CS598HS Advanced Social & Information Networks
Information Diffusion in Grid-like Social Networks

Guangyu Dong, Randolph Hill, Vasileios Livanos

Department of Computer Science
University of Illinois at Urbana-Champaign

gdong2@illinois.edu, *rwhill2@illinois.edu*, *livanos3@illinois.edu*

December 13, 2017

Overview

Problem Statement

Problem Statement

- We study how fast information propagates through variants of grid-like social networks.

Problem Statement

- We study how fast information propagates through variants of grid-like social networks.
- We assume a fixed-weight, directed graph G representing the social network, with nodes as agents and edges that represent the interaction between two agents.

Problem Statement

- We study how fast information propagates through variants of grid-like social networks.
- We assume a fixed-weight, directed graph G representing the social network, with nodes as agents and edges that represent the interaction between two agents.
- Each agent i has a specific type $\theta_i \in [0, 1]$, and interacts with another agent j at a time-varying frequency $f_{ij}^*(t)$.

Problem Statement

- We study how fast information propagates through variants of grid-like social networks.
- We assume a fixed-weight, directed graph G representing the social network, with nodes as agents and edges that represent the interaction between two agents.
- Each agent i has a specific type $\theta_i \in [0, 1]$, and interacts with another agent j at a time-varying frequency $f_{ij}^*(t)$.
- In our model, the information can only be transmitted between agents of the same type, and that there exists a source of information for each different agent type.

Problem Statement

- We study how fast information propagates through variants of grid-like social networks.
- We assume a fixed-weight, directed graph G representing the social network, with nodes as agents and edges that represent the interaction between two agents.
- Each agent i has a specific type $\theta_i \in [0, 1]$, and interacts with another agent j at a time-varying frequency $f^*_{ij}(t)$.
- In our model, the information can only be transmitted between agents of the same type, and that there exists a source of information for each different agent type.
- Finally, there is a constant probability p that the information will leak through a single interaction. This implies that an agent j who interacts with frequency $f^*_{ij}(t)$ at time t with another agent i who has the information, has a probability $pf^*_{ij}(t)$ of getting the information.

Motivation

Motivation

- All people interact with other people in real social networks, and most of the time we procure information through this interaction.

Motivation

- All people interact with other people in real social networks, and most of the time we procure information through this interaction.
- This information could be for example which individual is a good doctor, or a good car mechanic, and procuring this information could drastically affect our quality of life.

Motivation

- All people interact with other people in real social networks, and most of the time we procure information through this interaction.
- This information could be for example which individual is a good doctor, or a good car mechanic, and procuring this information could drastically affect our quality of life.
- Our problem is thus motivated by these examples, and we assume that there exist k external service providers, who provide services of different quality to agents of different type, but of the same quality to agents of the same type.

Motivation

- All people interact with other people in real social networks, and most of the time we procure information through this interaction.
- This information could be for example which individual is a good doctor, or a good car mechanic, and procuring this information could drastically affect our quality of life.
- Our problem is thus motivated by these examples, and we assume that there exist k external service providers, who provide services of different quality to agents of different type, but of the same quality to agents of the same type.
- Thus, the information that is propagated through the network is assumed to be the identity of the highest quality service provider for a specific type of agents.

Preliminaries

Preliminaries

- Our analysis focuses on 3 different graph structures:

Preliminaries

- Our analysis focuses on 3 different graph structures:
  1. Plain Grid (PG): A 4-regular graph where vertices are arranged in a grid pattern.

Preliminaries

- Our analysis focuses on 3 different graph structures:
  1. Plain Grid (PG): A 4-regular graph where vertices are arranged in a grid pattern.
  2. Unbiased Grid (UG): Similar to PG, but every node u forms one additional edge with some other node v, which is picked randomly, with probability proportional to $\frac{1}{d^2}$, where d is the Manhattan distance of u and v in the grid.

Preliminaries

- Our analysis focuses on 3 different graph structures:
  1. Plain Grid (PG): A 4-regular graph where vertices are arranged in a grid pattern.
  2. Unbiased Grid (UG): Similar to PG, but every node u forms one additional edge with some other node v, which is picked randomly, with probability proportional to $\frac{1}{d^2}$, where d is the Manhattan distance of u and v in the grid.
  3. Biased Grid (BG): Similar to UG but the probability of an agent u forming an edge with another agent v of the same type is proportional to $\frac{b}{d^2}$. b is called the *bias* of the graph.

Preliminaries

- Our analysis focuses on 3 different graph structures:
  1. Plain Grid (PG): A 4-regular graph where vertices are arranged in a grid pattern.
  2. Unbiased Grid (UG): Similar to PG, but every node u forms one additional edge with some other node v, which is picked randomly, with probability proportional to $\frac{1}{d^2}$, where d is the Manhattan distance of u and v in the grid.
  3. Biased Grid (BG): Similar to UG but the probability of an agent u forming an edge with another agent v of the same type is proportional to $\frac{b}{d^2}$. b is called the *bias* of the graph.
- We seek to compare the **Diffusion Rate (DR)** of information between these graphs, which is defined as the expected time it will take for the information to spread to all agents in the graph who can potentially learn it.

Preliminaries

- Our analysis focuses on 3 different graph structures:
  1. Plain Grid (PG): A 4-regular graph where vertices are arranged in a grid pattern.
  2. Unbiased Grid (UG): Similar to PG, but every node u forms one additional edge with some other node v, which is picked randomly, with probability proportional to $\frac{1}{d^2}$, where d is the Manhattan distance of u and v in the grid.
  3. Biased Grid (BG): Similar to UG but the probability of an agent u forming an edge with another agent v of the same type is proportional to $\frac{b}{d^2}$. b is called the *bias* of the graph.
- We seek to compare the **Diffusion Rate (DR)** of information between these graphs, which is defined as the expected time it will take for the information to spread to all agents in the graph who can potentially learn it.
- We seek to both develop a model and obtain some theoretical results about the DR between these graphs, as well as perform several simulations that possibly confirm our theoretical analysis.

Overview

Plain Grid

Plain Grid

- We model our system as a **Time Inhomogeneous Markov Chain**. While TIMCs are extremely important, not many theoretical results are known about them.

Plain Grid

- We model our system as a **Time Inhomogeneous Markov Chain**. While TIMCs are extremely important, not many theoretical results are known about them.
- To make the problem tractable, we assume that any non-zero interaction between two agents is lower-bounded by a constant $\delta > 0$, meaning that agents cannot have an arbitrarily small interaction.

Plain Grid

- We model our system as a **Time Inhomogeneous Markov Chain**. While TIMCs are extremely important, not many theoretical results are known about them.
- To make the problem tractable, we assume that any non-zero interaction between two agents is lower-bounded by a constant $\delta > 0$, meaning that agents cannot have an arbitrarily small interaction.

C_t: Possible Expansion Set

$$C_t = \{\{i, j\} \in E(G) \mid i \in \mathcal{I}(t), j \in \mathcal{F}(t), \theta_i = \theta_j\}$$

Plain Grid

- We model our system as a **Time Inhomogeneous Markov Chain**. While TIMCs are extremely important, not many theoretical results are known about them.
- To make the problem tractable, we assume that any non-zero interaction between two agents is lower-bounded by a constant $\delta > 0$, meaning that agents cannot have an arbitrarily small interaction.

C_t: Possible Expansion Set

$$C_t = \{\{i,j\} \in E(G) \mid i \in \mathcal{I}(t), j \in \mathcal{F}(t), \theta_i = \theta_j\}$$

Diffusion Rate

For any graph G

$$DR \leq n \sum_{t=1}^{\infty} \left(t \left(1 - (1-\delta)^{|C_t|} \right) \prod_{k=1}^{t-1} (1-\delta)^{|C_k|} \right)$$

Unbiased and Biased Grids

Unbiased and Biased Grids

Expected increase in $|C_t|$ in UG

$$\mathbb{E}\left[\left|C_t^{UG}\right| - \left|C_t^{PG}\right|\right] = \frac{\left(\sum_{k=2}^{2n} \frac{1}{k^2}\right) |\mathcal{I}(t)| \left(1 - \frac{|\mathcal{I}(t)|}{n}\right)}{N_\theta}$$

Unbiased and Biased Grids

Expected increase in $|C_t|$ in UG

$$\mathbb{E}\left[|C_t^{UG}| - |C_t^{PG}|\right] = \frac{\left(\sum_{k=2}^{2n} \frac{1}{k^2}\right) |\mathcal{I}(t)| \left(1 - \frac{|\mathcal{I}(t)|}{n}\right)}{N_\theta}$$

- We observe that the increase in DR between PG and UG is exponential!

Unbiased and Biased Grids

Expected increase in $|C_t|$ in UG

$$\mathbb{E}\left[|C_t^{UG}| - |C_t^{PG}|\right] = \frac{\left(\sum_{k=2}^{2n} \frac{1}{k^2}\right) |\mathcal{I}(t)| \left(1 - \frac{|\mathcal{I}(t)|}{n}\right)}{N_\theta}$$

- We observe that the increase in DR between PG and UG is exponential!

Expected increase in $|C_t|$ in BG

$$\mathbb{E}\left[|C_t^{BG}| - |C_t^{PG}|\right] = b \cdot \mathbb{E}\left[|C_t^{UG}| - |C_t^{PG}|\right]$$

Unbiased and Biased Grids

Expected increase in $|C_t|$ in UG

$$\mathbb{E}\left[|C_t^{UG}| - |C_t^{PG}|\right] = \frac{\left(\sum_{k=2}^{2n} \frac{1}{k^2}\right) |\mathcal{I}(t)| \left(1 - \frac{|\mathcal{I}(t)|}{n}\right)}{N_\theta}$$

- We observe that the increase in DR between PG and UG is exponential!

Expected increase in $|C_t|$ in BG

$$\mathbb{E}\left[|C_t^{BG}| - |C_t^{PG}|\right] = b \cdot \mathbb{E}\left[|C_t^{UG}| - |C_t^{PG}|\right]$$

- Similarly, the increase in DR between BG and UG is exponential, and is dependent on the bias b.

Unbiased and Biased Grids

Expected increase in $|C_t|$ in UG

$$\mathbb{E}\left[|C_t^{UG}| - |C_t^{PG}|\right] = \frac{\left(\sum_{k=2}^{2n} \frac{1}{k^2}\right) |\mathcal{I}(t)| \left(1 - \frac{|\mathcal{I}(t)|}{n}\right)}{N_\theta}$$

- We observe that the increase in DR between PG and UG is exponential!

Expected increase in $|C_t|$ in BG

$$\mathbb{E}\left[|C_t^{BG}| - |C_t^{PG}|\right] = b \cdot \mathbb{E}\left[|C_t^{UG}| - |C_t^{PG}|\right]$$

- Similarly, the increase in DR between BG and UG is exponential, and is dependent on the bias b.
- Specifically, for $b = 2$ we expect the difference in DR between BG and UG to be larger than the difference in DR between UG and PG.

Overview

Diffusion Rate Simulations

QUESTIONS ?