

コンテンツに基づくグループ情報流通の解析

門脇 千恵[†] 爰川 知宏^{††} 杉田 恵三^{††} 國藤 進[†]

[†]北陸先端科学技術大学院大学 情報科学研究科
〒 923-12 石川県能美郡辰口町旭台 15

^{††}NTT 通信網研究所
〒 238-03 神奈川県横須賀市武 1-2356-523A

オフィスにおける効率の良い情報流通は、グループの協調作業を支える要因の一つであり、通信サービスの担うところは大きい。提供された通信サービスがその役割を十分に果たし、ユーザの利用動向の変化に対応しているか否かを分析するには、サービス上を流れるコンテンツを長期的に観察する必要がある。本稿ではノウハウ蓄積支援システム FISH を解析対象として取り上げた。このサービスの長期に渡るコンテンツを基に、情報の流通の実態を分析し、また情報の参照時間間隔という点で特徴的な3つのパターンを見出した。これは情報のライフサイクルという観点から、情報流通モデルの基礎を与えるものとして有効である。

Contents-based Analysis of Information Propagation in Group

Chie KADOWAKI[†], Tomohiro KOKOGAWA^{††}, Keizo SUGITA^{††},
and Susumu KUNIFUJI[†]

[†] Graduate School of Information Science, Japan Advanced Institute of Science and
Technology, Hokuriku
e-mail: {kadowaki, kuni}@jaist.ac.jp
15 Asahidai, Tatsunokuchi, Ishikawa, 923-12, Japan

^{††} NTT Telecommunication Networks Laboratories
e-mail: {koko, sugita}@nttmhs.ntt.jp
1-2356-523A, Take, Yokosuka, Kanagawa, 238-03, Japan

An efficient information propagation in offices, where the telecommunication service plays a significant role, is an important factor to support cooperative works. In order to analyze the efficiency of a telecommunication service and its adaptability to user's preferences, a long-term observation focusing on contents flux through the service is necessary. We performed such an analysis using the telecommunication system called FISH (Flexible Information Sharing and Handling system), and verified three characteristic patterns in time between information refer. This result is important for the formulation of information propagation model incorporating the concept of information life cycles.

1 はじめに

オフィス業務は、スキルやバックグラウンドなどの異なる様々な人間から構成されたいくつかのサブグループの協調作業である。これら協調作業が円滑に行なわれるためには、技術力の差を補完する知識や、個々が持つ情報の効率よい流通が重要となる。

近年、通信技術の発達に加え、同期・非同期を問わず種々のシステムの開発が行なわれ、情報流通の方式に関する研究は盛んである [1]。これら通信サービス上での知識の獲得 [2] や情報の流通の可能性およびその限界について考察するには、長期的なビジョンが必要である。時間の経過とともにユーザの利用動向は変化してゆき、サービス自体も種々の環境変化を受けやすくなるからである。よって、導入時には予想もしなかった情報の流れや滞留が起こる可能性もある。長期的にこれらの変動を分析し、利用動向の特徴を明確に捉えることは、息の長い通信サービスへの第一歩となる。

この長期的観察の観点から行なわれた研究に、集団の性質や導入されるサービスの性質に基づいて提案された「定着段階モデル」[3] や、「サービスからの脱落」を考察したもの [4] がある。しかし、長期にわたる膨大なコンテンツを客観的に解析し、体系だった考察を与えた研究は行なわれていない。情報流通の実態をさらに詳細に分析するためには、サービス間の相互依存も踏まえたグローバルな視点に立つ上記の知見を骨組みとし、その中身に当たる情報の粒子、つまりサービスで生じるコンテンツまで見究める必要がある。なぜならば、利用動向の微妙な変化はコンテンツに反映されるからである。

本研究では通信サービスの長期に渡る利用動向を、サービスのコンテンツという切口を用いて解析し、グループ情報流通の実態を明らかにすることが目的である。『コンテンツ』とは通信サービス上を流れる内容と、その発信源、受信先であるユーザを指す。

以下、2 節では解析の対象としたグループ、通信サービス、データと仮説の説明を行なう。3 節ではコンテンツを内容もしくはユーザという視点から基礎的に解析し、キーワードに基づいたカード分類とユーザの利用動向について述べる。そして、これらコンテンツに関する仮説の検証を行なう。4 節ではあるコンテンツ（内容）がグループ情報の中でいかに活用されているか、またコンテンツ（ユーザ）をアクティブユーザという視点からとらえた際の流通に与える影響を考察する。最後 5 節にまとめを述べる。

2 解析方法

2.1 解析対象

対象とするグループは人数が 20 数名であり、メンバーの転入出によりグループの構成が数回変わっている。このグループでは 6 年以上前から情報共有／流通の支援のため、電子メールをはじめ、様々な通信サービスを採用、あるいは提案してきた。この中で FISH[5] というシステムに着目して解析を行った。FISH はノウハウ情報の蓄積と共有を目的に、このグループ内で独自に開発され、1991 年 1 月に導入された。その後、1994 年 4 月に分散環境への適応性を考慮した代替の新サービス GoldFISH[6] に移行されたが、現在でもまだ少数のユーザが根強く使用している。

FISH はユーザが自由文入力したノウハウカードにキーワードを付け、このキーワードを用いてカード間のリンクを張って管理することができ、カードは多方向からの検索が可能である。長期的な情報共有を目指して蓄積された情報の内容は豊富であり、今回の解析目的に適合すると判断した。

通信サービスの利用動向を調べるには、アンケートやインタビューといったユーザへの問い合わせと、ログなどサービスの利用履歴を解析する方法がある。前者はユーザ個々のニーズを細かく掴むことはできるが主観的評価に陥りやすく、またユーザの記憶という面から全てのコンテンツ（内容）を拾い出すことは困難である。それに較べて後者は、コンテンツ（内容）の収集や統計的处理が簡単であり、客観的評価を行なう上で有用である。そこで、今回はログ解析という方法を採用した。

対象グループではサービスの長期観察を目的に、複数サービスにわたる数種類のログを取っているが、FISH では検索の成功／失敗履歴と参照されたカードの履歴が残されている。参照履歴のデータは、導入当初から残っており（約 1 万 4 千件）、これと蓄積されたカードが解析主対象である。導入日から約 1 ヶ月間は運営テストなどが頻繁に行なわれているため試用期間として解析対象から除外した。

2.2 カード参照間隔に関する仮説

テストデータの確認を兼ね参照履歴データを概観した際、あるユーザが同一カードを数十分、数時間、あるいは数日後に再び参照するケースが多く認められた。そこで、以下の仮説を立てた。

仮説 1: あるコンテンツ（内容）のカード参照間隔には特徴があり、これは知識獲得の度合に関係する。

仮説 2: コンテンツ（ユーザ）とカード参照間隔には関連がある。

たとえば、技術知識に関するカードでは、最初は頻繁に

参照が起りその間隔は短い、獲得度が増すにつれて参照間隔が長くなるという仮説である。解析ではこれらの検証も含めた情報流通の実態を分析した。

3 コンテンツからのサービスの実態

3.1 カードの分類

まず、どのような情報が流通されたのか明確にするために、キーワードを基にカードの分類を行なった。方法としては最初に表1の項1-14に分け、登録カード数の多い項1-6に関しては、さらに細分した。たとえば、項1の教育・知識はノウハウ、新人、質問への回答の3分類に細分している。これら分類とそれに該当するカード数が複数付いているカードについては重複して計上されている。たとえば、'freshman,denpyou'の2つのキーワードが付いているカードでは、表1の項1中の新人と、項4中の業務(伝票関係)の2分類に数えられている。表1の最終行はその年に登録されたカード数である。

項	分類	細分数	91年	92年	93年
1	教育・知識	3	86	24	13
2	研究・学会	2	74	42	33
3	計算機環境	5	291	121	50
4	業務関連	8	97	84	45
5	電話・住所	2	52	33	24
6	FISH 関連	2	61	10	0
7	アイデア	-	7	2	1
8	登録テスト	-	10	14	1
9	個人プラン	-	12	6	0
10	福利厚生	-	21	3	5
11	趣味娯楽	-	15	5	5
12	一般教養	-	3	4	15
13	一般情報	-	8	5	7
14	時刻表	-	6	5	1
登録カード数			633	288	167

表1: 分類と該当カード枚数の推移

観測1: ノウハウとは言い難い分類(項5、項8-11、項14)のカードがある。

推測: 観測1より生活や業務に密着した利用頻度の高い情報が、そのニーズから本来のサービス目的以外で登録されると考えられる。

観測2: FISH 関連の登録カード数は91年には多いが、93年には0になっている。なお、これらカードが参照された件数についても、291件、113件、43件(年順)と推移している。

推測: 観測2は導入年にはFISHの利用方法の情報交換が頻繁で、またそのサービスの利用調査が行なわれたため件数が多かったものが、サービスの安定とともに登録数、参照件数ともに減少したと考えられる。

分類単位で見たカードの参照件数についても調べた(ユーザ区別なし)。上位3位は、91年計算機環境1570件、電話・住所1182件、教育・知識725件、92年計算機環境798件、業務関連724件、電話・住所700件、93年計算機環境584件、業務関連584件、電話・住所416件であった。

観測3: 毎年、計算機環境の参照件数が一番多く、電話・住所と業務関連(91年は4位)も常に上位を占めている。

次に、カード1枚づつの単位で見た参照件数も調べた(ユーザ区別なし)。

観測4: 参照件数が一番多かったのは時刻表Aで523件(全参照件数の4.3%)、2位が時刻表Bで218件、以下9位まで全て電話・連絡先、10位が電話と業務(勤務関係)であった。

推測: 観測4よりノウハウ情報以外の利用も盛んであったと言える。しかし、時刻表などは長期間記憶したい対象となるとは考えにくく、参照件数に反映する可能性がある。反対に長期記憶を望むような情報では数回の参照で、獲得が十分となるかも知れない。よって、参照件数の多さだけではFISHサービスの評価は難しい。

3.2 ユーザの利用動向

FISHではユーザはカードの登録/検索/更新/参照/削除を行なえるが、表2はカードを1度でも登録もしくは参照したユーザ数の推移である。表2を見ると積

処理内容	91年	92年	93年
登録	12	15	13
参照	24	25	21

表2: カード登録/参照ユーザ数推移

極的にカードを登録するユーザ数は、参照ユーザ数全体の約6割程度である。参照ユーザの4割は参照しかせず、その参照件数が全体参照件数に占める割合は9.0、9.1、12.2%(年順)にしかならない。さらに参照のみのユーザのうち、3年を通じてカードの登録は殆んど行なわず、しかも参照件数が非常に少ないユーザが数人いることが分かった。

また、ユーザが登録したカードと参照件数の関係についても考察した。ここで、カード登録者と参照者が同一ユーザの場合を自己カード参照と呼ぶ。

観測5: 全参照件数が多いユーザは、自己カード参照率が高い。

3.3 コンテンツと参照周期

次に仮説1および2を確かめるため、カード参照間隔についてユーザ毎に調査した。一ユーザが参照したカードをひとまとめに計算した場合と、選抜したカード一枚

について計算した場合とがある。

観測 6: 以下のような3つの参照間隔パターンが認められた。これらの混合パターンをとるものや、全く参照間隔がランダムでパターンを有しないものも少数あった。なお、グラフのX軸は参照間隔の順番で、Y軸は参照間隔時間（最小単位は1時間）である（たとえば、X軸上における最初のデータは、第1回目が参照されてから、次に2回目が参照されるまでの時間間隔を表している）。

- **逆L字型:** 参照は初回からコンスタントかつ短間隔で、時間の経過とともに参照間隔が開いてゆく度合が強いもの（図1）。

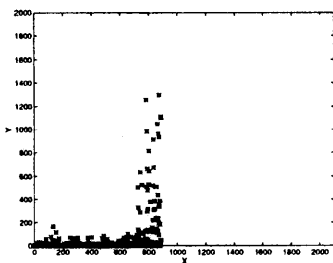


図1: 参照間隔パターン<逆L字型>

- **山型:** 序盤ではほぼ短間隔で参照し、その後一旦参照間隔が開いてゆき、また間隔が短くなる繰り返しである（図2）。

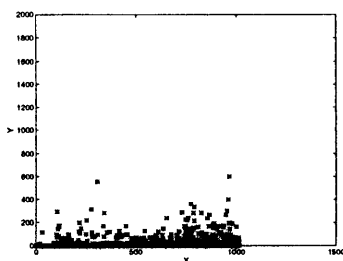


図2: 参照間隔パターン<山型>

- **一字型:** 初回（カード登録直後）から、参照間隔が殆んどコンスタントでかつ短く、グラフ上ではほぼ一直線になるもの（図3）。

これらの参照間隔パターンとコンテンツについて分かったことは以下のとおりである。

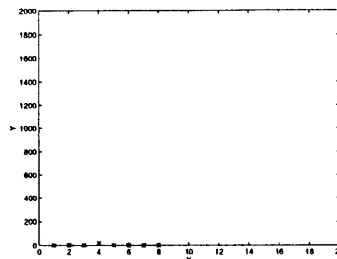


図3: 参照間隔パターン<一字型>

観測 7: 教育・知識関係のカードは一字型もしくはゆるやかな逆L字型をとる傾向がある。カード参照期間が短期間のものほど一字型になりやすい。反対に電話・連絡先や時刻表では一字型をとることはなく、山型か逆L字型をとる傾向がある（ユーザもしくはカード内容によりどちらかの型に変わる）。

観測 8: 自己カード参照率が高いユーザは参照間隔パターンは一字型に近付く傾向があり、逆に低いユーザは山型になり、さらに低いとパターンは認められなくなる（図4）。

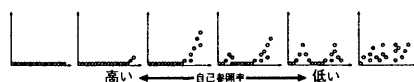


図4: 参照間隔パターンの変遷

推測: 観測5と8より、自己参照率の高いユーザは全参照件数も多く、かつカードも登録直後（まだ登録情報の記憶が確かな時期である）から利用していることより、検索したいカードのキーワード縦りがはっきりしており検索が容易であると考えられる。また、山型は自分の知りたい内容があるのかどうか、どういうキーワードが付けられているのかを知る手段に左右され参照間隔に波があると推測する。

観測7は仮説1、観測8は仮説2をおのおの裏付けるものである。

4 コンテンツが流通に及ぼす影響

本節では特徴的コンテンツ（内容）の流通例（4.1節）と、コンテンツ（ユーザ）の流通に及ぼす影響を考察する（4.2,4.3節）。

4.1 技術習熟を目的とする情報の流れ

このグループの利用動向の特徴として、新人社員の検索支援のために'freshman'というキーワード（表1項1中の分類）を付加していることが挙げられる。新入社員が業務ノウハウや計算機環境の情報から、さしあたって重要なものを効率良く参照できるように付けられていると考えられる。'freshman'のカードは、カード内容を端的に表す別のキーワードで検索できるようになっている。このキーワードの分類を調べてみたところ、業務（伝票関係）252枚が一番多く、次いで別キーワードなし183枚、学会関係76枚の順位であり、そのあとは計算機環境や業務関連の分類が多かった。

観測9:'freshman'カードの別キーワードには業務（伝票関係）が特に多い。

推測:観測9は伝票の切り方など業務ノウハウの伝達を意図した結果と思われる。

観測10:'freshman'カード登録者を見ると新人社員の上司にあたる人物が最も多く、新人社員の多くはこのユーザが登録した'freshman'カードをよく参照している。

推測:観測10より技術習熟度を上げるための上司から新入社員への情報の流れができていたと推測できる。

観測11:図5はある社員が入社後、'freshman'カード（カード区別なし）を参照した間隔をグラフ化したものであり、年数を経るに従い参照間隔が開く傾向が認められる。

観測12:図6はさらに上記からあるカードのみを取り出して参照間隔を求めたものである。この場合、入社後から参照し始め、約36日間の間に全部で9回参照し、その後は一切参照していない。内容は「電子メールの使い方」であった。

推測:観測11,12より、情報を獲得するまでは最初頻繁に参照が行なわれ、獲得度合が進むにつれて参照間隔が開き、獲得完了で参照が終了したと考えられる。これも仮説1を裏付けるものである。

4.2 アクティブユーザの流通への影響

3.2節ではユーザによってかなり利用状況に差があることが分かった。そこで、本節ではアクティブなユーザがグループ情報の流通に与える影響について考察する。

今回アクティブユーザを参照件数の多さ（表3：項目1,2,3）と、カード登録数の多さ（表3：項目4,5,6）という2つの視点から考察した。表3の1と4の『登録』の行は、各年にアクティブユーザ4人が登録したカード枚数の合計が、その年の登録カード全体に占める割合(%)である。表3の2と5の『参照』の行は、各年にアクティブユーザ4人が参照した件数の合計が、その年の参照件数全体に占める割合(%)である。ここで<登録⇒参照>の表記は、カードの登録者と参照者が限定された

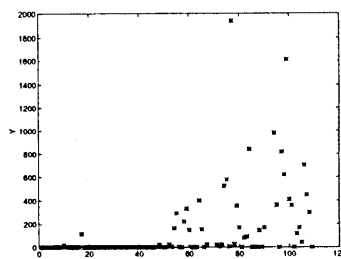


図5: freshman 全カードの参照間隔 (luser)

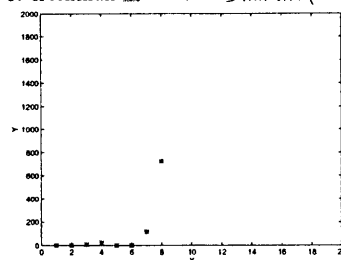


図6: freshman 1カードの参照間隔 (luser)

ユーザ間での情報流通（登録と参照者が同一も可）を表す。表3の3と6の『4者間』の行は、各年にアクティブユーザ4人間だけで発生した情報の流通、つまりアクティブユーザ4人に限った場合の<登録⇒参照>の件数がその年の参照件数全体に占める割合(%)である。

	91年	92年	93年
1. 登録	70.6	52.1	70.7
2. 参照	62.1	64.6	62.9
3. 4者間	48.6	42.9	36.8
4. 登録	78.8	65.6	80.8
5. 参照	61.1	38.6	48.2
6. 4者間	54.9	30.4	32.7

表3: アクティブユーザの影響度

観測13:アクティブユーザ4者間だけの<登録⇒参照>は、2視点とも3割を超えている。

推測:観測13よりアクティブユーザがサービスに与える影響が非常に大きいと言える。

さらに、表4は各年のアクティブユーザの顔ぶれの変遷を整理したものである。表4の見方はたとえば、92年の第4位の箇所では登録4位はuserA、参照4位はuserLであるという意味である。表4を見ると、userAとuserBは、3年間を通じて登録、参照ともに上位4位に常に入っていることが分かる。特にuserBは、91年、

	91年	92年	93年
1位	userB userB	userB userB	userB userA
2位	userA userC	userH userD	userC userS
3位	userK userA	userC userA	userA userB
4位	userC userD	userA userL	userD userD

※各要素の左側は登録順位、右側は参照順位である

表4: アクティブユーザの顔ぶれ

92年ともに登録枚数と参照件数が一番多い。

観測 14: アクティブユーザの中には常にカードの登録、参照件数が多い非常に活発なユーザがいる。

4.3 本人から本人への情報の流れ

全参照件数が多いユーザは自己カード参照率が高いことを3.2節で述べたが、この自己カード参照率の高いユーザがグループ情報の流れに与える影響について分析した。グループ内で発生した<登録⇒参照>のユーザの組合せ、つまり発生した情報の流れ路は389組あった。そのうち、参照件数の多い上位50組み中の1,2,4,9,11,14,24,34,38,46,47位は、ユーザ本人から本人への情報の流れである(表5)。1位の参照件数だけでグループ全体の参照件数の約1割を占める。ユーザ名が表4と一致するものは同一人物である。

観測 15: アクティブユーザ本人から本人の流れが、グループ情報の流通の上位を占める。

順位	参照件数	参照者	登録者
1	1210	userB <--	userB
2	799	userA <--	userA
4	306	userC <--	userC
9	174	userL <--	userL
11	160	userK <--	userK
14	145	userD <--	userD
24	87	userS <--	userS
34	67	userR <--	userR
38	64	userE <--	userE
46	51	userH <--	userH
47	50	userF <--	userF

表5: 発生した情報の流れ(1-389位)からの抜粋

5 おわりに

本研究では、グループにおける情報の蓄積と共有を目指したサービスの利用実態の解析から、情報の参照間隔パターンを抽出した。この参照間隔パターンは、『参照価値の推移』という情報のライフサイクルの観点から一般に適用可能なものである。また、得られた観測事項は、このグループ特有の性質に依存したものととは考えにくく、特に情報登録/参照に関して非アクティブ/アクティブ/さらにアクティブなユーザの3層は「利用者間

の教えあいの構造」[7]の典型例と言える。さらに、ある特定のキーワードに参照が集中するなどの利用者行動も考えあわせると、従来のようなサービスモデルでは、上記のような粒度の細かい実態まで分析するには不十分であり、木目の細かいサービスの支援を目指すにはコンテンツに基づく必要がある。

今回の解析により、サービス全体の利用動向はアクティブなユーザ群に左右されていることが分かった。今後、FISHのサービス評価をさらに正確に行なうためには、コンテンツの内容やユーザの違いを情報の参照間隔に反映したモデルを確立する必要がある。そのためにも、一字型においては知識獲得までの参照回数の個人差を考慮し、山型パターンではカード検索の成功/失敗体験が影響すると考えられることから、キーワード情報との相関についても検討する必要がある。逆し字型については、獲得した記憶の忘却特性が反映されていると考えられ、忘却曲線を考慮した検討も今後必要である。

謝辞

本研究は1995年3月にNTT通信研究所ネットワークインテグレーション研究部において、学外実習プログラムの一環として行なわれました。本研究の機会を与えて下さったNTT通信研究所の木下研作研究部長、中田寿主幹研究員に感謝致します。また、丁寧なご指導と貴重なご助言を頂きました山上俊彦主任研究員に感謝致します。

参考文献

- [1] Ishii, H., Kobayashi, M., and Arita K.: 'Iterative Design of Seamless Collaboration Media', CACM, pp.83-97 1994.
- [2] 國藤 進: '知識獲得と学習研究の新しい流れ', 人工知能学会誌, Vol.3 No.6 pp.741-747 1988.
- [3] 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.
- [4] 愛川 知宏, 山上 俊彦, 杉田 忠三: '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.
- [5] Seki, Y., Yamakami, T., and Shimizu, A.: 'Flexible Information Sharing and Handling system: Towards Knowledge Propagation', IEICE Trans. Commun., March 1994.
- [6] 関 良明, 藤木 直人: '分散型ノウハウ蓄積システム GoldFISHの検討', 情報処理学会研究会報告, GW3-3, pp.17-24 1993.
- [7] 野島 久雄: 'コンピュータ利用場面における他者の役割', 人工知能学会研究会資料, SIG-F/H/K-9001-9(12/6) pp.77-86 1990.