



Intent Based Association Modeling for E-commerce

Sailesh Kumar Sathish^(✉) and Anish Patankar

Samsung R&D Institute Bangalore – India, Bengaluru, India
{sailesh.sk, anish.p}@samsung.com

Abstract. Online e-commerce sites track user behavior through use of in-house analytics or by integrating with third party platforms such as Google Analytics. Understanding user behavioral data assists with strategies for user retention, buy-in loyalty and optimizing objective completions. One of the more difficult problems though is understanding user intent that can be dynamic or built over time. Knowing user intent is key to enabling user conversions - the term used to denote completion of a particular goal. Current industry approaches for intent inference have an inherent disadvantage of having the need for embedded tracking code per site-sections as well as the inability to track user's intent over longer periods. In this paper, we present our work on mining dynamic as well as evolving user's intents, using a latent multi-topic estimation approach over user's web browsing activity. Further, based on the intent patterns, we look at generating association rules that model purchasing behavior. Our studies show that users typically go through multiple states of intent behavior, dependent on key features of products under consideration. We test the behavioral model by coupling it with Google Analytics platform to augment a re-marketing campaign, analyzing purchasing behavior changes. We prove statistically that user conversions are possible, provided purchase category dependent associations are effectively used.

Keywords: Intent mining · Latent models · Association modeling

1 Introduction

The market for global online shopping, as of 2018, is estimated at US\$ 1.9 trillion [1]. A foothold into this ever-competitive market requires a thorough understanding of one's user base. E-commerce sites are now capable of monitoring every single of their user's online activities and are coming up with ingenious ways to make use of such information in order to translate them into valued transactions. The term used is "user conversions", where conversion means the execution of a series of steps that result in a goal completion. Using third party analytics services such as Google Analytics (GA) [2] and Microsoft Azure Analytics [3], services can now go beyond their own domains to track user activity through cross-domain linking. Such services can track and identify user events, building higher-level inferences, and simultaneously track the number of goal completions resulting from individual marketing campaigns. In the case of Google Analytics, by linking with Google Ads, entities can track full customer cycle

through ad impressions and enable better remarketing campaigns. Such cycles translate to short-term intent of users, which being generally dynamic, is difficult to predict based on past behavior. The difference between intent and long term profiling is that intent behavior does not conform to consistent preferences that are categorical and contextual as captured by a profile. From an industry perspective, services such as GA have made a big impact on the way user behavior and conversions are affected. The disadvantage is that such services require embedded scripts that address each page sections and events requiring tracking in order to do backend analysis. External marketing channels require proprietary campaign tagging with explicit tracking that makes scaling over other sites difficult. Furthermore, such systems only capture short-term intents, relying more on syntactic matches than a semantic behavioral understanding behind user intents.

The paper describes our intent mining framework that captures short and long-term user intent. The aim is to aid third party services (subject to user permission) with remarketing campaigns. By using web browser as application platform, we are able to model an original intent that may have evolved over multiple sessions and across content categories. Our goal is to recognize, understand and model accurately the evolution of user intent over time and study behavioral aspects on the different content associations that user makes based on their intent.

2 Prior Art

The study and analysis of online consumer behavior is a well-researched subject that can be traced back to the onset of e-commerce on web. The fields of research in this area mainly focus on profile-based behavior, prediction models based on user purchase expectations, social media based behavior correlations and works that model user intent itself. Work done by Kumar et al. [4] looks at how profile based behaviors including demography affects purchase behavior. Other works look at predicting purchase behavior. Several diverse approaches have been proposed such as predictions based on statistical purchase probability based on pre-modeled scales [5], visual feature based matching and recommendation [6], modeling customer attitudes towards products & companies [7], machine learning based approaches on customer segmentation with predictive analysis [8], looking at repeat behaviors to predict future purchases [9, 10] and using web search data [11]. Work done by Ioanas et al. [12] looks at social media behavior and its correlation with online purchase behavior. A study on user behavior over Pinterest [13] looks at predicting intent spread between applications through use of a cross application model. Our work differs in the way intents are inferred and modeled. We account for reinforcement and decay of intent and associate intents with a step transition model based on pre-determined threshold values. Our study goes a step further by generating unsupervised association models for user behavior based on intents inferred.

3 Intent Capture and Analysis Framework

We have implemented the intent mining software framework as a hybrid client-server model, with the client part provided as an extension to Samsung Internet Browser v6.2. To address the issue of cross-content analysis, we analyze all web content that are article qualified (over 200 words) that contribute to user intent. The framework creates intent objects for the user based on the topics user consumes through the web browser. An intent object expires when that intent results in a successful purchase.

3.1 Intent States

We look at only those intents that may result in purchase of a particular product or service. All other categories, which are unrelated to any commercial transactions are ignored as they are outside the scope of our study. We observed that users go through at least three phases between recognition of a first intent and a consequent purchase decision. When user browse a new product related topic, a “weak intent” is generated for that topic. The product or category is referenced through our ontology and intent recognition can be seen as the predominant topic identification for that page. We use a refined latent multi-topic classification that follows from our earlier work [14], which can perform multi-categorical inference within a single page.

The second phase of intent transition is “in-progress” that represents a stronger conviction on the intent state. In this state, we observed that typically, the user’s browsing habits indicated a more narrowed focus on a specific range of products as opposed to a topic related to that product. Example, when user starts browsing on “home theatre audio reviews”, the system generates a “weak intent” for “Technology => home audio”. The “weak intent” transitions into an “in-progress” state once user reinforces the intent through consistent browsing on that topic. It is also typical that user may have narrowed on particular brands at this stage.

The final intent phase is “strong” phase. This is the typical phase where users may make a purchase (conversion) decision. The modeling function takes into account reinforcement and decay of intent with respect to other active intents within a user dependent activity period. In calculating reinforcement and decay, we do not take into account the correlation or independence between the products and categories. Instead, we factor in topic dependence. This is because relational intent is hard to model, as product relations, even if modeled within the ontology, often results in erroneous inference.

Figure 1 shows the normalized intent progression time against intent level values for two product’s price range: less than and greater than 100 USD. We normalized progression time, as users have different time range for intent realization but on average, apart from outliers, tend to follow a similar transition pattern. We observed that the intent state times are highly dependent on price of the product under consideration. We observed variance between product classes but for brevity, we factor for price, which is the more significant influencer.

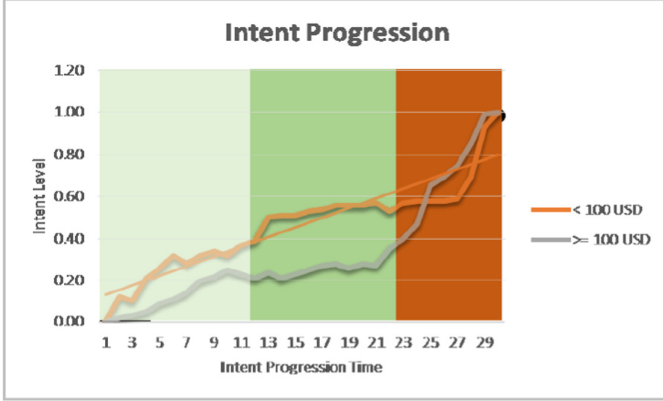


Fig. 1. User intent state progression. The three regions denote states of “weak”, “in-progress” and “strong” respectively. The two curves indicate, progression for low value (<100 USD) and high value priced (>=100 USD) products respectively

3.2 Modeling Intent

Figure 1 shows the progression of user intent states. The three regions denote states of “weak”, “in-progress” and “strong” respectively. The two curves indicate progression for low value (<100 USD) and high value priced (>=100 USD) products respectively. We chose to make our analysis over these two price points as we were able to cluster behavior patterns to within these brackets. Based on captured data points, we present an empirical model for user intent. We use co-occurrence of categories as well as specific product models within each category. Some categories have well defined product names while other product types have generic names. We observed that having specific product names resulted in stronger intent.

To normalize across such categories, we use a factor called *Product Specificity* for a category (C_{PS}). The categories having specific product names get lower weightage while other categories get higher weightage. In addition, extracting product category and product names is important for estimating price. It so happens that products with less monetary value move from weak to strong intent faster than products with high monetary value. We factor in the inverse relation with monetary value using parameter $Z = k/V$, where, V is price of product and k is a constant empirically taken as 100. Occurrences of category and products in the user topic vector are referred as C_C and P_C respectively. From e-commerce data [4], we understand that certain categories have more customers as compared to others. This variation across categories is captured as C_W . The intent capture model is given in (1).

$$I = f(C_W, C_C, P_C, Z) = \varphi(C_W(C_C C_{PS} + P_C)Z) \quad (1)$$

Where, φ is a constant, and I is the intent score. It should be noted that this intent score is for a category for a given time window and needs to be adapted as new data becomes available. We applied temporal learning to learn the intent score for category

over a period. Therefore, within a time window if intent score for a category is improved, it adds up to current aggregate intent score. For high monetary value items, linear learning works better while inverse learning is better suited for lower valued items.

3.3 Topic Modeling

We analyzed topic distributions within web content that users consumed to determine intent. We modeled web article content into topics using well-known Latent Dirichlet Allocation (LDA) based modeling [15]. For topic inference, we first built a supervised model for 36 categories (from our internal ontology) for English language resulting in a model size of 2 MB. We used a pre-categorized 6K URL corpus taken from our internal web proxy service for training. We built the models using a proprietary batch integration process including a hyper-parameter estimation method for accurate model convergence. This derives from our earlier work on building semantic indices [18], described briefly in following sub-sections.

Determination of Hyper-parameters

For LDA model, hyper-parameter α indicates the distribution of topics over a document and β indicates the distribution of words over a given topic. We need optimal α , and β values that give best converged model for a given set of topics. For this reason, we use an internally developed metric called Averaged Normalized Mode (ANM), given by Eq. 2 as a score to compare the purity of the mixture produced by LDA clustering.

$$ANM = \frac{1}{n} \sum \frac{\max\{T_i\}}{C_i} \quad (2)$$

Where C_i denotes the i^{th} cluster out of n and T_i represents the i^{th} topic. The ANM score is computed for $\forall \alpha \in [0.1, 3.0]$ and $\forall \beta \in [0.1, 1.0]$, with incremental steps of 0.1 each. In repeated runs with fixed α and β , if the ANM score of 1.0 repeats consistently, then the model is considered stable.

Incremental LDA

Incremental LDA (iLDA) performs supervised inference against a set of pre-built LDA batch models. Inference is performed through an incremental Gibbs sampler by applying sampling process to a pre-set of the sampled distribution and sampling for particular topic to which word i belongs conditioning on the previous word model ($i - n$) (as shown in (3)).

$$P(Z_i | Z_{i/j} w_i) \propto \frac{n_{Z_i, i/j}^{(w_j)} + \beta n_{Z_i, i/j}^{(d_j)} + \alpha}{n_{Z_i, i/j}^{(i)} + W \beta n_{i, i/j}^{(d_j)} + K \alpha} \quad (3)$$

Here, K is the number of topics, W is vocabulary size, Z_i represents i^{th} topic assignment; $n_{Z_i, i/j}^{(w_j)}$ represents word – topic Z_j assignment, and $n_{Z_i, i/j}^{(d_j)}$ is document-to-topic Z_j assignment. After incremental inference, we perform cluster process by aligning all vectors together that fall within a set threshold.

3.4 Product Name Extraction

We further developed a product name extractor to detect product name from web content. The product name extractor is based on Stanford NLP's Named Entity Recognition (NER) extractor software [16]. A rule-based module augments the NER to capture n-gram tokens as probable candidate for product names. The rules were written specifically for every base category. The title and URL tokens extracted helped in boosting the confidence level in the product names from the web page content.

3.5 User Intent Structure

A data structure, called User Intent Structure (UIS) hosts the identified user intent. The client generates the UIS and depending on the application, the UIS may be used locally or sent to an external service. When there is a change in intent data, the client sends an updated UIS to server for use by validated services. The UIS covers intent, topic vector, state values, categories and optional fields for URL values.

3.6 Observations on User Intent

The cut-off threshold values for the intent states were determined based on the observed drop-off probability values and the amount of time users spent within these states. We observed that users spent least time within the “strong” intent state, often marked by increased activity and lowest time intervals between topic reads. When in weak intent state, user concentration on topic was lower and on average, users created more intent-objects during this time. When user intent was in “in-progress” state, the number of weak intents created (for other intents than the current one) were lower. The number of new intent creation was another way of observing whether other user intents were in weak or in-progress state. The observations made here were for majority users and there were few outlier cases where users did not fit within this model. As Fig. 1 shows, purchase intent for higher priced products follows a lower level of intent between weak and in-progress states. We also observed that activities on search and reviews pick up at a much higher rate during the strong intent phase. Such observations were not applicable across categories. One example was “Fashion Apparels” where the transition states were not applicable for majority users.

4 Association Mining

Our association-mining engine looked at intent-based purchase paths that led to actual purchases on partner web sites. The association-mining engine is part of the intent mining and reinforcement framework.

4.1 Unsupervised Association Rule Mining

The association rule-mining engine on server generated unsupervised rules pertaining to user journey on specific topics and associated topic features similar to “if-then” rules. A typical rule may encompass one of the many user journeys possible from the

beginning of an identified intent to a goal completion (such as a purchase). Association rules or affinity rules look at identifying proper antecedents (X) and their corresponding consequents (Y). The antecedents and their corresponding consequents form item sets where the items within the sets are disjoint. We derived association rules from user behavioral patterns and thus, the probabilistic nature of behaviors also extend to association rules. The association rules were generated for each topic/product items identified as an intent (indicated via UIS).

Clients would send UIS updates on identified intent that contain fields for “recognized association journey” for that intent. Clients (in default mode) sent updates through UIS once a week to server and when connected over Wi-Fi. The association “journey” captured on the client side will identify the next “similar” topic (web page) user had consumed, either as search and read, direct read and/or bookmarked by the user. The capture and update to server continued until the intent (UIS) closed at user end.

The server combined all the UIS received to date (within a processing window) and computed association rules per intent. The first stage of the process was to group all similar intents together. This was done based on topics contained with the UIS that were used for grouping. Note that even though the topics were same, the underlying vectors that defined each topic could be different. At this stage though, the differences in topics were not considered. Topic based identification of frequent sets were done in second stage. The set of UIS with their unique IDs were stored within a table in the server database for a particular intent (topic). All subsequent updates received from clients for the same UIS were stored against the corresponding entry within the database. If a new UIS was received for an existing topic, a new entry for that UIS was made within the association content table for that topic. The further steps for identifying a set of unsupervised association rules per intent is described in subsection B. At present, we are only focusing on positive rule mining ($X \rightarrow Y$) and not considering negative rule mining. Considering that total product catalog can be huge, all types of negative rule associations ($\neg X \rightarrow Y$, $X \rightarrow \neg Y$, $\neg X \rightarrow \neg Y$) are computationally expensive and cumbersome to formulate. However, since, we have only 36 topics negative rule mining of type $\neg X \rightarrow Y$ is still possible to some extent with comparatively less computational complexity. That is absence of topic from browsing history can be a cue for intent. E.g., Person not interested in Science might be interested in fashion accessories.

4.2 Identifying Frequent Sets

After grouping of intent sets were made in first stage, the following stage looked to identify frequent item sets within the intent and its associations. The second stage for identifying frequent items consisted of two sub stages. The first sub-stage performed a time-windowed alignment of page jumps by a single user based on the unique UIS. What this did was, within a time window (2 days), all the UIS association updates received were considered as a single journey path so as to build a simple and reasonable corpus for rule mining. Any sites re-visited by the user within this window was ignored and the first visit time (or bookmark) of that URL (within the window) was used in building the disjoint path set. After this process, we ended up with a corpus for

a topic with sets ranging from single antecedent – single consequent to single antecedent – multi consequents.

For the second sub-stage, we used the Apriori algorithm [17] for reducing set complexity for generating frequent item sets. If we use all antecedent-consequents – combinations of single items, paired items, and triples, it would require high computational resources that would grow exponentially with each combination addressed. With the Apriori algorithm, we initially generated frequent item sets with just one item. The frequency of occurrence formed the support for that set. The one-item lists where the support was below our set threshold were dropped and only those above the threshold were chosen. The next step was to identify the most frequent sets containing at least two consequents by using the frequent one-item sets that were identified in the first iteration. We used the same one item set for identifying the next two-item set that contained the same antecedent and consequent sets and so on. Through this iterative process, we identified all k-item sets based on the frequent (k – 1) item sets identified through the preceding step with each step calculation requiring only a single database query. The second stage of association mining still produced a large set of association rules that did not indicate the level of bond between the antecedents and the consequents. In order to filter out the weak associations and to end up with a convincing rule set, we used two additional metrics: confidence score and lift ratio, described in subsection C.

4.3 Determining Association Strength

Confidence score measures the degree of uncertainty amongst the identified association set. Confidence score compares the item co-occurrences (transactions with both antecedent and consequent sets) to the total antecedent occurrences. Confidence is given by (4).

$$C_f = \frac{P(\text{antecedent AND consequent})}{P(\text{antecedent})} \quad (4)$$

Where C_f , denotes the conditional probability that a randomly selected rule corresponding to an antecedent will contain all consequent transactions.

Once confidence score C_f , was calculated for an association rule, we further filtered the rule based on lift ratio. Confidence score is a conditional probability score with the consequent dependent on an antecedent occurrence. Lift gives us a benchmark ratio that compares an independent probability score against the conditional probability. Lift metric thus indicates how valuable the conditional clause is as compared to the case where the antecedent set and the consequent sets are independent of each other. Lift is interpreted as a benchmark score (5).

$$\text{Lift ratio, } L_r = \frac{C_f}{P(\text{consequent})} \quad (5)$$

A value for $L_r > 1$ indicates that the level of antecedent \rightarrow consequent association is higher than what would be expected if the two sets were independent. It gives the

level of correlation between the two sets and is thus a useful metric for determining strength of associations. Table 1 shows a sample of the mined association rules. Note that some of the association rules gets repeated. For these rules, even though the topic association was the same, the underlying sub-topics or their corresponding token vectors would be different indicating the various user preferences on topic variations among users.

Table 1. Association rules example

Rule#	Antecedent (a)	Consequent (c)	Antecedent Vector (Va)	Consequent Vector (Vc)	Lift ratio
1	Photography	Camera, Smartphone	{mirrorless, speedlite, AI, BSI, image processing, autofocus, EV, AF, exposure, buffer, JPEG, log gamma, SD, UHS-C...}	{Google, clips, shutter, AI smarts, gesture, goofy, galaxy...}, {aperture, variable, dual lens, jack, snapdragon, android, exynos, IP68, emoji, iPhoneX...}	1.38
2	Photography	Camera, Smartphone	{macro, large aperture, landscape, focal points, horizon, sport metering, exposure lock...}	{dual pixel, cmos, eye detection, autofocus, dual lens, touch and drag, itu, cr3, digic, 4k...}, {OIS, lowlight, android, portrait, dualcamera, wireless, pixel, poled, amoled...}	1.265553

5 Evaluating Intent Behavior

We used the browser client extension running the intent mining engine for both user data collection and for validating behavioral activities. Users could invoke a mock screen which showed a dynamic selection UI with a purchase button next to the detected intent. Users, when they were ready to make a purchase, could invoke this UI and click the purchase icon. A modified version of this client was used for our

evaluation phase. The association engine, user account management and recommender (rule based) is server based running on an AWS (Amazon Web Service) instance.

The trial and evaluation period was spread over 3.5 months. We collected 873 intents from the 53 users over approximately 11K URLs. Using this set, we built 903 association rules. The second activity was to validate our hypothesis that intent behavior could be affected through effective content recommendations. 88 users participated in this trial over a 45-day period. We benchmarked user click through rate (CTR) against Google AdWords in order to determine whether provisioning additional intent information can help with user conversions. We omit implementation details due to space constraints.

AdWords can track user activity through Google ID across domain and service provided ID linking if coupled with Google Analytics backend. Google uses a combination of search keywords, extracted keys, keyword bids, quality score (for ads) and cost-per-click (CPC) factors to determine what advert to serve to user. Our end goal was to make this more relevant by feeding contextually, intent keywords at the right context through Google Ads or other services such as Yahoo Bing Network and Samsung’s own Ad platform. The program scope was limited to evaluating effectiveness of intent based recommendations through a simulated UI in addition to Google Ads. We base our observations on relative CTR with respect to Google Ads. Our comparative analysis only looks at the time when we do an intent based recommendation to user. This happened for under 5% of total user session time. Ad word interactions (as a non-recommended set) during the overlap time per user was collected over GA data by using Google Ad id of user while intent recommendations were collected as a custom dimension over GA. Figure 2 shows the CTR differences between the two sets. A second collection looked at a specific category of purchase. We figured that decision time taken for a purchase differed between categories. Seasonal factors can also affect decision times. We found that mobile category had an overlap of 28% on the topics searched, and so we analyzed specifically the topic of mobile related purchases. Twenty-seven users from the content recommended set ended up purchasing within the mobile category while corresponding number of buyers from the

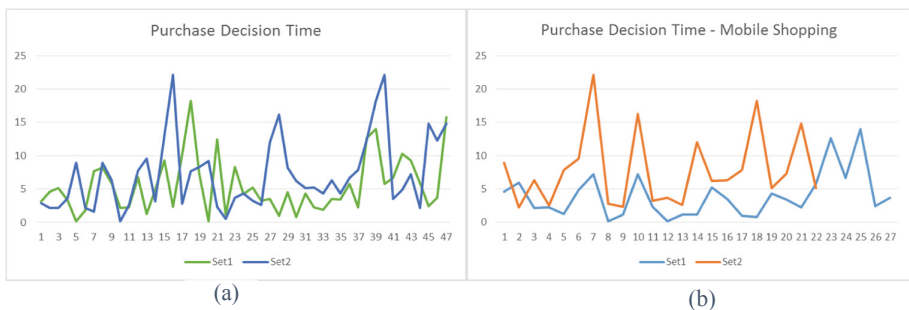


Fig. 2. (a) Purchase decision time plot: Set 1 represents recommended content group. Set 2 represents non-recommended group. (b) Purchase decision time plot for mobile category: Set 1 represents recommended content group. Set 2 represents non-recommended group.

non-recommended set was 22. We used the two different data sets to validate two hypotheses.

1. There is a significant difference on purchase decision time between the intent determined content recommendations set vs. the non-recommended ad-word set over all shopping categories.
2. There is a significant difference on purchase decision time between the intent determined content recommendations set vs. the non-recommended set on any particular shopping category (in this case mobile shopping category).

We used the two means t-test for statistically evaluating our hypothesis. The values within the two sets were independent and we assumed a normal distribution of the values. Set 1, in Fig. 2a, shows a sample plot for average CTRs for 44 users belonging to the recommended content group. Set 2 gives the corresponding plots (average CTR) for the users based on AdWords recommendation. Set 1 plots the purchase decision time, recorded as day counts, for the 44 user set who were provided with recommended content once the intent was captured based on the pre-modeled association rules. For Set 1, the mean value for the purchase decision time was 5.15 days with a standard deviation of 4.17. The average decision time for a purchase was 7.16 days with a standard deviation of 5.38 in Set 2. Our t-test evaluation over the two distributions, taking an un-paired two-tail analysis gave a p-value of 0.1002. Given that the p-value exceeds the alpha value of 0.05, we cannot claim statistical significance for our hypothesis 1. One reason for this may be that the data has a wide spread of category specific purchases. In addition, since the concentration of categories is non-uniform, this would affect the average purchase time as well as the deviation.

Figure 2b shows the plot for purchase decision time taken within the mobile shopping category by two different user sets. The mean value for the content recommended group (set 1) was 3.97 days as compared to decision time of 7.87 days for users within set 2. Set 1 also showed a smaller standard deviation of 3.44 as against 5.58 for users within set 2. The un-paired two-tailed p-value for our hypothesis 2 was 0.007268 indicating a high statistical significance. The value indicates a confidence > 99% that our hypothesis holds, meaning there is a significant difference between purchase decision times on mobile category between the two sets.

The average time and lower deviation also suggests acceleration in purchase decision times. We believe this validates our assumption that recommending latest content based on intent detection might work to reduce average purchase decision times. We inferred that keeping intent “alive” was key to boosting intent transitions achieved through our content recommendation service. The rate of keeping “alive” i.e. recommendation rate was however user specific (according to their preference), as was revealed through our user questionnaire. We acknowledge that this is still a small sample set and validated for a single category. It is difficult to validate hypothesis 1 given a mixture of purchases across categories. To do a full evaluation, we need to consider multiple category purchase timelines by users.

6 Conclusion

As part of our user insight activity, we have built an intent mining engine, as an experimental extension to our browser (Samsung Internet). The intents were created based on topic inference over web articles browsed by user and maintained within client till the user made a purchase decision based on the intent. Our experimental activity for intent behavioral inference, conducted using 53 volunteers gave interesting insights into each of these states, helping us to derive a user model that brought in a quantitative measure to intent calculations. Based on the behavioral data collection activity, we built a further unsupervised association rule set allowing us to test our hypothesis, using a further 88 volunteer set, as to whether targeted recommendations can affect intent behavior. We found that intent state transitions can be affected if done over specific categories. Our future task will to be increase our test user base through a beta release of the system with more user controls and permissions on mining intent categories. We aim to study correlations between demography and intent to check if intent dependent personalized content can further reinforce and assist in purchase decisions.

References

1. KPMG: Truth about online consumers, 2017 Global Online Consumer Report (2017)
2. Google Analytics. <https://analytics.google.com/analytics/web>
3. Microsoft Azure Analytics. <https://azure.microsoft.com/en-us/product-categories/analytics/>
4. Kumar, A.H., John, S.F., Senith, S.: A study on factors influencing consumer buying behavior in cosmetic products. *Int. J. Sci. Res. Publ.* **4**(9), 6 (2014)
5. Day, D., Gan, B., Gendall, P., Esslemont, D.: Predicting purchase behavior. *Market. Bull.* **2**, 18–30 (1991). Article 3
6. Bell, S., Bala, K.: Learning visual similarity for product design with convolutional neural networks. *ACM Trans. Graph. (TOG)* **34**(4) (2015). SIGGRAPH 2015. Article no 98
7. Cesar, A.C.: Impact of Consumer Attitude in Predicting Purchasing Behaviour (2007)
8. Arulkumar, S., Kannaiyah, D.: Predicting purchase intention of online consumers using discriminant analysis approach. *Eur. J. Bus. Manag.* **7**(4), 319–324 (2015)
9. Pal, S.: Know your buyer: a predictive approach to understand online buyer's behavior', white paper, Happiest Minds
10. Banerjee, N., Chakraborty, D., Joshi, A., Mittal, S., Rai, A., Ravindran, B.: Towards analyzing micro-blogs for detection and classification of real-time intentions. In: *International Conference on Web and Social Media* (2012)
11. Rose, D.E., Levinson, D.: Understanding user goals in web search. In: *Proceedings of the 13th Conference on World Wide Web* (2004)
12. Ioanas, E., Stoica, I.: Social media and its impact on consumers behavior. *Int. J. Econ. Pract. Theor.* **4**(2), 295–303 (2014)
13. Guo, S., Wang, M., Leskovec, J.: The role of social networks in online shopping: information passing, price of trust, and consumer choice. In: *Proceedings of the 12th ACM Conference on Electronic Commerce*, pp. 157–166 (2011)
14. Sathish, S., Patankar, A., Neema, N.: Semantics-based browsing using latent topic warped indexes. In: *International Conference on Semantic Computing, ICSC 2016*

15. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. *J. Mach. Learn. Res.* **3**, 993–1022 (2003)
16. Finkel, J.R., Manning, C.D.: Nested named entity recognition. In: *Conference on Empirical Methods in Natural Language Processing*, vol. 1, pp. 141–150 (2009)
17. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules. In: *Proceedings of the 20th VLDB Conference, Chile* (1994)
18. Sathish, S., Patankar, A., Priyodit, N.: Enabling multi-topic and cross-language browsing using web-semantic service. In: *International Conference on Web Services (ICWS)* (2017)