Exploiting Semantic Product Descriptions for Recommender Systems

CaiU255Nicolas Ziegler Lars SchmidtU255Thieme Georg Lausen

Institut f

blessings provided by recommender systems.

We intend to alleviate the information sparsity issue by exploiting those before-mentioned product classification taxonomies as powerful background knowledge. Hereby, our hybrid information filtering approach permits properly inferring profile similarity between two given users though both agents might not have rated any products in common. Making use of the "collaboration via content" paradigm [23], of dimensionality $|B|$, but by vectors of interest scores assigned to *topics* taken from taxonomy C over product categories d D.

Relevance $w_i(b_k)$ of product b_k for the active user a_i is then defined as follows:

$$w_i(b_k) = \frac{q \cdot c_b(a_i, b_k) \cdot \sum_{a_j \quad A_i(b_k)} c(a_i, a_j)}{|A_i(b_k)| + Y_R}, \qquad (7)$$

where

$$A_i(b_k) = \{a_j \quad \text{clique}(a_i) \mid b_k \quad R_j\}$$

and

$$q = \text{q}$$

procedure diversify (*P*

$$w_i(b_k) = \sum_{a_j \in A_i(b_k)} c(a_i, a_j) \tag{8}$$

Hereby, we measure user similarity $c(a_i, a_j)$ according to Pearson correlation, introduced in Section 3.3.1. Profile vectors v_i, v_j for agents a_i, a_j, respectively, represent implicit ratings for every product $b_k \in B$, hence $v_i, v_j \in \{0, 1\}^{|B|}$.

4.2.1.3 Hybrid Recommender Approach.

The third competing system exploits both collaborative and content-based filtering facilities, hence its hybrid nature. The algorithmic clockwork mimics Pazzani's "collaboration via content" proposal [23], representing user profiles v_i through

14	1.37057472	0.17335063	2.41760397	1.60240506
15	1.39215949	0.18358535	2.44001489	1.65442854
16	1.40580331	0.19158667	2.44087139	1.70458443
17	1.38639753	0.18093777	2.54088386	1.72781798
18	1.3813139	0.18297222	2.5909104	1.77596723
19	1.37089092	0.19004361	2.57706446	1.7171881
20	1.34556085	0.1827844	2.61306722	1.76000474
21	1.41032046	0.18140914	2.68021622	1.84471085
22	1.39184461	0.19165824	2.69981036	1.91931885
23	1.40053345	0.20192565	2.70627937	1.89828101
24	1.39420985	0.16253384	2.73844159	1.9214737

Figure 4: Results obtained from online evaluation

participating users. In both cases, the taxonomy-driven system performed best and the purely collaborative worst.

Second, we counted all those raters perceiving one specific system as best. Again, comparison was based upon the overall verdict and average recommendation rating, likewise. In order to guarantee fairness, we discarded users not having rated all three systems for each metric. The right chart of Figure 4 shows that the appreciation of the taxonomy-driven method significantly prevailed.

Eventually, we may conclude that results obtained from the online analysis back offline evaluation results. In both

Workshop on Personalization and Recommendation in E-Commerce (Malaga, Spain, May 2002), Springer-Verlag.

[11] HERLOCKER, J., KONSTAN, J., BORCHERS, A., AND RIEDL, J. An algorithmic framework for performing collaborative filtering. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*