Towards A Hybrid Approach of Primitive Cognitive Network Process and Agglomerative Hierarchical Clustering for Music Recommendation

Chun Guan
Department of Computer Science
University of Liverpool
Liverpool, UK
chunguan@liverpool.ac.uk
chun_guan@163.com

Kevin Kam Fung Yuen
Department of Computer Science and Software Engineering
Xi'an Jiaotong-Liverpool University
Suzhou, China
kevin.yuen@xjtlu.edu.cn
kevinkf.yuen@gmail.com

***Abstract*-Clustering algorithms have been used in many real world applications including recommendation systems. This paper proposes PCNP-AHC, which is a hybrid approach of Primitive Cognitive Network Process (PCNP) and Agglomerative Hierarchical Clustering (AHC) to cluster music pieces on the basis of user's preferences and similarities between music pieces. PCNP is an ideal alternative of Analytic Hierarchy Process (AHP) to quantify weights of attributes which are used in clustering process. The application of PCNP-AHC for music recommendation is demonstrated.**

Keywords—Primitive Cognitive Network Process; Hierarchical clustering; recommendation system.

I. Introduction

Hierarchical clustering methods [1-9] are popular clustering methods. The goal of hierarchical clustering is to build a hierarchical decomposition of the objects of data sets [2]. The formulation of hierarchical clustering algorithm was proposed by Joe H. WARD, JR in 1963 [3]. Over the next fifty years, hierarchical clustering methods have been progressively applied in many areas [4-7]. The advantage of hierarchical clustering is that the output of hierarchical clustering is demonstrated in a tree graph (named as dendrogram) which is easy to be interpreted [8].

Music recommendation is the popular topic. Agglomerative hierarchical clustering is one of the methods for music recommendation, for example [6]. There are two limitations when applied agglomerative hierarchical clustering to music recommendation. Firstly, the nominal scales of music attributes are difficult to be used in the clustering methods, as the number may be used to represent the nominal scale items but the numbers do not have ranking relationship or numerical value. User could rank the nominal scale which converts to ordinal/interval/ratio scale. Different users may, however, have different music tastes and therefore a technique is needed to reflect the user preferences for scale items. Secondly, the weight of each attribute is equally treated when computing the similarity between music pieces, but, users may have preference with different attributes in many cases.

To address these two limitations, this paper proposes Primitive Cognitive Network Process (PCNP) [10-13] to be applied to Agglomerative Hierarchical Clustering (AHC) [1, 2, 14] for music recommendation. Primitive Cognitive Network Process (PCNP), which is the basic type of Cognitive Network Process (CNP) [10-13], is used to scale the attributes and quantify the importance of each attribute. CNP [10-13] is an approach rectifying the mathematical representation problem of the perception of the paired differences in Analytic Hierarchy Process (AHP) [15]. [16] presented the combination of PCNP and K-means. On the basis of these studies, the proposed PCNP-AHC should be more reliable and feasible.

The rest of this paper is organized as follows. Section II describes the proposed PCNP-AHC algorithms. Section III presents a demonstration of applying PCNP-AHC to music recommendation. Section IV gives a summary of this research and proposed several further research directions.

II. PCNP-AHC

The proposed PCNP-AHC method has three steps: scales definition, weights determination and clustering algorithm.

A. *Scales Definition*

Generally, a music data matrix includes a set of music pieces, and each music piece includes some attributes of nominal scales such as genre, tempo and mood. The Cognitive Pairwise Comparisons (CPC) in Primitive Cognitive Network Process (PCNP) [10-13] are used to define the nominal scales of music attributes. The details are as below.

To represent the paired comparison between two entities, a measurement scale schema $(\aleph, \overline{\mathrm{X}})$ is used. $\aleph$ is the space of linguistic labels of the paired interval scales such as {Equally, Slightly, Moderately, Fairly, Highly, Strongly, Significantly, Outstandingly, Absolutely}. $\overline{X}$ is the numerical representation of $\aleph$ in the form below.

$$\overline{X} = \left\{ \alpha_i = {}^{i\kappa}\!/_{\tau} \mid \forall i \in \{-\tau, \dots, -1, 0, 1 \dots, \tau\}, \kappa > 0 \right\} \quad (1)$$

QSHINE 2015, August 19-20, Taipei, Taiwan

DOI 10.4108/eai.19-8-2015.2261344

κ is the normal utility indicating the subjective perception of the difference between pairs. By default setting, $Max\left(\bar{X}\right)=\kappa$. τ is the number of linguistic scale.

The Pairwise Opposite Matrix (POM) B is used to interpret the individual utilities of the entities.

$$\begin{aligned}\tilde{B}=\left[\tilde{b}_{ij}\right]&=\begin{bmatrix}0 & v_1-v_2 & \cdots & v_1-v_p\\ v_2-v_1 & 0 & \cdots & v_2-v_p\\ \vdots & \vdots & \ddots & \vdots\\ v_p-v_1 & v_p-v_2 & \cdots & 0\end{bmatrix}\\ &\cong\begin{bmatrix}0 & b_{12} & \cdots & b_{1p}\\ b_{21} & 0 & \cdots & b_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ b_{p1} & b_{p2} & \cdots & 0\end{bmatrix}=\left[b_{ij}\right]=B\end{aligned} \quad (2)$$

Equation (2) shows the definition of POM, where v_i means the importance value of entity x_i , and $\tilde{b}_{ij}\cong v_i-v_j$ is the approximate comparison value between entities x_i and x_j . b_{ij} is provided by the expert judgment. For instance, $b_{12}=2$ means entity 1 is moderately more important than entity 2.

To validate the matrix B, the Accordance Index (AI) is used as below:

$$\begin{aligned}&AI=\frac{1}{p^2}\sum_{i=1}^{p}\sum_{j=1}^{p}\delta_{ij}, where\\ &\delta_{ij}=\sqrt{Mean\left(\left(\frac{1}{k}\left(B_i+B_j^T-b_{ij}\right)\right)^2\right)}, i,j\in\left(1,\ldots,p\right)\end{aligned} \quad (3)$$

$AI\geq 0$ and $p\kappa$ is the population utility. If $AI=0$, then B is perfectly accordant; if $0 < AI \leq 0.1$, then B is satisfactory; if $AI > 0.1$, then B is unsatisfactory.

Whilst B is valid, the Row Average plus the normal Utility (RAU) shown as below is used to calculate the weights of entities.

$$v_i=\left(\frac{1}{p}\sum_{j=1}^{p}b_{ij}\right)+\kappa \quad (4)$$

In the above equation, the average of each row in B is calculated, and then the value is added to each average value. Finally, the individual factor weight is derived. The weights can be normalized by the function below.

$$W=\left\{w_i : w_i=\frac{v_i}{p\kappa}, \forall i\in\left\{1,\ldots,p\right\}\right\}, \sum_{i\in\{1,\ldots p\}}v_i=p\kappa \quad (5)$$

B. Weights Determination

In this step, the weights of attributes are determined by CPC of PCNP, which is presented in Step 1.

C. Agglomerative Hierarchical Clustering Process

Agglomerative hierarchical clustering [3] is a bottom-up strategy. [14] briefly described three steps of hierarchical clustering methods. It starts by regarding each object as atomic cluster and then merges them into larger and larger clusters, until all of the objects are in a single cluster or termination condition is satisfied [2]. As there are a number of types or variations of AHC, the detailed procedure used in this paper is described as below.

1) Initialize each object as an individual cluster.

2) Determine all dissimilarities between clusters.

Euclidean distance has been used to calculate the dissimilarities in AHC. In PCNP-AHC, weighted Euclidean distances are used to calculate the dissimilarities as formula (6).

$$d_{\alpha\beta}=\sqrt{\sum_{i=1}^{n}w_i^2\left(x_{\alpha i}-x_{\beta i}\right)^2} \quad (6)$$

where $d_{\alpha\beta}$ is the weighted Euclidean distance between clusters α and β . w_i is the weight of the i th attribute and is measured by CPC.

3) Combine the two closest clusters into a bigger cluster.

4) Compute dissimilarities between new cluster and other clusters (all other dissimilarities remaining unchanged).

Several types of measurement are suitable for measuring the distance between clusters. As a widely used measurement, average distance is used in PCNP-AHC as the form below.

$$d_{avg}(C_i,C_j)=\frac{1}{n_i n_j}\sum_{q\in C_i, q'\in C_j}d_{qq'} \quad (7)$$

where $d_{qq'}$ is the weighted Euclidean distance between objects q and q' ; and n_i is the number of objects in cluster C_i .

5) Repeat Steps 3 and 4 until all objects are in the one cluster or termination condition given by user is satisfied.

III. Application

Assume that 15 music pieces have been randomly chosen from a music retailing website. Each piece has three attributes: Genre, Tempo and Mood. The sample data are presented in Table I. PCNP-AHC is applied to produce a dengrogram of these 15 music pieces according to a user's preference. The music pieces data are processed by PCNP-AHC. According to a user's playlist history, the system will recommend some music pieces. The calculation steps of PCNP-AHC are demonstrated as follows.

A. Scales Definition

In Table I, the scales of three attributes (Genre, Tempo and Mood) should be justified, as they are nominal scales. The scale of Genre attribute has five values: rock, ballad, jazz, dance and pop. The scale of Tempo attribute has three values: fast, moderate and slow. The scale of Mood attribute has five values: cheerful, depressing, relaxing, disturbing and comforting.

PCNP is applied to quantify the user's preference of the five values of Genre scale. Table II shows the cognitive pairwise matrix in the form of (2), weights calculated by (4) and (5), and the accordant index calculated by (3). κ is set to 8 by default and the matrix is perfectly accordant.

TABLE I. A SAMPLE DATASET OF 15 MUSIC PIECES INFORMATION

Id	Genre	Tempo	Mood
1	Rock	Fast	Disturbing
2	Ballad	Moderate	Cheerful
3	Dance	Fast	Cheerful
4	Rock	Slow	Depressing
5	Dance	Moderate	Comforting
6	Pop	Moderate	Relaxing
7	Pop	Fast	Cheerful
8	Ballad	Slow	Relaxing
9	Jazz	Moderate	Relaxing
10	Rock	Moderate	Depressing
11	Pop	Slow	Depressing
12	Ballad	Moderate	Cheerful
13	Pop	Slow	Depressing
14	Rock	Fast	Relaxing
15	Ballad	Fast	Disturbing

TABLE II. COGNITIVE PAIRWISE MATRIX AND SCALE VALUES FOR GENRE

Genre	Rock	Ballad	Jazz	Dance	Pop	Scale Values
Rock	0	2	4	4	-3	0.235
Ballad	-2	0	2	2	-5	0.185
Jazz	-4	-2	0	0	-7	0.135
Dance	-4	-2	0	0	-7	0.135
Pop	3	5	7	7	0	0.310
AI=0						

TABLE III. COGNITIVE PAIRWISE MATRIX AND SCALE VALUES FOR TEMPO

Tempo	Fast	Moderate	Slow	Scale Values
Fast	0	5	7	0.500
Moderate	5	0	2	0.292
Slow	-7	-2	0	0.208
AI=0				

TABLE IV. COGNITIVE PAIRWISE MATRIX AND SCALE VALUES FOR MOOD

Mood	Che	Dep	Rel	Dis	Com	Scale Values
Che	0	1	2	1	-4	0.200
Dep	-1	0	1	0	-5	0.170
Rel	-2	-1	0	-1	-6	0.145
Dis	-1	0	1	0	-5	0.170
Com	4	5	6	5	0	0.315
AI=0						

Similarly, Table III shows the scale of Tempo attribute and Table IV presents the scale of Mood attribute. In table IV, the five scale values are denoted as Che, Dep, Rel, Dis, and Com.

After evaluated by CPC, the dataset has been transformed into a numerical data matrix as Table V shows.

TABLE V. A NUMERICAL DATA MATRIX OF 15 MUSIC PIECES INFORMATION

Id	Genre	Tempo	Mood
1	0.235	0.500	0.170
2	0.185	0.292	0.200
3	0.135	0.500	0.200
4	0.235	0.208	0.170
5	0.135	0.292	0.315
6	0.310	0.292	0.145
7	0.310	0.500	0.200
8	0.185	0.208	0.145
9	0.135	0.292	0.145
10	0.235	0.292	0.170
11	0.310	0.208	0.170
12	0.185	0.292	0.200
13	0.310	0.208	0.170
14	0.235	0.500	0.145
15	0.185	0.500	0.170

B. *Weights Determination*

To determining the weights of the three attributes, Table VI shows the cognitive pairwise matrix in the form of (2), weights calculated by (4) and (5), and the accordant index calculated by (3). κ is set to 8 by default and the matrix is perfectly accordant.

TABLE VI. COGNITIVE PAIRWISE MATRIX AND WEIGHTS FOR ATTRIBUTES

Attributes	Genre	Tempo	Mood	Attribute Weights
Genre	0	5	8	0.514
Tempo	-8	0	3	0.264
Mood	-5	-3	0	0.222
AI=0				

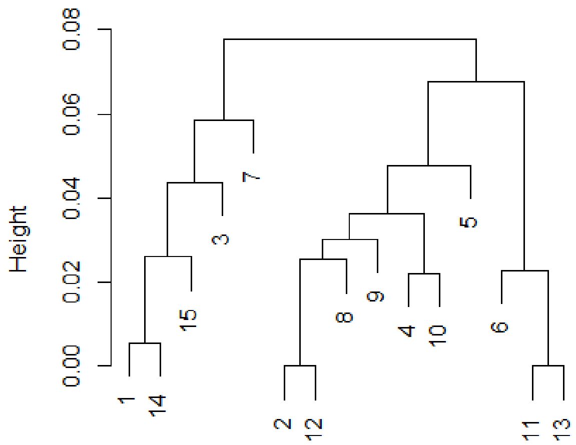

Fig. 1. Dengrogram produced by PCNP-AHC approach

C. *Agglomerative Hierarchical Clustering Process*

In this step, *R* language has been used to implement the proposed method. The data matrix shown in Table V is used as input of clustering process, the output is presented as a dengrogram as Fig. 1.

At the height of 0.06 of the dengrogram, the music pieces are separated into 3 clusters, {1, 14, 15, 3, 7}, {2, 12, 8, 9, 4, 10, 5} and {6, 11, 13}. Assume a user listened music piece 2 on website, the system will recommend the music pieces in the sequence of 12, 8, 9, 4, 10, and 5 to the user.

IV. CONLUSION

This paper proposes PCNP-AHC, a hybrid approach of Primitive Cognitive Network Process (PCNP) and classical Agglomerative Hierarchical Clustering (AHC). The scales and weights of attributes are evaluated by PCNP according to users' preferences. The scaled data and weights are used in AHC to produce a dengrogram. An application of music recommendation is demonstrated.

REFERENCES

[1] AK. Jain, MN. Murty and P J. Flynn, "Data clustering: a review," ACM computing surveys (CSUR), vol. 31, 1999, pp. 264-323.

[2] J. Han and M.Kamber, Data Mining: Concepts and Techniques, Southeast Asia ed., Morgan kaufmann, 2006.

[3] JH. Ward Jr, "Hierarchical grouping to optimize an objective function," Journal of the American statistical association, 1963, vol. 58, pp. 236-244.

[4] G. Karypis, EH. Han and V. Kumar, "Chameleon: Hierarchical clustering using dynamic modeling," Computer, 1999, vol. 32, pp. 68-75.

[5] Y. Zhao and G. Karypis, "Evaluation of hierarchical clustering algorithms for document datasets," Proceedings of the eleventh international conference on Information and knowledge managemen, ACM, 2002, pp.515-524.

[6] A. Shepitsen, J. Gemmell, B. Mobasher and R.Burke, "Personalized recommendation in social tagging systems using hierarchical clustering," Proceedings of the 2008 ACM conference on Recommender systems, ACM, 2008, pp: 259-266.

[7] D. Cai, X. He, Z. Li, W.Y. Ma and J.R. Wen, "Hierarchical clustering of WWW image search results using visual, textual and link information," Proceedings of the 12th annual ACM international conference on Multimedia, ACM, 2004, pp.952-959.

[8] S. Guha, R. Rastogi and K. Shim, "CURE: an efficient clustering algorithm for large databases," ACM SIGMOD Record, ACM, 1998, 1998, vol. 27, pp. 73-84.

[9] S. Guha, R. Rastogi and K. Shim, "ROCK: A robust clustering algorithm for categorical attributes," Data Engineering, 1999. Proceedings, 15th International Conference on IEEE, 1999, pp. 512-521.

[10] K.K.F. Yuen, "Cognitive network process with fuzzy soft computing technique in collective decision aiding," Ph.D. thesis, Hong Kong Polytechnic Univ., 2009, HK.

[11] K.K.F. Yuen, "The pairwise opposite matrix and its cognitive prioritization operations: the ideal alternatives of the pairwise reciprocal matrix and analytic prioritization operators," Journal of the Operational Research Society, 2012, vol. 63, pp. 322-338.

[12] K.K.F. Yuen, "The primitive cognitive network process in medical decision making: Comparisons with the Analytic Hierarchy Process," Applied Soft Computing, 2014, vol.14, pp. 109-199.

[13] K.K.F. Yuen, "Fuzzy Cognitive Network Process: Comparison with Fuzzy Analytic Hierarchy Process in New Product Development Strategy", IEEE Transactions on Fuzzy Systems, 2014, vol.22, pp.597-610.

[14] F. Murtagh, "A survey of recent advances in hierarchical clustering algorithms," The Computer Journal, 1983, vol. 26, pp. 354-359.

[15] T.L. Saaty, "How to make a decision: The analytic hierarchy process," European Journal of Operational Research, 1990, vol. 48, pp. 9-26.

[16] C.Guan and K.K.F. Yuen, "Toward A Hybrid Approach of Primitive Cognitive Network Process and K-Means Clustering for Social Network Analysis," Proceedings of 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, 2013, pp.1267- 1271.