**l-Injection: Toward Effective Collaborative Filtering Using Uninteresting Items**

**ABSTRACT:**

We develop a novel framework, named as l-injection, to address the sparsity problem of recommender systems. By carefully injecting low values to a selected set of unrated user-item pairs in a user-item matrix, we demonstrate that top-N recommendation accuracies of various collaborative filtering (CF) techniques can be significantly and consistently improved. We first adopt the notion of pre-use preferences of users toward a vast amount of unrated items. Using this notion, we identify uninteresting items that have not been rated yet but are likely to receive low ratings from users, and selectively impute them as low values. As our proposed approach is method-agnostic, it can be easily applied to a variety of CF algorithms. Through comprehensive experiments with three real-life datasets (e.g., Movielens, Ciao, and Watcha), we demonstrate that our solution consistently and universally enhances the accuracies of existing CF algorithms (e.g., item-based CF, SVD-based CF, and SVD++) by 2.5 to 5 times on average. Furthermore, our solution improves the running time of those CF methods by 1.2 to 2.3 times when its setting produces the best accuracy. The datasets and codes that we used in the experiments are available at: https://goo.gl/KUrmip.

**EXISTING SYSTEM:**

* Among existing solutions in recommender systems RS, in particular, collaborative filtering (CF) methods have been shown to be widely effective. Based on the past behavior of users such as explicit user ratings and implicit click logs, CF methods exploit the similarities between users’ behavior patterns.
* Most CF methods, despite their wide adoption in practice, suffer from low accuracy if most users rate only a few items (thus producing a very sparse rating matrix), called the data sparsity problem. This is because the number of unrated items is significantly more than that of rated items.
* To address this problem, some existing work attempted to infer users’ ratings on unrated items based on additional information such as clicks and bookmarks

**DISADVANTAGES OF EXISTING SYSTEM:**

* These works require an overhead of collecting extra data, which itself may have another data sparsity problem.
* 0-injection simply considers all uninteresting items as zero, it may neglect to the characteristics of users or items. In contrast, l-injection not only maximizes the impact of filling missing ratings but also considers the characteristics of users and items, by imputing uninteresting items with low peruse preferences.

**PROPOSED SYSTEM:**

* In this work, we develop a more general l-injection to infer different user preferences for uninteresting items for users, and show that l-injection mostly outperforms 0-injection.
* The proposed l-injection approach can improve the accuracy of top-N recommendation based on two strategies: (1) preventing uninteresting items from being included in the top-N recommendation, and (2) exploiting both uninteresting and rated items to predict the relative preferences of unrated items more accurately.
* With the first strategy, because users are aware of the existence of uninteresting items but do not like them, such uninteresting items are likely to be false positives if included in top-N recommendation. Therefore, it is effective to exclude uninteresting items from top-N recommendation results.
* Next, the second strategy can be interpreted using the concept of typical memory based CF methods.

**ADVANTAGES OF PROPOSED SYSTEM:**

* We introduce a new notion of uninteresting items, and classify user preferences into pre-use and post-use preferences to identify uninteresting items.
* We propose to identify uninteresting items via peruse preferences by solving the OCCF problem and show its implications and effectiveness.
* We propose low-value injection (called l-injection) to improve the accuracy of top-N recommendation in existing CF algorithms.
* While existing CF methods only employ user preferences on rated items, the proposed approach employs both peruse and post-use preferences. Specifically, the proposed approach first infers pre-use preferences of unrated items and identifies uninteresting items.

**SYSTEM REQUIREMENTS**

➢ **H/W System Configuration:-**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**Software Requirements:**

* Operating System - Windows XP
* Coding Language - Java/J2EE(JSP,Servlet)
* Front End - J2EE
* Back End - MySQL