A Greedy Algorithm for Budgeted Multiple-Product Profit Maximization in Social Network

Chun-Cheng Fang, Chia-Chun Ho, Bi-Ru Dai
Dept. of Computer Science and Information Engineering,
National Taiwan University of Science and Technology
Taipei, Taiwan

MobiSocial'22

220606 Chia-Chun Ho

Outline

Introduction

Related Works

Problem Formulation

Proposed Method

Experiments

Conclusions and Future Works

Introduction

Profit Maximization (PM) Problem

- Information can be disseminated widely and quickly through social networks
  - Viral marketing takes advantage of the word-of-mouth effects on social networks
- Profit Maximization (PM) problem
  - The goal is to maximize profits from viral marketing
  - The company needs to select potential influencers to propagate the product information
  - To maximize the profit for the company, the effectiveness of influence propagation and the cost of the influencers are both required to be considered

Introduction

PM Problem on Single / Multiple Products

- Most studies focus on the diffusion process for a single product
  - But in fact most companies provide several kinds of products for various demands
  - Not suitable to the business model in reality
- Therefore, we study the problem for multiple products
  - Moreover, the purchasing ability of a user should be different for different products

Introduction

Budgeted Multiple-Product Profit Maximization (BMPM) Problem

- Budgeted Profit Maximization (BPM) Problem (proposed by Zhang et al.)
  - Aimed at maximizing the overall profit of multiple products by selecting seeds with budget constraints
  - BPM problem has no limitation on the purchasing ability
- We propose the Budgeted Multiple-Product Profit Maximization (BMPM) Problem
  - Consider the consumers' purchasing ability among multiple products

Introduction

BG (BMPM Greedy)

- A greedy algorithm named BG is proposed to solve the BMPM problem
  - The purchasing ability distribution is taken consideration
    - Estimate the purchasing ability of users
    - Emphasize the importance of different products
  - The graph structure PWDAG (Product Weighted Directed Acyclic Graph) is designed
    - Approximate the influence propagation

Related Works

Influence Maximization (IM) Problem

- Select the seeds to maximize the influence
- Hill-climbing algorithms
  - Naïve Greedy
  - CELF
- Pros: high quality
- Cons: high time cost
- Heuristic algorithms
  - PMIA
  - DAGs
- Pros: less time cost
- Cons: unstable quality

Related Works

Profit Maximization (PM) Problem

- Select the seeds to maximize the profit
- Single Product
  - Without Budget Limitation
    - DGIP
  - With Budget Limitation
    - INFOCOM'18
- Multiple Products
  - Without Budget Limitation
    - RevMax-Separate
  - With Budget Limitation
    - PMIS

Problem Formulation

Expected Profit

- Given a social network graph $G = (V, E)$, a product set P, and a seed set S where each seed $s \in S$ is a pair consisting of a product $p \in P$ and a node $v \in V$

expect profit of seed set

profit of product p

$$\rho(S) = \sum_{p \in P} \sum_{v \in V} ap(S, p, v) \times profit(p)$$

activated probability that a node v will adopt a product p because of the influence from seed set S

Problem Formulation

BMPM Problem

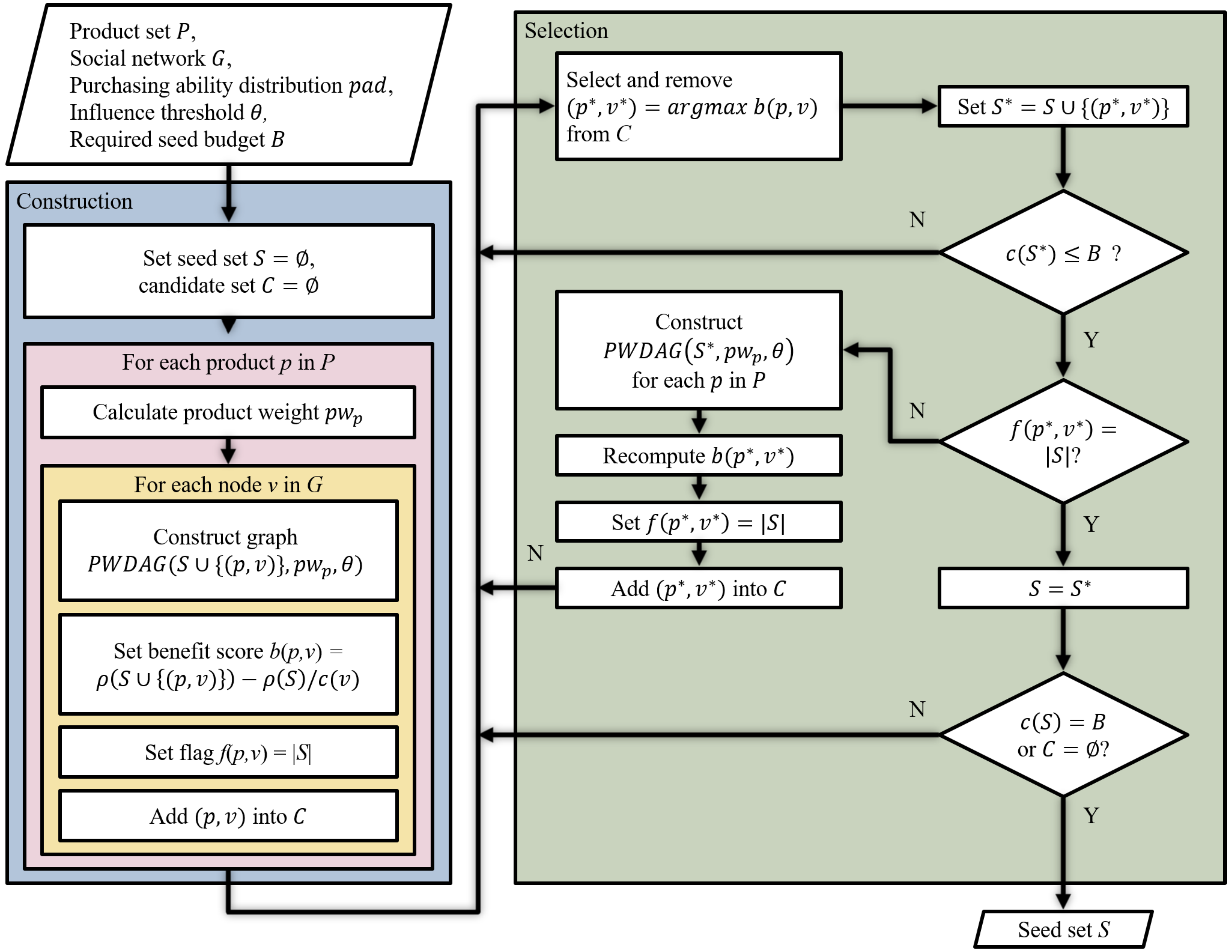
- For a given multiple-product diffusion model, such as MPIC, the goal of BMPM problem is finding a seed set for multiple products to maximize the overall profit
- Given a social network graph $G = (V, E)$, a product set P, and the seed budget limitation B, the BMPM problem aims to find the optimal seed set such that is maximized

$$\underbrace{S^*}_{\text{optimal seed set}} = \arg\max_{S \subset V} \underbrace{\rho(S)}_{\text{expect profit}} \ s.t.\ \underbrace{c(S)}_{\text{total cost of seeds in } S} \leq \underbrace{B}_{\text{seed budget limitation}}$$

Proposed Method

BG (BMPM Greedy)

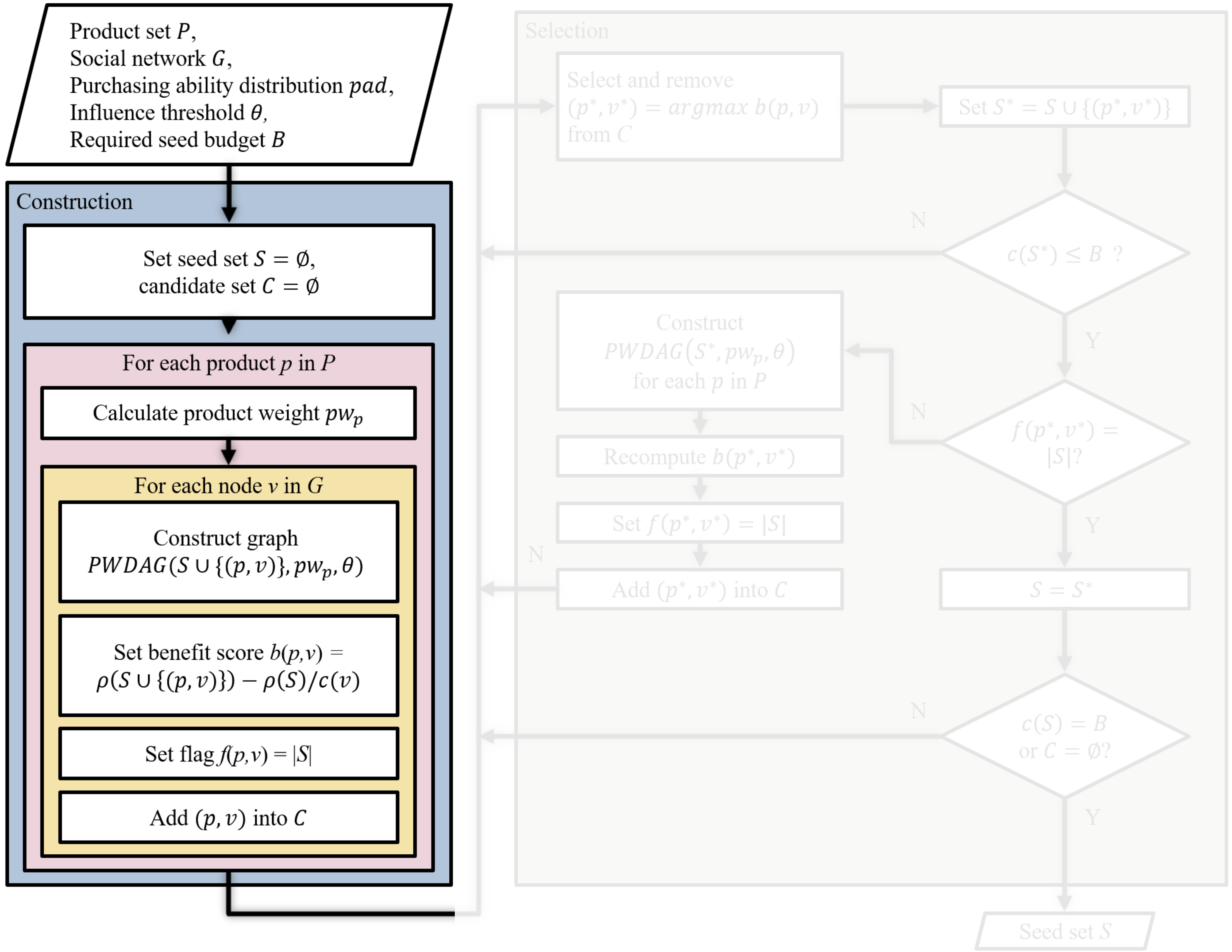
- Construction Phase
  - Calculate Product Weights
  - Construction of Influence Approximation Graphs
- Selection Phase
  - Benefit Score
  - Limitation of Selecting the Seed



Proposed Method

BG (BMPM Greedy)

- Construction Phase
  - Calculate Product Weights
  - Construction of Influence Approximation Graphs
- Selection Phase
  - Benefit Score
  - Limitation of Selecting the Seed

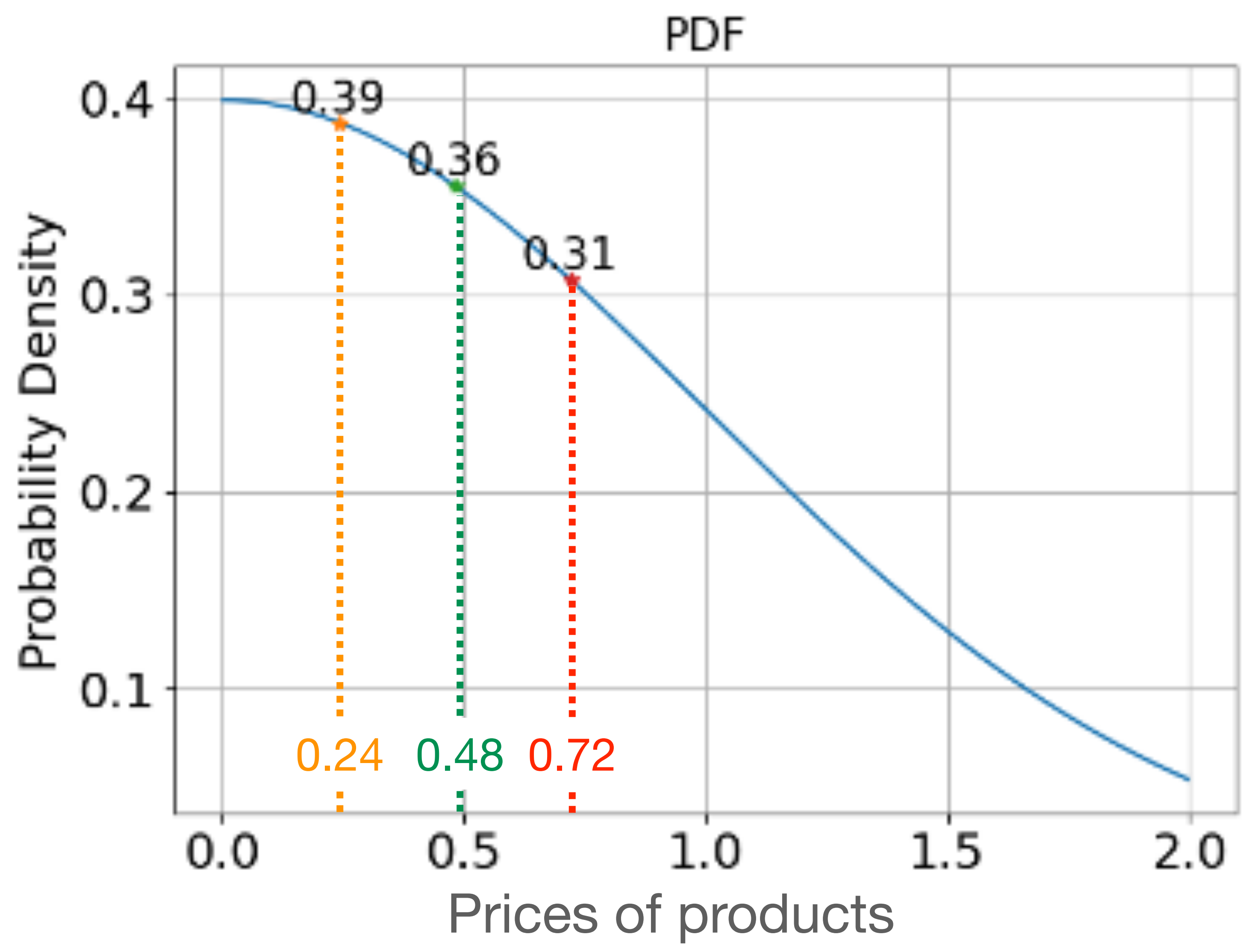


Proposed Method

Product Weights

- Emphasize the importance of different products in diffusion process
- Considered as probability of node purchasing product immediately, and propagating product information to its out-neighbors
- If price of product is higher, product weights will be lower, and vice versa

Proposed Method
Purchasing Ability



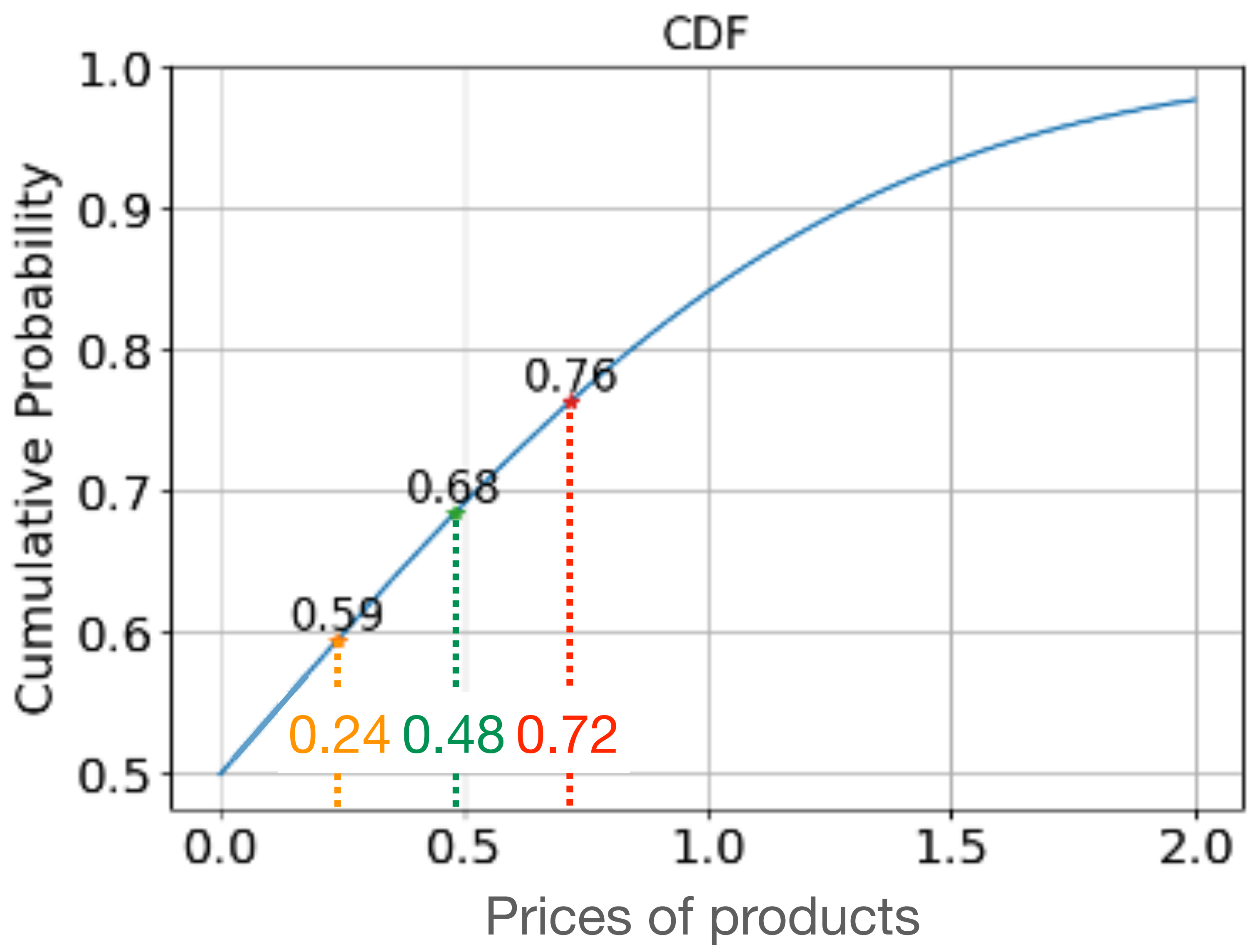
- We cannot actually know the purchasing ability of users
  - The purchasing ability distribution pad is considered as a probability density function (PDF)
  - To estimate the probabilities of users in the market to purchase given product
  - Assume pad can anticipate purchasing ability of users
- Example: product weight calculation
  - Product prices: 0.24, 0.48, 0.72
  - Purchasing ability of user: 0.39, 0.36, 0.31 (Percentage of users who can afford the price)

Proposed Method

Purchasing Ability

- Obtain CDF by integrating PDF, then calculate complementary cumulative distribution function (CCDF)

$$F_X(x) = P(X \leq x)$$



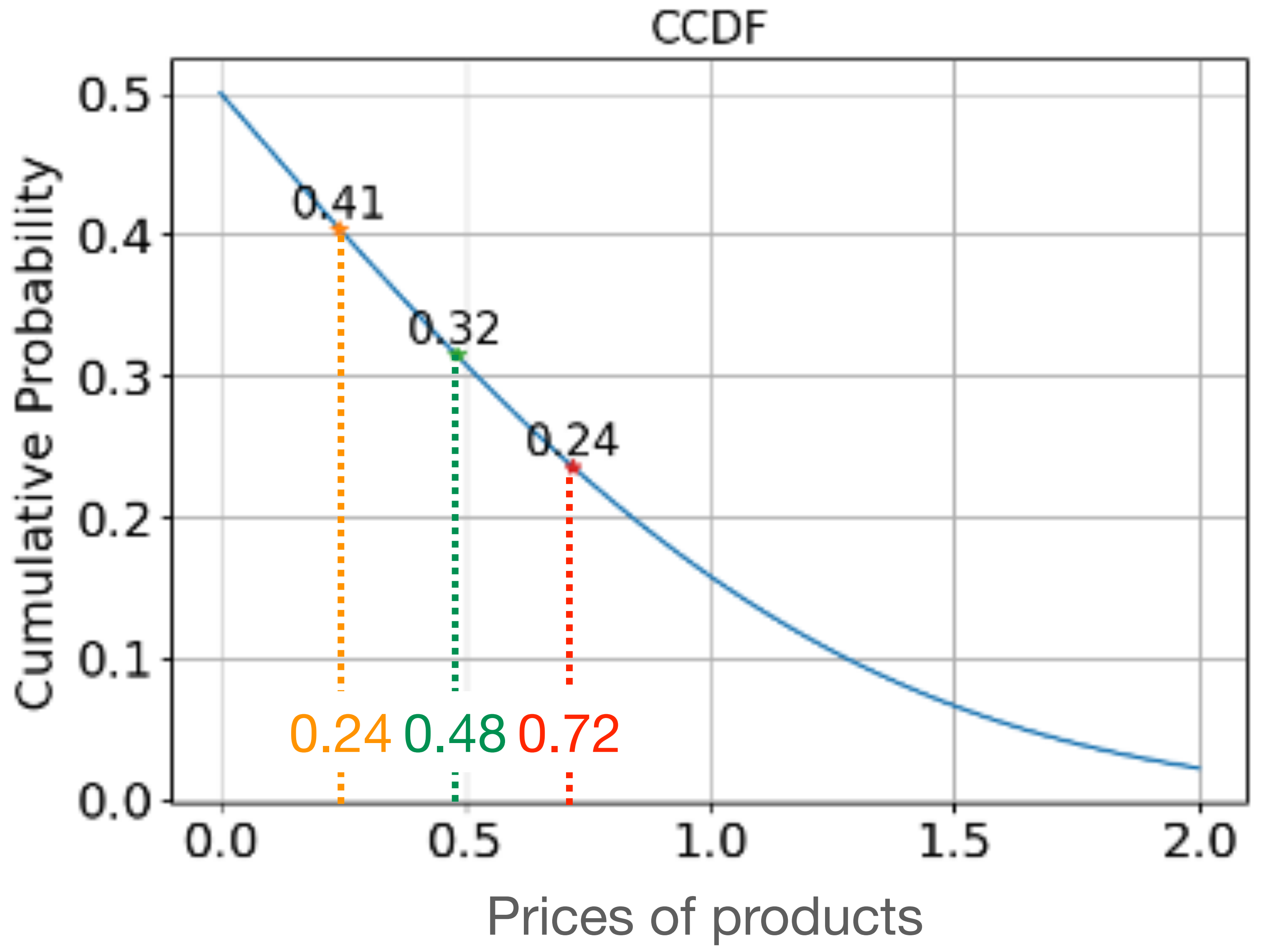
$$\bar{F}_X(x) = P(X > x)$$

random variable

price

probability that the user can afford the product at price x

$$\bar{F}_X(x) = P(X > x)$$



Proposed Method

Construction of Influence Approximation Graphs

- Product Weighted Directed Acyclic Graph (PWDAG)
  - Extension of DAG structures
    - Reduce the computing time for estimating the influence propagation
  - To approximate the influence propagation for different products
    - Small region of influence for high-priced priced products and vice versa

Proposed Method

Product Weighted Maximum Influence Path (PWMIP)

- Path from $u \rightarrow v$ with the maximum product weighted propagation probability
- Consider pw_p as purchasing probability of product p

$$PWMIP(\underbrace{p}_{\text{product}}, \underbrace{u, v}_{\text{users}}) = \underset{Path \in \underbrace{\mathbf{P}_G(u,v)}_{\text{set of all paths from } u \text{ to } v}}{\arg\max} \ \underbrace{pwpp(p, Path)}_{\text{product weighted propagation probability}}$$

$$pwpp(p, Path) = \prod_{i=0}^{|Path|-1} \underbrace{pp(v_i, v_{i+1})}_{\text{probability of edge } (v_i, v_{i+1})} \times \underbrace{pw_p}_{\text{product weight of } p}$$

Proposed Method

Product Weighted Maximum Influence Out-Arborescence ($PWMIOA$)

- Product weighted influence region of u

$$PWMIOA(p, u, \theta) = \bigcup_{pwpp(p, PWMIP(p,u,v)) \geq \theta} PWMIP(p, u, v)$$

product

user

influence threshold

user: $\forall v \in V/u$

Proposed Method

Steps of constructing the PWDAG

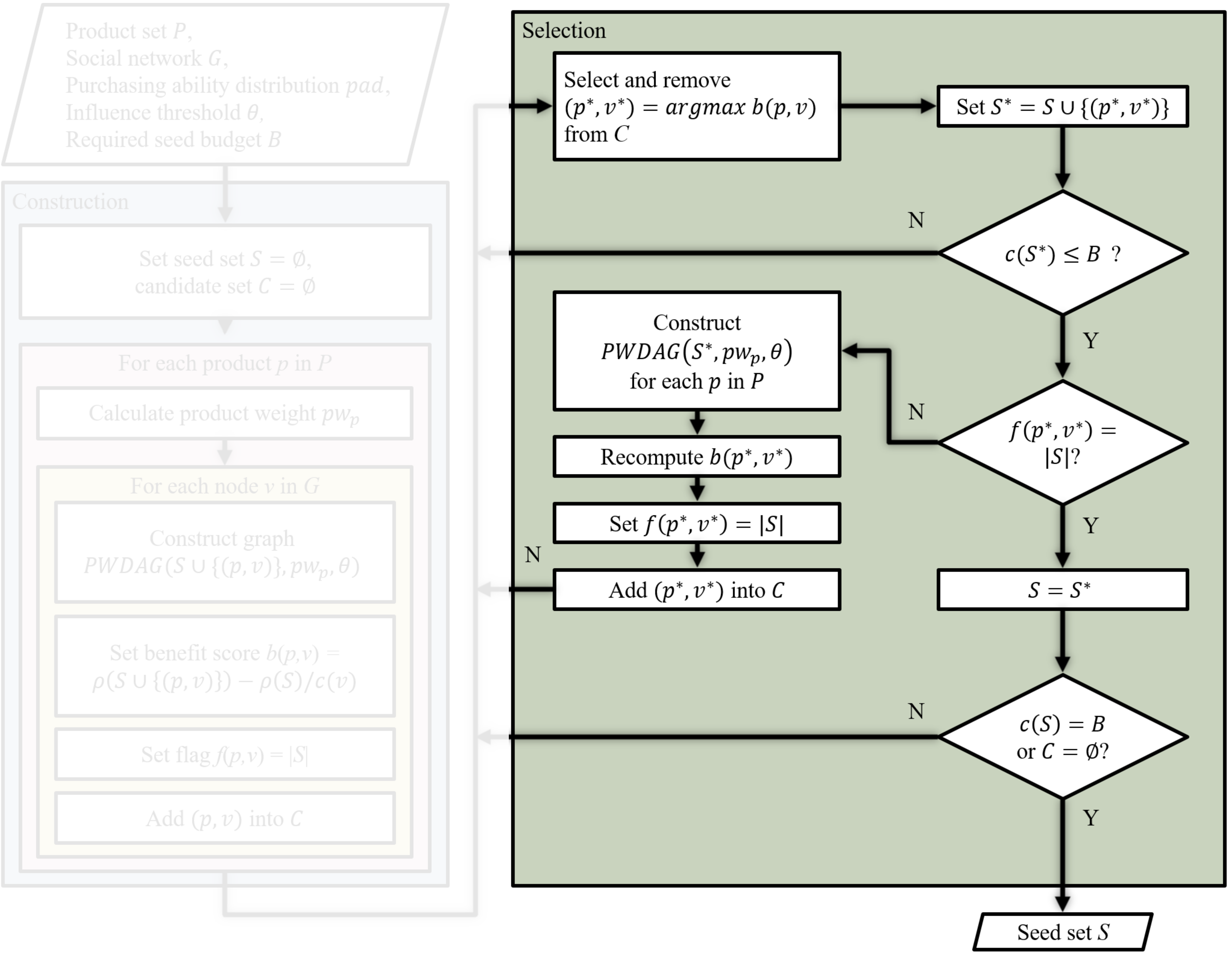
- PWDAG_1
- Regarding all seeds as a whole to calculate PWMIOA for estimating the influence propagation region of the seed set
  - Build a super root R connecting all seeds with probability 1
  - Construct PWMIOA of R
  - Remove R
  - Include some edges ending in PWDAG_1

- PWDAG_2
- Considers regions of propagation influence of different seeds and adopts their calculated propagation influence regions
  - Union PWMIOAs of seeds
  - Remove some edges

Proposed Method

BG (BMPM Greedy)

- Construction Phase
  - Calculate Product Weights
  - Construction of Influence Approximation Graphs
- Selection Phase
  - Benefit Score
  - Limitation of Selecting the Seed



Proposed Method

Benefit Score

- Select the cost-effective seeds
  - High profit seeds have high cost
    - A small amount of seeds under budget limitations, not suitable to BMPM problem

$$b(p, v) = \frac{\rho(S \cup \{(p, v)\}) - \rho(S)}{c(v)}, \forall v \in V \backslash S, p \in P$$

product

$\rho(\cdot)$: profit $c(\cdot)$: cost

Proposed Method

Limiting on Selecting the Seed

- Use Flag to record the validity of $b(p, v)$. (KDD'07)
- Record size of seed set at that time when calculating $b(p, v)$
  - $f(p, v) = |S|$
    - $b(p, v)$ is calculated with the current seed set
  - $f(p, v) < |S|$
    - $b(p, v)$ is not calculated with the current seed set

Experiments

Social Network Datasets

	Email	NetPHY
Direction	Undirected	Undirected
# of Node	1.1K	37.2K
# of Edge	5.5K	23.2K
Max Degree	105	218
Description	Node: an email address	Node: an author
	Edge: two people have communicate by email	Edge: two authors have a collaborative relationship

Experiments
Baselines

- DAG$_1$-LBP & DAG$_2$-SPBP: the best-performed algorithm on static networks and the fastest algorithm on rapidly changing communities respectively
- CELF: a greedy algorithm based on a lazy-forward optimization in selecting seeds and uses 1000 times of Monte Carlo simulation to estimate the influence
- CPHU: designed for the BPM problem that implements the cost performance update heuristics. It picks up the seed with the maximum expected number of cost performance and then updates other seeds until no more candidate seeds
- High Degree (HD): selects the nodes with the highest degree as seeds unless the seed budget limitation is exceeded
- Random: selects seeds randomly unless the seed budget limitation is exceeded.

Experiments

Proposed Methods

- BG_1
  - Use $PWDAG_1$
- BG_2
  - Use $PWDAG_2$

Experiments

Settings

- Budget of the seed set
  - 1/1024, 1/512, 1/256, 1/128 of total cost of all nodes selected as seeds
- Use the IC & WC model to generate the probability of each edge
  - IC model: $pp(u, v) = 0.1$
  - WC model: $pp(u, v) = 1/(|N_v^+|)$
- Monte Carlo simulations: 1000 times
- Influence threshold: $\theta = 0.001$

Experiments

Results

- BG_1
- BG_2
- DAG_1-LBP
- DAG_2-SPBP
- CELF
- CPUH
- HD
- Random

IC model propagation

WC model propagation

Email

NetPHY

Email

NetPHY

Conclusions and Future Works

of BG (BMPM Greedy)

- We proposed the greedy algorithm named BG to solve BMPM problem
  - Concept of purchasing ability distribution is applied to produce product weight
    - To emphasize the importance of different product
  - Graph structure using product weights reduce the time cost
    - Suitable to approximate the influence propagation for multiple products
- We will explore the relationship between products in the future work

Thanks for listening!

A Greedy Algorithm for Budgeted Multiple-Product Profit Maximization in Social Network

Chun-Cheng Fang, Chia-Chun Ho, Bi-Ru Dai
Dept. of Computer Science and Information Engineering,
National Taiwan University of Science and Technology
Taipei, Taiwan
{M10615054, M10915109}@mail.ntust.edu.tw, brdai@csie.ntust.edu.tw

Experiments

Product set & Purchasing Ability Distribution

	P
Profit	2.7, 4.2, 9
Cost	1.8, 2.8, 6
Price	4.5, 7, 15
Description	Profit is higher than cost. The products are the more profitable products in reality.

3 products from the book category in the Amazon dataset