

属性付きレビュー関係ネットワークを利用したレシピ共有サイト における高影響度ユーザ/レシピの特徴分析

平野 雅也[†] 豊田 哲也[†] 大原 剛三[†]

[†]青山学院大学 理工学部

1. はじめに

近年普及しているレシピ共有サイトでは、利用レシピに対する感想を投稿するレビュー機能が存在し、レビューを貰うことがレシピ投稿者の投稿意欲向上につながることで指摘されている。そのことから、レビュー数の多いユーザ、もしくはレシピの特徴を明らかにすることは重要である。そこで本研究では、属性付きレビュー関係ネットワークを構築し、それを通して多くのレビューを集める高影響度ユーザ/レシピの特徴を明らかにする。具体的には、ユーザをノード、レビューの投稿関係をリンクとしたネットワークに対し、さらにレシピの食材・カテゴリなどに基づきノードとリンクに属性を付与する。そして、ネットワーク構造から抽出されるコミュニティ構造と属性情報の類似性を対比させることで、ユーザやレシピの特徴分析を試みる。

2. 属性付きレビューネットワーク

本研究では、投稿型レシピ共有サイトにおけるユーザ間のレビュー投稿関係を有向ネットワーク $G(V, E)$ で表現する。ここで、 V は各ユーザに対応するノードの集合であり、ユーザ $u \in V$ がユーザ $v \in V$ の投稿したレシピに対してレビューを投稿したことを有向リンク $(u, v) \in E$ により表現する。ユーザ v が投稿した複数のレシピに対してユーザ u がレビューを投稿する可能性があるが、ここではノード u, v 間に多重リンクを生成するのではなく、リンク重み w_{uv} によりユーザ u がユーザ v の投稿したレシピに投稿したレビュー数を表現する。ユーザ v が投稿したレシピの総数を R_v としたとき、明らかに $w_{uv} \leq R_v$ である。また、

ノード集合 $F_v = \{u | u \in V, (u, v) \in E\}$ に対して、 $d(v) = |F_v|$ をノード v の入次数と呼び、一方、 $d_w(v) = \sum_{u \in F_v} w_{uv}$ をノード v の重みつき入次数と呼ぶ。ノード v の重みつき入次数 $d_w(v)$ は、 v が投稿したレシピに対するレビューの総数を表し、この数値が大きいユーザほど他のユーザに大きな影響を与えていると解釈できる。

次に、ノードに付与する属性を定義する。ノードに対応するユーザは、レシピ投稿者とレビュー投稿者という2つの役割をもつ。そして、ユーザが投稿するレシピ、およびレビューの対象となるレシピ、いずれのレシピにもそのユーザの嗜好や要求が反映されると考えられる。ただし、必ずしもそれらのレシピのもつ特徴が一致するとは限らない。そのため、ここでは各ユーザ v に対してレシピ特徴ベクトル $A_{rec}(v)$ とレビュー特徴ベクトル $A_{rev}(v)$ の2つの特徴ベクトルを定義する。いずれも、投稿者が自身であるかどうかに関わらず、レシピの特徴を表すものであり、その構成要素は共通のものを用いる。レシピの特徴としては、料理カテゴリ、目的、食材、調理方法など多様なものが考えられるが、ここでは、その中でも料理カテゴリの利用頻度を用いる。具体的には、 i 番目のベクトル要素は対象レシピ集合における料理カテゴリ $M(i)$ に属するレシピの相対出現頻度とする。たとえば、ユーザ v の投稿したレシピのうちパスタに属するものの数が p であった場合、パスタカテゴリに対応する $A_{rec}(v)$ の要素の値は p/R_v となる。

3. 分析結果

国立情報学研究所を通して公開されているクックパッドのデータを利用し、実際に属性付きレビューネットワークを構築し、そこから抽出したコミュニティに着目して、重みつき入次数、および入次数の高いユーザの特徴を分析した。今回は、クックパッドにおけるレビューに相当

Characterizing High-impact Users and Recipes based on Attributed Review Relation Networks on a Recipe Sharing Site

Masaya HIRANO[†], Tetsuya TOYOTA[†], and Kouzou OHARA[†]

[†]College of Science and Engineering, Aoyama Gakuin University

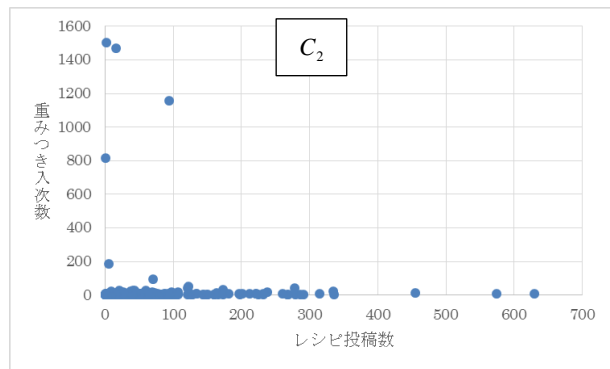
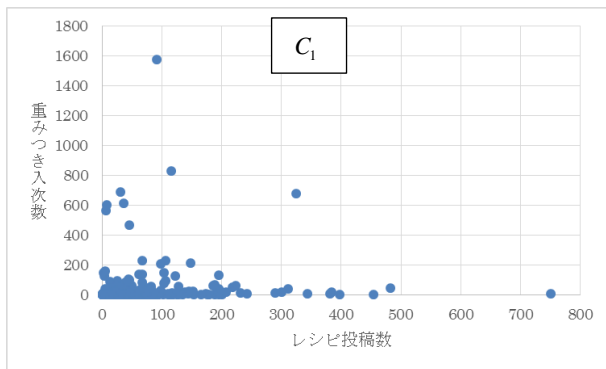


図1 各ユーザのレシピ投稿数と重みつき入次数の関係

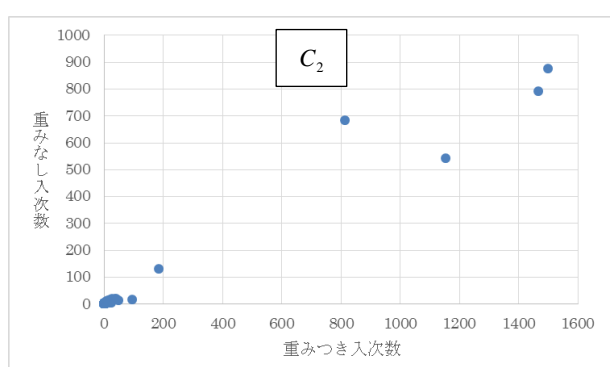
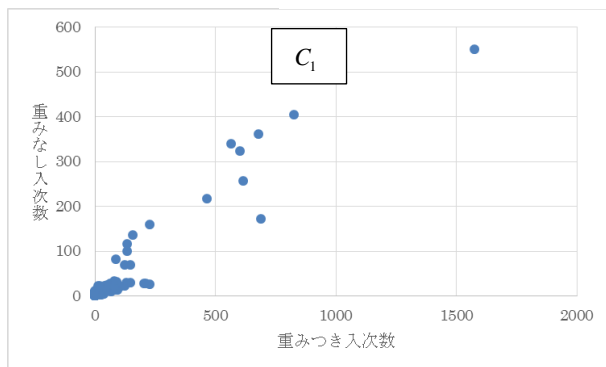


図2 重みつき入次数と重みなし入次数の関係

する「つくれば」を投稿したユーザのうち、投稿数上位 50 名を基準にネットワークを構築した。具体的には、それらのユーザと「つくれば」の対象レシピを投稿したユーザ間に重みつき有向リンクを生成した。得られた属性付きレビューネットワークのノード数は 6,510、リンク数は 32,109、平均次数（重みなし）は 4.9 となった。料理カテゴリはクックパッドのものをを用い、特徴ベクトルの次元数は 57 となった。

このネットワークを重みつき無向グラフとみなし、Louvain アルゴリズム[1]によりコミュニティを抽出した。その結果、抽出された 6 つのコミュニティのうち平均重みつき次数が最も高い (13.5) コミュニティ C_1 と、逆に平均重みつき次数が最小 (1.9) となるコミュニティ C_2 について考察する。

各コミュニティにおけるユーザのレシピ投稿数と重みつき入次数の関係を図 1 に、重みつき入次数と重みなし入次数の関係を図 2 に示す。図 1 から、多くの「つくれば」を得ているユーザは必ずしも多くのレシピを投稿しているわけではないことがわかる。また、図 2 から、いずれのコミュニティにおいても重みつき入次数と重みなし入次数の間に正の相関が見て取れるが、その比率は、 C_1 の方が 1 人から 2~3 件のレシピ

を受けているのに対し、 C_2 は 1 人から 1~2 件程度と差があることがわかる。また、特徴ベクトルの比較では、 C_1 では「お菓子」「野菜のおかず」などのカテゴリの出現頻度が相対的に高く、 C_2 では「お肉のおかず」「お弁当」などの出現頻度が相対的に高いことが観測できた。

4. おわりに

本稿では、投稿型レシピサイトにおいて影響力の高いユーザを分析するために、属性付きレビューネットワークを提案し、実際のデータに適用した。活性度の異なるコミュニティ間では、その特徴に違いが見られたものの、現段階では影響力の高いユーザを特徴づけるまでは至っていない。今後、詳細な分析を進める必要がある。

本研究では、クックパッド株式会社と国立情報学研究所が提供する「クックパッドデータ」を利用した。

参考文献

- [1]Fast unfolding of communities in large networks, V.D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, J. of Statistical Mechanics: Theory and Experiment, 2008(10) (2008).