*Human Interaction and Emerging Technologies (IHIET-FS 2025),* Vol. 196, 2025, 186–197
https://doi.org/10.54941/ahfe1005967

# Social Media and Internet Celebrity for Social Commerce Intentional and Behavioural Recommendations

**Shu-Hsien Liao and Yao-Hsuan Yang**

Department of Management Sciences, Tamkang University, New Taipei City, Taiwan 251, R.O.C.

## ABSTRACT

Social media is a virtual community and online platform that people use to create, share, and exchange opinions/experiences. Internet celebrities are people who become famous on the Internet, increasing their popularity through their social networking or video websites. Social commerce (s-ecommerce) is the combination of social relations and commercial transaction activities. The combination of social media and Internet celebrities is an emerging model for the development of s-ecommerce. With recent advances in system sciences, recommendation systems are gradually moving to develop intentional and behavioural recommendations. This background leads to the research issues regarding digital and social media in enterprises. Thus, this study implements data mining analytics, including clustering analysis and association rules, to investigate Taiwanese users (n = 2,102) to investigate social media and Internet celebrities' preferences to find knowledge profiles/patterns/rules for s-ecommerce intentional and behavioural recommendations.

**Keywords:** Social media, Internet celebrity, Social commerce, S-ecommerce, Data mining analytics, Intentional and behavioural recommendations

## INTRODUCTION

This study considers that the signal targeting of user data can reveal consumers' intentional and behavioural signal targeting and combine the two different types of signal targeting to provide intentional and behavioural recommendations for s-ecommerce. The recommendation systems typically generate recommendations through collaborative filtering, content-based recommendations, or personalized recommendations. Collaborative filtering records a user's historical behaviour in combination with similar decisions made by other users. Content-based recommendation uses a list of discrete characteristics of the items to recommend similar items to users with similar properties. Recommendation systems are an effective alternative to search algorithms because they help users find items that they are unlikely to find on their own (Lee and Wu, 2022). They are often implemented using search engines rather than traditional data indexing. However, in 2023, Google will discontinue the use of third-party cookie technology in the browser and will not use any alternative methods to track users' web browsing data (Google Workspace Update, 2023).

 

The next step in recommendation systems could relate to signal targeting. Many intentional signals are generated when consumers need a certain product to satisfy their needs or solve problems (Cheng et al., 2023). The Internet has numerous signals that goods are in demand. Therefore, the behavioural signal targeting of traditional recommendation systems differs from the intentional signal targeting for recommendations. It can be said that behavioural recommendation is a point-to-point marketing extension, where merchants find the people who want to buy a product and deliver that product (Billore et al., 2023). For example, behavioural recommendation occurs when a consumer clicks on a smartphone store catalogue. The system shows "This person is looking for smartphones". So, for the next two weeks, when the consumer clicks on a website, ads for smartphones will pop up. Behavioural recommendation can only provide marketing/promotion based on past behaviour records. On the other hand, intentional recommendation is a mindset that seeks to understand consumers' lives and intentions; by constantly collecting information about Internet users' behaviour and monitoring events and information in consumers' lives, it leads consumers to gradually explore their needs, wants and demand. Intentional recommendation depicts what a person's life is like through the things (signals) that the consumer clicks on. Based on this, it analyses the person's specific profile, and then further presents information that they may need (targeting). The information relates not only to commodities, but also includes smartphone apps, news, social media, socializing, gaming. Key Opinion Leaders (KOL) etc., on smartphones (Yang et al., 2023).

Social media is an online media platform based on interests and creative content formed by a group of Internet users through Web 2.0 network application technologies (Jia et al., 2023). Social media can also be the most frequent contact point for people to interact with each other. It is not only used to transmit information and knowledge, but also supports communication, digital content, and even social commerce (s-ecommerce) promotion (Dowling, 2023). With the rapid development of social media, the influence of Internet celebrities such as YouTubers has also increased. In addition to being an important channel for s-ecommerce, Internet celebrities also demonstrate the trend of monetizing an online personality (Hong, 2023).

Internet celebrities, abbreviated as influencers or KOL, are people who become famous on the Internet, increasing their popularity by their social networking or video websites to develop a career from this (Lu et al., 2022). More established Internet celebrities will regularly publish articles, photos, videos, music and other works on social platforms or audio and video websites to accumulate their own audiences or fans. They may have their own studios, employees, agencies, and some even have their own brands or products. Platforms such as Instagram, Snapchat, YouTube and VSCO are the most common social media through which media anchors build the following. The combination of social media and live streaming by Internet celebrities is an emerging model for the development of s-ecommerce (Jun et al., 2023).

Social commerce is the combination of social relations and commercial transaction activities, abbreviated as s-ecommerce and used to describe a new

online sales model or marketing strategy that is a mix of social networks and peer-to-peer communication (Zhao et al., 2023). It is also a means for consumers to communicate through social media while satisfying their shopping needs. Since its inception, s-ecommerce has taken many forms. Business opportunities derived from mobilizing family, friends, or netizens to make group purchases, or having users share information on social networking sites are also part of s-ecommerce (Liu et al., 2023).

Thus, this study implements data mining analytics, including clustering analysis and association rules, to investigate Taiwanese users (n = 2,102) to investigate social media, and Internet celebrities' behaviours and preferences to find meaningful profiles/patterns/rules for s-ecommerce intentional and behavioural recommendations.

## LITERATURE REVIEW

### Social Media Recommendations

Gruzd et al. (2023) explored the role of Facebook and YouTube in exposing people to misinformation related to the COVID-19 vaccine. Specifically, by examining potential levels of exposure, the paper modelled a one-way information-sharing pathway that began with Facebook users encountering a vaccine-related post with a link to a YouTube video, following the link to YouTube, and then seeing videos automatically recommended by the associated YouTube list.

### Internet Celebrity Recommendations

Park and Lin (2023) explored the effects of various contests on consumer attitudes in the context of live streaming on Chinese Netflix. Their results showed that product-source fit affected perceived source attractiveness and credibility, while product-content fit affected utilitarian and hedonic attitudes toward the content. That study also developed and tested a model of Netflix celebrities by investigating the congruence effect of a live-streaming audience.

### S-ecommerce Recommendations

Zhao et al. (2023) proposed cross-domain recommendations for e-commerce using user-level preference transfer network (CDR-ULPT). This process first incorporates an attention mechanism for resistant domain adaptation to personalize users' domain sharing preferences. Experiments on five public datasets showed that this method outperforms other cross-domain recommendation methods.

### Data Mining for Social Media, Internet Celebrity and S-ecommerce Recommendations

Chaturvedi et al. (2023) developed a conceptual framework for a comprehensive understanding of travel experiences with AI technologies, which required performance analysis of scientific participants, thematic analysis, and content analysis of research knowledge structures. Key technologies that shape the tourist experience in this context include service

robots, personalized recommendation systems, Internet of Things, chatbots, AR/VR, big data analytics, text prospecting, text analytics and natural language processing for recommending e-commerce.

## METHOD

### The Conceptual Model

Internet celebrities are now extending beyond social networking and studios, bringing their followers to new levels of user experience and usability interacting with social media/network on socializing and communication in addition to online platforms. Thus, this study implements data mining analytics, including clustering analysis and association rules, to investigate Taiwan users (n = 2,102) to investigate social media, and Internet celebrities' behaviours and preferences to find meaningful profiles/patterns/rules for s-ecommerce intentional and behavioural recommendations. Fig. 1 shows the conceptual model for data modelling and analytics.

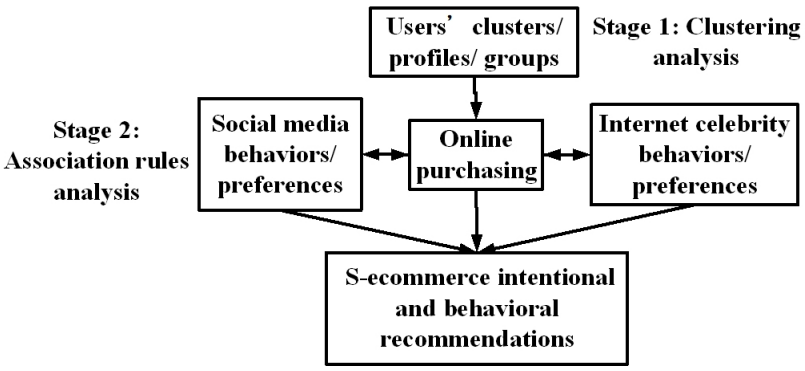


**Figure 1:** The conceptual model.

### Subject Background and Data Collection

From 12 noon to 12 am, questionnaires were distributed randomly on social media platforms daily from July 8, 2022, to October 27, 2022, for a total of 2,243 questionnaires. After collecting and sorting the questionnaires, there were 2,102 valid responses after deducting invalid questionnaires with repeated IPs, too long or too short a period, or with missing answers. The recovery rate of valid questionnaires was 93.7%.

### Database Design – Snowflake Schema

A snowflake schema contains a fact data table and a dimension data table, with the fact table in the centre of the data structure, surrounded by the dimension table. A snowflake schema normalizes each dimension table of the star schema and stores them in independent tables to form a multi-level data structure state. In this study, snowflake schema contains one fact table and 12-dimensional tables.

## Data Mining Analytics

### Clustering Analysis

This study uses the K-means algorithm and divides consumers into groups according to their consumption behaviour and preferences (Gustriansyah et al., 2022). The steps are as follows:

We assume there are N data sets {X1, X2,..., Xp}, and randomly extract k initial clusters from them.

The receiver uses Euclidean distance to calculate the distance between each data and the average value of each initial cluster and then assigns each data to the cluster with the closest distance. The calculation method is:

$$||Xi - Zi|| \leqq ||Xi - Zp||$$

Xi: each different data; Zi, Zp: the mean value of the initial cluster, i={1,2,3,…,K}, p={1,2,3,…,K}, i≠p

We divide the observed value (X_(1,) X_(2,)…,X_p) into k determined cluster initial centres (m_1^((1)),…,m_k^((1))), and the receiver uses the Euclidean distance to calculate the distance between each data and the average value of each initial cluster, and then assigns each data to the cluster with the closest distance. The algorithm is as follows:

$$S_i^{(t)} = \{x_p : ||x_p - m_i^{(t)}||^2 \leq ||x_p - m_j^{(t)}||^2 \forall j, 1 \leq j \leq k\} \quad (1)$$

When each cluster centre has already been classified observations value, and calculate the Euclidean distance from each observation value to k cluster which is assigned to the cluster with the nearest distance to it, the algorithm is as follows:

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j \quad (2)$$

Then we continue to repeat the above steps until the change of the cluster centre becomes smaller and smaller, and the observation value does not change, then the final structure is generated.

### Association Rules

Agrawal and Srikant (1994) first proposed that the main purpose is to find out the relationship between the data in the database, and to explore the meaning of the relationship. Association rules are often used to analyse the association of different commodity combinations in the database. This study assumes that consumers will purchase B (Consequent) because of the consumption behaviour of A (Antecedent), and the association rules are expressed by two parameters: Support and Confidence value. In the algorithm for finding association rules, the thresholds of Minimum support and Minimum confidence set by users must be met before the rules can be determined to be meaningful. The formulas for calculating support and confidence are as follows (Agrawal and Srikant, 1994):

A. Set the threshold value of Minimum support and Minimum Confidence.

B. The Apriori algorithm uses the concept of Candidate itemsets. If the support of the candidate item set is greater than or equal to the minimum support, the candidate item set is the Large itemsets.
C. Probability of all items in a rule which is divided by the product of probabilities of items X and Y occurring if there is no association among them found. The larger lift value the higher association between two items.

Using the data in the entire database is searched and analysed repeatedly, and then the analysed content and the minimum support are compared, and the above actions are repeated until no new candidate item sets are combined.

## RESULTS

### Clustering Analysis

According to the characteristics obtained from the clustering results, the clusters were named Cluster-1 (736 data) peer entertainment group, Cluster-2 (658 data) family life group, and Cluster-3 (708 data) personal leisure group. The group profile and feature data of the analysis results are described below.

*Cluster-1 Peer entertainment group:* This group of viewers is mainly single male students aged 19 to 24 years old, with a monthly disposable income of less than 2,000 NTD, and an average of 1 to 3 hours of following celebrities per day. Laptops are the main device, and YouTube is the main social media to follow. The celebrities for clister-1 most interested in are gaming celebrities. The motivation for continuing to follow celebrities is for game entertainment. The group of students who mainly pay attention to games do not have much disposable income, and they spend more time following celebrities to get chatting topics among their peers. This cluster usually focuses on celebrities about games and entertainment, so that named as the peer entertainment group.

*Cluster-2 Family life group:* This group is mainly composed of married female office workers aged 25 to 60 years old, with a monthly disposable income of more than 10,001 NTD, and the average daily time spent following celebrities is the least among the three groups. TV is the main device to follow celebrities, and the main social media is Facebook. The celebrities are most interested in singing type celebrities. Motivation for continuing to follow celebrities is to learn about different foods and restaurants. They usually follow singing type celebrities. This group is enjoying their life when traveling with family members.

*Cluster-3 Personal leisure group:* This group is mainly composed of single male office workers aged 25–60 years old, with monthly disposable income of more than 10,001 NTD, and the average daily time spent following celebrities is 1 to 3 hours. The main device is the desktop computer to follow celebrities, and the main social media is Facebook. The most interesting topic on celebrities is food, eating and streaming. The motivation for continuing to follow celebrities is to learn about different food and restaurants. The topic they are usually interested in is news topics, and they will also learn about celebrity news as their own interests. They all focus on personal leisure and entertainment, so that named as the personal leisure group.

## Association Rules Analysis

### Pattern 1 – Associations of Internet Celebrity Preferences and Intentional Signals

In this study, under the condition that the Minimum antecedent support value is greater than 2% and the Minimum confidence value is greater than 30%, five meaningful correlation rules are generated in pattern 1 (Fig. 2). The Lift values are all greater than 1. In terms of finding Internet celebrity preferences, signal targeting questions include: What topics are interested in? Why continue to follow celebrities? How to support celebrities? What kind of celebrity activities participated in? What factors focus on celebrities? For Rule 1, when the antecedent is a game type, the antecedent includes game entertainment, live messaging, game streaming media, and is associated with a celebrity. On the other hand, for Rule 2, when consequent is food eating type, the antecedents are learning about different cuisines and restaurants, subscription, live chat and recommended by friends and family.

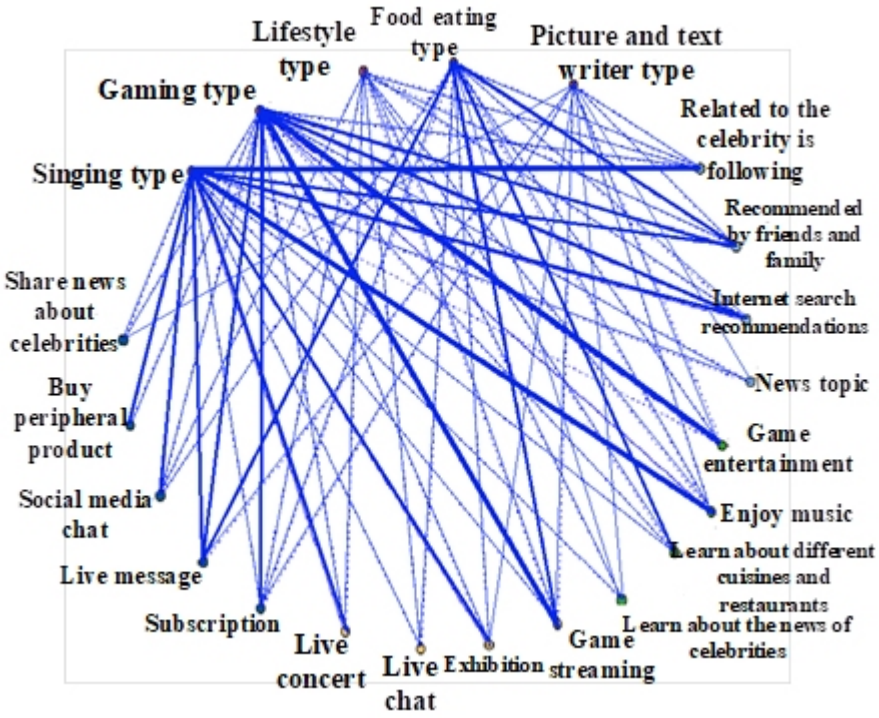


**Figure 2:** Associations of Internet celebrity preferences and intentional signals (Cluster-1).

### Pattern 2 – Associations of Social Media Preferences and Intentional Signals

For Rule 2, when consequent is personal life, the antecedents are high media frequency, providing overseas sales service for celebrity products, convenient online payment, and Instagram. Regarding Cluster – 2, for Rule 1, when consequent is current affairs discussion, the antecedents include high

community discussion, recommend celebrities frequently, save time, and Facebook. On the other hand, for Rule 3, when consequent is personal life, the antecedents are member-only videos/benefits, provide overseas sales service for celebrity products, and YouTube (Figure 3). Regarding Cluster – 3, for Rule 3, when consequent is singing performance, the antecedents include collaborating with other celebrities, recommend celebrities frequently, save time and Facebook.

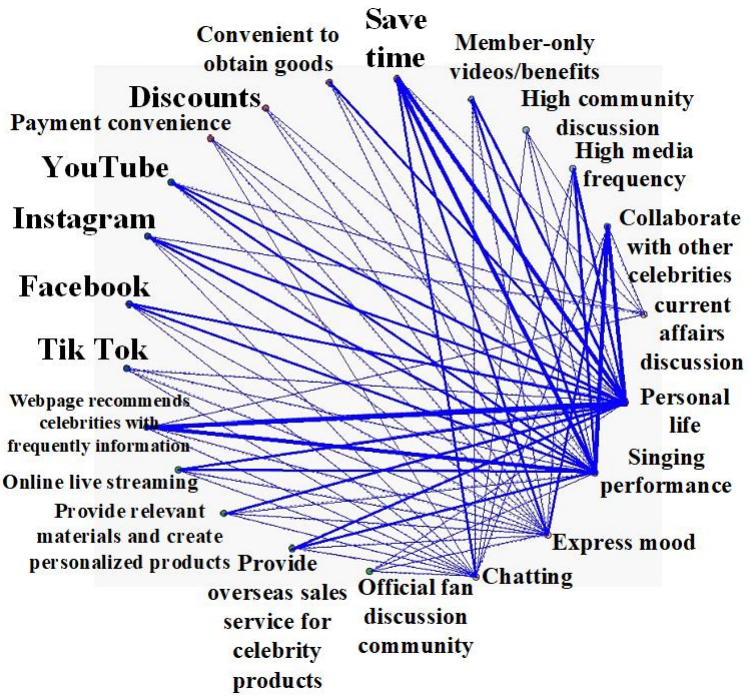


**Figure 3:** Associations of social media preferences and intentional signals (Cluster-2).

### Pattern 3 – Associations of S-ecommerce Preferences and Behavioural Signals

In this study, under the condition that the Minimum antecedent support value is greater than 2% and the Minimum rule confidence value is greater than 30%, five meaningful correlation rules are generated in the pattern 2. The Lift values are all greater than 1. In terms of finding social media preferences, signal targeting questions include what kind of products do you prefer to buy online? What online shopping themes would you be interested in? Motives for s-ecommerce goods, what situations will increase intention for online buying? which payment method is preferred for s-ecommerce? Regarding Cluster – 1, for Rule 1, when consequent is personal things, the antecedents include singing type, like music/show, related to celebrities, store pickup and payment (Fig. 4).

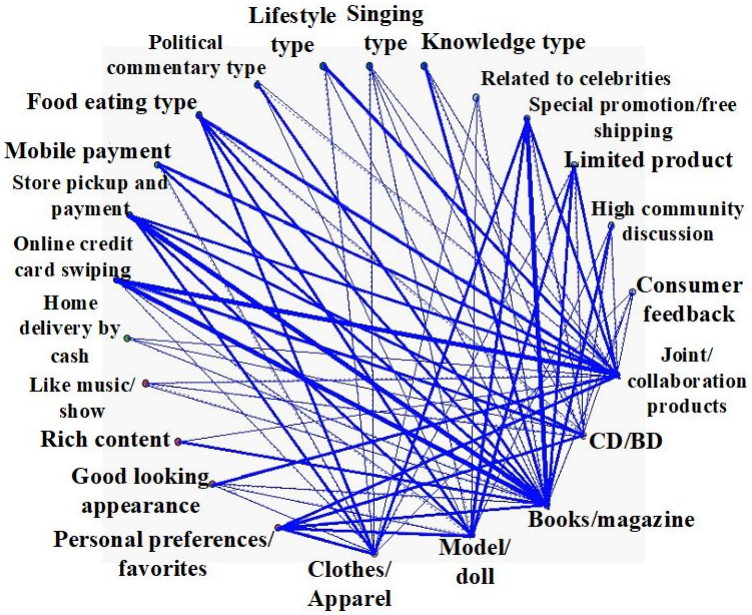


**Figure 4:** Associations of s-ecommerce preferences and intentional signals (Cluster-3).

## DISCUSSIONS

### Internet Celebrity Intentional Recommendations

Regarding marketers and e-commerce platforms, when a buyer looks for information about buying a camera, the system will also recommend camera batteries to the buyer. However, collaborative filtering will be limited due to sparse and incomplete data, which affects the accuracy of recommendations. Thus, by analysing users' social media behaviours and things (signal), this study suggests intentional signal targeting recommendation methods to determine the media usage intentions of active users based on integrated user similarity and then provide customized social media recommendations to match/predict the user's interests.

### Social Media Intentional Recommendations

Regarding marketers and e-commerce platforms, a business must have a comprehensive understanding of consumers to predict their future needs. At the same time, Internet operators must also clearly demonstrate what the recommendation system can provide to what type of consumers.

### S-ecommerce Behavioural Recommendations

Regarding marketers and e-commerce platforms, the consumer profile of signal targeting is no longer rigidly defined as simply gender, age, and location; rather, it is based on people's lives and collects richer and more diverse elements. For example, regarding coffee preference, does a consumer prefer to brew it by hand or buy takeaway? Through this element, merchants

can promote the appropriate coffee advertisement to lead the consumer to purchase. Thus, signal targeting provides a clearer picture of the consumer profile and a process for behavioural recommendations (Fig. 5).

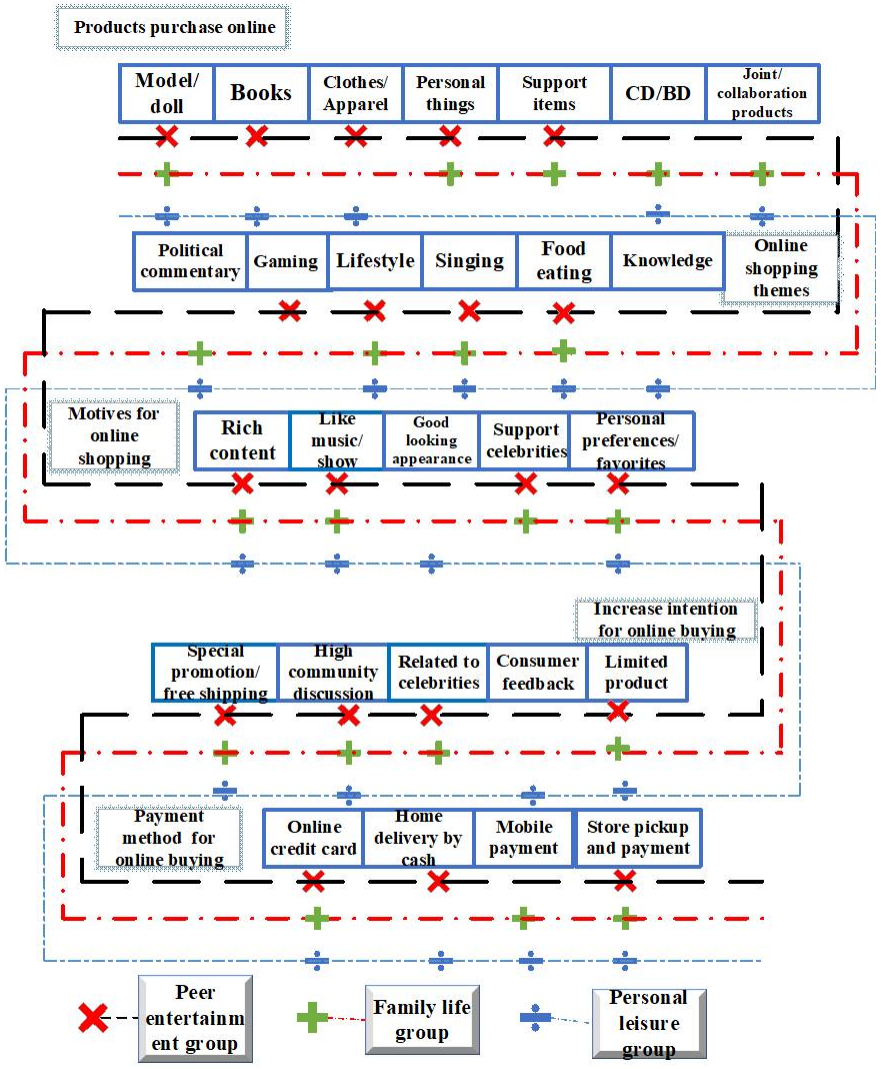


**Figure 5:** Knowledge map of s-ecommerce behavioural recommendations.

## CONCLUSION

The information relates not only to commodities, but also includes smartphone apps, news, social media, socializing, gaming KOL etc., on smartphones. In these regards, this study considers that signal targeting is a method by which social media operators and internet celebrities' platforms can understand consumers' media tools preferences and individual lifestyles so that operators can effectively recommend social commerce to the consumers. Thus, this study first investigates Taiwan

users to investigate social media, and Internet celebrities' preferences to find knowledge profiles/patterns/rules for s-ecommerce intentional and behavioural recommendations.

## ACKNOWLEDGMENT

The authors would like to acknowledge the National Science and Technology Council, Taiwan, Republic of China [MOST 113-2410-H-032-051-MY2].

## REFERENCES

Agrawal, R. Srikant, R. (1994) "Fast algorithms for mining association rules." Paper presented at the Proceedings of the 20th International Conference on very Large Database, Santiago, Chile, 20, pp. 487–499.

Billore, S. Anisimova, T. Vrontis, D. (2023) "Self-regulation and goal-directed behavior: A systematic literature review, public policy recommendations, and research agenda." Journal of Business Research. Volume 156, No. 4.

Chaturvedi, R. Verma, S. Ali, F. Kumar, S. (2023) "Reshaping Tourist Experience with AI-Enabled Technologies: A Comprehensive Review and Future Research Agenda." International Journal of Human–Computer Interaction, Published Online: 09 Aug 2023, https://doi-org.ezproxy.lib.tku.edu.tw/10.1080/10447318.2023.2238353

Cheng, J. Chen, B. Huang, Z. (2023) "Collective-based ad transparency in targeted hotel advertising: Consumers' regulatory focus underlying the crowd safety effect." Journal of Retailing and Consumer Services, Volume 72, No. 2.

Dowling, M. E. (2023) "News to me: far-right news sharing on social media." Information, Comminication & Society, Published Online: 30 Jan, https://doi.org/10.1080/1369118X.2023.2166796.

Google Workspace Update. (July 6, 2023) "Upcoming changes to third-party cookie requirements in Google Drive." Website: https://workspaceupdates.googleblog.com/2023/10/upcoming-changes-to-third-party-cookies-in-google-drive.html.

Gruzd, A. Abul-Fottouh, D. Yun, M. Song, J. Saiphoo, A. (2023) "From Facebook to YouTube: The Potential Exposure to COVID-19 Anti-Vaccine Videos on Social Media." Social Media + Society, Published online February 1, 2023, https://doi-org.ezproxy.lib.tku.edu.tw/10.1177/20563051221150.

Gustriansyah, R. Ermatita, E. Rini, D. P. (2022) An approach for sales forecasting. Expert Systems With Applications, Volume 207, No. 6.

Hong, F.-Y. (2023) "Structural Equation Model Analysis of Taiwanese Teenagers' Attitudes and Behaviors Related to Live Stream Watching." International Journal of Human–Computer Interaction, Published Online: 20 Aug 2023, https://doi-org.ezproxy.lib.tku.edu.tw/10.1080/10447318.2023.2247608.

Jia, M. Zhao, Y. Li, J. (2023) "How vicarious learning increases users' knowledge adoption in live streaming: The roles of parasocial interaction, social media affordances, and knowledge consensus." Information Processing & Management, Volume 61, No. 2.

Jun, M. Han, J. Eisingerich, B. (2023) "When is celebrity endorsement effective? Exploring the role of celebrity endorsers in enhancing key brand associations." Journal of Business Research, Volume 164, No. 4.

Lee, P.-J. Wu, T.-Y. (2022) "Mining relations between personality traits and learning styles." Information Processing & Management, Volume 59, No. 5.

Lu, X. Xie, Y. Zhang, B. Li, K. (2022) "Supply–Demand Docking Joint Decision of Cross-border E-commerce under the Influence of Internet Celebrity." Journal of Global Information Technology Management, Volume 25, No. 1.

Liu, L. Fang, J. Wen, C. (2023) "The power of talk: Exploring the effects of streamers' linguistic styles on sales performance in B2B livestreaming commerce." Information Processing & Management, Volume 60, No. 3.

Park, H. J. Lin, L. M. (2023) "The effects of match-ups on the consumer attitudes toward internet celebrities and their live streaming contents in the context of product endorsement." Journal of Retailing and Consumer Services, Volume 52, No. 6.

Yang, C. Zheng, R. Wang, H. (2023) "Content recommendation with two-level TransE predictors and interaction-aware embedding enhancement: An information seeking behavior perspective." Information Processing & Management, Volume 60, No. 4.

Zhao, J. Yao, L. Zhang, T. (2023) "CDR-ULPT: Cross-domain recommendation for e-commerce using user-level preferences transfer network." Electronic Commerce Research and Applications, Volume 61, No. 3.