Formal Concept Analysis and Pattern Structures for Recommendation Systems

Nyoman Juniarta

Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
nyoman.juniarta@loria.fr

Abstract. This work focuses on the application of Formal Concept Analysis and pattern structures in the problem of recommendation. We focus on the collaborative recommendation, where we give a suggestion to a new user based on a database of previous users. Here we can look the pattern in the database using two approaches: interval pattern structures or gradual pattern mining.

Keywords: biclustering, FCA, pattern structures, recommendation

1 Introduction

CrossCult (http://www.crosscult.eu) is a European project whose idea is to support the emergence of a European cultural heritage by allowing visitors in different cultural sites (e.g. museum, historic city, archaeological site) to improve the quality of their visit by using adapted computer-based devices and to consider the visit at a European level. Such improvement can be accomplished by studying, among others, the possibility to build a dynamic recommendation system. This system should be able to produce a relevant suggestion on which part of a cultural site may be interesting for a specific visitor. Here, our objective is to study a dynamic recommendation system for visitors in a museum. Given a new visitor V_n, the task is to suggest a museum item that may be interesting for him/her. Based on how a suggestion is made to a new visitor V_n, a recommendation system can be classified into one of the three following categories [1]:

- *Content-based recommendations*: The system makes a suggestion based only on the previous visited items of V_n. For example, if V_n visited mostly the items from prehistoric era, then the system recommends another item from that era.
- *Collaborative recommendations*: The system looks for previous users who have similar interest to V_n, and makes a suggestion based on their visited items. For example, if many of V_n's similar users have visited item I, then the system recommends this item.
- *Hybrid approaches*: The combination of content-based and collaborative approaches.

In this work, we explore the second category (collaborative recommendation) for museum visitors.

2 State-of-the-Art

2.1 Formal Concept Analysis and Pattern Structures

Formal Concept Analysis (FCA) is a mathematical framework based on lattice theory and used for classification, data analysis, and knowledge discovery, introduced in [4]. From a formal context, FCA detects all formal concepts, who can be arranged in a concept lattice.

Definition 1. A formal context is a triple (G, M, I), where G is a set of objects, M is a set of attributes, and I is a binary relation between G and M, i.e. $I \subseteq G \times M$.

If an object g has an attribute m, then $(g, m) \in I$. An example of a formal context is shown in Figure 1.

	m_1	m_2	m_3	m_4
g_1				×
g_2	×		×	
g_3	×	×	×	
g_4		×	×	×

Fig. 1. A formal context.

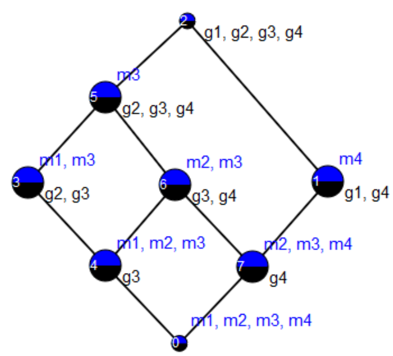

Fig. 2. Concept lattice for the formal context in Figure 1.

The Galois connection for a formal context (G, M, I) is defined as follows:

Definition 2. For a subset of objects $A \subseteq G$, A' is the set of attributes that are possessed by all objects in A, i.e.: $A' = \{m \in M | \forall g \in A, (g, m) \in I\}$

Dually, for a subset of attributes $B \subseteq M$, B' is the set of objects that have all attributes in B, i.e.: $B' = \{g \in G | \forall m \in B, (g, m) \in I\}$

A formal concept is the pair (A, B), where $A \subseteq G$ and $B \subseteq M$, and such that $A' = B$ and $B' = A$. For such concept, A is called its extent and B is its intent.

From Figure 1, we see that $\{g_3, g_4\}' = \{m_2, m_3\}$ and $\{m_2, m_3\}' = \{g_3, g_4\}$. Therefore, $(\{g_3, g_4\}, \{m_2, m_3\})$ is a formal concept. It should be noticed that the extent and the intent of a concept (A, B) are closed sets, i.e. $A = A''$ and $B = B''$. A formal concept (A, B) is a *subconcept* of a formal concept (C, D) – denoted by $(A, B) \leq (C, D)$ – if $A \subseteq C$ (or equivalently $D \subseteq B$).

If $(A, B) < (C, D)$ and there is no (E, F) such that $(A, B) < (E, F) < (C, D)$, then (A, B) is a cover of (C, D). A concept lattice can be formed using the $\leq$ relation which defines the partial order among concepts. For the context in Figure 1, the formal concepts and their corresponding lattice are shown in Figure 2 (extent and intent is below and above each concept respectively).

FCA is restricted to specific datasets where each attribute is binary (e.g. has only yes/no value). For more complex values (e.g. numbers, strings, trees, graphs...), FCA is then generalized into pattern structures [3].

Definition 3. A pattern structure is a triple $(G, (D, \sqcap), \delta)$, where G is a set of objects, $(D, \sqcap)$ is a complete meet-semilattice of descriptions, and $\delta : G \rightarrow D$ maps an object to a description.

The operator $\sqcap$ is a similarity operator that returns the common elements to two descriptions. A description can be a set, a sequence, or other complex structure. In the binary case where descriptions are sets, $\sqcap$ corresponds to set intersection ($\cap$), i.e. $\{a, b, c\} \sqcap \{a, b, d\} = \{a, b\}$, and $\sqsubseteq$ corresponds to subset inclusion ($\subseteq$). The order between any two descriptions is given by the subsumption relation:

$$d_1 \sqsubseteq d_2 \Longleftrightarrow d_1 \sqcap d_2 = d_1$$

Definition 4. The Galois connection for a pattern structure $(G, (D, \sqcap), \delta)$ is defined by:

$A^\diamond = \bigsqcap_{g \in A} \delta(g)$, $A \subseteq G$,
$d^\diamond = \{g \in G | d \sqsubseteq \delta(g)\}, d \in D$.

A pattern concept is a pair (A, d), $A \subseteq G$ and $d \in D$, where $A^\diamond = d$ and $d^\diamond = A$.

	m_1	m_2	m_3
g_1	5	7	6
g_2	6	8	4
g_3	4	8	5
g_4	4	9	8
g_5	5	8	5
g_t	4	7	?

Table 1. A numerical dataset with 6 objects (one as a target object).

Interval pattern structures (ip-structures) is one of possible extensions of FCA to study more complex descriptions. In ip-structures, the description of an object is an interval for each attribute. Consider the dataset given in Table 1 where $G = \{g_1, g_2, g_3\}$ is the set of objects and $M = \{m_1, m_2, m_3, m_4\}$ is the set of attributes. Here the description of g_1 is $\delta(g_1) = \langle [5, 5], [7, 7], [6, 6] \rangle$, while the description of g_4 is $\delta(g_4) = \langle [4, 4], [9, 9], [8, 8] \rangle$.

The similarity operator $\sqcap$ is the smallest interval containing both descriptions in each attribute. Therefore, $\delta(g_1) \sqcap \delta(g_4) = \langle [4, 5], [7, 9], [6, 8] \rangle$. A description with larger interval is "subsumed by" ($\sqsubseteq$) descriptions with smaller ones. The corresponding Galois connection is illustrated as follows:

$$\begin{aligned}
\{g_1, g_4\}^{\diamond} &= \delta(g_1) \sqcap \delta(g_4) \\
&= \langle [4,5], [7,9], [6,8] \rangle \\
\langle [4,5], [7,9], [6,8] \rangle^{\diamond} &= \{g \in G | \langle [4,5], [7,9], [6,8] \rangle \sqsubseteq \delta(g)\} \\
&= \{g_1, g_4\}
\end{aligned}$$

Similar to formal concept, interval pattern concept (ip-concept) can also be arranged in a lattice. The ip-concept lattice for Table 1 is illustrated in Figure 3. For further background concerning ip-structures, see [5].

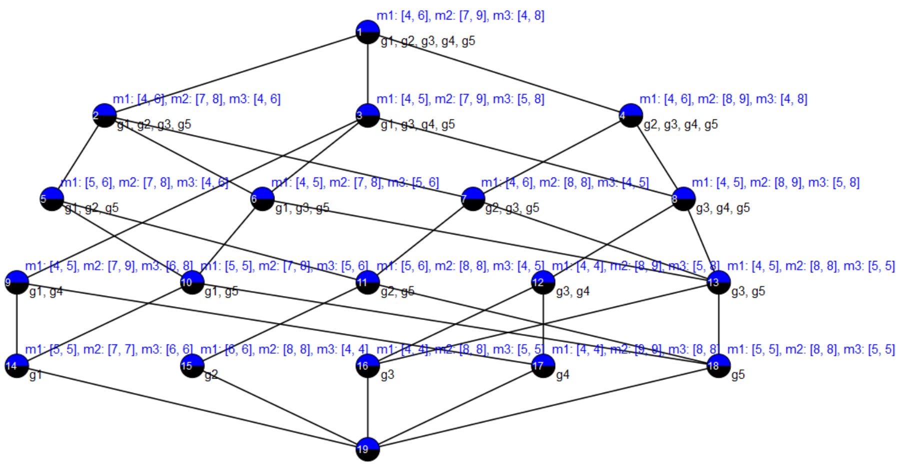

Fig. 3. Lattice of interval pattern concept for g_1, g_2, and g_3 in Table 1.

2.2 Gradual Pattern Mining

In a numerical dataset, a gradual pattern is a pattern that is observed from at least two objects across at least two attributes. Usually this pattern is described as a correlation between two (or more) attributes: "the more/less A corresponds to the more/less B". This type of pattern was originally proposed in [2] and studied by [7] as an instrument to aid fuzzy controllers.

Consider again the matrix in Table 1. This numerical matrix can be transformed into a sign matrix as shown in Table 2. It contains information about the attribute comparison between any pair of objects. For example, pair (g_1, g_2) has '$\nearrow$' in m_1 because according to this attribute, $g_1 < g_2$. Then, from this sign matrix, we can find a coherent-sign-changes bicluster (for detailed explanation about biclustering, please refer to [6]). One of such bicluster is $(\{p_3, p_5, p_6, p_7, p_8, p_9, p_{10}\}, \{m_1, m_2, m_3\})$. This bicluster represent the pattern "the more/less

m_1, the less/more m_2, the less/more m_3 (the '=' can be regarded as either '$\searrow$' or '$\nearrow$').

Pair	m_1	m_2	m_3
$p_1 = (g_1, g_2)$	$\nearrow$	$\nearrow$	$\searrow$
$p_2 = (g_1, g_3)$	$\searrow$	$\nearrow$	$\searrow$
$p_3 = (g_1, g_4)$	$\searrow$	$\nearrow$	$\nearrow$
$p_4 = (g_1, g_5)$	$=$	$\nearrow$	$\searrow$
$p_5 = (g_2, g_3)$	$\searrow$	$=$	$\nearrow$
$p_6 = (g_2, g_4)$	$\searrow$	$\nearrow$	$\nearrow$
$p_7 = (g_2, g_5)$	$\searrow$	$=$	$\nearrow$
$p_8 = (g_3, g_4)$	$=$	$\nearrow$	$\nearrow$
$p_9 = (g_3, g_5)$	$\nearrow$	$=$	$=$
$p_{10} = (g_4, g_5)$	$\nearrow$	$\searrow$	$\searrow$

Table 2. Sign matrix for Table 1.

3 Proposed Approach

3.1 Recommendation Based on Interval Pattern Structures

Table 1 can be regarded as a rating matrix, where G is the set of users and M is the set of movies. Here we have a target user g_t for which we will estimate his/her rating for movie m_3. Based on this estimation, we can decide whether we recommend m_3 to g_t.

To do that we can traverse the lattice in Figure 3 from the top node. Here we search the concept(s) with the most specific interval that contains $m_1 : 4$ and $m_2 : 7$. We will arrive at concept $(\{g_1, g_3, g_5\}, \langle[4,5], [7,8], [5,6]\rangle)$. Therefore, we can estimate that g_t will likely give a rating between 5 and 6 for m_3.

3.2 Recommendation Based on Gradual Pattern Mining

For explaining this approach, we will also use the example in Table 1 as user–movie rating matrix. The user g_t has rated the movie m_1 and m_2, and his/her rating for m_3 is to be estimated. Therefore, from the sign matrix in Table 2, we have to find the largest bicluster that contains m_1 and m_2.

Suppose that the largest bicluster is $(\{p_3, p_5, p_6, p_7, p_8, p_9, p_{10}\}, \{m_1, m_2, m_3\})$ that represents the pattern R_1 "the more/less m_1, the less/more m_2, the less/more m_3. Then, we have to compare the ratings from g_t to the ratings from every other users, as shown in Table 3. Here the pairs that follow R_1 are p_a, p_c, and p_d, and the estimation of m_3's rating from g_t is in the range [6,8].

Pairs	m_1	m_2	m_3	estimation
$p_a = (g_1, g_t)$	$\searrow$	$=$		≥ 6
$p_b = (g_2, g_t)$	$\searrow$	$\searrow$		–
$p_c = (g_3, g_t)$	$=$	$\searrow$		≤ 5
$p_d = (g_4, g_t)$	$=$	$\searrow$		≤ 8
$p_e = (g_5, g_t)$	$\searrow$	$\searrow$		–

Table 3. Comparison of g_t and every other users in Table 1.

4 Methodology

Our research problem is to build a more sophisticated collaborative recommendation system for visitors in a museum. The problem is then formulated as rating prediction, i.e. predicting the given rating from a new user based on patterns studied from previous users. We focus on whether FCA and pattern structures are applicable to our problem. In literature, certain types of pattern structures have been studied to solve the task of recommendation. On the other hand, the applicability of interval pattern structures for this task is not yet explored, and this becomes our objective.

Finally, the proposed recommendation system is evaluated based on the accuracy of recommended items. Given that the CrossCult project doesn't have any substantial visitor–item rating dataset, the approaches will be tested on a movie dataset.

Currently, the authors work on gradual pattern mining using coherent-sign-changes biclustering. In order to take into account the '=' sign, we will apply the interval pattern structures.

Acknowledgments

The thesis of the author is financed by the Région Grand-Est and the European project CrossCult (`http://www.crosscult.eu/`), under the supervision of Amedeo Napoli, Chedy Raïssi, and Miguel Couceiro.

References

1. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE transactions on knowledge and data engineering **17**(6), 734–749 (2005)
2. Dubois, D., Prade, H.: Gradual inference rules in approximate reasoning. Information Sciences **61**(1-2), 103–122 (1992)
3. Ganter, B., Kuznetsov, S.O.: Pattern structures and their projections. In: International Conference on Conceptual Structures. pp. 129–142. Springer (2001)
4. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer-Verlag, 2nd edn. (1999)
5. Kaytoue, M.: Traitement de données numériques par analyse formelle de concepts et structures de patrons. Ph.D. thesis, Université de Lorraine (2011)

6. Madeira, S.C., Oliveira, A.L.: Biclustering algorithms for biological data analysis: a survey. IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB) **1**(1), 24–45 (2004)
7. Subašić, P., Hirota, K.: Similarity rules and gradual rules for analogical and interpolative reasoning with imprecise data. Fuzzy Sets and Systems **96**(1), 53–75 (1998)