TITLE:
Supporting Entity-oriented Search with Fine-grained Information in Knowledge Graphs (Digest_要約)

AUTHOR(S):
Wiradee, Imrattanatrai

CITATION:
Wiradee, Imrattanatrai. Supporting Entity-oriented Search with Fine-grained Information in Knowledge Graphs. 京都大学, 2020, 博士(情報学)

ISSUE DATE:
2020-09-23

URL:
https://doi.org/10.14989/doctor.k22806

RIGHT:
学位規則第9条第2項により要約公開; 許諾条件により本文は2021-09-22に公開; 許諾条件により要約は2020-12-23に公開; 許諾条件により要旨は2020-12-23に公開
Supporting Entity-oriented Search with Fine-grained Information in Knowledge Graphs

WIRADEE IMRATTANATRAI

Department of Social Informatics, Graduate School of Informatics, Kyoto University

ABSTRACT

Users usually search for information related to entities such as celebrities, politicians, and companies via search engines. Many modern search engines have been trying to improve user’s search experience in terms of entity-oriented search by introducing search functionalities that directly present relevant entity-related information based on a user’s information need on the search engine result page. With such functionalities, the search users could complete their search tasks regarding their entities of interest easily with less efforts since they do not need to explore as many documents in organic search results. Normally, the entity-oriented search functionalities are empowered by a large-scale Knowledge Graph (KG) as the KG contains structured information of entities gathered from various sources. Nonetheless, although there exists a handful of entity-related information within the KG, the search engines sometimes fail to provide the direct information of entities due to the complexity and ambiguity of the search queries. Since some search queries are lengthy as they contain different specific terms or phrases regarding the entities or the entity types such as a query “top selling movies julia roberts during 80’s” which contains several specific phrases “top selling” and “during 80’s” with respects to the entity type “Movies by Julia Roberts”, without the potential to properly interpret these phrases, the search engines could end up providing irrelevant results to the users. In contrast, some search queries can be ambiguous and under-specified as the user might issue a plain and simple query “julia roberts” that does not indicate any specific terms representing aspects of the entity “Julia Roberts”. With this type of queries, it is difficult for the search engines to prepare the relevant entity-related information unless they can infer the actual user’s information need. Thus, in this thesis, we aim to support the users in searching for the entity-related information by leveraging fine-grained entity information in terms of different data representation and different complexity level of the information within the KG. We propose three research problems where each of them could help mitigating the difficulty in searching for entity-related information in heterogeneous ways. Each of the research proposals is listed as follows:

• First, we focus on the ranking of entities for queries with modifiers. For example, given a search query “ancient temples in kyoto”, the goal is to return a ranked list of entities temples (e.g. “Kennin-ji”, “Tenryu-ji”, and “Kinkaku-ji”) according to the modifiers “ancient” and “kyoto”. To achieve this, we utilize qualitative properties that indicate categorical information of entities (e.g. city) and quantitative properties that indicate numerical information of entities (e.g. founded date) in the KG to filter and rank the entities according to the modifier terms in the queries. Thus, we could increase the coverage of the search queries that expect a ranked list of related entities to be returned so that the users can directly obtain the relevant entities for a more number of queries.

• Second, we focus on the identification of entity properties from text with zero-shot learning. For example, given text “2012 album Red took Taylor Swift’s popularity to new levels and the universal appeal of I Knew You Were Trouble was a key part of that success.”, the task is to identify the property album, as well as track, of the entity “Taylor Swift”. Identifying entity properties from text is considered as a fundamental task that is essential and can be applied to several entity-oriented search functionalities. For instance, the identified entity properties can be used to construct a more relevant entity card, i.e. one
of the successful entity-oriented search functionalities. In addition, we incorporate a zero-shot learning (ZSL) that could help to solve a task despite not having enough training sentences for some properties, especially those properties that are quantitative properties and are represented by complex information within the KG. With the ZSL, the higher number of properties could be identified as we could integrate many types of properties for entities.

- Third, we focus on the explanation of relationship between entities. For example, given a set of entities including “The Lion King”, “Tim Rice”, “Elton John” and “Disney”, we aim to provide the explanation “Tim Rice wrote the songs with Elton John for Disney animated film The Lion King” as it expresses how the given entities are related. For this proposal, we attempt to make use of the complex information of multiple entities in the KG since those information could be useful for providing descriptive explanation of entity relationship. Providing the explanation of the relationship of entities to the users could reduce the difficulty of the users in understanding the relatedness between entities since there are multiple entity-oriented search functionalities involving multiple entities. One good example of those functionalities is entity recommendation where it recommends a set of entities relevant to the given entity. Without any explanation of how the suggested entities are related to the given entity, the users might hesitate to put any attentions to any suggested entities. Hence, good relationship explanation of entities would help relieving this complication between the search users and the search engines.