

# Tuning Graph2vec with Node Labels for Abuse Detection in Online Conversations

Noé Cécillon Vincent Labatut  
Richard Dufour Georges Linarès

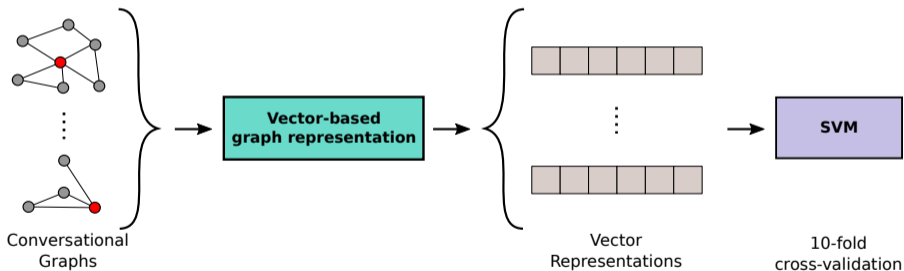
Laboratoire Informatique d'Avignon, Avignon Université – LIA EA 4128  
`{firstname.lastname}@univ-avignon.fr`

15/10/2020 – **MARAMI**

Modèles & Analyse des Réseaux :  
Approches Mathématiques & Informatiques

# Context

→ Classification problem: decide if a message is abusive or not



# Context

## Graph embedding

- Represent graphs as low-dimensional vectors
- Automatically learned
- Preserve at least a part of their topological properties
- 4 main categories of graph embedding methods
  - Node embedding
  - Edge embedding
  - Subgraph embedding
  - Whole-graph embedding
- Some methods can capture additional information
  - node labels
  - edge weights
  - edge directions

# Method

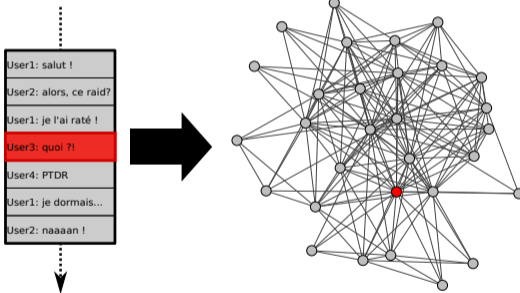
## Graph2vec

- Whole-graph embedding
  - Captures structural equivalence
  - Includes node labels in the representation learning process
- 1 Input: Set of graphs to represent
  - 2 Extracts **rooted subgraph** surrounding each node
  - 3 Fetches these subgraphs (vocabulary) to a **doc2vec SkipGram [LM14]** model.
  - 4 Output: Graph representations

# Experiments

## Data

- 1,320 messages (50% *Abuse*, 50% *Non-Abuse*)
- Constitute a conversation by leveraging the targeted message and the neighboring messages.
- Extraction of a conversational graph [Pap+19]
  - Nodes = users
  - Links = interactions



# Experiments

## Labeling strategies

- Study of the effectiveness of different node labeling strategies in the context of online abuse detection.
- **Degree**: degree of the node (default strategy)
  - **Random multiple**: random label with with a same label possibly assigned to distinct nodes over the whole corpus
  - **Random unique**: random label, each label being unique in the whole corpus
  - **Author ID**: ID of the author represented by the node
  - **Distance to target**: Distance to the targeted node in the graph
  - **Targeted**: Binary label depicting whether the node is the targeted node

# Experiments

## Results

<b>Node labeling strategy</b>	<b>Micro <math>F</math>-measure</b>	<b>Runtime</b>
Degree	80.47	8:05
Random multiple	78.26	6:59
Random unique	79.00	7:04
Author ID	81.79	7:41
Distance to target	84.03	8:28
Targeted	81.90	7:10

- **Random** labels do not bring any useful information
- Nodes with the same degree might have a similar role in the graph

# Experiments

## Results

<b>Node labeling strategy</b>	<b>Micro <math>F</math>-measure</b>	<b>Runtime</b>
Degree	80.47	8:05
Random multiple	78.26	6:59
Random unique	79.00	7:04
Author ID	81.79	7:41
Distance to target	84.03	8:28
Targeted	81.90	7:10

- Class of nodes improves the representation
- **Distance to target** combines generic and problem-related aspects



# Experiments

## Results

Node labeling strategy	Micro $F$ -measure	Runtime
Degree	80.47	8:05
Random multiple	78.26	6:59
Random unique	79.00	7:04
Author ID	81.79	7:41
Distance to target	84.03	8:28
Targeted	81.90	7:10
Topological measures [Céc+19]	88.08	8:19:10
Topological measures TF [Céc+19]	86.01	14:10

- **Topological measures** outperforms previous strategies but is much less time efficient
- **Topological measures TF** requires an important effort to determine the **top features**

# Conclusions & Perspectives

- Main results
  - Appropriate labels with Graph2vec allows improving the quality of the generated embeddings
  - Graph embeddings are much more time efficient than traditional topological measures
- Perspectives
  - Other graph embedding methods could benefit information such as node labels

# Conclusions & Perspectives

- Main results
  - Appropriate labels with Graph2vec allows improving the quality of the generated embeddings
  - Graph embeddings are much more time efficient than traditional topological measures
- Perspectives
  - Other graph embedding methods could benefit information such as node labels

**Questions?**

# References I

- [Céc+19] N. Cécillon, V. Labatut, R. Dufour, and G. Linarès. "Abusive Language Detection in Online Conversations by Combining Content- and Graph-Based Features". In: *Frontiers in Big Data* 2 (2019), p. 8. ISSN: 2624-909X. DOI: [10.3389/fdata.2019.00008](https://doi.org/10.3389/fdata.2019.00008). URL: <https://www.frontiersin.org/article/10.3389/fdata.2019.00008>.
- [LM14] Q. Le and T. Mikolov. "Distributed Representations of Sentences and Documents". In: *31st International Conference on International Conference on Machine Learning*. Vol. 32. 2014, pp. II-1188-II-1196. URL: <http://proceedings.mlr.press/v32/le14.html>.
- [Pap+19] E. Papegnies, V. Labatut, R. Dufour, and G. Linarès. "Conversational Networks for Automatic Online Moderation". In: *IEEE Trans. Comput. Social Systems* 6.1 (Feb. 2019), pp. 38-55. DOI: [10.1109/TCSS.2018.2887240](https://doi.org/10.1109/TCSS.2018.2887240).