

## 階層型クラスタリングを用いた 社会ネットワークと選好の関係性の分析

黒柳風人<sup>†</sup> 島孔介<sup>†</sup> 森山甲一<sup>†</sup> 武藤敦子<sup>†</sup> 松井藤五郎<sup>‡</sup> 犬塚信博<sup>†</sup>  
名古屋工業大学<sup>†</sup> 中部大学<sup>‡</sup>

### 1 はじめに

社会ネットワークとは、1つ以上の関係性を結んだ社会的構造であり、Ronald S Burt[1]が提唱した構造的空隙のように、その様相は個人の社会関係資本であるとされている。人間関係を記述するネットワークは様々だが、近年、その分析が盛んに行われており、桑島[2]は個人のミクロな社会ネットワークと消費者の購買行動の関係性を示した。

また細木ら[3]は行動特性の一例として選好を扱い、個人の社会ネットワーク構造と音楽選好の関係性を分析した。しかし、細木ら[3]の研究では社会ネットワークを特徴量とした音楽選好の分布の確認までしか行っておらず、具体的にどの音楽ジャンルが近いネットワーク構造を持つのか不明であった。そこで本研究では社会ネットワーク構造の似通った選好分野の抽出及び分析する手法を提案し、実際に稼働しているSNSに手法を適用させる実験を行った。

### 2 ネットワーク属性と選好の分析

細木ら[3]は個人とその隣接者から成る社会ネットワークと選好の関係性を分析するため、次の通り、抽出した属性値を分割し、主成分分析を行った。1. 個人の選好と社会ネットワークの属性値を抽出。2. 属性値を順に並べて分割。

3. ある選好における各属性の分割されたグループ上での分布割合を計算し、得られた結果に主成分分析を行い、結果を取得することである。

本論文では以下に述べる通り、この方法を一般的に定める。

ここでは社会ネットワークのノード集合を  $N$  とし、 $n \in N$  の属性集合として  $U$  を考える。また  $n$  は二値の選好を持つものとし、選好の集合を  $B$  とする。

属性  $u \in U$  に注目し、その値の範囲を  $k$  区分に分け、各区分 ( $k$  分位) に属するノードの選好  $b \in B$  の分布割合を考えたい。 $u$  の値に基づき  $N$  は昇順に  $u(n_1), \dots, u(n_{|N|})$  ( $u(n_i) \leq u(n_{i+1})$ ) と整列しているとする。すると、第  $l$  番目の区画には、 $n_{\lceil l/kN \rceil}$  から  $n_{\lceil (l+1)/kN \rceil}$  まだが含まれるが、境界における属性値  $t_l = u(n_{\lceil l/kN \rceil})$ ,  $t_h = u(n_{\lceil (l+1)/kN \rceil})$  は、これと同じ値を持つノードが複数存在し得るため、これを適切に扱う必要がある。そのため、次の通り代  $l$  区分の境界部分の割合を求める。

$$P_{low}(l) = \frac{l}{k} - \frac{|N_A|}{N} \quad (N_A = |\{n | u(n) < t_l\}|)$$

$$P_{high}(l) = \frac{l}{k} - \frac{|N_B|}{N} \quad (N_B = |\{n | u(n) > t_h\}|)$$

すると、この区分に含まれるノードの選好  $b$  が 1 となる割合  $X_l(b)$  は次のとおりである。

$$X_l(b) = P_{high}(l)\{b(n) | u(n) = t_l\} + \\ \{b(n) | t_l < u(n) < t_h\} + \\ P_{low}(l)\{b(n) | u(n) = t_h\}$$

$u$  の値に基づく第 0 から第  $k-1$  の各  $k$  分位の

Analysis of relationship between preference and social network using hierarchical clustering

<sup>†</sup> Nagoya Institute of Technology

<sup>‡</sup> Chubu University

選好  $b$  の割合を分布割合ベクトル  $Var(b)$  として以下の通り定める。

$$Var(b) = \left( \frac{X_1(b)}{b(n)}, \dots, \frac{X_k(b)}{b(n)} \right) \quad (1)$$

**階層型クラスタリングの適用** 得られた  $k$  次元ベクトルに群平均法を用いて階層型のクラスタリングを行った後、シルエット分析によってクラスタ数を決定し、分析結果を得る。

### 3 実験

提案手法を実際に稼働している SNS に適用させる。なお本論文では取得した次数中心性、エッジ数、島数の  $k$  次元ベクトルを一つに結合し、主成分分析を行った。

#### 3.1 利用データ

音楽ストリーミングサービス Deezer[4](ルーマニア国内限定) に手法を適用させる。Deezer はユーザが他のユーザをフォローする機能を持ち、各ユーザは一つ以上の好きな音楽ジャンルを設定している。フォロー、フォロワー間の情報から社会ネットワークを抽出し、好きな音楽ジャンルの設定から選好分野を取得する。

#### 3.2 実験結果

属性値の分割数は 4 とし、得られた主成分分析結果にシルエット分析を適用させたところ、最適な分割数は 5 と判定された。図 1 はその結果である。同一の枠線で囲まれた内部は選好者

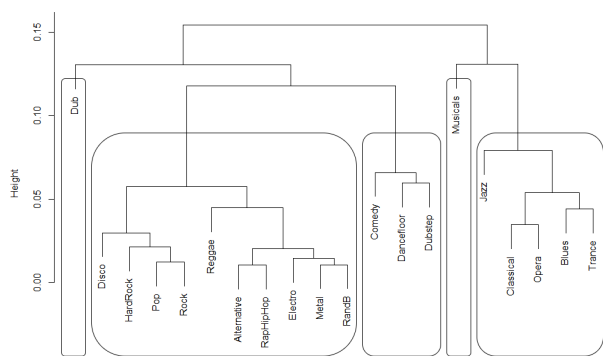


図 1 社会ネットワーク構造を基準に作成されたデンドログラム

の社会ネットワーク構造が類似している分野であることを示している。

左から 2 番目のグループには Rock が所属しており、その派生の HardRock や Alternative-Rock、Metal も見られ、音楽ジャンルの類似性が社会ネットワークに表れていると言える。またその考え方であれば、様々な系統の音楽ジャンルが混同している Musical が単離している点も理解できる。

### 4 まとめ

本研究では、社会ネットワークの構造と選好との関係性を分析する手法を提案し、実際に稼働している SNS、Deezer に適用させて音楽分野と社会ネットワーク構造の関係性を可視化した。

Rock、HardRock、AlternativeRock、Metal 等の派生形が同じグループにまとまっており、音楽ジャンルの選好と社会ネットワーク構造の傾向には関連性があると考えられる。

### 参考文献

- [1] Burt, Ronald S. (1995). Structural Holes: The Social Structure of Competition. Cambridge: Harvard University Press
- [2] 桑島 「関係性から見る購買行動 — ネットワーク分析を用いて—」 MMRC Discussion Paper No144
- [3] 細木 他 「行為者の選好と社会ネットワークの構造的属性との関係」 情報処理学会研究報告 2023-MPS-142 13 号 1-5
- [4] Deezer 「Brand Partnerships」 Who we are (<https://www.deezer-brandpartnerships.com/en/who-we-are/>)