Skip Graph Convolutional Networks (Skip-GCN)

A Framework for Hierarchical Graph Representation Learning

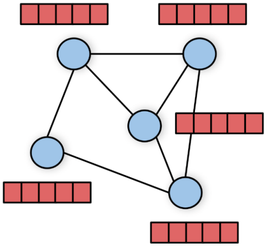
Jackson Cates, Justin Lewis, Dr. Randy Hoover, Dr. Kyle Caudle

Department of Electrical Engineering and Computer Science and Engineering,
South Dakota Mines

Feburary 7th, 2023

Introduction - Graph Theory

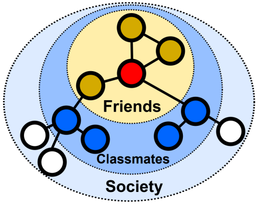
- Recently there has been great interest in graph representation learning. Graphs are known to be complex data structures because it contains both topological information and features.
- A graph $G = (V, E)$ where $|V| = n$ contains an adjacency matrix $A \in \mathbb{R}^{n \times n}$ that represents the edges.
- Node contain f features are represented with a feature matrix $X \in \mathbb{R}^{n \times f}$.



SOUTH DAKOTA MINES

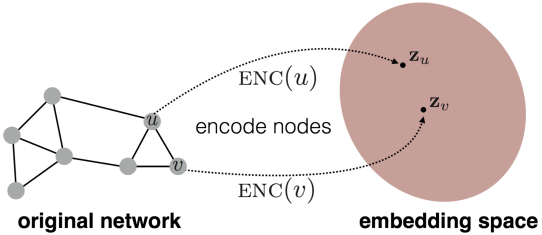
Applications

- One application of graph representation learning is the analysis and prediction of social networks. Examples include:
  - Contact tracing or contagion spread
  - Disinformation detection
  - Social media community detection
  - Covert network detection/prediction
- Social networks are known to be challenging to analyze because of topology. Local and global information can be equally important.



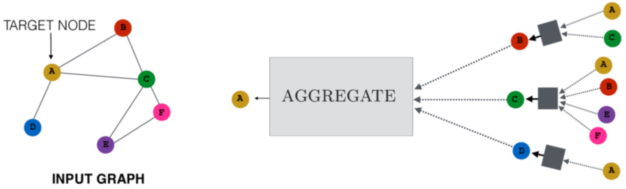
Node Embeddings

- Typically, node embeddings are used to perform machine learning tasks.
  - Node embedding is where we encode a node $u \in V$ to a vector $\mathrm{ENC}(u) = x \in \mathbb{R}^d$.
- The goal is that if nodes u and v are similar, then they will be close in the embedding space.



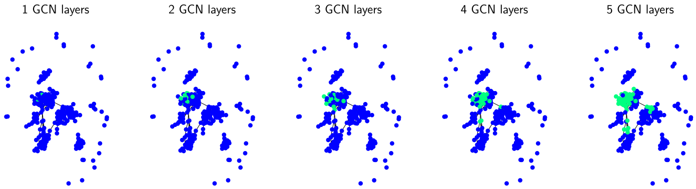
Graph Convolutional Networks (GCN)

- One popular method to compute node embeddings is through GCNs by averaging node features based on neighbors.
- A GCN layer is computed by $H_{i+1} = \sigma(\tilde{A} H_i W)$, where $\tilde{A} = D^{-\frac{1}{2}}(A + I)D^{-\frac{1}{2}}$ is the normalized adjacency matrix with D as the diagonal degree matrix of $A + I$, and $W \in \mathbb{R}^{f \times d}$ is a parameter matrix. For the first layer we use the features matrix $H_0 = X$.



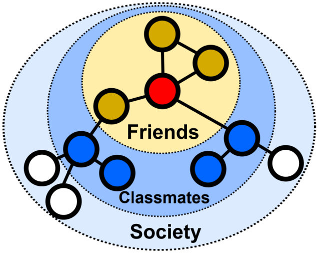
GCNs in Social Networks

- GCNs perform well for capturing local information, but lack in capturing both local and global information.
- More global information can be captured by adding more GCN layers, however at the sacrifice of diluting local node features.



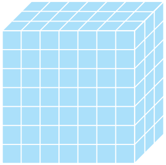
Our Proposed Method - Skip-GCN

- Our goal is to create a family of representations that capture the full range of an individual's community, both local and global.
- We aim to expand the GCN by allowing skipping in its convolutions, aptly named the skip graph convolutional network (Skip-GCN).



Multilinear Alegbra

- In order to fully capture the features along all ranges within an individual's community, we will utilize a multilinear algebra framework, utilizing *tensors*.
- A tensor $\mathcal{A} \in \mathbb{R}^{n \times n \times k}$ in this context is a multi-dimensional array, where we denote $\mathcal{A}^{(k)} \in \mathbb{R}^{n \times n}$ as a frontal slice at index k.



Tensor Products

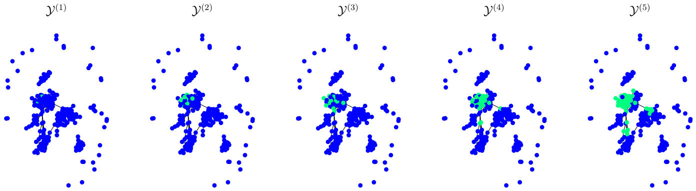
- Fundamental to the results of the current research is the product of two tensors.

Definition

Let $\mathcal{A} \in \mathbb{R}^{\ell \times n \times m}$ and $\mathcal{B} \in \mathbb{R}^{n \times p \times m}$ be tensors. The tensor-tensor product based on the L invertible linear transform $\mathcal{A} *_L \mathcal{B} \in \mathbb{R}^{\ell \times p \times m}$ is computed by applying matrix multiplication to the frontal slices of $\mathcal{A}$ and $\mathcal{B}$ in the transform domain.

Skipping

- We aim to capture the full representation of an individual's community by performing skipping in graph convolutions.
- To capture the full range, we will convolve nodes that have a walk length of i, up to a maximum parameter k, for $i = 1, \cdots, k$.
- We will then collect that into a tensor $\mathcal{Y} \in \mathbb{R}^{n \times f \times k}$, where each frontal slice is a convolution $\mathcal{Y}^{(i)} = \tilde{A}^i X$.



Skip-GCN Architecture

- After performing convolutions, we can also add a parameter tensor $\mathcal{W} \in \mathbb{R}^{f \times d \times k}$.
- Formally, our Skip-GCN layer is $\sigma(\mathcal{Y} *_L \mathcal{W}) = \mathcal{Z} \in \mathbb{R}^{n \times d \times k}$ where $\mathcal{Y}^{(i)} = \tilde{A}^i X$.
- $\mathcal{Z}$ is our node embedding for each node at each skipping size. To perform node classification, we can now flatten $\mathcal{Z}$ to a matrix in $\mathbb{R}^{n \times dk}$ and apply a simple dense layer for the embeddings.

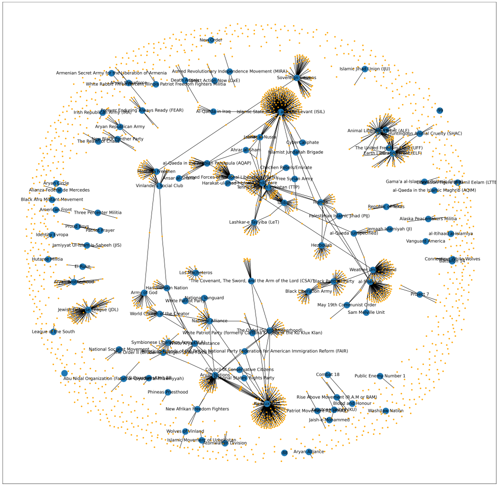
SOUTH DAKOTA MINES

Experimental Datasets

- We will compare our framework with other methods in the literature on benchmark social network datasets.
  - Profiles of Individual Radicalization in the United States (PIRUS) dataset
    - A 2226 node network that contains violent and non-violent extremists in the US from 1948-2018. Connections to individuals are based on their involvement of extremist groups. Node features contain ideology, crime history, and demographic information.

SOUTH DAKOTA MINES

PIRUS



New Order
Armenian Secret Army for the Liberation of Armenia
Armed Revolutionary Independence Movement (MIRA)
Direct Action Now (DxE)
Sovereign Citizens
Islamic Jihad Union (IJU)
Forever Enduring Always Ready (FEAR)
Irish Republican Army (IRA)
Al-Qaeda in Iraq
Islamic State of Iraq and the Levant (ISIL)
99
Aryan Republican Army
New Black Panther Party
The Realm of Chaos
Jabhat al-Nusra
Animal Liberation Front (ALF)
Stop Huntingdon Animal Cruelty (SHAC)
CyberCaliphate
The United Freedom Front (UFF)
Earth Liberation Front (ELF)
Ahrar al-Sham
Islamist Jundallah Brigade
al-Qaeda in the Arabian Peninsula (AQAP)
Chechen Rebels/Emirate
Vinlanders Social Club
Free Syrian Army
Tehrik-i-Taliban Pakistan (TTP)
al-Qaeda in the Islamic Maghrib (AQIM)
Aryan Circle
Alianza Federal de Mercedes
Black Afro Militant Movement
American Front
Three Percenter Militia
Taliban
Hamas
Republic of Texas
Palestinian Islamic Jihad (PIJ)
Alaska Peacemakers Militia
Proud Boys
Patriot Prayer
Lashkar-e-Taiyiba (LeT)
Jemaah Islamiyah (JI)
Identity Evropa
al-Qaeda (unspecified)
al-Ittihaad al-Islamiya
Vanguard America
Jamiyyat Ul-Islam Is-Saheeh (JIS)
Hezbollah
Hutaree Militia
Connecticut White Wolves
El Rukn
Los Macheteros
Weather Underground
Aryan Brotherhood
The Covenant, The Sword, and the Arm of the Lord (CSA)
Black Panther Party
al-Shabaab
Hammerskin Nation
Army of God
Black Liberation Army
Project 7
National Vanguard
White Patriot Party
May 19th Communist Order
Jewish Defense League (JDL)
Sam Melville Unit
World Church of the Creator
National Alliance
League of the South
White Patriot Party (formerly Carolina Knights of the Ku Klux Klan)
Symbionese Liberation Army (SLA)
White Aryan Resistance
National Socialist Movement
Federation for American Immigration Reform (FAIR)
Council of Conservative Citizens
Aryan Nations
Combat 18
Public Enemy Number 1
Abu Nidal Organization
Phineas Priesthood
Rise Above Movement (R.A.M or RAM)
Blood and Honour
New Afrikan Freedom Fighters
Washitaw Nation
Jaish-e-Mohammed
Wolves of Vinland
Islamic Movement of Uzbekistan
Atomwaffen Division
88
Aryan Alliance

Experimental Results

Table: K-fold cross-validation for the PIRUS dataset.

Fold	GCN	GraphSAGE	Skip-GCN
1	0.748	0.735	0.776
2	0.728	0.713	0.748
3	0.753	0.763	0.771
4	0.749	0.743	0.757
5	0.766	0.753	0.764
Avg. ± Std.	0.749 ± 0.014	0.742 ± 0.019	**0.763** ± 0.011

Note all numbers report F1 score. Higher is better.

Future Steps

- Perform experiments with more datasets.
- Explore how to incorporate edge features.
- Explore link prediction.
- Explore different aggregate methods, possibly multiple at once.
    - Min
    - Max
    - Pooling
- Explore temporal networks by adding LSTM layers.

Acknowledgments

NSF

NAVSEA
WARFARE CENTERS
DAHLGREN
NAVAL SURFACE WARFARE CENTER
DAHLGREN DIVISION

SOUTH DAKOTA
BOARD of
REGENTS

SOUTH DAKOTA MINES