A Data Mining Approach for Developing Online Streaming Recommendations

Shu-Hsien Liao, Retno Widowati & Hao-Yu Chang

Published online: 07 Jan 2022.

Submit your article to this journal ↗

Article views: 583

View related articles ↗

View Crossmark data ↗

Taylor & Francis
Taylor & Francis Group

A Data Mining Approach for Developing Online Streaming Recommendations

Shu-Hsien Liao [iD][a], Retno Widowati[b], and Hao-Yu Chang[a]

[a]Department of Management Sciences, Tamkang University, New Taipei City, Taiwan (R.O.C.);
[b]Department of Management, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia

ABSTRACT

Online streaming has become increasingly popular with the availability of broadband networks and the increase in computing power and electronic distribution. Online streaming operators have difficulty developing flexible business alternatives according to users' changing streaming behaviors in terms of generating a good and profitable business model. In terms of e-commerce development, the live-streaming platform that provides streaming of the main merchandise to users, allowing the users to directly consume via live-streaming become critical issues. In this regard, personalized recommendation systems can use the user's interests and purchasing behavior to recommend information and merchandise. Thus, this study investigates the online streaming experiences of Taiwanese consumers to evaluate online streaming users and their online purchase behaviors for developing online recommendations. This study uses a snowflake schema, which is an extension of the star schema. In addition, this study develops a rule-based recommendation approach for investigating online streaming and purchasing behaviors in terms of online recommendations. This study is the first to determine how online streaming proprietors and their affiliates are disseminated using online streaming consumer behaviors in terms of online recommendations for further electronic commerce development.

Introduction

With the increasing popularity of broadband networks and the improvement of computing power, live-streaming media that was originally only suitable for electronic media and enterprises can be extended to users and families on the Internet (Singh and Sharma 2020). The line stream uses and continues the advantages of the Internet, and uses video for online live streaming. This stream can display product demonstrations, conferences, background introductions, program evaluations, online surveys, dialog interviews, online training, online libraries, and other online applications (Anderson 2017). On the road, using the fast, intuitive, and rich content of the Internet, as well as strong interactivity and lack of geographical restrictions, audiences can be divided to

CONTACT Shu-Hsien Liao ✉ michael@mail.tku.edu.tw 🖃 Department of Management Sciences, Tamkang University, No. 151, Yingjuan Rd., Danshuei Dist, New Taipei City 251, Taiwan (R.O.C.).

develop the effect of communication (Wang 2019). After the live stream is over, creators can continue to provide replay or on-demand services at any time, thereby effectively expanding the time and space of the live stream. One is to provide TV signal viewing on the Internet, such as live broadcast of various sports games and cultural activities. The principle of this live streaming is to collect TV signals from analog (analog) signals and input them to the computer, thereby providing real-time uploading websites for people to watch, which is equivalent to Internet TV. The other is live streaming in the true sense: an independent signal acquisition device (audio + video) is set up on the site, and then the director (guidance device or platform) is imported, and then uploaded to the server through the console (Bilal, Erbad, and Hefeeda 2018). The biggest difference between this type of streaming media and the previous one is the autonomy of live streaming: independent and controllable collection of audio and video is completely different from a single broadcast of TV signals.

As a result, the huge business opportunities brought by online shopping have also turned retail companies selling various products into virtual e-commerce suppliers, which has led to more common online shopping behaviors among consumers (Smink et al. 2019). In particular, Facebook launched the fan page function in 2007 and the Facebook Live Stream function in 2015, which not only made Facebook a platform for interpersonal interaction but also expanded its popularity, making Facebook an essential tool for online marketing. Since online streaming allows multiple viewers to watch the same video file at the same time, and the viewer can play the file interactively at any time, for example, the current Facebook shopping live streaming model has been developed to include auction-based activities (Lim and Kim 2018). Regarding to interactivity, live streaming allows consumers to browse and participate in real-time, and enhances the characteristics of the connection between people. This is also a consumption attribute that other traditional e-commerce websites cannot achieve (Huang et al. 2015). In addition, in terms of transparency, online streaming allows online and offline operators to display sites and products to consumers, thereby enabling consumers to evaluate whether the products meet their needs. This marketing model overcomes the key limitation of online shopping, that is, consumers cannot see the physical product and get the texture. Live streaming enables consumers to better grasp the texture of the product and improve consumers' imagination when actually using the product (Yang et al. 2019). Thus, through live streaming of product features, functions, prices, advantages, and disadvantages product attributes can be fully presented to consumer groups.

The live-streaming industry in Taiwan has developed rapidly, and its output value has increased by more than 24 times from 2016 to 2018. In 2020, it will reach an economic market of 300 billion TWD. In terms of the Taiwanese online streaming market, the results of the Taiwan MIC survey conducted by

the Taiwan Council show that 31.3% of Taiwanese have the habit of tracking specific broadcasters. In this group of followers, most of them interact with the broadcasters through methods such as click to support (38.6%), pure viewing (34.8%) and barrage/comments (26.9%). However, 15% of followers have spent money rewarding live broadcasters, 14.3% have purchased peripheral goods, 10.9% have participated in offline activities, and male followers (18.6%) are more willing to spend money than women for rewards (10.8%)(Taiwan Market Intelligence & Consulting Institute (Taiwan MIC) 2019). The above findings indicate that an influential economy has great development potential. Generally speaking, live broadcasting has different development paths in different regions. For example, China has developed a unique live-streaming media market, and it also has a special high-frequency, high-value reward culture. Among the first-tier, second-tier, and third-tier cities, the types of live programs and age levels can be classified. On the other hand, live-streaming media in Europe and the United States mostly follow the operating brand model, with the ultimate goal of creating its own brand and e-commerce operations. For example, a live broadcast can start with teaching beauty, then endorsements, selling cosmetics, and finally creating an e-commerce model of his/her own brand. Therefore, rewards or platform sharing are only the initial profit model (Taiwan TSRI 2019).

Thus, this study is the first to investigate the online streaming experiences of Taiwanese consumers to evaluate online streaming users and their online purchase behaviors for developing online recommendations. This study uses a snowflake schema including 2,124 data record, which is an extension of the star schema and combines association rules with rough sets in terms of developing a data mining approach. In addition, this study develops a rule-based recommendation approach for investigating online streaming and purchasing behaviors in terms of online recommendations. This study is the first to determine how online streaming proprietors and their affiliates are disseminated using online streaming consumer behaviors in terms of online recommendations for further electronic commerce development.

Literature Review

Online Streaming

With the availability of broadband networks and the increase in computing power and electronic distribution, online streaming has become more and more popular (Aljukhadar and Senecal 2017). With the availability of broadband networks and the increase in computing power and electronic distribution, online streaming has become more and more popular (Hamid et al. 2016). Online streaming media is a kind of multimedia, which is continuously received by the end user and presented to the end user when it is continuously delivered

Table 1. Online streaming technologies and applications.

Author	Streaming technology	Streaming application
Mahanti et al. (2013)	Pareto and Zipf distributions method.	Power-laws.
Huang et al. (2015)	Heterogeneous vehicular networks	Adaptive live streaming mechanism.
Thomson, Mahanti, and Gong (2017)	One-click file hosting indexes	Copyright infringement.
Farrelly et al. (2017)	Video characteristics of xHamster.	Video streaming service
Thomson, Mahanti, and Gong (2018)	One-click file hosting services (OCHs).	Internet piracy.
Song and Mahanti (2019)	Mobile streaming.	Academic Web Server.
Yang et al. (2019)	Fuzzy Petri Nets	News streams
Wong, Song, and Mahanti (2020)	Video streaming; Tagging; Data mining.	Online social networks.
Takimoto, Omorib, and Kanagawa (2021)	Multi-stream and multi-task convolutional neural networks (CNNs).	Higher-level visual information.
Adib, Mahanti, and Naha (2021)	Video streaming; Data mining; Characterization.	Active measurement.

by the provider. The term of live streaming refers to the process of the medium's delivery method, not the medium itself. In addition, live-streaming media provides an alternative method of file download, in which the end user obtains the entire file of the content before watching or listening to the content. On the contrary, live streaming is the real-time transmission of Internet content, just like real-time TV broadcasts content via airwaves through TV signals. Live streaming requires a form of source media (e.g., a video camera, an audio interface, screen capture software, etc.), an encoder to digitize the content, a media publisher, and a content delivery network to distribute and deliver the content. Live-streaming media does not need to be recorded at the point of origin, although recording is often required. However, the development of live-streaming media on the Internet presents challenges. For example, in terms of long-term behavior observation, it is difficult to investigate user behavior through data collection and analysis. In addition, it is difficult for online streaming media operators to develop flexible business alternatives based on the changing streaming media behavior of users, thus failing to produce good and profitable business models.

In addition, this paper should also look at power-law behaviors reported in the literature and how online streaming can be used for business purposes on Table 1.

By examining the illustrated literature of online streaming technologies and applications, we infer that past studies have a gap on investigating users' behavior and preference on online streaming in terms of an e-commerce business model development. This is a main motivation of this paper to develop online streaming using recommendations.

Personalized Recommendation Systems

A personalized recommendation system (PRS) uses users' interests and purchasing behavior to recommend information and products. With the expansion of e-commerce and the increase in the number of products offered online,

customers need time to find the products they want to buy (Chang and Jung 2017). Browsing a large number of irrelevant information sources and product processes means that consumers will experience information overload. Therefore, a personalized recommendation system (PRS) is proposed (Modarresi 2016). PRS is an advanced business intelligence platform that uses massive data mining to allow e-commerce websites to provide its customers with complete personalized decision support and information services (Xu et al. 2016). The recommendation system of shopping websites recommends products to customers and automatically completes the process of personalized product selection to meet the individual needs of customers (Contreras, Salamó, and Pascual 2015). The main algorithms for e-commerce recommendation systems include association rule-based recommendation (Santos et al. 2018), content-based recommendation (Zheng et al. 2018) and collaborative filtering recommendation (Akcayol et al. 2018). The biggest advantage of personalized recommendation is that it can collect user data and provide online users with personalized recommendations based on user characteristics (such as interests or preferences) (Lee and Brusilovsky 2017). Various online platforms and social networks that provide personalized services also need the support of the recommendation system (Yin et al. 2018). In an increasingly competitive environment, personalized recommendation systems can retain customers and improve the services of e-commerce systems. A successful recommendation system (RS) can use accurate big data analysis to bring huge benefits (Gao et al. 2019). In the electronic market, more and more recommendation systems are adopted to improve the pre-selection of available products and services (Hwangbo, Kim, and Cha 2018). Determining user preferences is an important condition for effective operation of these automatic recommendation systems (Dixit, Gupta., and Jain 2018). This study develops a rule-based recommendation method for investigating online streaming and purchasing behavior based on online personalized recommendations.

Rough Set Theory

Rough set theory (RST) was introduced by Pawlak in the 1980s (Pawlak 1982, 2002) as a mathematical approach to aid decision-making in the presence of uncertainty. The rough set philosophy assumes that for every object there is associated a certain amount of information (data, knowledge), expressed by means of some attributes used for the object's description. Objects having the same description are indiscernible (similar) with respect to the available information. The indiscernibility relation thus generated constitutes the mathematical basis of RST. It induces a partition of the universe into blocks of indiscernible objects, called elementary sets that can be employed to build knowledge about a real or abstract world. The use of the indiscernibility relation

results in information granulation (Greco, Matarazzo, and Slowinski 2001). In other words, when the available knowledge is employed, boundary-line cases cannot be properly classified. Therefore, rough sets can be considered as uncertain or imprecise as illustrated in the following (Liao and Chang 2016).

An attribute a is a mapping $a : U \rightarrow Va$ where U is a non-empty finite set of objects (called the universe) and Va is the value set of a. An information system is a pair $T = (U, A)$ of the universe U and a non-empty finite set A of attributes. Let B be a subset of A. The B-indiscernibility relation is defined by an equivalence relation IB on U such that $IB = \{(x, y) \in U2| \quad a \in B. a(x) = a(y)\}$. The equivalence class of IB for each object $x(\in U)$ is denoted by $[x]B$. Let X be a subset of U. We define the lower and upper approximations of X by $B(X) = \{x \in U|[x]B \subseteq X\}$ and $B(X) = \{x \in U|[x]B \cap^X \neq \emptyset\}$. A subset B of A is a reduct of T if $IB = IA$ and there is no subset B of B with $IB = IA$ (i.e., B is a minimal subset of the condition attributes without losing discernibility). A decision table is an information system $T = (U, A \cup \{d\})$ such that each $a \in A$ is a condition attribute and $d \notin A$ is a decision attribute. Let Vd be the value set $\{d1, \ldots, du\}$ of the decision attribute d. For each value $di \in Vd$, we obtain a decision class $Ui = \{x \in U|d(x) = di\}$ where $U = U1 \cup \cdots \cup U|Vd|$ and for every $x, y \in Ui, d(x) = d(y)$. The B-positive region of d is defined by $PB(d) = B(U1) \cup \cdots \cup^B(U|Vd|)$. A subset B of A is a relative reduct of T if $PB(d) = PA(d)$ and there is no subset B of B with $PB(d) = PA(d)$. We define a formula $(a1 = v1) \wedge \cdots \wedge (an = vn)$ in T (denoting the condition of a rule) where $aj \in A$ and $vj \in Vaj(1 \leq j \leq n)$. The semantics of the formula in T is defined by $[[(a1 = v1) \wedge \cdots \wedge (an = vn)]]T = \{x \in U|a1(x) = v1, \ldots, an(x) = vn\}$. Let ϕ be a formula $(a1 = v1) \wedge \cdots \wedge (an = vn)$ in T. A decision rule for T is of the form $\phi \rightarrow (d = di)$, and it is true if $[[\phi]]T \subseteq [[(d = di)]]T(= Ui)$. The accuracy and coverage of a decision rule r of the form $\phi \rightarrow (d = di)$ are respectively defined as follows:

$$accuracy(T', r, U_i) = \frac{\left|U_i \cap^{[[\varphi]]_{T'}}\right|}{\left|[[\varphi]]_{T'}\right|}$$

$$coverage(T', r, U_i) = \frac{\left|U_i \cap^{[[\varphi]]_{T'}}\right|}{|U_i|}$$

In the evaluations, $|Ui|$ is the number of objects in a decision class Ui and $|[[\phi]]T'|$ is the number of objects in the universe $U = U1 \cup \cdots \cup^U |Vd|$ that satisfy condition ϕ of rule r. Therefore, $|Ui \cap [[\phi]]T'|$ is the number of objects satisfying the condition ϕ restricted to a decision class Ui.

Database Modeling – Snowflake Schema Database

Berson et al. (1997) noted that a snowflake schema database is constructed using dimensional data tables and a fact table, and the dimensional structure of the data is presented in a star schema, a snowflake schema, and a fact constellation schema. A multi-dimensional data model is necessary to organize and store a large amount of data in a data warehouse so that queries and analysis decisions can be optimized. This study uses a snowflake schema including 2,124 data record, which is an extension of the star schema. This schema queries the data and saves storage space and is easier to maintain than other schema modes that use a relational database. The snowflake schema database is shown in Figure 1 and includes 17 fact tables, 9 relationships and 76 dimensions.

The Proposed Rough Set Association Rules Recommendation Algorithm

Ordinal Scale Data Processing – Rough Set Processing

Ordinal scale is a scale that uses labels to classify cases into ordered categories. Traditional association rules ignore the rules found in ordered data. Ordinal scale means that the classes must be arranged in order so that each case in one class is considered greater (or less than) each case in the other class. Cases of the same category are considered equivalent. Data science is a critical issue in applied artificial intelligence in terms of developing information science (Jadhav, Pramod, and Ramanathan 2019). This study combines association rules with rough sets to create an application for ordinal-scale data. This study investigates a market survey using a questionnaire to collect data. A total of 2,124 valid questionnaires were divided into six sections with 29 items in the database design. All questions were designed using ordinal scales. All items used multiple-choice questions. An example is given as follows:

Which of the following live-streaming social platforms have you used to watch a live stream?

Facebook YouTube Instagram Twitch Live.me 17 Stream PikoLive Others
(Please list your three top choices in order/_)_ / _)._
The processing of ordinal-scale data is described below.

Definition1: Transform the questionnaire answers into information system $IS = (U, A)$, where $U = \{x_1, x_2, \cdots x_i\}$ is a finite set of objects and $i = 1 \cdots n$, $A = \{a_1, a_2, \cdots a_j\}$ is a finite set of general attributes/criteria and $j = 1$. $f_a = U \times A \to V_a$ called the information function, V_a is the domain of the attribute/ criterion a, and f_a is a ordinal function set such that $f(x, a) \in V_a$ for each $x_i \in U$.
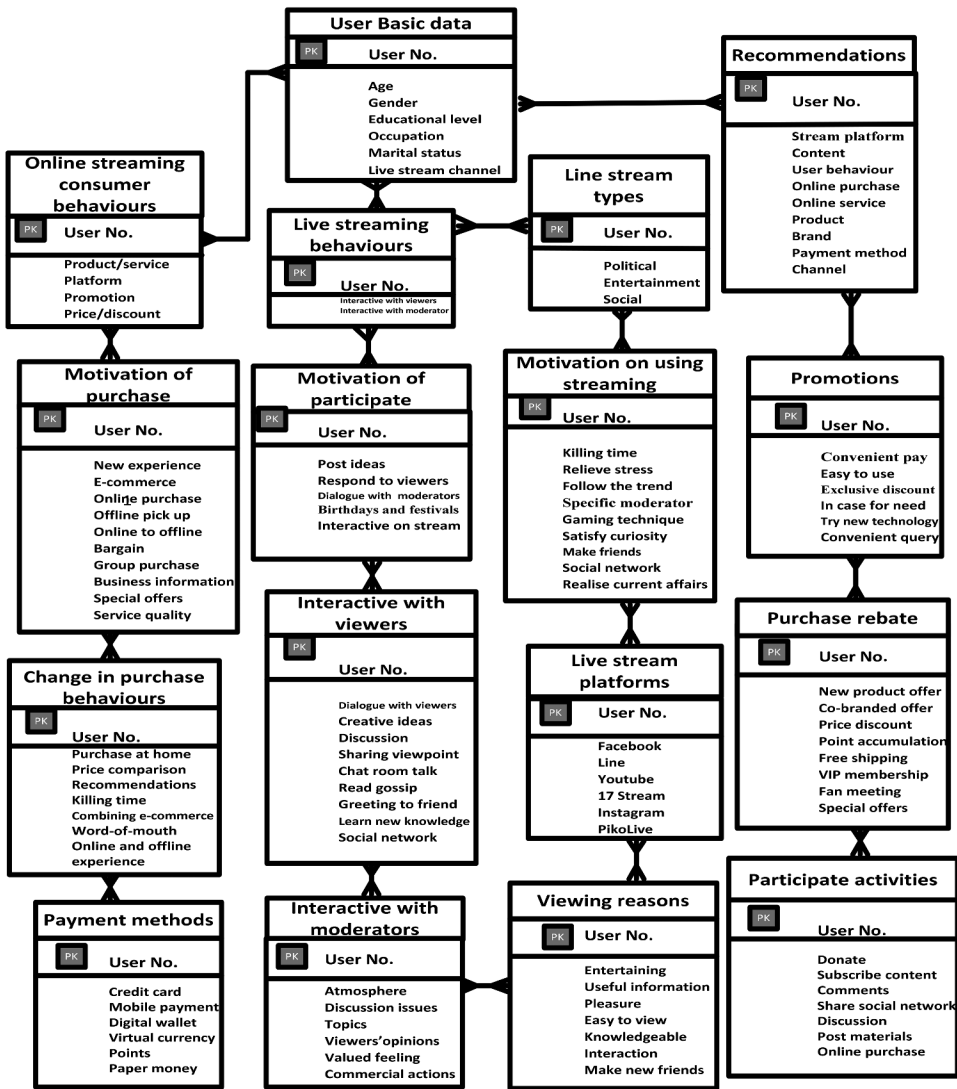
Figure 1. Database modeling – snowflake schema database.

Example: Table 2 shows the ranking of live-streaming platforms, from the first to eighth, by x_1, named Facebook, YouTube, Instagram, Twitch, Live.me, 17 Stream, PikoLive and Others.

Then:

$$f_{a_1} = \{1\}\, f_{a_2} = \{2, 3, 5, 6, 7\}\, f_{a_3} = \{2, 3, 6, 7, 8\}\, f_{a_4} = \{2, 4, 5\}$$

$$f_{a_5} = \{3, 4, 6, 7\} f_{a_6} = \{2, 5, 6\} f_{a_7} = \{3, 4, 8\} f_{a_8} = \{6, 7, 8\}$$

$$V_{a_1}^{x_1} = 1 V_{a_2}^{x_1} = 5 V_{a_3}^{x_1} = 7 V_{a_4}^{x_1} = 4$$

$$V_{a_5}^{x_1} = 3 V_{a_6}^{x_1} = 2 V_{a_7}^{x_1} = 8 V_{a_8}^{x_1} = 6$$

Table 2. Information system.

	A Ordinal Scale Data Sets							
U	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
x_1	1	5	7	4	3	2	8	6
x_2	1	6	8	2	4	5	3	7
x_3	1	7	2	4	6	5	3	8
x_4	1	2	3	5	7	6	4	8
x_5	1	3	6	2	4	5	8	7

Definition2: According to specific universe of discourse classification, a similarity relation of the general attributes $a \in A$, denoted by $\frac{U}{A}$. All of the similarity relation, denoted by $R(a_j)$.

$$\frac{U}{A} = \left\{ [x_i]_A | x_i \in U \right\}$$

Example:

$$R(a_3) = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5\}\}$$

$$R(a_5) = \{\{x_1\}, \{x_2, x_5\}, \{x_3\}, \{x_4\}\}$$

$$R(a_6) = \{\{x_1\}, \{x_2, x_3, x_5\}, \{x_4\}\}$$

$$R(a_7) = \{\{x_1, x_5\}, \{x_2, x_3\}, \{x_4\}\}$$

Definition3: The information system is composed of ordinal scale data. Therefore, an ordinal response occurs between the two attributes, where D_a presents the pair wise comparison results of ordinal scale data, which are defined as follows,

$$D_a^+ = \left\{ x_i | \frac{U}{a}, V_{a_1} > V_{a_j} \right\}$$

$$D_a^- = \left\{ x_i | \frac{U}{a}, V_{a_i} < V_{a_j} \right\}$$

Then, using the concept of similarity relation in rough set theory we find the value of the ordinal-scale data between a_i and a_j, where $ind(B)$ is the core attribute value of the ordinal-scale data in the first step, and B is the subset of A.

$$ind(B) = [f_a]_{B \subseteq A} = \bigcap_{B \subseteq A} \left[\frac{U}{a} \right]$$

Table 3. The core attribute values of the ordinal scale data for online streaming platforms.

$R(a_j)$	f_{a_5}	f_{a_6}	D_a
$\{x_1\}$	3	2	D_a^+
$\{x_2, x_5\}$	4	5	D_a^-
$\{x_3\}$	6	5	D_a^-
$\{x_4\}$	7	6	D_a^+
$R(a_j)$	f_{a_5}	f_{a_6}	D_a

Example: Example: According to the similarity relation and the fact that $R(a_5) = \{\{x_1\}, \{x_2, x_5\}, \{x_3\}, \{x_4\}\}$ and $R(a_6) = \{\{x_1\}, \{x_2, x_3, x_5\}, \{x_4\}\}$ both belong to the same fundamental set, the ordinal function set is $f_{a_5} = \{3, 4, 6, 7\}$ and $f_{a_6} = \{2, 5, 6\}$. Therefore, a_5 and a_6 are both core attribute values of the ordinal-scale data for live-streaming platforms and for user x_1, x_3 and x_4, a_5 always places after a_6, denoted by D_a^+. The pair-wise comparison of a_5 and a_6, as shown in Table 3.

$$ind(B) = [a_5, a_6]$$

Rough Set Association Rules

Definition 1: As a first step, this study identifies the core attribute values of ordinal-scale data. In this step, the object generates the rough associational rule. The consideration of other attributes and the core attributes of ordinal-scale data as the highest decision-making attributes is used to establish the decision table and to generate rules, as shown in Table 4.

$DT = (U, Q)$, where $U = \{x_1, x_2, \cdots x_i\}$ is a finite set of objects and $i = 1 \cdots n$, Q is usually divided into two parts. $G = \{g_1, g_2, \cdots g_j\}$ is a finite set of general attributes/criteria and $j = 1$, $D = \{d_1, d_2, \cdots d_l\}$ is a set of decision attributes and $k = 1 \cdots p$. $f_g = U \times G \rightarrow V_g$ is called the information function, V_g is the domain of the attribute/ criterion, g, and f_g is a total function, such that $f(x, g) \in V_g$, for each $g \in Q$; $x \in U$. $f_g = U \times G \rightarrow V_d$ is called the sorting decision-making information function, V_d is the domain of the decision attributes/criterion, d, and f_d is a total function, such that $f(x, g) \in V_d$, for each $d \in Q$; $x \in U$.

Then:
$$f_{g_1} = \{g_{1_1}, g_{1_2}\} f_{g_2} = \{g_{2_1}, g_{2_2}\}$$
$$f_{g_3} = \{g_{3_1}, g_{3_2}\} f_{g_4} = \{g_{4_1}, g_{4_2}\}$$

Table 4. Decision-making table in terms of users' preferences on online streaming platforms.

Q	General attributes				Decision attributes	
U	g_1	g_2	g_3	g_4	Stream ranking	
x_1	g_{1_1}	g_{2_1}	g_{3_1}	g_{4_1}	3	a_5
x_2	g_{1_1}	g_{2_2}	g_{3_1}	g_{4_2}	6	a_5
x_3	g_{1_2}	g_{2_1}	g_{3_2}	g_{4_1}	6	a_5
x_4	g_{1_2}	g_{2_1}	g_{3_2}	g_{4_1}	7	a_5
x_5	g_{1_1}	g_{2_2}	g_{3_1}	g_{4_2}	5	a_5

Definition2: According to the specific universe of discourse classification, a similarity relation for the general attributes is denoted by $\frac{U}{G}$. All of the similarity relations are denoted by $R(g_t)$ and t is the combination of all the general attributes.

$$R(g_t) = \frac{U}{G}\left\{[x_i]_G | x_i \in U\right\}$$

Example:

$$R_1 = \frac{U}{g_1}\left\{\{x_1, x_2, x_5\}, \{x_3, x_4\}\right\}$$

$$R_2 = \frac{U}{g_2}\left\{\{x_1, x_3, x_4\}, \{x_2, x_5\}\right\}$$

$$\vdots$$

$$R_5 = \frac{U}{g_2 g_4}\left\{\{x_1, x_3, x_4\}, \{x_2, x_5\}\right\}$$

$$\vdots$$

$$R_t = \frac{U}{g_1 g_2 g_3 g_4}\left\{\{x_1\}, \{x_2, x_5\}, \{x_3, x_4\}\right\}$$

Definition3: By the similarity relation, and determination of the reduct and core, the attribute, g, of G and the set G, which was ignored, has no effect, so g is an unnecessary attribute and can be reducted. $R \subseteq G$ and $g \in R$. A similarity relation for the general attributes of the decision table is denoted by $ind(G)$. If $ind(G) = ind(G - g_1)$, then g_1 is the reduct attribute and if $ind(G) \neq ind(G - g_1)$, then g_1 is the core attribute.

Example:

$$\frac{U}{ind(G)} = \{\{x_1\}, \{x_2, x_5\}, \{x_3, x_4\}\}$$

$$\frac{U}{ind(G - g_1)} = \frac{U}{g_2 g_3 g_4} = \{\{x_1\}, \{x_2, x_5\}, \{x_3, x_4\}\} = \frac{U}{ind(G)} = \frac{U}{g_1 g_2 g_3 g_4}$$

$$\frac{U}{ind(G - g_1 g_3)} = \frac{U}{g_2 g_4} = \{\{x_1, x_3, x_4\}, \{x_2, x_5\}\} \neq \frac{U}{ind(G)} = \frac{U}{g_1 g_2 g_3 g_4}$$

When considering g_1, alone, g_1 is the reduct attribute, but when considering g_1 and g_3, simultaneously, g_1 and g_3 are the core attributes.

Definition4: The lower approximation, denoted as $^-G(X)$, is defined as the union of all of the elementary sets that are contained in $[x_i]_G$. More formally:

$$^-G(X) = \cup \left\{ [x_i]_G \in \frac{U}{G} | [x_i]_G \subseteq X \right\}$$

The upper approximation, denoted as $\bar{G}(X)$, is the union of those elementary sets that have a non-empty intersection with $[x_i]_G$. More formally:

$$\bar{G}(X) = \cup \left\{ [x_i]_G \subseteq \frac{U}{G} | [x_i]_G \cap^X \neq \phi \right\}$$

The difference: $Bn_G(X) = \bar{G}(X) - ^-G(X)$ is called a boundary of $[x_i]_G$.

Example: $\{x_1, x_2, x_4\}$ are the customers of interest, so $^-G(X) = \{x_1\}$, $\bar{G}(X) = \{x_1, x_2, x_3, x_4, x_5\}$ and $Bn_G(X) = \{x_2, x_3, x_4, x_5\}$.

Definition5: Using the traditional association rule to calculate the value of Support, Confidence and Lift value, the formula is shown as follows:

$$Sup(ind(B)) = |\{ind(B)|^-G(X) \subseteq \bar{G}(X)\}| = \left| \frac{ind(B)|^-G(X)}{\bar{G}(X)} \right|$$

$$Conf(ind(B) \rightarrow d_{g_m}) = \left| \{ind(B) \cap^{d_{gm}} |Sup(ind(B))\} \right| = \left| \frac{Sup\left(ind(B) \cap_m^{d_g}\right)}{Sup(ind(B))} \right|$$

$$Lift = Confidence\,(X \rightarrow Y)/Sup\,(Y)$$

Definition6: Rough set association rules.

$$\frac{\{x_1\}}{g_1 g_3} : g_{1_1} \cap^{g_{3_1}} \Rightarrow d_{d_1}^1 = 4$$

$$\frac{\{x_1\}}{g_1 g_2 g_3 g_4} : g_{1_1} \cap^{g_{2_1}} \cap^{g_{3_1}} \cap^{g_{4_1}} \Rightarrow d_{d_1}^1 = 4$$

Empirical Study and Results

Pattern 1: Associations between Users' Behaviors and Live Stream Platform/content

In terms of the users' preferences to use an online stream platform and content, the types of stream platforms that are preferred by the users are determined by investigating the users' habits and behaviors, so an online stream operator can tailor information for users. This study uses a minimum antecedent support of 2% and a minimum rule confidence of 30% to generate ten meaningful association rules, as shown in Table 5 and Figure 2. The lift values are all greater than 1.

In terms of R1, the participation motivation of the users to watch a live stream platform, we found that most of the users like to interact with live streamers and viewers using the PikoLive platform. They like to donate to their favorite live stream moderator on the live-streaming platform with a hope of improving their online gaming technique. In terms of R6, the participation motivation of the users was to celebrate birthdays and festivals using the witch

Table 5. Associations of online stream platform use and content preferences.

|				Consequent	Antecedent			
Rule	Lift	Sup.	Conf.	Participation motivation	Online streaming platform	Participate in online streaming	Reason for using online streaming	Motivation for using streaming
R1	1.68	2.64	87.5	Interact with live streamers and viewers	PikoLive	Donate	Favorite live stream platform	Improved gaming technique
R2	3.11	3.97	100	Post ideas	17 Stream	Subscribe or track live shows	Entertaining of stream	Follow the trend
R3	2.28	3.3	40	Discuss various topics	Instagram	Recom-mend to friends	Entertaining stream	Kill time
R4	2.32	5.96	77.77	Dialogue with live streamers	YouTube	Participate in live community discussions	Favorite live stream platform	Kill time
R5	2.99	3.31	80,0	Respond to the viewers for discussion	Facebook	Donate	Useful information	Follow the trend
R6	3.74	10.8	100.0	Birthdays and festivals	Twitch	Participate in live community discussions	Interact with streamers and viewers	Like a specific streamer
R7	2.35	5.88	100.0	Respond to viewers' comment	KingKong	Share on social networks	Interact with streamers and viewers	Relieve stress
R8	3.28	2.62	60.0	Post ideas	YouTube	Subscribe or track live shows	Entertaining of stream	Satisfy curiosity
R9	2.54	5.43	100.0	Discuss various topics	Facebook	Donate	Good quality of information	Like a specific streamer
R10	2.23	2.71	33.33	Dialogue with live streamers	YouTube	Recom-mend to friends	Atmosphere is fun	Relieve stress
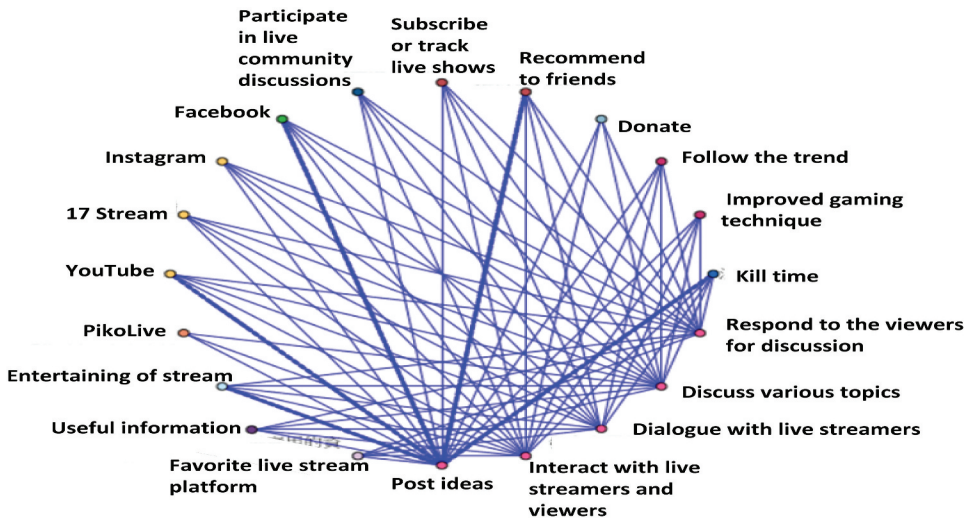
Figure 2. Associations of online stream platform and offline content.

platform. They like to participate in live community discussions and interact with streamers and viewers, spending time with a specific favorite streamer. In terms of R10, the participation motivation of the users was to engage in dialog with live streamers using the YouTube platform. They like to recommend stream content to their friends, enjoy the fun atmosphere and relieve stress.

Pattern 2: Associations between Users' Behaviors and Online Purchase Preferences

In regard to pattern 2, this study investigates live-streaming users' behaviors and their online purchase behaviors. This study uses a minimum antecedent support of 2% and a minimum rule confidence of 30% to generate nine meaningful association rules, as shown in Table 6 and Figure 3. The lift values are all greater than 1.

In terms of R1, users increase purchase intention by sharing their product/ service experience with live-streaming content. The reason for consumption of the live stream is to bargain for products and price. Reducing the offline shopping frequency is a main factor behind changing purchase behaviors in consumption patterns. In addition, users like to read messages and absorb new information in terms of stream interaction with other viewers. For R6, live shopping information increases purchase intention by solving problems through chat room talk, offers are promoted through stream shopping, special offers are provided by stream operators and users enjoy other viewers' sharing in terms of stream interaction with other viewers. On the other hand, in terms of R9, the main reason for stream users to increase purchase intention is that stream platforms can provide product classification. Also, increased viewer

Table 6. Associations of online stream use behaviors and online purchase preferences.

| Rule	Lift	Sup.	Conf.	Consequent			Antecedent	
				Increase purchase intention	Preference for chat room talk	Reason for watching a live stream	Changes in consumption patterns	Type of stream interaction
R1	1.68	2.64	87.5	Share product/ service experience	Live stream content	Bargain	Reduce offline shopping frequency	Like to read messages
R2	3.11	3.97	100	Actual clothing	Sharing experience	Combining with e-commerce	Online shopping deposit offer	Discussion and interaction on stream
R3	2.28	3.3	40	Product function and price comparison	Thematic interaction	Online to offline interaction	Reduce offline shopping frequency	Leave a message with other viewers
R4	2.32	5.96	77.77	Provide real-time shopping information	Issue discussion	Bargain	Online shopping deposit offer	Talk to the fans of the stream moderator
R5	2.99	3.31	80,0	Platform product classification	Issue discussion	Online to offline interaction	Discounts on various payment methods	Enjoy other viewers' sharing
R6	3.74	10.8	100.0	Provide real-time shopping information	Specific problem solving	Promotion-al offers on stream shopping	Special offers by stream operator	Enjoy other viewers' sharing
R7	2.35	5.88	100.0	Share product/ service experience	Share experience	Bargain	Reduce offline shopping frequency	Like to read messages
R8	3.28	2.62	60.0	Product function and price comparison	Thematic interaction	Combining with e-commerce	Online shopping deposit offer	Discussion and interaction on stream
R9	2.54	5.43	100.0	Platform product classification	Increase viewers' intention	Online to offline interaction	Discounts on various payment methods	Talk to the fans of the stream moderator
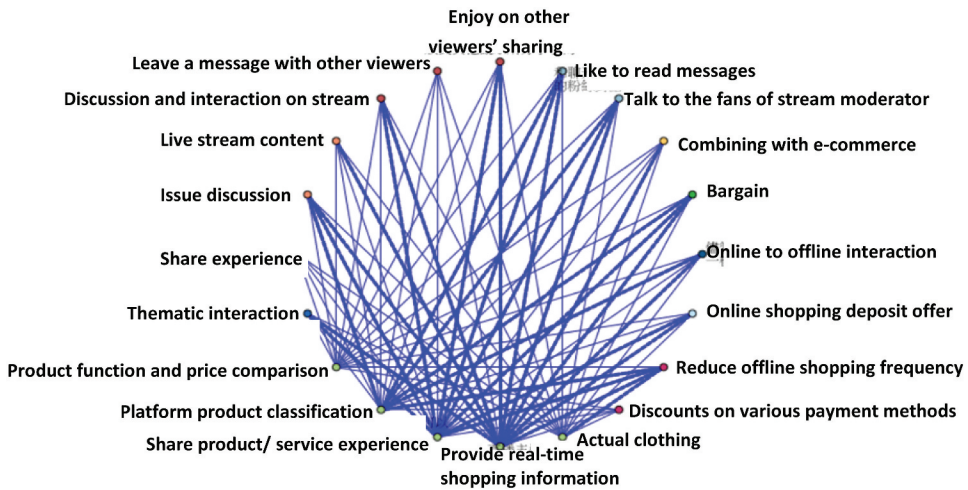
Figure 3. Associations of online stream use behaviors and online purchase preferences.

intentions, online to offline interaction, discounts on various payment methods and stream moderators talking to the fans will affect users' online purchase preferences.

Pattern 3: Associations between Online Stream Services and Online Recommendations

When it comes to online stream services and online recommendations, the content of stream services that are preferred by the users are determined by investigating the users' online streaming and purchase behaviors, so that an online stream operator can provide online recommendations to users so as to improve stream service and create online business opportunities. This study uses a minimum antecedent support of 1% and a minimum rule confidence of 30% to generate 10 meaningful association rules, as shown in Table 7 and Figure 4. The lift values are all greater than 1.

In terms of R2, live stream content should include differentiation of features and themes for. Value VIP users, daily chat, more product introductions and offering discounts are the antecedents in terms of associations of the rule the online stream services and online recommendations. For R4, organizing events, commenting on current affairs, make a good impression of the product and selecting the most sponsored fans are associated with improved message speed and social interaction. On the other hand, in terms of R9, increased placement marketing can incorporate events of new product development, comments on current affairs, recommending items trusted by users and discount offers by operators in terms of implementing online stream service and online recommendations to online streaming operators.

Table 7. Associations of online stream service and online recommendations.

| Rule	Lift	Sup.	Conf.	Consequent			Antecedent	
				Suggestions for live content	Promotion-al project preference	Increase stream viewing willingness	Placement marketing model	Consumer feedback preference
R1	3.51	3.97	100	Increase live content richness	Platform product discounts	Comment on current affairs	More product introductions	Shortening the distance
R2	1.94	2.32	100	Differentiation of features and themes	Value VIP users	Daily chat	More product introductions	Operators offer discounts
R3	4.08	3.97	100	Increase network stability	Keep creating topics	Comment on current affairs	Innovative live streams of businesses	Contests on platforms
R4	2.46	3.97	100	Improve message speed and social interaction	Organize events	Comment on current affairs	Make a good impression on the product	Select the most sponsored fans
R5	1.86	2.65	50	Increase placement marketing	Platform product discounts	Share novel events	Innovative live streams of businesses	Shortening the distance
R6	3.19	7.69	76.47	Differentiation of features and themes	Value VIP users	Daily chat	Make a good impression of the product	Operators offer discounts
R7	1.40	2.26	40	Improve message speed and social interaction	Theme innovation	Looking for viewer communication	More product introduction	Shortening the distance
R8	2.08	2.26	80	Increase network stability	Value VIP users	Share novel events	Recommend-ed items are trusted by users	Shortening the distance
R9	2.66	4.98	100	Increase placement marketing	New product development	Comment on current affairs	Recommend-ed items are trusted by users	Operators offer discounts
R10	2.64	7.64	81.25	Increase live content richness	Theme innovation	Looking for viewer communication	Innovative live stream on business	Operators offer discounts
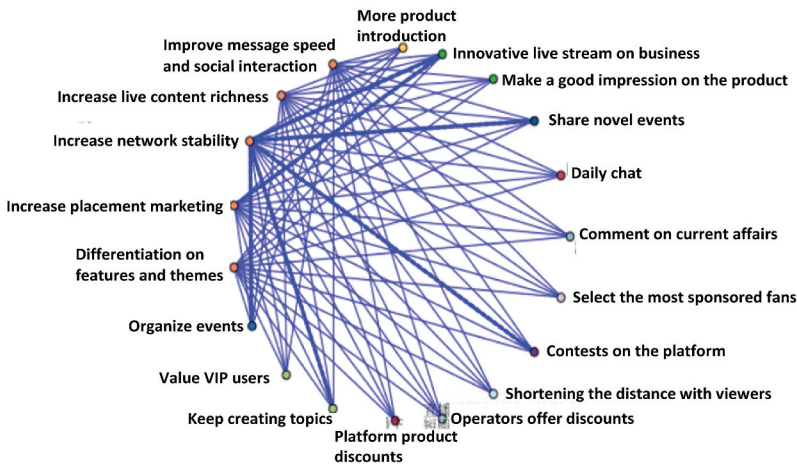
Figure 4. Associations of online stream service and online recommendations.

Implications

Live Streaming Platform/content Recommendations

In regard to platform recommendations, this study suggests that live-streaming platform operators can use the advantages of hardware facilities or featured live streamers to attract viewers. For example: 17 stream broadcasts have strong foundations of celebrity endorsements and beautiful live broadcasts to viewers. These broadcasts are dedicated to content production and interaction on the live-streaming platform, and thus have stable viewing groups. In regard to content recommendations, this study suggests that streaming platform operators or moderators should pay attention to the streaming program quality and content so that viewers can increase the exposure of the platform by sharing stream content on social networking sites. As for social media/network recommendations, this study suggests that streaming platform operators or moderators can try to provide current topics and issues that are popular in online networks. Instagram, for example, allows users to receive more timely information in the social media/network through live streaming. Furthermore, Facebook Live Stream does not only provide a streaming platform for online marketing and sales, but also develops membership and fans by expanding social media/network influence through online recommendations.

User's Behaviors and Online Purchase Recommendations

In regard to users' behavior recommendations, this study suggests that stream platform operators or moderators can provide users more integrated streaming activities by encouraging more live-streaming participation. Under the trend of live-streaming platforms that combine e-commerce activities, the

current AI technology will be used to classify consumer behaviors and stream viewing habits in order to effectively improve user behavior when using online streaming. As for online purchasing recommendations, this study suggests that streaming platforms and e-commerce operators target specific user groups and increase the exposure of online shopping brands and products. In addition, through e-commerce combined with live streaming, online events can be held and promotional discounts can be given to stream viewers, thereby increasing the exposure period of products and achieving the goal of advertising. Moreover, this study suggests that streaming and e-commerce operators can offer preferential promotions to increase users' intentions to purchase online. This can make up for the shortcomings of TV shopping and e-commerce, which fail to provide human interaction and physical service.

Online Stream Service Recommendations

As for online service recommendations, this study suggests that streaming platforms and e-commerce operators both to launch discount activities specific to products/services when selling placement products to stimulate buying momentum and contemplate new marketing techniques. Furthermore, operators can encourage stream users to integrate their consumer behaviors into the interaction of online and offline purchase services, in addition to increasing the loyalty of the fans to the platform by products/services recommendations, which can turn the streaming actions of fans into actual purchasing g behaviors. On the other hand, online live-streaming messages that promote consumer loyalty can also be sent using the online to offline business model. Restaurants and food outlets might allow reservations online and provide the final service offline by using streaming platforms to complete an online to offline business model. Thus, online to offline could be a business model that allows retailers to build on a transactional platform that co-operates with online live-streaming operators.

Conclusion and Future Work

The gift and sponsorship-streaming business model refers to the content provided by the live streamer, who broadcasts the content to the user through the live-streaming platform, and provides the user with the stored value to purchase virtual gifts from the live streamer or Internet celebrity, or provide services, such as higher viewing quality, extra features, and opportunities to interact with live streaming. In addition, one type of live streaming platform allows users to pay directly for sponsorship through the live streaming of a game, or to donate money to a fundraising proposal. The e-commerce and shopping guide streaming business model adopts a broader definition. It refers to a business model in which the live-

streaming platform provides streaming of the main merchandise to users, allowing the users to directly consume via live streaming. This business model is also integrated with e-commerce platforms. Through the live guide who is live streaming, users are directed to an e-commerce platform to purchase goods online, and e-commerce platforms are utilized in this process. In terms of future works, this study considers that different data mining methods might be investigate the online streaming issues on the further research for obtaining diversified knowledge and application results.

Disclosure Statement

Funding

ORCID

Shu-Hsien Liao (iD) http://orcid.org/0000-0002-3667-2867

References

Adib, S. A., A. Mahanti, and R. K. Naha. 2021. Characterisation and comparative analysis of thematic video portals. *Technology in Society* 4:101690. doi:10.1016/j.techsoc.2021.101690.

Akcayol, M. A., A. Utku, E. Aydoğan, and B. Mutlu. 2018. A weighted multi-attribute-based recommender system using extended user behavior analysis. *Electronic Commerce Research and Applications* 28 (3):86–93. doi:10.1016/j.elerap.2018.01.013.

Aljukhadar, M., and S. Senecal. 2017. Communicating online information via streaming video: The role of user goal. *Online Information Review* 41 (3):378–97. doi:10.1108/OIR-06-2016-0152.

Anderson, K. 2017. Getting acquainted with social networks and apps: Periscope up! Broadcasting via live streaming. *Library Hi Tech News* 34 (3):16–20. doi:10.1108/LHTN-03-2017-0015.

Berson, A., Dayal, U., & Smith, S.J. 1997. Data wsarehousing, Data mining, and OLAP. McGraw-Hill.

Bilal, K., A. Erbad, and M. Hefeeda. 2018. QoE-aware distributed cloud-based live streaming of multi-sourced multi-view videos. *Journal of Network and Computer Applications* 120 (2):130–44. doi:10.1016/j.jnca.2018.07.012.

Chang, W. L., and C. F. Jung. 2017. A hybrid approach for personalized service staff recommendation. *Information Systems Frontiers* 19 (1):149–63. doi:10.1007/s10796-015-9597-7.

Contreras, D., M. Salamó, and J. Pascual. 2015. A web-based environment to support online and collaborative group recommendation scenarios. *Applied Artificial Intelligence* 29 (5):480–99. doi:10.1080/08839514.2015.1026661.

Dixit, V. S., S. Gupta., and P. Jain. 2018. A propound hybrid approach for personalized online product recommendations. *Applied Artificial Intelligence* 32 (9–10):785–801. doi:10.1080/08839514.2018.1508773.

Farrelly, B., Y. Sun, A. Mahanti, and M. Gong, 2017. Video workload characteristics of online porn: Perspectives from a major video streaming service. 2017 IEEE 42nd Conference on Local Computer Networks (LCN), Singapore, 518–19.

Gao, M., B. Ling, L. Yang, J. Wen, Q. Xiong, and S. Li. 2019. From similarity perspective: A robust collaborative filtering approach for service recommendations. *Frontiers of Computer Science* 13 (2):231–46. doi:10.1007/s11704-017-6566-y.

Greco, S., B. Matarazzo, and R. Slowinski. 2001. Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research* 129 (1):1–47. doi:10.1016/S0377-2217(00)00167-3.

Hamid, S., S. Bukhari, D. Ravana, A. A. Norman, and M. T. Ijab. 2016. Role of social media in information-seeking behaviour of international students: A systematic literature review. *Aslib Journal of Information Management* 68 (5):643–66. doi:10.1108/AJIM-03-2016-0031.

Huang, C. J., S. C. Chang, C. T. Guan, Y. W. Wang, H. M. Chen, C. H. Weng, and S. Y. Chien. 2015. An adaptive live streaming mechanism for heterogeneous vehicular networks. *Applied Artificial Intelligence* 29 (7):637–59. doi:10.1080/08839514.2015.1051885.

Hwangbo, H., Y. S. Kim, and K. J. Cha. 2018. Recommendation system development for fashion retail e-commerce. *Electronic Commerce Research and Applications* 28 (2):94–101. doi:10.1016/j.elerap.2018.01.012.

Jadhav, A., D. Pramod, and K. Ramanathan. 2019. Comparison of performance of data imputation methods for numeric dataset. *Applied Artificial Intelligence* 33 (10):913–33. doi:10.1080/08839514.2019.1637138.

Lee, D. H., and P. Brusilovsky. 2017. Improving personalized recommendations using community membership information. *Information Processing & Management* 53 (5):1201–14. doi:10.1016/j.ipm.2017.05.005.

Liao, S. H., and H. K. Chang. 2016. A rough set-based association rule approach for a recommendation system for online consumers. *Information Processing & Management* 52 (6):1142–60. doi:10.1016/j.ipm.2016.05.003.

Lim, C., and K. J. Kim. 2018. Experience design board: A tool for visualizing and designing experience-centric service delivery processes. *Journal of Retailing and Consumer Services* 45 (2):142–51. doi:10.1016/j.jretconser.2018.07.021.

Mahanti, A., N. Carlsson, A. Mahanti, M. Arlitt, and C. Williamson. 2013. A tale of the tails: Power-laws in internet measurements. *IEEE Network* 27 (1):59–64. doi:10.1109/MNET.2013.6423193.

Modarresi, K. 2016. Recommendation system based on complete personalization. *Procedia Computer Science* 80 (6):2190–204. doi:10.1016/j.procs.2016.05.379.

Pawlak, Z. 1982. Rough sets. *International Journal of Computer and Information Sciences* 11 (3):341–56. doi:10.1007/BF01001956.

Pawlak, Z. 2002. Rough sets and intelligent data analysis. *Information Sciences* 147 (1):1–12. doi:10.1016/S0020-0255(02)00197-4.

Santos, F. F., M. A. Domingues, C. V. Sundermann, V. O. Carvalho, S. O. Rezende, and S. O. Rezende. 2018. Latent association rule cluster based model to extract topics for classification and recommendation applications. *Expert Systems with Applications* 112 (10):34–60. doi:10.1016/j.eswa.2018.06.021.

Singh, C., and A. Sharma. 2020. Online learning using multiple times weight updating. *Applied Artificial Intelligence* 34 (6):515–36. doi:10.1080/08839514.2020.1730623.

Smink, A. R., S. Frowijn, E. A. van Reijmersdal, G. van Noort, and P. C. Neijens. 2019. Try online before you buy: How does shopping with augmented reality affect brand responses and personal data disclosure. *Electronic Commerce Research and Applications* 35 (3):100854. doi:10.1016/j.elerap.2019.100854.

Song, Y. D., and A. Mahanti, 2019. Comparison of mobile and fixed device workloads in an academic web server. 2019 Network Traffic Measurement and Analysis Conference (TMA), France, 49–56.

Taiwan Market Intelligence & Consulting Institute (Taiwan MIC). 2019. 2018 Taiwan online streaming market survey. (30 Jun), available at: https://mic.iii.org.tw/Industryobservation_MIC01views.aspx

Takimoto, H., F. Omorib, and A. Kanagawa. 2021. Image aesthetics assessment based on multi-stream CNN architecture and saliency features. *Applied Artificial Intelligence* 35 (1):25–40. doi:10.1080/08839514.2020.1839197.

Thomson, W., A. Mahanti, and M. Gong. 2017. Understanding uploader motivations and sharing dynamics in the one-click hosting ecosystem. 2017 IEEE 42nd Conference on Local Computer Networks (LCN), Singapore, 520–22.

Thomson, W., A. Mahanti, and M. Gong. 2018. Uploader motivations and consumer dynamics in the one-click file hosting ecosystem. 2018 IEEE International Conference on Communications (ICC), Kansas City, USA, 1–7.

Wang, Y. 2019. User experiences in live video streaming: A netnography analysis. *Internet Research* 29 (4):638–58. doi:10.1108/IntR-01-2018-0029.

Wong, C., Y. D. Song, and A. Mahanti. 2020. YouTube of porn: Longitudinal measurement, analysis, and characterization of a large porn streaming service. *Social Network Analysis and Mining* 10 (1):62. doi:10.1007/s13278-020-00661-8.

Xu, W., J. Sun, J. Ma, and W. Du. 2016. A personalized information recommendation system for R&D project opportunity finding in big data contexts. *Journal of Network and Computer Applications* 59 (1):362–69. doi:10.1016/j.jnca.2015.01.003.

Yang, S. H., Y. N. Lin, G. J. Chiou, M. K. Chen, R. L. Shen, and H. Y. Tseng. 2019. Novel shot boundary detection in news streams based on fuzzy petri nets. *Applied Artificial Intelligence* 33 (12):1035–57. doi:10.1080/08839514.2019.1661118.

Yin, C., Y. Guo, J. Yang, and X. Ren. 2018. A new recommendation system on the basis of consumer initiative decision based on an associative classification approach. *Industrial Management & Data Systems* 118 (1):188–203. doi:10.1108/IMDS-02-2017-0057.

Zheng, X., Y. Luo, L. Sun, X. Ding, and J. Zhang. 2018. A novel social network hybrid recommender system based on hypergraph topologic structure. *World Wide Web* 21 (4):985–1013. doi:10.1007/s11280-017-0494-5.

Appendix – The algorithm

Algorithm: Rough set association rules for online streaming recommendation

Input:

Onbline Streaming Categories of Information System (MCIS), $MCIS = (U, M)$; user's contents x_k, attribute contents of merchandise category m_j;

Output

Online Streaming Purchase Recommendation (OSPR);

{Online Streaming Purchase Recommendations};

Method:

(1) Begin
(2) $MCIS = (U, M)$;
(3) **for** $k \leftarrow 1$ **to** n **do;**
(4) **if** $x_1, x_2, \ldots, x_k \in U$ **then**
(5) **for** $j \leftarrow 1$ **to** l **do;**
(6) **if** $m \sim U \rightarrow V_m, m \in M$ **then** $f_m = U \times M \rightarrow V_m$;
(7) compute $f(x, m)$, where the merchandise categories of information function in

    $MCIS$ as described in Step 1;

(1) **end if**
(2) **end for**
(3) **end if**
(4) **end for**
(5) $K = (U, B_1, B_2, \ldots, B_w)$;
(6) **for** $w \leftarrow 1$ **to** $m - 1$ **do;**
(7) **if** $B \subseteq M$ **then** $U/IND(B) = \{m \in B \sim U/IND(\{m\})\}$;
(8) $B(m_j) = \{(m_i, m_j) \in U^2 | \in B, m(m_i) = m(m_j)\}$;
(9) compute $B(m_j)$, where the indiscernibility relation in $MCIS$ as described in Step 2;
(10) **end if**
(11) **end for**
(12) **while** $IND(B) = {}^U/_M$ **do;**
(13) **if** $D_m^+ \equiv D_m^<$ or $D_m^- \equiv D_m^>$ **then**
(14) $IND(B) = \{[x_k]_m \in U^2 | \in B, m(m_i) = m(m_j)\}$;
(15) compute D_m, where the condition attributes in $B(m_j)$ as described in Step 3;
(16) **end if**
(17) Find the online streaming categories average weight set based on AHP as described in Step 4;
(18) $DR = (Y_l, \hat{\beta}_k MC_i)$;
(19) **for** $l \leftarrow 1$ **to** n **do;**
(20) **if** $Y_1, Y_2, \ldots, Y_l \in Y$ **then**
(21) **for** $k \leftarrow 1$ **to** 5 **do;**
(22) **if** $\hat{\beta}_k = \frac{1}{n} \sum MC_i$ **then**
(23) compute $\hat{\beta}_k$, where the average weights of online streaming categories value set as

    described in Step 5;

(1) **end if**
(2) **end for**
(3) **end if**
(4) **end for**
(5) $DT = (U, M = E \cup^D)$;
(6) **for** $g \leftarrow 1$ **to** v **do;**
(7) **if** $e \sim U \rightarrow V_e, e \in E$ **then** $f_e = U \times E \rightarrow V_e$;
(8) compute $f(x, e)$, where the information function in DT as described in Step 6;
(9) **end if**
(10) **end for**
(11) $K = (U, B_1, B_2, \ldots, B_w)$;
(12) **for** $w \leftarrow 1$ **to** $m - 1$ **do;**
(13) **if** $B \subseteq E$ **then** $IND(B) = \{[x_k]_e \in U^2 | \ e \in B, e(e_g) = e(e_v)\}$;

(14) $U/IND(B) = \{e \in B \sim U/IND(\{e\})\}$;

(15) compute $B(E_g)$, where the indiscernibility relation in DT as described in Step 7;

(16) compute $IND(E)$, where the relative reduct of DT as described in Step 8;

(17) compute $IND(E - e_g)$, where the relative reduct of the elements for element g as described in Step 8;

(18) **end if**

(19) **end for**

(20) **for each** $[x_k]_e$ **do;**

(21) **if** $x \subseteq U$ **then**

(22) $^-EX = \{x \in U | [x_k]_e \subseteq X\}$;

(23) $\bar{E}X = \{x \in U | [x_k]_e \cap^X \neq \phi\}$;

(24) compute ^-EX, where the lower-approximation of DT as described in Step 9;

(25) compute $\bar{E}X$, where the upper-approximation of DT as described in Step 9;

(26) compute $BND_e(D)$, where the bound of DT as described in Step 9;

(27) compute $Sup(IND(B))$, where the support as described in Step 10;

(28) compute $Conf\left(IND(B) \to d_{e_g}\right)$, where the confidence as described in Step 10;

(29) **end if**

(30) **end for**

(31) Output {*Online Streaming Purchase Recommendations*};

End