IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 16 April 2024, accepted 19 May 2024, date of publication 28 May 2024, date of current version 14 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3406533

RESEARCH ARTICLE

In the Arena of the Content War: A Social Network Analysis Approach for Content Differentiation in VOD Platforms

ALI NOROOZIAN[1], BABAK AMIRI[2], AND SAMIRA RAMEZANI[2]
[1]School of Management, Economics and Progress Engineering, Iran University of Science and Technology, Tehran 16846-13114, Iran
[2]School of Industrial Engineering, Iran University of Science and Technology, Tehran 16846-13114, Iran

Corresponding author: Babak Amiri (babakamiri@iust.ac.ir)

ABSTRACT The quest for content differentiation and audience engagement is paramount in the competitive landscape of Video-on-Demand (VOD) platforms. The paper proposes a system investigating VOD platform series data to recommend actors and themes to VOD vendors. To achieve this, we analyzed the storylines of 72 VOD movie series to extract keywords. Using these keywords, we generated a network using Social Network Analysis. By examining the network, we identified content gaps. After semantic clustering, we utilized TOPSIS to determine keywords with high centrality scores. These are a foundation for connecting peripheral keywords and generating new content. The paper's findings suggest that the proposed content-based recommender system can help VOD platforms create innovative storylines by leveraging structural holes in a TV series content keyword graph. The study also suggests that successful directors take risks by combining different keywords from various network parts. This approach allows them to attract a wider range of audiences and increase their chances of success.

INDEX TERMS Video on demand (VOD), movie recommender system, content analysis, bipartite graph, social network analysis (SNA), TV series.

I. INTRODUCTION

Nowadays, TV series have asserted a paramount position within the ensemble of cultural commodities, and the discussion of their storylines has transformed into a focal point for interpersonal discourse [1]. On the other hand, due to the fast pace of today's life and the advancement of technology, people prefer consuming these kinds of entertainment via video-on-demand (VOD) platforms rather than dedicating time to the cinema [2]. VOD platform is an online service that allows users to access video content such as movies and TV shows anytime. The advantage inherent in these platforms, relative to the act of attending cinemas, lies in the fact that:

- In the cinema, the audience must be present at a specific time and place to watch a film; however, on VOD platforms, they are empowered to access the desired content at their preferred time and location 24/7.
- In the cinema, you can only watch a single film for one ticket, whereas on VOD platforms, you can access a considerable amount of content.
- In the cinema, one ticket is sold per person, whereas by purchasing a subscription to VOD platforms, you can view content alongside others.

It can be inferred that VOD platforms offer advantages in terms of time and place and are also more cost-effective due to the cost ratio to the quantity of content.

It is worth mentioning that prominent cinema producers have also become aware of the audience preference change and are striving to have a share in the market of these platforms. Among the most renowned producers, one can mention Warner Bros. and Walt Disney, both of which have introduced their dedicated platforms, namely HBO Max and Disney Plus, to the market. Subscribing to these platforms allows users to access all of the films produced by the producers over time. Also, it provides an opportunity to watch on-screen movies produced by the same producers.

The associate editor coordinating the review of this manuscript and approving it for publication was Yin Zhang.

As the popularity of VOD platforms continues to grow, the competition among them has also increased. One of the initial risks arising from this competition is that if remarkable cinema producers intend to establish their exclusive platforms, other competitors will face content scarcity due to copyright regulations. Consequently, the content war becomes a central battleground in this competitive market.

For a better understanding of the current landscape of VOD platforms, it should be noted that according to a report by Acumen Research and Consulting, the market is anticipated to undergo substantial growth, projected to increase from \$64.3 billion in 2022 to approximately \$329.2 billion by the year 2032 [3]. In an interview with Ehsan Alikhani, a prominent Iranian producer and host of popular and successful large-scale production shows, it is highlighted that what sustains cinematic films is a well-crafted storyline, and accurate character portrayal engages the audience throughout a series. The audience must accompany themselves along the series' trajectory, witnessing the character's challenges. Also, a study regarding the perspectives of American viewers revealed that 55% of subscribers of VOD platforms actively seek out new shows every week. In comparison, 62% indicated challenges in finding content to watch [4]. So, If VOD platforms aim to secure a share of this market, they must emerge victorious in the content war.

Now that the major competition revolves around content and considering the formidable presence of influential film producers in this arena, coupled with the risk of content scarcity, platform vendors must strategically invest in novel and captivating storylines to safeguard their market share. This research aims to propose a recommender system for VOD vendors to create innovative storylines by leveraging structural holes in a keyword graph of TV series content, harnessing social network analysis (SNA) as a tool.

SNA involves a graph theory-based depiction in which entities under investigation are represented as nodes, and the connections between them are defined as edges. These edges can be strong or weak ties based on the nature of the relationship and directed or undirected depending on the direction of the connection [5]. Figure 1 illustrates an example of a collaboration network among actors. The figure shows that nodes represent actors and edges depict their collaborations.

A significant aspect of this approach is its transition from individual-centered attributes to the exploration of pairs of individuals and their interconnections [6]. This distinctive aspect of SNA allows us to scrutinize the characteristics of a network while also allowing us to discover structural holes, thereby revealing potential opportunities that might otherwise have remained unnoticed. Structural holes are defined as the gaps or intervals between nodes in a network where there is no connection between non-redundant nodes [7]. Non-redundant contacts are connections or interactions between individuals or entities within a network that are distinct and singular, meaning they are not repeated or duplicated. For instance, a non-redundant contact would be a friend not connected to your other friends. The significance of non-redundant contacts lies in their ability to offer exclusive information and viewpoints inaccessible through other contacts. Furthermore, they can aid in the dissemination of information and resources across various sections of a network, thereby enhancing its efficiency and resilience [8].

Now that we are familiar with the SNA method, we will summarize the story and raise the research questions. We suppose that the main competition in the VOD market is content innovation. To tackle the challenge, we aim to detect content gaps based on the graph theory to suggest novel topics to the VOD vendors. Based on what has been said, the research questions are as follows:

RQ.1: What are the content gaps in the VOD platforms?

RQ.2: How can we address this challenge to foster the creation of innovative ideas?

We suggest the generation of a network of keywords to address these challenges. By identifying the holes or empty spaces between keywords, the proposed system can provide recommendations for new and creative storylines that bridge those gaps and appeal to a wider audience. In this study, we define storyline keywords as a subset of words or phrases that capture the essence of a movie or series storyline and are associated with the audience's perception of that content. These keywords reflect what the audience thinks of when they hear the name of the movie or series and are used to identify patterns and connections between different storylines in our analysis.

In the following sections, we will illustrate the research gap and approach to creating our recommender system. This will include an explanation of the methodology employed, as well as a depiction of the data sources and evaluation metrics used. We will also present the results of our experiments. Finally, the implications of our research for the VOD market and future directions for research in this area will be discussed.

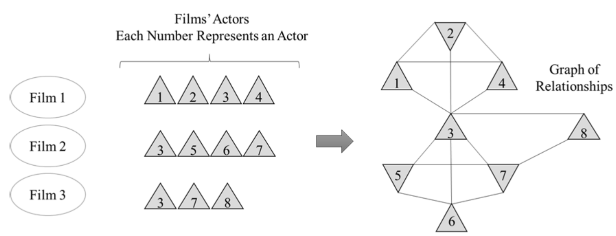

FIGURE 1. Actors network formation.

II. LITERATURE REVIEW

The realm of audiovisual media, particularly datasets derived from cinema, has consistently constituted a valuable asset for data scientists owing to its profound cultural and sociological revelations. Given that such data distinctly mirror societal dynamics and ambitions, many scholarly articles have employed cinematic datasets as a basis for examination. Furthermore, databases like Movie Lens and IMDB have become primary repositories for researchers pursuing data.

The majority of studies conducted in the realm of cinema have focused on the prediction of the success and profitability of movies through different data mining or machine learning methods [9], [10], [11], [12], [13], [14], [15]. Among the scarce studies that have directed their attention toward the content of cinematic works, one has considered references to a particular movie as the basis for generating a graph. This implies that any reference to a movie, such as alluding to a famous scene or a renowned dialogue from that film if utilized in another movie, is represented by a directed edge towards the reference film. The findings of this research illustrate that film festivals do not immediately embrace influential cinematic works upon their initial entry. Rather, a considerable amount of time needs to pass before these works gather the attention of audiences and critics [16]. In another study, by employing SNA and by defining the cinematic network as a collaboration network, two significant outcomes have been derived. Firstly, a direct relationship exists between the position of cinema actors in a collaboration network and their popularity on social media. In other words, the more popular an actor is, the better their position within the network. The second point is that people tend to show interest in genres where they can observe their daily life needs and concerns being portrayed [17].

In today's context, as VOD platforms have become a significant part of entertainment, alongside capturing the attention of investors, they have also become the subject of extensive research studies. Most of the studies carried out in this field aim to develop recommender systems. The main objective of a recommender system is the identification and extraction of consumer preferences [18]. It is shown that the quality of the recommender system has a stronger influence compared to social impacts like WOM [19]. Also, new studies tend to examine the influence of social behaviors on VOD trends [20].

During the initial investigations into recommendation systems, a research article presented an architecture for a recommendation system comprising three central modules: data processing, data serving, and a recommender database. The data processing phase encompasses tasks like filtering VOD items, forming item and user profiles, implementing recommendation algorithms, and generating data structures. The personalized recommendations aim to suggest VOD content, like movies or shows, by analyzing user profile preferences and utilizing a calculated utility function to predict user interest in specific items [21]. In another paper, it is stated that the concept behind the recommender system lies in recognizing events suitable for the target audience as a valuable foundation for suggesting pertinent movies. In this study, the researchers employed learning algorithms to detect events within news data streams suitable for recommending specific films. The authors observed that their approach achieved remarkable precision in recommending relevant movies based on events [22]. However, identifying events fitting for item recommendations poses various challenges, such as pinpointing appropriate sources and devising suitable methods for processing and storing the encompassed information. Recommendations centered around news events remain relatively unfamiliar to most users. A study introduced an offline evaluation procedure to enhance the effectiveness of recommendations on user pages, as opposed to conventional fixed options like the latest movies. This method quantifies how much a new model enhances existing sets of carousels. The paper additionally suggests expanding ranking metrics to consider position bias and assesses two methods for ranking carousels, assuming a layout agnostic to the algorithms applied. The authors conducted experiments on publicly accessible movie datasets, revealing considerable shifts in algorithm rankings within a carousel framework [23]. Most current methods for evaluating recommender systems focus only on accuracy and cannot consider user requirements or business models. However, a 3D Recommender System Benchmarking Model was proposed to address this issue. This model evaluates recommender systems based on real-world user needs, large and diverse data sets, and functional and non-functional requirements. It covers three evaluation aspects: business models, user requirements, and technical constraints [24].

It is worth mentioning that a recommender system is confronted with several challenges and issues, which include changing user preferences, sparsity, scalability, synonymy, and privacy. The dynamic nature of user preferences challenges the system since it is based on user interest and profile. The sparsity issue arises as most users only rate a few items, making it challenging to determine their preferences. Scalability becomes an issue as the number of users and items increases, requiring more resources for processing information. The problem of synonymy makes it difficult for the system to differentiate between similar items with different names. Finally, to provide accurate recommendations, the system must collect a substantial amount of user data, including demographic and location information, which raises privacy concerns [18], [25], [26].

Alongside the significance of content and marketing, some believe that user experience (UX) plays a crucial role in the market of VOD platforms [4]. Research suggests that a tailored user interface (UI) design can significantly improve the overall UX [27]. However, enhancing the recommender system, optimizing the UI and UX, implementing effective marketing strategies, and providing high-quality content are advantageous when a platform can retain its customers. Recent research highlights the significance of monitoring customer retention rates [28], [29]. Utilizing the Light GBM model for machine learning, it has been established that certain variables significantly impact customer retention prediction. These variables include the total number of user logins within the previous week, the duration of video playback on the current day, and the time elapsed since the last login. By incorporating these factors in customer management strategies, businesses can optimize their customer acquisition and retention efforts [30].

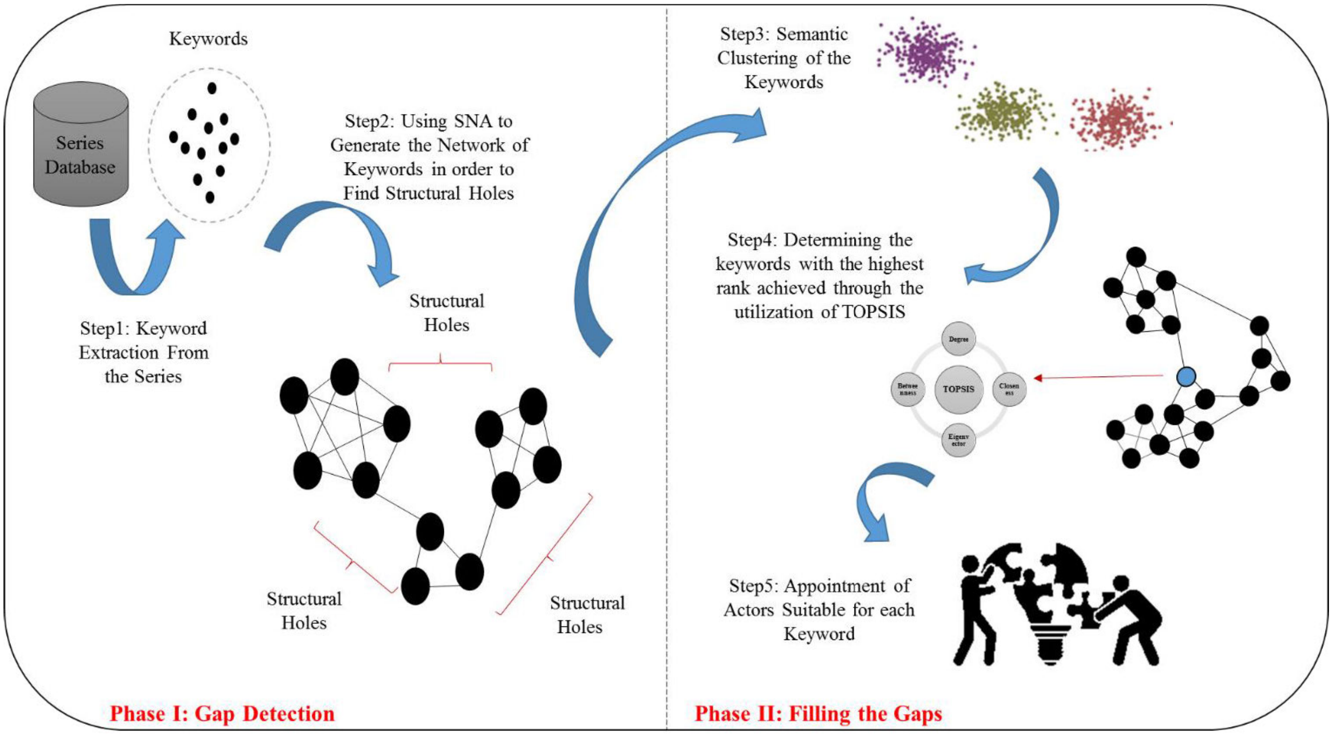

FIGURE 2. The proposed recommendation system mode.

Although machine learning methods have been widely used to analyze VOD platform datasets [31], [32], there has been a lack of attention given to other techniques, such as SNA. This is even though utilizing SNA enables researchers to obtain a comprehensive understanding of the relationships between various phenomena, assess audience preferences, explore the interplay between genres and actors, analyze information dissemination patterns to enhance digital marketing, and model and examine distribution velocity systematically and holistically [17], [33].

In one of the rare SNA-centric studies, an innovative approach was propounded to enhance UX on VOD platforms by leveraging users' implicit feedback, specifically their viewing percentages. The approach creates a video similarity network using SNA principles and modularity analysis to identify clusters of closely related videos. The study demonstrates that this method outperforms existing techniques in terms of accuracy and robustness [34]. As well as, another study focused on the marketing strategies adopted by two Chinese VOD platforms, VIU and iQiyi, and utilized SNA to analyze Twitter subscriber tweets. The study identified divergent marketing approaches between the two platforms, with VIU emphasizing content diversification and iQiyi prioritizing content quality. Furthermore, the study highlights the crucial role of user satisfaction in determining the success of these platforms [35].

To put it in a nutshell, based on the literature review, the research gap reveals that there is a lack of attention given to other techniques beyond machine learning despite their potential for providing valuable insights. Also, alongside cinema studies, a gap can be observed in research on the content of VOD platforms, particularly the storyline. To address this research gap, the current study uses the SNA approach to analyze the content-related keywords of series produced on Iran VOD platforms and provide a comprehensive understanding of this issue. Therefore, this research has the potential to contribute to both the academic literature and the practical implementation of data-driven content selection on VOD platforms.

III. METHODOLOGY

"Art is not a handicraft; it is the transmission of feeling the artist has experienced." This well-known quote by Tolstoy emphasizes the significance of delving into an artistic work's inner essence and content, rather than merely focusing on its superficial appearance. In the realm of literature, it is evident that researchers in the field of data science have not primarily focused on examining the content of artistic works.

As shown in Figure 2, the research method is based on presenting a recommender system that operates in two phases. In the first phase, we aim to establish the network of keywords by extracting relevant keywords from the series' plot. This network, which is generated using the SNA approach, will be assessed to detect structural holes and areas where the content may be lacking for the filmmakers. This strategy is critical for VOD platforms to remain competitive in the content market. In the second phase, we are looking to bring these peripheral keywords closer to the network's core so that creative content is formed, i.e., we aim to actively involve these keywords and incorporate them into the overall context. For this purpose, we first consider in which semantic clustering this keyword is located, that is, what other words it has semantic proximity with. After identifying the connectivity candidates of a keyword, we need to associate it with keywords that possess the highest centrality measures of degree, closeness, betweenness, and eigenvector. During this phase, we employ the TOPSIS method to select keywords that rank highest in all these four centrality measures. These selected keywords are considered strong acceptors for connection due to their high importance in the network. The unique feature of these centrality metrics allows the different parts of the network to be connected. In the final step, since the actors represent a movie and play the content, we seek to identify actors with the highest overlap with the target context.

A. DATA

Given this research gap, our main dataset was carefully selected from movie series produced by three prominent Iranian VOD platforms: Filimo, Namava, and Filmnet. According to information on imarketor.com, these platforms have invested in all series produced in the Iranian online film streaming industry, with Filimo holding a 65% market share, Namava with 25%, and Filmnet with 10%. However, approximately five years before the advent of VOD platforms in Iran, some pioneering producers initiated investment in movie series and distributed them on CDs. As the populace more favorably received these movie series than those produced by television, the VOD platforms emerged in this fledgling market. In this study, we also considered the information about these series. Furthermore, some of these series are currently available on these platforms.

To accomplish our research objective, we employed web crawling techniques in Python, specifically utilizing the BeautifulSoup library, to extract information about these series from their platforms. Additionally, the keywords of these series were extracted from the storylines available on Wikipedia using the same method. In conjunction with the Counter functionality from the collection's library, the re-library was employed to extract keywords from the storylines while excluding prepositions and pronouns. Indeed, the significance and frequency of words were considered when identifying keywords. Finally, the resulting keywords were organized with the help of the N-grams library, i.e., words with similar meanings were grouped, such as "love" and "romance," which both pertain to romantic relationships. Figure 3 shows the distribution of keywords among these three VOD platforms. Also, Additional information germane to these platforms is provided in TABLE 1.

To ensure a better comprehension of the subsequent steps, let's first clarify the terminology and algorithms of the SNA metrics before delving into the identification phase of structural holes.

B. NETWORK GENERATION

Upon completing the data-gathering step, we utilized SNA to construct a network of keywords to identify structural holes or areas in which the series has a dearth of content. As previously mentioned, SNA operates on the principles of graph theory. In this study, the graph G = (V, E) is utilized where V represents the number of keywords and E represents the number of edges or links that connect these keywords. Specifically, we established connections between keywords based on their presence within the same series, e.g., if a series contained five keywords, all of them were connected. Since the direction of the connection was deemed unimportant, we generated an undirected graph.

To depict the undirected graph, we utilized the adjacency matrix as input. As shown in Figure 4, the adjacency matrix displays the weight between every pair of nodes in a weight represents the frequency of edges drawn between two nodes.

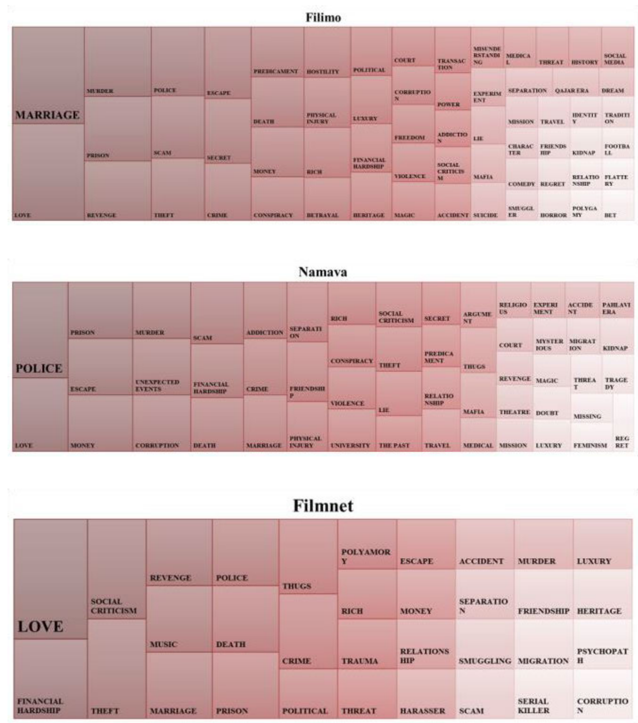

FIGURE 3. The distribution of keywords in the movie series created by Iranian VOD platforms.

TABLE 1. VOD platforms information.

VOD Platform	Starting Activity	Num. of Series	Most Frequent Genre	Most Frequent Keyword	Avg. IMDB Voters	Avg. IMDB Rate
Filimo	2015	36	Drama	Marriage	1682	5.6
Namava	2014	28	Drama	Police	2674	5.1
Filmnet	2015	13	Comedy	Love	1099	6.6
Total	-	76	Drama	Love	1617	4.8

Note(s): The number of series denotes the quantity of distinct series not affiliated with cable TV series generated or listed on the platform. Additionally, the remaining characteristics pertain to these particular series. Also, some series may repeat in more than one row, meaning they are distributed through multiple platforms.

In this study, the number of times two keywords appeared together in a series indicates the weight between those two nodes.

After obtaining the graph, it can be analyzed employing centrality measures. In SNA literature, the notion of node importance which is investigated through centrality measures, holds significant scholarly value. Numerous centrality measures have been propounded by researchers for this matter, but the four most crucial and commonly utilized ones are degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality [37]. However, the purpose of this study is not to identify the most significant node in the network using centrality measures. Instead, we aim to leverage the unique characteristics of these metrics; therefore, we want to identify those keywords that possess the highest level of these centrality measures.

To comprehend the idea of centrality measures, it is first necessary to define the concept of shortest path length. In a graph, there are numerous paths connecting two nodes. To incorporate the concept of a path in a network, the shortest path length is utilized, which represents the shortest distance between two nodes [36]. Therefore, we refer to the path as the shortest path length between two nodes. The four centrality measures in the SNA literature are defined below [17], [33], [36], [37], [38]:

- Degree Centrality: The node with the highest edge or link is the most important. This centrality measure is calculated according to (1) by dividing the number of edges of each node by the number of total nodes without considering that node itself. In Figure 1, node 3 possesses the highest degree of centrality since it possesses the most links.

$$C_i^{deg}(g) = \frac{d_i(g)}{n-1} \tag{1}$$

In this equation $d_i(g)$ is the edges of the node i and n is the total number of nodes.

- Closeness Centrality: The node that is closest to other nodes is the most important one. Equation (2) reveals that closeness centrality is achieved by calculating the distance of a specific node from all of the nodes in the network. The one that has the minimum sum of distance, is the most important node. In In Figure 1 node 3 possesses the highest closeness centrality as it is clearly closer to all nodes of the network.

$$C_i^{cls}(g) = \frac{1}{\sum_{j \neq i} \rho_g(i,j)} \tag{2}$$

In this equation $\rho_g(i,j)$ is the distance between the pair of nodes i and j in the network.

- Betweenness Centrality: According to the betweenness centrality definition, the node that connects more nodes is the most important. These nodes play the role of a bridge in a network and are placed more frequently than other nodes in the path that connects two nodes. In In Figure 1 Node 3 possesses the highest betweenness centrality as it acts as a bridge between non-redundant nodes, bringing them closer. Equation (3) reveals the calculation of betweenness centrality:

$$C_b(v_i) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{3}$$

In this formula, σ_{st} is the number of shortest paths between two nodes s and t while $\sigma_{st}(v)$ is the number of shortest paths between two nodes s and t passing through v.

- Eigenvector Centrality: In calculating this measure, a node's neighbors' importance is also considered. In other words, the more important a node is connected to, the more important it is. Equation (4) illustrates that this metric is calculated according to the linear algebra formula:

$$\lambda C_i = \sum\nolimits_j g_{ij} C_i \tag{4}$$

In this equation, λ is constant, positive, and is the eigenvalue of the adjacency matrix g_{ij}.

In this study, a node with a high value in all four centrality metrics means that not only is this node the key player of the network, but this node also has links to other keywords that are not directly connected. In other words, it serves as a bridge between different keywords, providing an opportunity for innovative utilization of hybrid keywords that blend genres. As well as it has also been proven that there is a positive correlation between the diversity of the genre or topic of a film and the level of public reception, as this breadth is responsible for satisfying different segments of the audience [17].

The Python NetworkX library has been used to construct the keywords network and calculate the centrality measures.

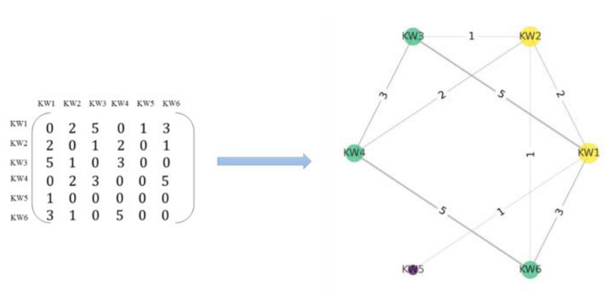

FIGURE 4. Adjacency matrix and generated graph. If a value greater than 0 is present in the adjacency matrix, it signifies a connection between two nodes. Conversely, if the value is 0, it indicates no connection between them. The strength of connections between nodes can be represented using any real number when generalized.

C. STRUCTURAL HOLES

The historical background of the structural holes concept is rooted in the work of sociologist Ronald S. Burt, whose research in the 1990s significantly popularized the idea. Burt's work emphasized the competitive advantages that accrue to entities or individuals occupying positions that allow them to bridge structural holes. Acting as

Input:
- A dataset represented as an undirected graph G

Output:
- List of nodes with structural holes, sorted by brokerage score

1. Initialize an empty dictionary called brokerage_scores.
2. For each node in the graph G:
   a. Create a set called neighbors and add all neighbors of the node to it.
   b. Create a subgraph by extracting the subgraph of G containing nodes in the 'neighbors' set.
   c. Calculate the number of external ties by counting the number of edges in the 'subgraph'.
   d. Calculate the number of internal ties by counting the number of edges in G connected to the current node.
   e. Calculate the brokerage score for the node using the formula:
      brokerage_score = external_ties – internal_ties
   f. Store the brokerage score in the 'brokerage_scores' dictionary with the node as the key.
3. Sort the nodes in 'brokerage_scores' in ascending order of their brokerage scores.
4. Initialize an empty list called structural_holes.
5. For each node in the sorted list of nodes:
   a. Append the node to the 'structural_holes' list.
6. Return the 'structural_holes' list, which contains the nodes with structural holes, sorted by brokerage score.
7. End of the algorithm.

FIGURE 5. Algorithm to calculate Structural Holes in an undirected graph.

intermediaries between disconnected groups, these individuals can facilitate the exchange of diverse information, exert control over the flow of information, and potentially broker valuable connections, illustrating the practical implications of understanding structural holes. In practice, the concept of structural holes finds broad applications such as innovation and information flow, entrepreneurship, and social capital [7], [39], [40].

In this paper, the structural holes are the gaps between clusters of keywords that are not directly connected. These gaps represent opportunities for a movie to differentiate itself from others by including a combination of keywords that bring innovation to movie storylines. We use the brokerage score to define structural holes, which is the subtraction between external and internal ties. According to Burt's definition, we establish an algorithm illustrated in Figure 5 in order to calculate the structural holes of the network of keywords [7], [40], [41].

In an undirected graph, internal ties have been considered as ties between the focal node and other nodes within the same community or cluster. Also, external ties can be defined as ties between the focal node and nodes in different communities or clusters. This adaptation helps preserve the spirit of Burt's theory in the context of undirected graphs.

Since brokerage scores indicate the ability of a node to bridge the different parts of the network together, the lower the brokerage score, the more the chance to find a structural hole.

D. SEMANTIC CLUSTERING

Semantic clustering refers to grouping words or concepts based on semantic similarity or relatedness. In other words, it involves identifying and organizing words or concepts that share similar meanings or themes into clusters or groups. This can be useful for various natural language processing tasks, such as text classification, topic modeling, and information retrieval [42].

In this research, Python and the Gensim and Scikit-learn libraries were utilized to cluster a curated list of 170 words. Leveraging pre-trained word embeddings from the FastText model, the script retrieved word vectors for the given words. It employed an Agglomerative Clustering algorithm from Scikit-learn to cluster these words. The goal was to explore semantic associations and similarities between the keywords, aiming to categorize them based on their underlying meanings. The clustering algorithm organized these words into distinct clusters based on their semantic relationships captured within the word embeddings. Various Python functionalities facilitated the data processing, clustering, and analysis, ultimately aiding in understanding the semantic structure and relationships within the provided list of words. Figure 6 represents the steps taken for semantic clustering.

E. TOPSIS

The Technique for Order Preference by Similarity to Ideal Solution, commonly referred to as TOPSIS, is regarded as a multi-criteria decision-making (MCDM) technique. MCDM techniques are employed to select the most suitable option among a set of available options, and their distinctive feature is that there are usually a countable number of pre-determined options. The best option will be one that provides the best value from each existing criterion [43].

The underlying logic of TOPSIS defines the positive and negative ideal solutions. The positive ideal solution is

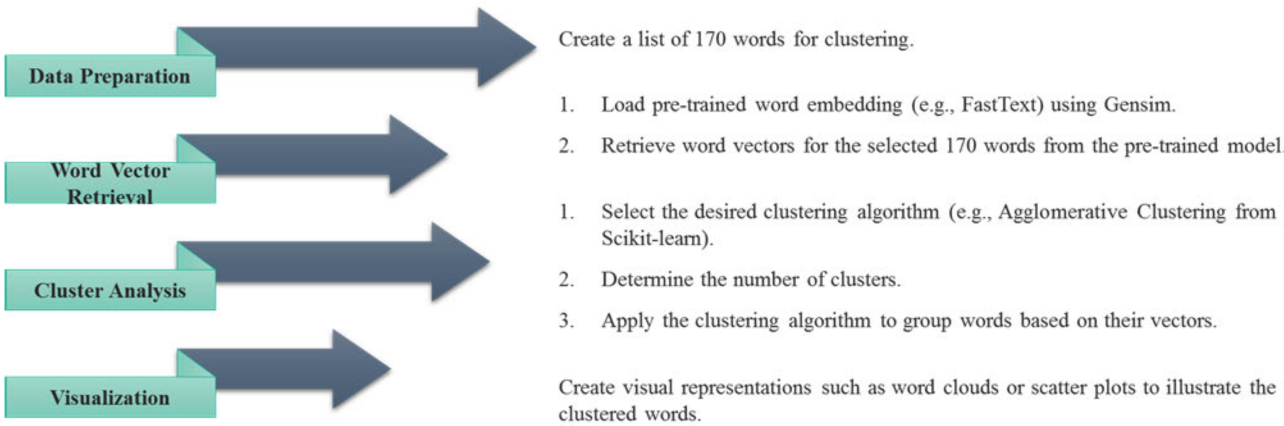

FIGURE 6. Semantic clustering steps.

a solution that increases the benefit criterion and decreases the cost criterion. The optimal choice is the one with the least distance from the positive ideal solution while simultaneously having the farthest distance from the negative ideal solution. In the TOPSIS method, options with the highest similarity with the ideal solution are ranked higher in the options ranking [44]. In Figure 7, the target space between the two criteria is shown. Here, A^+ and A^- represent the ideal positive and negative solutions, respectively. Option A_1 has a shorter distance to the ideal solution and a longer distance to the negative ideal solution compared to option A_2 which makes it more desirable.

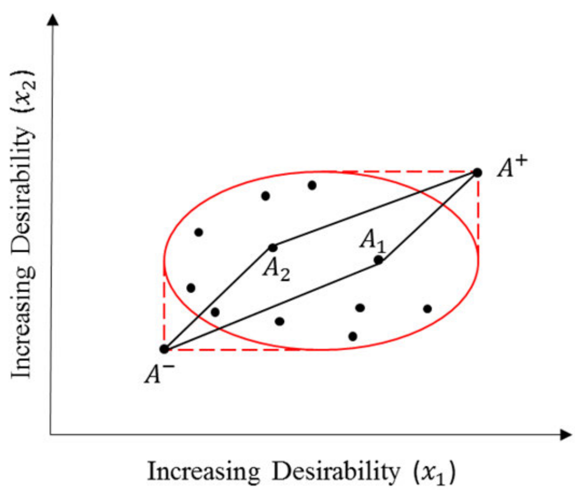

FIGURE 7. TOPSIS principle.

TOPSIS has been applied in numerous research studies that aim to identify the optimal choice from a set of criteria. SNA is one of the realms where TOPSIS has been utilized [38]. In this study, we intend to utilize this technique to identify the best option among four centrality measures: degree, closeness, betweenness, and eigenvector. Here, "option" refers to the number of keywords, and "criterion" refers to the four centrality measures. Since centrality is a positive criterion and we do not consider any specific differences, we assign equal weights to all so that the sum of weights should equal one.

The first step in TOPSIS is to form the decision matrix. The decision matrix evaluates several options based on a set of criteria. It is a matrix where each option is assigned scores based on the criteria. The decision matrix is denoted by X, and each ematrix elementis represented as x_{ij}. After Euclidean normalization, the weight of each criterion is multiplied by the same criterion in all matrix elements. Finally, the Euclidean distance between node i and the ideal positive and negative alternatives is performed based on the best and worst values for each centrality measure among all selected nodes, denoted as v_j^+ and v_j^- respectively according to the following equations:

$$d_i^+ = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^+)^2} \tag{5}$$

$$d_i^- = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^-)^2} \tag{6}$$

$$S_{TOPSIS} = \frac{d_i^-}{d_i^- + d_i^+} \tag{7}$$

The value of $\mathrm{S_{TOPSIS}}$ measure ranges between zero and one. The closer this value is to one, the closer the solution is to the ideal answer, indicating a better solution.

F. ALLOCATION OF PROPER ACTORS

In the literature of SNA, it is a term called bipartite graph. A bipartite graph is a graph that can be separated into two different sets of vertices in such a way that no two vertices within the same set are directly connected. This means that in a bipartite graph, all edges link vertices from one set to vertices in the other set. This characteristic gives bipartite graphs a unique structure and an easily identifiable pattern [45].

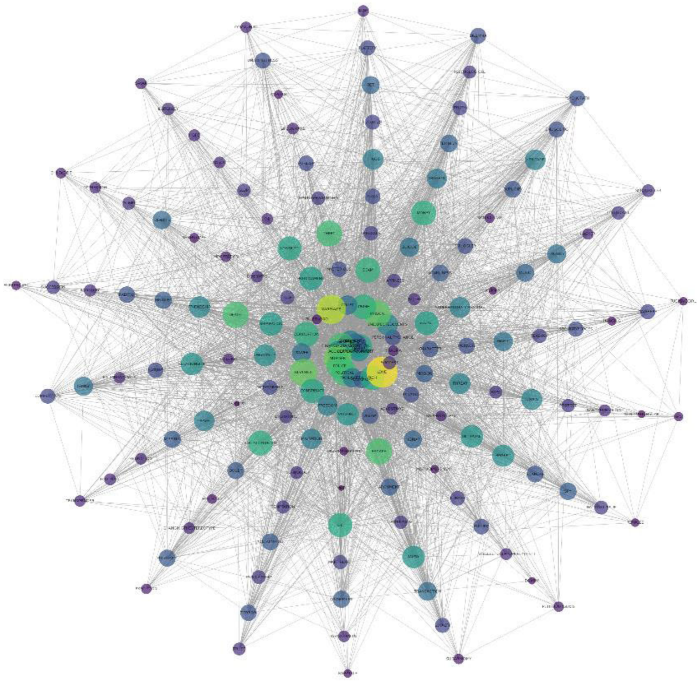

FIGURE 8. Network of keywords.

In this study, we can also consider a bipartite graph between actors and keywords. Since the film serves as the common ground between these two sets, meaning that each film consists of a series of actors and a series of keywords, we can construct a bipartite graph that connects actors to keywords based on their presence in a shared film. Finally, we introduce the actors who had the most links to a keyword as candidates for playing that role.

IV. RESULTS

As mentioned in the previous section we extracted the keywords of 76 movie series to generate a network. Figure 8 shows the graph of keywords with 170 nodes and 3586 edges. The graph in Figure 8 is an example of a core-periphery graph that demonstrates how the most frequent keywords are clustered in the center while others are dispersed in the periphery. This network also reveals the presence of structural holes, as some keywords are located far from the core, with only a few edges connecting them.

In the next step, based on structural holes algorithm, we calculated the brokerage score for all the keywords in the network. TABLE 2 reports this amount. Burt's structural holes theory suggests that individuals who bridge different network clusters have higher brokerage scores. So, individuals with lower brokerage scores are likely to be located in structural holes.

This table lists the brokerage score of 171 keywords obtained through the movie series aired on Iranian VOD platforms. These keywords have been extracted from the movie's plot lines. Furthermore, semantic clustering categorizes the keywords into 15 clusters, some of which include certain words in more than one cluster due to their semantic relatedness and establish semantic relationships with other keywords. Figure 9 illustrates the connections between the keywords and Figure 10 shows how the clusters are connected.

As depicted in Figure 9 Love, political issues, death, and money are among the keywords related to the other words. Keywords related to sports, supernatural events, and technology are not integrated into the giant part of the network. This suggests that when creating a movie based on these topics, the central theme of these movies cannot be linked to a keyword that can connect with other topics. For instance, ''magic'' can be associated with supernatural events and entertainment. This word serves as a bridge between the two topics.

Figure 10 depicts another form of Figure 9 and shows the relationships between the clusters. It is important to note that we did not limit the number of clusters. Rather, we developed a code to categorize all extracted keywords into the desired distinct clusters semantically. We then thoroughly reviewed the resulting clusters to ensure logical and meaningful categorization.

Figure 10 depicts 15 groups:

• Group 1- Emotional Relationships and Commitment: Love, Marriage, Relationship, Friendship, Trust, Partnership, Loyalty, Betrayal, Separation, Family, Polyamory, Polygamy, Amnesia.

• Group 2- Legal and Justice: Court, Crime, Police, Theft, Scam, Corruption, Spy, Bet, Identity, Transaction, Execution, Succession, Power, Identity, Success, Failure, Dignity.

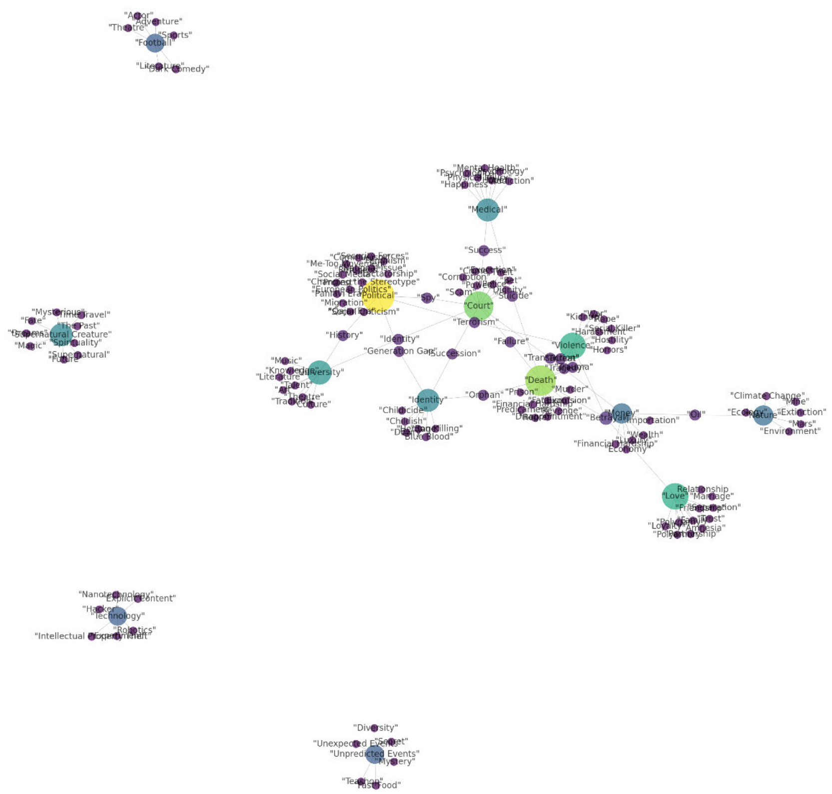

FIGURE 9. Semantic clustering of keywords.

FIGURE 10. The network of clusters.

- Group 3- Adversity and Misfortune: Death, Murder, Revenge, Prison, Financial Hardship, Predicament, Tragedy, Betrayal, Regret, Threat, Suicide, Orphan, Disappointment, Expulsion, Trauma, Failure, Fatalism.
- Group 4-Social Issues and Movements: Political, Social Criticism, Feminism, History, Migration, National Issue, Protest, Changing the Stereotype, Terrorism, Refugees, Generation Gap, Me-Too Movement.
- Group 5-Conflict and Violence: Violence, Hostility, Threat, War, Terrorism, Tragedy, Kidnap, Horrors, Betrayal, Serial Killer, Harassment, Rape, Tragedy, Trauma.
- Group 6: Supernatural and Mystical: Supernatural Creatures, Magic, Fate, Mystery, Time Travel, the Future, The Past, Dreams, and Spirituality.
- Group 7 Health and Well-being: Medical, Physical Injury, Addiction, Suicide, Psychology, Happiness, Psychological, Mental Health, Hope, Success.
- Group 8-Education and Culture: University, Literature, Tradition, Music, Theatre, Culture, Art, Knowledge, History, Talent, Identity.
- Group 9-Financial Aspects: Money, Financial Hardship, Wealth, Luxury, Economy, Transaction, Importation, Oil.
- Group 10-Technological and Scientific: Technology, Nanotechnology, Robotics, Experiment, Hacker, Intellectual Property Theft, Explicit Content.
- Group 11-Nature and Environment: Nature, Mine, Mars, Extinction, Climate Change, Environment, Oil.
- Group 12-Entertainment and Sports: Football, Theatre, Literature, Sports, Actor, Adventure, Dark Comedy.
- Group 13-Family and Social Status: Identity, Heritage, Orphan, Succession, Blue Blood, Generation Gap, Infanticide, Son-Killing, Dowry, Childish.

TABLE 2. Brokerage scores of movie series keywords.

Keyword	Brokerage Score	Keyword	Brokerage Score	Keyword	Brokerage Score	Keyword	Brokerage Score
War	2	Dogmatic	155	Loyalty	375	Tragedy	883
Dowry	14	Economy	159	Mysterious	382	Travel	924
Generations Gap	20	Technology	168	Security Forces	385	Heritage	927
Fairytale	28	Intellectual Property Theft	172	Unworthiness	385	Thugs	941
Mummy	29	Nanotechnology	172	Smuggler	391	Luxury	995
Fugitive Girl	35	Talent	172	Dilemma	395	Freedom	1001
Housing	35	Trust	191	Failure	400	Accident	1022
Refugee	35	Illiteracy	195	Religious	413	Threat	1070
Disappointment	36	Psychological	204	Tradition	418	Unexpected Events	1079
Me-Too Movement	37	Rape	204	Duplicity	432	Friendship	1114
Isis	44	Serial Killer	206	Horror	439	University	1162
Orphan	45	Success	208	Character	440	Physical Injury	1175
Fajr Festival	55	Happiness	209	Terror	441	Betrayal	1229
Blueblood	65	Protest	209	Psychopath	451	Regret	1229
Fast Food	65	Literature	211	Polygamy	459	Violence	1298
Shahnameh	65	Temptation	212	Missing	460	Relationship	1321
Childish	76	Slum	213	Medical	461	Conspiracy	1346
Son-Killing	77	Europe	214	History	463	Predicament	1349
Trauma	80	Expulsion	215	The Past	468	Mafia	1371
Petit Bourgeois	90	Time Travel	232	Loneliness	480	Hostility	1374
Sugar Mommy	90	Amnesia	233	Scandal	502	Secret	1435
Execution	91	Misunderstanding	236	Music	504	Lie	1438
Supernatural	93	National Issue	237	Mission	507	Separation	1439
Hope	99	Help	246	Doubt	530	Addiction	1480
Chalcidice	104	Succession	254	Spy	531	Rich	1519
Importation	104	Partnership	262	Polyamory	540	Social Criticism	1550
Transgender	119	Competition	276	Bet	542	Crime	1644
Hacker	127	Martial Art	282	Feminism	547	Political	1668
Adventure	128	Dictatorship	283	Argument	549	Money	1688
Oil	130	Teashop	285	Atonement	561	Corruption	1706
Explicit Content	131	Nature	295	Identity	619	Scam	1745
Football	132	Truth	303	Kidnap	632	Financial Hardship	1803
Communist	135	Futurism	314	Family	653	Theft	1841
Mine	135	Harasser	315	Magic	705	Police	1916
Savak	135	Social Media	315	Transaction	741	Death	1949
Actor	143	Qajar Era	316	Migration	744	Escape	2028
Changing the Stereotype	152	Smuggling	327	Pahlavi Era	795	Revenge	2096
Dark Comedy	152	Personality Change	338	Dignity	802	Prison	2169
Extinction	152	Theatre	345	Court	809	Murder	2185
Life on Mars	152	Dream	354	Power	835	Marriage	2402
Malignant Virus	152	News	359	Super Natural Creature	841	Love	2800
Oppression	152	Flattery	367	Suicide	863		
Robot	152	Fatalism	369	Experiment	879		

- Group 14-Government and Politics: Political, Dictatorship, Qajar Era, European Politics, National Issue, Social Media, Spy, Security Forces, Communism, Pahlavi Era.
- Group 15-Miscellaneous and Diverse: Secret, Mystery, Unexpected Events, Diversity, Teashop, Fast Food

Figure 10 illustrates the relationships between semantic clusters. The concepts of crime, emotions, and social issues, representing the thematic triangle in Iranian movie series, act as a bridge between other semantic clusters at the center of the network. This monopolization of content, which has become a rule in Iranian movie series, hinders the growth of creativity in exploring other themes.

In the next phase, we calculated the centrality measures of the keywords obtained. The report of the results for degree, closeness, betweenness, and eigenvector centrality measures can be found in TABLE 3, TABLE 4, TABLE 5, and TABLE 6 respectively.

These tables infer that emotional subjects such as love and marriage and criminal topics such as murder have the highest importance among keywords. As it is evident, there is no significant difference between the centrality measures, and their results are close to each other. This closeness of results confirms the assumption of equal weighting of centrality measures in TOPSIS. Therefore, TOPSIS was employed

TABLE 3. Degree centrality of movie series keywords.

Keyword	Degree Centrality	Keyword	Degree Centrality	Keyword	Degree Centrality	Keyword	Degree Centrality
Love	140	Tragedy	57	Mysterious	35	Nanotechnology	20
Marriage	125	Pahlavi era	55	News	34	Dogmatic	20
Murder	110	Magic	54	Security forces	34	Hope	20
Revenge	108	Court	54	Fatalism	34	Oil	19
Prison	106	Suicide	54	Nature	34	Life on Mars	19
Escape	100	Power	54	Flattery	33	Robot	19
Financial hardship	99	Family	53	Qajar era	32	Malignant virus	19
Death	99	Dignity	52	Futurism	32	Extinction	19
Police	98	Transaction	51	Dictatorship	32	Changing the stereotype	19
Political	97	Identity	49	Competition	32	Dark comedy	19
Theft	97	Kidnap	47	Loyalty	32	Oppression	19
Scam	92	Music	46	Truth	31	Hacker	18
Crime	91	Bet	46	Europe	30	Transgender	18
Corruption	90	Feminism	45	Harasser	30	Mine	18
Money	90	History	43	Teashop	29	Communist	18
Social criticism	89	Argument	43	Martial art	29	Savak	18
Rich	87	Mission	43	Help	28	Childish	17
Lie	86	Spy	43	Misunderstanding	27	Importation	16
Addiction	84	Missing	42	National issue	27	Supernatural	16
Conspiracy	84	Polyamory	42	Happiness	27	Trauma	16
Hostility	83	Atonement	42	Partnership	27	Infanticide	16
Secret	81	Social media	41	Amnesia	26	Sugar mommy	15
Predicament	80	Polygamy	41	Time travel	26	Petit bourgeois	15
Separation	78	Loneliness	41	Literature	26	Execution	15
Betrayal	74	Doubt	41	Slum	26	Son-Killing	14
Violence	73	Medical	40	Succession	26	Shahnameh	13
Mafia	73	Scandal	40	Temptation	25	Fast food	13
Relationship	72	Duplicity	40	Expulsion	25	Blueblood	13
Regret	70	Tradition	39	Protest	25	Fajr festival	12
University	70	Dream	39	Psychological	24	Isis	12
Physical injury	68	The past	38	Success	24	Housing	11
Unexpected events	66	Character	38	Talent	24	Orphan	11
Friendship	66	Terror	38	Rape	24	Disappointment	10
Accident	64	Horror	37	Trust	23	Fairytale	10
Luxury	64	Smuggling	37	Serial killer	23	Me-Too movement	10
Threat	64	Psychopath	37	Technology	23	Refugee	10
Supernatural creature	62	Religious	36	Illiteracy	23	Fugitive girl	10
Experiment	61	Theatre	36	Economy	22	Mummy	9
Freedom	61	Smuggler	36	Explicit content	21	Generation gap	8
Heritage	61	Dilemma	36	Actor	21	Dowry	7
Travel	61	Unworthiness	36	Football	20	War	4
Thugs	61	Personality change	35	Adventure	20		
Migration	60	Failure	35	Intellectual property theft	20		

to rank the keywords to identify the keywords with the most desirable positions based on four centrality measures. TABLE 7 illustrates the ranking of keywords based on their importance. The results of TABLE 7 negatively correlated with the identification of structural holes in TABLE 2. This finding indicates that our method has correctly identified the structural holes. By definition, structural holes are mainly found around the network periphery, which means the locations where nodes have lower centrality.

The final question in the model is who should play the role of the linking keywords. Here, we have suggested a bipartite graph of actor-films, where we recommend the actor with the highest degree as a suitable option for each keyword. Since the number of actors and keywords is large, in Table 8, we only provide recommendations for the top 15 keywords with the highest actor centrality.

The results of TABLE 7 negatively correlated with the identification of structural holes in TABLE 2 This finding indicates that our method has correctly identified the structural holes. By definition, structural holes are mainly found around the network periphery, which means the locations where nodes have lower centrality.

TABLE 4. Closeness centrality of movie series keywords. The values have been rounded to the nearest thousandth.

Keyword	Closeness Centrality	Keyword	Closeness Centrality	Keyword	Closeness Centrality	Keyword	Closeness Centrality
Love	0.85	Tragedy	0.601	Horror	0.554	Adventure	0.523
Marriage	0.791	Pahlavi era	0.596	Personality change	0.554	Importation	0.52
Murder	0.739	Magic	0.594	Mysterious	0.554	Trauma	0.52
Revenge	0.733	Court	0.594	Flattery	0.554	Explicit content	0.518
Prison	0.726	Family	0.592	News	0.552	Mine	0.518
Financial hardship	0.705	Suicide	0.59	Fatalism	0.552	Communist	0.518
Death	0.705	Power	0.59	Qajar era	0.552	Savak	0.518
Escape	0.702	Dignity	0.59	Futurism	0.552	Supernatural	0.518
Police	0.702	Transaction	0.584	Dictatorship	0.552	Infanticide	0.518
Political	0.7	Identity	0.584	Competition	0.552	Trust	0.517
Theft	0.7	Music	0.578	Loyalty	0.548	Life on Mars	0.517
Crime	0.683	Bet	0.578	Europe	0.548	Robot	0.517
Scam	0.68	Kidnap	0.576	Truth	0.547	Malignant virus	0.517
Corruption	0.68	Feminism	0.576	Martial art	0.547	Extinction	0.517
Money	0.68	History	0.572	Harasser	0.545	Changing the stereotype	0.517
Social criticism	0.677	Missing	0.57	Teashop	0.543	Dark comedy	0.517
Rich	0.672	Atonement	0.57	Misunderstanding	0.543	Oppression	0.517
Lie	0.669	Argument	0.569	Happiness	0.543	Childish	0.515
Addiction	0.664	Mission	0.569	Help	0.541	Blue blood	0.515
Conspiracy	0.664	Spy	0.569	Slum	0.541	Housing	0.514
Hostility	0.661	Social media	0.569	Succession	0.541	Intellectual property theft	0.512
Secret	0.656	Polygamy	0.569	Partnership	0.54	Nanotechnology	0.512
Predicament	0.649	Loneliness	0.569	Protest	0.54	Hacker	0.512
Separation	0.644	Polyamory	0.567	National issue	0.538	Sugar mommy	0.509
Betrayal	0.639	Scandal	0.567	Amnesia	0.538	Petit bourgeois	0.509
Violence	0.637	Duplicity	0.567	Time travel	0.538	Isis	0.509
Mafia	0.637	Doubt	0.565	Literature	0.538	Shahnameh	0.507
Relationship	0.634	Tradition	0.565	Temptation	0.536	Fast food	0.507
Regret	0.63	Medical	0.563	Expulsion	0.536	Me-Too movement	0.504
University	0.63	Terror	0.563	Serial killer	0.536	Son-Killing	0.503
Physical injury	0.625	Dream	0.561	Psychological	0.535	Fairy tale	0.503
Unexpected events	0.62	Psychopath	0.561	Success	0.535	Dowry	0.503
Friendship	0.62	The past	0.559	Talent	0.535	Execution	0.5
Accident	0.616	Character	0.559	Illiteracy	0.535	Mummy	0.5
Luxury	0.616	Religious	0.559	Economy	0.535	Generation gap	0.5
Threat	0.612	Theatre	0.559	Rape	0.53	Refugee	0.497
Supernatural creature	0.612	Smuggler	0.559	Actor	0.53	Fugitive girl	0.497
Experiment	0.609	Dilemma	0.559	Hope	0.53	Fajr festival	0.493
Heritage	0.609	Unworthiness	0.559	Football	0.528	Orphan	0.493
Travel	0.609	Failure	0.557	Dogmatic	0.528	Disappointment	0.493
Thugs	0.609	Smuggling	0.556	Transgender	0.528	War	0.42
Migration	0.607	Security forces	0.556	Oil	0.525		
Freedom	0.605	Nature	0.556	Technology	0.523		

The final question in the model is who should play the role of the linking keywords. Here, we have suggested a bipartite graph of actor-films, where we recommend the actor with the highest degree as a suitable option for each keyword. Since the number of actors and keywords is large, in Table 8, we only provide recommendations for the top 15 keywords with the highest actor centrality.

The obtained results are acceptable for the Iranian audience. An interesting point to note is that the work history of these actors in cinema or other works outside of this dataset has also influenced these results. For example, Mehdi Hosseini Nia has appeared in cinema in roles related to criminals and revenge, reflected in the highest degree for the revenge keyword in the table. Siamak Ansari is a well-known figure in films with a social critique theme, and Mehran Modiri is an actor-director who has extensively addressed the issue of corruption in his works. Additionally, playing criminal and police roles has made Amir Aghaei a recognizable figure in this context. Yekta Naser recently focused on prestigious roles and has achieved the highest score for portraying a glamorous actress. In the following section, we will scrutinize and assess the results acquired.

TABLE 5. Betweenness centrality of movie series keywords. The values have been rounded to the nearest thousandth.

Keyword	Betweenness Centrality	Keyword	Betweenness Centrality	Keyword	Betweenness Centrality	Keyword	Betweenness Centrality
Love	967.082	History	66.613	Futurism	15.481	Trust	2.189
Marriage	746.031	Thugs	65.46	Dilemma	14.289	Trauma	1.473
Social Criticism	478.066	Music	61.069	Atonement	14.196	Serial Killer	1.349
Murder	430.038	Threat	59.783	Qajar Era	13.341	Isis	1.257
Revenge	378.746	Freedom	54.962	Fatalism	12.958	Dogmatic	0.81
Financial Hardship	376.829	Identity	51.554	Horror	12.599	Transgender	0.79
Political	344.706	Smuggling	50.326	The Past	12.503	Housing	0.781
Prison	295.167	Pahlavi Era	49.825	Psychopath	12.373	Supernatural	0.684
Escape	283.144	Dream	49.209	Mysterious	12.136	Fairytale	0.657
Crime	273.684	Bet	46.214	Religious	11.622	Hacker	0.438
Theft	267.689	Tragedy	46.162	News	10.389	Intellectual Property Theft	0
Lie	266.787	Missing	45.625	Flattery	9.192	Nanotechnology	0
Police	250.167	Technology	44.993	Failure	9.155	Life on Mars	0
Conspiracy	244.197	Court	43.585	Truth	9.051	Robot	0
Death	239.951	Feminism	38.776	Help	8.852	Malignant Virus	0
Rich	231.648	Transaction	38.601	Hope	8.409	Extinction	0
Corruption	219.979	Power	38.268	Misunderstanding	7.605	Changing the Stereotype	0
Scam	218.209	Mission	36.68	Literature	7.403	Dark Comedy	0
Money	212.172	Kidnap	36.184	Security Forces	7.321	Oppression	0
Hostility	201.641	Dignity	36.005	Slum	6.769	Mine	0
Addiction	190.084	Nature	35.658	Teashop	6.547	Communist	0
Secret	167.853	Childish	33.532	Talent	6.325	Savak	0
Migration	163.039	Suicide	32.115	Time Travel	5.972	Importation	0
Predicament	161.851	Duplicity	30.514	Martial Art	5.896	Childcide	0
Betrayal	146.262	Tradition	30.137	National Issue	5.573	Sugar Mommy	0
Supernatural Creature	137.266	Loneliness	29.266	Protest	5.564	Petit Bourgeois	0
Separation	110.87	Polygamy	27.628	Temptation	5.525	Execution	0
University	108.866	Europe	27.098	Adventure	5.379	Son-Killing	0
Experiment	100.356	Dictatorship	26.957	Amnesia	5.242	Shahnameh	0
Travel	98.456	Personality Change	24.553	Harasser	4.737	Fast Food	0
Regret	94.858	Spy	23.451	Loyalty	4.305	Blueblood	0
Violence	90.782	Argument	22.6	Explicit Content	3.965	Fajr Festival	0
Luxury	89.134	Medical	22.42	Rape	3.927	Orphan	0
Friendship	88.386	Unworthiness	22.074	Actor	3.691	Disappointment	0
Family	87.777	Theatre	21.779	Partnership	3.565	Me-Too Movement	0
Relationship	85.14	Competition	20.512	Economy	3.291	Refugee	0
Unexpected Events	84.235	Scandal	19.945	Succession	3.16	Fugitive Girl	0
Social Media	84.21	Character	18.92	Expulsion	2.958	Mummy	0
Magic	83.821	Smuggler	18.141	Illiteracy	2.879	Generations Gap	0
Heritage	78.74	Polyamory	17.223	Oil	2.533	Dowry	0
Mafia	77.766	Doubt	17.075	Success	2.303	War	0
Physical Injury	74.102	Terror	16.229	Football	2.269		
Accident	72.584	Happiness	16.014	Psychological	2.215		

V. DISCUSSION

Artistic work should mirror social problems and society's concerns, especially in cinema. Art has the power to reflect and comment on the issues and challenges faced by society, allowing for a deeper understanding and dialogue [46]. By investigating the Iranian series, it is obvious that general topics such as love, marriage, and murder, which can be characteristics of every society, are bold. However, domestic problems, such as the involvement and threat of war, generation gaps that have divided society into modern and traditional lifestyles (common in developing countries), and dowry problems, are not the focus of these movies. According to Islamic Republic News Agency (IRNA), nearly 20 percent of the 12,008 prisoners who have committed non-intentional crimes are captured due to dowry-related issues. Dowry is an amount of money that men are required to pay for their wives

TABLE 6. Eigenvector centrality of movie series keywords. The values have been rounded to the nearest thousandth.

Keyword	Eigenvector Centrality	Keyword	Eigenvector Centrality	Keyword	Eigenvector Centrality	Keyword	Eigenvector Centrality
Love	1	Suicide	0.523	Harasser	0.322	Explicit content	0.194
Marriage	0.916	Supernatural creature	0.518	Dream	0.321	Actor	0.189
Murder	0.87	Court	0.517	Truth	0.318	Illiteracy	0.187
Prison	0.867	Dignity	0.505	Theatre	0.316	Importation	0.179
Revenge	0.853	Pahlavi era	0.499	Flattery	0.315	Mine	0.179
Escape	0.828	Transaction	0.494	Smuggling	0.303	Communist	0.179
Death	0.819	Migration	0.476	Social media	0.297	Savak	0.179
Police	0.813	Magic	0.47	Teashop	0.296	Execution	0.177
Theft	0.798	Family	0.458	Martial art	0.295	Adventure	0.175
Financial hardship	0.785	Kidnap	0.447	Unworthiness	0.294	Supernatural	0.173
Scam	0.761	Identity	0.43	Partnership	0.286	Hope	0.173
Money	0.758	Atonement	0.424	Futurism	0.282	Blue blood	0.168
Corruption	0.756	Argument	0.42	Misunderstanding	0.28	Technology	0.164
Political	0.747	Doubt	0.411	Amnesia	0.279	Chalcidice	0.154
Crime	0.746	Spy	0.409	National issue	0.279	Trauma	0.152
Social criticism	0.723	Polyamory	0.409	Help	0.274	Dark comedy	0.15
Rich	0.718	Bet	0.408	Competition	0.271	Oppression	0.15
Addiction	0.704	Feminism	0.405	Dictatorship	0.269	Life on Mars	0.149
Separation	0.695	Scandal	0.4	Qajar era	0.266	Robot	0.149
Secret	0.688	Loneliness	0.396	Literature	0.265	Malignant virus	0.149
Lie	0.687	Music	0.396	Serial killer	0.262	Extinction	0.149
Mafia	0.678	Mission	0.387	Expulsion	0.261	Changing the stereotype	0.149
Hostility	0.675	The past	0.385	Success	0.258	Son-Killing	0.134
Predicament	0.669	Missing	0.381	Rape	0.256	Fajr festival	0.13
Relationship	0.667	Psychopath	0.378	Temptation	0.253	Sugar mommy	0.127
Conspiracy	0.666	Medical	0.375	Nature	0.251	Petit bourgeois	0.127
Violence	0.651	Character	0.368	Europe	0.249	Housing	0.118
Regret	0.645	Horror	0.367	Psychological	0.249	Me-Too movement	0.112
Betrayal	0.638	Religious	0.359	Dogmatic	0.236	Childish	0.11
Physical injury	0.629	Terror	0.356	Slum	0.232	Orphan	0.104
University	0.619	Failure	0.353	Trust	0.231	Mummy	0.104
Friendship	0.603	Security forces	0.348	Protest	0.23	Shahnameh	0.094
Unexpected events	0.601	Smuggler	0.346	Succession	0.225	Fast food	0.094
Threat	0.6	Tradition	0.345	Economy	0.224	Fairytale	0.093
Accident	0.58	Polygamy	0.345	Football	0.219	Disappointment	0.091
Freedom	0.575	Loyalty	0.345	Happiness	0.215	Isis	0.089
Luxury	0.559	Dilemma	0.345	Time travel	0.214	Dowry	0.089
Heritage	0.556	Mysterious	0.341	Talent	0.209	Generation gap	0.083
Thugs	0.551	History	0.337	Intellectual property theft	0.205	Refugee	0.064
Travel	0.548	News	0.337	Nanotechnology	0.205	Fugitive girl	0.064
Tragedy	0.531	Duplicity	0.333	Oil	0.204	War	0.003
Experiment	0.526	Fatalism	0.332	Hacker	0.203		
Power	0.525	Personality change	0.322	Transgender	0.198		

according to Islamic law in return for their marriage proposals. While many brides ignore this money, some force their husbands to pay a large sum, which can become a social problem. When husbands are unable to pay, they may face arrest. This issue has caused many divorces in the country during the last decade. A part of the dissatisfaction in Iran with the government stems from unworthiness, in which a group of petty bourgeoisie individuals, due to their connections with high-ranking politicians, have obtained positions of power. Many of their children, referred to as ''blue-blooded'' in Iran, have taken on important political and social positions.

In addition to political and social issues, other narrative themes can be a successful experience for movie series. One of these themes is serial killers and psychopathic criminals, which have great excitement for audiences around the world. Furthermore, the successful experience of the series ''Game of Thrones'' shows that people are interested in historical legends. Strangely, with the availability of powerful

TABLE 7. TOPSIS results.

Keyword	S_{TOPSIS}	Keyword	S_{TOPSIS}	Keyword	S_{TOPSIS}	Keyword	S_{TOPSIS}
Love	1	Tragedy	0.183	Theatre	0.105	Actor	0.052
Marriage	0.792	Magic	0.182	News	0.104	Oil	0.052
Social criticism	0.526	Family	0.181	Fatalism	0.103	Hacker	0.05
Murder	0.518	Pahlavi era	0.176	Loyalty	0.102	Transgender	0.049
Revenge	0.477	Court	0.175	Unworthiness	0.102	Hope	0.048
Financial hardship	0.461	Power	0.175	Flattery	0.098	Adventure	0.047
Political	0.432	Suicide	0.173	Truth	0.095	Mine	0.044
Prison	0.418	Dignity	0.168	Nature	0.095	Communist	0.044
Escape	0.4	Transaction	0.165	Harasser	0.094	Savak	0.044
Theft	0.384	Identity	0.155	Dictatorship	0.091	Childish	0.043
Crime	0.376	Kidnap	0.15	Futurism	0.091	Importation	0.042
Police	0.376	Music	0.148	Competition	0.09	Dark comedy	0.04
Death	0.371	Bet	0.145	Qajar era	0.088	Oppression	0.04
Lie	0.359	Feminism	0.141	Teashop	0.087	Supernatural	0.04
Scam	0.343	History	0.136	Martial art	0.087	Life on mars	0.04
Corruption	0.342	Argument	0.136	Europe	0.084	Robot	0.04
Rich	0.341	Social media	0.136	Partnership	0.082	Malignant virus	0.04
Conspiracy	0.339	Spy	0.134	Help	0.082	Extinction	0.04
Money	0.337	Atonement	0.134	Misunderstanding	0.081	Changing the stereotype	0.04
Hostility	0.313	Missing	0.134	National issue	0.081	Execution	0.039
Addiction	0.311	Mission	0.134	Amnesia	0.079	Infanticide	0.036
Secret	0.293	Polyamory	0.131	Literature	0.076	Trauma	0.035
Predicament	0.285	Doubt	0.13	Expulsion	0.074	Blue blood	0.035
Betrayal	0.265	Loneliness	0.13	Temptation	0.072	Sugar mommy	0.028
Separation	0.262	Scandal	0.127	Success	0.071	Petit bourgeois	0.028
Violence	0.24	Medical	0.123	Rape	0.071	Son-Killing	0.028
Mafia	0.239	Polygamy	0.12	Serial killer	0.071	Fajr festival	0.024
Relationship	0.239	The past	0.12	Happiness	0.07	Housing	0.02
University	0.239	Dream	0.12	Slum	0.07	Shahnameh	0.018
Regret	0.237	Tradition	0.118	Psychological	0.069	Fast food	0.018
Migration	0.234	Character	0.118	Succession	0.068	Me-Too movement	0.016
Supernatural creature	0.227	Duplicity	0.117	Protest	0.067	Orphan	0.016
Physical injury	0.224	Psychopath	0.117	Technology	0.066	Isis	0.015
Friendship	0.222	Terror	0.115	Time travel	0.066	Mummy	0.013
Unexpected events	0.22	Horror	0.115	Trust	0.064	Fairytale	0.012
Luxury	0.213	Smuggling	0.114	Talent	0.062	Disappointment	0.011
Travel	0.211	Religious	0.112	Economy	0.061	Refugee	0.008
Accident	0.21	Smuggler	0.11	Dogmatic	0.061	Fugitive girl	0.008
Experiment	0.208	Dilemma	0.109	Football	0.057	Dowry	0.008
Threat	0.208	Failure	0.109	Illiteracy	0.055	Generation gap	0.007
Heritage	0.204	Mysterious	0.107	Intellectual property theft	0.053	War	0
Freedom	0.199	Security forces	0.106	Nanotechnology	0.053		
Thugs	0.198	Personality change	0.105	Explicit content	0.053		

storytelling literature such as the Shahnameh in Iranian literature, no series has been made or focused on this subject.

According to the methodology, if we want to bring one of the peripheral keywords closer to the network's core to create innovative content, we should look for other keywords semantically categorized within the same cluster as the considered keyword. Then, we choose a keyword from among them with a high TOPSIS score. For example, if a director wants to make a film about the generation gap, a marginal topic in the series but a fundamental issue in Iranian society, he or she can use keywords from the same group in the ‘‘Social Issues and Movements’’. In this category, social criticism has the highest TOPSIS rank and can work as a linkage connecting the generation gap to other keywords. As well as, Siamak Ansari is the proposed actor who can play the main role of social criticism.

Semantic clustering also helps directors to identify other keywords in the vicinity of a keyword. For example, alongside the generation gap, other words such as migration, changing the stereotype, feminism, and national issues are also observed. All of these keywords reflect problems that Iranians and all developing countries have struggled with, and a significant portion of these problems has roots in the movement of society from a traditional to a modern style.

Another interesting point that can be examined in the analysis of the series' storyline keywords is investors' approach. In cinema, similar to the business world, there are two types of investors: conservative investors who follow conventional practices, and risk-taking investors who venture into new experiences. In this context, new and creative experiences can involve combining keywords from different categories that have the least connection in the obtained graphs. It can

TABLE 8. The result of the bipartite graph for keywords with the highest value of TOPSIS.

Keyword	Actor(s)	Degree
Love	Hooman Barghnavard, Masoud Rayegan, Setareh Eskandari, Hossein Yari, Abbas Jamshidifar	4
Marriage	Siamak Ansari, Hadi Kazemi	5
Social criticism	Siamak Ansari	6
Murder	Hanieh Tavasoli	5
Revenge	Mehdi Hosseini Nia	4
Financial hardship	Sahar Ghoreishi, Alireza Khamseh, Giti Ghasemi	3
Political	Ghazal Shakeri	3
Prison	Hanieh Tavasoli, Hossein Yari	4
Escape	Behnam Tashakor	4
Theft	Shabnam Moghaddami	4
Crime	Hanieh Tavasoli, Amir Aghaei	3
Police	Hanieh Tavasoli, Amir Aghaei, Ahmad Mehranfar	4
Death	Kazem Sayahi	4
Corruption	Mehran Modiri, Nima Shabannejad	3
Rich	Yekta Naser	3

also involve using an actor in a different role than they usually play, such as casting a comedian in a tragic role. One of the well-known figures in this field is Mehran Modiri, who uses actors from different genres and integrates various themes. According to our results, there is no direct connection between sports and legal issues. However, Modiri has made a film in his career that addresses financial corruption in football. This can provide valuable insights and attract a wider range of audiences. Among the series directed by Mehran Modiri that have been showcased on the Iranian VOD platforms and collected in our dataset, we found 6 films. The IMDB ratings of all these six series are higher than the average ratings in TABLE 1. "Dracula" with a rating of 6.1, "Just Kidding" with a rating of 6.9, "Bitter Coffee" with a rating of 7.4, "Mozaffar's Treasure" with a rating of 6.1, "My Villa" with a rating of 6.4 and "The Monster" with a rating of 7 are all considered successful experiences in series production in Iran. These results show that paying attention to the content significantly affects the success of movies.

Certainly, content is not everything and the actors hold significance in portraying a production. In this study, we proposed a methodology to recommend actors based on a bipartite graph. Another approach to actor selection is to evaluate his or her popularity in a reality show and assess the feedback from the viewership. Reality shows are a type of TV programming that portrays real situations and events with real people, without a script. They usually track individuals' or groups' daily lives, challenges, competitions, or social interactions. The main purpose of reality shows is to offer entertainment, drama, and a window into ordinary people's and celebrities' lives. Abbas Jamshidifar is an exemplary actor who, after achieving success in a comedy reality show called "Joker" and receiving abundant positive feedback, expanded his presence in seven other reality shows or series on the VOD platforms. Before this, he had been active on television, and the producers of the VOD platforms did not show much interest in him.

Also, since reality shows are less expensive than series and are usually made in one episode for each person, they can be used to invest in people who are not actors based on the audience's feedback.

VI. CONCLUSION

In Fihi Ma Fihi, Rumi believes that the quality of the container is less important than the substance it contains. He continues that if a precious golden chalice is filled with something unpleasant like vinegar, it will still be unpleasant to consume. Conversely, if a simple, ordinary container like a clay cup is filled with a delightful wine, it will still bring joy and pleasure. Therefore, the focus should be on the inherent value and nature of what is contained; in other words, the content; Thus, this research endeavors to emphasize the content alongside technical considerations in creating television series and audience engagement. Ultimately, what remains in people's minds is what they have learned from the content of films, as techniques naturally improve over time.

Based on the identified research gap, it is evident that a significant portion of the existing literature focuses on analyzing VOD platforms to develop recommender systems. These systems pay more attention to technical topics and do not consider the importance of content in the competition. This lack of attention is limited to VOD platforms and observed in other media. This paper aimed to emphasize the significance of topics and content within the realm of recommender systems. So, in this study, we introduced a two-phase recommender system. In the first phase, uncovered content gaps were identified, and in the second phase, a solution was proposed to fill these gaps based on centrality measures in the SNA.

This study introduces a novel approach using the SNA technique to examine the relationship between keywords. This analysis helps identify content gaps and determine how keywords can effectively connect to create innovative stories. It is worth noting that the proposed method can be readily applied in practical scenarios. VOD platforms can leverage this technique to enhance their competitive edge in the market.

Additionally, this research aspires to provide a fresh perspective on emphasizing the importance of content alongside other issues related to technique within the realm of cultural matters. This approach can be applied to various artistic domains such as cinema, theater, literature, and even content creation by influencers on platforms like YouTube or Instagram.

In the future, we aim to add a third phase to this recommender system, introducing a strategy for customer retention.

Also, future studies may be able to build upon this methodological approach and apply it to a wider, global population. Additionally, scholars could explore the viability of implementing this technique to analyze the content of other cultural artifacts and expressions. Such investigations could yield valuable insights by examining the adaptability and broader applicability of the described method across diverse cultural contexts and data sources. Alternative methods for implementing a recommender system can also be explored to suggest content and identify content gaps. For instance, leveraging techniques like associative rule mining could establish connections between keywords, elucidating each keyword's predominant associations with others.

To put it in a nutshell, this paper presents a pioneering, content-based two-phase recommendations system, addressing a gap in the existing literature. By leveraging the intrinsic properties of movies, rather than relying solely on technical features or demographics, our innovative approach offers a more comprehensive understanding of user interests. This novel design endeavors to cover all audiences' tastes and represents a significant advancement in the field of recommendation systems. On the other hand, this trailblazing recommender system increases the competitive power of VOD vendors, as it enables a content-driven war for user satisfaction.

VII. LIMITATIONS

Many people access VOD content through unofficial and illegal channels, so we cannot analyze the real societal feedback on movies. Also, the IMDB contribution for Iranian series is very small compared to the global reaction, which did not provide us with actual feedback. Also, besides the specific movies exclusively presented on a single VOD platform, they consist of many movies accessible across multiple platforms. In this paper, we limited our study to only the exclusive series, while we could have also utilized all the movies made available by these platforms.

REFERENCES

[1] M. V. Alvarez, J. M. T. López, and M. J. U. Ruíz, "What are you offering: An overview of VODs and recommender systems in European public service media," in *Information Technology and Systems*, 2020, pp. 725–732.
[2] J. Savage, "The effect of video on demand services on the cinema industry," Tech. Rep., 2017.
[3] *Video on Demand Market Size—Global Industry, Share, Analysis, Trends and Forecast 2023–2032*, Acumen Res. Consulting, Pune, Maharashtra, Jul. 2023.
[4] *PWC, How Tech Will Transform Content Discovery*. Pricewaterhouse Coopers Consumer Intelligence Series. Accessed: Sep. 2, 2019. [Online]. Available: https://tinyurl.com/r69zrq5
[5] M. O. Jackson, "An overview of social networks and economic applications," *Handbook social Econ.*, vol. 1, pp. 511–585, Jan. 2011.
[6] L. Hagen, T. Keller, S. Neely, N. DePaula, and C. Robert-Cooperman, "Crisis communications in the age of social media: A network analysis of zika-related tweets," *Social Sci. Comput. Rev.*, vol. 36, no. 5, pp. 523–541, Oct. 2018.
[7] R. S. Burt, "Structural holes and good ideas," *Amer. J. Sociol.*, vol. 110, no. 2, pp. 349–399, Sep. 2004.
[8] V. Latora, V. Nicosia, and P. Panzarasa, "Social cohesion, structural holes, and a tale of two measures," *J. Stat. Phys.*, vol. 151, nos. 3–4, pp. 745–764, May 2013.
[9] T. L. D. E. Souza, M. Nishijima, and R. Pires, "Revisiting predictions of movie economic success: Random forest applied to profits," *Multimedia Tools Appl.*, vol. 82, no. 25, pp. 38397–38420, Oct. 2023.
[10] S. Sahu, R. Kumar, H. V. Long, and P. M. Shafi, "Early-production stage prediction of movies success using K-fold hybrid deep ensemble learning model," *Multimedia Tools Appl.*, vol. 82, no. 3, pp. 4031–4061, Jan. 2023.
[11] K. Lee, J. Park, I. Kim, and Y. Choi, "Predicting movie success with machine learning techniques: Ways to improve accuracy," *Inf. Syst. Frontiers*, vol. 20, no. 3, pp. 577–588, Jun. 2018.
[12] S. M. R. Abidi, Y. Xu, J. Ni, X. Wang, and W. Zhang, "Popularity prediction of movies: From statistical modeling to machine learning techniques," *Multimedia Tools Appl.*, vol. 79, nos. 47–48, pp. 35583–35617, Dec. 2020.
[13] I. S. Ahmad, A. A. Bakar, and M. R. Yaakub, "Movie revenue prediction based on purchase intention mining using Youtube trailer reviews," *Inf. Process. Manag.*, vol. 57, no. 5, Sep. 2020, Art. no. 102278.
[14] W. Bai, Y. Zhang, W. Huang, Y. Zhou, D. Wu, G. Liu, and L. Xiao, "DeepFusion: Predicting movie popularity via cross-platform feature fusion," *Multimedia Tools Appl.*, vol. 79, nos. 27–28, pp. 19289–19306, Jul. 2020.
[15] D. Xi, W. Xu, R. Chen, Y. Zhou, and Z. Yang, "Sending or not? A multimodal framework for danmaku comment prediction," *Inf. Process. Manag.*, vol. 58, no. 6, Nov. 2021, Art. no. 102687.
[16] A. Spitz and E.-Á. Horvát, "Measuring long-term impact based on network centrality: Unraveling cinematic citations," *PLoS ONE*, vol. 9, no. 10, Oct. 2014, Art. no. e108857.
[17] A. Noroozian, B. Amiri, and M. A. M. A. Kermani, "Investigation of cinematic genre diversity based on social network analysis: The lost ring of the Iranian cinema industry," *Kybernetes*, vol. 52, no. 12, pp. 6321–6347, Nov. 2023.
[18] B. Patel, P. Desai, and U. Panchal, "Methods of recommender system: A review," in *Proc. Int. Conf. Innov. Inf., Embedded Commun. Syst. (ICIIECS)*, Mar. 2017, pp. 1–4.
[19] R. Weidhaas, S. Schlogl, V. Halttunen, and T. Spieß, "Watch this! The influence of recommender systems and social factors on the content choices of streaming video on demand consumers," in *Innovation Through Information Systems*. Berlin, Germany: Springer, pp. 738–753.
[20] J. Fang, L. Liu, M. A. Hossin, C. Wen, and G. Xia, "Market competition as a moderator of the effect of social signals on viewership in video-sharing platforms," *Inf. Process. Manag.*, vol. 60, no. 3, May 2023, Art. no. 103329.
[21] K. Pripužic, I. P. Žarko, V. Podobnik, I. Lovrek, M. Cavka, I. Petkovic, P. Stulic, and M. Gojceta, "Building an IPTV VoD recommender system: An experience report," in *Proc. 12th Int. Conf. Telecommun.*, Jun. 2013, pp. 155–162.
[22] T. Dörsch, A. Lommatzsch, and C. Rakow, "Topical video-on-demand recommendations based on event detection," in *Proc. LWDA*, 2016, pp. 186–193.
[23] M. Ferrari Dacrema, N. Felicioni, and P. Cremonesi, "Offline evaluation of recommender systems in a user interface with multiple carousels," *Frontiers Big Data*, vol. 5, Jun. 2022, Art. no. 910030.
[24] A. Said, D. Tikk, K. Stumpf, Y. Shi, M. A. Larson, and P. Cremonesi, "Recommender systems evaluation: A 3D benchmark," in *Proc. RUE@ RecSys*, 2012, pp. 21–23.
[25] F. Ricci, L. Rokach, and B. Shapira, "Introduction to recommender systems handbook," in *Recommender Systems Handbook*. Berlin, Germany: Springer, 2010, pp. 1–35.
[26] S. S. Lakshmi and T. A. Lakshmi, "Recommendation systems: Issues and challenges," *Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 4, pp. 5771–5772, 2014.
[27] S. C. Guntuku, S. Roy, W. Lin, K. Ng, N. Wee Keong, and V. Jakhetiya, "Personalizing user interfaces for improving quality of experience in VoD recommender systems," in *Proc. 8th Int. Conf. Quality Multimedia Exper. (QoMEX)*, Jun. 2016, pp. 1–6.

[28] R. Pereira and C. Tam, "Impact of enjoyment on the usage continuance intention of video-on-demand services," *Inf. Manag.*, vol. 58, no. 7, Nov. 2021, Art. no. 103501.
[29] G. Kour and B. Chhabria, "Understanding platform strategies for consumer stickiness on OTT platforms," *J. Indian Bus. Res.*, vol. 14, no. 4, pp. 540–555, Nov. 2022.
[30] Q. Zhao, Z. Zhao, L. Yang, L. Hong, and W. Han, "Research on customer retention prediction model of VOD platform based on machine learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 4, 2023.
[31] F. Loh, K. Hildebrand, F. Wamser, S. Geißler, and T. Hoßfeld, "Machine learning based study of QoE metrics in Twitch.Tv live streaming," in *Proc. IEEE/IFIP Netw. Operations Manage. Symp.*, May 2023, pp. 1–7.
[32] C. Ruiz and M. Quaresma, "UX aspects of AI principles: The recommender system of VoD platforms," in *Proc. Int. Conf. Human-Comput. Interact.*, Springer, 2021, pp. 535–552.
[33] S. H. Ghasemi, A. Arshadi Khamseh, and M. V. Sebt, "Identifying critical supply chain risks through social network analysis: ICT company in Iran," *J. Ind. Syst. Eng.*, vol. 15, no. 1, pp. 179–202, 2023.
[34] M. Maghsoudi, M. Valikhan, and M. Zohdi, "Video recommendation using social network analysis and user viewing patterns," 2023, *arXiv:2308.12743*.
[35] N. Kartikasari, R. Ambarwati, and A. Pebrianggara, "Marketing content analysis of Korean drama streaming platforms: Social network analysis on Twitter," *Budapest Int. Res. Critics Inst.*, vol. 6, no. 2, pp. 798–809, 2023.
[36] R. Zafarani, M. A. Abbasi, and H. Liu, *Social Media Mining: An Introduction*. Cambridge, U.K.: Cambridge Univ. Press, 2014.
[37] A. Saxena and S. Iyengar, "Centrality measures in complex networks: A survey," 2020, *arXiv:2011.07190*.
[38] M. Agha Mohammad Ali Kermani, A. Badiee, A. Aliahmadi, M. Ghazanfari, and H. Kalantari, "Introducing a procedure for developing a novel centrality measure (sociability centrality) for social networks using TOPSIS method and genetic algorithm," *Comput. Hum. Behav.*, vol. 56, pp. 295–305, Mar. 2016.
[39] E. Lazega, "Structural holes: The social structure of competition," *JSTOR*, vol. 36, no. 4, pp. 779–781, 1995.
[40] R. S. Burt, R. M. Hogarth, and C. Michaud, "The social capital of French and American managers," *Org. Sci.*, vol. 11, no. 2, pp. 123–147, Apr. 2000.
[41] R. S. Burt, *Structural Holes: The Social Structure of Competition*. Harvard Univ. Press, 1995.
[42] K. Shannon and R. T. Snodgrass, "Semantic clustering," in *Proc. POS*, 1990, pp. 389–402.
[43] M. R. Asadabadi, E. Chang, and M. Saberi, "Are MCDM methods useful? A critical review of analytic hierarchy process (AHP) and analytic network process (ANP)," *Cogent Eng.*, vol. 6, no. 1, Jan. 2019, Art. no. 1623153.
[44] C.-L. Hwang, K. Yoon, C.-L. Hwang, and K. Yoon, "Methods for multiple attribute decision making," in *Multiple Attribute Decision Making*, 1981, pp. 58–191.
[45] J. Salvatore, "Bipartite graphs and problem solving," Univ. Chicago, Tech. Rep., 2007.
[46] D. P. Mines and S. Lamb, *Everyday Life in South Asia*. Indiana Univ. Press, 2010.

ALI NOROOZIAN received the B.Sc. degree from the Sharif University of Technology and the M.Sc. degree in industrial engineering from Iran University of Science and Technology (IUST). He has experience in data and network science. He brings over four years of managerial and product development expertise in both B2B and B2C markets and shows a deep fascination with the study of consumer behavior. He has further shaped his multifaceted interests, spanning marketing, cultural artifact analysis, data science, and societal-economic factors. This diverse scholarly background uniquely positions him to undertake comprehensive, data-driven investigations into consumer behavior, and market dynamics.

BABAK AMIRI received the Ph.D. degree in information technology from The University of Sydney, in 2014. He is currently an Assistant Professor with Iran University of Science and Technology. He has published extensively in highly regarded scientific journals and has contributed to the advancement of knowledge in complex systems, machine learning, and data science. His research interests include artificial intelligence, big data analytics, and social computing, with a particular focus on the application of complex network theories to various business and engineering problems.

SAMIRA RAMEZANI is currently pursuing the M.S. degree in industrial engineering with Iran University of Science and Technology (IUST). She is skilled in machine learning and data science. Her research interests include supply chain management and human resources.