Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 5, pp. 2016–2026, DOI 10.12694/scpe.v26i5.4851

RESEARCH ON THE PATH OF EMPOWERING RURAL REVITALIZATION WITH E-COMMERCE UNDER THE BACKGROUND OF DIGITAL ECONOMY

ZHENYA ZHAO* AND XIAOLAN FENG †

Abstract. Effective utilization of marketing data can enhance market mobility and competitiveness, thus increasing the number of users. In order to help enterprises proactively deal with competition in the Shaanxi Province market, this study validates and products a game theory-assisted machine learning scheme based on the theory of game decision-making and conditional relaxation. The technique helps enterprises identify users' churn tendencies and adopt differentiated handling strategies through machine learning models. Specifically, it enables enterprises to formulate targeted migration strategies and accurately identify "abnormal" users that may flow in or out. According to the experimental results, the program has been successfully transformed into a product, which greatly improves the marketing effect of enterprises. For example, in one company, the number of users increased by about 50%. In addition, this study analyzes the interest preferences of moviegoers on several dimensions by comparing the seed sets of the Movie Consumer Network and the Yahoo Answer Network.

Key words: Increase in users; Game theory; Machine learning; Marketing data; Comparative analysis

1. Introduction. With the development of Internet technology, the marketing medium has become more diverse, encompassing the emerging Internet, social networks, and mobile applications in addition to the more traditional media of radio, television, newspapers, and magazines. The new communication media also bring new marketing ideas and a range of innovative marketing techniques[1]. Two of these characteristic marketing techniques are particularly highlighted in this essay: viral marketing, which is centered on social networking sites, and online marketing, which connects well-known companies with the Internet.

Viral marketing leverages existing social networking platforms or other technology to distribute content in a way similar to a virus replicating and spreading itself in order to promote a brand or other marketing objectives [2]. Viral marketing has garnered significant interest from both industry and academia due to its often low cost and ability to disseminate information quickly and widely. Mathematically, viral marketing can be expressed as an influence maximization problem, or more precisely, how to select a subset of nodes of size K in a social network so that its influence is maximized (activating the most nodes in the network) [3]. In the real world, we might seek to influence as many nodes as we can, as well as to be influenced by as many different nodes as we can, in order to increase the likelihood of reaching the target audience and reduce marketing risk.

Meituan and Dianping in the catering and entertainment industry, Didi Express in the taxi industry, Ctrip and Qunar in the travel and tourism industry, and Soufun and chain home real estate in the real estate industry are just a few examples of traditional industries that have been seriously affected by the Internet in recent years [4]. When it comes to consumer consumption and review data, trajectory data or real estate market data, the Internet process makes it easier to collect and obtain data from traditional industries [5]. We can improve user experience and make marketing more intelligent and targeted by mining this data.

We acquired transaction data, such as dates, costs, specifics of property qualities, and pertinent geographic location information, for a commercial real estate company in Shaanxi Province. We made an effort to estimate the time needed for the house's sale using the information supplied [6]. By helping sellers set fair asking prices for properties that sell at the right time and helping buyers ascertain the property's popularity and the most effective bidding strategies, these pieces of information can assist both buyers and sellers in making decisions.

*School of Economics and Management, Gansu Vocational College of Communications,Lanzhou,Gansu, 730207, China; School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou, Gansu, 730070, China (Corresponding author, jilly_ya@163.com)
†School of Information Engineering, Xi'an Translation Institute, Xi'an, Shaanxi, 710105, China (15389050978@163.com).

2. Game-theoretic based decision algorithms.

2.1. Game theoretic algorithms. Game theory has been effectively utilized in various domains, such as analyzing transaction selection, pricing competition in sales, and fraud penalties [7]. Building on this success, we propose applying a game theory approach to strategically evaluate the offensive and defensive marketing behaviors of mobile network operators (MNOs). This framework is designed to optimize the marketing decisions made by enterprises operating within the mobile network operations mode, considering the complex interplay between competitors and market conditions. In this context, MNOs must make critical marketing choices, including pricing strategies, customer acquisition efforts, and retention tactics, all while anticipating the reactions of competing companies. By modeling these decisions through game theory, operators can predict the outcomes of different strategies and select the optimal course of action. Specifically, the offensive strategies focus on gaining market share and attracting new customers, while defensive strategies aim at retaining existing users and minimizing churn. This game-theoretic framework allows MNOs to account for both short-term and long-term market dynamics. For example, it helps businesses evaluate the effectiveness of price cuts, promotional offers, and loyalty programs while anticipating potential counteractions from competitors. By understanding the strategic responses of other players in the market, MNOs can better position themselves to maximize profitability and reduce the risk of adverse outcomes, such as customer defections or revenue loss. Moreover, this approach offers the flexibility to incorporate various market variables, such as customer preferences, competitor behaviors, and external factors like regulatory changes. Ultimately, the proposed game theory-based marketing strategy enables MNOs to navigate the competitive landscape with a more informed and proactive approach, leading to better decision-making and improved financial performance. Fig.2.1 depicts the workflow of the suggested game theory-based offensive and defensive marketing strategy approach, which entails feature extraction of the most important business data, thorough scoring of the worth of the company, a market game with rivals, and the production of the optimal decision strategy. For Chinese mobile network operators, a decision algorithm based on game theory has been created to maximize the effectiveness of mobile network operations [8]. The algorithm dynamically outputs the best current marketing decision, including maximum carry-in ($\max_{in}$), minimum carry-out ($\min_{out}$) or a combination of maximum carry-in and minimum carry-out $\left(\max_{in/out}\right)$, depending on the current marketing objectives and market conditions.

Algorithm 1 gives the algorithmic process of the game in the MNP business of the firm. First, the important features $X_i, i = 1, 2, ..., n$ of the firm's public and/or private business data are selected as inputs and the relative importance matrix β between X_i and X_j is calculated as follows:

$$\beta_{i,j} = \begin{cases} 1, X_i \text{ is more important than } \mathrm{X}_j \\ 0, X_i \text{ is not as important as } \mathrm{X}_j \end{cases} \tag{2.1}$$

The score of feature importance a_i and weight ω_i can be calculated using Eq.2.2 andEq.2.3:

$$a_i = \sum_{j=1}^{n} \beta_{i,j} \tag{2.2}$$

$$\omega_i = \frac{a_i}{\sum\limits_{j=1}^{n} a_i} \tag{2.3}$$

Assuming a marketing objective of $\tilde{x}_i^v$ for feature i and marketing strategy v in the MNP business, where v denotes $\max_{in}, \min_{out}$ or $\max_{in/out}$, the baseline score for the marketing objective is shown in Table 2.1, $\tilde{s}_i^v$ is the baseline score for $\tilde{x}_i^v$ and the combined value of marketing strategy v is:

$$Z_v = \sum_{i=1}^{n} \omega_i \tilde{s}_i^v \tag{2.4}$$

V represents the combination value of marketing strategies, used to measure the effectiveness of different marketing strategies. S represents the benchmark score for features, used to evaluate marketing objectives for

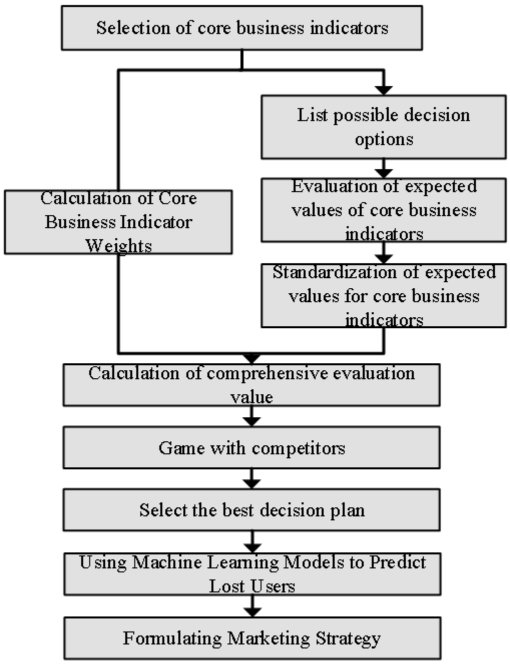

Fig. 2.1: Workflow of Attack and Defense Marketing Decision Evaluation Method Based on Game Theory.

Algorithm 1 Game algorithm for Chinese enterprises to port their numbers to other networks

Selecting core indicators $X_i, i = 1, 2, ..., n$.
The parameters $\beta_{i,j}, a_i$ and $\omega_i for v \in \left\{\max_{in}, \min_{out}, \max_{in/out}\right\}$ *do* are calculated by Eq.2.1 to Eq.2.3 respectively.
For i^{th} characteristics and v^{th} marketing strategy, $i = 1, 2, ..., n$,set operational target expectations $\widetilde{x}_i^v$.
The scores were calculated using Table 2.1, $\tilde{s}_i^v, i = 1, 2, ..., n$ to calculate the overall evaluation value $Z_v = \sum_{i=1}^{n} \omega_i \tilde{s}_i^v$.
For All $v \in \left\{\max_{in}, \min_{out}, \max_{in/out}\right\}$ *do.*
For All $o \in \{A, B, C\}$ *do.*
Calculate the score by Eq.2.5: $S_{v(i)}^{o(m)}$.
Calculate mathematical expectations for all $v \in \left\{\max_{in}, \min_{out}, \max_{in/out}\right\}$ and $o \in \{A, B, C\}$ all $E\left\{S_{v(i)}^{o(m)}\right\}$.
Output the best i for all $o \in \{A, B, C\}$.

different features. J may represent an index of a specific marketing strategy or market segment, and its specific role needs to be further explained.

Due to the fact that the three major telecommunications companies A, B, C in China have three possible marketing strategies in actual games, namely $\max_{in}, \min_{out}$ and $\max_{in/out}$, and any two companies will play games with each other, there are 27 possible permutations and combinations of their game strategies. In order to obtain the optimal solution of each player, all possible game theory results are calculated by Eq.2.5.

$$S_{v(i)}^{o(m)} = 2Z_{v(i)}^{o(m)} - Z_{v(j)}^{o(n)} - Z_{v(k)}^{o(l)} \tag{2.5}$$

Among them, $S_{v(i)}^{o(m)}$ is the final score of Enterprise $o\,(m)$ after selecting Strategy $v\,(i)$ and playing games with the other two enterprises. Z_v^o is the comprehensive score of enterprise o marketing strategy v calculated by Eq.2.4, where the values of $o \in \{A, B, C\}$ $v \in \left\{\max_{in}, \min_{out}, \max_{in/out}\right\}, i, j, k, m, n, l$ are 1, 2, or 3, and

Table 2.1: Scoring benchmarks.

Characteristic	$Score \geq 5$	$Score \geq 4$	$Score < 4$	$Score \geq 3$	$Score < 3$	$Score \geq 2$	$Score < 2$	$Score \geq 1$	$Score < 1$	$Score = 0$
$\tilde{x}_1$	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}	S_{19}
$\tilde{x}_2$	S_{20}	S_{21}	S_{22}	S_{23}	S_{24}	S_{25}	S_{26}	S_{27}	S_{28}	S_{29}
$\tilde{x}_3$	S_{30}	S_{31}	S_{32}	S_{33}	S_{34}	S_{35}	S_{36}	S_{37}	S_{38}	S_{39}
					$\vdots$					
$\tilde{x}_n$	S_{n0}	S_{n1}	S_{n2}	S_{n3}	S_{n4}	S_{n5}	S_{n6}	S_{n7}	S_{n8}	S_{n9}

$m \neq n \neq l$.

By calculating the expected value $E\left\{S_{v(i)}^{o(m)}\right\}$ for all v and o, the optimal strategy for enterprise $o(m)$ can be obtained.

$$v\{i\} \leftarrow \max\left[E\left\{S_{v(i)}^{o(m)}\right\}\right] \tag{2.6}$$

All $v \in \left\{\max_{in}, \min_{out}, \max_{in/out}\right\}$ and all $o \in \{A, B, C\}$ are considered here.

2.2. Subscriber churn prediction algorithm. In the Chinese telecommunications industry, subscribers are limited to three major service providers: China Mobile, China Unicom, and China Telecom. With the implementation of the Mobile Number Portability (MNP) policy [9], subscribers are now able to switch between these providers while retaining their original phone numbers. This policy introduces new challenges for telecom companies, as they must now proactively anticipate and manage subscriber turnover in order to maximize their revenue. To stay competitive, these businesses must develop strategies that not only attract new customers but also reduce churn by understanding the factors that influence customers' decisions to switch providers. By leveraging data analytics and predictive modeling, telecom companies can optimize their marketing efforts and improve customer retention rates.

Assuming that user i decides to move from firm m to firm n, the transfer decision $p_{m,n}^i$ can be expressed as:

$$p_{m,n}^i = \sum_k \delta_{m,k} p_{k,n}^i \tag{2.7}$$

where $\delta_{m,k}$ is the Dirac Delta formula. The transfer matrix can be constructed from Eq.2.7. For any given user, his choice is deterministic and unique. Thus, the user's transfer decision $P_{m,n}$ can be written as:

$$P_{m,n} = \begin{cases} 1, \text{Usertransitionfromton} \\ 0, \text{other} \end{cases} \tag{2.8}$$

From Eq.2.8, it can be seen that only one element of the transfer matrix is 1 and all other elements are 0. Assuming that there are N subscribers in the entire communications market, at time $t+1$ (in months), the total number of subscribers carrying into s firm is:

$$in\,(m, t+1) = \sum_{i=1}^{N} \sum_{n \neq m} p_{nm}^i\,(t) \tag{2.9}$$

and the total number of users carried out from the enterprise m:

$$out\,(m, t+1) = \sum_{i=1}^{N} \sum_{n \neq m} p_{nm}^i\,(t) \tag{2.10}$$

In order to attract users to bring in and reduce users to bring out, the transfer decision matrix of any user i needs to be known in advance M^i. According to previous research, there are many factors that influence a user's transfer decision, including education level, package bundle status, portability costs, call quality, monthly fees and call duration [10]. Therefore, a user's transfer decision will depend on the attributes of the user, i.e.:

$$p^i_{m,n}\left(t\right) = p^i_{m,n}\left(t, D^i\right) \tag{2.11}$$

where D^i is the characteristic of user i.

The challenge now is to obtain $p^i_{m,n}$ from $\left(t, D^i\right)$ in Eq.2.11. Recently, some studies have proposed to predict $p^i_{m,n}$ using Markov chains, logistic, social network analysis and some other statistical methods [11]. Given that consumers have three options—remaining with the original firm and moving to the other two firms—the transfer decision problem is essentially a triple classification problem. In order to achieve great prediction performance and interpretability, users' transfer decisions are predicted in this research using a stacked neural network model (SNN) [12, 13]. Additionally, some additional models were employed for comparisons, including Random Forest (RF), Decision Tree (DT), Long Short Term Memory Network (LSTM), Gradient Boosted Tree (GBDT), XG Boost, Extra Tree, and Fully Connected Neural Network (NN) [14, 15]. Due to the relatively small number of user carry-outs, resampling and down-sampling techniques were used to address the problem of data imbalance. After a user churn prediction model is constructed, the expected return for each strategy can be expressed as:

$$\max\left\{in\left(m, t+1\right)\right\} = \max\left\{\sum_{i=1}^{N}\sum_{n\neq m}\left(p^i_{nm}\right)^*\left(t, D^i\right)\right\} \tag{2.12}$$

$$\min\left\{out\left(m, t+1\right)\right\} = \min\left\{\sum_{i=1}^{N}\sum_{n\neq m}\left(p^i_{nm}\right)^*\left(t, D^i\right)\right\} \tag{2.13}$$

$$\max\left\{\frac{in\left(m, t+1\right)}{out\left(m, t+1\right)}\right\} = \max\left\{\frac{\sum\limits_{i=1}^{N}\sum\limits_{n\neq m}\left(p^i_{n,m}\right)^*\left(t, D^i\right)}{\sum\limits_{i=1}^{N}\sum\limits_{n\neq m}\left(p^i_{m,n}\right)^*\left(t, D^i\right)}\right\} \tag{2.14}$$

where p^* is the machine learning prediction of user transfer decisions.

Using this approach, whatever the strategy obtained from game theory, the transfer decision matrix first needs to be obtained by a machine learning model. For strategies $\max_{in}, \min_{out}$ and $\max_{in/out}$, the machine learning models for predicting the propensity of users to shift are the same. After identifying the users that will be lost, companies can target their marketing activities. At the same time, the results of the machine learning model can be used as a basis for game-theoretic decisions.

3. Condition loosening and expansion. According to the general framework mentioned earlier, we need to compute μ^s_j (i.e., the probability that node j is activated when the seed set is S) to calculate the diversity of the affected population. However, a common way to compute μ^s_j is to run Monte Carlo simulations enough times (e.g., 20,000 times), which is very time-consuming, even for medium-sized networks (e.g., a graph with 10k nodes and 100k edges. In fact, some studies give methods on how to estimate μ^s_j, however, most of these studies address special cases, such as assuming that influence propagates along the shortest path, or that nodes have a small probability of influencing each other [16]. Furthermore, most of these methods do not have theoretical performance guarantees.

This motivates us to define a direct diversity measure that can be effectively computed. To simplify the problem, we focus on optimizing the diversity of the seed set nodes instead of directly optimizing the diversity of the affected population. This approach is inspired by the concept of "homogeneity" in social networks [17], where the spread of influence often results in a concentration of similar behaviors or opinions. By targeting the seed set, we aim to maximize the potential for diversity within the broader network, thus indirectly achieving the desired diversity in the affected population. This approach is referred to as a relaxation problem, where the original, more complex problem is simplified to facilitate computational efficiency while maintaining the core objectives.

$$\begin{array}{l} \max F_s\left(S\right) = \left(1-\gamma\right)\frac{\sigma(S)}{\overline{\sigma}} + \gamma\frac{D(S)}{\overline{D}}, \\ s.t. \quad S \subseteq V, |S| = K \end{array} \tag{3.1}$$

where $D\left(S\right)$ is the value of the diversity of the seed set S.

In particular, we choose two forms of $D\left(S\right)$:

$$D\left(S\right) = \sum_{i=1}^{C} f\left(\sum_{j\in S} w_{ji} * 1\right) \tag{3.2}$$

$$D\left(S\right) = \sum_{i=1}^{C} f\left(\sum_{j\in S} w_{ji}\sigma\left(\{j\}\right)\right) \tag{3.3}$$

We call Eq.3.2 a consistent diversity metric because each node's contribution to diversity is the same.

We call Eq.3.3 a weighted diversity metric because the contribution of seed node j to diversity is weighted by the influence of that node.

However, the time complexity of Eq.3.1 is still high, as the calculation of $\sigma\left(S\right)$ still requires a Monte Carlo simulation. Nevertheless, this relaxed thinking can guide us to incorporate diversity into some impact maximization heuristics. Next, we present specific methods using the degree centrality metric (degree centrality) and PageRank as examples.

The degree centrality approach is very simple: the node with the largest K degrees is selected, while PageRank selects the node with the largest K PageRank values. These heuristics can also be seen as optimizing specific submodular functions. For example, the degree centrality approach can be seen as finding a set S of nodes of size K to maximize $\sum_{i\in S} \deg\left(i\right)$, where $\deg\left(i\right)$ is the degree of node i. Thus, we can easily incorporate diversity into the degree centrality approach using this relaxed form (Eq.3.1):

$$\begin{array}{l} \max F_d\left(S\right) = \left(1-\gamma\right)\frac{\sum_{i\in S}\deg}{\overline{\deg}} + \gamma\frac{D(S)}{\overline{D}}, \\ s.t. \quad S \subseteq V, |S| = K \end{array} \tag{3.4}$$

Similarly, we have consistent indicators of diversity:

$$D\left(S\right) = \sum_{j\in S}^{C} f\left(\sum_{j\in S} w_{ji} * 1\right) \tag{3.5}$$

and weighted diversity indicators:

$$D\left(S\right) = \sum_{i=1}^{C} f\left(\sum_{j\in S} w_{ji} \deg\left(j\right)\right) \tag{3.6}$$

Replacing the degree of a node with the PageRank value of the node, we formalize the diverse PageRank heuristic as follows:

$$\begin{array}{l} \max F_P\left(S\right) = \left(1-\gamma\right)\frac{\sum_{i\in S}PR(i)}{\overline{PR}} + \gamma\frac{D(S)}{\overline{D}}, \\ s.t. \quad S \subseteq V, |S| = K \end{array} \tag{3.7}$$

Table 4.1: Some statistics for the graph data.

Figure	Number of nodes	Number of edges	Number of categories
Film consumption network	2,113	276,676	20
Yahoo Q&A Network	42,043	194,619	31

where $PR(i)$ is the PageRank value of node i. Similarly, we have consistent diversity metrics:

$$D(S) = \sum_{i=1}^{C} f\left(\sum_{j \in S} w_{ji} * 1\right) \tag{3.8}$$

and weighted diversity indicators:

$$D(S) = \sum_{i=1}^{C} f\left(\sum_{j \in S} w_{ji} PR(j)\right) \tag{3.9}$$

4. Analysis of experimental results. Movie Lens and Yahoo Answers data were included in the experimental dataset we used. The benchmarking technique we employed and the evaluation metrics are then presented. Finally, the approaches' performance comparison and efficiency comparison findings are shown.

4.1. Experimental datasets. We use two real datasets, one from Movie Lens * and the other from Yahoo! Answer). Especially, we have built a movie consumption network based on the Movie Lens dataset and a Yahoo!

In a movie consumption network, each node represents a movie. If a user evaluates movie A and then movie B, we insert a directed edge $(A \to B)$ between A and B. We then replace parallel edges (heavy edges) with edges with weights, which are the number of parallel edges. Then, because these edges are probably noisy, we filter away edges with weights below a predetermined threshold (in the trials, this threshold was set at 10). Then, we multiply the incoming edges of every node by a weighting factor such that the sum of the incoming edge weights for every node is a number between 0 and 1 (representing the probability of the node being activated, which in our experiments was set to 0.1), so the edge weights can be thought of as activation probabilities. So an edge with weight $P\ (A \to B)$ means that the probability of a user watching movie A and then going to see movie B is P.

In the Yahoo Answers network, we selected the 30 categories with the most responses as our experimental data. Nodes represent users, both askers and answerers. The category of a node is the category of the question asked by that user. If a user A answers a question asked by user B, we insert a directed edge $(A \to B)$ between A and B. These edges are then assigned weights in a similar way to the movie consumption network and are not repeated.

We assume that the distribution of nodes across the tagged categories is uniformly collinear due to the lack of detailed information regarding the exact affiliation of nodes to categories. For instance, a movie may be labeled under both "love" and "war" categories, but it is unclear how the movie aligns with each category or to what degree. In the context of films, the category is defined by the genre of the film; for Yahoo Answers data, the category corresponds to the topic area in which the user asked the question, such as "technology," "health," or "entertainment." This uniform assumption simplifies the analysis but also limits the accuracy of category-specific inferences. Table 4.1 presents the basic details and properties of the two datasets used in this study, providing further insights into their structure and the relationships between nodes and categories. By considering the assumptions about the uniform distribution and categorization, we can attempt to better understand the patterns in user behavior, content interaction, and their impact on network dynamics. However, future work could explore incorporating more granular data on category affiliation to improve model precision and offer a deeper understanding of cross-category influence.

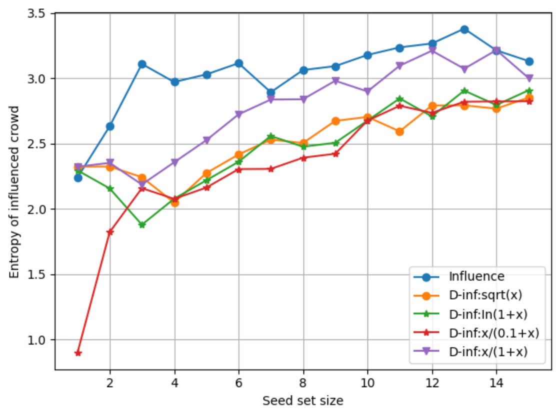

Fig. 4.1: Influence diversity on the diversity of affected populations in two datasets (Movies).

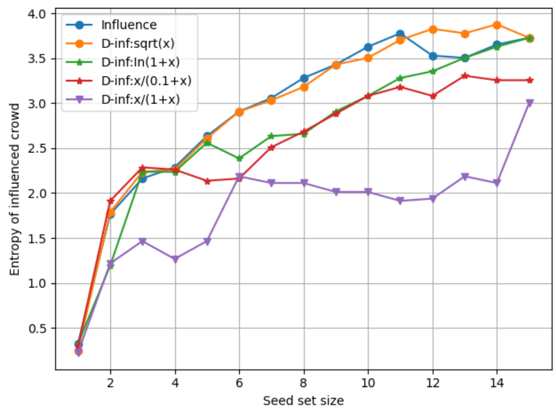

Fig. 4.2: Influence diversity on the diversity of affected populations in two datasets (Yahoo).

4.2. Baseline methods for experiments. In order to evaluate the effectiveness of our proposed diverse impact maximization approach, we chose the following baseline methodology:

Influence: Basic greedy algorithm for maximizing influence (Eq.3.1).

PageRank: greedy algorithm for maximizing the sum of the PageRank values of the set of seed nodes.

Degree: greedy algorithm for maximizing the sum of the degrees of the set of seed nodes (degree centrality method).

The methods we proposed in the previous subsections are summarized as follows:

D-mf: greedy algorithm for maximizing diversity impact (Eq.2.14).

Seed-DU(W): greedy algorithm for maximizing the diversity and influence of the set of seed nodes with a consistent (weighted) diversity measure (Eq.3.1).

Deg-DU(W): a greedy algorithm for maximizing the sum of the diversity of the set of seed nodes with a consistent (weighted) diversity measure (Eq.3.4).

PR-DU(W): a greedy algorithm to maximize the sum of the diversity PageRank values of the set of seed nodes with a consistent (weighted) diversity measure (Eq.3.7).

Although our formalism is general for both the independent cascade model and the linear threshold model, the implementations of $In-fluence, D-Inf$ and $Seed-DU\ (W)$ are different under different influence models. We chose the independent cascade model for the experimental validation of our proposed strategy because it is more widely used and shares similarities with the linear threshold model, such as the fact that they both have equivalent live-edge graph models and their generalised versions are equivalent.

4.3. Evaluation indicators. We chose the widely used Shannon entropy to measure the diversity of the influenced population. After obtaining the seed set using each method, we run 20,000 Monte Carlo simulations under an independent cascade model to estimate the influence of the seed set $\sigma\left(S\right)$ and a vector consisting of the probability of each node being activated μ^S. The Shannon entropy of the influenced population is defined as:

$$Shannonentropy = \sum_{i=1}^{C} -p_i \log_2 p_i \tag{4.1}$$

where $p_i = \frac{\sum_{j=1}^{|V|} w_{ji}\mu_i}{\sum_{j=1}^{|V|}\mu_i} p_i$, has an intuitive physical meaning, i.e. the number of people activated in category i as a proportion of all activated people. Please refer to Fig.4.1, Fig.4.2, Fig.4.3 and Fig.4.4 for details.

4.4. Diversification of the seed set. Fig.4.5 to Fig.4.12 respectively show the results of seed set diversification on the movie consumption network and Yahoo Q&A network.

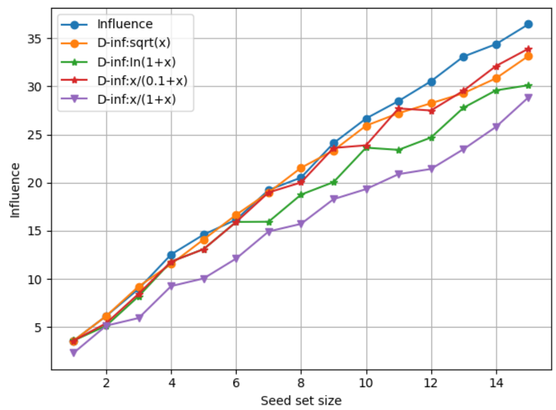

Fig. 4.3: The impact of diversification on two datasets (Movie).

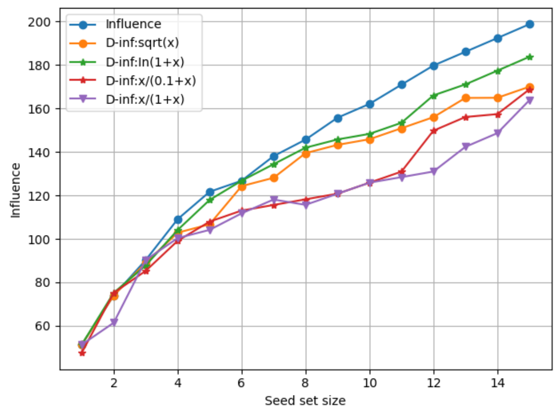

Fig. 4.4: The impact of diversification on two datasets (Yahoo).

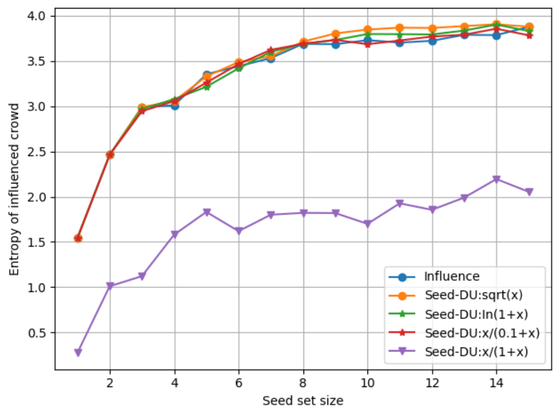

Fig. 4.5: Seed diversification affects the diversity of the affected population in the film consumption network (Unified).

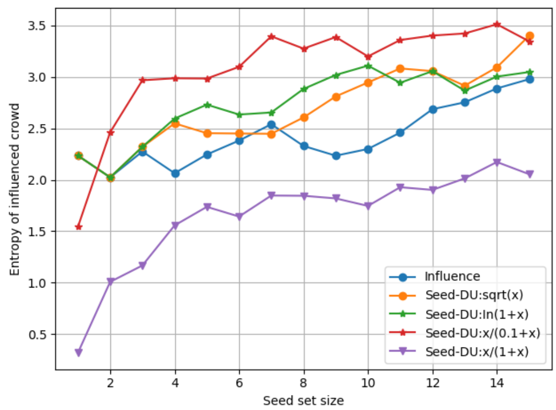

Fig. 4.6: Seed diversification affects the diversity of the affected population in the film consumption network (Weighted).

In both figures, γ is also set to 1, as in the previous subsection. Two things can be seen from the experimental results: first, diversifying the seed set (i.e., maximizing $F_s(S)$ in Eq.3.1) also leads to a more diverse population, which confirms the validity and reasonableness of the model. Second, we can see that in the $\gamma = 1$ setting, the consistent diversity indicator generally corresponds to higher diversity (and lower influence) than the weighted diversity indicator. Since $\gamma = 1$ corresponds to the maximum value of diversity we can obtain, these plots show that the consistent diversity indicator is more tunable in terms of balancing diversity and influence.

5. Conclusion. In response to the increasing market rivalry and the need for businesses to adapt quickly, this study introduces a novel framework that combines game theory and machine learning to enhance marketing strategies. By applying decision-making principles from game theory and incorporating conditional relaxation techniques, the proposed framework helps businesses proactively identify user behavior patterns, such as the likelihood of user churn or migration. Using a machine learning algorithm, the framework allows businesses to tailor their marketing strategies to different user segments, improving engagement and retention. This framework has been successfully validated through experimental data and subsequently productized, leading

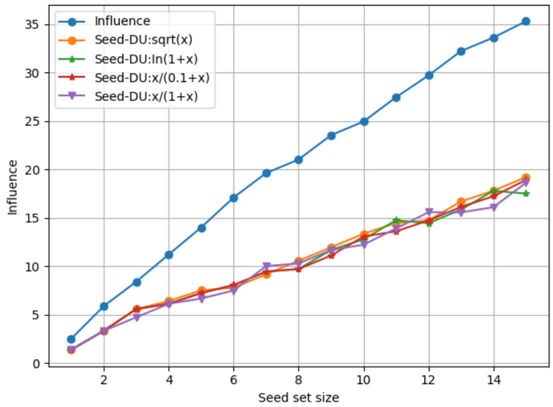

Fig. 4.7: Seed diversification affects the diversity of the affected population in the film consumption network (Unified).

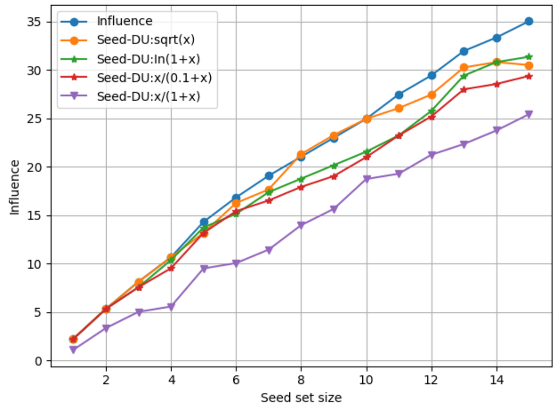

Fig. 4.8: The impact of seed diversification on the film consumption network (Weighted).

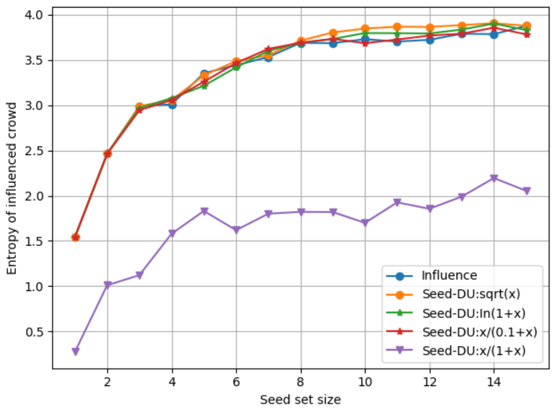

Fig. 4.9: Diversity of Seed Sets and the Affected Population of Yahoo Q&A Network (Unified).

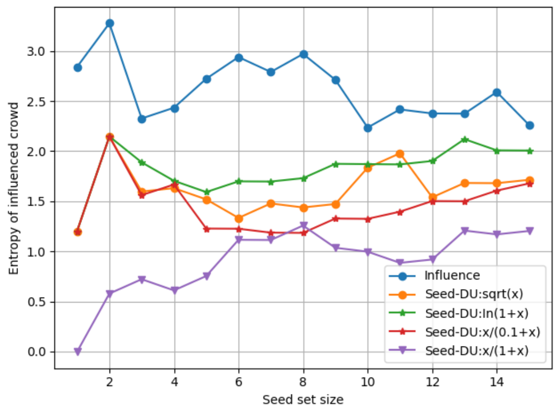

Fig. 4.10: Diversity of Seed Sets and the Affected Population of Yahoo Q&A Network (Weighted).

to a significant boost in marketing effectiveness. For instance, one large company implemented the solution and saw a 50% increase in its subscriber base, demonstrating the practical benefits of applying these advanced techniques to real-world marketing challenges.

Data Availability. The experimental data used to support the findings of this study are available from the corresponding author upon request.

Funding Statement. This work was supported by Provincial Youth Talent Project of Gansu Province (2024QNGR32) and Research on the Path of Empowering Rural Revitalization with E-commerce in Shaanxi Province under the Background of Digital Economy (2024ZD462).

REFERENCES

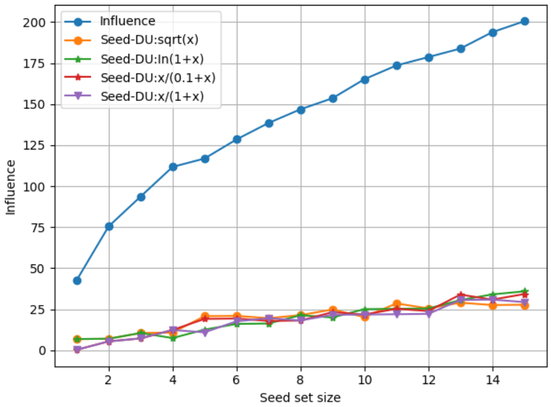

Fig. 4.11: The impact of diversified seed sets on Yahoo Q&A network (Uniform).

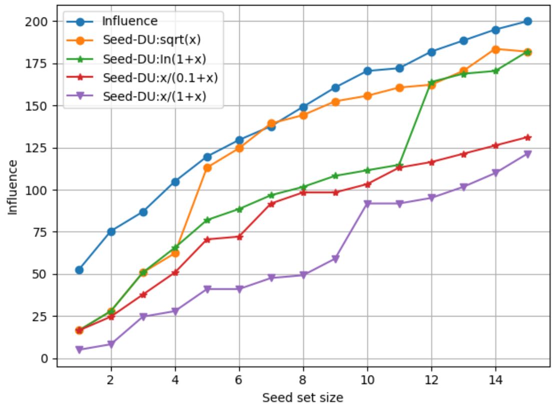

Fig. 4.12: The impact of diversified seed sets on Yahoo Q&A network (Weighted).

[1] Abreu, M., Ferreira, F., & Silva, J. *To be or not to be sustainable in an emerging market? conjoint analysis of customers' behavior in purchasing denim jeans.* Journal of Fashion Marketing and Management: An International Journal, (2022), 26(3), 452-472.

[2] Ying, K., & Zhong, W. *A new data-driven robust optimization approach to multi-item newsboy problems.* Journal of Industrial and Management Optimization, (2023), 19(1), 197-223.

[3] Han, W., & Bai, B. *Pricing research in hospitality and tourism and marketing literature: a systematic review and research agenda.* International journal of contemporary hospitality management, (2022), (5), 34.

[4] Petersen, J. A., & Schmid, F. *Leveraging stakeholder networks with outside-in marketing.* Industrial Marketing Management, (2021), 92(Suppl), 72-75.

[5] Ibrahimov, Z., Hajiyeva, S., Nazarov, V., Qasimova, L., & Ahadov, V. *Bank efficiency analysis of financial innovations: dea model application for the institutional concept.* Marketing & Management of Innovations, (2021), (1), 290-303.

[6] Pollak, F., Vavrek, R., J Váchal, Markovi, P., & Konen, M. *Analysis of digital customer communities in terms of their interactions during the first wave of the covid-19 pandemic.* Management & Marketing: Challenges for the Knowledge Society, (2021), 16(2), 134-151.

[7] Almuraqab, N., & Andleeb, N. *Special issue 2, 2021 1 marketing management and strategic planning.* Academy of Strategic Management Journal, (2021), 20(2), 1-19.

[8] Jiang, H., & Cheng, Y. *Customer–brand relationship in the era of artificial intelligence: understanding the role of chatbot marketing efforts.* Journal of Product & Brand Management, (2022), 31(2), 252-264.

[9] Le, L. H., Kim, J., & Min, J. E. *Impacts of brand familiarity and brand responses on perceived brand credibility, similarity, and blog recommendation intention: a study of corporate blogs.* Journal of Fashion Marketing and Management: An International Journal, (2022), 26(2), 328-343.

[10] Rossi, M., Festa, G., Carli, M. R., & Kolte, A. *Envisioning the challenges of the pharmaceutical sector in the indian health-care industry: a scenario analysis.* Journal of Business & Industrial Marketing, (2022), 37(8), 1662-1674.

[11] Shafiee, M. M. *Knowledge-based marketing and competitive advantage: developing new scales using mixed method approach.* Journal of modelling in management, (2021), (4), 16.

[12] Irfan, I., Chan, F., Khurshid, F., & Sumbal, M.*Toward a resilient supply chain model: critical role of knowledge management and dynamic capabilities.* Industrial Management & Data Systems, (2022), 122(5), 1153-1182.

[13] Urbanski, M. *Company global competitive strategy and employee's awareness.* Marketing and Management of Innovations, (2021), 5(2), 49-64.

[14] Smith, C., Ogutu, M., Munjuri, M., & Kagwe, J. *The effects of foreign market entry strategies on financial performance of listed multinational firms in kenya.* European Journal of Business Management and Research, (2021), 6(3), 216-225.

[15] Acharya, S., Tajane, V., & Prof. Shubhangi. *Novel technologies for processing mushrooms and its marketing strategies.* International Journal of Engineering and Management Research, (2021), 11(1), 93-96.

[16] Abou-Shouk, M., & Soliman, M.*The impact of gamification adoption intention on brand awareness and loyalty in tourism: the mediating effect of customer engagement.* Journal of Destination Marketing and Management, (2021), 20(2), 100559.

[17] Crick, J. M., & Crick, D.*Rising up to the challenge of our rivals: unpacking the drivers and outcomes of coopetition activities.* Industrial Marketing Management, (2021), 96(3), 71-85.

Edited by: Bradha Madhavan
Special issue on: High-performance Computing Algorithms for Material Sciences
Received: Sep 9, 2024
Accepted: Feb 27, 2025