1. Structural Deep Network Embedding
   1. Nueva representación para network embedding
   2. Usan proximidad de 1er orden y 2do orden para modelar la información local y global, respectivamente.
   3. Utilizan Deep learning para “aprender” la representación propuesta
2. Sparse network embedding for community detection and sign prediction in signed social networks (OJO)
   1. Nueva representación orientada a la detección de comunidades en redes con signo.
   2. Abordar las redes con signo lleva el estudio de teorías relacionadas con “balance”, ….
   3. Hacen referencia a dos trabajos anteriores. Estos se basan en entrenar mediante Deep learning.
   4. Destacan dos tipos de información relevante en el análisis de redes con signo: (a) relaciones de vecindario y (b) relaciones de vecindario en común.
   5. Experimentan la nueva representación en el problema de CD y link prediction en redes con signo.
   6. Buen trabajo para abordar el problema de detección de comunidades en redes con signo mediante una representación embedding networks.
3. Sim2vec. Node similarity preserving network embedding
   1. Proponen una nueva representación basada en similitud entre vértices donde se contempla diferentes propiedades estructurales: adyacencias, cercanía, vecindario. Según su opinión es la 1era propuesta basada en similitud entre vértices.
   2. Evalúan la propuesta en varias tareas del SNA: clasificación de vértices, predicción de enlaces, agrupamiento de vértices y visualización. Estas tareas son evaluadas respecto a los otros modelos de networks embedding, notar que se evalúan en tareas supervisadas y no supervisadas.
   3. Definen tres “dimensiones” para modelar la similitud entre vértices de la red: (a) mediante las adyacencias directas, (b) mediante la cercanía y (c) mediante el vecindario compartido.
   4. Definen una matriz de similitud basada en las dimensiones mencionadas.
   5. Definen una red neuronal basada en Autoencoder para el aprendizaje de la representación subyacente.
4. SIDE: Representation Learning in Signed Directed Networks
   1. Nueva representación orientada a redes dirigidas y con signo.
   2. Realiza una extensión/adaptación de la “fislosofía” Random walk. Añade una función de probabilidad para modelar los signos en la red.
   3. Abordan los aspectos sociales relacionados con las redes con signo: afinidad, teoría del balance, etc.
   4. Proponen aspectos para mejorar la optimización del modelo: eliminan nodos de grado 1, ya que no aportan al aprendizaje, los Random walk que tienen este tipo de vértices son redundantes. Además, proponen muestrear con los vértices de alto grado. Ambas acciones van en correspondencia a la característica de este tipo de redes de tener distribución de grado de vértices basado en Ley de potencia.
   5. Realizan un análisis asintótico riguroso introduciendo teoremas y lemas.
   6. Evalúan el método en la tarea de predicción de enlaces en redes con signo y dirigidas. Además, evalúan aspectos de rendimiento, parametrización, etc.
5. Semi supervised Community Preserving Network Embedding (OJO: revisar la idea de introducir información local de pertenencia básica a comunidades, seed)
   1. Propone una nueva representación utilizando información de comunidades.
   2. Propone un modelo Semi-supervisado basado en factorización no negativa de matriz, preservando en el embedding la información de comunidad en la red.
   3. Defienden la hipótesis de que las estructuras de comunidades ayudan a obtener una representación (embedding) más discriminatoria de cara a las tareas en el SNA.
   4. Hacen referencia a los métodos anteriores de network embedding que hacen uso de la información de comunidades. La crítica a los métodos del estado del arte está orientada a que son no supervisados (revisar porqué esto es deficiencia).
   5. La idea general es ayudar con información previa relacionada con la pertinencia o no de pares de vértices a la misma comunidad. Este preprocesamiento es lo que da el carácter Semi-supervisado del método. Hacen alusión a métodos anteriores que usan información de pares de vértices en el aprendizaje de la representación pero que se limitan a buscar similitud entre vértices sin tomar en cuenta la información de comunidad.
   6. La nueva representación, que utiliza información previa asociada a la pertenencia o no de vértices a comunidades similares, y es “aprendida” mediante un método Semi-supervisado, es evaluada en tareas de clasificación de nodos, agrupamiento de nodos y visualización.
   7. Dan una revisión del estado del arte clasificando los tipos de métodos según la supervisión y si son estructuras micro o macro. Hace pertenecer node2vec a la clase de Semi-supervisado (revisar el porqué)
   8. Evalúan la nueva representación en tareas de clasificación de nodos, agrupamiento de nodos, predicción de enlaces y visualización.
6. Scalable Multiplex Network Embedding
   1. Orientado a proponer una nueva representación para redes multi-características.
   2. Evalúan sobre problemas de predicción de enlaces y clasificación de nodos. Se testea sobre métodos tradicionales como node2vec.
7. A Survey on Network Embedding (OJO: dan muy Buena taxonomía de los métodos existentes)
   1. Da un review del estado del arte. Clasifica los métodos 1eramente en dos grandes clases: network embedding para reconstruir la red y network embedding para inferir en las redes. Definen que los métodos para reconstruir la red son una especie de subconjunto de los métodos para inferir.
   2. La clasificación está agrupara por: (a) información topológica presente en la red, (b) información atributada a los vértices y enlaces y (c) información específica a un dominio de aplicación.
      1. En (a) se tienen métodos orientados a preservar la información topológica de 1er y 2do orden: node2vec, deepWalk, LINE y métodos para preservar información de mayor grado como puede ser comunidades. Así mismo da una revisión de métodos en función de propiedades tales como: dirigidos o no, con signo o no. En este último punto la idea es preservar propiedades de la red como: no transitividad en redes dirigidas o preservar teoría del balance estructural en redes con signo. Tener en cuenta estas propiedades es importante dado que se requiere que el embedding mantenga presente estas propiedades. (los estudios relacionados con mantener las propiedades (redes dirigidas, con signo…) comentadas son de menor número que los orientados a preservar las estructuras (1er orden, 2do orden, estructuras de mayor grado como comunidades …))
      2. En el caso de (b) se tienen métodos que integran información atributada de nodos en los modelos propuestos. Estos modelos pueden ser optimizados mediante enfoque semi supervisado, donde se agrega la información presente en los nodos como información de preprocesamiento. Además, existen métodos orientados a trabajar redes atributadas en nodos y enlaces.
      3. En el caso de (c) se tienen trabajos orientados a dominios específicos, con lo cual se utiliza la información presente en estos dominios para mejorar la representación embedding. Entre los dominios referenciados están: difusión de información (se conoce acá el problema de predicción en cascada), detección de anomalías, alineamiento de redes (OJO).
   3. Los modelos referentes sobre los cuales descansan las diferentes propuestas son: matriz de factorización (ej: SVD, matriz no negativa de factorización), Random walk, Deep learning.
   4. Dan una excelente asociación de tareas abordadas mediante network embedding.
   5. Muestran una figura resumen que describe los diferentes tipos de modelos. Argumentan como la preservación de las propiedades y estructuras resulta la base del resto de los enfoques y como según el dominio de aplicación e información disponible resulta de utilidad agregar la misma en la representación propuesta.
   6. Entre los trabajos futuros están agrupados en las siguientes clases:
      1. Según las propiedades y estructuras a preservar: proponen estudiar otras estructuras complejas y útiles como redes motif, así mismo estudiar las propiedades como distribución del grado según ley de potencia, además del estudio de redes con hyper-enlaces, o sea un enlace está compuesto por más de dos vértices.
      2. Relacionado con las redes atributadas, se plantea estudiar la correlación que existe entre la información topológica y la información atributada.
      3. En cuanto a los dominios de aplicación, resulta de utilidad estudiar modelos orientados a dominios, se busca aprovechar las propiedades, características de los dominios de aplicación para obtener representaciones más eficaces.
      4. Caso particular tienen las redes dinámicas así como los espacios vectoriales. En el 1er punto poco estudio hay en abordar los modelos para redes que cambian y en el 2do punto la mayoría de los modelos mapean las redes hacia un espacio vectorial euclidiano.
8. Graph Embedding Techniques, Applications, and Performance. A Survey
   1. Dan en la introducción una explicación de los enfoques generales de las principales tareas dentro del análisis de grafo: clasificación de nodos, agrupamientos, predicción y visualización.
   2. Enumeran los desafíos a enfrentar en el network embedding: (1) selección de las propiedades útiles a preservar, (2) escalabilidad, y (3) dimensionalidad del embedding.
   3. Dan una descripción de una librería de network embedding desarrollada por ellos, así mismo comentan que es el 1er survey de las técnicas y sus aplicaciones.
   4. Excelentes definiciones relacionadas con la temática. Así mismo, dan una buena explicación de las clases de métodos según los desafíos a abordar.
   5. Agrupan los métodos en tres clases: (a) basados en factorización matricial, (b) basados en Random walk y (c) basados en Deep learning.
      1. Los métodos basados en (a) realizan transformación sobre la matriz de los datos (existen diferentes tipos, la más usada es la matriz de adyacencias) para obtener el embedding. Utilizan métodos del análisis de matemática numérica. Existen diferentes técnicas de factorización matricial, cada una con su ventaja y desventaja.
      2. Los métodos basados en (b) basan su comportamiento en estudiar o visualizar una porción de la red.
      3. Los basados en (c) hacen uso de las tecnologías Deep Learning para aprender la representación. Existen trabajos mayoritariamente que utilizan los enfoques Autoencoder y otros que usan los enfoques más supervisados. Hacen alusión a un trabajo basado en covolucional para mejorar eficiencia.
      4. Realizan una sección de discusión muy interesante, mostrando los tipos de estructuras obtenidas según los enfoques utilizados.
   6. Tienen toda una sección a los dominios de aplicación. Muy útil para enmarcar nuestro campo de estudio.
      1. En el dominio de aplicación orientado al agrupamiento de nodos se menciona una clase de métodos de agrupamientos basados en equivalencia estructural, la cual busca identificar grupos de vértices con “roles” similares (ej: nodos puentes, nodos outliers)
      2. En el dominio de aplicación orientado a predicción de enlaces hacen alusión a trabajos en este contexto y los enfoques de solución: (a) basados en similaridad gloab y/o local, (b) basados en máxima probabilidad y (c) basados en probabilidad.
      3. Para la clasificación de nodos se dan los enfoques abordados: (a) basados en extracción de rasgos y (b) basados en Random walk.
   7. Dan una serie de trabajos futuros.
9. Survey of network embedding techniques for social networks
   1. Orientado a dar una descripción de los métodos network embedding en el contexto de las redes sociales.
   2. En la sección 3 dan una formulación general de como se optimiza según la clase de método.
   3. Se define como el marco general de cualquier tipo de método consiste en tres componentes: un codificador que mapea los objetos de la red original en un vector numérico (embedding), una métrica de similitud en el espacio de la red que preserve cercanía en el espacio latente y un proceso de optimización que busca encontrar la similitud en el espacio latente.
   4. Da una taxonomía de los métodos en función de la información preservada.
   5. Tiene una tendencia a preferenciar los métodos basados en Deep Learning.
10. Community aware random walk for network embedding (OJO OJO)
    1. Introducen información de comunidades (identificadas previamente con Louvein) para mejor el embedding.
    2. Orientan el estudio hacia redes dirigidas y ponderadas.
    3. Evalúan la representación en tarea de clasificación multi clase de nodos y predicción de enlaces.
    4. Buscan preservar la información local y global, esta última mediante la información de las comunidades identificadas.
    5. Skip-gram es el modelo de optimización usado en Deep Walk y en este trabajo: de manera general la idea es que el modelo skip-gram aprende mediante optimización la mejor representación del nodo a partir de la información estructural representada en el vector información de vecindario que contempla la información de comunidades.
    6. 1eramente identifican las comunidades con Louvain. Esta información es adicionada a la información estructural construida mediante ramdom walk y luego estos vectores de información son optimizados por el modelo Skip-gram para dar la representación final (el embedding).
    7. Para obtener la estructura de información de vecindario del nodo se construyen varios (según un parámetro) Random walk adaptados a la información adicionada de comunidades.
    8. Se explica con facilidad la construcción del Random walk “adaptado”.
    9. Existen 3 momentos del método: (a) detección de las comunidades, (b) generación de los Random walk “adaptados” y (c) aprendizaje del embedding.
    10. El enfoque de este trabajo va orientado a mejorar el embedding haciendo uso de la información de las comunidades obtenidas previamente.
11. Community Preserving Network Embedding Based on Memetic Algorithm
    1. Usa información de comunidades para mejorar el embedding. O más bien busca una representación óptima para preservar la información de comunidad presente en la red (OJO).
    2. Señala un trabajo anterior por utilizar un método de CD basado en modularidad debido al problema de límite de resolución.
    3. Utiliza un enfoque evolutivo para aprender la representación: el argumento es que la optimización de una función objetivo como la densidad del modularidad es un problema NP y por tanto las metaheurísticas resultan útiles para abordar este problema en un tiempo razonable. En este caso, es utilizada los algoritmos meménticos.
    4. Combinan algoritmo genético con estrategia de aprendizaje dos niveles de heurística específica como búsqueda local. El AG es usado como método de búsqueda global y la estrategia local para acelerar la optimización.