

## Graduiertenkolloquium Angewandte Informatik

### Embedding based Link Prediction for Knowledge Graph Completion

**M. Sc. Russa Biswas (FIZ Karlsruhe – Leibniz Institute for Information Infrastructure)**

Knowledge Graphs (KGs) have recently gained attention for representing structured knowledge. Since the advent of the Linked Open Data (LOD) cloud, it has constantly been growing containing many KGs about numerous different domains such as government, scholarly data, biomedical domain, etc. KGs are used in Natural Language Processing (NLP) based applications such as entity linking, question answering, recommender systems, etc. Some KGs are heuristically created, automatically generated from heterogeneous resources such as text, images, etc., whereas the others are human-curated. These KGs consist of huge amounts of facts in the form of entities (nodes) and relations (edges) between them. The entities which exhibit similar characteristics are grouped together into classes or types in the KGs and are known as the entity type information. However, these KGs are often incomplete, i.e., there are missing links between the entities and missing links between the entities and their corresponding entity type information.

This thesis focuses on addressing these two challenges in Knowledge Graph Completion:

- Predicting the missing links between entities in KG via head, and tail prediction as well as triple classification.
- Predicting the semantic type of an entity in a KG.

To date, many algorithms are proposed to learn the embeddings of the entities and relations into the same vector space. However, none of the baseline models considers the contextual information of the KGs along with the textual entity descriptions to learn the latent representation of the entities and relations within the KG. This thesis proposes a multi-hop attentive KG embedding model MADLINK for predicting the missing links between the entities.

On the other hand, recent research on entity typing focuses on exploiting the structural information of the KGs, and keywords from the entity description for the prediction task. A novel entity typing model CAT2Type, to predict the missing types of the entities in a KG leveraging the Wikipedia Category information, is proposed in this thesis. Furthermore, to predict the multi-label, multi-class, and hierarchical entity types in KG, a GRAND framework is proposed that exploits graph walks and entity descriptions. The results show that the proposed methods, MADLINK, CAT2Type and GRAND outperform the current state-of-the-art models for the benchmark datasets in the corresponding tasks.

**Termin: Mittwoch, 29.06.2022, 15:45 Uhr**

Ort: Kaiserstr. 89, 76133 Karlsruhe  
Kollegiengebäude am Kronenplatz (Geb. 05.20), 1. OG, Raum 1C-04  
(Hinweise für Besucher: <http://www.aifb.kit.edu/web/Kontakt>)

Veranstalter: Institut AIFB, Forschungsgruppe Information Service Engineering

Zu diesem Vortrag lädt das Institut für Angewandte Informatik und Formale Beschreibungsverfahren alle Interessierten herzlich ein.

M. Färber, A. Oberweis, H. Sack (Org.), A. Sunyaev, Y. Sure-Vetter, M. Volkamer, J. M. Zöllner