

ソーシャルリンクの特徴と優位性

Characteristics and Advantage of Social Linking

長屋 彩子*¹ 松村 敦*¹ 宇陀 則彦*¹
 NAGAYA Ayako MATSUMURA Atsushi UDA Norihiko

*¹筑波大学大学院 図書館情報メディア研究科

Graduate School of Library, Information and Media Studies, University of Tsukuba

Social Tagging is a method for sharing and organizing contents by collaborative tagging. However, edges of the content networks organized by Social Tagging do not always express correct relationships, because they are organized only through ambiguous keywords added by users within various contexts. In order to improve the quality of content networks, we propose Social Linking which organizes contents by collaborative linking and investigated the characteristics of the two undergraduate curriculum networks organized by both methods. In this paper, we show the advantage of the Social Linking to the social tagging, according to network analysis of these two curriculum networks.

1. はじめに

Social Tagging (以下, ST) とは, 利用者が共有したコンテンツにタグと呼ばれるキーワードを付与することで, 注釈や索引語として利用できるほか, タグをたどりコンテンツを次々と閲覧できる. しかし, そのつながりは付与されたタグが一致するかどうかによって依存するため, 利用者の意図に反する可能性がある. 主な原因として, タグの多義性, 表記揺れ, 様々なレベルのタグが混在する Basic level がある [Golder 05].

これに対し本研究では, 利用者がコンテンツ間の関係性を判断し, 明示的にリンクを付与する Social Linking (以下, SL) を提案している. SL では関係性のある 2 つのコンテンツに対してつながりを直接付与するため, 利用者の意図が反映され, 精度の高いリンクで構成されたコンテンツネットワークが形成されることが期待される. 本稿では, ネットワーク分析とタグの分類を行い, ST と比べ SL の方がより精度の高いつながりで組織化できることを示す.

2. 授業科目の組織化実験

2.1 実験の概要

SL の特徴を明らかにするために, ST との比較実験を行った. 実験対象は, 筑波大学図書館情報専門学群の授業科目 155 科目である. 学群学生 111 名に, web 上に構築した実験システムで, 授業科目に対してリンクとタグを付与してもらった. なお, 他の学生の付与したタグやリンクを見ることはできない.

2.2 授業科目ネットワークの特徴 [長屋 07]

実験で形成された 2 つの授業科目ネットワークの比較を行った. 全体及び 3 種類のエッジ毎のサブグラフにおける, ノード数, エッジ数, クラスタ係数 C を表 1 に示す. ネットワーク全体を構成するノード数は同程度であるが, エッジの数は ST の方が 4 倍程度多く, クラスタ係数 C を見ると ST が非常に高い凝集性を持つことが分かった. 科目同士をつなぐエッジには, 内容, 印象, 履修に関するものの 3 種類があり, 内容に関するエッジのみからなるサブグラフを取り出すと, エッジ数, クラスタ係数の値が近づき, ネットワーク構造が類似したものになることが分かった.

連絡先: 長屋彩子, 筑波大学 図書館情報メディア研究科,
 茨城県つくば市春日 1-2, ayako@slis.tsukuba.ac.jp

表 1: 授業科目ネットワークの特徴

	ノード数	SL エッジ数	C	ノード数	ST エッジ数	C
全体	120	729	0.435	124	3152	0.782
内容	118	620	0.442	113	957	0.504
印象	45	68	0.161	87	2019	0.772
履修	66	156	0.255	105	2285	0.797

3. 距離に着目した分析

今回は内容に関するエッジのみに注目し, SL と ST のより詳細な比較分析を試みた. ST における内容に関するエッジは付与されたタグの一致によるものであり, 共通の内容を持つことを示す. 一方で, SL では共通の内容を持つという関係性の他に [基礎と発展の関係] などといった ST では表現できない関係性も表現できる. 今回の分析では, SL と ST で共通に表現可能な, 共通の内容を持つという関係性のみに着目し, SL の特徴と優位性を考察する. その際, 科目間の距離と関係性に着目して分析をおこなった.

3.1 最短パス長の分布

授業科目のネットワークにおいて, 最短パス長の平均は, ST で 2.257, SL で 2.796 となった. また, 最大の最短パス長, すなわちネットワークの直径はどちらも 6 であった. 図 1 は, ST と SL における各科目組の最短パス長の分布を示したグラフである. グラフを見ると, ST ではパス長が 2 の科目の組が最も多く, 4 科目以上離れているものは少ない. 一方で, SL ではパス長が 2, 3 の科目組が多いが, ST に比べると 2 から 6 まですらゆるやかに分布している.

科目間の最短パス長は, 関係性の高い科目間で短く, 低い科目間で長くなるのが望ましい. そこで, 各ネットワークにおいてどのような関係性でパス長が短くなっているかを調べた.

図 2 は, ST と SL における各科目組の最短パス長の差の分布を示したグラフである. 科目間の距離が SL と ST で等しい科目組が最も多く, 距離の差が大きくなるほどその数は減る. グラフが左側に膨らんでいるのは, ST で距離が近くなる科目組が多いことを示している. パス長の差が最も大きいのは, 3, -4 の科目組である. ST または SL の一方では近くに

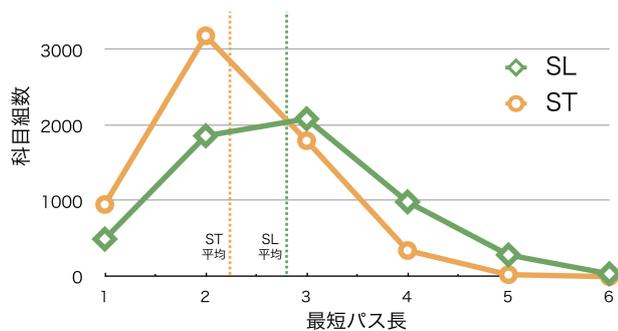


図 1: 最短パス長の分布



図 2: 各科目組の最短パス長の差 (ST - SL)

配置されているのにも関わらず、もう一方で遠くに配置されているこれらの科目組は、ST と SL のそれぞれにおいて特徴的な関係性であると考えられる。

3.2 ST で距離が近い関係

最短パス長の差 (ST - SL) が -4、すなわち、ST では距離が近いのに対し、SL では距離が遠くなった科目の組を表 2 に示す。

表 2: ST で距離が近い科目の組

最短パス長 ST	最短パス長 SL	科目名	タグ	科目名
1	5	日本古文入門	[歴史]	科学技術史: 発想法
1	5	学術情報流通論	[書店]	出版流通論
1	5	学術情報流通論	[図書館]	図書館文化史
1	5	図書館文化史	[図書館]	学校図書館メディアの構成
2	6	科学技術史: 発想法	[歴史], [崩し字]	古文獻講読

『日本古文入門』と『科学技術史: 発想法』の科目組は、SL では最短パス長が 5 なのに対し、ST では 1 である。この科目間をつなげていたのは、[歴史] というタグであり、それぞれ付与した被験者は異なっていた。『日本古文入門』と『科学技術史: 発想法』はどちらも歴史に関するものではあるが、一方は日本の古文、他方は科学史であり、[歴史] というタグの多義性により ST での距離が近くなっていたが、SL では遠くに配置されていた。このように、多義性のあるタグは、科目間の関係がないにも関わらず、ST では距離を縮めてしまっていた。

また『図書館文化史』と『学校図書館メディアの構成』、『学術情報流通論』は各組とも、SL では最短パス長が 5 なのに対し、ST では 1 である。これらの科目間を距離 1 でつなげていたのはどちらも [図書館] というタグであり、ST で 19 科目と最も多くの科目に付与されているタグであった。いずれの科目も図書館が関係するが [図書館] というタグは図書館情報学

を学ぶ学生のコミュニティにおいては一般的であり、重要性を持たないと考えられる。このような一般的な語のタグは、ネットワーク構造を利用するには、不当に科目間の距離を近づけてしまうのでノイズとなり得る。

3.3 SL で距離が近い関係

最短パス長の差 (ST - SL) が 3、すなわち、SL では距離が近いのに対し、ST では距離が遠くなった科目の組を表 3 に示す。

表 3: SL で距離が近い科目の組

最短パス長 ST	最短パス長 SL	科目名	リンクのラベル	科目名
4	1	法律学	[社会情勢][文系の基本]	経済学
4	1	政治学	[文系の基本][社会]	経済学
5	2	知覚情報処理	[どちらも認知心理学に関連している分野である]・情報デザイン・[XML]	マークアップ言語

『知覚情報処理』と『マークアップ言語』の科目組は、ST では最短パス長が 5 なのに対し、SL では 2 であった。これらのリンクを見ると、『知覚情報処理』と『情報デザイン』が [どちらも認知心理学に関連している分野である] というラベル、『情報デザイン』と『マークアップ言語』が [XML] というラベルでつながっていた。一方タグは、『情報デザイン』には [アフォーダンス][インターフェース]、『知覚情報処理』は [色知覚][受容器]などが付与されていた。

このように、ST に比べ SL での距離が近い科目組のリンクのラベルは、2 単語以上からなる句や文による表現が見られ、より詳細な共通の内容記述でつながっている。ST のタグにおいても詳細な内容記述があるが、タグが一致しなかったために科目間の距離は遠くなったと考えられる。このように、SL ではキーワードに依存せず、関係を表現することができ、関係のある科目間の距離を明示的に近づけるためには有利である。

4. おわりに

本稿では、SL と ST による授業科目ネットワークそれぞれにおける科目組の最短パス長を比較した。その結果、ST では一般的なタグや多義性のあるタグがノイズとなり得るつながりを作り、関係性の少ない科目間の距離を縮めることが分かった。一方、SL で関係性を判断することでより精度の高いつながりを付与できる。また、SL では句や文などのより詳細なつながりを表現できる例も見られた。したがって、ネットワークの構造を用いて科目を見ていく際に、SL ではより精度の高いつながりをたどることができ、近くに関係性の高い科目が配置できる。しかしながら、ST で付与された詳細な内容を表すタグは注釈としては有用であり、一般的な語であっても大まかにカテゴリ分けするには便利である。今後の課題は、このような SL と ST の異なった特徴を理解した上で、システムを実装することである。

参考文献

[Golder 05] Golder, S. and Huberman, B. A.: The Structure of Collaborative Tagging Systems. Journal of Information Science. vol. 32, no. 2, 2006, p. 198-208.

[長屋 07] 長屋彩子, 松村敦, 宇陀則彦. Social Linking で形成されるコンテンツネットワーク構造の特徴. 第 10 回 Web インテリジェンスとインタラクション研究会, 2007, p. 31-32.