

NAACL-HLT 2021

**Graph-Based Methods
for Natural Language Processing**

Proceedings of the Fifteenth Workshop

June 11, 2021

©2021 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-954085-38-1

Introduction

Welcome to TextGraphs, the Workshop on Graph-Based Methods for Natural Language Processing. The fifteenth edition of our workshop is being organized online on June 11, 2021, in conjunction with the 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-2021).

The workshops in the TextGraphs series have published and promoted the synergy between the field of Graph Theory (GT) and Natural Language Processing (NLP). The mix between the two started small, with graph theoretical frameworks providing efficient and elegant solutions for NLP applications. Graph-based solutions initially focused on single-document part-of-speech tagging, word sense disambiguation, and semantic role labeling, and became progressively larger to include ontology learning and information extraction from large text collections. Nowadays, graph-based solutions also target on Web-scale applications such as information propagation in social networks, rumor proliferation, e-reputation, multiple entity detection, language dynamics learning, and future events prediction, to name a few.

The target audience comprises researchers working on problems related to either Graph Theory or graph-based algorithms applied to Natural Language Processing, Social Media, and the Semantic Web.

This year, we received 22 submissions and accepted 17 of them for oral presentation (12 long papers and 5 short papers). Similarly to the last year, we organized a shared task on Multi-Hop Inference for Explanation Regeneration. The goal of the task was to provide detailed gold explanations for standardized elementary science exam questions by selecting facts from a knowledge base. This year's shared task on multi-hop explanation regeneration attracted four teams. Three participants' reports along with the shared task overview by its organizers are also presented at the workshop.

We would like to thank our invited speakers Laura Dietz (University of New Hampshire) and Jure Leskovec (Stanford University) for their talks. We are also thankful to the members of the program committee for their valuable and high-quality reviews. All submissions have benefited from their expert feedback. Their timely contribution was the basis for accepting an excellent list of papers and making the fourteenth edition of TextGraphs a success.

Alexander Panchenko, Fragkiskos D. Malliaros, Varvara Logacheva, Abhik Jana, Dmitry Ustalov, Peter Jansen.

TextGraphs-15 Organizers

June 2021

Program Committee

Željko Agić, Unity Technologies, Denmark
Ilseyar Alimova, Kazan Federal University, Russian Federation
Martin Andrews, Red Dragon AI, Singapore
Amir Bakarov, Behavox, Russian Federation
Tomáš Brychcín, University of West Bohemia, Czech Republic
Ruben Cartuyvels, Catholic University of Leuven, Belgium
Flavio Massimiliano Cecchini, Università Cattolica del Sacro Cuore, Italy
Mihail Chernoskutov, IMM UB RAS, Russian Federation
Yew Ken Chia, Singapore University of Technology and Design, Singapore
Rajarshi Das, University of Massachusetts, Amherst, USA
Stefano Faralli, University of Rome Unitelma Sapienza, Italy
Deborah Ferreira, University of Manchester, UK
Michael Flor, Educational Testing Service, USA
Debanjan Ghosh, Educational Testing Service, USA
Goran Glavaš, University of Mannheim, Germany
Natalia Grabar, CNRS STL UMR8163, Université de Lille, France
Aayushee Gupta, International Institute of Information Technology, Bangalore, India
Binod Gyawali, Educational Testing Service, USA
Carlos Gómez-Rodríguez, Universidade da Coruña, Spain
Rima Hazra, Indian Institute of Technology, Kharagpur, India
Tomáš Hercig, University of West Bohemia, Czech Republic
Dmitry Ilvovsky, National Research University Higher School of Economics, Russian Federation
Mohammad Javad Hosseini, University of Edinburgh, Informatics, UK
Abhik Jana, Universität Hamburg, Germany
Ming Jiang, University of Illinois at Urbana-Champaign, USA
Sammy Khalife, Ecole Polytechnique, France
Fragkiskos Malliaros, CentraleSupélec, University of Paris-Saclay, France
Valentin Malykh, Huawei Noah's Ark Lab / Kazan Federal University, Russian Federation
Gabor Melli, OpenGov, USA
Alireza Mohammadshahi, IDIAP, Switzerland
Clayton Morrison, University of Arizona, USA
Matthew Mulholland, Educational Testing Service, USA
Giannis Nikolentzos, Ecole Polytechnique, France
Enrique Noriega-Atala, The University of Arizona, USA
Damien Nouvel, INALCO ERTIM, France
Chunguang Pan, DeepBlue Technology (Shanghai) Co., Ltd, China
Aditya Girish Pawate, Indian Institute of Technology, Kharagpur, India
Jan Wira Gotama Putra, Tokyo Institute of Technology, Japan
Zimeng Qiu, Amazon Alexa AI, USA
Leonardo F. R. Ribeiro, TU Darmstadt, Germany
Michael Richter, Leipzig University, Germany
Stephen Roller, Facebook AI Research, USA
Minoru Sasaki, Ibaraki University, Japan
Viktor Schlegel, University of Manchester, UK
Rebecca Sharp, University of Arizona, USA
Artem Shelmanov, Skolkovo Institute of Science and Technology, Russian Federation
Konstantinos Skianis, BLUAI, Greece
Mark Steedman, University of Edinburgh, UK
Mokanarangan Thayaparan, University of Manchester, UK

Antoine Tixier, Ecole Polytechnique, Palaiseau, France, France
Nicolas Turenne, BNU HKBU United International College (UIC), China
Adrian Ulges, RheinMain University of Applied Sciences, Germany
Vaibhav Vaibhav, Apple, USA
Anssi Yli-Jyrä, University of Helsinki, Finland
Xiang Zhao, National University of Defense Technology, China

Organizing Committee

Alexander Panchenko, Skoltech
Fragkiskos D. Malliaros, CentraleSupélec, University of Paris-Saclay
Varvara Logacheva, Skoltech
Abhik Jana, University of Hamburg
Dmitry Ustalov, Yandex
Peter Jansen, School of Information, University of Arizona

Table of Contents

<i>Bootstrapping Large-Scale Fine-Grained Contextual Advertising Classifier from Wikipedia</i> Yiping Jin, Vishakha Kadam and Dittaya Wanvarie	1
<i>Modeling Graph Structure via Relative Position for Text Generation from Knowledge Graphs</i> Martin Schmitt, Leonardo F. R. Ribeiro, Philipp Dufter, Iryna Gurevych and Hinrich Schütze ..	10
<i>Entity Prediction in Knowledge Graphs with Joint Embeddings</i> Matthias Baumgartner, Daniele Dell’Aglia and Abraham Bernstein	22
<i>Hierarchical Graph Convolutional Networks for Jointly Resolving Cross-document Coreference of Entity and Event Mentions</i> Duy Phung, Tuan Ngo Nguyen and Thien Huu Nguyen	32
<i>GENE: Global Event Network Embedding</i> Qi Zeng, Manling Li, Tuan Lai, Heng Ji, Mohit Bansal and Hanghang Tong	42
<i>Learning Clause Representation from Dependency-Anchor Graph for Connective Prediction</i> Yanjun Gao, Ting-Hao Huang and Rebecca J. Passonneau	54
<i>WikiGraphs: A Wikipedia Text - Knowledge Graph Paired Dataset</i> Luyu Wang, Yujia Li, Ozlem Aslan and Oriol Vinyals	67
<i>Selective Attention Based Graph Convolutional Networks for Aspect-Level Sentiment Classification</i> Xiaochen Hou, Jing Huang, Guangtao Wang, Peng Qi, Xiaodong He and Bowen Zhou	83
<i>Keyword Extraction Using Unsupervised Learning on the Document’s Adjacency Matrix</i> Eirini Papagiannopoulou, Grigorios Tsoumakias and Apostolos Papadopoulos	94
<i>Improving Human Text Simplification with Sentence Fusion</i> Max Schwarzer, Teerapaun Tanprasert and David Kauchak	106
<i>Structural Realization with GGNNs</i> Jinman Zhao, Gerald Penn and huan ling	115
<i>MG-BERT: Multi-Graph Augmented BERT for Masked Language Modeling</i> Parishad BehnamGhader, Hossein Zakerinia and Mahdieh Soleymani Baghshah	125
<i>GTN-ED: Event Detection Using Graph Transformer Networks</i> Sanghamitra Dutta, Liang Ma, Tanay Kumar Saha, Di Liu, Joel Tetreault and Alejandro Jaimes	132
<i>Fine-grained General Entity Typing in German using GermaNet</i> Sabine Weber and Mark Steedman	138
<i>On Geodesic Distances and Contextual Embedding Compression for Text Classification</i> Rishi Jha and Kai Mihata	144
<i>Semi-Supervised Joint Estimation of Word and Document Readability</i> Yoshinari Fujinuma and Masato Hagiwara	150
<i>TextGraphs 2021 Shared Task on Multi-Hop Inference for Explanation Regeneration</i> Peter Jansen, Mokanarangan Thayaparan, Marco Valentino and Dmitry Ustalov	156

<i>DeepBlueAI at TextGraphs 2021 Shared Task: Treating Multi-Hop Inference Explanation Regeneration as A Ranking Problem</i>	
Chunguang Pan, Bingyan Song and Zhipeng Luo	166
<i>A Three-step Method for Multi-Hop Inference Explanation Regeneration</i>	
Yuejia Xiang, Yunyan Zhang, Xiaoming Shi, Bo Liu, Wandu Xu and Xi Chen	171
<i>Textgraphs-15 Shared Task System Description : Multi-Hop Inference Explanation Regeneration by Matching Expert Ratings</i>	
Sureshkumar Vivek Kalyan, Sam Witteveen and Martin Andrews	176

Workshop Program

Friday, June 11, 2021

10:00–10:15 **Opening Session**

10:15–11:15 **Invited Talk by Prof. Jure Leskovec (Stanford University)**

11:15–11:30 *Break*

11:30–13:10 **Oral Presentation Session - 1**

11:30–11:50 *Bootstrapping Large-Scale Fine-Grained Contextual Advertising Classifier from Wikipedia*

Yiping Jin, Vishakha Kadam and Dittaya Wanvarie

11:50–12:10 *Modeling Graph Structure via Relative Position for Text Generation from Knowledge Graphs*

Martin Schmitt, Leonardo F. R. Ribeiro, Philipp Dufter, Iryna Gurevych and Hinrich Schütze

12:10–12:30 *Entity Prediction in Knowledge Graphs with Joint Embeddings*

Matthias Baumgartner, Daniele Dell’Aglío and Abraham Bernstein

12:30–12:50 *Hierarchical Graph Convolutional Networks for Jointly Resolving Cross-document Coreference of Entity and Event Mentions*

Duy Phung, Tuan Ngo Nguyen and Thien Huu Nguyen

12:50–13:10 *GENE: Global Event Network Embedding*

Qi Zeng, Manling Li, Tuan Lai, Heng Ji, Mohit Bansal and Hanghang Tong

13:10–13:25 *Break*

Friday, June 11, 2021 (continued)

13:25–15:25 Oral Presentation Session - 2

13:25–13:45 *Learning Clause Representation from Dependency-Anchor Graph for Connective Prediction*

YanJun Gao, Ting-Hao Huang and Rebecca J. Passonneau

13:45–14:05 *WikiGraphs: A Wikipedia Text - Knowledge Graph Paired Dataset*

Luyu Wang, Yujia Li, Ozlem Aslan and Oriol Vinyals

14:05–14:20 *Selective Attention Based Graph Convolutional Networks for Aspect-Level Sentiment Classification*

Xiaochen Hou, Jing Huang, Guangtao Wang, Peng Qi, Xiaodong He and Bowen Zhou

14:20–14:45 *Keyword Extraction Using Unsupervised Learning on the Document's Adjacency Matrix*

Eirini Papagiannopoulou, Grigorios Tsoumakas and Apostolos Papadopoulos

14:45–15:05 *Improving Human Text Simplification with Sentence Fusion*

Max Schwarzer, Teerapaun Tanprasert and David Kauchak

15:05–15:25 *Structural Realization with GGNNs*

Jinman Zhao, Gerald Penn and huan ling

15:25–15:40 Break

Friday, June 11, 2021 (continued)

15:40–16:40 Invited Talk by Prof. Laura Dietz (University of New Hampshire)

16:40–16:50 *Break*

16:50–18:05 **Oral Presentation Session - 3**

16:50–17:05 *MG-BERT: Multi-Graph Augmented BERT for Masked Language Modeling*
Parishad BehnamGhader, Hossein Zakerinia and Mahdieh Soleymani Baghshah

17:05–17:20 *GTN-ED: Event Detection Using Graph Transformer Networks*
Sanghamitra Dutta, Liang Ma, Tanay Kumar Saha, Di Liu, Joel Tetreault and Alejandro Jaimes

17:20–17:35 *Fine-grained General Entity Typing in German using GermaNet*
Sabine Weber and Mark Steedman

17:35–17:50 *On Geodesic Distances and Contextual Embedding Compression for Text Classification*
Rishi Jha and Kai Mihata

17:50–18:05 *Semi-Supervised Joint Estimation of Word and Document Readability*
Yoshinari Fujinuma and Masato Hagiwara

18:05–18:10 *Break*

Friday, June 11, 2021 (continued)

18:10–18:50 Poster Session

18:10–18:50 *TextGraphs 2021 Shared Task on Multi-Hop Inference for Explanation Regeneration*
Peter Jansen, Mokanarangan Thayaparan, Marco Valentino and Dmitry Ustalov

18:10–18:50 *DeepBlueAI at TextGraphs 2021 Shared Task: Treating Multi-Hop Inference Explanation Regeneration as A Ranking Problem*
Chunguang Pan, Bingyan Song and Zhipeng Luo

18:10–18:50 *A Three-step Method for Multi-Hop Inference Explanation Regeneration*
Yuejia Xiang, Yunyan Zhang, Xiaoming Shi, Bo Liu, Wandu Xu and Xi Chen

18:10–18:50 *Textgraphs-15 Shared Task System Description : Multi-Hop Inference Explanation Regeneration by Matching Expert Ratings*
Sureshkumar Vivek Kalyan, Sam Witteveen and Martin Andrews

18:50–19:00 Closing Remarks