

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

**Φανουράκης Νικόλαος
Μεταπτυχιακός Φοιτητής**

**Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης
Επόπτης Μεταπτυχιακής Εργασίας: Καθηγητής, Β. Χριστοφίδης**

Τρίτη, 21 Σεπτεμβρίου 2021, ώρα 11:00 π.μ.

**Join Zoom Meeting
<https://zoom.us/j/99736946175>**

**“ Μέθοδοι Διανυμαστικής Αναπαράστασης Γράφων Γνώσης για Αντιστοίχιση
Οντοτήτων”**

Περίληψη

Τα τελευταία χρόνια, έχουμε γίνει μάρτυρες του πολλαπλασιασμού των γράφων γνώσης σε διάφορους τομείς, έχοντας ως στόχο την υποστήριξη εφαρμογών όπως απάντηση ερωτήσεων, συστάσεις, κλπ. Μία συχνή εργασία όταν ενσωματώνουμε γνώση από διαφορετικούς γράφους γνώσης είναι να βρούμε ποιοι υπογράφοι αναφέρονται στην ίδια οντότητα του πραγματικού κόσμου. Πρόσφατα, νευρωνικά δίκτυα και τεχνικές βαθειάς μάθησης έχουν προταθεί για εργασίες αντιστοίχισης οντοτήτων, που μαθαίνουν αναπαράστασεις οντοτήτων στο διανυσματικό χώρο που διατηρούν την ομοιότητα που έχουν στους αρχικούς γράφους. Μια μεγάλη ποικιλία μεθόδων ενσωμάτωσης με εποπτεία, χωρίς εποπτεία και με ημι-εποπτεία έχουν προταθεί στη βιβλιογραφία που εκμεταλλεύονται τις δομικές πληροφορίες των οντοτήτων στους γράφους γνώσης ή τις τιμές τους. Ακόμα, λείπει μια ποσοτική εκτίμηση των πλεονεκτημάτων και των αδυναμιών τους σε γράφους γνώσης πραγματικού κόσμου χρησιμοποιώντας διαφορετικές μετρήσεις απόδοσης και χαρακτηριστικά γράφων γνώσης. Σε αυτή τη

διατριβή, διεξάγουμε μια μετα-ανάλυση δημοφιλών μεθόδων ενσωμάτωσης για σκοπούς αντιστοίχισης οντοτήτων βασισμένη σε μια στατιστικά ορθή μεθοδολογία. Η ανάλυσή μας αποκαλύπτει ενδιαφέρουσες, στατιστικά σημαντικές συσχετίσεις διαφορετικών μεθόδων ενσωμάτωσης με διάφορα μετα-χαρακτηριστικά που εξάγονται από γράφους γνώσης. Τέλος, είμαστε σε θέση να κατατάξουμε με στατιστικά σημαντικό τρόπο την απόδοσή τους σε όλους τους πραγματικούς γράφους γνώσης που έχουμε στη διάθεσή μας και να παρουσιάσουμε τι κερδίζουμε και τι χάνουμε όσον αφορά την αποτελεσματικότητα και την απόδοση. Από όσο γνωρίζουμε, δεν υπάρχουν πειραματικές μελέτες που να παρέχουν τέτοιου είδους μετα-ανάλυση.

University of Crete

Computer Science Department

M.Sc. Thesis

Fanourakis Nikolaos

Master's Thesis Supervisor: Professor, V. Christophides

Tuesday, 21 September 2021, 11:00 a.m.

Join Zoom Meeting

<https://zoom.us/j/99736946175>

“Knowledge Graph Embedding Methods for Entity Alignment”

Abstract

In recent years, we have witnessed the proliferation of knowledge graphs (KG) in various domains, aiming to support applications like question answering, recommendations, etc. A frequent task when integrating knowledge from different KGs is to find which subgraphs refer to the same real-world entity. Recently, neural networks and deep learning techniques have been proposed for entity alignment tasks, that learn a vector-space representation (i.e., embedding) of entities which preserves their similarity in the original KGs. A wide variety of supervised, unsupervised, and semi-supervised embedding

methods have been proposed in the literature that exploit the structural information (relation based) of entities in the KGs, or their attribute values (attribute based). Still, a quantitative assessment of their strengths and weaknesses on real-world KGs using different performance metrics and KG characteristics is missing. In this thesis, we conduct a meta-level analysis of popular embedding methods for entity alignment based on a statistically sound methodology. Our analysis reveals interesting statistically significant correlations of different embedding methods with various meta-features extracted by KGs. Finally, we are able to rank in a statistically significant way their performance across all real-world KGs of our testbed, while also to demonstrate their trade-offs in terms of effectiveness and efficiency. To the best of our knowledge, there are no experimental studies that provide such a type of meta-level analysis.