Crowdsourced Evaluation of Semantic Patterns for Recommendations

Recommender systems help people cope with the amount of information available on the Internet. Standard methods tend to suggest only similar items, preventing diversity.

Extend existing approaches with semantic patterns to improve diversity in recommendation results. Linked Data enables us to discover connections between items that otherwise do not surface.

The goal of this study is to identify statistics of the patterns that indicate their suitability for the (diversity enabled) recommendation process.

**SEMANTIC PATTERNS** connect a source type with a target type through property-type pairs.

**PATTERN STATISTICS**

**Global Frequency**: how many times a pattern appears in the source.

**Local Frequency**: how many times a pattern appears in the source considering the instantiation of the source type (i.e. Schindler’s List).

**Type Frequency**: how many times a pattern appears in the source considering the instantiation of the type(s) involved (i.e. person).

**Pattern Length**: number of properties involved in the pattern.

**EXPERIMENT**

CrowdFlower workers choose a movie recommendation from 5 options:

- 4 recommendations generated by DBpedia patterns.
- One is from Amazon recommendations.

For each choice we collect also the motivation.

**Some numbers**

- 36 tests
- 12 movies
- 3 genres
- 12 patterns (people-based)
- 722 contributions
- 28 spammers
- 157 hours

**RESULTS**

Global frequency and pattern length are the statistics which correlate higher with the users’ choices.

**RECOMMENDATION DIVERSITY** is calculated as the semantic distance between genres and topics of the recommended items and the items in the user profile.