

Association for Information Systems AIS Electronic Library (AISeL)

PACIS 2014 Proceedings

Pacific Asia Conference on Information Systems
(PACIS)

2014

AN IMPROVEMENT TO E-COMMERCE RECOMMENDATION USING PRODUCT NETWORK ANALYSIS

Jiaheng Xie

Renmin University of China, aaron.xiejh@gmail.com

Wangsheng Zhun

Renmin University of China, Zhuws1992@163.com

Kanliang Wang

Renmin University of China, kanliang.wang@gmail.com

Follow this and additional works at: <http://aisel.aisnet.org/pacis2014>

Recommended Citation

Xie, Jiaheng; Zhun, Wangsheng; and Wang, Kanliang, "AN IMPROVEMENT TO E-COMMERCE RECOMMENDATION USING PRODUCT NETWORK ANALYSIS" (2014). *PACIS 2014 Proceedings*. 182.

<http://aisel.aisnet.org/pacis2014/182>

This material is brought to you by the Pacific Asia Conference on Information Systems (PACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in PACIS 2014 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

AN IMPROVEMENT TO E-COMMERCE RECOMMENDATION USING PRODUCT NETWORK ANALYSIS

Jiaheng Xie, School of Business, Renmin University of China, Beijing, China
aaron.xiejh@gmail.com

Wangsheng Zhu, School of Business, Renmin University of China, Beijing, China
Zhuws1992@163.com

Kanliang Wang, School of Business, Renmin University of China, Beijing, China
kanliang.wang@gmail.com

Abstract

To help consumers find the most wanted products effectively, e-commerce recommendation saves lots of time spent on viewing unnecessary web pages and increases revenue for e-commerce websites. Because of this significance, this paper concentrates on technology of recommendation and makes an improvement to recommender system. To solve the problem of current technology that only concerns association between two products, this model considers the product network and guides consumers to view products following the intended path. The object is to maximize revenue of the entire product network. An empirical study of yhd.com shows our model is more effective than current model.

Keywords: E-commerce recommendation, Product network analysis, Centrality

1 INTRODUCTION

Recommender systems have been traditionally used by most e-commerce websites to solve the personalization problem by guiding customers to find the products they would like to purchase (Yong et al, 2005). For e-commerce websites, recommendation is an important means to guide consumers to buy products by predicting consumers' preferences. For consumers, recommendation saves unnecessary time spent on viewing web pages and helps them find the most wanted products effectively, thus providing consumers with better shopping experiences. As a result, the effectiveness of recommendation is a key point for both e-tailers and consumers. Recommendation technologies fall into two distinct categories: content based technologies and collaborative technologies (Yu et al, 2005). Content-based technologies make recommendations by analyzing the descriptions of the items that have been rated by the user and the descriptions of items to be recommended (PAZZANI, 1999) while collaborative filtering is a technology of making automatic predictions about the interests of a user by collecting preferences or taste information from many users (Wikipedia). Compared with content-based information filtering approaches, collaborative filtering has also the salient advantage that a user may benefit from other people's experience, thereby being exposed to potentially novel recommendations beyond her own experience (Adomavicius and Tuzhilin 2005). Collaborative filtering based on user (Resnick et al, 1994; Sarwar et al., 2000; Shardanand and Maes, 1995) is the most successful recommendation technology to date. This method relies on the fact that each person belongs to a larger group of similarly behaving individuals. Consequently, items frequently purchased by various members of the group can be used to form the basis of the recommended items. For example, amazon.com, the first website to use item-based filtering, uses the ratio of "people who viewed also bought"¹ as the criterion to make recommendation, as most e-tailers do now. However, this method only considers the association between two products. What if there are new changes when we view the whole product network? Besides, the sparse transaction (or rating) data condition makes predicting accurate recommendations difficult. To solve this problem, this paper proposes an improved model. By drawing the product network and by using social network analysis, the improved model takes all products and their complicated associations into consideration. Since collaborative filtering is the most widely used and most effective traditional method. We just compare the efficiency of the improved model and collaborative method.

Product network analysis is derived from social network analysis (SNA) which is used to analyze mutual relations in groups so as to study phenomena and structures in the society. SNA takes individuals as nodes and associations between two individuals as links. While viewing an e-commerce web page, you might notice that there are some products marked "x% of people who viewed this product also bought" in the left column. Unlike traditional SNA which studies the relations among a group of people, our model takes this viewed & bought ratio as associations between products and each product as a node (like an individual) in the network. By constructing a product network, we can see the features of the whole products and their structures, so these features serve as a way to make recommendation.

SNA includes 3 levels: individual node level, subgroup level and the entire network level. This

¹ To simplify, we take it as the ratio of viewed & bought

paper mainly uses individual level methods to optimize the effectiveness of the entire network level. Individual level concerns centrality of nodes and is a quantified study of the power of nodes. The widely used centrality are degree centrality, betweenness centrality and closeness centrality. Degree centrality is the number of nodes that link to a specific node, betweenness centrality reveals the possibility that one node occupies the links between other nodes and closeness centrality is the degree that a node is free from other nodes' control.

As described above, SNA is used to study the relations in a group of people and most SNA studies focus on human behaviors. But this paper applies this method to a different system—product network. SNA is to study a set of individuals and their various associations. Nodes refer to the participants in the social network and their associations are taken as a certain kind of links in a period of time or in a certain area (Sundaresan and Yi, 2000). In the e-commerce product network, flows among products turn out to be flows among consumers. The practical meaning that one product flows to another is that consumers who viewed the first product bought the other one, so it's a transfer of consumers' preferences. The flows among products are similar to relations among people, so it's suitable to apply SNA to product network with which we can study e-commerce recommendation.

This paper is organized as follows: Section 2 describes the general method of e-commerce recommendation and its flaw. Section 3 solves the problem. Section 4 gives an empirical study. Section 5 compares the improved model with traditional model. Section 6 concludes this paper.

2 THE GENERAL METHOD

2.1 An example of a simplified product network

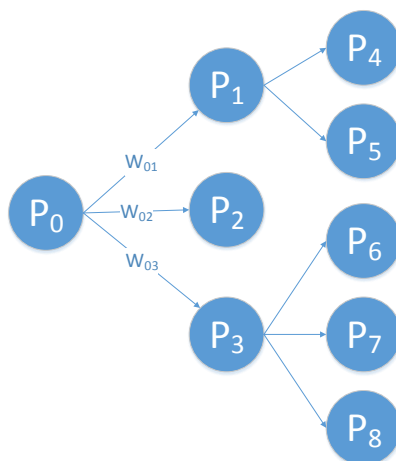


Figure 1. Viewing path starting from P0

Figure 1 is a simplified product network example which records the path consumers view products starting from P_0 . P_1 , P_2 and P_3 are products recommended in P_0 's web page, P_4 and P_5 are recommended in P_1 's web page and P_6 , P_7 and P_8 are recommended in P_3 's web page. Consumers will view P_0 first and will follow this path guided by recommendation. The weight of each edge denotes the ratio of viewed & bought between two products. For example, the edge weighted W_{01} between P_0 and P_1 refers to W_{01} percent of people who viewed P_0 finally

bought P_1 . To simplify, define P_0 as upstream product², P_1 as downstream product³, and W_{01} as the strength of association⁴ between P_0 and P_1 .

The strength of association reveals the probability that consumers would buy P_1 after viewing P_0 , of course these consumers would view P_1 as well, so the product of number of people who viewed P_1 and strength of association W_{01} can be taken as the number of people who bought P_1 . Obviously, strength of association is very important to recommendation because if the strength of association W_{01} is stronger than W_{02} , it's better to recommend P_1 (the *revenue* recommendation creates is greater).

2.2 Criteria to evaluate effectiveness of recommendation

The aims of recommendation are to convert viewers into buyers, to improve cross-sell by suggesting additional products for the customer to purchase, and to improve loyalty by creating a value-added relationship between the site and the customer (Schafer et al., 2001). Whatever, effective recommendation should help consumers find proper products and make purchasing decisions. On websites' side, the amount of revenue that recommendation creates is a good criterion to evaluate effectiveness of recommendation.

We use the following equations to evaluate recommendation. Downstream product of P_0 are P_1, P_2, \dots, P_i , $W_{01}, W_{02}, \dots, W_{0i}$ are their strength of associations, and Pr_k is the price of P_k . Define R_0 and TR as

$$R_0 = \sum_{k=1}^{k=i} (W_{0k} \times N_{0k} \times Pr_k) \quad (1)$$

$$TR = \sum_x I = \sum_{x=1}^{x=n} \sum_{k=1}^{k=i} (W_{xk} \times N_{xk} \times Pr_k) \quad (2)$$

Where N_{xk} refers to the number of consumers who viewed P_k which is recommended in P_x 's web page, but it doesn't mean these consumers all bought P_k . As mentioned above, the strength of association W_{xk} can be taken as the probability of purchasing P_k , so $W_{xk} \times N_{xk}$ can be taken as the number of people who bought P_k through recommendation in P_x 's web page, and $W_{xk} \times N_{xk} \times Pr_k$ is the revenue recommendation creates. R is the sum of the anticipated revenue of all the products through recommendation after viewing P_x . In another word, R is the revenue recommendation creates in P_x 's web page. TR is the total revenue that recommendation creates in the entire network.

At last, we use an equation to evaluate N_{xk} . Assume that V_x is a set of consumers who viewed the recommended product in P_x web page but didn't buy P_x . So V_x equals to the sum of N_{xk} ($k=1, 2, 3, \dots, i$). Here we assume that N_{xk} is only related to the rank of recommended products which are ranked as P_1, P_2, \dots, P_i . To simplify, we take the ratio of each N_{xk} from P_1 to P_i as $i:(i-1):\dots:2:1$. Though N_{xk} may actually follow a specific distribution, we here just want to distinguish which N_{xk} is higher. So the real distribution is not a necessity, but the proportional distribution can be easier to compute. So the number of consumers who viewed P_k is

² Upstream product is recommended before a specific product along the path.

³ Downstream product is recommended after a specific product along the path

⁴ Strength of association is the ratio of "people who viewed also bought".

$$N_{xk} = V_x \frac{i+1-k}{1+2+\dots+i} \quad (3)$$

2.3 The collaborative method

The critical step of collaborative filtering approach lies in searching the similar preference customers with the active customer, that is, find the similar customers. After finding similar customers, it then presents recommendation for active customer according to the preference of similar ones (Wei et al, 2007). The distance between two products is used to determine the their similarity. Since our goal is to maximize revenue, closer distance serves for higher revenue. Higher ratio of viewed & bought means more people will buy the recommended products after view a specific product. So this ratio is a perfect distance to represent similarity between two products.

An application is yhd.com which takes the ratio of viewed & bought as the similarity between two products as can be seen in yhd.com's web page. This technique ranks each item according to how many similar customers purchased it (Greg et al, 2003). As can be seen in equation (2), this method aims to maximize W_{xk} . Downstream products with high W_{xk} are ranked on the top of the recommended products, thus having a high number of viewers (N_{xk}). By combining equations (1) and (3), we can see this method is an effective way to increase R.

Figure 2 is a simplified product network which records the viewing path starting from P_0 . Assume V_0 is 90 and the number on each edge is the strength of association. Here we mean to distinguish R and TR, but price is unknown. We simply assume that all the prices are 1.

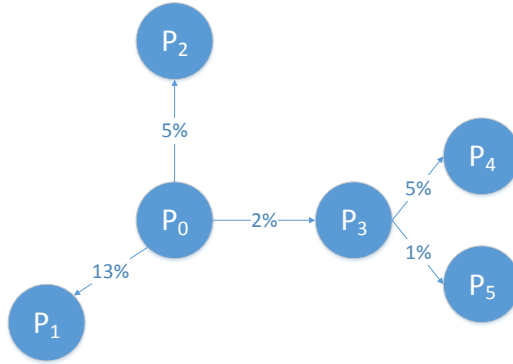


Figure 2. Relation of R and TR

In P_0 's web page, yhd.com would recommend P_1 , P_2 and P_3 . The number of consumers who viewed P_0 and viewed P_1 is

$$N_{01} = V_0 \frac{i+1-k}{1+2+\dots+i} = 90 * \frac{3}{3+2+1} = 45. \text{ So } N_{02} \text{ and } N_{03} \text{ are}$$

$$N_{02} = 90 * \frac{2}{3+2+1} = 30, \quad N_{03} = 90 * \frac{1}{3+2+1} = 15.$$

The number of consumers who flowed to P_3 from P_0 and flowed to P_4 and P_5 is

$$N_{34} = N_{03}(1 - W_{03}) \frac{i+1-k}{1+2+\dots+i} = 15 * (1 - 2%) * \frac{2}{2+1} = 9.8,$$

$$N_{35} = N_{03}(1 - W_{03}) \frac{i+1-k}{1+2+\dots+i} = 15 * (1 - 2%) * \frac{1}{2+1} = 4.9.$$

The amount of revenue recommendation creates is

$$R_0 = W_{01} \times N_{01} \times 1 + W_{02} \times N_{02} \times 1 + W_{03} \times N_{03} \times 1 = 45 * 13\% + 30 * 5\% + 15 * 2\% = 7.65$$

The amount of total revenue recommendation creates in the network is

$$TR = W_{01} \times N_{01} \times 1 + W_{02} \times N_{02} \times 1 + W_{03} \times N_{03} \times 1 + W_{34} \times N_{34} \times 1 + W_{35} \times N_{35} \times 1 = 45 * 13\% + 30 * 5\% + 15 * 2\% + 9.8 * 5\% + 4.9 * 1\% = 8.189$$

Using the strength of association as a criterion, the current method only recommends closely associated product, so customers are limited to a small area. If high betweenness products are recommended, the variety of products is increased, leading to higher revenue. This method can maximize R, but it doesn't mean maximum TR, where R means revenue of one product recommendation creates and TR means the total revenue recommendation creates in the entire network. There are differences between R and TR: most consumers will follow the recommended products until purchasing or giving up viewing products. TR computes the total revenue along the whole path while R only counts revenue from one product to its downstream products. There might be a case: if lower revenue at the beginning of the path can make higher revenue in the following paths, TR will increase although at the cost of the decrease of R.

3 SOLVING THE PROBLEM

Following the previous example, if the sequence is changed to P₁, P₃ and P₂, the number of consumers who view the first subscript also viewed the second subscript is

$$N_{01}' = 90 * \frac{3}{3+2+1} = 45, \quad N_{03}' = 90 * \frac{2}{3+2+1} = 30, \quad N_{02}' = 90 * \frac{1}{3+2+1} = 15,$$

$$N_{34}' = 30 * (1 - 2\%) * \frac{2}{2+1} = 19.6, \quad N_{35}' = 30 * (1 - 2\%) * \frac{1}{2+1} = 9.8$$

The amount of revenue recommendation creates is

$$R_0' = 45 * 13\% + 15 * 5\% + 30 * 2\% = 7.2$$

The amount of total revenue recommendation creates in the network is

$$TR' = 45 * 13\% + 15 * 5\% + 30 * 2\% + 19.6 * 5\% + 9.8 * 1\% = 8.278$$

In the revised method, TR' increased although R₀' decreased, because the revenue created by the recommendation in P₃ webpage increased.

This example shows that using strength of association as the simple criterion fails to achieve sustainable increase in revenue in the entire network. When some downstream products have strong ability to prolong the path that consumers viewing products, this simple method is not a good strategy, because after viewing P₀ and its downstream product P₃, few consumers would buy P₃, but most consumers who didn't buy P₃ would buy other products later. So P₃ occupies important path in this network which means P₃ can make higher revenue in the later path. To avoid the current problem, recommending products that occupy important paths is a good strategy. By guiding consumers to view important nodes, their viewing area is expanded, so the revenue is also increased

In the view of SNA, one criterion to measure whether P_K "occupies important paths" is betweenness centrality. High betweenness centrality means consumers can view more downstream products, which is showed in Figure 3.

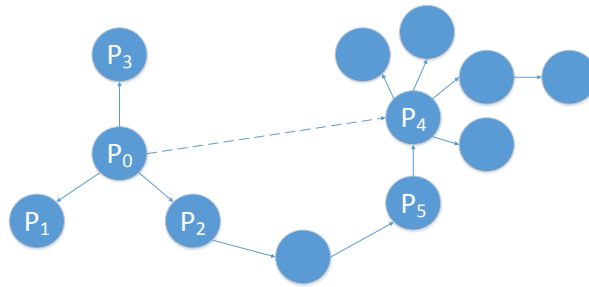


Figure 3. Application of betweenness centrality

According to current model which uses similarity as the simple criterion to make recommendation, P₁, P₂ and P₃ would be recommended in P₀'s web page. As the length of path consumers view products is limited, some consumers would give up viewing further products before viewing P₅. But if P₄ is recommended, consumers are likely to view more downstream products, so high-betweenness centrality products can boost revenue.

Products with high-betweenness centrality is a complement to products with high strength of association. If combining this two strategies, revenue in the entire network will be optimized.

4 AN EMPIRICAL STUDY

Figure 4 is the product network of 12 air cleaners. To simplify, we omitted the percent sign of the weight on each edge in Figure 4. Assume 1000 consumers viewed P₀ but didn't purchase it. If these consumers follow recommended products, we can compute the anticipated revenue that different recommendation creates.

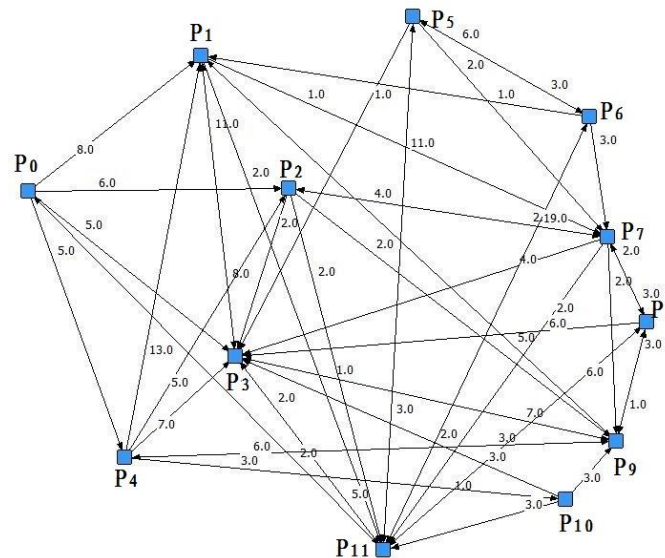


Figure 4. Product network of 12 air cleaners

Traditional method would recommend P₁, P₂, P₃ and P₄. The total revenue along the path is:

$$TR = R_0 + R_1 + R_2 + R_3 + R_4 = 152316.4 \text{ (the computation process can be seen in appendix 1).}$$

Ranking them using betweenness centrality, we would recommend P₄, P₃, P₁ and P₂.

Their betweenness centrality are listed in Table 1.

| Name | Betweenness |
|----------------|-------------|
| P ₄ | 12.292 |
| P ₃ | 9.750 |
| P ₁ | 7.000 |
| P ₂ | 3.708 |

Table 1. Betweenness centrality of air cleaners

The total revenue along the path is:

$TR' = R_0' + R_1' + R_2' + R_3' + R_4' = 159439.1$ (the computation process can be seen in appendix 2).

Combining the two criteria (betweenness centrality and strength of association), P₁, P₄, P₃ and P₂ will be recommended. Here let the first product with the highest strength of association. From the second product, the criterion changes to betweenness centrality. $TR'' = R_0'' + R_1'' + R_2'' + R_3'' + R_4'' = 165865.4$ (the computation process can be seen in appendix 3).

| Strategy | Sequence of products | R0 | TR |
|-------------------------|----------------------|-------|----------|
| Strength of association | P1, P2, P3, P4 | 94947 | 152316.4 |
| Betweenness centrality | P4, P3, P1, P2 | 81299 | 159439.1 |
| Combined strategy | P1, P4, P3, P2 | 96029 | 165865.4 |

Table 2. TR and R₀ of three strategies

As can be seen in Table 2, if we rank them using betweenness centrality, the revenue of P₀ that recommendation creates is reduced, but the total revenue is increased and if the two criteria is combined, the revenue is maximized. So by guiding consumers to view high betweenness centrality products, the total revenue in the entire network can be maximized.

5 DISCUSSION

Contrary to current model which uses similarity to rank recommended products, our model uses the strength of association to construct a product network and uses SNA to analyze product network which breaks the limit of SNA that only concerns relations among people. Taking the network as a whole, key nodes can be distinguished, as a result, solving the problem of current model that neglects the entire complicated associations. The improved model guides consumers to view products that occupy important paths, without which the associations of the entire network would be weakened. Although revenue on some paths is not maximized, the revenue in the entire network can be optimized. Using the improved model, e-tailers can recommend high-betweenness centrality products as a complement to high strength of association products so as to improve effectiveness of recommendation. Personalized recommendation can also be used as a complement to meet the need of a specific consumer.

Based on this paper, there are some expectations for future research:

1. The data we can access is the ratio of viewed & bought. If more accurate data can be

obtained, the result would be more promising.

2. The principle of combined strategy in this paper is let the first product be with the highest strength of association. If the first two, three or even more products use strength of association as the criterion and the left products use betweenness centrality as the criterion, revenue may increase again. The best principle to combine the two method can be reconsidered.

6 CONCLUSIONS

E-commerce recommendation plays a vital role in online shopping, so the effect of recommender technology is very important to e-commerce websites. Traditional technologies include content-based methods and collaborative methods, both of which fail to consider the relations in the product network. This paper improved recommendation model to solve the problem of traditional method and, as a result, to improve effectiveness of recommendation. Social network analysis are used in this model to guide consumers to view important products in the associated product network.

ACKNOWLEDGEMENT

This research was supported by the National Natural Foundation of China under the grant #71331007 and #70971107.

REFERENCES

- Kim, Y. S., Yum, B. J., Song, J., & Kim, S. M. (2005). Development of a recommender system based on navigational and behavioral patterns of customers in e-commerce sites. *Expert Systems with Applications*, 28(2), 381-393.
- Li, Y., Liu L., & Li X. (2005). A hybrid collaborative filtering method for multiple-interests and multiple-content recommendation in E-Commerce. *Expert Systems with Applications*, 28(1), 67-77.
- Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*, 13(5-6), 393-408.
http://en.wikipedia.org/wiki/Collaborative_filtering
- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE*, 17(6), 734-749.
- Xu, Y., Ma, J., Sun, Y. H., Hao, J., Sun, Y., & Zhao, Y. (2009). Using Social Network Analysis as a Strategy for E-Commerce Recommendation. In *PACIS* (p. 106).
- Sundaresan N. and Yi J. (2000). *Computer Networks: Mining the Web for relations*. 233(1-6):699-711.
- Schafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications. In *Applications of Data Mining to Electronic Commerce* (pp. 115-153). Springer US.
http://en.wikipedia.org/wiki/Collaborative_filtering

- Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. In Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence (pp. 43-52). Morgan Kaufmann Publishers Inc.
- Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *Internet Computing, IEEE*, 7(1), 76-80.
- Liu J. (2009). Lectures on Whole network Approach-A Practical Guide to UCINET. Shanghai: 1st Edition. Truth & Wisdom Press, 100-103.
- Wei, K., Huang, J., & Fu, S. (2007). A survey of e-commerce recommender systems. In *Service Systems and Service Management, 2007 International Conference on* (pp. 1-5). IEEE.

APPENDIX

1. Computation of TR in section 4

The number of consumers who viewed P₀ and viewed P₁ is

$$N_{01} = V_0 \frac{i+1-k}{1+2+\dots+i} = 1000 * \frac{4}{1+2+3+4} = 400. \text{ So } N_{02}, N_{03} \text{ and } N_{04} \text{ are}$$

$$N_{02} = 1000 * \frac{3}{10} = 300, \quad N_{03} = 1000 * \frac{2}{10} = 200, \quad N_{04} = 1000 * \frac{1}{10} = 100$$

Each of the amount of revenue recommendation creates after viewing P₀ is

$$R_0 = W_{01} \times N_{01} \times Pr1 + W_{02} \times N_{02} \times Pr2 + W_{03} \times N_{03} \times Pr3 + W_{04} \times N_{04} \times Pr4 = 400 * 8\% * 1690 + 300 * 6\% * 1399 + 200 * 5\% * 675 + 100 * 5\% * 1787 = 94947$$

R₀, R₁, R₂, R₃, R₄ are

$$R_1 = 400 * 92\% \left(\frac{3}{6} * 11\% * 675 + \frac{2}{6} * 2\% * 1980 + \frac{1}{6} * 1\% * 1599 \right) = 19500.32$$

$$R_2 = 300 * 94\% \left(\frac{4}{10} * 4\% * 1599 + \frac{3}{10} * 2\% * 1980 + \frac{2}{10} * 2\% * 928 + \frac{1}{10} * 2\% * 675 \right) = 11992.33$$

$$R_3 = 200 * 95\% \left(\frac{3}{6} * 8\% * 1690 + \frac{2}{6} * 2\% * 1980 + \frac{1}{6} * 1\% * 928 \right) = 15645.87$$

$$R_4 = 100 * 95\% \left(\frac{5}{15} * 13\% * 1690 + \frac{4}{15} * 7\% * 675 + \frac{3}{15} * 6\% * 928 + \frac{2}{15} * 5\% * 1690 + \frac{1}{15} * 3\% * 699 \right) = 10230.93$$

So the amount of total revenue along the path is

$$TR = R_0 + R_1 + R_2 + R_3 + R_4 = 152316.4$$

2. Computation of TR' in section 4

The computation process is the same as appendix 1.

The number of consumers who viewed P₀ and viewed the second subscript is

$$N_{01}' = 1000 * \frac{2}{10} = 200, \quad N_{02}' = 1000 * \frac{1}{10} = 100,$$

$$N_{03}' = 1000 * \frac{3}{10} = 300, \quad N_{04}' = 1000 * \frac{4}{10} = 400$$

Each of the amount of revenue recommendation creates after viewing P₀ is

$$R_0' = 200 * 8\% * 1690 + 100 * 6\% * 1399 + 300 * 5\% * 675 + 400 * 5\% * 1787 = 81299$$

$$R_1' = 200 * 92\% \left(\frac{3}{6} * 11\% * 675 + \frac{2}{6} * 2\% * 1980 + \frac{1}{6} * 1\% * 1599 \right) = 9750.16$$

$$R_2' = 100 * 94\% \left(\frac{4}{10} * 4\% * 1599 + \frac{3}{10} * 2\% * 1980 + \frac{2}{10} * 2\% * 928 + \frac{1}{10} * 2\% * 675 \right) = 3997.44$$

$$R_3' = 200 * 95\% \left(\frac{3}{6} * 8\% * 1690 + \frac{2}{6} * 2\% * 1980 + \frac{1}{6} * 1\% * 928 \right) = 23468.8$$

$$R_4' = 400 * 95\% \left(\frac{5}{15} * 13\% * 1690 + \frac{4}{15} * 7\% * 675 + \frac{3}{15} * 6\% * 928 + \frac{2}{15} * 5\% * 1399 + \frac{1}{15} * 3\% * 699 \right) \\ = 40923.72$$

So the amount of total revenue along the path is

$$TR' = R_0' + R_1' + R_2' + R_3' + R_4' = 159439.1$$

3. Computation of TR'' in section 4

The computation process is the same as appendix 1.

The number of consumers who viewed P_0 and viewed the second subscript is

$$N_{01}'' = 1000 * \frac{4}{10} = 400, \quad N_{02}'' = 1000 * \frac{1}{10} = 100,$$

$$N_{03}'' = 1000 * \frac{2}{10} = 200, \quad N_{04}'' = 1000 * \frac{3}{10} = 300$$

Each of the amount of revenue recommendation creates after viewing P_0 is

$$R_0'' = 400 * 8\% * 1690 + 100 * 6\% * 1399 + 200 * 5\% * 675 + 300 * 5\% * 1787 = 96029$$

$$R_1'' = 400 * 92\% \left(\frac{3}{6} * 11\% * 675 + \frac{2}{6} * 2\% * 1980 + \frac{1}{6} * 1\% * 1599 \right) = 19500.32$$

$$R_2'' = 100 * 94\% \left(\frac{4}{10} * 4\% * 1599 + \frac{3}{10} * 2\% * 1980 + \frac{2}{10} * 2\% * 928 + \frac{1}{10} * 2\% * 675 \right) = 3997.44$$

$$R_3'' = 200 * 95\% \left(\frac{3}{6} * 8\% * 1690 + \frac{2}{6} * 2\% * 1980 + \frac{1}{6} * 1\% * 928 \right) = 15645.87$$

$$R_4'' = 300 * 95\% \left(\frac{5}{15} * 13\% * 1690 + \frac{4}{15} * 7\% * 675 + \frac{3}{15} * 6\% * 928 + \frac{2}{15} * 5\% * 1399 + \frac{1}{15} * 3\% * 699 \right) \\ = 30692.79$$

So the amount of total revenue along the path is

$$TR'' = R_0'' + R_1'' + R_2'' + R_3'' + R_4'' = 165865.4$$