

投稿型サイトにおける影響伝播を利用したコンテンツのクオリティを考慮したランキング手法

金子 鷹弥† 牛尼剛聡‡

† 九州大学芸術工学部 〒815-8504 福岡県福岡市南区塩原 4-9-1

‡ 九州大学芸術工学研究院 〒815-8504 福岡県福岡市南区塩原 4-9-1

E-mail: † takaya@gospel.aid.design.kyushu-u.ac.jp, ‡ ushiama@ design.kyushu-u.ac.jp

あらまし 近年、コンテンツのデジタル化が進み、Web 上に存在するコンテンツは増加の一途を辿っている。一般のユーザが投稿者となる各種投稿型サイトにおいては、視聴に値しないクオリティの低いコンテンツも多数存在する。こうした中で、投稿型サイトに投稿された多数のコンテンツの中から、クオリティの高いコンテンツを発見することが重要である。本研究では、投稿型サイトにおいて、クオリティの高いコンテンツは他の新しい創作活動を誘発するという点に注目し、クオリティに基づくランキング手法を提案する。あるコンテンツに基づいて他のコンテンツが制作されたことを影響伝播と名付け、影響を受けたコンテンツに対する評価を、影響を与えたコンテンツに伝播させる。コンテンツに対する評価として既存のランキングを利用することで、クオリティを考慮したランキングを実現することを目指す。

キーワード コンテンツ管理、再ランキング、影響伝播、投稿型ウェブサイト

A Ranking Algorithm for Shared Contents Based on Propagation of Creative Influence

Takaya KANEKO† Taketoshi USHIAMA‡

† Department of Design, Kyushu University 4-9-1 Shiobaru, Minami-ku, Fukuoka, 815-8540, Japan

‡ Faculty of Design, Kyushu University 4-9-1 Shiobaru, Minami-ku, Fukuoka, 815-8540, Japan

E-mail: † takaya@gospel.aid.design.kyushu-u.ac.jp, ‡ ushiama@ design.kyushu-u.ac.jp

Abstract Recently, the number of web contents on the web are increasing with the attendant on the growth of digital content technology. On content sharing websites, everyone is allowed to post his/her content, so many low quality contents, which are not worth watching, may exist. It is important to discover high quality contents on content sharing sites. In this paper, we focus on the assumption that high quality contents give some creative influence on other contents, and propose a ranking algorithm based on their quality. We named the relationship that a new content is created under the influence of another content as “creative influence propagation”. Our algorithm propagate the evaluate value of a content to the content which have given creative influence on it our purpose is to find contents which have high quality but unpopular.

Keyword digital content management, re-ranking, creative influence propagation, content sharing website

1. はじめに

現在、Web 上には大量のデジタルコンテンツが存在し、それらを効率的かつ効果的に利用可能とすることが重要になっている。一口に Web 上のデジタルコンテンツと言っても、その種類は様々である。Web 上のデジタルコンテンツが飛躍的に増加している背景の一つに投稿型サイトの存在が挙げられる。投稿型サイトは、デジタルコンテンツを投稿し閲覧するためのサービスを提供するが、そこには 2 種類のユーザが存在する。一つは閲覧者としてのユーザであり、もう一つは投稿者としてのユーザである。これらの役割は固定的ではなく、閲覧者が投稿者となることもある。投稿型サイ

トに存在する多くのコンテンツは、既製のコンテンツ（プロの写真家やアーティストの作品）に比べ、クオリティが低い物が多く存在する一方で、閲覧者に高い満足を与えるクオリティの高い作品も数多く存在する。

投稿型サイトに投稿されるコンテンツの種類は幅広く、画像や動画を始めとして、現在では小説や詩の投稿サイトが運営されている。表 1 に代表的な投稿サイトの例を示す。多くの投稿型サイトの中で、ブロードバンドインターネット接続が整備され、爆発的に普及したのが動画共有サイトの YouTube である。日本国内において、最大の利用者数を持つのは YouTube であり、利用者は 2009 年 10 月の時点で、2248 万人である。

一方、ニコニコ動画のユーザは 800 万人を上回っている。ニコニコ動画は日本国内でもっとも著名な動画投稿サイトであり、現在では約 350 万本の動画コンテンツが投稿されている。

ニコニコ動画の特徴的な機能として、動画上にコメントを重畳表示できることがあり、閲覧者は動画再生中の任意の時刻、任意の場所にコメントを付与することができる。この機能により、閲覧者にとっては、同一の動画を他の閲覧者と同時に視聴しているような感覚を味わうことができる。一方、投稿者にとっては、視聴者が動画のどの部分に興味を持つのかを知ることができる。このことによって、ニコニコ動画には、単に動画を共有する他の投稿サイトとは異なり、創作物が多数投稿されるようになったと考えられる。

表 1 投稿型サイトの例

コンテンツの種類	主な Web サイト名
画像	pixiv(イラスト),Flickr(写真)
動画	YouTube, veoh, ニコニコ動画,
小説	Novelist.jp, 小説家になろう
音楽	MySound, NEXTMUSIC
漫画	ACOMI

一方で、投稿型サイトの利用者の増加に伴い投稿されるコンテンツが増大するにつれ、クオリティの高いコンテンツの発見が困難になっている。莫大なコンテンツが存在する中で、ユーザは検索やランキングといった方法で、満足感の得られるコンテンツを探さなければならない。現在の検索システムでは、ユーザがあらかじめ、自分の要求を満足するコンテンツが明確に想定できている場合は、ユーザはクエリとして対象となるコンテンツの特徴を表すキーワードを入力し、希望のコンテンツを発見できることが多い。しかし、投稿型サイトにおいては、毎日莫大な量のコンテンツが投稿され、ユーザは未知のコンテンツに対して適切なクエリを発することが困難である。こうしたことから、投稿型サイトでコンテンツを見つけるためには、ランキングが重要な役割を果たしている。

従来から、ランキングは未知のコンテンツを発見する際に利用されてきた。Web ページをはじめとする各種検索システムにおいても、検索結果を表示する際に、より適切なものを上位に配置するランキングが利用されている。通常、ランキングの基準は、閲覧数やコメント数といった一般的な普及度を表している基準が利用されることが多い。しかし、コンテンツのクオリティを考慮する際、普及しているもののクオリティが高いとは限らない。

投稿型サイトの特徴として、創作物の発表の場として利用されることが多いことが挙げられる。これは、

Web 上で自作のコンテンツを投稿し、閲覧者からの評価を得ることができることが理由である。例えば、Flickr には、ユーザが制作した絵画や、撮影した写真等が投稿され、ニコニコ動画には個人で制作したアニメーションや、個人の演奏等が動画として投稿されている。また、そういった創作物の投稿が増加するにつれ、ユーザが他の投稿者と協力してコンテンツを作る傾向がある。楽曲を制作するユーザが投稿したコンテンツを引用し、別のユーザがプロモーションビデオを制作するといった引用も数多く行われ、特にニコニコ動画は、こういった創作物の発表の場として利用されることが多く、投稿者の創作物と判断できるコンテンツが多い。

我々は、個人の創作物が投稿される投稿型サイトにおいて、コンテンツのクオリティがどのように評価されているかに注目する。大量に存在する投稿コンテンツに対して、一人のユーザが全てのコンテンツを閲覧して、クオリティを評価することは難しい。本研究ではコンテンツの引用関係に注目し、引用数の多いコンテンツや、他からの評価の高いコンテンツから引用されているコンテンツはクオリティが高いと言えるのではないかと考えた。

上記の考察に基づき、本研究では、あるコンテンツが他のコンテンツの制作に影響を与える関係を影響伝播関係と名付け、影響伝播関係を利用することにより、コンテンツのクオリティを考慮したランキングを決定する手法を提案する。対象はニコニコ動画とし、アクセス数に基づいて決定されたランキング情報を利用して、影響伝播関係を利用してクオリティを考慮したランキングを導出する。

2. コンテンツのクオリティと発見手法

2.1. コンテンツのクオリティ

本研究におけるクオリティとは何かを述べる。従来のランキングでは、ランキングの基準として閲覧数や販売数といった直接的な評価基準が用いられることが多かった。これらの評価基準は「普及度」ということができ、普及の程度を評価の高さと考えている。しかし、普及度の高さがクオリティの高さであるとは一概には言えない。なぜなら、普及してはいないがクオリティが高いコンテンツが存在するにも課から割らず、普及度に基づくランキングでは、それらを発見することができないからである。図 1 にコンテンツに対する評価の関係の概念図を示す。

本研究では、クオリティが高いコンテンツには、ユーザがコンテンツを閲覧または視聴した際に、他のユーザの新しい創作活動を誘発するものが多い、という点に注目する。これは閲覧数等とは違い、直接的に集

計することは困難であるが、本研究ではユーザの振る舞いから抽出を試みる。

2.2. 検索とランキング

コンテンツを発見する手法としてよく使われる手法は「検索」と「ランキング」である。「検索」と「ランキング」の大きな違いは「クエリ」の有無である。「検索」では、見つけたいコンテンツに対しての明確なイメージをキーワードとして表現できれば、それをクエリとしてシステムに与えることで、ユーザの要求を満足できる結果が得られる。しかし、対象に対するイメージを言語化できない場合や、希望するコンテンツがあるわけではなく、広くコンテンツを見渡したい場合には、「ランキング」が有効である。「ランキング」は「クエリ」を指定する必要がなく、なんらかの基準をもとに作成されているため、対象を具体的にイメージできない場合でも有効である。特に、投稿型サイトにおいては、毎日大量のコンテンツが投稿される中で、ユーザはどういったコンテンツが投稿されたかを把握することも困難である。そういった場合は検索ではなく、新規に投稿されたコンテンツをランキングとして表示することで、ユーザは新規のコンテンツを視聴するかどうかに関して有用な情報を取得可能である。

現在では、コンテンツ量の増加量が多いサイトにはランキング機能が付いていることが多い。ランキングは Web に限らず、様々な場面で用いられている。例えば、楽曲の販売枚数に基づいてオリコン社が提供するランキングは楽曲ランキングの中では特に重要視されており、ユーザのランキングへの信頼が高いと考えられる。

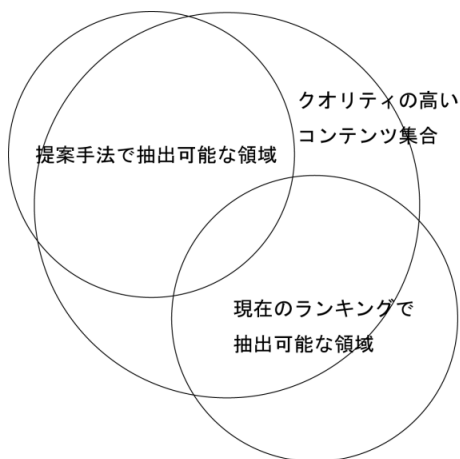


図1 コンテンツに対する評価の関係

2.3. ランキングの問題点

2.1 で述べたように、従来のランキングでは、普及してはいないがクオリティが高いコンテンツを発見することができない。しかし、特に投稿型サイトにおいては、上位にランキングされたことがないコンテンツの中にも、クオリティの高いコンテンツが数多く存在

すると考えられる。それらの中には、1度もランキング上位になることがなく、本来受けるべき正当な評価を受けることができないまま埋もれてしまうことも多い。投稿型サイトのように、対象とするコンテンツが多い場合には、コンテンツを探しているユーザもランキングをメインに使用する傾向があり、上位にランキングされないコンテンツをユーザが発見することは困難であると考えられる。

また、今回クオリティの基準とした主観的な評価は集計することが難しく、ランキングの基準とすることはできない。そこで、本研究では影響伝播という考え方を導入し、影響伝播を考慮した重みづけ手法を提案する。

3. アプローチ

3.1. コンテンツ間の影響

本研究では、コンテンツが他の制作に与える影響に注目する。例えば、子ども時代にプロ野球の中継を良く見ていた少年は、将来の夢としてプロ野球選手と答えることがある。それは少年が見ていた選手、もしくは球団が優秀であったからだ。そしてプロ野球選手を夢みる少年達の中から、夢を達成する少年が存在し、彼らを見た少年達もまたプロ野球選手を志す。これを投稿型のコンテンツに当てはめて考えると、ユーザは優秀なコンテンツを視聴し、新しいコンテンツを投稿する。そして新しいコンテンツを視聴した別のユーザは、そのコンテンツからの影響を受けた、更に新しいコンテンツを投稿する。このように、投稿型サイトではクオリティの高いコンテンツから派生したコンテンツの流れを構成する(図2)。この関連を用いて、コンテンツのクオリティを評価することを考える。

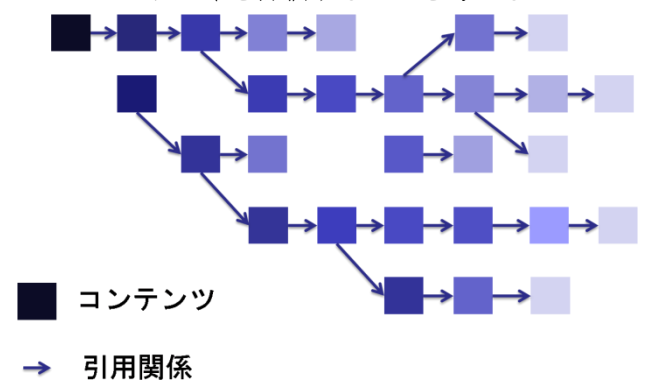


図2 コンテンツの派生図

3.2. 影響伝播関係

影響伝播関係とは、あるコンテンツに対してそのコンテンツの内容に影響を受けたコンテンツが生まれる、という関係を示している。

我々は、影響伝播関係を導出するためにコンテンツの引用関係を利用することを考える。Web 上では

様々なコンテンツが存在するが、クオリティの高いコンテンツは頻繁に引用されることが多い。例として、論文の引用がある。よい論文は数多くの引用をされながら、やがて自分を中心とした1つの分野を形成する。つまり、よく引用される論文は優秀である、と考えられる。本研究ではクオリティの高いコンテンツを消費したユーザが、そのコンテンツから影響を受けたことを引用という形で明示して新しいコンテンツを投稿する、という振る舞い(図3)に注目し、投稿型サイトにおける引用関係に基づいて影響伝播関係を抽出する。

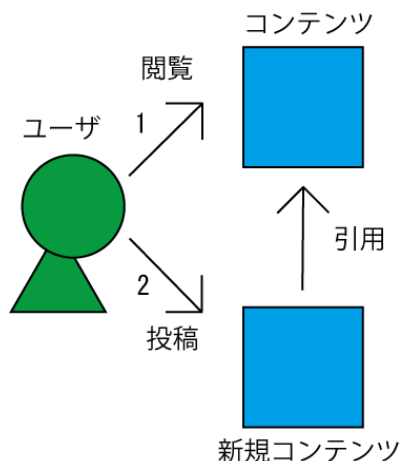


図3 既存コンテンツを引用した新規コンテンツの投稿

このような影響伝播関係を考慮することにより、他のコンテンツの中で引用され、自分自身は高い評価を得られなかったコンテンツを評価することができるようになる。具体的には、あるコンテンツに対して評価が与えられた際に、そのコンテンツに影響を与えたコンテンツにも何らかの評価を与える、というものである(図4)。

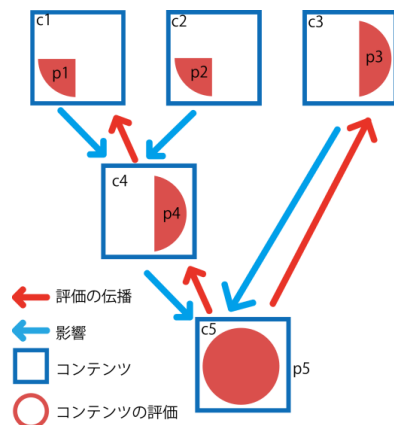


図4 影響伝播模式図

具体的な計算方法を示す。あるコンテンツの評価値 P_n は、 C_n を参照しているコンテンツの数を m 、それぞれのコンテンツが参照しているすべてのコンテンツの

数を x_m として、

$$P_n = \sum_{k=1}^m \left(\frac{1}{x_k} \cdot P_k \right)$$

となる。この計算式により、高い評価値を得たコンテンツだけでなく、クオリティの高いコンテンツに引用されたコンテンツの評価値が高くなる。

本手法は、基本的に **PageRank** における評価値の決定方法と同様な考えである。**PageRank** と本手法の違いは、ネットワークの構造と初期値にある。本論文では作品の影響関係であるために、閉ループは存在しないものとする。また、ダンピングファクタも考えないために、初期値が決まらないう値を計算することができない。本手法では、別のランキングによる評価値を初期値として利用する。すなわち、本手法は、普及度などによるある視点からのランキングと影響の伝播の構造という異なる視点からの情報を融合してコンテンツの評価を行う。

3.3. 影響伝播の取得

普及度を元にした評価法に比べ、影響伝播は抽出することが困難である。これは、1つのコンテンツの評価を定める際に、影響を与えた可能性のあるコンテンツをすべて考慮する必要があるからである。また、投稿型サイトにおいても、この関係を抽出することは困難であるが、本研究では投稿者のコメント等から抽出することを考える。ニコニコ動画では、投稿時に参照したコンテンツや、関連するコンテンツをハイパーリンクとして記述することが一般的であるため、本研究では前述の理由と、投稿型コンテンツにおいて、動画コンテンツは画像や音楽といった様々なコンテンツが存在することから、対象をニコニコ動画とした。ニコニコ動画における影響伝播を抽出する方法として、本手法では動画の再生ページ中に現れる投稿者コメントを利用して影響伝播を抽出する。ニコニコ動画では、投稿した動画を制作する際に参照した動画が存在する場合、参照した動画へのリンクを付与することが一般的に行われている(図5)。本手法では、このリンク構造を辿ることで、影響伝播関係を抽出する。



図5 ニコニコ動画における影響伝播取得

再生数を初期の評価値とした.

本研究では、影響伝播により与えられた評価値を元にリランキングを行い、有効性を評価することを目指す。元のランキングから、それぞれのコンテンツの影響伝播を考慮し、与えられた評価値を元にランキングを再構成する。集計する期間を1日、1週間とし、それぞれ元のランキングとの比較を行う。投稿型サイトでは毎日のランキングに大きな差があることから、我々が求めるクオリティの高いコンテンツは集計期間が長い場合に高い評価を受けると考えられる。つまり、長期間に渡って他のユーザのコンテンツ制作に影響を与えたコンテンツを抽出することができる。

上記のアプローチに基づいて、ランキングを行うプロトタイプシステムを実装した。システムは Perl と SQLite を利用して実装し、Windows XP 上で動作する。このシステムでは、ニコニコ動画のデイリー再生数ランキング評価値の初期値とし、影響伝播を考慮し、コンテンツのクオリティに基づく新しいランキングを構成する。具体的な手順を以下に示す。

1. ニコニコ動画のデイリー再生数ランキングの情報を取得する

2. 100 位までにランク付けされたそれぞれの動画に関して再帰的に影響伝播関係を抽出する

3. 100 位までにランキングした動画に対して、順位に応じて評価値を付与する

4. 影響伝播関係に基づいて、引用された動画に評価値を伝播させる。

5. ポイントを一定期間の集計を行い，評価値をもとに再ランキングする

本システムでは、付与されるポイントは上位 100 位までにランクインした動画に対し、それぞれの当日の

提案手法の有効性を評価するために、実験を行った。ニコニコ動画のデイリーマイリストランキングの情報を取得し、ランクインした動画それぞれに影響伝播を抽出するプログラムを用いた。ここでは、1日、1週間の2つのランキングを元に影響伝播を考慮し、比較した。現在ニコニコ動画では過去のランキング情報を取得するためには有料会員である必要があるため、システムでは有志の Web サイト（ニコニコチャート <http://www.nicochart.jp>）の情報をもとにして処理を行った。

1日分の処理は2009年11月15日のデイリー再生数ランキング(表2)を対象とした。影響伝播を考慮したランキングを表3に示す。

表 2 影響伝播考慮前(1 日)

順位	動画タイトル	投稿日時
1	【東方】 Bad Apple!! PV 【影絵】	2009/10/27
2	【鏡音リン】炉心融解【オリジナル】	2009/08/30
3	【オリジナル曲 P V 】 clock lock work 【初音ミク】	2009/11/27
4	2009年シェリルメドレー完成版 私の歌を聴けええええええええええ！	2009/11/26
5	クマさんチャーハン	2009/12/04

結果としては、元のランキングで1位にランキングされた「【東方】Bad Apple!! P V【影絵】」という動画の関連動画が、影響伝播考慮後も上位にランキングされた。この「【東方】Bad Apple!! P V【影絵】」と

表 3 影響伝播考慮後(1 日)

順位	動画 ID	動画タイトル	評価点	投稿日時	最高順位
1	sm8628149	【東方】Bad Apple!! P V【影絵】	305635.5	2009/10/27	1
2	nm3601701	【UP主が見たい】Bad Apple!! P V【誰か描いてくれ】	242403.37	2008/06/08	2
3	sm7284045	【UP主が見たい】Bad Apple!! PV 全比較版【みんなが描いてくれた】	231602.37	2009/06/08	13
4	sm2077177	【Alstroemeria Records × Bad Apple!! 】-Bad Apple!! feat. nomico-	221291.5	2008/01/19	7
5	sm6105642	クマさんチャーハン	206394.75	2009/02/10	132

いう動画は、影響伝播考慮後の2位にランキングされた「【UP主が見たい】Bad Apple!! PV【誰か描いてくれ】」という動画を引用した動画である。影響伝播考慮後にランクインした、「【東方】Bad Apple!! PV【影絵】」と関連する4つの動画の時系列と関連性を述べる。

- 2008年01月19日
【Alstroemeria Records × Bad Apple!!】-Bad Apple!! feat. nomico-という動画が投稿される。
- 2008年06月08日
1を視聴したユーザが「【UP主が見たい】Bad Apple!! PV【誰か描いてくれ】」という動画を投稿する。
- 2009年06月08日
2を引用した動画が多数投稿されたため、2を投稿したユーザが「【UP主が見たい】Bad Apple!! PV 全比較版【みんなが描いてくれた】」を投稿。
- 2009年10月27日
2を引用した動画として、「【東方】Bad Apple!! PV【影絵】」が投稿される。

元のランキングでは4の「【東方】Bad Apple!! PV【影絵】」という動画が大きく評価されていたが、影響伝播考慮後では、それらの関連動画が大きな評価を受けた。この影響伝播関係の元となったのは1の「【Alstroemeria Records × Bad Apple!!】-Bad Apple!! feat. nomico-」であり、この動画は楽曲である。この動画を見たユーザが2を投稿し、のちのPV動画の投稿につながったと考えられる。

このモデルは、まさに本手法で提案している影響伝播と言える。「【東方】Bad Apple!! PV【影絵】」という動画が高く評価されることにより、その動画が引用した関連動画も評価された。

5.3. 1週間分

1週間分のランキング情報に関しては、2009年11月30日から12月6日までのランキング(表4)を元に処理を行った。影響伝播を考慮したランキングを表5に示す。

表4 影響伝播考慮前(週間)

順位	動画タイトル	投稿日時
1	大人には聴こえない音で演奏してみた【耳年齢チェック有】	2009/10/23
2	【東方】Bad Apple!! PV【影絵】	2009/10/27
3	【スラムダンク】山王vs湘北 2Q [NBA2K9]	2009/11/12
4	スーパーマリオの声真似	2009/11/11
5	『低燃費少女ハイジ』まとめ 第01話～第05話	2009/11/09

結果は1日分の結果と大きな差異がみられなかった。これは「【東方】Bad Apple!! PV【影絵】」という動画が長期間にわたり元のランキング上位に存在したため、視聴者の多くが関連する動画を視聴していたことが理由と考えられる。

5.4. 実験のまとめ

今回行った実験では、影響伝播考慮後のランキングに「【東方】Bad Apple!! PV【影絵】」に関連する動画が上位に多数ランキングされた。元のランキングの2位以下が大きく変化したことから、影響伝播を考慮することにより評価される動画は多数存在すると考えられる。

今回の実験結果において最も興味深いのは、1日、1週間と短い期間での実験を行ったにも関わらず、2008年1月に投稿された動画が影響伝播考慮後に上位にランクインしたことである。これは元のランキングにおいて1位であった、「【東方】Bad Apple!! PV【影絵】」が非常に高い評価を得ていたからであるが、影響伝播を考慮することにより、深い引用関係を抽出できることが分かった。

今回行った実験の対象期間では、1つの引用構造が大きく評価されてしまったため、異なる期間を対象として実験を行うことで、新しい事実が発見される可能性がある。

表5 影響伝播考慮後(週間)

順位	動画ID	動画タイトル	評価点	投稿日時	最高順位
1	sm8628149	【東方】Bad Apple!! PV【影絵】	1249868	2009/10/27	1
2	nm3601701	【UP主が見たい】Bad Apple!! PV【誰か描いてくれ】	980082.1	2008/06/08	2
3	sm2077177	【Alstroemeria Records × Bad Apple!!】-Bad Apple!! feat. nomico-	802529.1	2008/01/19	7
4	sm7284045	【UP主が見たい】Bad Apple!! PV 全比較版【みんなが描いてくれた】	735987.1	2009/06/08	13
5	nm4595473	【UP主が見たい】Bad Apple!! PV full【誰か描いてくれ】	593431	2008/09/12	順位なし

6. 関連研究

本研究で提案した影響伝播と関連する研究として、PageRank[15]が挙げられる。PageRank では重要なウェブページは多くのページからリンクされると考えられることから、被リンクに基づいて Web ページの重要度を判定している。PageRank は Google 社の検索エンジンに用いられている。

また、PageRank を利用した研究も数多くなされ、その 1 つには TwitterRank[16]がある。TwitterRank は Twitter 版の PageRank と言うことができ、(1)多くのユーザから Follow されているユーザは価値が高い。(2)価値の高いユーザが Follow しているユーザは価値が高い。という論理に基づき、独自のアルゴリズムを利用している。

PageRankと同様にWebページを対象とする手法として、HITS[17]が挙げられる。HITS はハブ (Hub) とオーソリティ (Authority) というコンセプトを元にし、オーソリティ値を利用して Web ページの順位付けを行う。これらに対し、本研究で提案した影響伝播は、投稿型サイトにおけるコンテンツを対象としており、さらに、与えられた評価値を元に、引用関係にあるコンテンツを評価するというものである。しかし、これらは基本的にコンテンツ間のリンク構造のみに基づいてコンテンツの重み付けを行う物である。それに対して提案手法では、ランキングと参照構造を複合的に利用してランク付けを行う。

また、風間ら[18]は、ブログ情報を用いて、注目されている情報源の推薦を行った。トピックを表す検索語でブログエントリを収集し、本文中のハイパーリンクを分析することで、情報源を発見する。そして、それらを起点とする情報伝播ネットワークを抽出し、ブログの情報源を多面的にランキングする手法を提案した。

一方で、濱崎ら[19]はニコニコ動画における大規模協調的な創作活動に関する研究を行った。彼らは動画の引用関係ネットワークに注目し、また、「初音ミク」というタグのつく動画については、作曲が創作活動において牽引力を持っていること、そして作画がその間口を広げることに貢献していることを発見した。投稿されたコンテンツに創作物が多いという点は興味深く、今後の研究において考慮すべき事項と言える。

7. まとめ

現在のランキングは普及度を元にしたものが多く、コンテンツのクオリティを測ることができない。本研究ではコンテンツのクオリティを、他のユーザの創作物に与えた影響という視点からとらえし、影響伝播という概念の下でコンテンツのクオリティを反映したラ

ンキングを構成することを試みた。

影響伝播関係とは、あるコンテンツが、他のコンテンツの創作活動に影響を与えた関係である。投稿型サイトにおいて、多くのユーザがコンテンツに影響を受けたとき、そのコンテンツを引用した新規コンテンツを投稿することが多いということに注目した。

本研究の目標は影響伝播関係を利用することにより、コンテンツのクオリティを考慮したランキングを制作、評価することであった。また、ニコニコ動画のデイリー再生数ランキングを対象とし、実験を行った。出力された影響伝播考慮後のランキングでは、1つの引用関係が大きく評価されてしまったが、元のランキングでは評価されていなかったコンテンツを発見することができた。

本研究においては、クオリティの高いコンテンツを、他のユーザの新しい創作活動に影響を与えるコンテンツであることと定義した。定義に基づき影響伝播を用いてのリランキングを行った。しかし、クオリティの高いコンテンツが他のユーザの新しい創作活動に影響を与えるコンテンツであることと定義した点で、出力されるデータは一律のものになってしまう。つまり、個々のユーザの要求を考慮することができない。この点は本研究において最も大きな障壁となった。例えば、本研究において提案した影響伝播を考慮したランキング手法の考え方で最もクオリティの高いミュージシャンは誰か、という問いを考えてみる。おそらく、出力されるデータの上位にはビートルズが挙がるだろう。ユーザの多くは、ビートルズは有名なロックミュージシャンであり、またクオリティの高いバンドであると考えられる。しかし、ユーザが求める出力として、ビートルズが必ずしも要求を満たした結果であるとは限らない。コンテンツの推薦に対して、何らかの方法でユーザの状況や興味を考慮することが必要であると思われる。

本研究ではコンテンツのクオリティを評価するために、他のユーザの新しい創作活動を誘発するコンテンツを、クオリティが高いと定義した。しかし、この定義はクオリティの高いコンテンツの振る舞いの1つであるにすぎない。本当にクオリティの高いコンテンツを見つけるためには、さらなる議論が必要である。この問題を解消する1つの例として、ユーザがコンテンツに興味を持つかどうか、を推定することが考えられる。現状のランキングは視聴数、閲覧数を元に行っていることは前述の通りである。この視聴数、閲覧数は文字通りコンテンツを視聴したユーザの数を示しているが、このユーザの中でどれほどのユーザがこのコンテンツに興味を持ったのか、という点に注目する。今回の実験の対象としたニコニコ動画においては、Web

ページのブックマークと同じような扱いをされる「マイリスト」が存在し、また個々のコンテンツにおいてマイリストの数を取得することも可能である。この時、再生数をマイリスト数で除算すれば、視聴したユーザの中でどれほどのユーザがコンテンツをブックマークしたかを得ることができる。マイリスト、もしくはブックマークするという行動は興味を持つこととも判断でき、興味を持ったユーザの割合を出すことができる。このような手法も考慮していくべきだと考えられる。

また、今回の手法では意識的なリンクに基づいて影響伝播の抽出を行ったが、投稿者が無意識に親となるべき動画を参照している作品を投稿した場合（つまりリンクの記述がない場合）や、投稿者の記述忘れといった場合に影響伝播を抽出することができない。影響伝播抽出の手法についても、今後改善の必要がある。

今回の研究結果を踏まえ、今後の展開としては、ユーザの Web 上での振る舞いについて考えていくことが挙げられるブロードバンドインターネット環境が整備され、また、本研究で扱った投稿型サイト等が普及したことにより、ユーザが Web 上で他のユーザとコミュニケーションをとる機会が増加した。特に、投稿型サイトでは演奏者であるユーザが存在し、同様に歌手や、映像を作るユーザも存在する。その中で、一部のユーザ達は協力してコンテンツを制作しており、今後そういった協調創作活動は増加すると思われる。このような協調創作活動が行われるようになったことの背景には、SNS やブログの存在が挙げられる。ユーザはそれらを利用して Web 上での友好関係を結んでいる。そういったユーザの Web 上での振る舞いに注目し、Web を介したユーザのコミュニケーションや協調創作活動について考えていくことも重要である。

文 献

- [1] YouTube <http://www.youtube.com/>
- [2] ニコニコ動画 <http://www.nicovideo.jp/>
- [3] Flickr <http://www.flickr.com/>
- [4] Pixiv <http://www.pixiv.net/>
- [5] Veoh <http://www.veoh.com/>
- [6] Pandora TV <http://www.pandora.tv/>
- [7] Novelist.jp <http://novelist.jp/>
- [8] 小説家になろう <http://syosetu.com/>
- [9] ブルーキャンパス <http://jouti.sakura.ne.jp/>
- [10] MySound <http://players.music-eclub.com/>
- [11] NEXTMUSIC <http://www.nextmusic.net/>
- [12] ACOMI <http://www.acomi.jp/>
- [13] オリコンスタイル <http://www.oricon.co.jp/>
- [14] ニコニコチャート <http://www.nicochart.jp>
- [15] S.Brin and L.Page, “The anatomy of a large-scale hypertextual web search engine,” Computer Network and ISDN Systems, 30(1-7):pp.107-117, 1998
- [16] Kleinberg, J.M. “Authoritative sources in a hyperlinked environment,” Journal of the ACM, Vol.46, No.5, pp.604-632, 1999
- [17] Jianshu Weng, Ee-Peng Lim, Jing Jiang, Qi He, “TwitterRank: Finding Topic-sensitive Influential Twitterers,” WSDM.
- [18] 風間一洋, 今田美幸, 柏木啓一郎, “ブログ空間の情報伝播特性を用いた情報源の多面的ランキング,” WebDB Forum 2009 予稿集, CD-ROM, Nov,2009.
- [19] 濱崎雅弘, 武田英明, 西村拓一, “動画共有サイトにおける大規模協調的な創造活動の分析,” 情報処理学会研究報告, Vol.2009-MUS-80 No.11, pp.1-6, May, 2009.