2)で得た再利用過程1217各々の分析結果に(3)の表現方法を適用した結果、3つの再利用傾向が抽出された。図6は、抽出された3つの再利用傾向をイメージとしてつかみやすくするために、抽象化した図である。各傾向のグラフ上のプロット点の特徴は以下のとおりである。

- 傾向1：図6(a)の傾向では、初回の利用から後、再利用間隔のほとんどは短間隔である。グラフ上では、序盤からのプロット点の8割以上がほぼ一

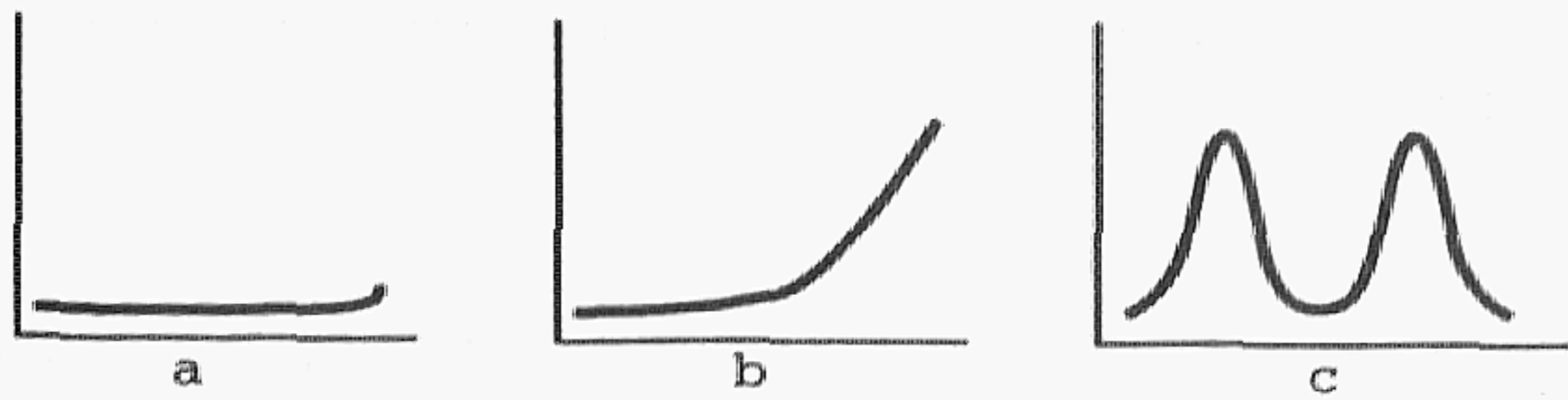


図6 抽出された再利用傾向の特徴

Fig. 6 Extracted characteristic patterns of reusing content.

直線に並ぶ傾向を，“一字型”傾向と呼ぶ。

- 傾向2：図6(b)の傾向では，再利用の間隔は，序盤では短間隔の傾向が強いが，中盤以降は長間隔の度合が次第に強くなってゆく。グラフ上では，プロット点は序盤では一字型に近付くが，後半は次第に間隔が開いてゆき，プロット点の趨勢が結果的には右上がりになる傾向を，“逆L字型”傾向と呼ぶ。
- 傾向3：図6(c)の傾向では，再利用の間隔は，序盤・中盤・終盤とは関係なく短間隔と長間隔を繰り返す。グラフ上でプロット点の趨勢が結果的には右上がりとは右下がりを経る傾向を，“波型”傾向と呼ぶ。

上記の特徴的な傾向以外は，傾向が認められないものと，プロット点が少な過ぎる，などの理由で傾向が判定できないものがあつた。非長期の利用期間に再利用が無作為に行われ，グラフ上でプロット点の傾向が認められないものを，“ランダム”傾向と呼ぶ。

以下に，各傾向を示す実測例を紹介する。図7は一字型傾向を示す実測データである。ある利用者は約2日間にわたり，このコンテンツを9回利用している。コンテンツの内容は「perlのレクチャスクリプト」であり，プログラミングなどの特定の作業中に頻繁に参照される情報である。図8は逆L字型傾向を示す実測データである。ある利用者は約95日間にわたり，このコンテンツを18回利用している。コンテンツの内容は「高速マルチメディア通信」であり，業務遂行過程でたびたび参照が必要となった補足的な情報である。業務完了後は参照の必要性はそれほどない情報である。図9は波型傾向を示す実測データであり，ある利用者は約1020日間にわたり，このコンテンツを32回利用している。コンテンツの内容は「会議室の電話番号一覧」であり，通常業務の形態に応じて使用/不使用になる情報である。図10はランダム傾向を示す実測データであり，ある利用者は約116日間にわたり，このコンテンツを5回利用している。コンテンツの内容は「飲食店の場所」であり，業務には直接関係しない情報である。

以上は，各傾向の特徴を顕著に表す実測データであ

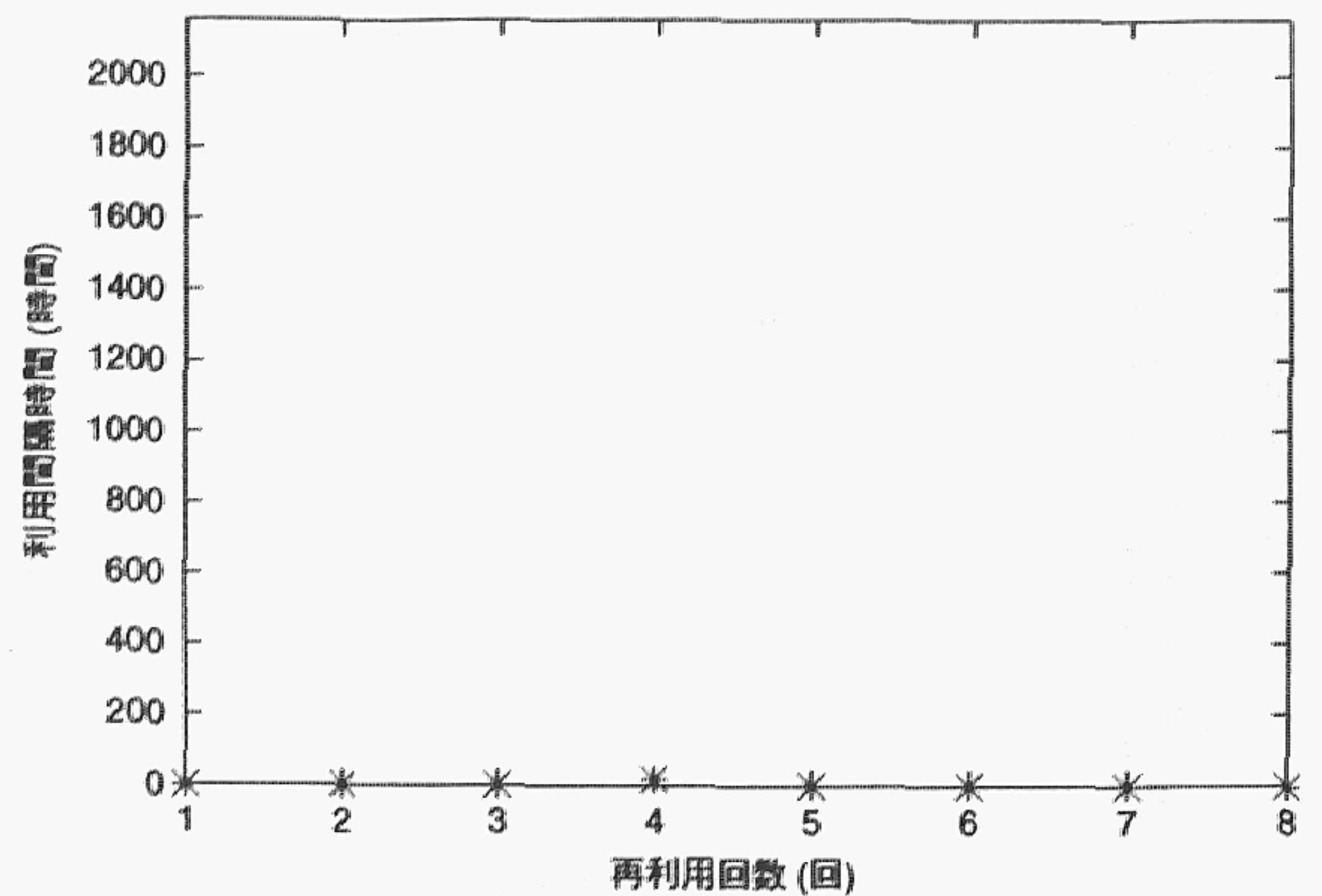


図7 一字型傾向を示す実測データ

Fig. 7 An instance of a “straight line” pattern.

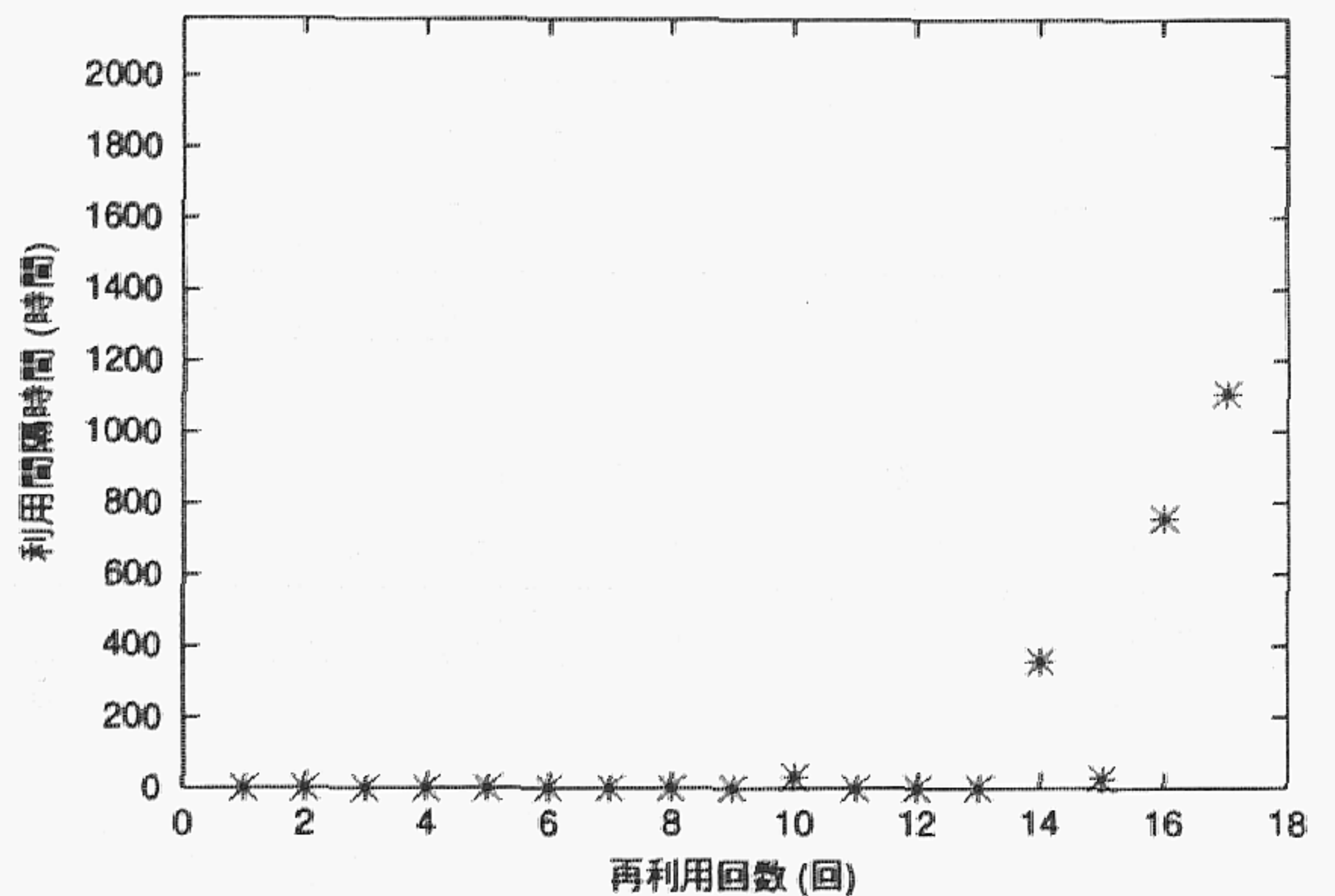


図8 逆L字型傾向を示す実測データ

Fig. 8 An instance of a “rising line” pattern.

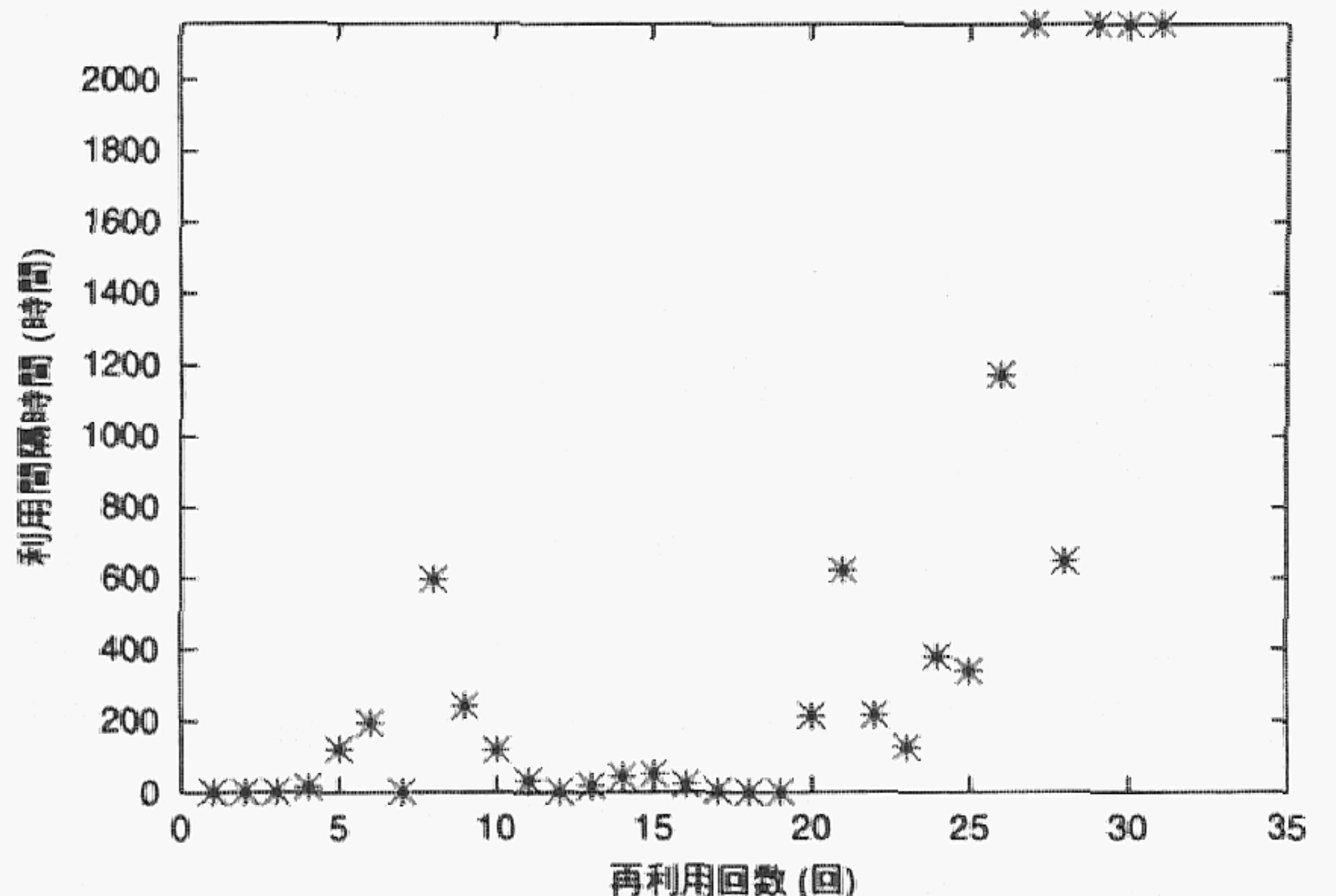


図9 波型傾向を示す実測データ

Fig. 9 An instance of a “wavy line” pattern.

る。また，データによっては，グラフ上でプロット点の傾向が，一字型と逆L字型，逆L字型と波型，波型とランダム，の各中間傾向を示すものも認められた。

表1は，利用回数が5回以上の場合の再利用過程321に占める各再利用傾向の割合を示したものである。なお，利用回数が少ないデータは傾向がつかみにくいいため，再利用過程1217の利用回数の平均値5.50を参

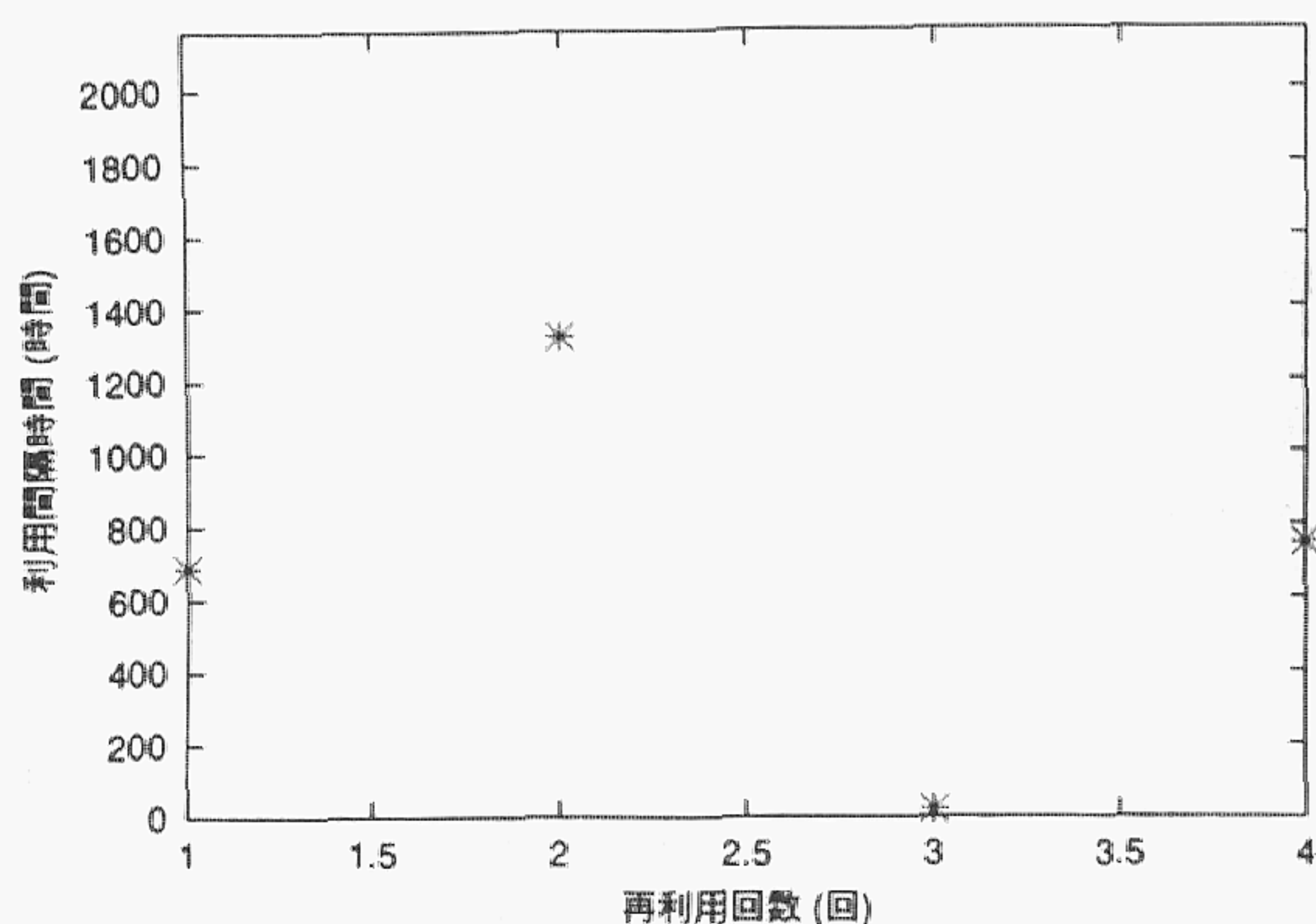


図 10 ランダム傾向を示す実測データ

Fig. 10 An instance of a "random dots" pattern.

表 1 各再利用傾向の割合

Table 1 Appearance of reusing patterns.

再利用傾向				
一字型	逆L字型	波型	ランダム	その他
16.5%	20.0%	19.0%	10.9%	33.6%

表 2 再利用傾向のまとめ

Table 2 Statistical values of reusing patterns.

再利用傾向		一字型	逆L字型	波型	ランダム
再利用回数 (回)	最小	5	5	5	5
	最大	46	46	209	99
	平均	9.6	9.2	30.9	11.2
再利用期間 (日)	最小	0.0	7.2	30.8	10.1
	最大	953.1	938.8	1339.6	559.0
	平均	68.6	270.9	447.8	122.7
再利用間隔 (時間)	最小	0.0	0.0	0.0	0.0
	最大	22858.2	22372.1	18584.4	2159.9
	平均	190.9	791.5	359.9	287.9

考に利用回数を5回以上とした。

(5) 特徴のまとめ: 表2は、各傾向ごとに前述の(4)と(2)の結果を総合したものである。

ここで、上記までの分析条件(観察条件a)による結果と、「物」というクラスをある視点で包括されるカード群に変えた場合(観察条件b)の結果とを比較する。図11は、ある利用者が「電子メールの使い方」という内容を持つカードを再利用した過程をグラフ化したものである(観察条件a)。この利用者は、入社後からこのカードを利用し始め、約36日間の間に全部で9回利用し、その後はいっさい利用していない。また、このカードには、「email」というキーワードのほかに、「freshman」という別のキーワードがつけられている。freshmanキーワードとは、新入社員や中途転入者がさしあたって必要となる重要な業務ノウハウや計算機環境の情報などが記載されたカードに、本来のキーワードとは別にこのキーワードをつけて検索

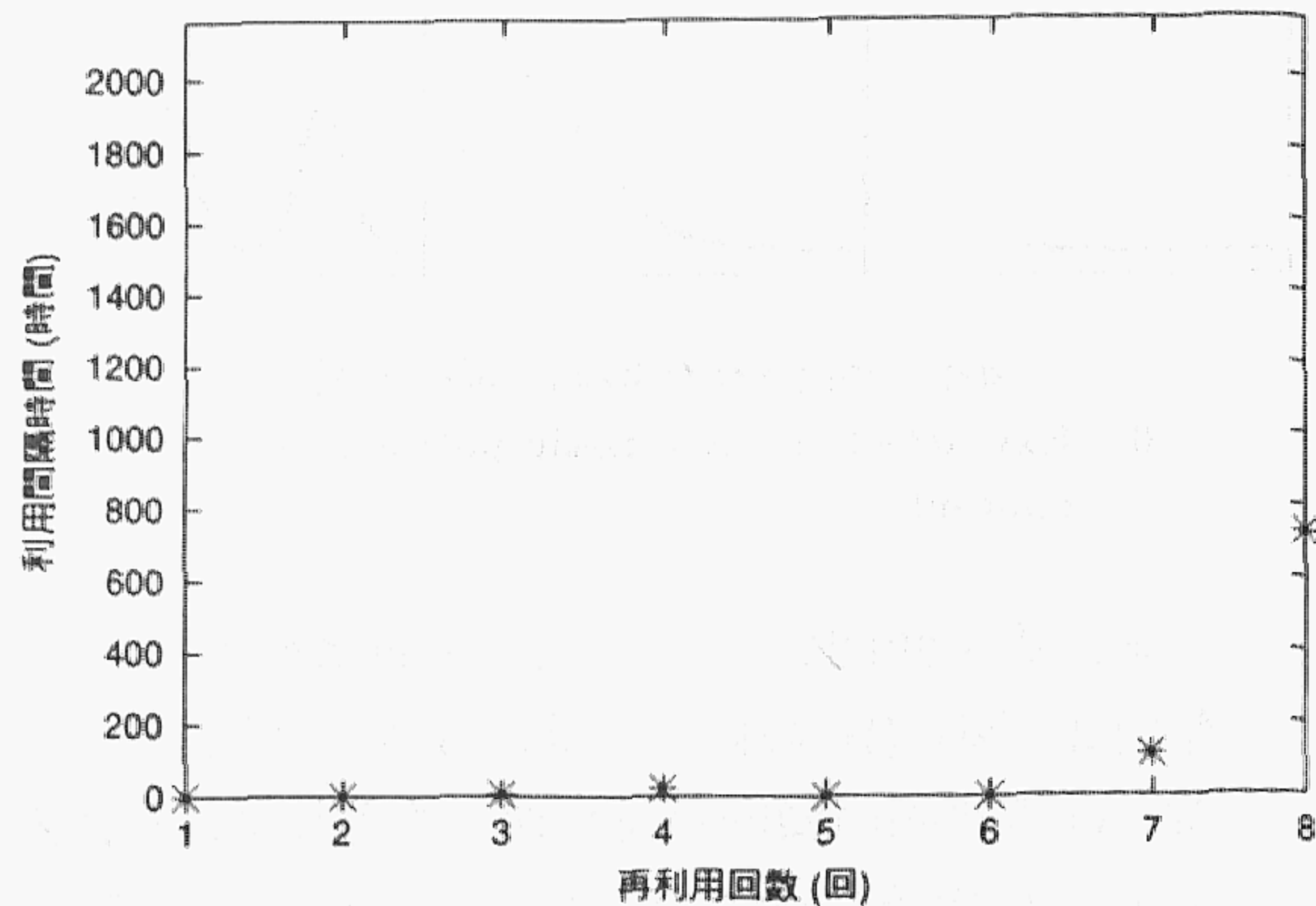


図 11 freshman カード 1 枚の利用推移

Fig. 11 A life cycle of one "freshman" card.

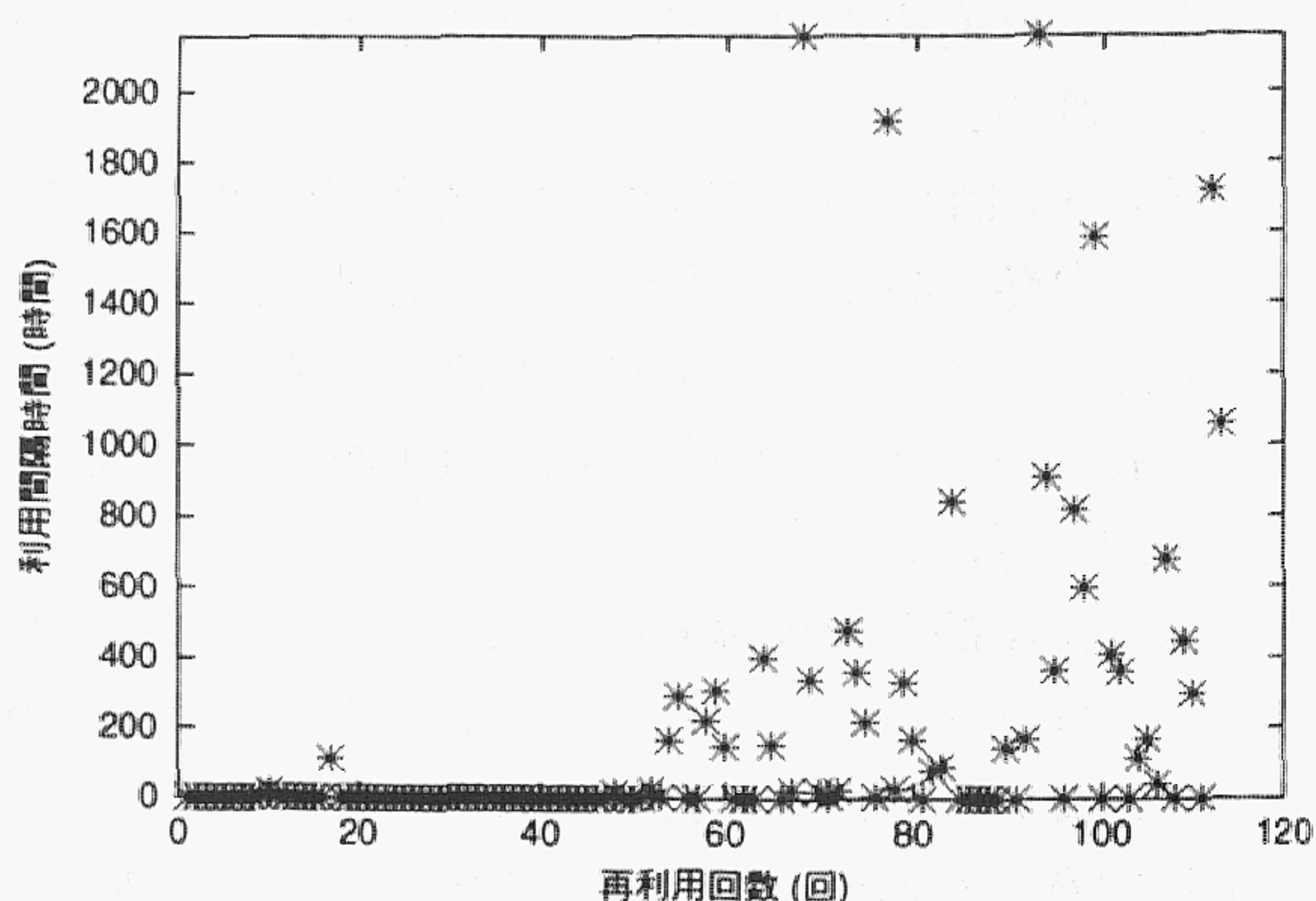


図 12 freshman カード群の利用推移

Fig. 12 A life cycle of "freshman" cards.

しやすくしたものである。そこで、この利用者が再利用したfreshmanキーワードを持つカード群を1つの再利用過程として(観察条件b)、利用履歴から取り出し観察条件aと同じ方法で分析した。その結果をグラフ化したものが図12であり、freshmanという視点からの利用推移を表している。このfreshmanカード群は、1017日間に114回利用されている。カード群は、43枚のカードからなり、それらのカード内容は伝票処理手順などの事務処理内容が多い。図12のグラフからは、入社後の利用者が年数を経るに従い、freshmanという視点から見たカード群の再利用間隔が開いていくことが分かる。上記のように、ある利用者(たとえば、新入社員)において、ある視点から見た場合に同一範疇に入る情報(たとえば、freshmanカード群)の再利用傾向は、おおむね似た傾向をとると考えられる。つまり、各利用者において、再利用傾向が同様な傾向を示すコンテンツは、それらコンテンツの取得の必要性も似ているものと考えられる。

以上、分析によって抽出された再利用傾向を区別するための閾値は、対象となるシステムや参照頻度、業

務の性質などによって変化するため、一般的な尺度を定めることは難しいと考えられる。しかしながら、各再利用傾向は比較的明確に判別できるものであり、主観による判別結果の差異はほとんど認められない。さらに、今回得られた分析結果は、分析ケーススタディに特化したものではなく、グループダイナミクスにおける集団の凝集性²²⁾を持つメンバ、つまり、ある目的のために組織化された不特定多数ではないメンバによって蓄積・共有されている組織情報に対しても適用可能と考える。

4. 分析からの考察とモデルの提案

4.1 考察

まず、再利用傾向と情報の分類との相関について考察する。ある分類に属する情報のすべてが、まったく同一の再利用傾向としては抽出されないことが分析により分かった。これは、同じ分類に属する情報でも、情報の利用価値によって再利用傾向が変わること、ノウハウは形式化しにくく明確な分類の対象とはならないこと、利用場面によってノウハウのどの部分が利用されるかわかること、情報の分類方法自体が確立されておらず完全な分類結果は望めないこと、コンテンツがテキストで書かれているため複数分類の内容を含む場合があること、などの理由が考えられる。ここでは、各再利用の傾向にどのような情報が該当するかを考察するための補助として、あえて利用が5回以上の場合における各再利用傾向に多く認められたコンテンツの上位3分類を表3に示す。

次に、上記までの分析結果をもとに、再利用傾向について考察を行う。

一字型傾向と逆L字型傾向には、教育・知識関係のカードなどのように業務に直結した知識として覚える情報やFISH関連や計算機環境のカードが多く認められる。FISH関連のカードにはFISHを効率的に使うためのコマンドオプションなどが記載されており、また、計算機環境のカードはネットワーク化されたオフィスにおける業務遂行を補足するために利用価値の高い情報である。逆L字型傾向では、カードの利用期間は短期ではないことから、これらのカードは、業務

の遂行とともに徐々に習得されて、再利用間隔が開き始めると考えられる。一字型傾向では、ほとんどが短期間に集中して再利用されていることから、業務遂行にあたってすぐ必要となる情報、あるいは、毎日の業務で使う情報を短期間で覚えるため、利用回数も少なくグラフは直線的になる。これは、逆L字型の序盤の部分(一直線の部分)で利用が終結したと考えられる。

波型は、時刻表や連絡先のカードに多く認められ、これらカードの参照された回数は多く、利用期間も長い。しかし、利用頻度は一定ではない。これは、時刻表や連絡先のカードは利用者が取得すべき情報ではなく、ニーズが生じた時点で参照だけすればよい情報であるためと考えられる。さらに、ニーズの発生は業務の周期性とも関係があると考えられる。また、時刻表や連絡先のカードは他のカードに比べて、内容が変更される可能性が高く、それらの変更が周期と関係している可能性もある。

再利用間隔がランダムなカードは、利用されるタイミングに規則性は見られず、偶発的にしか利用されない情報か、あるいは、業務とは関係なく関心の惹くまま散策的に利用される情報である。

4.2 利用推移モデル

以上の考察から、コンテンツの再利用間隔を特徴付ける要因は、コンテンツ取得の“急迫度”であると考え、“利用頻度”と“利用期間”によって表される再利用の傾向と情報取得の急迫度の関係を表す利用推移モデルを提案する(図13)。急迫度とは、あるコンテンツが業務遂行にあたって、必要な情報としてどれだけ取得を急がれているかという概念である。“利用頻度”と“利用期間”の2つの要素によって2次元で表現可能な利用推移の概念を、1つの要素「急迫度」で再表現したものであり、情報の利用推移を特徴付ける指標となる。このモデルは、以下の(1)~(4)の順にコンテンツが取得される急迫度が低くなるととらえる。急迫度は、(1)と(4)を両極値として連続した値をとる概念である。

- (1) 図13左上の一字型傾向をとるコンテンツは、序盤に利用が集中し、かつ利用期間が短いため、情報取得の急迫度が高いコンテンツととらえる。
- (2) 図13右上の逆L字型傾向をとるコンテンツは、序盤から中盤にかけて利用の集中度がやや高く、かつ利用期間は短くないため、上記(1)よりは情報取得の急迫度がやや低いコンテンツととらえる。
- (3) 図13右下の波型傾向をとるコンテンツは、序盤から中盤にかけての利用集中傾向は低く、か

表3 再利用傾向におけるコンテンツ3種

Table 3 Well-observed 3 contents in reusing patterns.

再利用傾向	上位3種		
一字型	FISH 関連	教育・知識	計算機
逆L字型	教育・知識	FISH 関連	計算機
波型	連絡先	FISH 関連	時刻表
ランダム	連絡先	時刻表	研究・学会

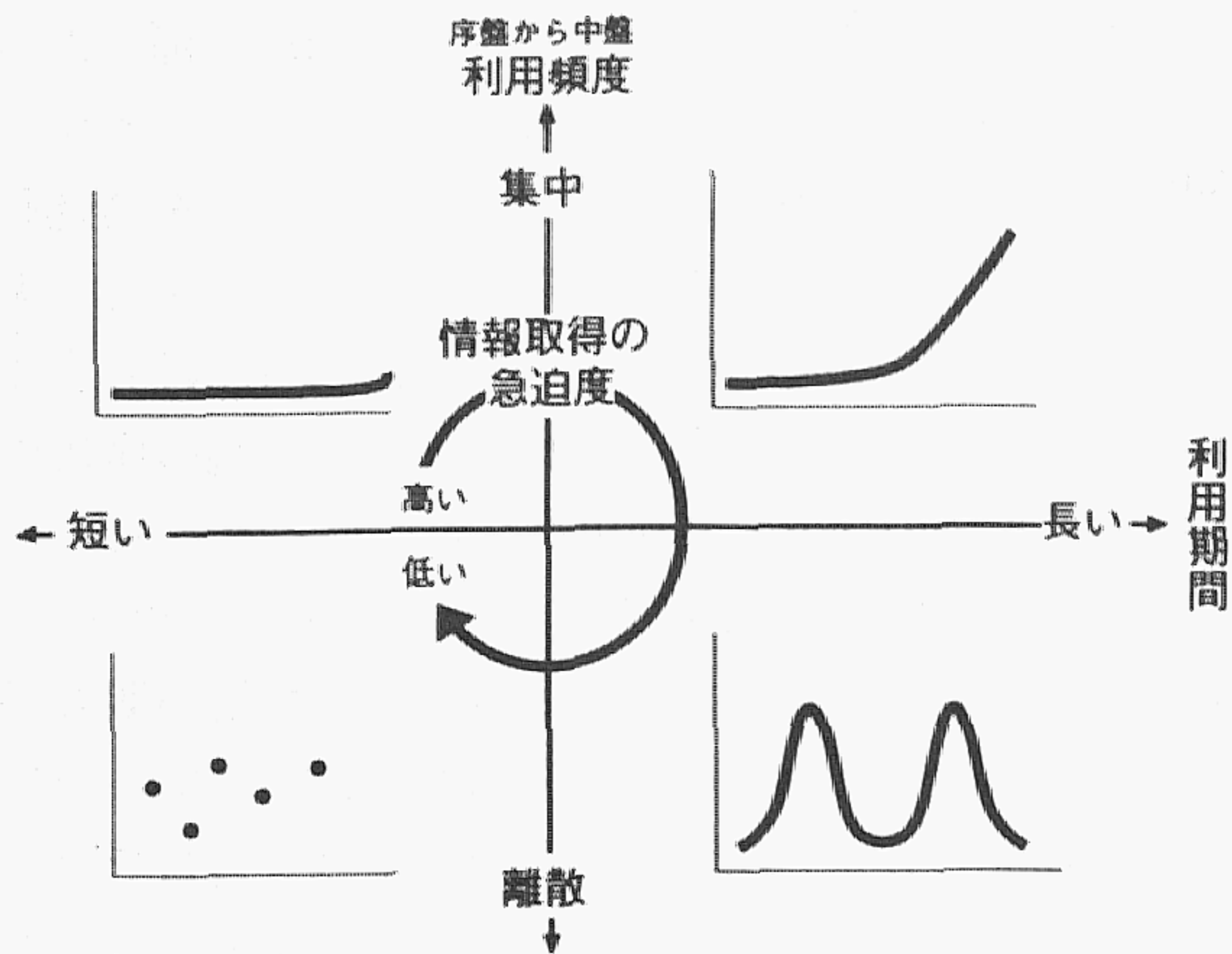


図13 利用推移モデル

Fig. 13 An information life cycle model.

つ利用期間が長いため、上記(2)よりさらに情報取得の急迫度が低いコンテンツにとらえる。

- (4) 図13左下のランダム傾向をとるコンテンツは、利用の集中傾向がなく、かつ利用期間が短いため、最も情報取得の急迫度が低いコンテンツである。

このモデルの目的は、利用推移の特徴をもとにコンテンツを4種類に明確に分けることではなく、各傾向の典型的な特徴をもとに、コンテンツがどの傾向に位置するのかによって、コンテンツの持つ急迫度の推測材料を提供することにある。よって、各傾向の中間傾向をとるデータでは、急迫度も中間の値をとるという考え方である。そして、推測された急迫度により、組織におけるコンテンツの保守管理のための材料を与えることである。急迫度をもとにコンテンツの有用/不用の基準をどこに定めるかは、その組織固有のコンテキストに依存するものである。

5. ディスカッション：再利用支援への適用

5.1 WWWを用いた組織内情報共有支援システムへの適用例

組織内情報環境（イントラネット）の1つに、組織内部やグループ内の情報共有のために、組織情報を収集して自動的に構成し、WWW上で見せるシステムが最近多く見られる。これらのシステムは、情報埋没化への対策として選別した情報の提供を行っているが、実際には本質的な解決策とはなっていない。本論文で提案したモデルをこの種のシステムに適用すれば、コンテンツの利用推移に適した蓄積情報の再利用支援のための情報提供が自動化できる。ここでは、蓄積した情報をWWWブラウザを介して再利用するシステムを例にとり、再利用の傾向に応じた蓄積情報の再利用

支援の方法を示す。

従来の問題点：組織内情報をWWWを用いて共有するシステムでは、情報洪水への対策として全文検索機能などを持っている。各々のページはおおむね階層構造であるが、ページ数が多く、たどるリンク数が多い。そのため、目的のページにたどり着きやすくするために、別のリンクが必要となる。従来、別リンクの主な方法は、全文検索と新着情報などのリンクページの2つである。検索機能では、語による全文検索に更新日などの属性を付加した検索を行う。これらはコンテンツが持つ特質と必ずしも一致しないため、検索結果には有用なページだけでなく、不用なページも多く含まれる。検索結果は各ページのスコアや信頼度によって整列されるが、これは語の出現頻度や出現場所を用いて計算されているものが多く、有用度とは必ずしも一致しない。また、更新情報から新着情報ページを自動生成することは可能であるが、新しいページが必ずしも有用とは限らず、また古くても有用なページも多いため、自動生成で有用な情報へのリンクページを作成するのは容易ではない。また、参照履歴を用いるものでも、単純に参照数から判断した人気ページでは、やはりコンテンツが持つ特質を無視しているため、有用度とは一致しない。このため、有用な情報の場所へ誘導するページを作成するのは容易でなかった。

支援システムの設計例：WWWサーバの参照履歴を用いる。週に一度、もしくは月に一度程度、定期的に過去の履歴を解析し、各コンテンツの再利用傾向を抽出する。長期的に利用されるページなどで、再利用過程が終了している利用者がいた場合には、いったん明確な判別ができれば再判別の必要性は低い。また、急迫度の高い一字型や逆L字型のページについては、各利用者が再利用過程の途中に位置しているかどうかを、夜間などを利用して毎日調査しておく。

〈支援機能1：急迫度による検索機能〉：語の全文検索によるスコア計算に各コンテンツの急迫度を加える。すなわち、語の出現頻度がある程度低くても、急迫度が高く、また検索利用者がそのコンテンツを利用していなかった場合には、利用を強く推奨できるページの可能性が高いため、高いスコアを付ける。モデルの急迫度をもとに、スコアに反映させる計算法は、その組織や業務の性格などのコンテキストや、対象システムが利用される頻度、などによって異なるため、その組織で組織固有の判断基準を定める。

〈支援機能2：有用リンク集ページの作成機能〉：リンクページを開いた時点で動的に作成することにより、各利用者に特化したページを作成する。一字型傾向の

ページの場合、その利用者が参照していなければ、参照を勧める情報として掲示する。参照期間の途中であれば、最も重要な情報として掲示する。

〈支援機能3：情報ページ修飾機能〉：コンテンツが記述された個々のページやメニューページには、新着マーク、人気マーク、もしくはアクセスカウンターなどの補助的な装飾が施されることが一般的である。参照履歴から求めた急迫度を利用することにより、有用なページに有用性を示すマークを付けることが可能になる。また、各利用者が再利用過程のどの段階に位置しているかによって、一字型や逆L字型の序盤、終盤、過程終了後などの区別が可能なマークを付加できる。波型傾向の場合では、ニーズが生じるたびに内容を再利用すると考えられるため、abc順などの目次にそったコンテンツの提示が考えられる。さらに、急迫度や、参照利用者の位置する段階に応じて、メニューページ内の順序や表示形式を変更することにより、急迫度の高いコンテンツへの到達の確実性を上げることができる。

以上、システムによって自動提供される情報は、利用履歴から客観的な視点で抽出された再利用傾向に基づいている。よって、支援方法による再利用過程の促進効果を確認するためには、利用者の再利用目的と支援内容がどの程度一致していたかを、インタビューなどで利用者の主観を調査し、直接的な評価を行う必要がある。それにより、本研究で提案した手法で推測される利用推移と、利用者の持つ再利用目的の推移との同一点/相違点を明確にすることが可能であると考ええる。

5.2 ディスカッション

以上のように、本論文で提案した方法は、情報共有の過程全般を支援するものではなく、再利用の過程に焦点を当てた支援により、情報共有促進を目指すものである。そして、支援の範疇に入る情報共有システムは、共有する情報をいったん蓄積しておき、利用者に再利用目的が発生した時点で、検索などの方法によって再利用目的の対象情報を利用者が探し出して使うシステムである。支援の対象となる情報は、組織情報としての再利用価値が長く、一過性でない情報である。また、本論文で分析対象となったコンテンツはテキストから構成されていた。テキストの場合は文字を通じて情報を利用するが、画像データなどを含むコンテンツの場合は、イメージを通じて情報を利用する、という情報の媒介方法が違うだけで、どちらの場合も取得の急迫度という概念の対象となるという点では、同一範疇に入る情報である。よって、画像データなどを含むコンテンツに対しても本モデルは適用可能であると

考える。

6. む す び

情報コンテンツの利用推移の把握は、研究面、ビジネス面から着目されている。本論文では、組織情報の利用推移を特徴付けるための指標の抽出と、その指標を包括する概念モデルの提案を目的に、ノウハウが実組織で長年再利用された履歴を、コンテンツの再利用過程の視点から分析した。分析の結果、組織情報の中でも再利用性の高いノウハウなどの情報では、その情報が必要とされる急迫度に応じて特徴的な再利用傾向があることが分かった。さらに、それらの傾向を抽出し、視覚的に表現した。そして、この分析結果をもとに、情報の利用推移を表現するモデルを提案した。本論文の手法により、多量のコンテンツを対象に情報の利用推移を自動的に抽出可能であり、抽出結果へのモデルの適用により、コンテンツの情報生命の推測材料を得ることができる。

本論文で提案したモデルは、分析事例の場合に特化したものではなく、集団の凝集性を持つメンバによって蓄積・共有されている組織情報に対しても適用可能である。そこで、情報共有促進の支援を目的に、コンテンツの利用推移に応じた蓄積情報の再利用支援について、モデルに基づく方法を示した。

今後の課題として、本研究で抽出した再利用の傾向と、利用者が一度取得したコンテンツがどのくらいの期間記憶されているのかという忘却曲線²³⁾との関係も考察する必要がある。これら再利用の傾向は、定性的な傾向であるが、統計解析における特定の分布による検定など定量的な傾向を調べ、また、各傾向の度合いを定式化する必要がある。さらに、マルチメディアコンテンツの場合においても分析・比較を行いつつ、5章で提案した支援方法を実現し、評価することによってモデルの妥当性を検証してゆく予定である。

謝辞 本研究で用いた分析データの収集は、NTTマルチメディアネットワーク研究所ネットワークング研究部のご協力により実現いたしました。この機会を与えてくださったNTTの木下研作氏（現在、NTTアドバンステクノロジー）に感謝いたします。また、有益なコメントをいただきましたNTTデータの山上俊彦氏（現在、株式会社アクセス）に心よりお礼申し上げます。

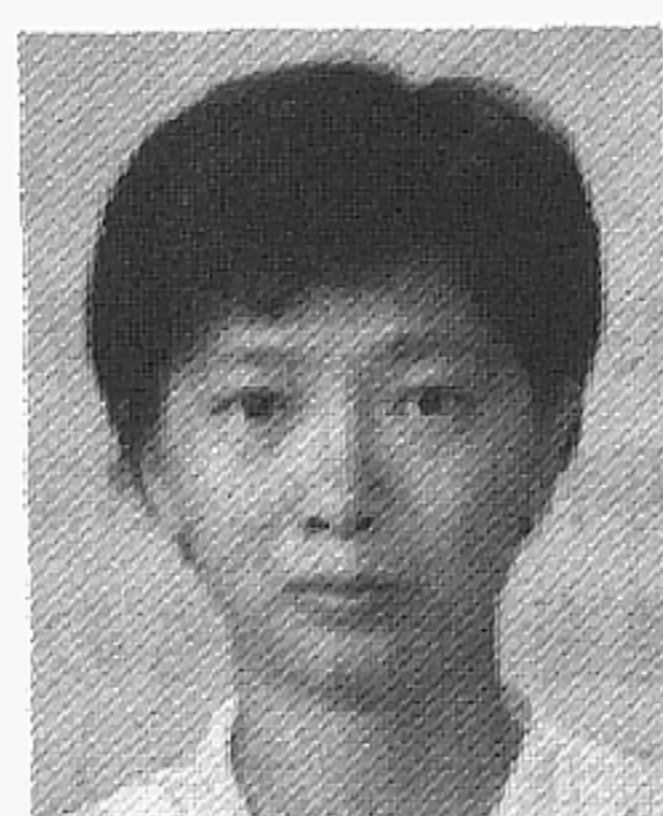
参 考 文 献

- 1) Ackerman, M.S. and Mandel, E.: Memory in the Small: An Application to Provide Task-

- Based Organizational Memory for a Scientific Community, *Proc. Hawaii International Conference of System Sciences*, Vol.IV, pp.323-332 (1995).
- 2) 野中郁次郎, 竹内弘高: 知識創造企業, 東洋経済新報社 (1996).
 - 3) Ackerman, M.S. and Malone, T.W.: Answer Garden: A Tool for Growing Organizational Memory, *Proc. ACM Conference on Office Information Systems*, pp.31-39, ACM (1990).
 - 4) Conklin, E.J.: Capturing Organizational Memory, *GroupWare '92*, pp.133-137, Morgan Kaufmann (1992).
 - 5) 門脇千恵, 爰川知宏, 山上俊彦, 杉田恵三, 國藤進: 情報取得ウェアネスによる組織情報の共有促進支援, *人工知能学会誌*, Vol.14, No.1, pp.111-121 (1999).
 - 6) 中山康子, 真鍋俊彦, 竹林洋一: 知識情報共有システム (Advice/Help on Demand) の開発と実践: 知識ベースとノウハウベースの構築, *情報処理学会論文誌*, Vol.39, No.5, pp.1186-1194 (1998).
 - 7) 村永哲郎, 守安 隆: グループワークのための情報共有技術, *情報処理*, Vol.34, No.8, pp.1006-1016 (1993).
 - 8) 爰川知宏: インフォーマル情報の共有支援に向けて, マルチメディア, 分散, 協調とモバイル (DiCoMo) ワークショッププロシーディング, pp.639-644 (1997).
 - 9) Resnick, P. and Varian, H.R.: Recommender Systems, *Comm. ACM*, Vol.40, No.3, pp.56-58 (1997).
 - 10) 森田昌宏, 速水治夫: 情報フィルタリングシステム, *情報処理*, Vol.37, No.8, pp.751-758 (1996).
 - 11) Balabanović, M. and Shoham, Y.: Fab: Content-Based, Collaborative Recommendation, *Comm. ACM*, Vol.40, No.3, pp.66-72 (1997).
 - 12) Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. and Riedl, J.: GroupLens: An Open Architecture for Collaborative Filtering of Netnews, *Proc. CSCW '94*, pp.175-186, ACM (1994).
 - 13) Terveen, L., Hill, W., Amento, B., McDonald, D. and Creter, J.: PHOAKS: A system for Sharing Recommendations, *Comm. ACM*, Vol.40, No.3, pp.59-62 (1997).
 - 14) 斎藤典明, 水澤純一, 山本平一, 山口 英: 話題の自動抽出による電子メールの情報組織化手法, *情報処理学会論文誌*, Vol.39, No.10, pp.2907-2913 (1998).
 - 15) ロジャーズ, E.M.: 普及学, 産能大学出版部 (1998).
 - 16) 後藤滋樹, 野島久雄: 人間社会の情報流通における三段構造の分析, *人工知能学会誌*, Vol.8, No.3, pp.348-356 (1993).
 - 17) 山上俊彦: 電子通信の導入と定着に関する長期観察結果の考察, *信学技報*, IN92-116, pp.49-54 (1993).
 - 18) 爰川知宏, 山上俊彦, 杉田恵三: グループ情報共有の長期的行動遷移モデルの提案, *情報処理学会研究会報告*, GW9-19, pp.105-110 (1995).
 - 19) 塚田晃司, 岡田謙一, 松下 温: 情報へのアクセスの形態に着目した情報共有モデルの一提案, *情報処理学会論文誌*, Vol.39, No.10, pp.2802-2810 (1998).
 - 20) Seki, Y., Yamakami, T. and Shimizu, A.: Flexible Information Sharing and Handling System: Towards Knowledge Propagation, *IEICE Trans. Commun.*, Vol.E 77-B, No.3, pp.404-410 (1994).
 - 21) 関 良明: 分散型ノウハウ蓄積システム Gold-FISHにおける分散環境への適応, *情報処理学会論文誌*, Vol.36, No.6, pp.1359-1366 (1995).
 - 22) 村杉 健: 作業組織の行動科学, *税務経理協会* (1987).
 - 23) Baddeley, A.: 記憶力, chapter 4, 誠信書房 (1988).

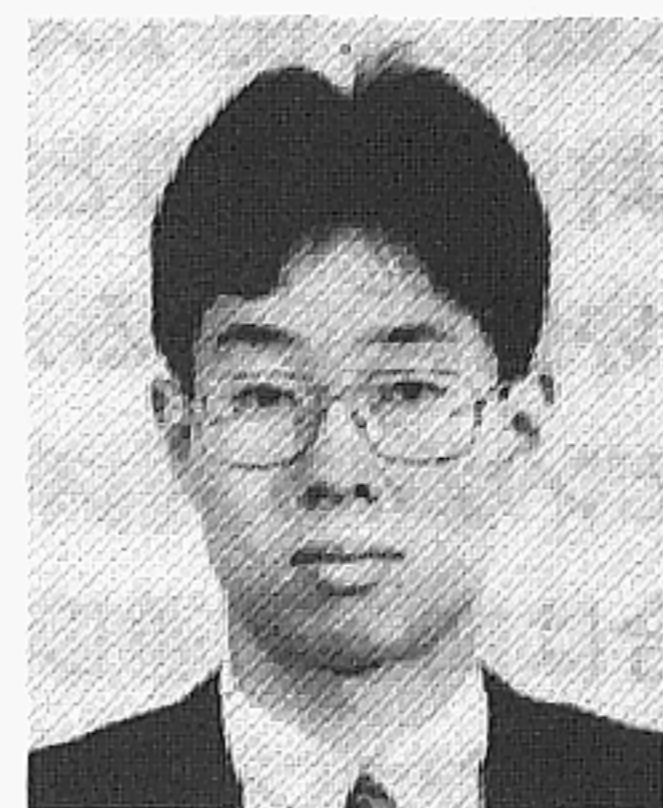
(平成11年3月23日受付)

(平成11年10月7日採録)



門脇 千恵 (正会員)

1998年北陸先端科学技術大学院大学情報科学研究科博士後期課程修了。博士(情報科学)。現在、同大学情報科学研究科助手。グループウェア, CSCW, 組織知識の共有に関する研究に従事。人工知能学会会員。



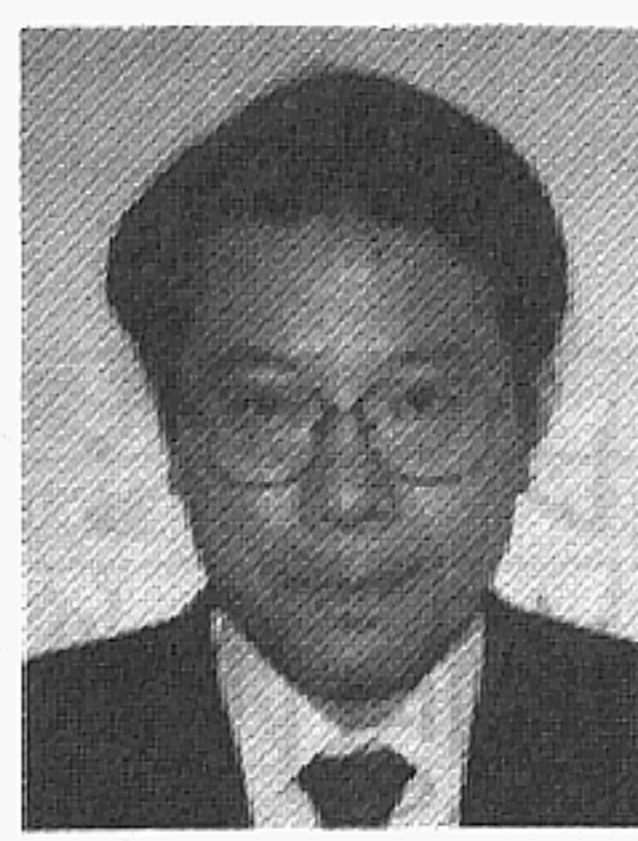
爰川 知宏 (正会員)

1991年大阪大学基礎工学部制御工学科卒業。1993年同大学大学院基礎工学研究科制御工学分野前期課程修了。同年日本電信電話(株)入社。グループウェア, 情報共有システム, 電子図書館に関する研究開発に従事。現在、西日本電信電話(株)研究開発センタ主査。IEEE会員。



杉田 恵三

1976年信州大学工学部電気工学科卒業。1978年同大学工学系大学院修士課程修了。同年日本電信電話公社入社。端末駆動系、情報共有システム、ネットワークインテグレーションに関する研究開発に従事。現在、NTTソフトウェア(株)経営企画部担当部長。電子情報通信学会会員。



國藤 進(正会員)

1974年東京工業大学大学院理工学研究科修士課程修了。同年、富士通(株)国際情報社会科学研究所入所。1982~1986年、ICOT 出向。1992年より北陸先端科学技術大学院大学情報科学研究科教授。1998年より知識科学研究科教授。情報処理学会創立25周年記念論文賞、人工知能学会1996年度研究奨励賞各受賞。工学博士。人工知能学会、計測自動制御学会、電子情報通信学会、日本創造学会等各会員。