BINGHAMTON UNIVERSITY SCHOOL OF COMPUTING SEMINAR SERIES

Ulugbek Ergashev - PhD Student
Resource selection in federated search

April 25, 2025
Noon-1
EB T1 or on Zoom

Abstract: In today's digital landscape, users often face challenges in accessing relevant information scattered across multiple, disparate sources. Federated Search (also known as Distributed Information Retrieval, DIR) offers a promising solution by enabling the retrieval of integrated results from heterogeneous repositories without the need for centralized indexing. However, existing resource selection methods in Federated Search are limited in capturing the complex semantic relationships between queries and resources. This research proposes novel approaches to enhance resource selection by leveraging advanced representation learning techniques. The primary contributions include FedGNN, a framework that utilizes Graph Neural Networks (GNNs) combined with pre-trained language models to capture both semantic and structural relationships between queries and resources, and Resource2Box, a method that models resources as hyperrectangular boxes in a latent space, enabling a richer representation of resource diversity. These methods aim to improve the accuracy and efficiency of resource ranking by integrating both query-resource and resource-resource relationships. Experimental results demonstrate that the proposed methods significantly outperform existing baseline models in various Federated Search tasks. This work contributes to the scalability and adaptability of Federated Search systems, making them more effective in dynamic and decentralized environments.

Bio: Ulugbek Ergashev is a Ph.D. student in the School of Computing at Binghamton University. He earned his M.S. in Computer Science from Binghamton University in 2019. His current research focuses on improving resource selection in Federated Search systems, particularly in decentralized environments where data is dispersed across multiple, heterogeneous sources. His work addresses the limitations of traditional search systems by leveraging advanced representation learning techniques, such as Graph Neural Networks (GNNs) and box embeddings, to enhance resource ranking accuracy.