

---

## クチコミサイトを用いた効率的販売促進のための考察 ～多重ネットワークによる ABS モデリングを用いて～

---

経済学研究科 李 詩 雨\*  
経済学部教授 谷 田 則 幸

### 1 はじめに

ウェブサイト上で公開されている商品に関する情報、消費者が実際に商品に対して使用後の感想と評価など情報が大量に発信されている。そのような情報は消費者同士の間で共有され、消費者に対してきわめて重要な情報と考えられる。このような情報は消費者だけではなく、企業においても有効的に利用しようとする動きが強まっている。

本論ではクチコミサイトに公開される既存の情報データを用いることで、サイト上の情報が消費者にどのような影響を与え、その結果として商品の話題性と商品に関するクチコミの情報がどのように変化するか、また商品に関する新しいクチコミがサイトに投稿された際にはユーザー同士の間で互いにどのように影響するかを研究するため、商品、ユーザーとクチコミに関するそれぞれのネットワークを関連付けた3重ネットワークを構築してシミュレーションを行った。クチコミサイトにおけるユーザー（消費者）行動全体の振る舞いを表現できる ABS モデル（Agent-Based-Simulation）を用いて人工社会を構築し、ミクロな行動から創発されたマクロな現象を観察する。

### 2 研究背景

これまでクチコミの情報に関する消費者の購買行動研究では、インターネットを利用した SNS、クチコミサイトなどソーシャルメディアへのレビュー投稿、消費者のニーズなどを分析することが広く行われてきた。商品に対する購買行動を行う時に他人のクチコミを参考すると答えた消費者は非常に多いことが示されている [田中・清水（2006）]。クチコミが発生する時には、マスメディアによる広告などとは異なって非計画的であり、意図的ではないことが多い。クチコミの受信者が商品に関する知識を充分に持っていないときには、クチコミ情報を利用することによって購買決定が可能となる [杉本（1997）]。

---

\* 現在の所属先：IBM Dalian Global Delivery Company

クチコミに関するテキスト・マイニングの研究も多く行われてきた。クチコミは、友人、家族など消費者同士が交換する意見であり、製品またはサービスを実際に使用したか、あるいは少なくとも関心を持つ人々が発信するものであるため、消費者に対して信頼しやすい情報と考えられる。香港のホテルを利用した顧客に対して実施された調査のテキスト情報を利用することで、テキスト・マイニングにより分析している研究もある [Lee, et al. (2011)]。また単語のコミュニティ性に基づいて重み付け指標である TF-ICF (Term Frequency-Inverse Corpus Frequency) を用いて、論文検索に用いることでクエリの関連語を提示する手法を提案する研究もある [岡崎ほか (2014)]。

以上の先行研究から、本研究ではクチコミサイトの消費者行動に対してクチコミの情報とユーザーの情報が、どのような影響を消費者に与えるのかを検証するとともに、またクチコミサイトにおける消費者は商品に対する情報探索を行う際に、この商品の販売にどのような違いが見られるのかを解明する。さらに、テキスト・マイニングを用いてクチコミサイトにおける商品の効果に関するクラスターに基づく特徴からシミュレーションの結果により消費者の購買行動を分析する。

### 3 シミュレーションの準備

#### 3.1 モデルの仕組み

クチコミサイトにおける商品、ユーザーとクチコミ三者の間のつながりに関する研究をするため、エージェント・ベース・シミュレーション (ABS) によりこの三つの対象をエージェントとして二次元空間で可視的に行動させる。エージェントとはネットワーク上を移動し、ユーザーが指定した情報を自動的・効率的に受送信するソフトモジュールである。そこで、クチコミサイトにおける商品、サイトを登録しているユーザーとクチコミをエージェント種として3重構造のクチコミサイト空間を想定する。図3.1のように、様々な商品エージェントは仮想のクチコミサイト上で存在する。クチコミサイトに登録しているユーザーもエージェントとしてクチコミサイト上に活動している。登録しているユーザーの中で、一部分の人はそちらの商品エージェントを選んでクチコミを投稿する。投稿したクチコミは他のユーザーエージェントに何らかの意味で影響を与える。商品、ユーザーとクチコミエージェントにはそれぞれのネットワークが形成しているし、互いにリンクが張られるため3重モデルと呼ばれる。本論ではクチコミサイトにおけるユーザーは商品により投稿された大量のクチコミを参考にすることとする。

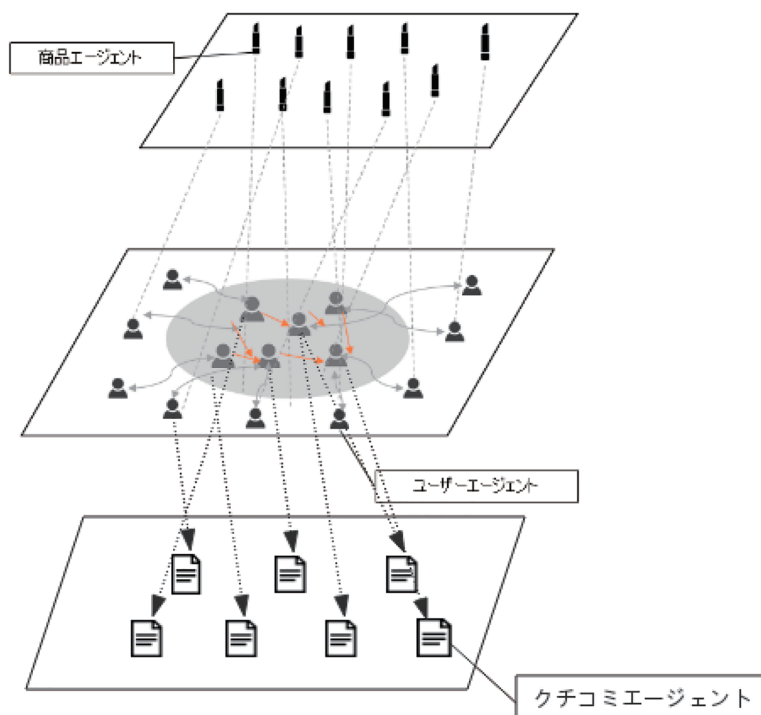


図3.1 クチコミサイトにおける3重構造の仕組み

## 3.2 情報による影響

### 3.2.1 クチコミによる影響

(A) クチコミサイト上のあるユーザーは他のユーザーが投稿したクチコミに一定の程度まで影響されることと仮定する。ここに、各ユーザーが商品の効果・機能に関する六つのキーワードによりクチコミに影響されることとする。ユーザー  $i$  が商品  $k$  に関するクチコミのキーワードにより影響される被影響度  $R_{(i,k)}$  は、以下のように定義する。

$$R_{(i,k)} = h_{(i)} \times EF_{(i,k)} \quad (1)$$

ここで、

$h_{(i)}$  はユーザー  $i$  が個人としてクチコミに影響されるウェイトである。 $h_{(i)}$  はユーザー  $i$  が書いた全クチコミの中で「評価」に関するキーワードの出現比率とみなすこととする。

$EF_{(i,k)}$  はユーザー  $i$  がクチコミから影響を受けられる閾値である。これにより、クチコミに影響されるウェイトが大きいほど、このユーザーに対してクチコミに影響される程度は大きくなる。 $EF_{(i,k)}$  は以下のように定義する。

$$EF_{(i,k)} = \sum_{m=0}^5 (\delta_{R(m,i,k)} \times \delta_{U(m,i)}) \quad (2)$$

ここで、

ユーザー  $i$  がクチコミサイトに投稿されたクチコミから影響を受けることは、商品に対するクチコミの中で六つのキーワードの出現確率とユーザー自身からクチコミに対する重視度

に基づいて判断される。

$\delta_{R(m,i,k)}$  はユーザー  $i$  が商品  $k$  に対してキーワード  $m$  により影響されるかどうかを判断する条件式である。以下のように定義する。

$$\delta_{R(m,i,k)} = \begin{cases} 0 & (EF_{k(m)} \leq \alpha) \\ 1 & (EF_{k(m)} > \alpha) \end{cases} \quad (3)$$

ここで、

$\alpha$  はユーザー  $i$  が商品  $k$  のクチコミに影響される閾値、シミュレーションよりコントロールできるパラメータとする。

$EF_{k(m)}$  は商品  $k$  のクチコミに対するキーワード  $m$  の出現確率である。

$\delta_{U(m,i)}$  はユーザー  $i$  自身がキーワード  $m$  に関する効果に影響されるかどうかを判断する条件式である。以下のように定義する。

$$\delta_{U(m,i)} = \begin{cases} 0 & (ef_{i(m)} = 0) \\ 1 & (ef_{i(m)} \neq 0) \end{cases} \quad (4)$$

ここで、

$ef_{i(m)}$  はユーザー  $i$  がクチコミにかかわるキーワード  $m$  に対する重視度である。本論では、これをユーザーが投稿した全クチコミに対してキーワード  $m$  の出現確率とみなす。

(B) ユーザー  $i$  は各商品に対するクチコミに書かれた総合点数にも影響を受けることとする。ある商品  $k$  に対する総合点数の平均をユーザー  $i$  がクチコミに影響されるウェイト  $h_{(i)}$  で除することで、影響された程度の大きさに対する実際の評価点数を被影響度といい、 $P_{(i,k)}$  と表す。被影響度  $P_{(i,k)}$  を以下のように定義する：

$$P_{(i,k)} = \frac{SP_{(k)}}{ef_{i(6)}} \times \delta_{P_{(i,k)}} \times D \quad (5)$$

ここで、

$SP_{(k)}$  は商品  $k$  に対する投稿したクチコミにかかる総合点数の平均値である。

$ef_{i(6)}$  は式(4)の定義で、ユーザー  $i$  がクチコミにかかわるキーワード「評価」に対する重視度である。

$P_{(i,k)}$  は、商品  $k$  の総合点数の平均値  $SP_{(k)}$  が高いほど影響させられる確率が高くなる。

$D$  はキャリブレーションより、0.01とする。

注：(5)式で、 $ef_{i(6)} = 0$  のときは、 $P_{(i,k)} = 0$  とする。

なお、ユーザーエージェントが総合点数に関する情報に影響させられる機会は必ずしも一定ではなく、商品  $k$  に対する投稿したクチコミにかかる総合点数の平均値より判断する。

$\delta_{P_{(i,k)}}$  は以下のように定義する。



$$\delta_{P(i,k)} = \begin{cases} 0 & (SP_{(k)} \leq \beta) \\ 1 & (SP_{(k)} > \beta) \end{cases} \quad (6)$$

ここで、

$\beta$ は総合点数がユーザーに影響を与える閾値であり、シミュレーションによりコントロールできるパラメータとする。

### 3.2.2 他のユーザーによる影響

クチコミサイトにおけるユーザーがサイトを閲覧するときに他のユーザーに関する情報にも一定の影響を受ける。また、各ユーザーは年齢層毎に行動による影響を受けることとする。当商品の購買者（ユーザー）の年齢にかかわる情報によってユーザーの被影響度  $A_{(i,k)}$  は以下のように定義する。

$$A_{(i,k)} = E \times \delta_{A_{(i,k)}} \quad (7)$$

ここで、

$E$ はユーザーが投稿者の年齢層に影響されるウェイトである（本シミュレーションでは、アンケート結果［ポーラ文化研究所（2016）］に基づき、口紅使用者が口紅を選ぶときに重視することの中で「年齢に見合う」という選択の確率を0.25とした）。

$\delta_{A_{(i,k)}}$ はユーザー  $i$ が商品  $k$ に対してユーザーの年齢層  $n$ に関する影響される閾値である。商品に対するユーザーがいる年齢層のユーザー数が多いほどこの値は高くなることとし、 $\delta_{A_{(i,k)}}$ は以下のように定義する。

$$\delta_{A_{(i,k)}} = \begin{cases} 0 & (U_{(i,n)} \leq \gamma) \\ 1 & (U_{(i,n)} > \gamma) \end{cases} \quad (8)$$

ここで、

$\gamma$ はユーザー  $i$ が個人として年齢にかかわる情報に影響される閾値であり、シミュレーションによりコントロールできるパラメータとする。

$U_{(i,n)}$ は商品  $i$ に対する年齢層  $n$ に属するユーザー数の比率である。以下のように定義する。

$$U_{(i,n)} = \frac{u_{(i,n)}}{\sum_{n=1}^7 u_{(i,n)}} \quad (9)$$

ここで、

$u_{(i,n)}$ は商品  $i$ に対する年齢層  $n$ に属するユーザー数である。

ユーザーはクチコミ情報とユーザー情報に影響される。影響される総被影響度  $I_{(i,k)}$  は以下のように定義する。

$$I_{(i,k)} = R_{(i,k)} + P_{(i,k)} + A_{(i,k)} \quad (10)$$

### 3.3 分析対象データ

本論の分析対象としては、化粧品のクチコミサイト「@cosme」における「口紅・グロス・リップライナー」の商品、商品に関するクチコミ及び投稿者を用いる。@cosme サイトに登録するユーザーの評価と総合点数などデータに基づく商品のクチコミランキングが公開されている。本論では、@cosme クチコミサイト上の2007年7月から2018年4月まで「口紅・グロス・リップライナー」というカテゴリに属する商品のデータを収集した。「クチコミランキング」の前十位の商品に関して、商品情報、投稿したユーザー情報とクチコミのテキスト内容を抽出した。抽出したクチコミの総件数は38422である。抽出した前十位の商品に関する基本情報は表3.1のようにする。

表3.1 クチコミランキング前十位の商品情報

商品 ID	ブランド	名称	クチコミ件数	発売時間	価格
10119418	オペラ	リップティント	1890	2016/11/1	1,500
2913937	ディオール	ディオール アディクト リップ マキシマイザー	557	2007/7/27	3,600
10133338	フローフシ	LIP38℃ リップトリートルージュ	687	2017/7/7	1,600
10055980	キャンメイク	ステイオンバームルージュ	458	2013/4/19	580
10143733	クレ・ド・ポーボーテ	ブリアンアレーブルエクラ	180	2018/1/21	4,000
10051660	イブ・サンローラン	ルージュ ヴォリュプテ シャイン	264	2013/4/19	4,100
10118017	オペラ	シアーリップカラー N	344	2016/11/1	1,200
10144698	SHISEIDO	ルージュ ピコ	132	2018/2/1	1,800
10123639	イブ・サンローラン	ヴォリュプテ ティンとインバーム	186	2017/2/3	4,300
10019914	シャネル	ルージュ ココ シャイン	132	2011/3/30	3,900

出所：2018年4月13日 @cosme 口紅ランキングより作成<sup>1)</sup>[@cosme (2018)]

### 3.4 データの前処理

#### 3.4.1 テキスト・マイニング

テキスト・マイニングとは、自由記述などのテキストデータを体系的、計量的に分析し、一次資料のみでは見出すことができない有効な情報を発掘（Mining）する手法の総称である。この手法を用いることで、テキストデータの分析を行う際に研究者による主観の影響をコントロールすることが可能となる。

本論では、TTM というテキスト・マイニングの前処理のフリーウェアを使用する。TTM はフリーソフトであり、Mecab を形態素解析エンジンとして採用している。Mecab とはオープンソースの形態素解析エンジンである<sup>2)</sup>。TTM を利用してクチコミのテキストを CSV 形式で読み込んで、単語の出現頻度・件数の集計結果を6種類のデータファイルとして作ること

1) クチコミ件数の集計期間は2018/1/12～2018/4/11である。

2) Mecab パッケージ <http://taku910.github.io/mecab/>

ができる。各ユーザーに対してある商品に対してクチコミを検索したいときはサイトのページでいくつかの絞り込み条件が提示されている。本論で使用するデータは「口紅・グロス・リップライナー」というカテゴリーの商品を対象とするため、この中で、カテゴリー口紅に関する「発色」、「ツヤ」、「落ちにくい」、「乾燥・荒れ」と「うるおい」という五項の効果をキーワードとする。また、各クチコミの中で頻繁的に出現する「心地」と「評価」もキーワードとする。TTM より以下の七つのキーワードの出現頻度のデータベースを作った。

7つのキーワード<sup>3)</sup>：発色、ツヤ、心地、落ちにくい、乾燥・荒れ、うるおい、評価。

### 3.4.2 クラスター分析

クチコミの全体傾向から商品の特徴的なものを把握するため、クラスター分析を用いて商品をいくつかのクラスターに分ける。クラスター分析とは、異なる性質を持つものを個体間の類似度あるいは非類似度（距離）に基づく同じクラスターにまとめ、そうでないものを異なるクラスターに集めるデータの分析方法である。

クラスター分析の前に、データセットを成分負荷行列に整理することは必要があるため、主成分分析を行う。主成分分析 PCA（Principal Component Analysis）とは、多くの変数により記述された量的データについて、複数の変数間の相関を少数個の合成変数に縮約し、データの解釈を容易にするための分析手法である。TTM による処理を行った出力結果であれば、抽出された語を変数とみなし、これらが各テキスト・マイニングで出現する頻度のデータを用いて、語どうしの関連を解析する際に用いることができる。得られたキーワードの出現頻度に関するデータベースを、R により関数 PCA を使って成分負荷行列が得られる。次にクラスター分析を適用してみる。クラスター分析の結果からは図3.5に示される。

クラスター分析の結果として、任意の縦軸のレベルで樹形図の枝を切って、近い関係にある語のグルーピングを行うことによって解釈できる。図3.5であれば、縦軸の値が0.6のところで枝を切ると、この10個の商品を三つのクラスターに分かれる。クラスター分析の結果によって同じクラスターに属する商品はキーワードに関する類似度が近いことを示す。

ここに、十個の商品の価格を四つの段階に分ける（A:500～1500；B:1500～2500；C:2500～3500；D:3500～4500；単位：円）。各商品に対してキーワードの出現頻度の平均得点を以下の表3.3に示す。一行目は商品の ID 番号である。こちらの各商品の平均得点は商品ごとにクチコミの中で、七つのキーワードの出現頻度の平均値と標準偏差に基づく偏差値とする。クラスター分析の結果からキーワードに基づき各商品の類似度が分かれる。各商品の平均得点から七つのキーワードにより各商品の差異が分かれる。そこで、この二つの結果からクラスターごとに特徴がまとめられる。

3) 日本語で異なる表記の同義語があるため、「落ちにくい」を「落ちにくい」と「落ない」に分ける、「うるおい」を「うるおい」と「潤い」に分ける、「評価」を「評価」、「評判」、「クチコミ」と「口コミ」に分けることにしてそれぞれの出現頻度をまとめる。

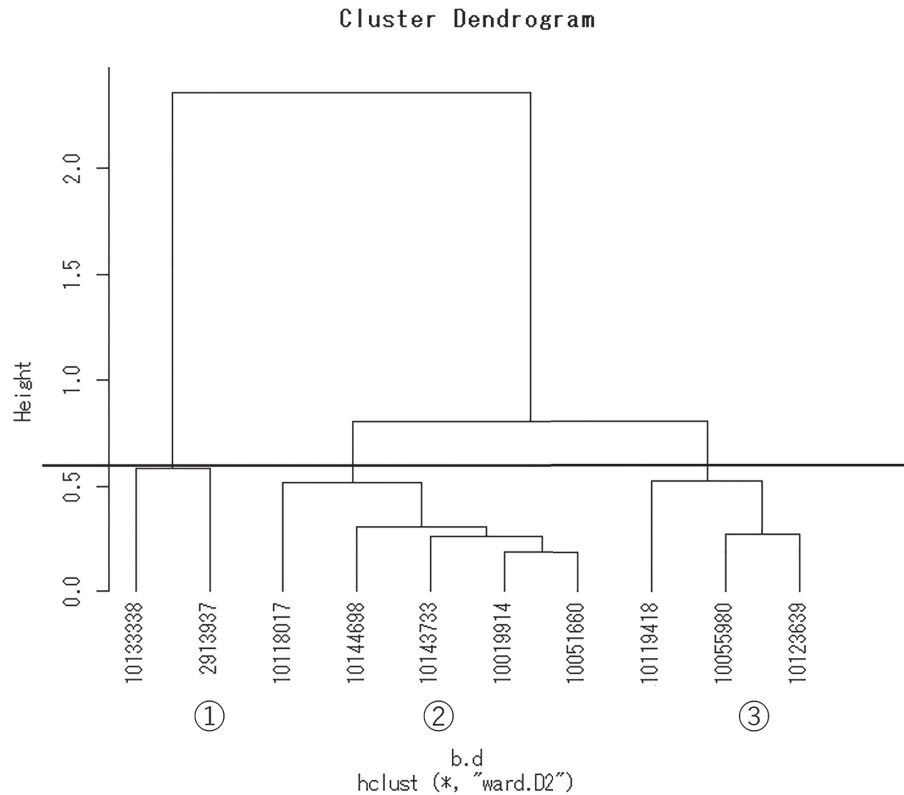


図3.5 Rのクラスター分析で得られた樹形図

表3.3 各商品の平均得点

語	10133338	2913937	10019914	10118017	10051660	10143733	10144698	10055980	10119418	10123639
価格段階	B	D	D	A	D	D	B	A	A	D
クラスター	1	1	2	2	2	2	2	3	3	3
発色	50.24	42.52	71.73	67.73	70.76	71.23	70.68	69.82	65.76	71.99
ツヤ	46.78	49.00	51.08	54.61	49.31	49.68	52.00	45.71	46.58	46.25
心地	44.71	44.79	44.31	47.28	46.19	49.68	47.33	41.87	42.61	44.58
落ちにくい	39.36	38.91	37.98	34.89	36.96	36.69	39.69	39.89	41.49	40.48
乾燥・荒れ	73.21	70.47	53.49	56.33	56.21	53.49	55.82	59.76	65.43	55.97
うるおい	47.53	46.48	45.12	48.45	44.58	43.66	41.81	46.06	42.60	46.74
評価	48.17	57.83	46.29	40.71	45.99	45.56	42.66	46.89	45.53	43.99

クラスター1：2つの商品がこのクラスターに入っている。各キーワードの平均得点からみると、キーワード「乾燥・荒れ」と「評価」の値が他より高いことが分かれる。二番目の商品ディオールのリップは2007年から発売する商品であるが、持続的に何年間でランキングのトップに占める。この商品の各キーワードの平均得点を見ると、「乾燥・荒れ」の値は他より高いである事が分かれる。

クラスター2：クラスター2には五個の商品からなるグループである。この中で三つの商品が価格段階Dに属する、二つはそれぞれ価格段階A、Bに属する。各商品のクチコミに書かれたキーワードの平均得点をみると、「発色」と「ツヤ」は他より高い値をもつことが分か

る。この中で「発色」は平均60を超えたことになる。

クラスター3：三の商品はクラスター3に入っている。このグループで「落ちにくい」というキーワードの平均得点は他のグループより高いことが分かれる。この中でID番号10119418の商品は一定の時期で一番多くのクチコミ件数をもっていた。値段はAの価格段階に属する。この商品に対してキーワードの平均得点から言うと、「落ちにくい」というキーワードの数値は他の商品に比べて高いである。これの以外に、顕著な特徴はあまり見えない。一瞬に話題が爆発する商品は一般的な効果があるということは、消費者にとっては「群集心理」が働く可能性があるのではないかと考えられる。

### 3.4.3 初期値の整理

人工市場により、シミュレーションでクチコミサイトにかかわる既存のデータに基づいて作ったルールに従って将来の消費者購買行動をとらえるため、各エージェントに初期値を入力することは必要である。ユーザーエージェントの初期値の中で、"effect"というユーザーが商品を重視する程度（7つのキーワードに基づく）を表す変数がある。各ユーザーが投稿したクチコミの中で、全キーワードの出現頻度の総数に占める七つのキーワードそれぞれの出現頻度の割合になる（0から1までの実数）。こちらは各ユーザーが投稿した全クチコミによって計算した結果とする。研究対象としてクチコミの件数は38422件（2018年4月13日より）があった。本研究では、コンピュータの処理能力の制約により、ランダムサンプリングで384件を抽出した。既存のユーザーエージェント数は381である。

## 4 シミュレーションのモデル

### 4.1 artisoc の概要

artisocとは「マルチエージェント・シミュレーション」を構築する環境であり、様々な研究分野に用いられる意思決定支援・分析ツールである。仮想の社会の中で一定の行動をさせることにより、人間の意思決定をモデル化して個々対象の相互作用により生じる様々な現象を分析することができる[山影（2007）]。ネットワークに関する用語を「マルチエージェント・シミュレーション」の言葉で解釈すると、頂点（ノード）がエージェントになり、枝（リンク）がエージェント同士の間での関係になる。artisocによるシミュレーションの基本要素は次のようにまとめている。

- ① ツリー（Tree）画面：Universeと呼ばれる「環境」、「空間」と「エージェント」という3階層からなる。
- ② 変数：Universe、空間とエージェントのそれぞれ属性について指定する。数値または文字列になる。
- ③ ルールエディタ：Universeとエージェントにシミュレーションのルールを書き込む。

- ④ ルール：エージェントの行動ルール、またシミュレーションの結果による観察と解析するため集計方法を書き込む。

## 4.2 シミュレーションモデルにおけるエージェント

本論では、クチコミサイトに登録しているユーザーの購買行動を中心として研究を行うため、ユーザーエージェントの行動を主要な対象とする。まず、artisoc 上に Universe という大きな環境がある。その下で、cosme という名前で作った空間とシミュレーションの結果を集計するための変数を配置していく。空間 cosme の下で product (商品)、user (ユーザー)、review (クチコミ) という三つのエージェントを artisoc 上に定義する。シミュレーションでは、1 ステップを一週間とする。シミュレーションの全ステップ数は52で、すなわち一年間のクチコミサイトにおけるユーザーエージェントの行動を観察する。

### 4.2.1 商品エージェント (Product Agent)

クチコミサイトにおける公開されている商品を代表する。既存の商品エージェントと新しい商品エージェントという二つのタイプに分けている。新しい商品エージェントはシミュレーションのステップ数が12、24、36、48になる際に、マップの真ん中に出てくる。各商品エージェントはシミュレーションの結果を集計する変数も持っている。

@cosme クチコミサイトに関する従来的一年間のデータに基づき、三か月間のごとに新しい商品がランキングに入ることが分かる [@cosme (2017)]。ですから、12ステップに三つの商品エージェントを追加、24ステップに二つの商品エージェントを追加、36ステップに一つの商品エージェントを追加、48ステップに二つの商品エージェントを追加する。新しく追加される商品エージェントにランダムでクラスター変数を設定する (cluster: 1, 2, 3)。

### 4.2.2 ユーザーエージェント (User Agent)

クチコミサイトのユーザーを代表する。シミュレーションでは1ステップごとに10個の新しいユーザーエージェントをマップ上に追加させる。新しく生成されたユーザーエージェントが最初にどの商品を選択するかは、既存のデータ (各クラスターに属する商品のクチコミ投稿率) に基づいて決める。マップに新しく追加されるユーザーがクラスター1に属する商品を選ぶことの可能性を43%、クラスター2を選ぶ可能性を26%、クラスター3を選ぶ可能性を31%とする。その後で、各ユーザー (既存のユーザーも含む) の購買行動は三つの指標に影響を受けるものとする。

#### ① キーワードによる影響

商品の効果に関する六つのキーワードにより重視度は各ユーザーが投稿したクチコミにかかわる各キーワードの出現頻度に依存する (小数)。ユーザーエージェントが個人としてクチコミに影響されるウェイト式(1)の  $h_{(i)}$  は0より大きければこのキーワードにより影響される



程度が大きくなることとする。六つのキーワードによりユーザーエージェントはそれぞれの被影響度を計算し、合計した数値式(1)の  $EF_{(i,k)}$  として使用する。また、六つのキーワードによりユーザーエージェントに影響する基準は商品に対する六つのキーワードの出現確率とユーザー自身からクチコミに対する重視度に基づいて判断される。式(3)の  $\alpha$  はシミュレーションよりコントロールできるパラメータ "keyword\_effected" として設定する。範囲は 0 以上 1 以下とする。

## ② 総合点数 star による影響

各ユーザーが総合点数に影響される程度は各商品の総合点数の平均値にかかわる。商品ごとにクチコミに書かれた総合点数を表す変数 star の合計を得ることで平均値式(5)の  $SP_{(k)}$  を計算し、総合点数がユーザーに影響を与える式(5)の閾値  $\beta$  より大きいならば、ユーザーに影響を与えることとする。 $\beta$  はシミュレーションによりコントロールできるパラメータ "star\_average" という変数に入れる。範囲は 0 以上 7 以下とする。

## ③ ユーザーの年齢による影響

各商品のクチコミ投稿者の人数を年齢層によって整理する。商品に対して各ユーザーが自分になりの年齢層の人数が多いほど購買行動する可能性が高くなることとする。ユーザーの年齢を 7 つの年齢層に分けている (10代: 20以前; 20代前半: 20~24; 20代後半: 25~29; 30代前半: 30~34; 30代後半: 35~39; 40代: 40~49; 50代: 50以後)。各商品によってユーザー自身の年齢がいる年齢層のユーザー数の比率  $U_{(i,n)}$  (式(8)) を計算し、式(8)の  $\gamma$  より大きいならユーザーに影響を与えることとする。 $\gamma$  はシミュレーションよりコントロールできるパラメータ "age\_effected" として設定する。範囲は 0 から 1 までとする。

以上の三つの要素で各ユーザーエージェントの行動は影響される。各要素に対してユーザーの行動に影響させる閾値を計算することで合計した総閾値  $I_{(i,k)}$  (式(9)) のは Universe の下でパラメータ "action\_Index" として設定し、コントロールすることができる。図4.1のようにユーザーエージェントに対して各被影響度は設定した閾値を超えたら購買行動が決まれる。ユーザーが被影響度によって購買した後で、0.91の確率 (@cosme クチコミサイトにおける従来の一年間のデータに基づき、「現品」または「購入品」を選択する投稿率とした) で商品にクチコミを投稿する。

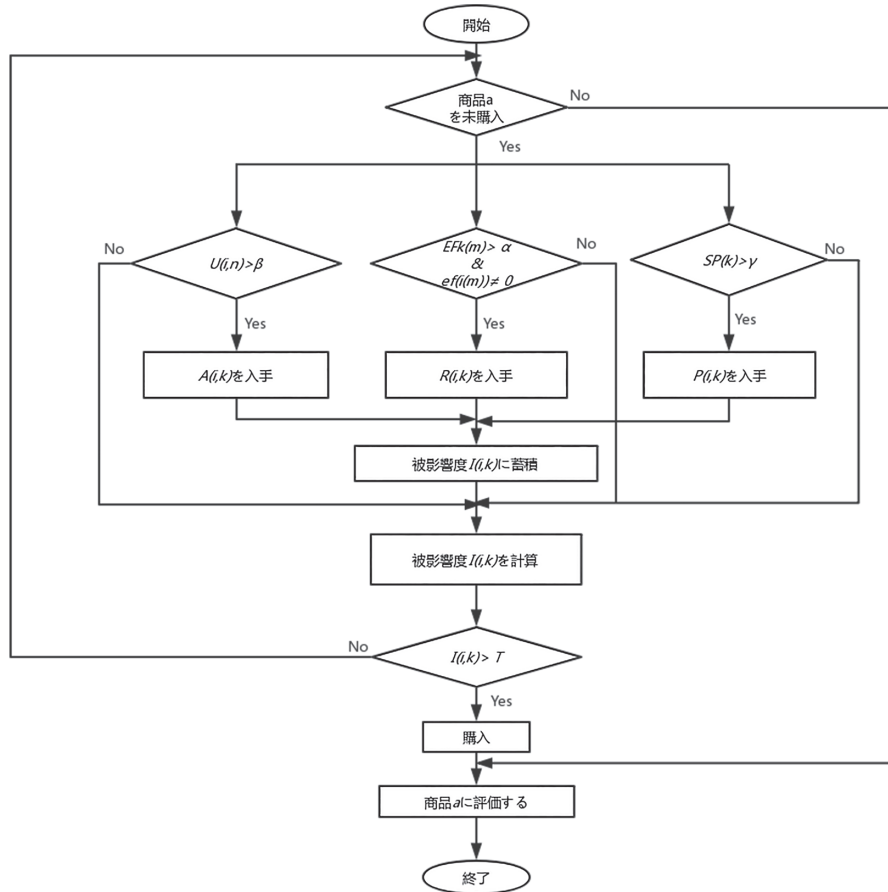


図4.1 ユーザーの購買行動のフローチャート

#### 4.2.3 クチコミエージェント (Review Agent)

クチコミサイトに投稿されたクチコミを代表する。シミュレーションのステップごとに新しいクチコミエージェントをマップの上に追加する。ユーザーエージェントは新しい商品エージェントにリンクを貼っていくところで、クチコミエージェントは同時に生成されてこのユーザーエージェントにつながる。クチコミの変数 "star" はユーザーより商品に対する総合点数である。新しいクチコミエージェントの変数 "star" にランダムで数値を与える。整数型の変数 "keyword" は七つのキーワードの出現頻度である。新しい生成させたユーザーエージェントの変数 "keyword" は既存のデータに基づいて一定の確率で決める。

#### 4.3 マップの出力

artisoc によりシミュレーションを動かせ、空間で三つのエージェントがそれぞれに行動させた結果はグラフとして出力することができる。図4.2のように従来のデータに基づいて初期状態をマップに示している。マップ上では、商品エージェントは真ん中にある数字付きの四角形として存在する。これらの商品エージェントは三つのクラスターに分け、クラスター1の商品を赤、クラスター2の商品を青、クラスター3の商品を黄色につける。既存のユーザーエージェントは青いノードとして、新しい追加されたのは赤いノードとする。マップ上で

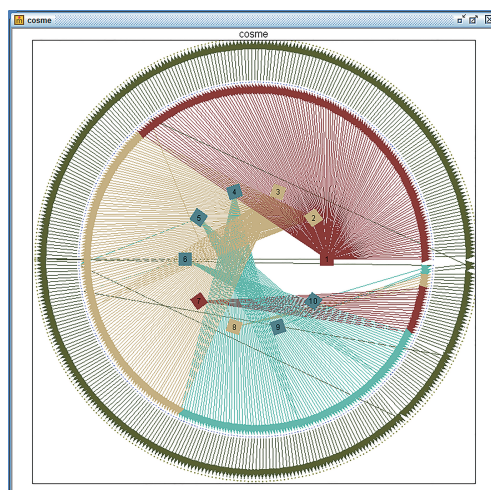


図4.2 空間 cosme の初期状態

一番外にいるノードはクチコミエージェントである。

## 5 シミュレーションのシナリオ

現実のクチコミサイトでは、サイト上の情報とユーザー自身の欲求によってユーザーの購買行動に影響を及ぼしあっている。そのため、クチコミサイトにおけるユーザーの購買行動に至る各要素単独の影響の度合いを現実のクチコミサイトにおける情報だけで知ることはできない。そこで、仮想空間上において各要素がそれぞれにユーザーに影響を与えるように設定し、異なる数値で各指標における影響の及ぼし方の違いを分析する。

シミュレーションにより様々なシナリオを仮想して各要素のパラメータを設定する。この中で、キーワードに関する閾値は商品によりユーザーが重視しているキーワードの出現率（範囲：0～1）とみなす。各閾値の基準値は、商品により半分以上の口コミの中でユーザーが重視するキーワードが出るとキーワードに影響されるとするため、0.5に設定する。総合点数（範囲：1～7）に関する閾値は、商品によりクチコミに書かれた総合点数の平均値とみなす。この基準値は@cosmeクチコミサイトにおける既存の10個の商品のstar数（各商品のクチコミ件数とユーザーが書いた総合点数により計算するもの）の平均値5とする。ユーザー情報として年齢に関する閾値は、商品よりユーザーがいる年齢層の投稿比率（範囲：0～1）とみなす。この基準値は@cosmeクチコミサイトにおける、ユーザーの年齢層により、既存の10個の商品のクチコミ件数に基づきのランキングの上位三番までの投稿比率の平均値を0.69とする。購買行動に至る総閾値  $T$  は総被影響度  $I_{(i,k)}$  により適度に行動が発生させるまでの数値2を設定した。

表5.1のように各閾値をそれぞれの基準値に設定し、三つのシナリオのパターンで設定する。使用したパラメータは3.2節の定義で以下のように示す。各パラメータを用いて、一年間の時期においてシミュレーションを実施し、その間に商品の購買率をクチコミサイトにおけ

る情報の影響による効果とした。

表5.1 パラメータ設定

パラメータ	意味	基準値
$\alpha$	キーワードに関する閾値 (ユーザーが重視するキーワードの出現率)	0.5
$\beta$	総合点数に関する閾値 (商品により評価の総合点数の平均値)	5
$\gamma$	年齢に関する閾値 (商品によりユーザーがいる年齢層の投稿比率)	0.69
$T$	購買行動に至る総閾値	2

### 5.1 シナリオ 0 (基準値により)

ユーザーの購買行動に影響を与える各指標 $\alpha$ 、 $\beta$ 、 $\gamma$ の数値の変化により、シミュレーション結果がどのように変化するかを観察する。そのために、ベンチマークとして表5.1に示した基準値に基づくシミュレーションをシナリオ 0 として設定する。

### 5.2 シナリオ 1 (商品に関するキーワードのみから影響を受ける)

サイトにおける各ユーザーは特に他人が投稿したクチコミに影響を受けやすい状態になるが、他の情報を気にせず購買行動をすることと仮定する。すなわち、ユーザーに対して購買行動が総合点数と年齢に殆どかわらず、主に商品のクチコミに書かれた六つのキーワードに影響されることとする。表5.2のようにシミュレーションによりキーワードに関するパラメータ $\alpha$ を基準値より高い数値0.8を設定する。総合点数に関する閾値 $\beta$ 、とユーザー年齢に関するパラメータ $\gamma$ はそれぞれの基準値を設定する。

クチコミサイトに公開されるクチコミレビューは、大勢のユーザーに対して発信される。多くのクチコミで評価された商品は、情報伝達において大きな役割を果たしているため、サイトで消費者に対して購買の意思決定を行うことの可能性が高いと考えられる。消費者の意思決定過程によると自己の欲求を認識し、情報を探索し、購買前の選択肢を評価して購買するといったように流れていくため、他人のクチコミ内容を購買行動に駆り立てられる心理的メカニズムは自然ではないと考えられる。

表5.2 シナリオ 1 における各パラメータの設定値

パラメータ	設定値
$\alpha$	0.8
$\beta$	5
$\gamma$	0.69
T	2

### 5.3 シナリオ 2 (クチコミに書かれた総合点数からのみ影響を受ける)

サイトにおける各ユーザーはサイトにおける他人が投稿したクチコミのキーワードとクチ

コミに書かれた総合点数より購買行動に影響を受けることと仮定する。すなわち、クチコミサイトに登録している他のユーザーの情報（年齢）を気にせず他人のクチコミ情報から影響を受けやすい状態である。表5.3のように総合点数に関するパラメータ $\beta$ は基準値より高い数値6を設定する。キーワードに関するパラメータ $\alpha$ とユーザー年齢に関するパラメータ $\gamma$ はそれぞれの基準値を設定する。

クチコミサイトのページに公開される商品のクチコミは大量に存在しているため、ユーザーに対して全部のクチコミ内容を一つずつチェックすることは困難である。クチコミに書かれた総合点数とはクチコミ投稿者が商品に対する総合的な評価を表す数値になる。そこで、他のユーザーはクチコミページを見るときに総合点数をチェックし、商品に購買行動を決めるという流れも自然だと考えられる。

表5.3 シナリオ2における各パラメータの設定値

パラメータ	設定値
$\alpha$	0.5
$\beta$	6
$\gamma$	0.69
T	2

#### 5.4 シナリオ3（ユーザー情報のみから影響を受ける）

クチコミサイトにおけるユーザーは自身の欲求によってサイトにおける他のユーザーの情報のみ購買行動に影響を及ぼすことと仮定する。本シナリオでは、表5.4のようにユーザー年齢に関するパラメータ $\gamma$ を基準値より高い数値0.8を設定する。キーワードに関するパラメータ $\alpha$ とクチコミに書かれた総合点数に関するパラメータ $\beta$ はそれぞれの基準値を設定する。

本論で用いるデータは@cosmeサイト上の「口紅・グロス・リップライナー」商品のクチコミ及び投稿者に関する情報とする。サイトページに年齢層によって口紅のランキングが公開されるため、ユーザーにとって購買するときに自分の年齢に合ったものを選ぶことが容易である。シナリオ3では、クチコミサイトにおけるユーザーも同様に、他のユーザーが投稿したクチコミは気にせずユーザーの情報（年齢）だけを気にすることを想定している。すなわち、ユーザーが自分の年齢層によって商品を検索して購買行動に影響を受けることと仮定する。

表5.4 シナリオ3における各パラメータの設定値

パラメータ	設定値
$\alpha$	0.5
$\beta$	5
$\gamma$	0.8
T	2

## 6 シミュレーションの結果

### 6.1 シナリオ 0 の結果

5.1節のシナリオ 0 に基づき、シミュレーションのパラメータを設定した。キャリブレーションのため、シミュレーションを30回で実行させた。シミュレーションの結果は図6.1に示される。横軸はステップ数、縦軸は全商品のクチコミ投稿数になる。

全商品の投稿数から見ると、基準値により影響された効果は約2.5倍増えたことが分かる。図6.2は各クラスターにいる商品の購買数である。この図からクラスター 1 の商品は持続的に一番多くの購買数を持つことが分かる。二番目はクラスター 3 の商品、三番目はクラスター 2 の商品である。

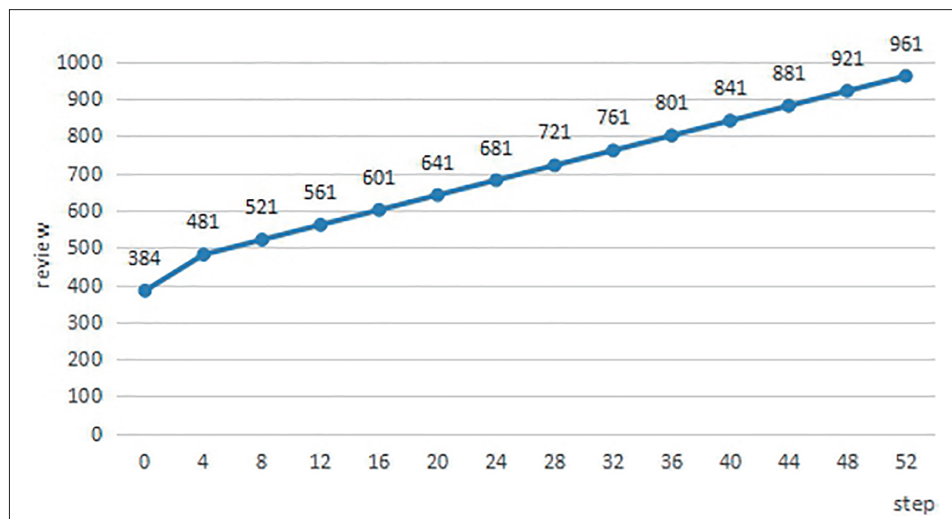


図6.1 基準値により影響される効果（全商品の投稿数）

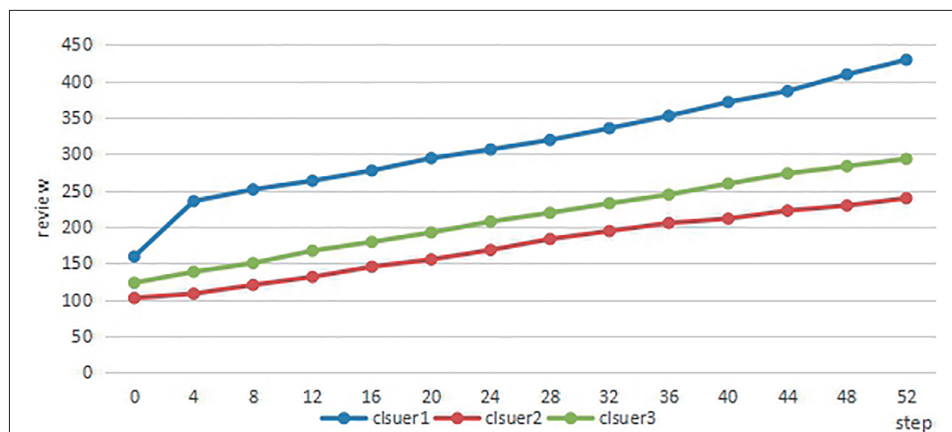


図6.2 各クラスターの商品の投稿数



## 6.2 シナリオ 1 の結果

5.2節のシナリオ 1 に基づきシミュレーションで各パラメータを設定した。シミュレーションの結果は図6.3のように示される。シナリオ 0 と比べると、ユーザーがキーワードに影響を受けた効果が約2.6倍増えたことから、 $\alpha$ を大きくしても投稿数に大きな変化が見られないことが分かる。図6.4は各クラスターにいる商品の購買数である。各クラスターの投稿数はシナリオ 0 の結果と大差が無いことが分かる。

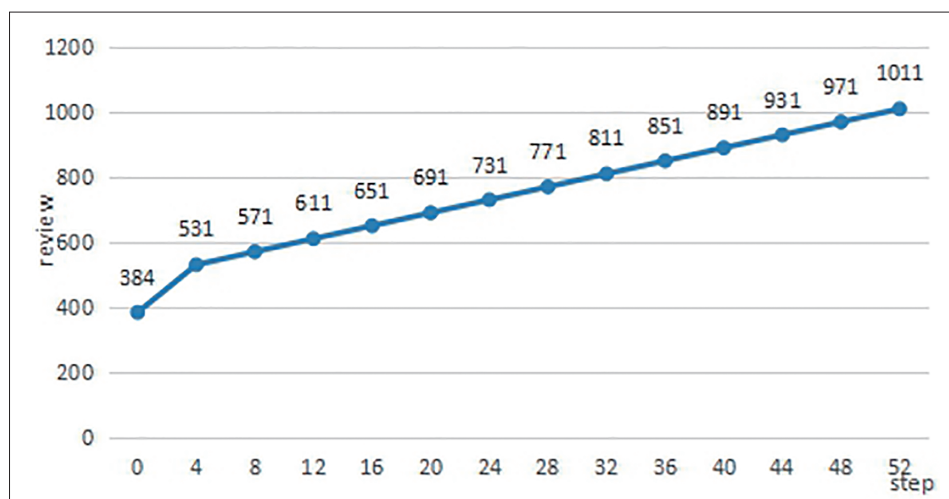


図6.3 キーワードに影響される効果（全商品の投稿数）

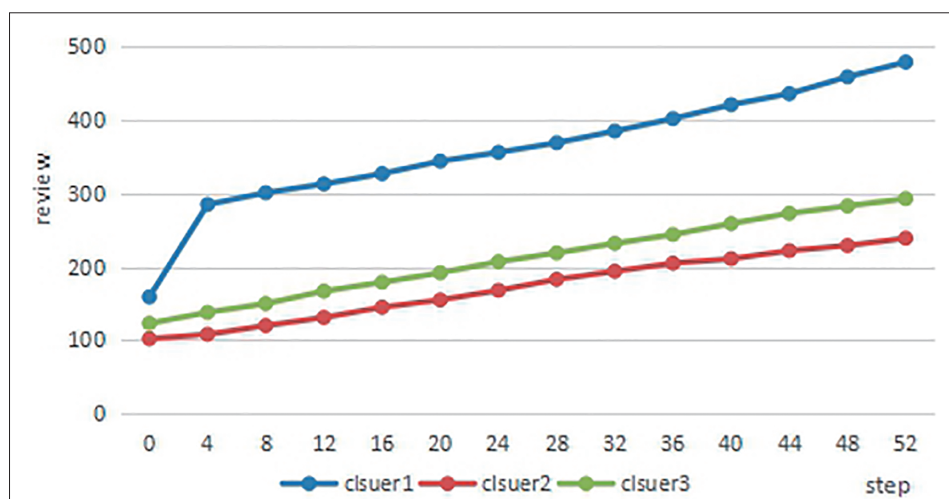


図6.4 各クラスターの商品の投稿数

## 6.3 シナリオ 2 の結果

各パラメータの数値を5.3節のシナリオ 2 に基づき設定し、実行された結果は図6.5のように示される。ユーザーが総合点数に影響を受けた効果が約3.2倍増えた。シナリオ 0 より、影響された効果は見えやすい。各クラスターの商品の投稿数（図6.6）を見ると、シナリオ 0 の投稿数の変化とほとんど同様の結果となった。

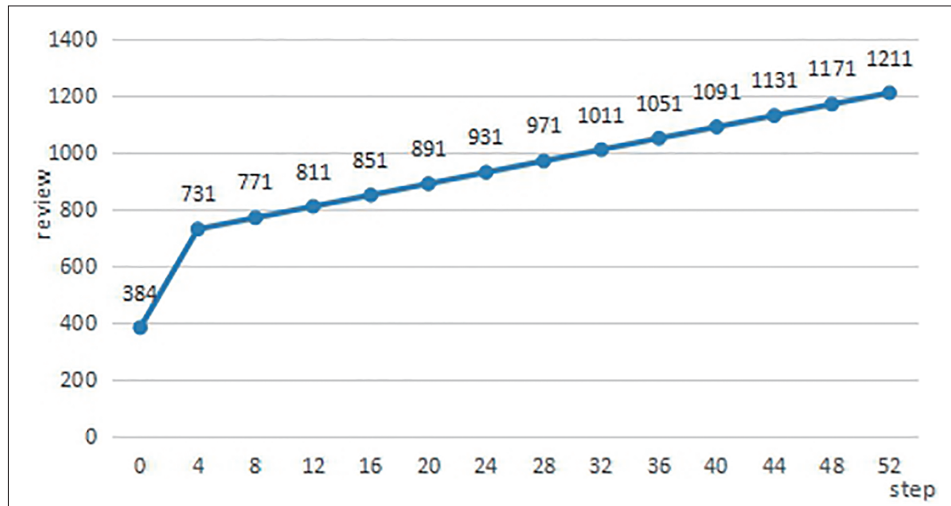


図6.5 総合点数に影響される効果（全商品の投稿数）

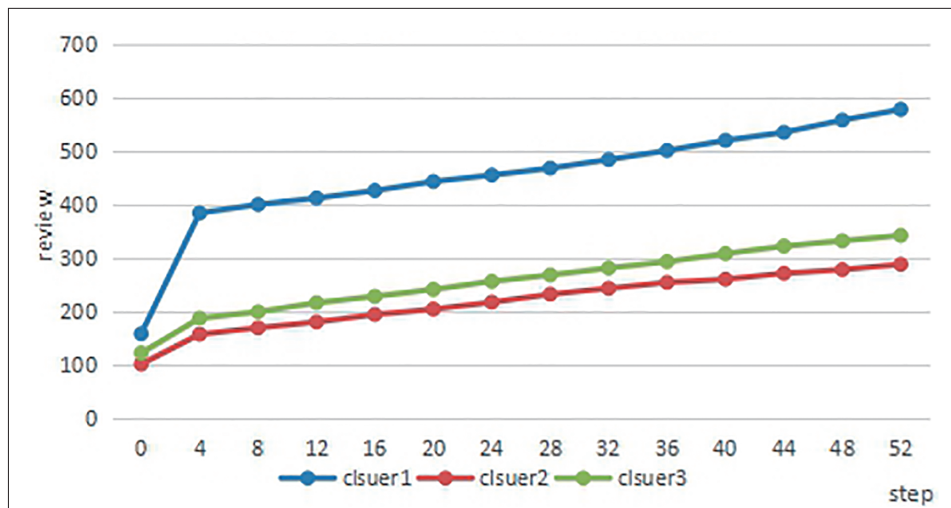


図6.6 各クラスターの商品の投稿数

#### 6.4 シナリオ 3 の結果

各要素の数値を5.4節のシナリオ 3 に基づき設定し、シミュレーションを30回実行させた結果を図6.7に示す。全商品の投稿数から見ると、ユーザーがユーザー情報（年齢）に影響を受けた効果は約3.3倍増えたことがわかる。シナリオ 0 と比較すると、より大きな影響を受けたことが分かる。また、図6.8を見ると、クラスター 0 の投稿数変化と同様に、クラスター 1 の投稿数が一番多くなった。

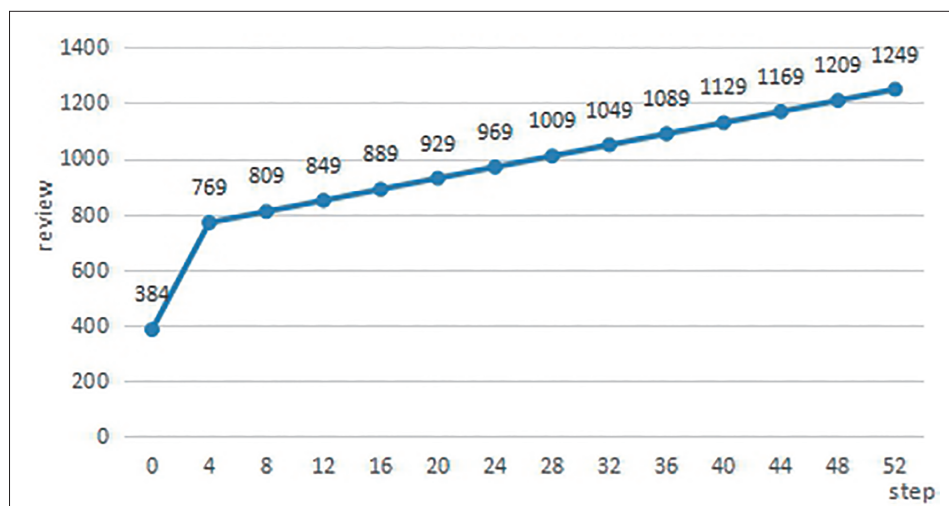


図6.7 ユーザー情報に影響される効果（全商品の投稿数）

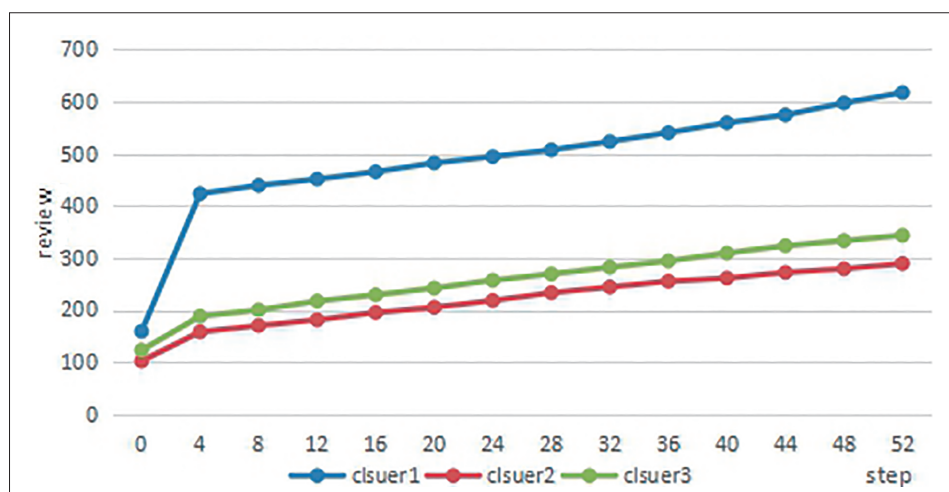


図6.8 各クラスターの商品の投稿数

## 7 最適な「販促強化」方法の探索的発見

本研究で使われる三つの指標により消費者の購買行動に変化が見られることを確認した。本シミュレーションにおいて、企業が直接的であれ、間接的であれ、影響を与えることが出来るとすると、 $\beta$ しかないことは自明である。また、企業の立場からすると、企業が市場に何らかの働きかけをすることにより、自社の売上増加につなげたいところである。本章では、この観点から、 $\beta$ を段階的に変化させることで、投稿数にどのように影響するかを観察する。

ここでは、パラメータ $\beta$ を引数として、それぞれ5回ごとのシミュレーションを実行し、その結果を平均値として求めた（図7.1）。

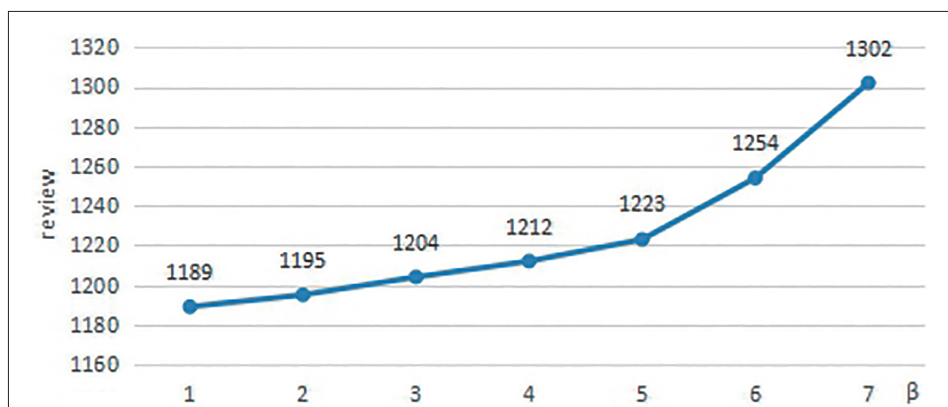


図7.1 総合点数に関する $\beta$ の影響

上の図から見ると、総合点数に関するパラメータ $\beta$ の値が5までは投稿数が線形的な変化を示しているのに対して、それ以降では指数関数的に増加しているように見える。@cosme クチコミサイトのクチコミランキングの順位は各商品の総合点数とクチコミ件数に基づいており、またランキングの上位にいる商品はマスメディアに対して露出度が高く、そのため消費者の購買意欲が高くなるものと思われる。化粧品に対してサンプル、商品抽選など販売促進に力を入れると全体的に総合点数が上がると思われる。一方で、クチコミサイトの情報を活用して商品の「話題性」を盛り上げると、クチコミ件数は短期間で増加するものと考えられる。したがって、企業側からの低いコストでの投資に対して、高いリターンが見込めるのではないかと考えられる。また、図7.1において、企業の投入コストに対して $\beta$ が比例的に変化するものと仮定すると、 $\beta$ が5以上のときに働きかけをする方が効率的であることは明らかである。

しかしながら、企業の投入コストに対する $\beta$ の変化は、現時点では明らかに出来ていない。これについては、継続して明らかにしたいと考えている。

## 8 考察とまとめ

本論では、クチコミサイトにおける情報が消費者の購買行動にどのように影響を与えることかを四つのシナリオでシミュレーションを実施した。そして、どのようなクラスターの特徴を持つ商品がクチコミサイトにおいて消費者の購買意欲を高めるのかについて検討した。そのために、artisoc 上に3重ネットワークモデルを構築し、テキストマイニングで得られたデータベースに基づき、従来のパターンからクチコミサイトにおける情報と消費者の購買行動の関係性に関する可視化を行った。

四つのシナリオで実行した結果を見ると、シナリオ3の結果でユーザー情報がユーザーの購買行動に影響を与える効果が一番大きいことが見える。次に、シナリオ2のクチコミに書かれた総合点数も近い効果を持つことが分かった。どちらのシナリオでもクラスター1の商品が一番多くの購買者を持つことが示された。各商品の購買数を見ると、シナリオ1で影響

された購買率はシナリオ 2 と 3 より変化の幅は大きい。各クラスターに属する商品の購買数を見ると、クラスター 1 に属する商品は既存のデータに基づき長時間でランキングの上位 3 位を占め、他の商品より数十倍のクチコミ件数を持っている。すなわち、クチコミサイトにおけるクチコミランキングの上位にいる商品は持続的に人気を持つことが確認できる。クチコミサイトのランキングページに存在する時間が多いほどクチコミがもたらす消費者の購買行動への効果が大きくなるのではないかと解釈できるように思える。

本論で使用した @cosme クチコミサイトにおける商品に関するデータと表 7.1 のクチコミランキングの上位 10 位の商品は、8 カ月を経過してもまだ五つの商品がランキングの上位 10 位に存在することがわかる（表で灰色の部分）。さらに、一年を経過したときの商品の全クチコミの集計件数を見ると（図 7.1）、「10119418」の商品の増加分が一番大きい。シミュレーションのシナリオ 1 の結果により、この商品を代表するエージェントが多くの購買数を持つことも見る事ができた。この商品を購入または評価したユーザーは、クチコミサイトにおいてクチコミに書かれた効果に関するキーワードにより購買行動に影響を受けたのではないかと考えられる。

表 7.1 @cosme サイトにおけるクチコミランキングの前十位の商品情報

商品 ID	ブランド	名称	クチコミ件数	発売時間	価格
10119418	オペラ	リップティント	1007	2016/11/1	1,500
2913937	ディオール	ディオール アディクト リップ マキシマイザー	411	2007/7/27	3,600
10159925	マキアージュ	ドラマティックルージュ EX	346	2018/11/21	2,800
10133338	フローフシ	LIP38℃ リップトリートメント	219	2017/7/7	1,600
2923643	ちふれ	口紅	229	2018/3/1	350
10159353	ステラシード	ブランピングピンク メルティーリップセラム	312	2018/9/25	1,300
10118017	オペラ	シアーリップカラー N	294	2018/11/30	1,200
10156141	リップザカラー	リップザカラー	179	2018/8/13	600
10083731	セザンヌ	ラスティンググロスリップ	171	2014/9/26	480
10051660	イブ・サンローラン	ルージュ ヴォリュプテ シャイン	153	2013/4/19	4,100

出所：2018年12月31日の @cosme 口紅ランキングより作成 [ @cosme (2018) ]<sup>4)</sup>

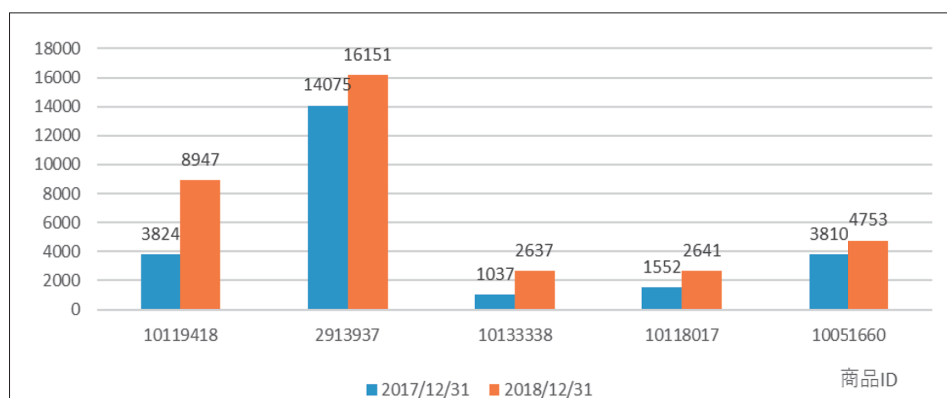


図 7.2 五つ商品の全クチコミ件数

出所：2017年12月31日と2018年12月31日の @cosme 口紅ランキングより作成 [ @cosme (2018) ]

4) ここのクチコミ件数の集計期間は2018/9/27～2018/12/26となる。

クチコミサイトに公開されている情報に影響を受けたことで、クチコミサイトにおける消費者は商品を選択して購買行動を決める。本論では、消費者の購買行動に影響される情報をクチコミ情報とユーザー情報という二つの情報源に分けることで、シミュレーションにより影響される三つの要素を異なる設定で結果を考察した。結果として行動への効果はクチコミサイトにおける総合点数とユーザーの情報により影響を受けることは相対的に大きい。企業側に対してクチコミサイトの情報を効率的に利用すると、低いコストで販促などの活動に力を入れることが可能であり、その結果としてクチコミの総合点数が間接的にコントロールできる可能性は高いと思われる。

クチコミサイトにおけるユーザー向けの情報は多様であるため、現実のクチコミサイトにおける情報と消費者の購買行動とのつながりはもっと複雑なものとなる。クチコミサイトにおける様々な検索の機能を利用することで商品に関する情報のやり取りも可能である。また、本研究では企業が総合点数を間接的にコントロールできるものと仮定したが、具体的にどのぐらいのコストをかけるとどの程度総合点数が上がるのか、すなわちコストと総合点数の関係性の解明については今後の課題としたい。

## 参考文献

- [istyle (2008)] <https://www.istyle.co.jp/>, ユーザーの最新動向, 2008年11月版.
- [istyle (2018)] <https://www.istyle.co.jp/>, 2018年12月26日閲覧.
- [Lee, et al. (1999)] Lee, Myong Jae, Neha Singh and Eric S.W. Chan, "Service failures and recovery actions in the hotel industry: A text-mining approach", Journal of Vacation Marketing, 17 (3), pp.197-207, 2011.
- [@cosme (2017)] 2007年7月から2018年4月まで @cosme "口紅・グロス・リップライナー" 口コミランキング.
- [@cosme (2018)] @cosme, <https://www.cosme.net/>, 2018年4月13日閲覧.
- [岡崎ほか (2014)] 岡崎伸也・風間一洋・篠田孝祐・大向一輝, 「単語のコミュニティ性に基づいたクエリの関連語推薦」, 第8回データ工学と情報マネジメントに関するフォーラム (DEIM2016), C8-4, 2016.
- [唐門ほか (2006)] 唐門準・松尾豊・石塚満, 「リンクに基づく分類のためのネットワーク構造を用いた属性生成」, 情報処理学会論文誌, Vol.49, No.6, pp.2212-2223, 2006.
- [構造計画研究所 (2018)] 構造計画研究所ホームページ, <https://www.kke.co.jp/>, (2018年12月26日閲覧).
- [田中・清水 (2006)] 田中洋・清水聡, 「消費者・コミュニケーション戦略—現代のマーケティング戦略」, 有斐閣アルマ, pp.25-28, 2006.
- [ポーラ文化研究所 (2016)] ポーラ文化研究所, 「女性の化粧行動・意識に関する実態調査2016口紅の流行色—認知と使用状況—」, 2016.
- [山影 (2007)] 山影進, 「人工社会構築指南」, 書籍工房早山, 2007.