Combinatorial optimization based recommender systems

Fabio Roda

DisMoiOu, 33 rue des Jeuneurs, 75002 Paris, France, and
LIX, École Polytechnique, 91128 Palaiseau, France

Leo Liberti

LIX, École Polytechnique, 91128 Palaiseau, France

Franco Raimondi

Deptartment of Computer Science, UCL, London, UK

Key words: collaborative filtering, maximum capacity path

1 Introduction

Recommender systems exploit a set of established user preferences to predict topics or products that a new user might like [2]. Recommender systems have become an important research area in the field of information retrieval. Many approaches have been developed in recent years and the interest is very high. However, despite all the efforts, recommender systems are still in need of further development and more advanced recommendation modelling methods, as these systems must take into account additional requirements on user preferences, such as geographic search and social networking. This fact, in particular, implies that the recommendation must be much more "personalized" than it used to be.

In this paper, we describe the recommender system used in the "DisMoiOu" ("TellMeWhere" in French) on-line service (`http://dismoiou.fr`), which provides the user with advice on places that may be of interest to him/her; the

Email addresses: `fronline@libero.it` (Fabio Roda), `liberti@lix.polytechnique.fr` (Leo Liberti), `f.raimondi@cs.ucl.ac.uk` (Franco Raimondi).

definition of “interest” in this context is personalized taking into account the geographical position of the user (for example when the service is used with portable phones such as the Apple iPhone), his/her past ratings, and the ratings of his/her neighbourhood in a known social network.

Using the accepted terminology [6], DisMoiOu is mainly a Collaborative Filtering System (CFS): it employs opinions collected from similar users to suggest likely places. By contrast with existing recommender systems, ours puts together the use of a graph theoretical model [4] and that of combinatorial optimization methods [1]. Broadly speaking, we encode known relations between users and places and users and other users by means of weighted graphs. We then define essential components of the system by means of combinatorial optimization problems on a reformulation of these graphs, which are finally used to derive a ranking on the recommendations associated to pairs (user,place).

Preliminary computational results on the three classical evaluation parameters for recommender systems (accuracy, recall, precision [3]) show that our system performs well with respect to accuracy and recall, but precision results need to be improved.

2 Formalization of the problem

We employ the usual graph-theoretical notation, e.g. for a vertex v of a graph G, $\delta^+_G(v), \delta^-_G(v)$ are the set of vertices adjacent to incoming and respectively outgoing arcs. For vertices u, v of G we also let $\Delta_G(u, v) = \delta^+_G(u) \cap \delta^+_G(v)$.

We are given two finite sets U (the users) and P (the places), and a vertex set $V = U \cup P$. We are also given two directed graphs as follows.

- A ratings bipartite digraph $R = (V, A)$ where $A \subseteq U \times P$ is weighted by a function $\rho : A \to [-1, 1]$, which expresses the ratings of users with respect to the places.
- A social network $S = (U, B)$ weighted by a function $\gamma : B \to [0, 1]$ which encodes a confidence coefficient between users.

The union of the two graphs $G = R \cup S$ is a mixed ratings/social network which is used to establish new arcs in $U \times U$ or to change the values that γ takes on existing arcs: a missing relation of confidence between two users can be established if both like (almost) the same places in (almost) the same way. Moreover, even when a confidence relation is already part of B, its strength can change according to similar shared preferences situations. This is encoded by the reformulated graph G' described below.

We define a graph G' with vertex set $V' = U \cup P$ and arc set B' (weighted by a function $\gamma' : B' \to [0,1]$) defined in the following way.

(1) For every $k, \ell \in U$ such that $(k,\ell) \notin B$ and subgraph $H = (V_H, A_H)$ of R induced by the vertex set $V_H = \{k,\ell\} \cup \Delta_R(k,\ell)$ (see Fig. 1) such that $A_H \neq \emptyset$, B' contains the arc (k,ℓ) weighted by $\gamma'_{k\ell} = f(\vartheta)$, where

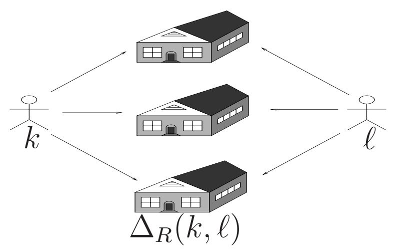

Fig. 1. A subgraph H of R.

$$\vartheta = \frac{1}{|\Delta_R(k,\ell)|} \sum_{i \in \Delta_R(k,\ell)} |\rho_{ki} - \rho_{\ell i}|. \quad (1)$$

ϑ represents the difference between users. The bigger it is, the lower the confidence $\gamma'_{k\ell}$. $\gamma'_{k\ell}$ is obtained as a function f of ϑ.

(2) For every $k, \ell \in U$ such that $(k,\ell) \in B$ and subgraph $H = (V_H, A_H)$ of R induced by the vertex set $V_H = \{k,\ell\} \cup \Delta_R(k,\ell)$ such that $A_H \neq \emptyset$, B' contains the arc (k,ℓ) weighted by $\gamma'_{k\ell} = g(\gamma_{k\ell}, \vartheta)$.

We let $X = (U \times P) \smallsetminus A$ be the set of all recommendations that the system is supposed to be able to make.

2.1 Identification of maximum confidence paths

Given $(k^*, i^*) \in X$, we consider the graph $Z = (W, C)$ where $W = U \cup \{i^*\}$ and $C = B' \cup \{(k, i^*) \mid k \in \delta^-_R(i^*)\}$. Our aim is to compute a ranking for the known ratings $\{\rho_{ki^*} \mid k \in \delta^-_R(i^*)\}$ by means of the confidence relations encoded in the network Z, using paths (or sets thereof) ensuring maximum confidence. By convention, we extend the confidence function γ to arcs in C adjacent to i^* as follows: $\forall k \in \delta^-_R(i^*)$ $(\gamma_{ki^*} = 1)$.

We make the assumption that for a path $p \subseteq C$ in Z, $\gamma(p) = \min\limits_{(k,\ell)\in p} \gamma_{k\ell}$, i.e. that the confidence on a path is defined by the lowest confidence arc in the path. This implies that finding the maximum confidence path between k^* and i^* is the same as finding a path whose arc of minimum weight γ is maximum (among all paths $k^* \to i^*$). Considering Z as a network where γ are capacities on the arcs, a maximum confidence path is the same as a *maximum capacity path* between k^* and i^*, for which there exists an algorithm linear in the number of arcs [5]. The mathematical programming formulation for the

MAXIMUM CAPACITY PATH (MCP) problem is:

$$\left.\begin{array}{rl} \max\limits_{x,t} & t \\ \text{s.t.} & \sum\limits_{\ell\in\delta_R^+(k^*)} x_{k^*\ell} = 1 \\ \forall \ell \in W \smallsetminus \{k^*, i^*\} & \sum\limits_{h\in\delta_R^-(\ell)} x_{h\ell} = \sum\limits_{h\in\delta_R^+(\ell)} x_{\ell h} \\ \forall (k,\ell)\in C & t \le \gamma_{k\ell}x_{k\ell} + M(1-x_{k\ell}) \\ & x\in\{0,1\},\ t\ge 0, \end{array}\right\} \quad (2)$$

where $M \geq \max\limits_{(k,\ell)\in C} \gamma_{k\ell}$. Let $\bar{p} \subseteq C$ be the maximum confidence path (i.e. the set of arcs (k,ℓ) such that $x_{k\ell} = 1$), and $\alpha(\bar{p}) = \text{argmin}\{\gamma_{k\ell} \mid (k,\ell) \in \bar{p}\}$. Removing $\alpha(\bar{p})$ from $C^1 = C$ yields a different set of arcs C^2 with associated network $Z^2 = (W, C^2)$, in which we can re-solve (2) to obtain a path $\bar{p}^2$ as long as Z^2 is connected (otherwise, define $\bar{p}^2 = \emptyset$): this defines an iterative process for obtaining a sequence of triplets $(Z^r, \bar{p}^r)$. Given a confidence threshold $\Gamma \in [0,1]$ and an integer $q > 0$, we define the set $\Omega = \{\bar{p}^r \mid \bar{p}^r \neq \emptyset \wedge r \leq q \wedge \gamma_{\alpha(\bar{p}^r)} \geq \Gamma\}$ of all high confidence paths from k^* to i^*.

2.2 Ranking the ratings

Recall each $p \in \Omega$ ends in i^*, so we can define $\lambda : \Omega \to \delta_R^-(i^*)$ such that $\lambda(p)$ is the last arc of p. Thus, we extend ρ to Ω as follows:

$$\rho(p) = \rho(\lambda(p)).$$

Let $\Theta = \{\sigma \in [-1,1] \mid \exists p \in \Omega\ (\sigma = \rho(p))\}$ be the set of ratings for i^* available to k^*. We evaluate each rating by assigning it the sum of the confidences along the corresponding paths. Let $v : \Theta \to \mathbb{R}_+$ be given by

$$\forall \sigma \in \Theta \quad v(\sigma) = \sum_{\substack{p\in\Omega \\ \rho(p)=\sigma}} \gamma(p).$$

We use v to define a ranking on Θ (i.e. an order $<$ on Θ): for all $\sigma, \tau \in \Theta$ $(\sigma < \tau \leftrightarrow v(\sigma) < v(\tau))$. Naturally, this set-up rests on the fact that $|\Theta| < |\Omega|$, which is exactly what happens in DisMoiOu's implementation. The recommender system then picks the greatest σ in Θ (i.e. the rating with highest associated cumulative confidence) as the recommendations to user k^* concerning the place i^*. Finally, the output of the recommender system is a set of high confidence recommendations for user k^* as i^* ranges in P.

3 Extensions

One of the troubles with the recommender system described in Sect. 2 is that paths in Ω might be too long: although in our formalization paths are only weighted by the value of the arc of minimum confidence, in practice it also makes sense to require that the paths should either be shortest or at least of constrained cardinality, for confidence usually wanes with distance in social networks. Enforcement of the first idea yields a bi-criterion path problem as (2) with an additional objective function:

$$\min_{x,t} \sum_{(k,\ell)\in C} x_{k\ell}. \tag{3}$$

Enforcement of the second idea (say with paths having cardinality at most K) yields the corresponding constraint:

$$\sum_{(k,\ell)\in C} x_{k\ell} \leq K. \tag{4}$$

References

[1] G. Adomavicius and A. Tuzhilin. 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] J. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In G. Cooper and S. Moral, editors, *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, pages 43–52, San Francisco, 1998. Morgan Kaufmann.

[3] C. Cleverdon, J. Mills, and M. Keen. *Factors Determining the Performance of Indexing Systems: ASLIB Cranfield Research Project. Volume 1: Design*. ASLIB Cranfield Research Project, Cranfield, 1966.

[4] Z. Huang, W. Chung, and H. Chen. A graph model for e-commerce recommender systems. *Journal of the American Society for Information Science and Technology*, 55(3):259–274, 2004.

[5] A. Punnen. A linear time algorithm for the maximum capacity path problem. *European Journal of Operational Research*, 53:402–404, 1991.

[6] E. Vozalis and K. Margaritis. Analysis of recommender systems algorithms. In E. Lipitakis, editor, *The 6th Hellenic European Conference on Computer Mathematics & its Applications*, pages 732–745, Athens, 2003. Athens University of Economics and Business.