VISVESVARAYA TECHNOLOGICAL UNIVERSITY

BELAGAVI-590018

Synopsis
for
Project Phase – I
(22ISP65)

Maximizing Influence In Social Networks

Submitted in the partial fulfillment of the requirement for the Project Phase – I

Bachelor of Engineering
In
Information Science and Engineering

Submitted by

Bhavya T S	**1NC22IS008**
Kavyashree C T	**1NC22IS024**
Megha Ari	**1NC22IS033**
Swaroop Patil	**1NC22IS057**

Under the Guidance

Mrs. Ankita Shukla

Assistant Professor

Dept. of ISE, NCET

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

NAGARJUNA COLLEGE OF ENGINEERING AND TECHNOLOGY
(An Autonomous Institution under VTU, Belgavi-590018)

VENKATAGIRIKOTE, DEVANAHALLI, BENGALURU– 562164

2024-2025

TABLE OF CONTENTS

Title: Maximizing Influence In Social Networks

1. Introduction

Maximizing influence in social networks can be explained in a broader sense. To begin, consider a social network as a directed graph with users as nodes and their connections to one another as directed edges.

This product is promoted throughout the social network graph via the word-of-mouth effect, which is the effect of human-to-human transmission. Promotion can result in either rapid decline or rapid growth.

These two outcomes depend on three key questions.

The first question is how to choose the people who are initially given the right to try the product. This is the seed node set S mentioned later.

The second question is, how does this small group of people get other people to be influenced to buy the product? In fact, this process is the capability of different diffusion models.

The last question is how to estimate the number of users affected by this small group of people, i.e., the most influential people. This is actually the problem of estimating the propagation influence of a given seed set.

According to the three problems mentioned above, the influence maximization problem in social networks can be defined as finding the seed set composed of the smallest number of nodes in the social network so that they can influence the largest number of users in the social network; that is, influence maximization.

2. Problem Statement

In the era of digital communication and online social platforms, maximizing the spread of information, products, or ideas through social networks has become a significant research area. This challenge, known as **Influence Maximization**, focuses on identifying a small subset of influential individuals within a social network who can trigger a large cascade of further adoptions through word-of-mouth propagation.

A social network can be modeled as a **directed graph** G=(V,E), where nodes V represent users and directed edges E denote relationships or connections through which influence can flow. Given a limited budget or constraint on the number of initial users (seed set S), the objective is to select the most influential k users such that the spread of influence throughout the network is maximized according to a given **diffusion model** (such as the Independent Cascade or Linear Threshold model).

3. Motivation, Justification, and Scope

Motivation:

The widespread use of social networks has revolutionized how information spreads.Organizations aim to maximize impact with minimal promotional effort.Influence Maximization (IM) helps identify key individuals for optimal diffusion.It is critical for applications like viral marketing, health awareness, and politics.The challenge lies in predicting how influence spreads and selecting ideal seed users.Understanding and solving the IM problem offers both practical and academic value.

Justification:

IM is widely applicable in business, healthcare, politics, and digital campaigns.It addresses core challenges in maximizing reach with limited resources.Theoretical challenges like NP-hardness make it an exciting algorithmic problem.Recent advances in deep learning and context-aware models demand new approaches.Traditional models fail to capture the dynamic nature of real-world networks.Hence, research in IM is both practically valuable and scientifically significant.

Scope:

This work studies the IM problem within the context of social networks modeled as graphs.It explores classical, context-aware, and deep learning-based diffusion models.It analyzes major algorithmic strategies: greedy, heuristic, sampling, and metaheuristics.The study evaluates performance, complexity, and approximation guarantees.Contextual aspects like time, location, and topic based propagation are also considered.Overall, it aims to provide a comprehensive understanding of IM methods and challenges.

4. Literature Survey and Background Study

S. No	Title of Research Work	Author(s)	Year	Key Contribution
1	Context-Aware IM (Time, Location, Topic)	Ye et al.	2022	Extends IM by integrating context like location and time into seed selection.
2	Classification of IM Algorithms	Ye et al.	2022	Categorizes IM methods into 4 types: simulation, heuristic, sampling, and learning.
3	DeepCas: Deep Learning for Cascade Prediction	Li et al.	2017	End-to-end deep learning model to predict cascade sizes using GRU.
4	IMM: Influence Maximization with Martingales	Tang et al.	2015	Improved scalability with lower memory usage and similar guarantees as TIM.
5	TIM and TIM+: Reverse Reachability Sampling Methods	Tang et al.	2014	Efficient sampling-based approach with approximation guarantees.
6	CELF++: Improved Greedy Optimization	Goyal et al.	2011	Speeds up CELF by minimizing marginal gain recalculations.
7	HEURISTICS: Degree Discount & PageRank for IM	Chen et al.	2009	Introduced scalable influence approximation techniques for large networks.
8	CELF: Cost-Effective Lazy Forward Selection	Leskovec et al.	2007	Optimizes greedy algorithm by reducing influence spread evaluations.
9	Influence Maximization Problem Definition	Kempe et al.	2003	Formally defines IM as a discrete optimization problem; proves NP-hardness.
10	Linear Threshold & Independent Cascade Models	Kempe et al.	2003	Introduced two foundational diffusion models used widely in IM research.

5. Objectives of the Project Work

- To formally define the Influence Maximization (IM) problem in the context of social networks using graph-based representations.
- To study and analyze various diffusion models such as Independent Cascade, Linear Threshold, and their context-aware extensions.
- To evaluate existing algorithmic approaches for solving the IM problem, including greedy, heuristic, sampling-based, and meta-heuristic methods.
- To compare the performance and efficiency of classical and modern IM algorithms based on spread, time complexity, and scalability.
- To explore recent advancements in context-aware and deep learning-based influence maximization models.
- To identify research gaps and propose future directions for more realistic, scalable, and context-sensitive IM solutions.

6. Methodology

The project methodology consists of the following key phases:

1. Problem Definition: Define the Influence Maximization (IM) problem as selecting a seed set of users in a social network graph $G = (V, E)$ such that the spread of influence $\sigma(S)$ is maximized under a given diffusion model.

2. Diffusion Model Selection: Analyze and select appropriate diffusion models such as Independent Cascade (IC), Linear Threshold (LT), or Triggering models to simulate influence propagation across the network.

3. Algorithm Classification: Categorize existing IM algorithms into simulation-based, heuristic-based, sampling-based, and meta-heuristic-based approaches. Understand their working principles and evaluate their strengths and weaknesses.

4. Context-aware Modeling: Incorporate real-world factors like user location, time of information dissemination, and topic relevance to improve the accuracy and effectiveness of influence maximization.

5. Comparative Analysis: Perform a theoretical and empirical comparison of algorithms based on parameters like time complexity, approximation guarantees, and scalability. Use standard benchmarks or datasets where applicable.

6. Tool Implementation: Simulate influence spread using computational tools or frameworks that support Monte Carlo simulations or reverse reachable set sampling (e.g., TIM, IMM). Visualize results using graphs or tables.

7. Result Evaluation: Compare the influence spread and efficiency of different algorithms before and after incorporating context-aware factors or optimization techniques to assess improvements in performance.

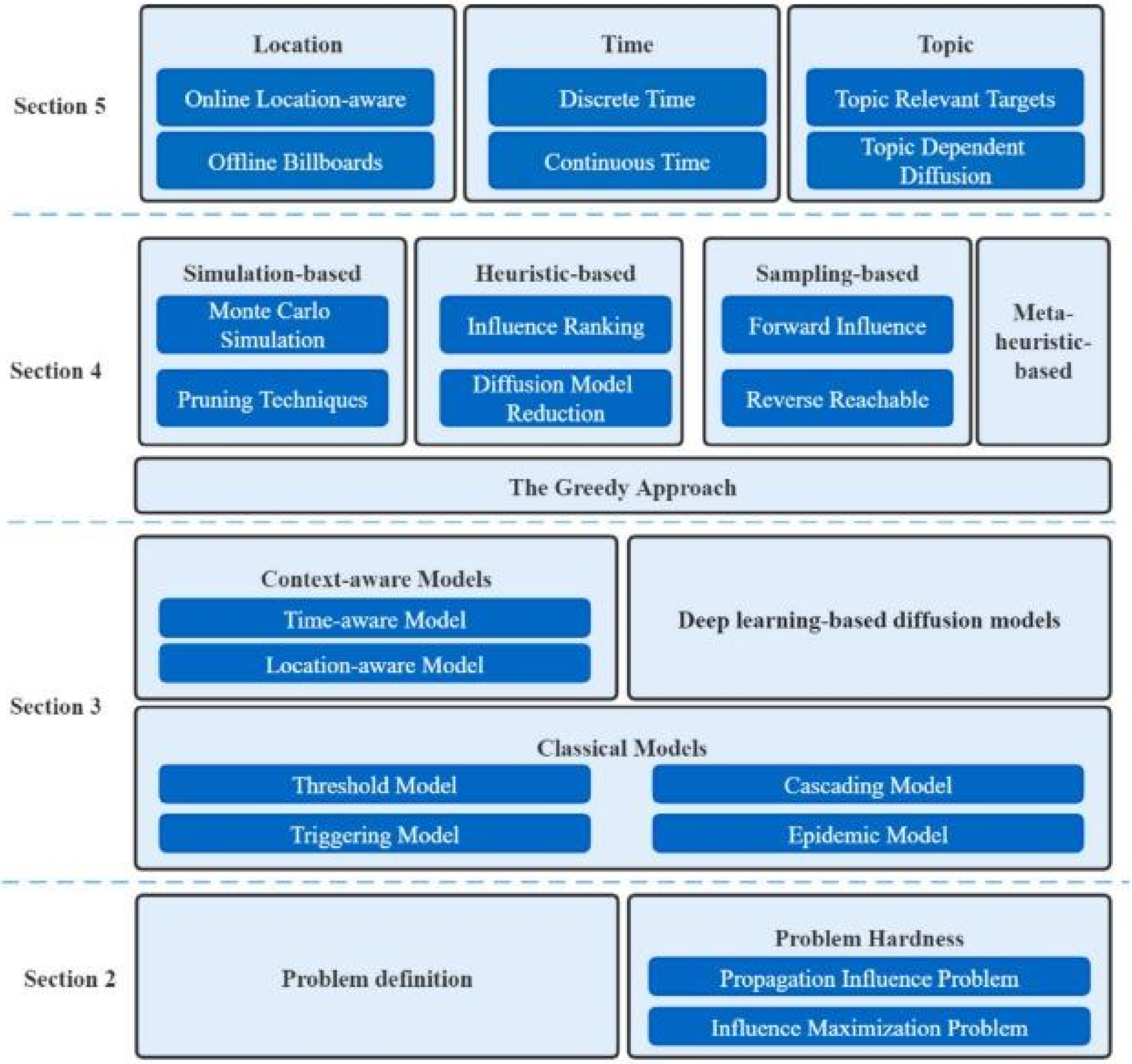

Fig 6.1 Block Diagram for methodology.

7. Proposed System / Methodology

Step 1: Define the problem and network structure: Model the social network as a directed graph $G = (V, E)$, where nodes represent users and edges represent influence paths.

Step 2: Select or simulate diffusion models: Choose appropriate diffusion models such as Independent Cascade (ICM), Linear Threshold (LTM), or Triggering Model to simulate how influence propagates.

Step 3: Identify seed nodes using IM algorithms: Apply Influence Maximization algorithms—categorized into simulation-based, heuristic-based, sampling-based, and meta-heuristic-based approaches—to select a seed set $S \subseteq V$.

Step 4: Evaluate influence spread for each algorithm: Estimate the influence spread $\sigma(S)$ of the selected seed set using Monte Carlo simulations or Reverse Reachable Sets (RRS), considering computational complexity and accuracy.

Step 5: Incorporate context-aware factors: Enhance realism and precision by including context-aware parameters like user location, time of dissemination, and content/topic relevance in the propagation models.

Step 6: Compare and analyze algorithm performance: Use metrics such as spread size, time complexity, and approximation ratio to evaluate the effectiveness of each algorithm. Visualize findings using comparative tables or influence graphs.

8. Expected Outcomes

- A formal definition and structured understanding of the Influence Maximization (IM) problem in social networks.
- A comprehensive comparison of diffusion models including classical, context-aware, and deep learning-based approaches.
- Categorized implementation and analysis of IM algorithms (simulation-based, heuristic-based, sampling-based, and meta-heuristic-based).
- Evaluation results demonstrating influence spread and performance efficiency of various algorithms across different models.
- A scalable and reproducible framework that can be adapted to real-world social networks and extended to dynamic, time-sensitive, or topic-specific contexts.

9. References

[1]. Ye, Y., Chen, Y., & Han, W. (2022). Influence Maximization in Social Networks: Theories, Methods and Challenges. Array, 16, 100264.

[2]. Han, W., & Ye, Y., Chen, Y., (2022). Influence Maximization : Theories, Methods and Challenges. Array, 16, 100264.

[3]. Li, C., Ma, J., & Wang, Y. (2017). DeepCas: An End-to-End Predictor of Information Cascades. Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI), 1587–1593.

[4]. Tang, Y., Shi, Y., & Xiao, X. (2015). Influence Maximization in Near-Linear Time: A Martingale Approach. Proceedings of the 2015 ACM SIGMOD International Conference, 1539–1554.

[5]. Tang, Y., Xiao, X., & Shi, Y. (2014). Influence Maximization: Near-Optimal Time Complexity Meets Practical Efficiency. Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data, 75–86.

[6]. Goyal, A., Lu, W., & Lakshmanan, L. V. S. (2011). CELF++: Optimizing the Greedy Algorithm for Influence Maximization in Social Networks. Proceedings of the 20th International Conference Companion on World Wide Web, 47–48.

[7]. Chen, W., Wang, Y., & Yang, S. (2009). Efficient Influence Maximization in Social Networks. Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 199–208.

[8]. Leskovec, J., Krause, A., Guestrin, C., Faloutsos, C., VanBriesen, J., & Glance, N. (2007). Cost-Effective Outbreak Detection in Networks. Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 420–429.

[9]. Kempe, D., Kleinberg, J., & Tardos, É. (2003). Maximizing the Spread of Influence through a Social Network. Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 137–146.

[10]. Tardos, É. & Kempe, D., Kleinberg, J.,(2003). Maximizing the Spread of Influence through a Social Network. Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 137–146.