Ph.D. in Information Technology: Theses Defenses

July 17th, 2018

Room Seminari Alessandra Alario (building 21) - 10.30 am

First Ph.D. presentation and discussion:

Yashar DELDJOO – XXX Cycle

"Video Recommendation by Exploiting the Multimedia Content"

Advisor: Prof. **Paolo Cremonesi**

Abstract:

Video recordings are complex media types. For example, when we watch a movie, we can effortlessly register a lot of details conveyed to us (by the author) through different multimedia channels, in particular, the audio and visual channels. To date, the majority of content-based movie recommender systems (CBMRS) base their recommendations on metadata (e.g., editorial metadata such as genre or wisdom of the crowd such as user-generated tags) since they are human-generated and are assumed to cover the content semantics of movies by a great degree. Multimedia features, on the other hand, provide the means to identify videos that look similar or sound similar. These discerning characteristics of heterogeneous feature sets meet users' differing information needs.

In the context of this PhD thesis, methods for automatically extracting video-related information from the multimedia content (i.e., audio and visual channels) have been elaborated, implemented, and analyzed. Novel techniques have been developed as well as existing ones refined in order to extract useful information from the video content and incorporate them in recommendation systems. Different video recommendation tasks are solved using the extracted multimedia information under recommendation models based on content-based filtering (CBF) models and the ones based on combination of CBF and collaborative filtering (CF).

As a branch of recommender systems, this thesis investigates a particular area in the design space of recommender system algorithm in which the generic recommender algorithm needs to be optimized in order to use a wealth of information encoded in the actual image and audio signals. The results and main findings of these assessments are reported via several offline studies or with user-studies involving real users testing a prototype of developed movie recommender systems powered by multimedia content.

Second Ph.D. presentation and discussion:

Shima ZAHMATKESH – XXIX Cycle

"On Relevant Query Answering over Streaming and Distributed Data"

Advisor: Prof. **Emanuele Della Valle**

Abstract:

Web applications that join streaming with distributed data to provide relevant answers are getting a growing attention in recent years. Answering in a timely fashion, i.e., reactively, is one of the most important performance indicators for those applications.

The Semantic Web community showed that RDF Stream Processing (RSP) is an adequate framework to develop this type of applications. However, remaining reactive can be challenging, especially when the distributed data is slowly evolving, because accessing the distributed data can be highly time consuming as well as rate-limited.

State-of-the-art work addresses this problem by proposing an architectural approach that keeps a local replica of the distributed data. The local replica progressively becomes stale if not updated to reflect the changes in the remote distributed data. For this reason, recently, the RSP community investigated maintenance policies of the local replica that guarantee reactiveness while maximizing the freshness of the replica. The investigated maintenance policies focus on a class of queries that join a data stream with a distributed data source.

This thesis goes beyond the state of the art, focusing on finding the most relevant answers by continuously answering query over streaming and distributed data, while considering the reactiveness constraints imposed by the users. The contributions of this study are various maintenance policies, which are tailored for two classes of queries: i) queries that have to filter data in the distributed dataset before joining it with streaming data, and ii) top-k queries where the scoring function involves data that appears both in the streaming and the distributed datasets.

The contributions of this doctoral thesis are advance policies that let RSP engines continuously answer the two classes of queries described above. In particular, the proposed policies focus on refreshing only the data in the replica that contributes to the relevancy of the results.

PhD Committee:

Prof. **Paolo Cremonesi**, DEIB

Prof. **David Carrera**, Universitat Politècnica de Catalunya (UPC)

Prof. **Oscar Corcho**, Universidad Politécnica de Madrid