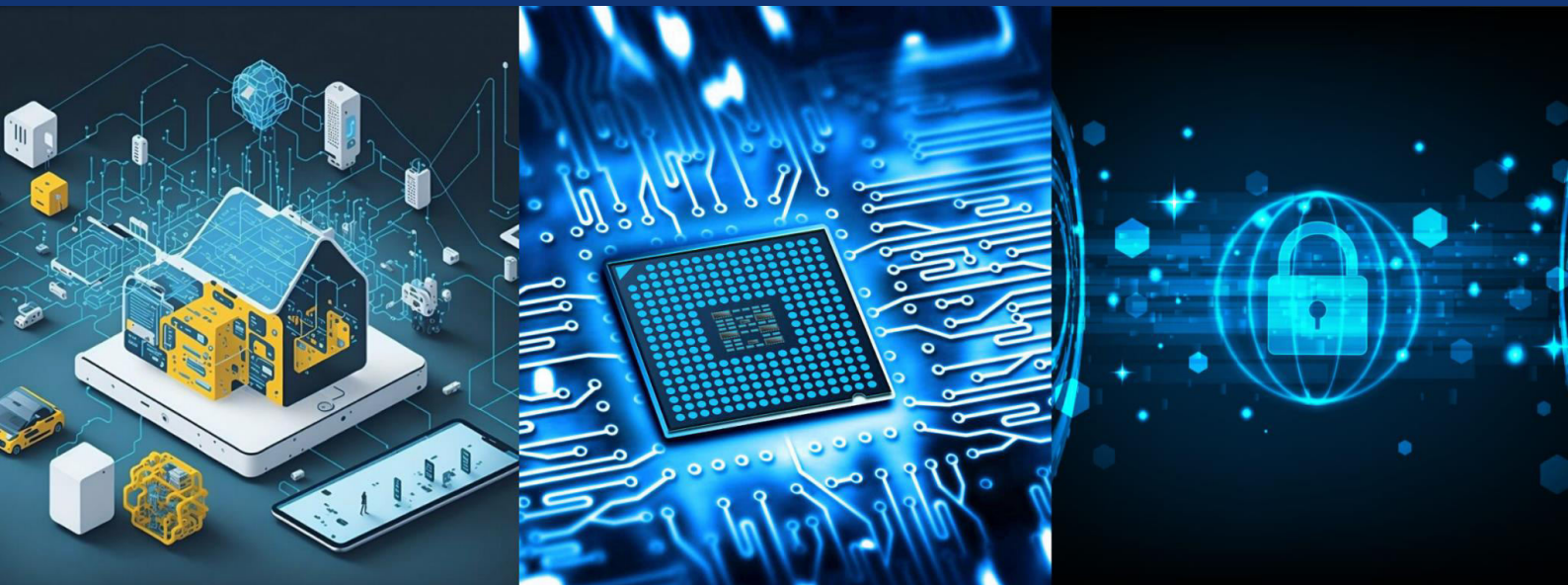
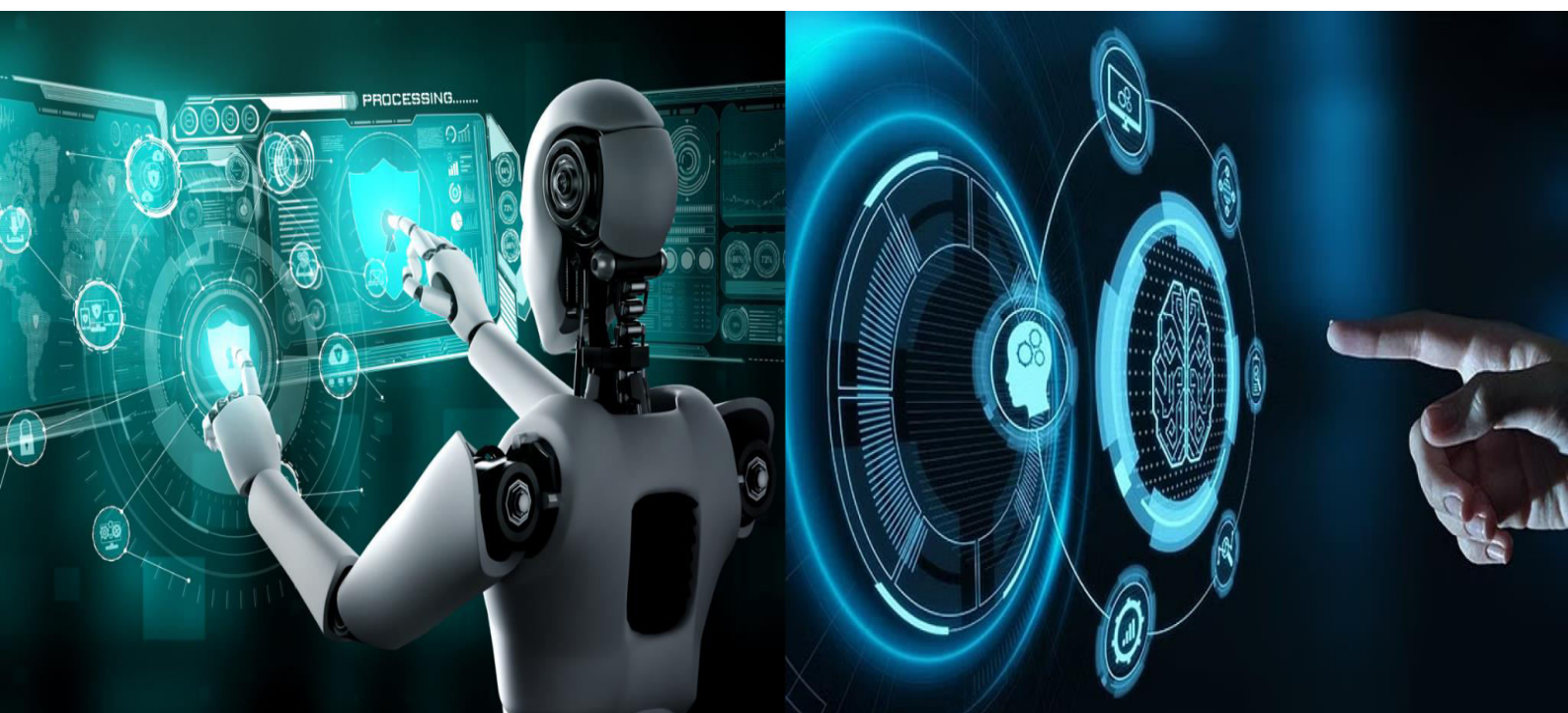




International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





A Learning-Based Adaptive Hybrid Framework for Dynamic Seed Selection in Social Network Influence Maximization

Nidhi Gupta, Dr. Sohit Agarwal

Research Scholar, Department of Computer Engineering and Information Technology, Suresh Gyan Vihar University, Jaipur, Rajasthan, India

Principal, ASTER School of Computer Science & Professor, Department of Computer Engineering and Information Technology, Suresh Gyan Vihar University, Jaipur, Rajasthan, India

ABSTRACT: Influence maximization looks for a select group of people who can start the spread of knowledge widely. Complementary influencer signals can be combined using fixed hybrid ranking, however it is predicated on feature weights being consistent across networks and diffusion conditions. In this paper, a learning-based adaptive hybrid framework for diversity-aware sequential seed selection is proposed, which learns seven feature weights from downstream diffusion feedback. Feature weighting is linked to a diffusion-and-overlap objective under Independent Cascade and Linear Threshold models through a simplex-constrained derivative-free search. Repeatable datasets, baselines, metrics, ablations, and complexity analysis are defined in the paper; numerical performance is purposefully withheld until the experiments are carried out.

KEYWORDS: Influence Maximization; Adaptive Weight Learning; Diversity-Aware Seed Selection; Independent Cascade; Social Network Analysis

I. INTRODUCTION

Through user-to-user interactions, social networks offer a graph-based environment where information can spread. The practical question is not whether all users should be activated in applications like product diffusion, public information dissemination, recommendations, and online campaigns, but rather which small group of users should be chosen first so that subsequent diffusion reaches a significant portion of the network. Influence maximization (IM), which selects a seed set of bounded size to maximize expected influence spread, is a typical formulation of this optimization issue [1], [7].

Propagation is modeled by probabilistic diffusion processes in the standard IM formulation, particularly in the Independent Cascade (IC) and Linear Threshold (LT) models. Because predicted spread has advantageous submodularity qualities under these models, greedy seed selection is theoretically well motivated. However, direct evaluation of marginal gains may necessitate repeated diffusion simulations. Through sluggish evaluation and more effective spread estimate, scalable techniques like CELF and TIM lessen the computing load [2], [3]. Later techniques like IMM further enhance the practical scalability of influence estimation [4].

By employing structural indicators like out-degree, in-degree, PageRank, or betweenness, lightweight ranking techniques prevent recurrent diffusion estimation. The degree to which a node is on shortest pathways is captured by betweenness [6], whereas PageRank quantifies global link-structure prominence [5]. Although these indications are computationally appealing, they only capture one aspect of a user's possible effect. Additionally, behavior-aware research demonstrates that message qualities, relationship attributes, and user behavior can significantly impact influence processes [7].

In order to overcome this restriction, the previous methodology provided with this study combined seven signals—out-degree, out-strength, PageRank, in-degree, reach, engagement, and betweenness—into an interpretable fixed hybrid



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

score before implementing diversity-aware selection. The weights that were recorded were 0.40, 0.25, 0.15, 0.08, 0.05, 0.04, and 0.03. Instead of using those values as universal parameters, the current approach considers them as a baseline. Whether feature weights may be directly connected to downstream diffusion results and modified through feedback is the methodological question.

As a result, this research suggests an adaptable hybrid framework based on learning. The term "dynamic" does not imply that the graph topology itself changes over time; rather, it refers particularly to the adaptation of feature weights and seed-selection utility across feedback rounds. Temporal snapshots or time-indexed edges, activities, or propagation parameters would be necessary for a truly time-evolving network, but these are outside the scope of the current experimental setup. Studies on adaptive seeding in dynamic diffusion environments have been conducted independently [8].

The proposed study makes the following contributions: (i) an interpretable seven-feature adaptive scoring model with nonnegative simplex-constrained weights; (ii) a diffusion-linked derivative-free learning mechanism that optimizes the downstream objective instead of predicting a manually fixed score; (iii) a sequential seed-selection rule that penalizes two-hop neighborhood overlap; (iv) a reproducible evaluation protocol under IC and LT; and (v) a multi-dataset experimental design that distinguishes interaction-aware settings from structural-only data. Since the suggested experiments have not yet been carried out, no numerical performance claim is provided here.

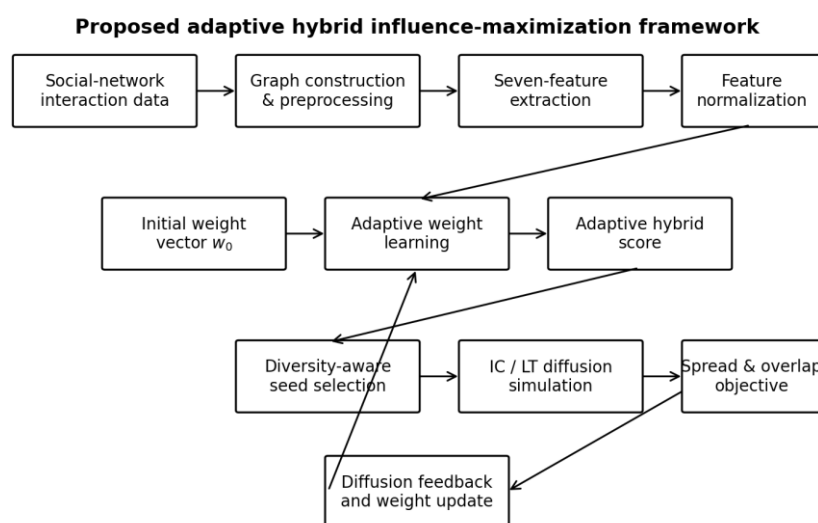


Fig. 1. Proposed adaptive hybrid framework. The feedback loop updates feature weights from diffusion outcomes; "dynamic" refers to this adaptation rather than an assumed time-evolving graph.

II. LITERATURE SURVEY

Kempe, Kleinberg, and Tardos established the importance of greedy maximization under suitable sub modularity constraints by analyzing the IC and LT diffusion models and formulating classical influence maximization as a discrete optimization issue [1]. The conceptual basis for assessing candidate seed sets based on predicted spread is still this formulation.

The goal of further studies was to lower the expense of repeated spread estimation. A low-cost lazy-forward approach that takes advantage of declining returns to prevent needless recomputation was presented by Leskovec et al. [2]. Tang et al. created TIM, a scalable influence-maximization technique that offers near-linear expected-time guaranties under the IC model and is based on reverse-reachable-style estimate [3].



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IMM enhanced martingale-based estimation to attain significant efficiency while preserving approximation guaranties [4]. These methods increase influence spread more directly than simple centrality rankings, but when analyzing a large number of potential weight combinations, their computational cost is still substantial. Another lightweight strategy is provided by centrality-based methods. PageRank describes global structural significance using a random-surfer formulation [5], whereas betweenness finds nodes that bridge shortest pathways [6]. Degree-based methods are simpler and often function as baselines.

However, engagement, interaction intensity, and other behavioral data may not always be captured by structural rankings. Zareie and Sakellariou's behavior-aware survey highlights the extra information available beyond topology by classifying approaches that take user behavior, relationship attributes, and message characteristics into account [7].

A different approach is offered by adaptive influence maximization. In order to enable seed choices to react to observed diffusion states, Tong et al. defined adaptive seeding under a Dynamic Independent Cascade model [8]. While the current system focuses on learning the relative weights of interpretable influencer traits and then applying those weights in diversity-aware seed selection, that line of work covers adaptive judgments during shifting diffusion. The distinction is crucial because in this case, adaptation takes place in the scoring parameters rather than by asserting that the graph is always changing over time.

Influencer prediction using machine learning has also been studied. Gradient-boosted decision trees were employed by Subramani et al. to forecast influencers [9]. Although this model shows the value of learning from influencer-related data, learning feature weights against the downstream diffusion aim is conceptually different from predicting an existing target score. Therefore, rather than using Random Forest or Gradient Boosting as a stand-in for a manually defined score, the proposed approach assesses candidate weight vectors based on the influence they produce following seed selection.

Further proof that derivative-free optimization can be helpful for the computationally challenging IM issue comes from meta-heuristic investigations. For influential-node selection, Tang et al. suggested an enhanced discrete particle-swarm optimization strategy [10], while current reviews use a variety of heuristic and meta-heuristic methods [11]. The current approach has a different optimization goal: each candidate weight vector is assessed using the downstream diffusion-and-diversity objective, and the decision variables are interpretable feature weights instead of the seed nodes themselves.

III. METHODOLOGY

3.1 Overview of the Framework

Network building, feature extraction, normalization, adaptive weight learning, adaptive hybrid scoring, diversity-aware seed selection, diffusion evaluation, and weight updating comprise the feedback-driven pipeline of the suggested system. Instead of replicating the fixed score employed by the prior framework, the key design choice is to optimize the seven feature weights with respect to the downstream diffusion aim. Fig. 1 depicts the entire flow.

3.2 Data Preparation and Network Representation

A directed graph $G=(V,E)$, where V is the set of users, E is the set of directed relationships, $n=|V|$, and $m=|E|$, is used to depict the social network. An edge weight a_{uv} may indicate the observed interaction strength between users u and v when interaction data is available. Self-loops are eliminated unless mandated by a dataset design, duplicate records are combined, and isolated nodes may be left out of diffusion studies if they are unable to contribute to propagation. Every technique is assessed using the same preprocessed graphs.

It is suggested to use three publicly available datasets. There are 508,837 directed trust edges and 75,879 nodes in SNAP's Epinions network [13]. There are 4,039 nodes and 88,234 total undirected edges in the Facebook ego-network dataset [14]. There are multiple directed interaction layers in the Higgs Twitter collection; the retweet, respond, and mention layers offer weighted interaction records, while the largest social layer has 456,631 nodes and 14,855,875 edges [15]. Because its activity data includes timestamps and interaction kinds, the Higgs collection is especially well-suited for the interaction-aware version.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Missing behavioral variables shouldn't be created by the seven-feature framework. The study should provide a six-feature structural variant or a separately defined proxy for structural-only datasets where engagement is noted as unavailable. While out-strength can employ edge weights, engagement on the Higgs interaction networks can be operationalized using observed interaction counts over a specified observation window. A structural statistic cannot be presented as witnessed human participation because of this separation.

3.3 Influencer Feature Extraction

Feature	Definition	Interpretation
f1: Out-Degree	Number of outgoing edges	Direct forwarding opportunities
f2: Out-Strength	Sum of outgoing edge weights	Weighted transmission opportunity
f3: PageRank	Random-walk centrality	Global structural prominence
f4: In-Degree	Number of incoming edges	Audience/popularity proxy
f5: Reach	Size of two-hop outgoing neighborhood	Potential local exposure
f6: Engagement	Observed interaction activity or explicit proxy	Behavioural response indicator
f7: Betweenness	Fraction of shortest paths through node	Bridge/community role

Table 2. Influencer features used by the proposed adaptive hybrid framework.

The raw feature vector for each candidate node v is $F(v)=[f1(v),...,f7(v)]^T$. Each accessible feature is normalized using min-max normalization with a small $\epsilon>0$ to prevent division by zero because the features have distinct scales:

$$\tilde{f}_i(v) = \frac{f_i(v) - \min_{u \in V} f_i(u)}{\max_{u \in V} f_i(u) - \min_{u \in V} f_i(u) + \epsilon}$$

For a partially constructed seed set S , the candidate-to-set overlap is $O(v,S)=0$ when S is empty; otherwise it is the maximum Jaccard overlap with any selected seed. The sequential utility is:

$$U(v | S, \mathbf{w}) = AH(v; \mathbf{w}) - \lambda_d O(v, S), \quad O(v, S) = \max_{s \in S} J(v, s)$$

Here $\lambda_d \geq 0$ controls the trade-off between the learned influencer score and local redundancy. The candidate with maximum utility is selected and the process continues until $|S|=B$. There is no guaranty of sub modularity or approximation for the combined utility; this is an explicit heuristic selection rule.

3.6 Diffusion Models and Learning Objective

Diffusion have two models. With an assumed propagation probability of p_{uv} , each newly active node in IC has one independent chance to activate every inactive outgoing neighbor. Each node in LT has a threshold, and it becomes active when the overall impact from active inbound neighbors exceeds that level. These values are experimental assumptions, and unless p_{uv} or LT weights are calculated from past diffusion logs, simulated cascades cannot be compared to observed human behavior [1].

For a weight vector w , the diversity-aware procedure produces a seed set S_w . Let $\sigma_M(S_w)$ be the mean number of activated nodes over R Monte Carlo runs under model M . Let the selected-seed redundancy be the mean pairwise two-hop Jaccard overlap:

$$R(S) = \frac{2}{B(B-1)} \sum_{\{u,v\} \subseteq S} J(u,v) \quad (B > 1)$$



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Normalized diffusion coverage under IC and LT is combined with a redundancy penalty in the learning objective:

$$Q(\mathbf{w}) = \frac{1}{|\mathcal{M}|} \sum_{M \in \mathcal{M}} \frac{\bar{\sigma}_M(S_{\mathbf{w}})}{n} - \lambda_r R(S_{\mathbf{w}})$$

Here $M=\{IC,LT\}$, $\lambda_r \geq 0$ is a learning-objective penalty, and $R(S_{\mathbf{w}})$ is the redundancy term. Weight learning and the downstream goal are intimately linked in this formulation: a weight vector is kept only if the resulting seed set offers a favorable trade-off between redundancy and diffusion coverage. **3.7 Feedback and Weight Updating**

Feedback-driven adaptation loop

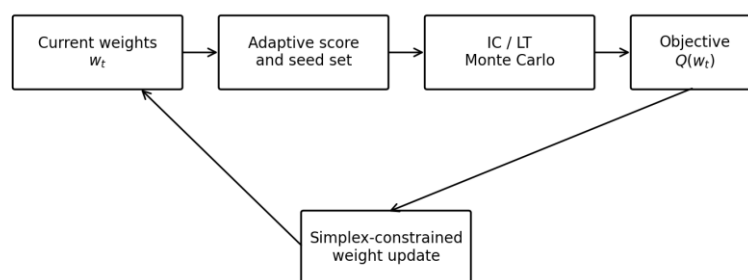


Fig. 2. Feedback-driven weight-learning cycle. Diffusion outcomes determine whether a candidate weight vector is retained and how the search step is updated.

The adaptation cycle begins with w_t , constructs the seed set S_{w_t} , estimates diffusion spread, computes $Q(w_t)$, and evaluates feasible neighbouring vectors. If an improved candidate is found, the vector becomes w_{t+1} ; otherwise the search step is reduced. The process stops when the maximum number of iterations T is reached, the step size falls below $\delta_{\min}$, or the objective improvement remains below ϵ for a predefined number of iterations.

3.8 Proposed Algorithm

Algorithm 1: Adaptive Hybrid Diversity-Aware Seed Selection

Input: Directed graph $G(V,E)$; candidate set C ; seed budget B ; feature matrix F ; initial weights w_0 ; diversity penalty λ_d ; learning penalty λ_r ; diffusion models $M=\{IC,LT\}$; simulations R ; maximum iterations T .

1. G should be preprocessed in order to obtain the seven influencer features.
2. Normalize each feature column and initialize $w \leftarrow w_0$.
3. For $t=1, \dots, T$, compute $AH(v;w)$ for all candidates.
4. Set $S \leftarrow \emptyset$. While $|S| < B$, compute $O(v,S)$ and $U(v|S,w)$ for every $v \in C \setminus S$; add the candidate with maximum utility.
5. Run R IC simulations and R LT simulations using the same seed set and prescribed diffusion parameters.
6. Estimate mean spread, normalized coverage, and redundancy; compute $Q(w)$.
7. Generate viable simplex neighbors by transferring step δ between weight coordinates.
8. Accept the best improving vector after evaluating nearby weight vectors using Steps 3–6.
9. If no improvement occurs, reduce δ ; stop when $\delta < \delta_{\min}$, $\text{improvement} < \epsilon$, or $t=T$.
10. Return the best observed seed set S^* and weight vector w^* , together with spread, coverage, runtime, overlap, and diversity measurements

3.9 Computational Complexity

Stage	Order / Cost	Notes
Graph construction	$O(n+m)$	Depends on input representation and preprocessing
Degree / strength	$O(n+m)$	Single graph pass
PageRank	$O(I \cdot PR(n+m))$	$I \cdot PR$ = PageRank iterations
Exact betweenness	$O(nm)$ unweighted	Brandes-style computation [6]



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Two-hop neighbourhoods	$O(\sum_v N_2^+(v))$	Can approach $O(n^2)$ in dense graphs
One seed-selection round	$O(Bn)$ after overlap precomputation	Naive candidate scan
Monte Carlo diffusion	$O(R \cdot B \cdot (n+m))$ worst-case order	Depends on cascade size and implementation
Weight search	$O(T \cdot K \cdot C_{diff})$	K = evaluated weight candidates per iteration; C_{diff} = diffusion-evaluation cost

Table 3. Computational complexity by primary stage of processing

The main element is usually repeated diffusion assessment, particularly when the weight search looks at a lot of potential vectors. On big graphs, exact betweenness can also become costly; consequently, approximate betweenness is a feasible scalability alternative as long as the same approximation approach is applied to all compared methods. Although the derivative-free search controls the investigation of weight possibilities, diffusion simulation costs are still incurred.

3.10 Protocol for Reproducibility

When paired comparisons are appropriate, all approaches must employ the same preprocessing, candidate sets, seed budgets, diffusion assumptions, simulation counts, and random-number seeds. A machine-readable description of the node count, edge count, directionality, weight semantics, feature availability, and preprocessing choices should be included with every dataset. Initial and final weight vectors, objective trajectory, random and chosen seeds, runtime, and per-model diffusion statistics should all be documented by the implementation. Instead of reporting a single Monte Carlo estimate, experiments should be repeated across independent random batches and provided with mean and variability.

IV. RESULTS & DISCUSSION

Although the suggested experiments have not been carried out in the current study package, the methodology and evaluation protocol were established using the provided research resources. As a result, no results are reported for influence-spread, coverage, runtime, significance, or improvement. Submission-time result templates are shown in the tables below. After implementation and several experiments, they have to be replaced with measured values.

4.1 Recommended Datasets and Preprocessing

Dataset	Nodes	Edges	Type	Edge meaning	Recommended use	Caution
Epinions [13]	75,879	508,837	Directed	Trust relationships	Structural influence; scalability	No validated engagement field; use structural-only variant
Facebook ego [14]	4,039	88,234	Undirected combined graph	Friendship circles	Small/medium structural benchmark	Direction must not be invented; document any directed conversion
Higgs Twitter [15]	456,631	14,855,875*	Directed	Follower + interaction layers	Large, interaction-aware experiment	Use appropriate layer and reverse retweet direction only if justified

Table 4. Recommended public datasets and preprocessing considerations.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

*The 14,855,875 figure is the largest social-layer count reported by SNAP; the Higgs collection also provides smaller retweet, reply, and mention networks. The exact layer used in an experiment must be stated because the networks have different semantics [15].

4.2 Baselines

Method	Selection rule	Purpose
Random	Uniformly random seed set; repeated trials	Reference for non-structured selection
Out-Degree	Highest normalized out-degree	Direct-connectivity baseline
PageRank	Highest PageRank	Global-centrality baseline
Fixed Hybrid	Previous seven-feature weights	Continuity baseline
Adaptive Hybrid, no diversity	Learned weights; top-score selection	Tests weight learning alone
Adaptive Hybrid, diversity-aware	Learned weights + overlap penalty	Proposed full framework
Greedy / CELF (if feasible)	Diffusion-oriented selection	Reference for direct spread optimization

Table 5. Baseline methods and their experimental purpose.

4.3 Performance Metrics

Metric	Definition	Operationalization
Influence Spread	Mean activated nodes over R runs	$\hat{\sigma}_M(S) = 1/R \sum_r X_r$
Diffusion Coverage	Activated nodes divided by n	$\hat{\sigma}_M(S)/n$
Runtime	Wall-clock scoring + selection + evaluation time	Seconds under fixed hardware/software
Seed Overlap	Mean pairwise overlap of selected seeds	$R(S)$ as defined in Section 3.6
Seed Diversity	One minus normalized redundancy, or a clearly stated community measure	Must be fixed before testing
Scalability	Change in runtime/quality with n and m	Compare across network sizes
Weight Stability	Variation of learned weights across runs	Mean SD or coefficient of variation
Marginal Gain	Change in spread after adding a seed	$\sigma(SU\{v\}) - \sigma(S)$, where estimated
Convergence	Objective improvement or search-step reduction	Stop using predefined ε and $\delta_{\min}$

Table 6. Performance metrics and operational definitions.

4.4 Experimental Result Template

Method	IC Spread	LT Spread	Avg. Coverage	Runtime (s)	Overlap
Random	To be measured	To be measured	To be measured	To be measured	To be measured
Out-Degree	To be measured	To be measured	To be measured	To be measured	To be measured
PageRank	To be measured	To be measured	To be measured	To be measured	To be measured
Fixed Hybrid	To be measured	To be measured	To be measured	To be measured	To be measured
Adaptive Hybrid, no diversity	To be measured	To be measured	To be measured	To be measured	To be measured
Adaptive Hybrid, diversity-aware	To be measured	To be measured	To be measured	To be measured	To be measured
Greedy/CELF	To be measured	To be measured	To be measured	To be measured	To be measured

Table 7. Experimental result template. Replace every “To be measured” entry only after the corresponding experiment is executed.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4.5 Weight Evolution Template

Iter.	Out-Deg.	Out-Str.	PageRank	In-Deg.	Reach	Engage.	Between.	Objective
0	0.40	0.25	0.15	0.08	0.05	0.04	0.03	Baseline initialization
1	To be measured	To be measured	To be measured	To be measured	To be measured	To be measured	To be measured	Measured
...	To be measured	To be measured	To be measured	To be measured	To be measured	To be measured	To be measured	Measured
Final	To be measured	To be measured	To be measured	To be measured	To be measured	To be measured	To be measured	Measured

Table 8. Template for learned feature-weight evolution.

The documented values from the prior framework serve as the beginning weights. A large learned weight does not necessarily indicate universal causal importance; rather, learned weights should be regarded as dependent on the dataset, feature creation, diffusion assumptions, seed budget, and optimization process.

4.6 Ablation Study

Variant	Removed/modified component	IC Spread	LT Spread	Overlap	Purpose
Full model	None	To be measured	To be measured	To be measured	Reference
A1	No Out-Degree	To be measured	To be measured	To be measured	Direct-connectivity contribution
A2	No Out-Strength	To be measured	To be measured	To be measured	Weighted interaction contribution
A3	No PageRank	To be measured	To be measured	To be measured	Global-structure contribution
A4	No In-Degree	To be measured	To be measured	To be measured	Popularity contribution
A5	No Reach	To be measured	To be measured	To be measured	Exposure contribution
A6	No Engagement	To be measured	To be measured	To be measured	Behavioural contribution
A7	No Betweenness	To be measured	To be measured	To be measured	Bridge-position contribution
A8	Fixed weights	To be measured	To be measured	To be measured	Adaptive-learning contribution
A9	No diversity penalty	To be measured	To be measured	To be measured	Diversity mechanism contribution

Table 9. Ablation-study design.

4.7 Discussion and Expected Interpretation

Whether learning feature weights against downstream diffusion alters seed sets and objective values in comparison to the fixed hybrid baseline is the main empirical question. One mean value should not be used to indicate success in a valid analysis. Paired simulation variability, objective convergence, weight trajectories, overlap, runtime, and behavior across seed budgets and datasets should all be examined. If, in a controlled comparison, the adaptive approach modifies



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

weights but does not enhance diffusion coverage or redundancy, the result is still instructive since it would show that, in the studied conditions, the additional learning stage does not justify its computational cost.

Adaptive weighting and the diversity component should be regarded independently. It is possible to determine if sequential diversification alters seed redundancy without confusing it with weight learning by comparing the adaptive approach with and without the overlap penalty. Likewise, it is not appropriate to combine the structural-only and interaction-aware datasets as if their feature semantics were the same. While the Facebook and Epinions graphs mostly enable structural evaluation, the Higgs interaction layers can allow interaction-based features [13]–[15].

The same diffusion random seeds should be used for paired comparisons in statistical analysis. The mean spread should be accompanied with standard deviations or confidence intervals. Declaring a difference from raw means alone should not be used when comparing numerous techniques at once; instead, a suitable multiple-comparison approach should be employed. Before the experiments are carried out, no such statistical conclusion can be drawn.

V. CONCLUSION

In this paper, a fixed seven-feature influencer-ranking method is extended with a learning-based adaptive hybrid framework for influence maximization. The suggested approach uses a downstream diffusion-and-diversity objective to acquire nonnegative, normalized feature weights, which are then used for sequential overlap-aware seed selection. The resulting seed sets are then assessed under IC and LT diffusion. Since the goal is discrete and Monte Carlo based, a derivative-free simplex search is used. Additionally, the study establishes reproducible baselines, metrics, ablations, complexity analysis, and dataset-specific feature availability. The publication purposefully offers no numerical performance conclusion because the trials have not yet been carried out. Until it is empirically validated, the framework should be regarded as a methodological research design.

REFERENCES

- [1] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in Proc. 9th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2003, pp. 137–146, doi: 10.1145/956750.956769.
- [2] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, "Cost-effective outbreak detection in networks," in Proc. 13th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2007, pp. 420–429, doi: 10.1145/1281192.1281239.
- [3] Y. Tang, X. Xiao, and Y. Shi, "Influence maximization: Near-optimal time complexity meets practical efficiency," in Proc. ACM SIGMOD Int. Conf. Management of Data, 2014, pp. 75–86, doi: 10.1145/2588555.2593670.
- [4] Y. Tang, Y. Shi, and X. Xiao, "Influence maximization in near-linear time: A martingale approach," in Proc. ACM SIGMOD Int. Conf. Management of Data, 2015, pp. 1539–1554, doi: 10.1145/2723372.2723734.
- [5] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: Bringing order to the Web," Stanford Digital Library Technologies Project, Technical Report, 1998.
- [6] U. Brandes, "A faster algorithm for betweenness centrality," Journal of Mathematical Sociology, vol. 25, no. 2, pp. 163–177, 2001, doi: 10.1080/0022250X.2001.9990249.
- [7] A. Zareie and R. Sakellariou, "Influence maximization in social networks: A survey of behaviour-aware methods," Social Network Analysis and Mining, vol. 13, art. 78, 2023, doi: 10.1007/s13278-023-01078-9.
- [8] G. Tong, W. Wu, S. Tang, and D.-Z. Du, "Adaptive influence maximization in dynamic social networks," IEEE/ACM Transactions on Networking, vol. 25, no. 1, pp. 112–125, 2017, doi: 10.1109/TNET.2016.2563397.
- [9] N. Subramani, S. Veerappampalayam Easwaramoorthy, P. Mohan, M. Subramanian, and V. Sambath, "A gradient boosted decision tree-based influencer prediction in social network analysis," Big Data and Cognitive Computing, vol. 7, no. 1, art. 6, 2023, doi: 10.3390/bdcc7010006.
- [10] J. Tang, H. Zhu, J. Lan, S. Song, J. Qu, and Q. Du, "Identifying influential nodes for influence maximization problem in social networks using an improved discrete particle swarm optimization," Social Network Analysis and Mining, vol. 13, art. 94, 2023, doi: 10.1007/s13278-023-01098-5.
- [11] C. Fan, Z. Wang, J. Zhang, J. Zhao, and X. Meng, "Meta-heuristic algorithms for influence maximization: A survey," Evolving Systems, vol. 16, art. 24, 2025, doi: 10.1007/s12530-024-09640-2.
- [12] N. Gupta and S. Agarwal, "A Scalable Hybrid Framework for Continuous Influencer Intelligence," unpublished/working manuscript supplied as the methodological basis for the present extension, 2026.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- [13] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford Large Network Dataset Collection," Stanford Network Analysis Project, dataset documentation for soc-Epinions1, accessed 2026.
- [14] J. McAuley and J. Leskovec, "Learning to discover social circles in ego networks," in Advances in Neural Information Processing Systems 25, 2012, pp. 548–556.
- [15] M. De Domenico, A. Lima, P. Mougél, and M. Musolesi, "The anatomy of a scientific rumor," Scientific Reports, vol. 3, art. 2980, 2013, doi: 10.1038/srep02980.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details