Abstract

A person may have multiple personal name aliases and that same thing might available be on the web. Identifying aliases of a name is useful in information retrieval, investigating about things, knowledge management, sentiment analysis, relation extraction and name disambiguation. Extracting aliases of an entity is important task in various jobs such as identification of relations among entities, web search and entity disambiguation. The objective of detecting aliases from the web is to retrieve all the information pertaining to a personal name whose content is described with different nick names called aliases in different documents of web. The Web contains aliases of popular personalities in various domains like sports, politics,
music, cinema etc., and the problem is that it does not contain alias information or information related to its alias names. Recently, there are proven methods of extracting aliases through lexical pattern based retrieval tested using real-world name-alias pairs in Japanese and English as training data related to limited domains. In this paper, a method is proposed in which a lexical-pattern-based approach to extract aliases of a given name that is helpful in information retrieval. A set of names is used and their aliases to extract lexical patterns that describe numerous ways in which information related to aliases of a name is presented on the web the aliases which will be extracted using the proposed method can be successfully utilized in an information retrieval task and will improve in a relation detection task.

References


Index Terms

Computer Science
Algorithms
Keywords
Information Retrieval  Web Mining  Information Extraction  Web Text Analysis.