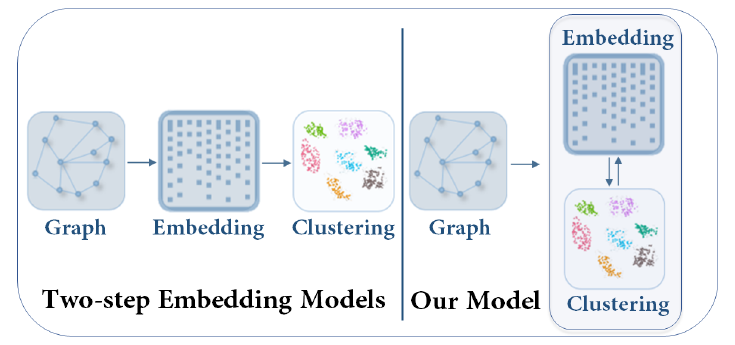
**Motivation:**

1. An overwhelming number of scenarios in which data is naturally represented in graph format rather than flat-table or vector format. Graph-based representation characterizes individual properties through node attributes, and at the same time captures the pairwise relationship through the graph structure
2. These two-step frameworks are difficult to manipulate and usually lead to suboptimal performance, mainly because the graph embedding is not goal-directed, designed for the specific clustering task. (The drawback is that the learned embedding may not be the best fit for the subsequent graph clustering task, and the graph clustering task is not beneficial to the graph embedding learning)



1. traditional goal-directed training model has GCN. Fewer studies on goal-directed embedding methods for graph clustering exist, to the best of our knowledge

**model:**

1. a goal directed graph attentional autoencoder based attributed graph clustering framework
2. encoder exploits both graph structure and node content with a graph attention network
3. decoder on the other side, reconstruct the topological graph information and manipulates the latent graph representation
4. a self-training module, which takes the ‘confident’ clustering assignments as soft labels to guide the optimizing procedure
5. specialized clustering component simultaneously learns the embedding and performs clustering in a unified framework

**Related Work:**

1. To handle both content and structure information, relational topic models, co-clustering method and content propagation
2. Limitations (not effective):
   1. only capture either parts of the network information or shallow relationships between the content and structure data,
   2. directly applied on sparse original graphs.
3. the defined clustering loss could corrupt the feature space and lead to non-representative features, so they added back the decoder and optimized the reconstruction error together with the clustering loss

**the same cluster:**

1. close distant of graph structure
2. similar attribute values

detail model:

1. Encode: learn hidden representations of each node by attending over its neighbors, to combine the attribute values with the graph structure in the latent representation

