



Entity Selection and Ranking for Data Mining applications

Εβιμαρία Τερζή,
Αναπληρώτρια Καθηγήτρια
Department of Computer Science,
Πανεπιστήμιο Βοστόνης,
ΗΠΑ



Περίληψη – Abstract

In many data-mining applications, the input consists of a collection of entities (e.g., reviews about a product, experts that declare certain skills, network nodes or edges) and the goal is to identify a subset of important entities (e.g., useful reviews, competent experts, influential nodes respectively). Existing work solves this problem either by entity ranking or by entity selection. Entity-ranking methods associate a score with every entity. The main drawback of these approaches is that they ignore the redundancy between the highly scored entities. Entity-selection methods try to overcome this drawback by evaluating the goodness of a group of entities collectively. These methods identify the best set of entities, implying that all entities not in the group are unimportant. Such dichotomy of entities conceals the fact that there may be other subsets of entities with equally-good (or almost as good) goodness scores.

In this talk, we will discuss how the drawbacks of the above methods can be overcome by integrating the ranking and selection paradigms. That is, we will introduce ranking mechanisms that are based on entity selection and selection mechanisms that are based on entity ranking. In this framework, the importance scores of individual entities are determined by how many good groups of entities they participate in. Consequently, a good group of entities consists of entities with high importance scores. The main challenge we will discuss is how to explore the solution space of combinatorial problems in order to identify many entities that participate in many good solutions. In the talk, we will describe how our methods can be applied to applications related to expert management systems, management of online product reviews, and network analysis (including physical and social networks).

[Evimaria Terzi](#) is an Associate Professor at the Computer Science Department at Boston University. Before joining BU in 2009, she was a research scientist at IBM Almaden Research Center. Evimaria has received her Ph.D. from University of Helsinki, Finland and her MSc from Purdue University. Evimaria is a recipient of the Microsoft Faculty Fellowship (2010) and the NSF CAREER award (2012). Her research interests span a wide range of data-mining topics including algorithmic problems arising in online social networks, social media and recommender systems.

Παρασκευή 01/04/2016 – 12:00

**Αίθουσα Σεμιναρίων,
Κτίριο Μηχανικών Η/Υ & Πληροφορικής
Πανεπιστήμιο Ιωαννίνων**