Multi-Round Influence Maximization: A Variable Neighborhood Search Approach

9th International Conference on Variable Neighborhood Search

Isaac Lozano-Osorio (isaac.lozano@urjc.es)
Jesús Sánchez-Oro (jesus.sanchezoro@urjc.es)
Abraham Duarte (abraham.duarte@urjc.es)

Universidad Rey Juan Carlos

International Conference on Variable Neighborhood Search VNS 2022 organized by Khalifa University

grafo RESEARCH

Index

Introduction

- The **problem** addressed belongs to the family of problems of the **Social Network Influence**.
- **Maximize** the **influence** propagated in **multiple rounds** independently from possibly different seed sets.
- Real-world applications:
  - Viral Marketing.
  - Disease analysis.

Introduction

Solution Representation

A solution consists of selecting seed sets R of size l, one for each round, i.e., $S = \{S_1, S_2, \ldots, S_R\}$.

Notice that, since l nodes can be selected for each round, the total number of nodes that fit the final seed set S is equal to $l \cdot R$.

The aim of MRIM is to **maximize** the number of active nodes following a specific Influence Diffusion Model (IDM).

Introduction

Objective function

The objective function value is then evaluated as:

$$MRIM(S) = IDM(S_1 \cup S_2 \cup \ldots \cup S_R)$$

The objective of the MRIM is to find the seed set for each round that maximizes the value of the objective function value. In mathematical terms,

$$S^\star \leftarrow \arg\max_{S \in \mathbb{S}} MRIM(S)$$

where $\mathbb{S}$ is the set of all possible combinations of seed sets for the problem under consideration.

Introduction

Influence Diffusion Model

Evaluation of the influence of a given seed set S requires the definition of an Influence Diffusion Model (IDM). This model is responsible for deciding which nodes are affected by the information received from their neighboring nodes in the SN.

The most extended IDMs are:

- Independent Cascade Model (ICM)
- Weighted Cascade Model (WCM)
- Linear Threshold Model (LTM)
- Triggering Model (TM)

All of them are based on assigning an influence probability to each relational link in the SN.

Introduction

Influence Diffusion Model

- ICM, which is one of the most widely used IDMs, considers that the probability of influence is the same for each link and is usually a small probability, being 1 % a widely accepted value.
- WCM considers that the probability that a user v will be influenced by an user u is proportional to the in-degree of user v, i.e., the number of users that can eventually influence the user v. Therefore, the probability of influencing the user v is defined as $1/d_{in}(v)$, where $d_{in}(v)$ is the in-degree of user v.
- LTM requires a specific activation weight for each link in the SN. Given these weights, a user will be influenced if and only if the sum of the weights of its neighbors is larger than or equal to a given threshold.

Introduction

Influence Diffusion Model

This work considers the Triggering Model as IDM, as it is the IDM used in the best algorithm found in the literature, with the aim of providing a fair comparison.

TM is a generalization of ICM and LT where every node v independently chooses a random trigger set according to some distribution over subsets of its neighbors and is influenced if any of the nodes in its trigger sets are influenced.

The trigger set for each user v is selected in each round, conformed with those users u whose probability of directly influencing v is larger than or equal to p_{uv}, being this probability selected at random for each round.

Introduction

Adaptative vs Non Adaptative

- **Non-adaptive** approach consisting of selecting a number l of nodes per round without having information about the influenced users in each round.
- **Adaptive** approach, users are influenced after the selection of each round is known, increasing the information available for the next rounds.

Introduction

Adaptative vs Non Adaptative

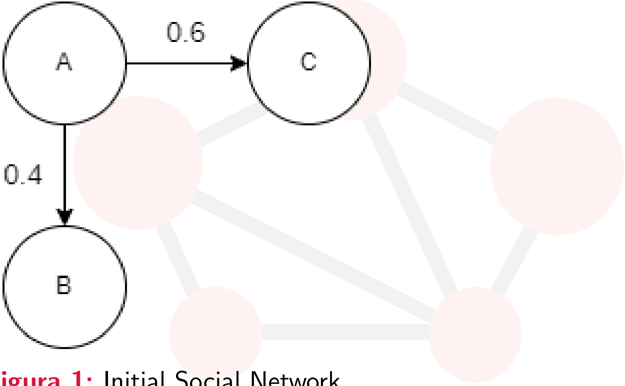


Figura 1: Initial Social Network.

Introduction

Adaptative vs Non Adaptative

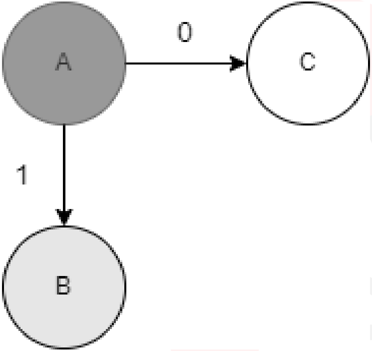


Figura 2: Real world influence selecting A

Introduction

Literature Review

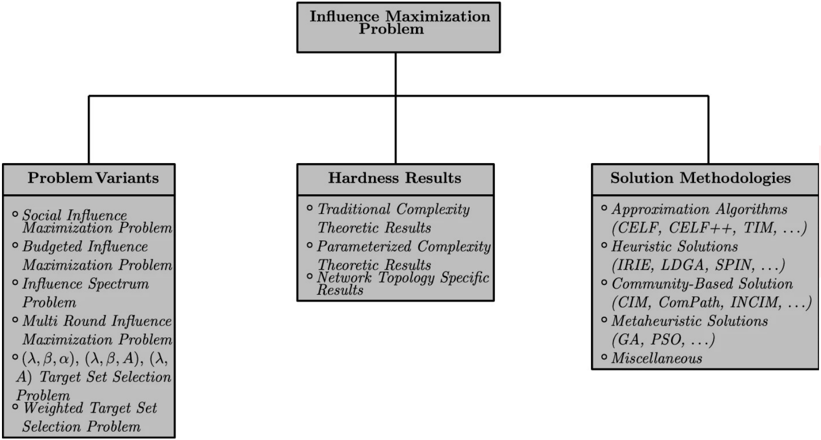
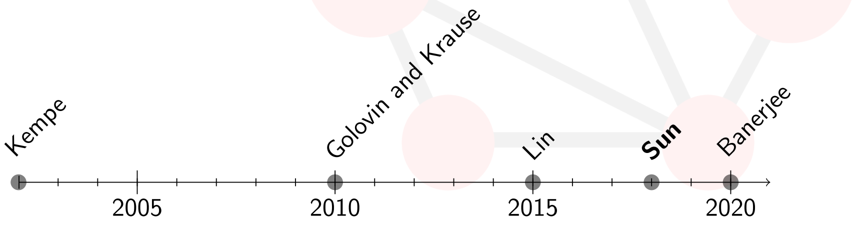


Figura 3: A survey on influence maximization in a social network, Banerjee et al (2020).

Introduction

Literature Review

If a single round is considered, i.e., $R = 1$, the problem is equivalent to the well-studied Social Network Influence Maximization Problem (SNIMP).
The classical SNIMP is $\mathcal{NP}$-hard (Kempe, 2015), so both the non-adaptive and adaptive versions of MRIM are also $\mathcal{NP}$-hard.



Proposal

Variable Neighborhood Search (VNS)

- Proposed by Nenad Mladenovic and Pierre Hansen in 1997.
- Main contributions:
  1. **Consider several neighborhoods during the search**
  2. **Perform systematic changes in the neighborhood structures**

Proposal

BVNS

Algorithm 1 $BVNS(k_{\text{máx}}, R, I)$

$S \leftarrow \emptyset$
while $R > 0$ **do**
  $S \leftarrow Construct(S, I)$
  $S \leftarrow LocalSearch(S)$
  $k \leftarrow 1$
  while $k \leq k_{\text{máx}}$ **do**
    $S' \leftarrow Shake(S, k)$
    $S'' \leftarrow LocalSearch(S')$
    $k \leftarrow NeighborhoodChange(S, S'', k)$
  end while
  $R \leftarrow R - 1$
end while
return S

Proposal

Constructive

Greedy algorithm which uses the objective function as greedy function value.

- Calculate the number of activated nodes when selecting the next seed node u, $MRIM(u)$.
- With the aim of reducing the computational effort, the number of iterations required by Monte Carlo is reduced.

Selecting nodes for the next rounds depends on the approximation considered.

- Non-adaptive model or adaptive model. In the former, the method selects the next l nodes that are the most influential.
- Adaptative model select those which are able to influence a larger number of non-previously influenced nodes.

Proposal

Local Search

The move considered is a swap move $Swap(S, u, v)$ where node u is removed from the seed set, being replaced by v, with $u \in S$ and $v \notin S$. This swap move is formally defined as:

$$Swap(S, u, v) = (S \setminus \{u\}) \cup \{v\}$$

Thus, the neighborhood $N_s(S)$ of a given solution S consists of the set of solutions that can be reached from S by performing a single swap move. More formally,

$$N_s(S) = \{Swap(S, u, v) \quad \forall u \in S, \forall v \in V \setminus S\}$$

Proposal

Local Search

The size of the resulting neighborhood, $l \cdot (n - l)$, makes the complete exploration of the neighborhood not suitable for MRIM, even considering an efficient implementation of objective function evaluation.

Then, the intelligent neighborhood exploration strategy proposed in Lozano-Osorio 2021 is followed, with the aim of reducing the number of solutions explored within each neighborhood (so limit the number of IDM simulations). This reduction in the size of the search space is performed by exploring just a small fraction δ of the available nodes for the swap move.

Proposal

Shake

The perturbation mechanism in VNS is usually called the Shake procedure.

We propose a method that modifies the structure of the solution according to a parameter k. Its value ranges from 1 to k_{max}, which is an input parameter of the complete procedure.

The proposed shake method performs k swaps moves to the incumbent solution. These elements are selected at random.

Computational Results

- Programming Language: **Java 17**.
- *Metaheuristic Optimization frameworRK* (MORK) 13.
- Experimental machine features: Intel Core i7-9750H (2.6 GHz) with 16GB RAM.
- Instances: 2 (Flixter with 95969 nodes and 484865 directed edges, NetHEPT 15233 nodes and 32235 directed edges).
- Performance metrics:
  - **Avg.**: the average objective function value (i.e., the number of nodes influenced, on average, after 100 simulations).
  - **Dev.**: average deviation with respect to the best known solution.
  - **Time (s)**: execution time measured in seconds.
  - **#B**: the number of times that the algorithm is able to reach the best solution in the experiment.

Computational Results

Parameters used

- Number of rounds: 5.
- Influence Diffusion model iterations: 100.
- Local search δ parameter: 25.
- Seed set nodes S size (l): 10.
- k parameter to modify the structure of the solution: 0.1.

Computational Results

Non-adaptative version

	R	BVNS Avg.	BVNS Dev.	BVNS Time (s)	BVNS #B	AdaIMM Avg.	AdaIMM Dev.	AdaIMM Time (s)	AdaIMM #B
NetHEPT	1	**355.89**	0.00	0.34	1	302.40	15.03	0.21	0
	2	**584.66**	0.00	0.49	1	544.40	6.89	0.28	0
	3	719.36	5.57	0.61	0	**761.80**	0.00	0.33	1
	4	900.89	6.72	0.75	0	**965.80**	0.00	0.51	1
	5	1015.53	11.41	0.82	0	**1146.30**	0.00	0.60	1
Flixster	1	**16637.78**	0.00	78.21	1	13560.51	18.50	6.59	0
	2	**17964.82**	0.00	89.16	1	13349.29	25.69	7.88	0
	3	**18288.07**	0.00	98.57	1	13655.16	25.33	9.50	0
	4	**18603.10**	0.00	106.82	1	14033.71	24.56	13.33	0
	5	**19120.03**	0.00	114.90	1	14820.86	22.49	17.22	0
		9419.01	**2.37**	49.07	**7**	7312.96	14.15	**5.64**	3

Tabla 1: Results of the BVNS algorithm versus the state-of-the-art procedure in the non-adaptive version. The best results are highlighted with bold font.

Computational Results

Adaptative version

		BVNS				AdaIMM			
	R	Avg.	Dev.	Time (s)	#B	Avg.	Dev.	Time (s)	#B
NetHEPT	1	**355.89**	0.00	0.34	1	302.40	15.03 %	0.21	0
	2	**584.66**	0.00	0.52	1	557.20	4.70 %	0.35	0
	3	**806.20**	0.00	0.68	1	776.90	3.63 %	0.38	0
	4	913.99	6.93	0.81	0	**982.00**	0.00 %	0.41	1
	5	1021.09	12.24	0.99	0	**1163.50**	0.00 %	0.56	1
Flixster	1	**16637.78**	0.00	78.21	1	13611.37	18.19 %	7.11	0
	2	**17831.39**	0.00	95.31	1	13749.83	22.89	7.99	0
	3	**18364.20**	0.00	106.87	1	13855.42	24.55	9.21	0
	4	**18929.40**	0.00	109.22	1	14213.69	24.91	13.24	0
	5	**19310.31**	0.00	121.23	1	14863.01	23.03	16.15	0
		9475.49	**1.92**	51.42	**8**	7407.53	13.69	**5.56**	2

Tabla 2: Results of the BVNS algorithm versus the state-of-the-art procedure in the adaptive version. The best results are highlighted with bold font.

Conclusions

- The proposed algorithm obtains competitive results with **state-of-the-art**.

- **Future work**: increased number of instances, work on each version individually, test other metaheuristics such as VND, GVNS, and improve computational time.

Multi-Round Influence Maximization: A Variable Neighborhood Search Approach

9th International Conference on Variable Neighborhood Search

Isaac Lozano-Osorio (isaac.lozano@urjc.es)
Jesús Sánchez-Oro (jesus.sanchezoro@urjc.es)
Abraham Duarte (abraham.duarte@urjc.es)

Universidad Rey Juan Carlos

International Conference on Variable Neighborhood Search VNS 2022 organized by جامعة خليفة Khalifa University

grafo RESEARCH