Artificial Intelligence Methods for Social Good

M1-4 [Optimization]:

Influence Maximization

08-537 (9-unit) and 08-737 (12-unit)

Instructor: Fei Fang

feifang@cmu.edu

Wean Hall 4126

Outline

- Propagation Process
- Influence Propagation Models
  - Independent Cascade Model
  - Linear Threshold Model
- Influence Maximization Problem

Learning Objectives

- Understand the concept of
  - Submodular function
- Describe
  - Independent Cascade Model
  - Linear Threshold Model
  - Influence Maximization Problem
  - Greedy Algorithm for Influence Maximization Problem

Propagation Process

- Viral propagation
  - Virus/Rumors
  - Get infected immediately and spread automatically
  - Individual agent does not make decisions
- Decision based models
  - Individual agent makes decisions
  - Influence and adoption

Influence Response Function

- Discuss: when would you adopt a recommendation from your friends?

Influence Response Function

- Influence Response Function
  - Independent Draws
    - n friends recommend it to me
    - $P(n) = 1 - (1 - p)^n$
    - Diminishing return (concave function)
  - Linear Threshold
    - b percentage of my friends bought the item
    - $P(b) = \delta(b > b_0)$
    - Critical Mass

Influence Propagation Models

- Independent Cascade Model (Goldenberg, 2001)
  - Model I
  - Initial set of active nodes
  - Discrete time steps
  - On every step, an active node can activate connected neighbor with a probability $p_{v,w}$ (single chance, if failed, no more trial on this edge)
  - If v succeeds, w becomes active on the next time step
  - Process runs until no more activations possible

Influence Propagation Models

- Independent Cascade Model (Goldenberg, 2001)
  - Exp 1

Quiz 1

- How many time steps are needed to achieve global cascade in Exp 1?
  - 2
  - 3
  - 4
  - 5

Influence Propagation Models

- Linear Threshold Model (M. Granovetter, 1978, T. Schelling, 1970, 1978)
  - Each node i has a threshold θ_i
  - Each edge has a weight w_{ij} indicating the influence of node i to node j
  - Activated if total weight of active neighbors exceeds threshold
  - Given initial set of active nodes, proceed iteratively with discrete time steps
  - Once activated, keep active
  - Model 2

Quiz 2

- Let θ_0 = common threshold, N_0 = common number of neighbors. $w_{ij} = \frac{1}{N_0}$. Consider the following three scenarios
  - S1: $\theta_0 = a, N_0 = b$
  - S2: $\theta_0 = a + 0.1, N_0 = b$
  - S3: $\theta_0 = a, N_0 = b + 1$
- When $b > 1$, what is ordering of the probability of getting global cascade following the LTM model under this three scenarios?
  - A: $S1 \geq S2 \geq S3$
  - B: $S3 \geq S2 \geq S1$
  - C: $S2 \geq S1, S3 \geq S1$, relationship between $S2, S3$ is unknown

Influence Maximization Problem

- How to select initial nodes A_0 to maximize influence $\sigma(A_0)$, under the constraint that A_0 has no more than k nodes
  - Problem 1
- NP-Hard (reduction from *Set Cover*, Kempe, Kleinberg & Tardos, 2003, 2005)

Greedy Algorithm

- Submodular Functions
  - Def 1
  - Diminishing return (similar to concave function)
  - Exp: Team of defensive resources
- Greedy algorithm leads to $1-\frac{1}{e}$ approximation for submodular monotone function
  - Exp: Maximum Coverage problem
- Theorem: In both LTM and ICM, $\sigma(A_0)$ is a submodular function (Kempe, Kleinberg & Tardos, 2003)
- Alg 1

Extensions

- Further propagation
  - If I bought the product, then I need to decide whether or not to recommend to others
    - May choose the level of advocating effort based on my satisfaction, e.g., twit about it, talk about it to my friend etc
- Compete with other sources of influence
  - Quit drinking/unhealthy behavior

Summary

- Propagation Process
- Influence Propagation Models
  - Independent Cascade Model
  - Linear Threshold Model
- Influence Maximization Problem

Acknowledgment

- The slides are prepared based on lecture slides of Leonid Zhukov.