A Trilateral Influence Model for Online Shopping

[1]S.A.K.G.Samaraweera, [2]N.G.H.P.Gamage, [3]I.G.Gallage, [4]D.D.T.M.Gunathilaka, [5]Nimalika Fernando, [6]Dharshana Kasthurirathna
Faculty of Computing
Department of Software Engineering
Sri Lanka Institute of Information Technology, Malabe, Sri Lanka
[1]kalanags331@gmail.com, [2]hasinipgamage@gmail.com, [3]isurigallage@gmail.com, [4]thushan.ddtmg@gmail.com,[5]nimalikaf@gmail.com, [6]dkasthurirathna@gmail.com

***Abstract*—Application of social influence toward E-commerce has brought a significant benefit for the stakeholders. Consequently, it has enhanced the consumer satisfaction as well as spread of experiences. However, even with the collaboration of social influence there are some visible short comings potentially appearing in such systems. In fact, the contribution of social influence is still in an evolving state. The reliability of products is such recognized key issue that still appears in exiting social E-commerce systems. In this context we introduce a social influence model combined with a built in social network which further improves the customer reliability and satisfaction on available products. Thus, it can propagate reliable knowledge among community and optimize product recommendation process. The implemented model considers the personal preferences of respective consumers, their social influences in social network and external social influences to the system for the execution. Furthermore, it operates as a multi-agent system. The model has been validated by two sample data sets of consumers and products. As the results, majority have picked products suggested by combining external influences, internal social influences, and personal preferences. Therefore it has concluded that recommendation of products considering above three combinations is more effective.**

***Keywords*– social influence, E-commerce, multi-agent**

I. INTRODUCTION

Online shopping is a rapidly expanding area in e-commerce that has attracted most of the consumers and vendors all over the world. Along with massive growth of Internet usage purchasing products through online shopping has become a regular habit in today's society. Therefore with the growth of purchasing products and services via online, there are some appearances of well recognized key issues in online shopping. The people who use online shopping without knowing what they purchase exactly, but they want it fast and accurately is one such issue [1]. Furthermore technical issues such as difficulties in cancelling purchases increase the severity of this problem. As a solution, offering too many choices based on bunch of assumptions is a source of great consumer frustration. Consequently, once consumers are stressed out and unable to make the ultimate choice, there is a higher possibility of winding up their shopping sessions. In worst case tend to use alternative solutions without any attempt in purchasing at all [2].Furthermore, according to the market research done in 2016 for 23000 people from 25 different countries, it was found that, 54% of users buy products via online weekly or monthly, and most importantly majority (67%) of them prefer to read reviews regarding products and get influenced before they commit any purchasing activity [3].

As solutions, there are researches conducted to maximize the influence propagation in social networking. Therefore diffusion models such as *Linear Threshold* and *Independent Cascade* models have been introduced. Currently, these basic models are also further extended to optimize the productivity [4] of influence propagation. However in current systems they suggest products for consumers based on analyzing personal preferences and internal social influence to support consumer decision making process [5][6].

Therefore this research was conducted to enhance the confident and knowledge of consumers by recommending the most suitable products with the aid of personal preferences and social influences gathered from internal as well as external sources.

II. METHODOLOGY

In order to generate recommendations, the model extracts social influences on available products. These social influences are gathered from the internal social network and the external influence source; *Twitter*. As the next step, it analyzes the gathered information qualitatively. This analysis is archived by natural language processing technique called *Sentiment analysis* which is discussed later in this context. Once the analysis is concluded, it returns quantitative results to the product recommendation system. These results contain the *attitude* and *repeatability* values of products. The *attitude* defines a quantitative value of interest by consumers toward products while the *repeatability*

978-1-5090-4729-1/17/$31.00 ©2017 IEEE

describes the tendency to re-purchase same product by consumers. The internal social network helps to track down closest friends of user and grab their personal preferences in recommendation process. Therefore it uses customized edge ranking mechanism and consumer experiences to quantify the relationships. The personal preference analysis of respective users provides a set of *attitude* and *repeatability* values on products using the purchase history of those individual users. The product recommendation system is a multi-agent system and it maps all the products received from above personal preferences and social influences into a single grid by taking *attitude* and *repeatability* as the axes. The following multi-agent system suggests selected products from grid to consumers. The validation of model is conducted by recommending products from different combinations of above three influence sources (i.e. past behavior, internal social influence and external social influence). Hence, it has decided to monitor the impact for consumer selections done by each combination from above sources as the verification. According to the results it has shown more tendencies toward product selection with aggregation of above three influences.

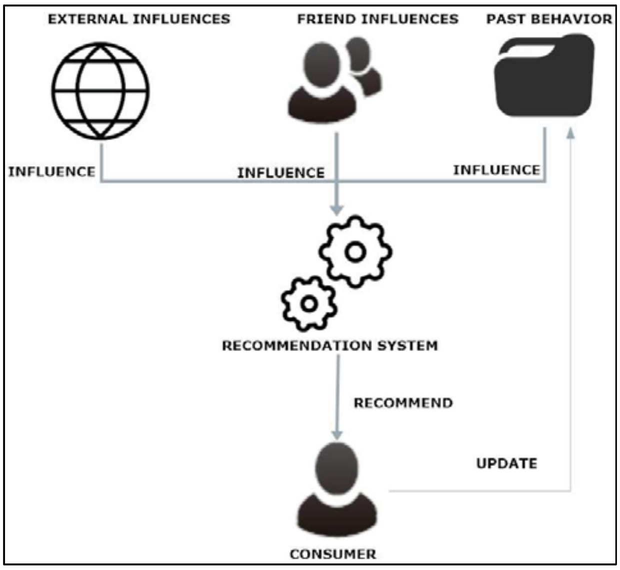

Fig. 1. Product recommendation with influences

A. *Personal Prefernce Analysis*

In this context, the *Personal-preference* is a term which describes the attitude of consumers towards products. Furthermore, the consumer attitude is a composite of their beliefs, feelings and behavioral intentions. These facts are considered together since they are highly interdependent and facilitate in determining how much the consumer will react to certain product or product categories.

The people have various beliefs and they directly affects to their selections. However consumers usually express their ideas in subjective manner. Therefore this research component is focused to convert those subjective measurements into quantitative values.

Even though it is possible collect quantitative ideas by a survey, it will not capture most accurate information due to invalid feedback given by consumers. Therefore monitoring user interaction via an actual system is more natural and accurate way of abstracting *personal-preferences*.

Therefore it has introduced multi attribute belief measuring mechanism to measure consumer beliefs. These attributes are namely; consumer beliefs on color, brand, price and quality of products. These attributes are identified using implemented online shopping application where consumers are facilitated to prioritize them accordingly.

In order to measure consumer beliefs, it has used *Elaboration-Likelihood* model [7].Therefore based on above adjusted attributes by consumer; it has given a common scale for each above belief attribute. Then using those beliefs it has measured the attitude by each individuals toward products. The *Fishbein* model is used to map the overall attitude values [8]. The following equation shows the application of *Fishbein* model to measure attitude.

Ao	= Attitude toward the product (Score)
b_i	= belief about the product possession from the attribute
e_i	= evaluation of the attribute as being good or bad
n	= the number of attributes

$$A_0 = \sum_{i=1}^{n} b_i e_i \qquad (1)$$

The belief is returned from *Elaboration-Likelihood* model, which is multiplied by the products in the consumer's purchase history considering same attribute under respective product category.

Finally, the summations of all those attributes are taken as total attitude of consumer for a certain product. Furthermore the repeatability of those products is calculated using standard deviation [9] where it measures the deviation of product attitude respective to other products under same category in consumer's purchase history.

s = Repeatability of attitude
x = attitude of selected item
n = Total attitude about the selected product category possession of the attribute

$$s = \sqrt{\frac{\frac{x}{n}\left(1-\frac{x}{n}\right)}{n}} \qquad (2)$$

The higher stranded deviation emphasizes the tendency to be repeated because the value is deviated slightly from stranded attitude value. Finally the *attitude* and *repeatability* values of products from purchase history will be transferred to recommendation system as the personal preferences of the consumer.

B. *External Social Influence Extraction*

The analysis of product related data which are published in social media networks is an approach to evaluate external social influence or the global influence for a product. Therefore Twitter API is used as the source of extracting product related data for further analysis. That process executes sequentially in three steps.

In extraction of external information, the product related *tweets* are extracted from Twitter through Twitter API. In this process product description is provided to filter the related tweets those have discussed about product.

Those extracted tweets are sentimentally analyzed to categorize them as positive tweets and negative tweets. This is archived through *Alchemy API* which is based on natural language processing, machine learning algorithms and now deep learning algorithms. *Alchemy API*'s *Sentiment analysis* algorithm seeks for words that carry a positive or negative connotation and captures the specific person, place and the fact that they have discussed. It also sense *negations* (i.e. *this is good* vs. *this is not good*) and *modifiers* (i.e. *this is good* vs. *this is really good*) [10]. Therefore it determines the sentimentality of the overall tweet to determine whether it is positive or negative and returns that information. These sentimentally analyzed *tweets* are used to determine the *attitude* and *repeatability* of respective products.

The *attitude* value indicates global *attitude* towards a product. The overall global attitude for a product can either *positive* or *negative*. According to the Equation 3, the overall external social influence score value is calculated using above analyzed data and generate overall attitude for the product.

p= total positive tweets of the product
n= total negative tweets of the product
A= product attitude

$$A = ((p - n)/(p + n)) \times 100\% \qquad (3)$$

The *repeatability* is defined by the global popularity of products. Therefore it is assumed that popularity is directly proportional to the positivity (i.e. popularity ∞ positive attitude). Then according to the Equation 4, using total number of positive tweets (p) over summation of both positive and negative (n) tweets, is considered as the repeatability of product.

$$R = (p / (p + n)) \times 100\% \qquad (4)$$

p= total positive tweets for the product
n= total negative tweets for the product
R= product repeatability

C. *Internal Social Network*

The implemented social network supports the product recommendation system with coordinating the internal social influence sources (i.e. consumers). Therefore it facilitates product recommendation system to identify most appropriate influence sources in personal preference extraction (i.e. *attitude* and *repeatability*). Hence it has implemented an internal measuring mechanism to rank user relationships and their gathered experiences on products. The experiences which are gathered by consumers on each product category rank users in social network. Therefore a separate retail application is implemented to measure user experience. The following application allows users to discuss product related concerns in real-time by conducting chat sessions on respective product categories. The number of *effective* chat sessions conducted for a particular product category is considered as the experience. An *effective* chat session is defined as a conversation conducted involving minimum of one participant, (i.e. friend) having at least a single feedback from that particular participant.

The relationships are ranked using edge ranking mechanism. Therefore it has abstracted the three ingredients of *Facebook* edge ranking algorithm in order to rank relationships [11]. The edge rank is calculated based on three factors namely; *affinity, weight* & *decay.*

The *affinity* score defines the connection strength of user toward a relationship. In the implemented social network it defines affinity score as the number of chat sessions conducted by certain consumer with respective participant of the relationship.

The *weight* is defined as the importance of consumer node in the social network. Therefore it is measured by sentimentally analyzing all the chat sessions executed between respective two consumers using sentiment analysis. If the overall discussions are more positive it is assumed that there is a high potentiality of propagating information among other connected nodes [12] via respective relationship and respective nodes become significant. Therefore more weight score is assigned for such relationships.

The reduction of relationship rank with respective to idle time of connected nodes without any interaction is considered as *decay.* In the implemented application, it has created a *cron-job* to reduce the decay score with respective to idle days between users in internal social network.

Finally the summation of *affinity*, *decay* and *weight* scores are considered as the relationship rank. Also, since there are two affinity scores per relationship it considers the average value as final affinity score.

$$R = \left(\frac{(a1+a2)}{2} + w + d\right) x100 \qquad (5)$$

d:decay score
w: weight score
a1: affinity score 1
a2: affinity score 2
R: Edge rank

These measured ranks are used to prioritize the influence sources when extracting products.

e: experience of a certain product category
y:percentage from total experience (i.e.8 product categories)
i: product category
R: Relationship rank
P: Priority

$$P = R.\left(\frac{e_i}{\sum_{i=1}^{8} e_i}\right) x\ 100\% \quad (6)$$

As conclusion this component returns the connected user nodes which are sorted according to above calculated value in Equation 6 for given product limit. Therefore recommendation system extracts the personal preferences of those nodes as internal social influence.

D. *Product Recommendation*

The integration and filtering of products received from above components are manipulated by a multi agent system with three types of agents and a grid environment [13]. The customer satisfaction for a product is used to measure the suggestions. Hence it refers the outputs (i.e. Score = *attitudinal strength*, Weight = *repeatability*) of above mentioned components to map the products for recommendation on the grid. These values are scaled from 0 to 100. The products with higher *weight* and *score* are considered as the best products due to their higher *attitude* and *repeatability*.

Product agent represents a point in the grid. Each of these agents can have multiple products at a time. Each suggestion source is mapped into the grid as they are, since each suggestion source can suggest the same product with their own point values. The *product* agents as a whole are responsible for picking the unique products.
The *Sweeper* agents start picking best products from the grid by sweeping from 100 to 0 in each axis. Once a product is picked it gets duplicated and that duplicated sweeper agents move towards the assigned *picker* agent in the shortest path with the picked product. Thus it enforces to select most likely products by consumer (i.e. high attitude and repeatability) to reach the picker first.

The *Picker* agent represents an imaginary most likely to buy product point for each category in the consumer's purchase history. However, when there is no purchase history available the above mentioned suggestion sources (i.e. *internal social network* or *tweeter* extraction*)* are considered to extract products information.

In cooperation with *Internal social network* list of recommended products is returned, extracted accessing the personal preferences of each friend of the user. These product lists are prioritized by Equation 6 and mapped in above mentioned grid with respective to attitude and repeatability. Also these mapped products will be evaluated with respective to external influence source *Twitter* as mentioned above.

The product selection process starts once a required number of suggestion count and requested categories are being set. Suggestions are personalized for the user and unique suggestions are generated every time. This module has the ability to manage conditions and inputs. The customer actions are recorded into a *MongoDB* collection. This includes consumer selection from suggested products. Consequently, they are used to validate the accuracy of recommendation. The options are provided to configure filtering by views, likes, dislikes, commented products and wish list as well as add or remove products from cart. Also deductible data such as current wish list, cart and suggestions can be filtered.

The Multi agent system is developed using *Python Mesa* [14] framework as a light weight solution.

III. RESULTS AND DISCUSSIONS

Even though the suggestions are generated from above three sources, there should be a mechanism to identify the most suitable combination out of above three sources to recommend products effectively. Also, in order to treat the consumers more effectively the recommendation process needs to be personalized for all product categories.

Therefore, in the proposed model, it has used Agent Based Modeling [11] (ABM) approach to generate results. Hence for each consumer separate grid with 100x100 of size with step based multi-agents (i.e. *picker, product and sweeper*) has been generated in runtime. However it raises the question about attitude of consumer toward generated recommendations. In order to validate and measure the accuracy of the proposed model following validation procedure has been followed.

Therefore using above mentioned sources (i.e. personal, social and external) model has been tested with 70 products which have been suggested for 16 users. There were seven combinations that could generate from above three influence sources and per each combination there were 10 products.

The following equation is used to measure the percentage of product selection attitude by consumers for each combination.

$$combination(i) = \left(\frac{\sum_{i=1}^{n} \left(\frac{x_i}{y} \right) x\ 100}{n} \right) x\ 100 \qquad (7)$$

x = number of picked products
y= total products from combination
n= total participants

As the results, following table was generated according to each combination of product selection. In the results, it has converted the actual number of selections into percentages.

TABLE 3.1. PRODUCT SELECTION OUTPUT

Criteria	Percentage
PP[1] only	6 %
ISN[2] only	3%
ESI[3] only	7%
PP+ISN	32%
PP+ESI	12%
ISN+ESI	5%
PP+ ISN+ESI	35 %

[1]PP- Personal Preferences [2]ESI - External Social Influence [3]ISN-Internal Social Network

According to the above results, it shows that majority of consumers have picked products suggested from the combination of all three components (35%). However it also indicates a higher rating (32%) when only internal social influence and personal preference are used for recommendation. Therefore it shows 3% of increment in satisfaction with the involvement of external social influence for product recommendation. The results also emphasize that recommending products only involving individual components has become less value compared to combined situations. In addition, a product selection once personal preference is disabled shows significant drop to 5% in customer satisfaction.
Furthermore, we have also conducted a survey to verify the concept of this research. Therefore it has made sample questionnaire and conducted a survey participating 20 users. These users are online consumers who use online shopping on daily basis.

According to survey results 17 users accepted product recommendation with combination of external social influences. In general, it shows that around 85 % voted for recommendation of products together with personal preferences and external & internal social influence.

IV. CONCLUSION AND FUTURE WORK

According to conducted evaluation process it has shown that majority of consumers (35%) have selected products suggested by aggregation of external & internal social influences as well as their personal preferences. Therefore it is possible to declare that majority of consumers prefer recommendation of products from above aggregation.

In the process of product recommendation, there are some other techniques such as data mining that can be applied for model execution. Therefore as future work, it is possible to combine those techniques as well to enrich the product recommendation process. However initially, it has used agent based modeling to increase adaptability of model toward various consumer behaviors. Thereafter once the behaviors of each individual have clearly identified above mentioned data-mining techniques also become eligible to apply.
The implemented model is capable to apply for any newly created or existing retail application, since it is highly adaptable to such environments. The ability of extracting external social influences and prioritize products accordingly is considered as major advantage that can be archived by this model.

In addition the implemented model is capable of identifying each individual's personal behavior. Therefore it supports personalization and performing specific reactions for each individual accordingly. Most importantly it is able to measure the dependability of individual from society and the adjust recommendation criteria.

Furthermore the implemented model is eligible to apply for various case studies related with social networking and influence propagation. Therefore it can be used as a tool of monitoring and controlling flow of influences in such social networks and communities. For example it can be used for; political campaigns, disease acknowledgement programs and etc.

How-ever there are some barriers in the implemented model. One such example is extraction of external social influences. Since most of the time external information is gathered via third party APIs (i.e. *Twitter*), it has to depend on practical limitations and their constraints. Consequently it may create an impact on the accuracy and reliability of suggestions generated by the model. Therefore in order to overcome this situation it has decided to get aggregate result from multiple external social influence sources to enhance the productivity.

REFERENCES

[1] Brian J. Corbitt, , Theerasak Thanasankit , Han Yi," Trust and e-commerce: a study of consumer perceptions ",vol.2, Electronic Commerce Research and Applications 2,203-215,April 2003

[2] "Solve five biggest problems in online shopping", 2014. [Online]. Available: http://www.guided-selling.org/solve-the-5-biggest-problems-of-online-shoppers, [Accessed on:8-Feb-2016]

[3] J. Maxwell, "Total Retail Survey 2016", 2016. [Online] Available:http://www.pwc.com/gx/en/industries/retail-consumer/global-total-retail.html, [Accessed on:10-Feb-2016]

[4] David Kempe, Jon Kleinberg, E´ va Tardos, "Maximizing the Spread of Influence through a Social Network" Department of Computer Science Cornell University, Ithaca, New York.

[5] Young Ae Kim "Impact of Social Influence in E-Commerce Decision Making" KAIST Business School, Kore

[6] Bo Xiao, Izak Benbasat "E-Commerce Product Recommendation Agents: Use, Characteristics and Impact" Sauder School of Business University of British Columbia, Canada

[7] Petty, Richard E; Cacioppo, John T (1986). "The elaboration likelihood model of persuasion". *Advances in experimental social psychology*: 124–125.

[8] Ajzen, Icek; Fishbein, Martin (1977). "Attitude-behavior relations: A theoretical analysis and review of empirical research." *Psychological Bulletin, American Psychological Association*. 84 (5): 888–918. doi:10.1037/0033-2909.84.5.88

[9] Koji Horie, Yusuke Tsutsumi, Yukio Takao, Tomomichi Suzuki. "Calculation of Repeatability and Reproducibility for Qualitative data". Department of Industrial Administration, Tokyo University of Science

[10] Sentiment Analysis API, 2015. [Online]. Available:http://www.alchemyapi.com/products/alchemylanguage/sentiment-analysis", [Accessed on:20-Feb-2016]

[11] M. Patterson, "Edge-rank: A Guide to the Facebook News Feed Algorithm", 2015. [Online]. Available: http://sproutsocial.com/insights/facebook-news-feed-algorithm-guide/, [Accessed on:5-Feb-2016]

[12] A.Borodin, Y. Filmus, J.Oren , "Threshold Models for Competitive Influence in Social Networks". Department of Computer Science, University of Toronto, Canada

[13] N. Gilbert, Agent Based Models, University of Surrey, Guildford, UK, 2008.

[14] David Masad, Jacqueline Kazil,Mesa, "An Agent-Based Modeling Framework", 2015:proc. of the 14th python in science conf. (scipy 2015)