Optimal Recommendation and Long-tail Provision Strategies for Content Monetization

Ting-Kai Hwang
Department of Journalism
Ming Chuan University, Taiwan
tkhwang@mail.mcu.edu.tw

Yung-Ming Li
Institute of Information Management
National Chiao Tung University, Taiwan
yml@mail.nctu.edu.tw

Abstract
This paper examines the optimal strategies for pricing, contents variety supply and recommendation system investment by digital contents providers. With the fast development of digitalization technology and social participation in recent years, the ways to create and access information contents become diverse with greater convenience and much lower cost. How to attract more customers of different segments and raise sales revenue becomes the most essential issue for content providers as the long tail phenomenon becomes significant. From the supply side, increasing and maintaining a wide variety of content can attract more users. From the demand side, adapting suitable recommender systems is considered as an effective implementation for content sale promotion. However, they both require the providers to make more efforts on information acquisition and balancing the budget allocated on various types of recommender systems, which leads to differentiated changes of sales patterns. In this paper, we propose an economic model to capture the technological and market factors affecting the categorization of sales pattern and develop the proper business strategies of content provision and content recommendation for supporting the operations of digital content providers.

1. Introduction

Due to the fast development of digitalization technology and internet, the ways to access information contents become diverse with greater convenience and much lower cost. Without the restrictions of physical space, delivery and inventory, consumers are more self-conscious to evaluate and obtain the expanded varieties of digital contents. Under the interactive network structure of web 2.0, user participation and content provision become inter-dependent. Content providers are therefore encouraging or supporting users to create valuable user-generated content [17, 23]. As consumers become with higher differentiated preferences and have more freedom of content access, the consumption patterns become diversified. This induces the significance of the long-tail phenomenon, which describes a kind of customer demand pattern for personalized service or niche products (tail-products). Anderson [2] claims that in a market with an amount of product varieties, the aggregation of small niche markets can compete against a market of mass common/popular hit-products. In order to attract more customers of different segments and raise sales revenue, increasing and maintaining a wide variety of products becomes an important business strategy for content providers [4]. In the supply side, besides continuously providing professional quality content to fit the common demand, it often gets along with crossover cooperation or utilization of valuable user-generated content to offer customers more varieties of content produced at lower cost [8]. Users expect a wide variety and mass amount of content environment, but they have to make more efforts for information acquisition. The products sold in the niche market are generally not easy to be found by the public. Therefore, besides providing a greater variety of content, it is necessary for content providers to help users to reduce the search costs of the product information. Adapting suitable recommender systems is considered as an effective implementation [5, 6, 18]. In addition, it was shown that recommendation systems can reshape the customers’ purchasing pattern [4]. As illustrated in the Figure 1, the x-axis denotes the popularity rank of content varieties and the y-axis represents the popularity. The content with higher popularity has a smaller rank, and the sales pattern of content is segmented to hit products and niche products. With recommender systems, the shape of content sale distribution is changed and the sale is improved [25]. However, as shown in the Figure 2, for a saturated market, the substitution effect may occur between hit products and niche products. Such effect may also lead to the reduction in the varieties of content [11, 12]. Since different types of digital content (e.g. movie,
music, news, or app software) have some distinct characteristics, and a hit product market and a niche market can be coexist, only focusing on the development of hit market but ignoring provision tail-end products may instead result in the shrink of profit [10].

Recommender systems usually adopt user’s profile information to suggest suitable (hit or niche) products to appropriate consumers [1]. Theoretically, recommender systems are basically divided to content-based and collaborative filtering-based systems. In practice, a content-based system assists customers to discover products of personal preference, which can be viewed as a kind of personalized recommender, which help increasing sales of niche products [5, 11]. For the collaborative filtering-based system, it captures the intentions of following social behaviors for some customers, which is an extension of social recommender application with the promotion effect of hit products [24]. For example, the listing the most popular items on a webpage can be treated as a kind of social recommender application.

To improve the profitability, from the supply side, content providers have to consider the issues of optimal strategies for provisioning product varieties and balancing the effect of professional content and user-generated content for different customer segments. From the demand side, the budget allocation on various types of recommender systems leads to differentiated changes of sales patterns. In addition, the impact of changing the customers’ purchasing pattern can be further explored to develop content provision and pricing strategies [7]. Indeed, digital content is a kind of content-based digital information goods. The key characteristics of these information goods are with low or zero reproduction marginal cost and distribution cost. Since the features of digital content are different from the traditional industrial goods, it is inappropriate for content providers to follow the traditional cost-based pricing strategy [14]. To capture the appropriate categorization of sales pattern and develop the proper business strategies, the potentially complex relationships between the demand and supply side factors should be clarified.

In this paper, we propose an economic model based on the integrated views of the supply and demand sides for digital content business operations. The business strategies including content pricing, provision (volume of varieties), and recommendation (social recommender and personalized recommender) are analyzed. The results derived from the proposed model can provide useful insights for supporting the content provider in developing appropriate business models of profitability enhancement.

The rest of this paper is organized as follows. In section 2, we review previous literature related to the current research. In section 3, we propose an economic model for digital content business operations and examine the business strategies (content pricing, provision, and recommendation) under different market structures. Finally, section 4 concludes our findings and discusses future research directions.

2. Literature Review

This section reviews the literature related to this research, including long tail theory, recommendation mechanism, and content pricing strategies.

2.1. Long-tail Theory

Lots of researches have been conducted on examining the importance of long-tail phenomenon for online commerce and its influence from both the supply (producers/retailers) and demand (consumers)
sides. In particular, the long-tail phenomenon has impacts on affecting the varieties of products and change of demand distribution [4-6]. Zhou & Duan [28] studied the impact of product variety on long-tail effects. They found the influences of online user reviews are decreased by the greater number of product varieties. In the online and transparency environment, it is easy for customers to find out a lower-priced product offered by a competing company. To avoid price comparisons and enhance consumer repeat patronage, Clemons & Nunes [9] suggested resonating with customers to offer what they really need. Porcel et al. [24] pointed out that provision of niche products can help to increase consumer satisfaction and create customer loyalty. Nevertheless, Elberse [10] found that the profitability from the long-tail is not guaranteed, and long-tail and hit-product markets can coexist for various consumption patterns. Lee et al. [18] noted the impact of electronic word-of-mouth (e-WOM) for online commerce would either promote the long-tail phenomenon or facilitate the sales of the hit products based on different types of e-WOM mechanism. To promote consumer demand distribution and improve the profitability, many studies have shifted the focus on long-tail phenomenon to both long-tail and hit-product market with adoption of recommendation mechanism.

2.2. Recommendation Mechanism

Adoption of recommendation mechanism is a key factor affecting the demand of digital content market. For example, Zhong & Michahelles [27] used empirical analysis to show that without the activation of recommendation features, Google Play app market is strongly dominated by hit products. Felder & Hosanagar [11] adopted a simulation which indicates that recommendations on hit products can affect the diversity of sales. Yin et al. [25] pointed out that successful recommendation of niche items to the interest of right users can expand the niche market and also boost hit-product sales. Various types of recommender systems have been developed to deal with long-tail phenomenon of making personalized recommendations [22, 25], or to raise the concentration of hit-product sales for social recommendations [11, 20]. To provide a better user experience for consumers with differentiated preferences, Jambor & Wang [13] pointed out the importance of developing a recommender system with flexible frameworks and multiple goals. Furthermore, Oestreicher-Singer & Sundararajan [21] showed that individual product demand and product sales pattern can be reflected by recommendations.

2.3. Content Pricing

Pricing is another important issue of business strategy for digital content enterprises. Lang & Vragov [17] compared pricing schemes and profits for distributing digital content over client-server and decentralized networks and showed that decentralized networks are more profitable by introducing incentives to users for content distributing. Chellappa & Kumar [7] examined pricing and customer retention strategies of online retailers with free product-augmenting services. They showed the implementation of efficient search service, which cannot be replicated easily by other competing retailers, can help online retailers to raise prices. Kannan, Pope, & Jain [15] implemented a demand model for content providers. Through examining optimal pricing strategies for the bundling of digital and printing content formats, they find bundle discounts can increase profitability. As digital content is tied with digital device, Yu, Hu, & Fan [26] analyzed the effect of content and device pricing policy on firm profits. In addition, the impact of product differentiation on pricing strategy is investigated. Bhargava and Choudhary [3] found that for information goods, price differentiation strategy is not optimal when the highest quality product has the best value-to-cost ratios. As the variance in the utility of non-price-related features is greater than that in the utility of price [19] and it becomes inconvenient and costly for setting differentiated prices as the number of content varieties largely increases, many online content providers choose the approach of uniform pricing. In practice, a popular example is Apple’s iTunes music store sold digital content at a single price of 99 cents for songs in the early phase of the service. By keeping a uniform and competitive price, the iTunes music store brought the principles of simplicity and easy-to-understand for consumers. It induces the increasing of consumers’ purchasing activities when people were just getting used to buy the on-line digital content. Therefore, a uniform price strategy is adopted in the proposed model.

3. The Model

We consider a content provider which retails or rents digital content (e.g. news, music, and movie). There are totally $n$ varieties of digital content available on the online channel or e-store. The sale amounts of various varieties are heterogeneous and following a power-law-like distribution. Each variety of content has a popularity rank $x$, $0 \leq x \leq n-1=m$ (the content with the highest popularity has rank 0). Denote the unit price of a content variety as $p$. The demand function of
content \( x \) is described as \( \lambda(x) = a\left(\lambda_0 - \beta x^\varepsilon\right) - \rho \), where \( \lambda_0 \) is a demand base coefficient; \( a > 0 \) is the network influence parameter, which can be improved by social recommender systems and \( \beta > 0 \) is the content fitness parameter, which can be enhanced by personalized recommender systems. The demand of a typical content variety increases as the recommender systems are exploited. Thus, \( \partial\lambda(x)/\partial x > 0 \), \( \partial\lambda(x)/\partial \varepsilon < 0 \). As the \( \alpha \) increases, the sale of hit content will increase more and the content sale distribution becomes more asymmetric (\( \partial^2\lambda(x)/\partial x^2 \varepsilon < 0 \)). As the \( \beta \) decreases, the sale of niche content will increase and the content sale distribution becomes more symmetric and balanced. Parameter \( \varepsilon \geq 0 \) is used to represent the significance level of sale decreasing on the content ranking (\( \partial\lambda(x)/\partial \varepsilon < 0 \)). Therefore, the shape of sale distribution becomes more asymmetric as \( \varepsilon \) becomes larger.

The total content sales are obtained as \( \Lambda(m) = \int_0^m \lambda(x)dx \) and the total revenue collected is written as \( p\Lambda(m) \). The cost structure of a content provider includes two parts. The first part is the cost of content provision. We assume the content provision cost \( C(m) \) is majorly determined by the digital content creation or acquisition and neglects the marginal cost. Due to the extra effort of content quality and marketing, the average cost for a variety of content with a higher popularity is not smaller than the average cost for a content variety with a lower popularity (\( \partial C(m)/\partial m \varepsilon \leq 0 \)). The second part is the technology investment required to develop the search tools or recommender systems to help customers find desirable content and enhance the content sale. The technology cost incurred in the development of a recommender system is denoted as \( R(\alpha, \beta) \). A recommender system includes the techniques for utilizing the power of social influence (\( \alpha \)) and mining the individual preference (\( \beta \)). Therefore, we have the properties \( \partial R(\alpha, \beta)/\partial \alpha > 0 \) and \( \partial R(\alpha, \beta)/\partial \beta < 0 \).

The profit function of the content provider is written as \( \pi = p\Lambda(m) - R(\alpha, \beta) - C(m) \), where

\[
\Lambda(m) = a\left(\lambda_0 m - \beta m^{\varepsilon+1}\right) - \rho m.
\]

The optimal content price is obtained by solving \( \partial\pi/\partial p = 0 \). We have

\[
p^* = \frac{\alpha}{2} \left(\lambda_0 - \frac{\beta}{\varepsilon+1} m^\varepsilon\right).
\]

The resulting profit function can be rewritten as

\[
\pi^* = \frac{\alpha^2 m}{4} \left(\lambda_0 - \frac{\beta}{\varepsilon+1} m^\varepsilon\right)^2 - R(\alpha, \beta) - C(m).
\]

According to (2), to generate non-negative revenue, the maximal number of content varieties provided is \( m_{\max} \leq \left((\varepsilon+1)\lambda_0 / \beta\right)^{1/\varepsilon} \). Besides, we also have the following observation:

**Result 1.** The optimal content price decreases with the number of content varieties. However the impact of the number of content varieties on the revenue generation is non-monotonic.

### 3.1. Optimal Content Provision

In the subsection, we analyze the issue of content provision decision - determining the optimal tail length. One main benefit of the niche contents is the lower provision cost, compared with the hit contents. The lower cost could be attributed to the lower acquisition cost (e.g. old, delayed content or user generated content) or the lower effort of promotion. Assume the content with a popularity rank higher than \( \kappa \) (i.e. \( x \leq \kappa \)) is classified as a hit content, while the one with a popularity rank lower than \( \kappa \) (i.e. \( x > \kappa \)) is classified as a niche content. The average provision cost for hit and niche contents is \( c_1 \) and \( c_2 \) respectively (\( c_1 > c_2 \)). The total cost of content provision is

\[
C(m) = \min(m, \kappa) c_1 + (m - \kappa)^+ c_2,
\]

where

\[
(m - \kappa)^+ = \begin{cases} (m - \kappa), & \text{if } (m - \kappa) \geq 0 \\ 0, & \text{if } (m - \kappa) < 0. \end{cases}
\]

The problem of deciding the optimal number of content provision can be formulated as

\[
\max m \pi^* = \frac{m}{4} \left(\alpha \lambda_0 - \frac{\alpha \beta}{\varepsilon+1} m^\varepsilon\right)^2 - R(\alpha, \beta) - \left(\min(m, \kappa) c_1 + (m - \kappa)^+ c_2\right).
\]

The optimal content provision is given by solving
\[ \Phi = \frac{\alpha^2}{4} \left( \lambda_0^2 - 2\lambda_0 \beta m' + \frac{(2\varepsilon + 1) \beta^2 m'^2}{(\varepsilon + 1)^2} \right) - c = 0 \], where

\[ c = \begin{cases} c_i & \text{if } m \leq \kappa \\ c_o & \text{if } m > \kappa \end{cases} \tag{5} \]

For \( \varepsilon = 1 \) the optimal content provision has a closed form expressed as

\[ m' = \frac{4\alpha \lambda_0 - 2\sqrt{\alpha^2 \lambda_0^2 + 12c}}{3\alpha \beta}. \tag{6} \]

**Result 2.** The optimal number of content varieties increases with the fitness of customer preference.

Result 2 reveals that the monetization of long tail content are determined by the lower content creation (e.g. user-generated content) and acquisition (i.e. old or delayed content) and the effectiveness of the developed recommendation system with accurate preference analysis (\( \beta \)). Notice that the ratio of hit content provision is measured by \( \kappa / m' \). For the sale distribution (\( \varepsilon = 1 \)), the niche content should be provided (\( \kappa / m' < 1 \)) when the condition

\[ c_i < \hat{c}_i = \frac{1}{12} \left( \frac{34\alpha \lambda_0 - 6\lambda_0 \beta}{2} - \alpha^2 \lambda_0 \right)^2 \]

**3.2. Optimal Recommender Development**

The content provider can develop a recommender system to promote the content sale (i.e. increasing \( \alpha \) and/or decreasing \( \beta \)). For example, developing a social recommender, utilizing the power of social influence and information cascade, can enhance the content value perceived and the corresponding purchasing willingness of a user. Owing to the network effect, a social recommender system will particularly benefit the sale of hit content. On the other hand, the content provider can develop a personalized recommender system to promote the sale of niche content. In the subsection, we will analyze the appropriate budget allocation strategy for developing these two types of recommender systems.

The development cost for the recommender system is formulated as \( R(\alpha, \beta) = \gamma_1 \alpha + \gamma_2 (\beta_0 - \beta) \), where \( \beta_0 \) is the initial natural level of content demand curve, without any support of a personalized recommender system. The total available budget for recommendation technology investment is \( I \). The optimal investment problem of the recommender system development can be formulated as

\[ \max_{\alpha, \beta} \pi = \frac{m}{4} (\alpha \lambda_0 - \frac{\alpha \beta}{\varepsilon + 1} m')^2 \]

\[ - (\gamma_1 \alpha + \gamma_2 (\beta_0 - \beta)) - C(m) \]

s.t. \( \gamma_1 \alpha + \gamma_2 (\beta_0 - \beta) \leq I \)

(7)

Substituting \( \alpha \) with \( (I - \gamma_2 (\beta_0 - \beta)) / \gamma_1 \) and solving \( \hat{\pi} = \beta \) \( \hat{\pi} = \frac{\beta_0}{2} \), we have the optimal design levels of social and personalized recommendation:

\[ \alpha^* = \frac{1}{\gamma_1} \frac{\beta_0}{2} \left( \frac{\varepsilon + 1}{2m'^2} \right), \]

\[ \beta^* = \frac{\beta_0}{2} \left( \frac{\varepsilon + 1}{2m'^2} \right). \tag{8} \]

The portfolio share of the budget allocated for a social recommender and personalized recommender development respectively is

\[ (\theta_s, \theta_p) = \left( \frac{\gamma_2 m' \beta_0 - (\varepsilon + 1) \lambda_0}{2m' I}, \frac{\gamma_2 m' \beta_0 - (\varepsilon + 1) \lambda_0}{2m' I} \right) \tag{9} \], \( \theta_s + \theta_p = 1 \).

Examining (8) and (9), we have the following findings

**Result 3.** As the number of content varieties increases, the optimal effectiveness level of the social recommender decreases but the optimal effectiveness level of the personalized recommender increases. However the portion of the budget allocated for personalized recommender will decrease as the total budget enlarges.

**3.3. Competing Content Providers**

We extend the model to consider a market with two competing content providers \( A \) and \( B \). The content pricing, recommendation, and provision of one provider can affect not only its own sale but also its opponent’s sale. The profit function of a content provider \( i \in \{A, B\} \) is formulated as:

\[ \pi_i = p_i \int p_j^\infty \left( (\alpha_j - \delta \alpha_j) (\lambda_j - (\beta_j - \eta \beta_j) x) - p_j + \mu p_j \right) dx - R(\alpha, \beta) - C(m_i) \],

(10)

where \( j = \{A, B\} - i \), \( \delta, \eta \geq 0 \) is the negative effect on the demand on a provider from the social and personalized recommendation strategies of its opponent respectively. \( \mu \geq 0 \) represents the opponent’s price effect.
Solving $\frac{\partial \pi_i}{\partial p_i} = 0$, $i \in \{A, B\}$, the best response pricing function of provider $i$ to provider $j$ is derived as:

$$p^*_i = \frac{1}{2} \left( \alpha_i - \delta \alpha \right) \left( \lambda_0 - \frac{\beta_i - \eta_\beta \mu}{\varepsilon + 1} m^*_i \right) + \mu p^*_j,$$

$i \neq j \in \{A, B\}$.

Further solving the equations expressed in (11) simultaneously, we have

$$p^*_i = \frac{2 \left( \alpha_i - \delta \alpha \right) \lambda_o - \left( \beta_i - \eta \beta \right) m^*_i}{4 - \mu^2} \left( \lambda_o - \frac{\beta_i - \eta \beta}{\varepsilon + 1} m^*_i \right) + \frac{\mu \left( \alpha_i - \delta \alpha \right) \lambda_o - \left( \beta_i - \eta \beta \right) m^*_i}{4 - \mu^2}.$$

For the symmetric setting ($\alpha_i = \alpha, \beta_i = \beta$, $m_i = m_j = m$), the equilibrium price and profit can be obtained as:

$$p^*_i = \frac{2 \left( \alpha - \delta \alpha \right) \lambda_o - \left( \beta - \eta \beta \right) m^*_i}{4 - \mu^2} \left( \lambda_o - \frac{\beta - \eta \beta}{\varepsilon + 1} m^*_i \right) + \frac{\mu \left( \alpha - \delta \alpha \right) \lambda_o - \left( \beta - \eta \beta \right) m^*_i}{4 - \mu^2}.$$

Further solving the equations expressed in (11) simultaneously, we have

$$\Phi_j = \frac{1}{4 - \mu^2} \left( \lambda_o - \frac{\beta - \eta \beta}{\varepsilon + 1} m^*_i \right) + \frac{\mu \left( \lambda_o - \frac{\beta - \eta \beta}{\varepsilon + 1} m^*_i \right)}{4 - \mu^2}.$$

The symmetric content provision can be obtained by solving the equation (4) with the setting of $m_i = m_j = m$. For $\varepsilon = 1$ the equilibrium optimal content provision can be expressed as a closed form:

$$m^* = \frac{(16 + 14 \mu) \overline{\alpha} \lambda_o}{(12 + 9 \mu) \overline{\beta}^2} - \frac{2 \sqrt{(8 + 7 \mu) \overline{\alpha}^2 \lambda_o^2 - (12 + 9 \mu) \left( 4 + 5 \mu \overline{\alpha} \lambda_o^2 - c \left( 4 - \mu^2 \right) \right)}}{(12 + 9 \mu) \overline{\beta}^2}.$$

**Result 4.** The equilibrium content price increases with the price ($\mu$) and personalization effect ($\eta$) but decreases with the social influence effect ($\delta$) from the opponent.

We further consider the scenario that the number of content varieties is a strategic decision variable. The profit-maximizing problem of content provider $i$ is rewritten as:

$$\max_m \pi_i^* = \frac{\overline{\alpha}^2 m}{4 - \mu^2} \left( 2 \left( \lambda_o - \frac{\beta - \eta \beta}{\varepsilon + 1} m^*_i \right) + \mu \left( \lambda_o - \frac{\beta - \eta \beta}{\varepsilon + 1} m^*_i \right) \right)^2,$$

where $\overline{\alpha} = \alpha (1 - \delta)$ and $\overline{\beta} = \beta (1 - \eta)$.

**Result 5.** In a competing market, the number of content varieties increases when the social influence effect from the opponent on the sale becomes less significant.

### 4. Conclusion

In this paper, we study the provision (volume of varieties) and recommendation strategies (social and personalized recommender design) for improving the sale and profitability of the digital content with asymmetric popularity distribution. From the supply side strategy, the decision of the optimal number of content varieties is analyzed. From the demand side strategy, the problem of budget allocation on various types of recommender systems is examined. The distinct economic roles of the two popular recommendation approaches are formulated and the impact of the network (social influence) and content (varieties and preference fitness) factors on developing the pricing strategies under different scenarios of market interactions are also analyzed.

We find that (1) content price will decrease with the number of content variety and (2) the optimal of content varieties increases with the fitness of customer preference. Besides, (3) the number of content varieties increases, the optimal effectiveness level of the social
(personalized) recommender decreases (increases). However, (4) the portion of budget allocated for personalized recommender development will decrease as the total budget enlarges. Finally, we also observe that (5) in a market with competing providers, the equilibrium content price and variety amount increases with the price and personalization effect but decreases with the social influence effect from the opponent.

The results can be used to make sense of some industry advancements. For examples, with more choices on products but lack of experience and information, customers are usually choosing on price. In order to attract customers to buy more varieties of content and raise sales revenue, increasing a wider variety of products and maintaining a lower unit price of a content variety becomes an important business strategy for content providers. As consumers become with higher differentiated preferences, the effectiveness of the developed recommendation system with accurate preference analysis and more content variety provision are essential. This principle is experienced by iTunes, Amazon, Netlix and other online content providers.

There are several studies which can be further extended. First, in the research, we focus on the “quantity” decision issue of the content provision. The “quality” issue can be further modeled and examined. A natural question is how to balance the effect of professional content and user-generated content on improving the profitability. Second, from the perspective of modeling, the current demand is aggregately measured without considering the purchase choice dynamics of customers. How the aggregated demand is formed from the choice dynamics can be further studied. Third, besides the usage based (per download or view) pricing scheme, the impact of non-linear pricing scheme on the resulting content provision and recommendation strategies can also be compared. Lastly, different types of digital content (e.g. movie, music, news, or app software) have some distinct characteristics which may affects the provision and marketing strategies and their impact can be explored.

5. Acknowledgements
The authors are grateful for the financial support of National Science Council of Taiwan (NSC 102-2410-H-009-054).

6. References


