

# A Study of Force2Vec Graph Embedding Method

Zoey Zhang & Md. Khaledur Rahman  
Indiana University Bloomington

## Introduction

Force2Vec is a parallel graph embedding algorithm that uses a force-directed model as an underlying method. The linear-algebraic formulation of the method makes it run faster on a multi-core CPU compared to other state-of-the-art methods. In this study, we explore the Force2Vec tool and run some experiments to revisit results for some scale-free graphs.

## Methodology

My research mainly used Python language. The two tables below show the quality scores of embedding by performing prediction task like link prediction, multilabel classification and clustering.

	5%	10%	15%	20%	25%
F1-Macro	0.62	0.77	0.75	0.78	0.78
F1-Micro	0.69	0.78	0.76	0.78	0.79

Table 1: Scores of node classification and clustering measures

	5%	10%	15%	20%	25%
F1-Macro	0.62	0.77	0.75	0.77	0.77
F1-Micro	0.69	0.78	0.77	0.78	0.79

Table 2: Scores of link prediction task

```
[zz13@silo Force2Vec]$ python3 -u performancescores/runnodeclassclust.py datasets/input/cora.mtx 1 datasets/output/cora.mtxF2VNS384D128IT1200NS5.embd 16 datasets/input/cora.nodes.labels
Reading graph!
Done!
Running native... 2708
Size of X: 2708
Making prediction data!!
Multilabel-classification: 0.05 F1-macro: 0.6884211986494319 F1-micro: 0.7279440342013215
Multilabel-classification: 0.1 F1-macro: 0.7373685967868278 F1-micro: 0.7584085315832649
Multilabel-classification: 0.15 F1-macro: 0.7612360013414262 F1-micro: 0.7706342311033884
Multilabel-classification: 0.2 F1-macro: 0.7676121538915339 F1-micro: 0.7798800184586987
Multilabel-classification: 0.25 F1-macro: 0.7763104913407358 F1-micro: 0.7838503200393895
```

Figure 1: code for tasks

## Results

We show the results of Force2Vec using the Cora dataset in Tables 1 and 2. These results are for node classification and link prediction tasks. We observe that the results are impressive even with a small fraction of the training dataset. In Fig. 2, we show the visualization of the Pubmed dataset using Force2Vec and DeepWalk methods. We can observe that the visualization of Force2Vec is more readable and natural.

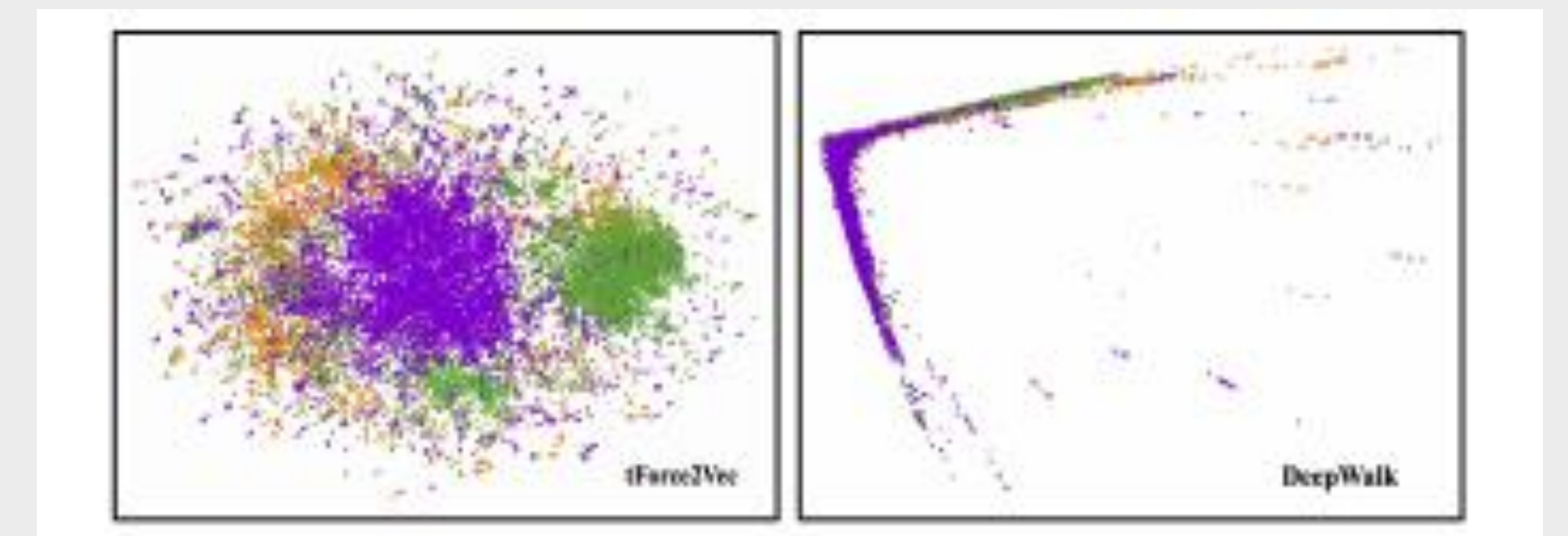


Figure 2: Visualization of 2D embeddings for Pubmed dataset generated by tForce2Vec (left) and DeepWalk (right). Colors represent respective classes in the dataset. From "Force2Vec: Parallel force-directed graph embedding".

## References

### Source1:

Rahman, M., Sujon, M., & Azad, A. (2020, September 17). Force2Vec: Parallel force-directed graph embedding. Retrieved December 10, 2020, from <https://arxiv.org/abs/2009.10035>

### Source2:

HipGraph. (n.d.). HipGraph/Force2Vec. Retrieved December 10, 2020, from <https://github.com/HipGraph/Force2Vec/tree/39cd96cb1faca3b54aaad360e682c70640b713d4>