On the von Luxburg Approximation for Hitting and Commute Times in Large Dense Digraphs

Gyan Ranjan, Zhi-Li Zhang and Daniel Boley
Department of Computer Science and Engineering, University of Minnesota, Twin Cities

Hitting and Commute Times in Random Walks

Given a directed graph $G(V, E)$, where $V(G)$ and $E(G)$ respectively represent the set of vertices and edges in the graph G, we define:

Hitting time ($H(i, j)$): The expected number of steps in a random walk that starts at vertex $i \in V(G)$ and stops when it reaches vertex $j \in V(G)$.

Commute time ($C(i, j)$): The expected number of steps in a random walk that starts at vertex $i \in V(G)$ and stops when it reaches vertex i again having visited vertex $j \in V(G)$ at least once.

Note, $C(i, j) = H(i, j) + H(j, i)$.

A popular assumption: Vertices in the same cluster of the graph have *small* commute distance, whereas two vertices in different clusters of the graph have a *large* commute distance [3].

Question: But is this always true?

The von Luxburg Approximation for Undirected Graphs

Given a *large* undirected graph, the following approximations hold[3]:

$$H(i, j) \approx \frac{Vol(G)}{d(j)} \tag{1}$$

and

$$C(i, j) = H(i, j) + H(j, i) \approx \frac{Vol(G)}{d(j)} + \frac{Vol(G)}{d(i)} \tag{2}$$

where $d(i)$ is the degree of vertex $i \in V(G)$ and $Vol(G) = \sum_{i \in V(G)} d(i)$, is the so called *volume* of the graph G.

Implication: Commute time is a *meaningless* distance function for many machine learning tasks.

In Directed Graphs

Given a directed graph $G(V, E)$, let $\boldsymbol{\pi} \in \Re^{n \times 1}$, represent the vector of steady state stationary probabilities for the set of vertices in G. If the graph is strongly connected, $\pi(i) > 0$ for all $i \in V(G)$.

Modification: The modified von Luxburg approximation for G is stated below

$$H(i, j) \approx \frac{1}{\pi(j)} \tag{3}$$

and

$$C(i, j) = H(i, j) + H(j, i) \approx \frac{1}{\pi(j)} + \frac{1}{\pi(i)} \tag{4}$$

Observation: If G is undirected $\pi(i) = \frac{d(i)}{Vol(G)}$, which is a special case of the proposed modification.

A Generalized Theoretical Framework

Let $\mathbf{A}$ be the adjacency matrix and $\mathbf{D}$ be diagonal matrix of out-degrees for the directed graph G. Then, the transition probability matrix is given as:

$$\mathbf{P} = \mathbf{D}^{-1}\mathbf{A} \tag{5}$$

From Perron-Frobenius theory [2] $\boldsymbol{\pi}'\mathbf{P} = \boldsymbol{\pi}'$, i.e. $\boldsymbol{\pi}'$ is the left eigen vector of $\mathbf{P}$ for the eigen value of 1.

The Fundamental Matrix

The so called *fundamental matrix* associated with the irreducible Markov chain representing the strong digraph, is given as [1]:

$$\mathbf{Z} = (\mathbf{I} - \mathbf{P} + \mathbf{1}\boldsymbol{\pi}')^{-1} \tag{6}$$

The hitting time between a pair of vertices i.e. source-destination pair, (i, j) can be expressed in terms of the elements of $\mathbf{Z}$ as follows:

$$H(i, j) = \frac{z_{jj} - z_{ij}}{\pi(j)} \tag{7}$$

and

$$C(i, j) = H(i, j) + H(j, i) = \frac{z_{jj} - z_{ij}}{\pi(j)} + \frac{z_{ii} - z_{ji}}{\pi(i)} \tag{8}$$

A Modified Problem

Note that:

$$\mathbf{Z} = (\mathbf{I} - \mathbf{P} + \mathbf{1}\boldsymbol{\pi}')^{-1} = \mathbf{I} + \sum_{k=1}^{\infty} (\mathbf{P} - \mathbf{1}\boldsymbol{\pi}')^k \tag{9}$$

For the von Luxburg approximation to have low errors in the case of a strong directed graph, $z_{ii} - z_{ij} \to 1$, for all $i \in V(G)$ i.e. $\mathbf{Z} \to \mathbf{I}$, the identity matrix. In other words,

$$\sum_{k=1}^{\infty} (\mathbf{P} - \mathbf{1}\boldsymbol{\pi}')^k \to 0 \tag{10}$$

Problem changes to mixing times of an irreducible Markov chain.

When is the Approximation Exact?

Does the approximation ever become exact?

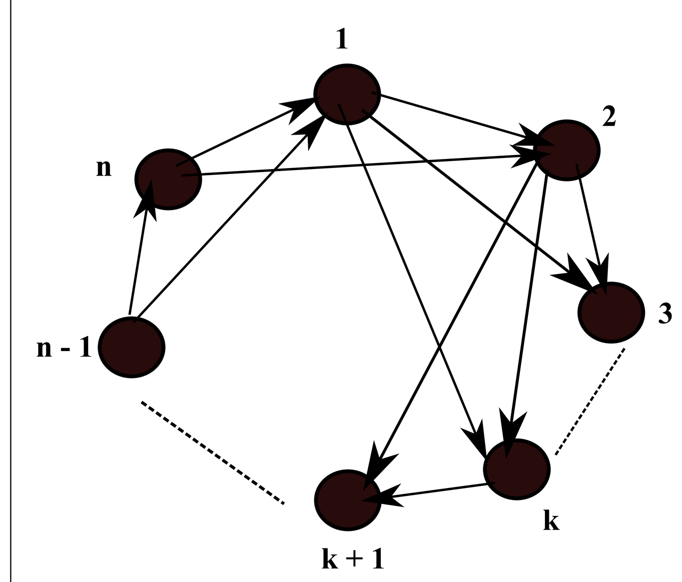
The Complete Graph (K_n)

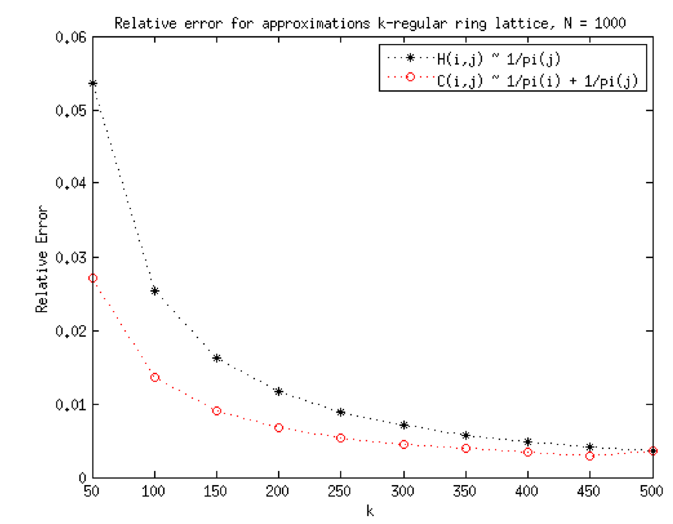
K_n is the **densest** graph of n vertices. As each vertex in K_n is connected to all the other $n - 1$ vertices;

- $[(\mathbf{P} - \mathbf{1}\boldsymbol{\pi}')]_{ii} = \frac{-1}{n}$
- $[(\mathbf{P} - \mathbf{1}\boldsymbol{\pi}')]_{ij} = \frac{1}{n(n-1)}$

As $n \to \infty$, $(\mathbf{P} - \mathbf{1}\boldsymbol{\pi}') \to 0$, and the von Luxburg approximation is exact.

An Example: k-Regular Lattice





- Even for a small directed graph (k-regular) with $n = 1000$ vertices, the error incurred seems to decrease as k increases.

Discussion

- An extension of the von Luxburg approximation is possible for the case of directed graphs.
- Modifies the problem to that of mixing times in *irreducible* Markov chains.
- The proposed framework does not depend on the interplay of random walks and electrical properties that apply to undirected graphs but do not have parallels in the case of directed graphs.
- Need to further extend the results and assess possible implications.

Literatur

[1] C. M. Grinstead and J. L. Snell. *Introduction to Probability.* American Mathematical Society, 2006.

[2] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1985.

[3] U. V. Luxburg, A. Radl, and M. Hein. Getting lost in space: Large sample analysis of the commute distance. *NIPS*, 2010.