

メンバー情報に基づく バンドネットワークの構築と利用

吉谷幹人[†] 宇佐美敦志[†] 浜中雅俊[†]

本稿では、バンドとミュージシャンの所属関係を表すネットワークを Web 上から自動的に構築し、バンドの推薦や分類に利用する方法を述べる。バンドとミュージシャンの関係はバンドのメンバーチェンジやゲスト出演など、複雑なネットワーク状の構造になっている。本研究では、Google の検索 API を用いて収集した Web ページからパターンマッチングによりメンバー名を自動抽出しネットワークの構築をおこなった。さらに、構築したネットワークを用い、メンバー名の重複数からバンド間の関係の強さを定義し、バンドの推薦や階層的クラスタリングへの応用を試みた。

Construction and Application of Band-Network using Member Information

Mikito Yoshiya[†] Atsushi Usami[†]
and Masatoshi Hamanaka[†]

In this paper, we describe the technique for automatically constructing the network which indicates the relationship between bands and musicians from the Web pages. And we also present the applications of the network for the recommendation and the classification of bands. The relation between bands and musicians make complex network structures according to the repetition of member change and appearance of the guest musician. We constructed the network by automatic acquisition of member name from Web pages collected by using Google Search API. In addition, we defined strength of the relations between bands from the number of duplication of members by using the constructed network. And we attempted to construct systems of the recommendation and the hierarchical clustering of bands.

1. はじめに

本研究では、Web 上から自動収集したメンバー情報に基づきバンド間の関係を表すネットワークを構築し、そのネットワークをバンドの推薦や分類などに応用する手法を提案する。ミュージシャンとバンドの関係を結んでいくと複雑なネットワークが構築される。これは、バンドのメンバーチェンジやゲスト出演などにより、ミュージシャンはしばしば複数のバンドと関係を持つためである。本研究では、バンドとミュージシャンの所属関係を線で結ぶことで生まれるネットワークをバンドネットワークと呼ぶこととし、このネットワークを利用した新たなアーティストの推薦手法を提案する。本手法の特徴は、バンドのメンバーという事実に基づいた関係を利用するという点である。事実に基づいた関係性を利用することで、好きなミュージシャンと関係があるアーティストを推薦することが可能となる。

バンドネットワークを構築するためには、Web 上のページから自動的にメンバー情報の収集をおこなう必要がある。メンバー情報の収集には、まず Google の検索 API を用い、「バンド名 + band + member」というクエリーで検索し、バンドごとに Web ページを 50 ずつ収集する。次に、収集した Web ページから N-gram で切り出した単語列のうち、人物名である可能性が高い文字列を抽出する。さらに、ミュージシャン名と楽器名の出現パターンに注目し guitar や drum などの楽器名と特定のパターンで出現する人物名をミュージシャン名として抽出する。最後に、同じ URL、同じパターンで出現する名前をメンバー名とみなすことで、不要な文字列の除去をおこなう。上記の処理により実際にレコード会社の Web サイト[a]に登録されている 354 バンドを対象としてメンバー名を収集しバンドネットワークを構築したところ、3793 人のミュージシャン名を収集できた。さらに、収集したメンバー情報を解析したところ、全バンド中 250 バンドがミュージシャンを介してつながっていることがわかった。

構築したバンドネットワークの利用法の提案として、関係の強いバンドを探索するシステムの構築と、バンドのグルーピングを行った。まず、注目する 2 つのバンドを合わせた全体のメンバー数と、共通するメンバー数から関連度を定義した。そして、注目する 1 つのバンドから各バンドとの関連度を評価し、関連度が高い順にバンドを提示するシステムを構築した。その際、例えばギターの関係性が強いバンドを探したいという場合にはギターに重みを置くなど、担当楽器ごとの重みの指定を可能にした。これによりユーザの楽器の好みを反映した探索を行えるようになった。バンドのグルーピングでは特定のバンドに注目せず、全体のバンド集合から最も関連度の高いバンド同士を順次グループとして結合する。こうして得られたグループは階層構造を持つため、ユーザはこれを浅い階層から辿っていくことで、好みのバンドを探ることが可能となる。提案手法を利用することにより、バンドのメンバー情報という事実に基づいたバンド探索を実現した。

[†] 筑波大学大学院 システム情報工学研究科
University of Tsukuba, Graduate School of System and Information Engineering
a HMV ONLINE - <http://www.hmv.co.jp/>

2. 関連研究

Web 上からメンバー情報を抽出しアーティストの推薦や分類に利用する本研究では、楽曲推薦や Web マイニングの分野において関連する研究が様々ある。

2.1 楽曲推薦手法

従来の楽曲の推薦や検索の研究では、楽曲間の音響的類似度に基づく手法が提案されている。文献[1,2]では、楽曲の音響信号からリズムや音色などの特徴を抽出し、特徴量が似ている楽曲をユーザに推薦することを可能としていた。しかし、音響的特徴に基づいて楽曲を推薦しても、嗜好にあわない可能性がある。なぜなら、ミュージシャン個人が好きという理由で楽曲を選択するユーザも存在するからである。

一方、協調フィルタリングを用いた楽曲推薦手法[3]では、他のユーザの楽曲に対する評価を利用することで推薦を行うため、大勢の人が評価をおこなったメジャーな楽曲は推薦されやすく、評価が少ないマイナーなアーティストの楽曲については推薦が困難という問題があった。

そこで本研究では、メンバー情報により構築されるバンドネットワークを用いた推薦手法を提案する。メンバー情報を利用することで、バンドのメンバー個人が好きユーザが、その人物の楽曲をもっと聞きたいと思ったとき、バンドに所属したという事実に基づいて新しいバンドを提示することができる。また、マイナーなバンドであってもメンバーの入れ替わりや他のバンドとの交流した経歴が存在すれば推薦をおこなうことができる。

2.2 Web マイニング

Web 上の情報を利用する研究として、文献[4]では研究者同士の関係の強さを名前の共起から求め、人間関係を表すネットワークを構築している。さらに、文献 [5]では、アーティスト名と共起している単語の頻度をアーティストの特徴量とし類似度を計算している。これらの手法では、名前のリストさえあれば、関係性を表す大規模なネットワークを構築することができる。しかし、本研究で求めるバンド間のネットワークは、バンドとメンバーの存在する関係を求める必要があり、文献[4,5]のような名前や単語の共起だけでは実際に関係があるかどうかは判定できないという問題がある。

また、特定のサイトの構造に特化することで情報を抽出する Web ラッパーを自動生成する研究[6]もおこなわれている。バンドのメンバー情報に関しては、Wikipedia[b]等の Web サイトを利用することで多くのバンドの情報を知ることができる。しかし、Wikipedia は有志により情報が書き込まれるため、個々のバンドにより Wikipedia 上のページの情報の詳しさには差があり、ページ自体がないマイナーなバンドも多く存在する。そのため、Wikipedia 等の特定のサイトのみを対象としてメンバー情報を収集すると、十分な情報を集めることができないバンドが出てきてしまう。

一方、文献[7]は、Web 上から集めた複数のページから、メンバー名と演奏する楽器名とのパターンマッチングにより、メンバー名の抽出を行う手法を提案していた。しかし、文献[7]で提案された手法では、名前の出現頻度に閾値設けることで名前のフィルタリングをおこない収集精度を高めようとしていたが、実際にはメンバーである名前も除去してしまい、あまり多くのメンバー名を集めることができないという問題があった。

本研究では、複数の Web ページからパターンマッチングを用いることでメンバー名を抽出する。その際、メンバー名は同じ URL 中に同じパターンで出現する可能性が高いという事実から、パターン共起を判定することで実際にはメンバーでない人物の名前や単なる文字列の除去をおこなう手法を提案する。

3. メンバー名の収集

バンドネットワークを構築するために Web 上から自動的にメンバー情報を抽出する。メンバー名の抽出は以下の四つの段階から成る (図 1)。まず、Web ページの収集部では Google の検索 API を用いてバンドに関する Web ページを収集し、前処理をおこなう (3.1 節)。次に、人物名抽出部では、簡単な文字列の判定により人物名の発見をおこなう (3.2 節)。そして、ミュージシャン名判定部では、楽器名との出現パターンによりミュージシャン名かの判定をおこなう (3.3 節)。最後に、不要ワード除去部では、実際にはメンバーでない人物の名前や、単なる文字列などの不要な要素を除去する (3.4 節)。なお、本手法は英語表記のバンド・Web ページを対象としている。これは海外のバンドではメンバーチェンジが多く行われることから複雑なネットワークを構築することが可能であると考えたためである。

3.1 Web ページの収集部

Web ページの収集には Google の検索 API を利用する。B というバンド名のメンバーを抽出するとき、B の後に次のような検索ワードを加えたクエリーで検索し、Web ページを収集する。

- A) "B" + music
- B) "B" + band
- C) "B" + band + member
- D) "B" + band + lineup

予備実験の結果、検索ワード C) が最も精度が高くなることがわかった。そこで、Google API で「"B" + band + member」というクエリーで検索し、検索結果の上位 50 件を収集するとし、検索結果が 50 件に満たないバンドに関しては、次に精度の高かった検索ワード D) の「band + lineup」を用い検索し、足りないページを補った。

Web ページの収集後、前処理として、収集したページの HTML のタグを全て取り除く。また、ミュージシャンの名前は「Björn "Speed" Strid」のようにステージネームをダブルクォーテーションではさむ形で出現する場合がある。ステージネームの表記の有無で、同一人物が他の人物であ

b Wikipedia - <http://www.wikipedia.org/>

ると判断されるのを防ぐために、15文字以内のダブルクォーテートではさまれた文字列をクォーテートごと全て削除する。

3.2 人物名抽出部

収集したページから、人物名である可能性が高い文字列を抽出する。まずテキストから N-gram を利用し単語列を抽出する。ほとんどの人物名は2から4単語で構成されているため、Nは2, 3, 4とする。また、人物名の多くはファーストネーム、ミドルネーム、ラストネームの先頭文字が全て大文字である。そこで、先頭文字が大文字でない単語を持つ単語列は全て削除する。さらに、人物名に使用される文字は限られているため、抽出した単語列で、ドット[.]、シングルクォーテート[']、ハイフン[-]以外の記号や数字を含んでいるものはすべて削除する。

また、バンドやミュージシャンと関係が深いページには、楽器名や member, review, lyric などの特定の単語が多く出現するため、これらの単語を StopWord とし、出現する単語列は削除する。そして、残った単語列を人物名である文字列の集合とみなす。

3.3 ミュージシャン名判定部

楽器名との出現パターンにより人物名集合の中からミュージシャン名を抽出する。ミュージシャン名は担当の楽器名と共に出現することが多い。本手法では、この性質を用いて、ミュージシャン名の判定を行う。調査の結果、以下の三つのパターンが最も多く出現することがわかった。

- (a) 役割名 名前
- (b) 名前 (楽器名)
- (c) 名前 - 楽器名

ここで(b)(c)の楽器名は, guitar, drum など主にバンド中で使用されるものとした。また(a)の役割名とは guitarist, drummer など、楽器名に人を表す修飾子が付いたものである。

収集した Web ページ上で、(a)(b)(c)のパターンのいずれかとマッチする人物名をミュージシャン名として判定する。

3.4 不要ワード除去部

ミュージシャン名集合には、実際にはバンドメンバーではない人物の名前も含まれている。例えば、交流が深いだけで、たまたま話題に出た人物や、偶然パターンにマッチした単なる文字列などである。そこで、ミュージシャン名の集合から、実際のメンバー名でないものを除去する。まず Document

Frequency を計算し、メンバー名である可能性が十分に低いとみなせる名前を除去する。次にバンド内での中心的な人物である重要メンバーを見つけ、その重要メンバーとのパターンの共起を判定することにより、メンバー名の抽出を行う。

3.4.1 Document Frequency の判定

収集した Web ページ中でミュージシャン名が何ページに出現したかを表す DF (Document Frequency) を計算する。このとき、DF が1、つまり1ページにしか名前が出現していないものは、偶然話題に出たか、単なる文字列であると考えられるため、実際のメンバー名ではないとみなし除去する。

3.4.2 重要メンバーの抽出

実際にメンバーである人物を最低二人見つけ、そのバンドでの重要メンバーとする。まず、収集した Web ページ上でミュージシャン名が(a)(b)(c)のパターンに何回マッチするかを計算し PF (Pattern Frequency) とする。次に PF の最大値が1になるように正規化する。その後 PF>0.7 となるミュージシャン名を重要メンバーとする。PF>0.7 という閾値は、予備実験上で実際のメンバー名を85%の精度で見発できることを確認した上で決定した。メンバーが一人しか見つからなかった場合は、PF の値が高い上位二人を重要メンバーとする。

3.4.3 パターンの共起による段階的抽出

少数の重要メンバー名を用い、メンバー名を段階的に抽出していく。同じバンドのメンバーは楽器名と共にリスト形式で書かれるなど、同じページに同じパターンで出現することが多いため、次のようなメンバー名抽出アルゴリズムを用いる。メンバー名抽出アルゴリズムでは、まず、重要メンバー2人以上と同じ URL、同じパターンで出現するミュージシャン名をメンバー名と確定する。その後、すでに確定したメンバー名と、まだメンバーと確定していないミュージシャン名との URL、パターンを比較し、確定メンバー名を段階的に増やしていく。

[メンバー名抽出アルゴリズム]

1. 抽出したミュージシャン名に対し、どのページ中でどのパターン[(a)(b)(c)]にマッチするかをタプル (Name, URL, Pattern) として全て記録する。
2. 全てのタプル (Name, URL, Pattern) を同じ URL, Pattern をもつもので分割し、n 個のミュージシャン名集合の部分集合 O_i ($1 \leq i \leq n$) を得る。
3. 確定メンバー集合 M を空にする。
4. M に重要メンバー (二人以上) を加える。
5. 新規メンバー集合 NM を空にする。
6. M のメンバーが二人以上含まれる O_i で、M に含まれない名前があったら NM に追加する。
7. NM を M に追加し 5 に戻る。このとき NM が空ならば終了する。M を最終的なメンバー名とする。



図1 メンバー名抽出処理の流れ

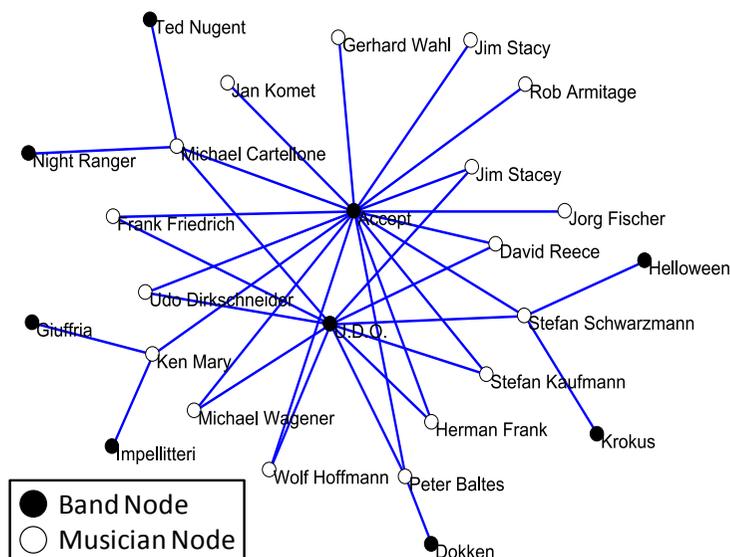


図 21 バンド周辺のネットワーク構造

3.5 バンドネットワークの構築

上記の手法で、実際にレコード会社の Web サイトにあるヘヴィメタルのベストセラーアーティスト 354 バンドのメンバー名を収集しバンドネットワークを構築した。メンバー名収集の結果 3793 人のミュージシャン名が抽出できた。

構築したバンドネットワークでは、ミュージシャン名が 3793 人であるのに対し、バンドとミュージシャンの所属関係は 4727 であった。このことから、複数のバンドに所属しているミュージシャンが存在することが分かる。また、ネットワーク解析ツール Pajek[c]で解析したところ、354 バンド中 250 バンドがメンバーの情報を辿って行けばどこかでつながっているということが分かった。また、図 2 は一つのバンド (Accept) を中心として、そのバンドとメンバーを共有しているバンドの関係をネットワーク状に可視化したものである。中心としたバンドが、メンバーのノードを介することにより、複数のバンドとのつながりを持つことが確認できる。

その反面、93 バンドが他のバンドとメンバーを共有せずにつながりを持つことができなかった。この問題を解決するためには、作曲、録音、編集に関わったスタッフやプロデューサーなどの人物、またはアルバムのレーベルなどの情報を利用してネットワークを構築することが考えられる。これにより孤立した 93 バンドに対してもつながりを持たせることが可能になると考えている。

c Pajek - <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>

4. バンドネットワークの利用

ここでは、バンド間の関連度の計算方法を定義し、構築したネットワークの利用方法を提案する。

バンドのつながりには、直接メンバーを共有してつながるものと、他のバンドを介して間接的につながるものがある。図 2 の例では、2 つのバンド「Accept」と「Giuffria」の間でメンバー「Ken Mary」を共有してつながっている。また、バンド「Accept」と「U.D.O.」に注目すると、メンバーの「Udo Dirkschneider」、「Wolf Hoffmann」をはじめ、11 名ものメンバーが共通しており、つながりが強いことがわかる。

一方、バンド「U.D.O.」と「Giuffria」の間では、直接共有するメンバーはいないが、「U.D.O.」は「Accept」と関連が強いので、「Giuffria」と「U.D.O.」の関連もある程度強いものと考えられる。そこで、このように他のバンドを介してバンド同士が関係する場合についても関連度を定義する必要がある。

4.1 メンバーを直接共有するバンド間の関連度

関連度は以下のような特徴を持つべきである。

- バンド間で共通するメンバーが多いほど、関連度は高い。
- バンドを構成するメンバー総数が多いほど、関連度は低い。

これらを満たすよう、バンド A, B のメンバー集合 A, B を用いて、関連度 R_{AB} を次の(4-1)式のように定義する。また、 A, B は空集合ではないため、(4-2)式が成り立つ。

$$R_{AB} = \frac{|A \cap B|}{|A \cup B|} \quad (4-1)$$

$$0 \leq R_{AB} \leq 1 \quad (4-2)$$

4.2 他のバンドを介して間接的につながりを持つバンド間の関連度

図 3 はバンド間の共有メンバーによるつながりを一本の線で簡略化して表した図である。各バンド間の関連度 $R_{AB}, R_{BC}, R_{CD}, R_{BE}, R_{ED}$ は、(4-1)より求まる。

このとき、バンド A と、バンド B を介して関係を持っているバンド C との関連度 R_{AC} は、明らかに R_{AB} や R_{BC} と比べて小さい。そこで、 R_{AC} は $R_{AC} < R_{AB}$ かつ $R_{AC} < R_{BC}$ を満たす必要がある。したがって、 R_{AC} を R_{AB} と R_{BC} の和や相加平均、相乗平均とすると、この条件を満たさない場合が発生してしまうため、適切ではない。したがって、本稿では、次式のように経路する関連度の積として R_{AC} を定義する。

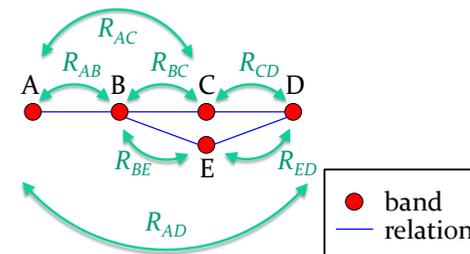


図 3 バンド間の関連度

$$R_{AC} = R_{AB} \times R_{BC} \quad (4-3)$$

図3のバンドAとバンドDのように、途中で複数の経路が考えられるような場合の関連度 R_{AD} について考える。(4-3)式と同様に考えると、 $R_{AD} = R_{AB} \times R_{BC} \times R_{CD}$ と、 $R_{AD} = R_{AB} \times R_{BE} \times R_{ED}$ の二通りが考えられる。今回の場合、関係の強いところをユーザに提示するのが目的であるため、考えられる関連度のうち、最大値を取るものを採用することとする。すなわち、

$$R_{AD} = \max(R_{AB} \times R_{BC} \times R_{CD}, R_{AB} \times R_{BE} \times R_{ED}) \quad (4-4)$$

とする。同様に、間の経路が複数考えられるバンド間の関連度は、関連度を最大とする経路のものを採用することとする。

4.3 関連度を最大とする経路の算出

解析するバンド数が増えてくると、間接的なつながりを持つバンド間におけるつながりの経路は膨大なパターンが考えられ、現実的に計算が不可能になってしまう。そこで、本稿では、バンドのネットワークを図3のようにバンドを頂点、共有メンバーによるつながりをエッジ、バンド間の関連度をエッジの重みとして考えることでグラフ理論を適用し、最短経路を効率よく計算するアルゴリズムであるダイクストラ法[8]を応用した。本手法では、エッジの重みが頂点間の距離ではないため、表1のような点を変更して適用した。

表1 ダイクストラ法と本手法の変更点

	ダイクストラ法	本手法
出発地点の値	0	1
各辺の値	頂点間の距離	頂点間のつながりの強さ
各頂点の値	出発地点との最短距離 (0~)	出発地点とのつながりの強さ (0~1)
各頂点の値の算出方法	ひとつ前の各頂点と辺の値の 和の最小値	ひとつ前の各頂点と辺の値の 積の最大値

4.4 関係の強いバンドの探索

前項のダイクストラ法の応用により、注目するバンドからつながりを持つ各バンドに対して関連度を求める。そして、関連度の高い順にユーザに提示することにより、指定したバンドに関係の強いバンドを検索することが可能となる。この際、図4のように、ボーカルやギターなど、各メンバーの担当パートによって、重みづけをできるようにすることで、ユーザの好みを反映した、バンド間のつながり探索を可能とした。

4.5 バンドのグルーピング

特定のバンドに注目せずに、あるバンドの集合を関連度によってグルーピングする方法を提示する。ユーザは、色々なバンド同士の関係を見て楽しむことができる。

4.1 項で述べたメンバーの直接共有による関連度を、つながりのある全てのバンド間において求め、そのうち R が最大となる2つのバンドをひとつのグループとして結合し、関連度の再

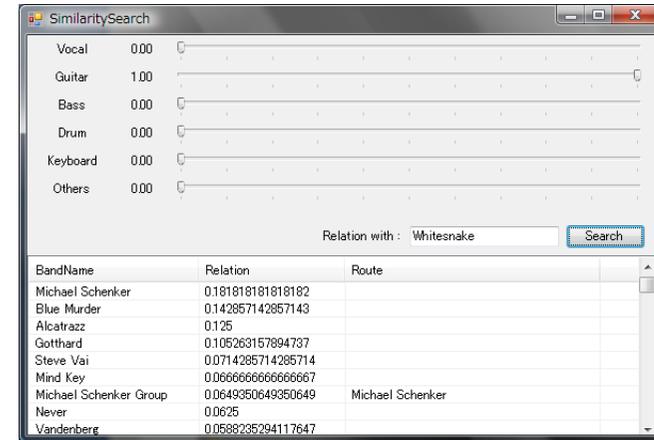


図4 関係の強いバンドの探索システム

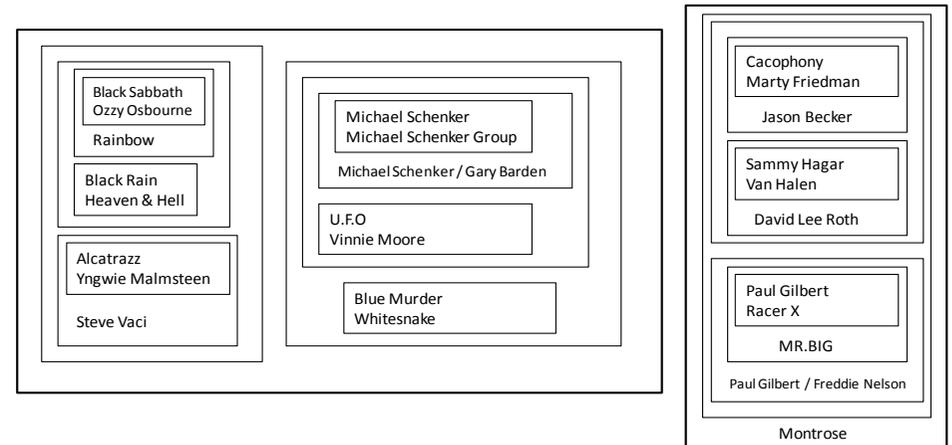


図5 バンドのグルーピング結果

計算を行う。同様の処理を、関連度の最大値がある閾値 R_{min} を下回るまで再帰的に繰り返すことにより、バンドのグルーピングを行う。

図5は実際に収集したデータからグルーピングを行った結果の一部である。

5. メンバー名収集手法の評価

メンバー名収集手法の精度を評価する。評価の対象としてバンドネットワークの構築に用いた354バンドから50バンドをランダムに選んだ。また、Wikipediaを用いなくてもバンドメンバーがどれくらいの精度で抽出できるかを調べるために、収集対象のページはWikipediaを含まないようにした。評価には、再現率、適合率、およびその両者を統合したF値を用いた。収集したメンバー名集合をM、正解のメンバー名集合をCとすると、再現率R、適合率P、F値は次式で表される。

$$\text{再現率(R)} = \frac{|M \cap C|}{|C|}, \quad \text{適合率(P)} = \frac{|M \cap C|}{|M|}, \quad \text{F値} = \frac{2PR}{R+P}$$

5.1 正解データの作成

正解データとして50バンドのメンバー名を手作業で収集した。このとき、現在のメンバー以外にも、すでに脱退しているメンバーやゲストミュージシャンなど、一度でもバンドと関わりを持ったミュージシャン名を全て正解とした。その結果623のミュージシャン名が集まった。

5.2 収集精度

メンバー名抽出アルゴリズムの有効性を示すため、次の四つの処理を行ったときの精度を比較した(表2)。

1. **処理なし**
適合率を上げる処理を行わず、パターンにマッチした人物名は全てメンバー名とした。
2. **DF=1を削除**
1ページのみしか出現していない名前はメンバー名に含まないようにした。
3. **DFに閾値を設定**
全ミュージシャン名のDFの最大値が1になるように正規化後、DFが閾値th以下のものはメンバー名に含まないようにした(th=0.15, 0.2を使用)。
4. **メンバー名抽出アルゴリズムを使う**
DF=1のものを削除後、メンバー名抽出アルゴリズムを適用する。

表2 処理の違いによる精度の差

処理	再現率(R)	適合率(P)	F値
処理なし	74.05	21.11	32.85
DF=1を削除	71.15	50.11	58.80
DFに閾値を設定(th=0.15)	59.39	62.44	60.88
DFに閾値を設定(th=0.2)	55.57	68.03	61.17
メンバー名抽出アルゴリズム使用	67.78	62.35	64.95

表2よりDF=1のものを削除すると、適合率が大幅に上昇し、不正解のデータを除去できることがわかった。またDFに閾値を設定すると再現率が大幅に下がり、正解であるデータも除

去されることが分かった。その反面、メンバー名抽出アルゴリズムを使用した場合の再現率、適合率は、正解データを大きく除去することなく、不正解データを除去できることがわかった。

6. まとめ

本研究では、バンドのメンバー情報をWeb上から自動収集することで、バンド間の関係を表すバンドネットワークを構築し、そのバンドネットワークを楽曲推薦やバンドの分類に利用する手法を提案した。メンバー名の収集では、GoogleのAPIを利用し収集した複数のWebページから楽器名との出現パターンを判定することでメンバー名の抽出をおこなった。また、メンバーである可能性が高い人物名とのパターンの共起を判定する手法を提案し、メンバー名の収集精度を向上させた。そして構築したバンドネットワークを利用し、関連度を基にバンドを探索するシステムを実装した。また、関連度を基にバンドのグルーピングを行った。

現在、メンバー名収集手法に関しては、同名のバンドや同名の人物を区別出来ないという問題がある。複数のジャンルにわたるような大規模のネットワークの構築を目指す場合には、同名のバンドも多く存在するため、この問題を解決する必要がある。

さらに、メンバー名の修正や入力をおこなえるシステムを構築することも考えている。あるバンドに詳しいユーザがこのシステムにより情報の不足や間違いを修正することで、今までにない詳細なメンバー情報を収集することも可能であると考えている。

参考文献

- 1) G.Tzanetakis and P.Cook, "Musical Genre Classification of Audio Signals," IEEE Transactions on Speech and Audio Processing, Vol. 10, No. 5 (2002)
- 2) J.Foote and S.Uchihashi, "The Beat Spectrum: A New Approach to Rhythm Analysis," in Proc. International Conference on Multimedia and Expo (2001)
- 3) U.Shardanand and P.Maes, "Social information filtering: Algorithms for automating "word of mouth"," In Proceedings of ACMCHI95 Conference on Human Factors in Computing Systems, Vol. 1, pp. 210-217 (1995)
- 4) 松尾 豊, 友部 博教, 橋田 浩一, 中島 秀之, 石塚 満, "Web上の情報からの人間関係ネットワークの抽出," 人工知能学会論文誌, Vol. 20, No. 1E, pp. 46-56 (2005)
- 5) E.Pampalk, A.Flexer, and G.Widmer, "Hierarchical Organization and Description of Music Collections at the Artist Level," In the Proceedings of the 9th European Conference on Research and Advanced Technology for Digital Libraries, pp. 37-48 (2005)
- 6) 山田 泰寛, 池田 大輔, 坂本 比呂志, 有村 博紀, "WWWからの情報抽出-Webラッパーの自動構築," 人工知能学会誌, Vol. 19, No. 3, pp. 302-309 (2004)
- 7) M.Schedl and G.Widmer, "Automatically Detecting Members and Instrumentation of Music Bands via Web Content Mining," Proceedings of the 5th Workshop on Adaptive Multimedia Retrieval (2007)
- 8) E.W.Dijkstra, "A note on two problems in connexion with graphs," In Numerische Mathematik, Vol.1, pp. 269-271 (1959)