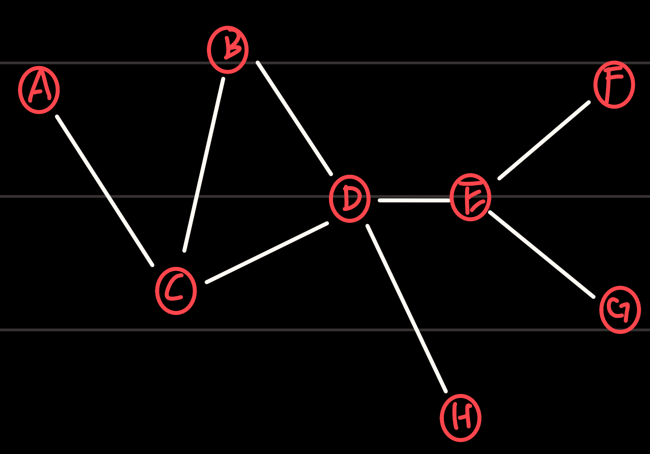
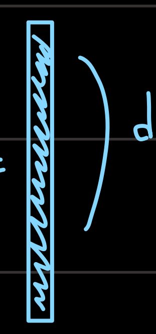
Node Embedding

Motivation

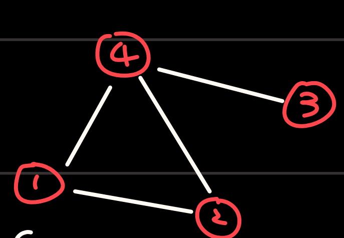


$f: node \to \mathbb{R}^d$

$F_{emb} =$

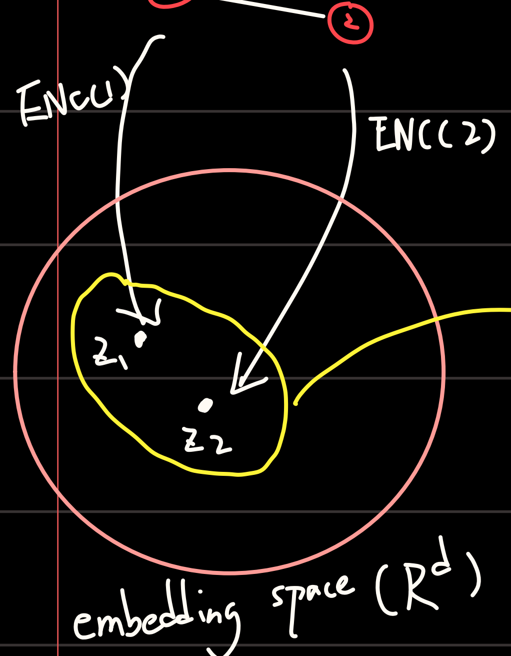


Key Point



$V = \{1, 2, 3, 4\}$

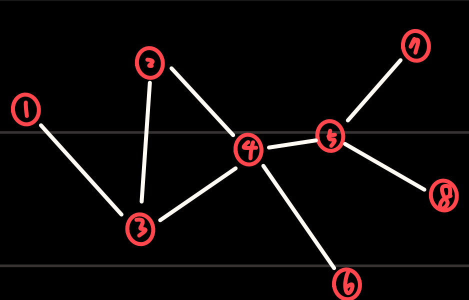
$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$



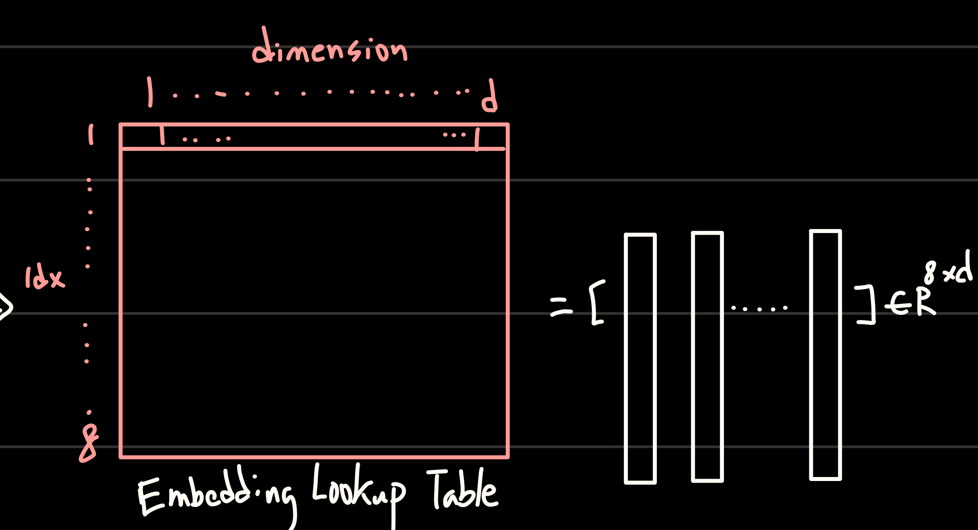
the similarity in embedding space ≈ the similarity in the graph
(e.g. cosine similarity) (e.g. link distance)

$sim(①, ②) \approx z_1^T z_2$

Def Shallow Encoding



$\to [1, \dots, 8] \to$



How to learn embedding lookup table → DeepWalk, Node2Vec

Def Random Walk for Node Embeddings

Notation:

Node u → table → emb z_u

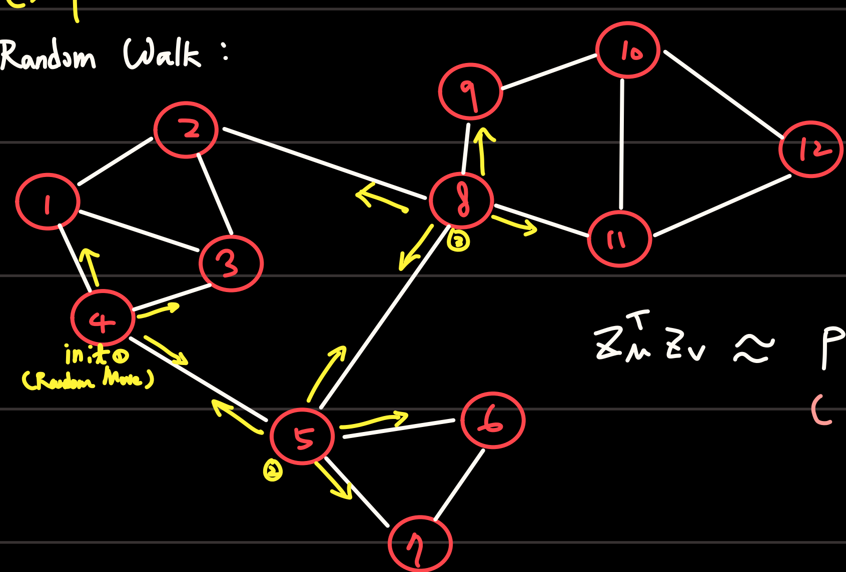
$P(v|z_u)$ = Prob of node u visit node v by Random Walk

① softmax : $z = \boxed{\ }\,k$, $\sigma(z_i) = \frac{e^{z_i}}{\sum_j^k \exp(z_j)}$, $\sigma(z) = \boxed{\ }\,k \in [0,1]_{1\times k}$

② sigmoid : $S(x) = \frac{1}{1+e^{-x}}$

$N_R(u)$: Given node u, neighborhood of u obtained by some Random Walk Strategy R
(Sequence of nodes)

Random Walk :



$z_u^T z_v \approx P(u$ and v co-occur on a random walk over the graph$)$
(if u, v close ∴ they have high frequency visiting each others)

Optimization

Vertex : All Nodes, # links

Given graph $G: G(V,E)$, learn $f: u \to \mathbb{R}^d$ denoted by $f(u) = z_u$

Optimize z_v, z_u, $\theta \to \cos(\theta) \propto P_R(v|u)$ from Random Walk Statistics where R is Random Walk Strategy

objective : $\max_f \sum_{u \in V} \log P(N_R(u) | z_u)$

Algorithm :

① Starting from each node u in graph

② $\forall N_R(u)$: optimize embeddings from → Given u, predict its neighbors $N_R(u)$

and $\max_f \sum_{u\in V} \log P(N_R(u)|z_u) \longleftrightarrow \min \sum_{u \in V} \sum_{v \in N_R(u)} -\log(P(v|z_u))$, where $P(v|z_u) = \frac{\exp(z_u^T z_v)}{\sum_{n \in V} \exp(z_u^T z_n)}$

(minimize cross entropy : $\sum_{u \in V} E_{v \in N_R(u)} [-\ln P(v|z_u)]$)

$\sum_V \sum_{N_R(u)} \frac{\cdots}{\sum_V}$ too expensive : Instead of normalizing w.r.t all nodes, normalize against k Negative Sample n_i

$\log\left(\frac{\exp(z_u^T z_v)}{\sum_{n \in V} \exp(z_u^T z_n)}\right) \approx \log(S(z_u^T z_v)) - \sum_{i=1}^{k} \log(S(z_u^T z_{n_i}))$, $n_i \sim P_V$ (Random Distribution over nodes)

Negative Sampling for n_i : Sample k negative nodes each with prob proportional to its degree ($k = 5 \sim 20$)

minimize Obj by SGD ∇_z Loss