Exploring User Engagement in Information Networks: Behavioural – based Navigation Modelling, Ideas and Directions

Vesna Kumbaroska
University of Information Science and Technology, St. Paul the Apostle, Ohrid, Macedonia
vesna.gega@uist.edu.mk

Pece Mitrevski
Faculty of Information and Communication Technologies, University St. Clement of Ohrid, Bitola, Macedonia
pece.mitrevski@fikt.uklo.edu.mk

Abstract

Revealing an endless array of user behaviors in an online environment is a very good indicator of the user's interests either in the process of browsing or in purchasing. One such behavior is the navigation behavior, so detected user navigation patterns are able to be used for practical purposes such as: improving user engagement, turning most browsers into buyers, personalize content or interface, etc. In this regard, our research represents a connection between navigation modelling and user engagement. A usage of the Generalized Stochastic Petri Nets concept for stochastic behavioral-based modelling of the navigation process is proposed for measuring user engagement components. Different types of users are automatically identified and clustered according to their navigation behaviors, thus the developed model gives great insight into the navigation process. As part of this study, Peterson's model for measuring the user engagement is explored and a direct calculation of its components is illustrated. At the same time, asssuming that several user sessions/visits are initialized in a certain time frame, following the Petri Nets dynamics is indicating that the proposed behavioral – based model could be used for user engagement metrics calculation, thus some basic ideas are discussed, and initial directions are given.

Keywords: user behaviour, navigation behavior patterns, Generalized Stochastic Petri Net, Continuous Time Markov Chain, user engagement.

1. Introduction and background

There are plenty of definitions related to the user engagement (Arapakis et al., 2014; O'Brien & Toms, 2008), but we give prominence to the following (Lehmann et al., 2012): "User engagement is the quality of the user experience that emphasizes the positive aspects of the interaction, and in particular the phenomena associated with being captivated by a web application, and so being motivated to use it." Measuring user engagement is a complex process that is strictly dependent on the web site application, thus there are not unique measures for it (Lehmann et al., 2012). Basic analytic metrics such as session duration, number of visits per session, number of sessions/visits per user, clickthrough rates, time spent on site, number of page views, etc., are mentioned in: Antovski & Armenski, 2012; Arapakis et al., 2014; Lehmann et al., 2012; O'Brien & Toms, 2008. The conversion rate and the abandonment rate are other metrics concerned with user engagement, loyalty and experience. The first one refers to the percentage of users who successfully completed a desired action. The second one is more related to the electronic commerce, meaning the percentage of users who did not become buyers, in other words, left their carts without a purchase. Another approach where user (visitor) engagement is defined as a function of several indexes (components) is described in Peterson & Carrabis (2008): "Visitor Engagement is a function of the number of clicks (C_i), the visit duration (D_i), the rate at which the visitor returns to the site over time (R_i), their overall loyalty to the site (L_i), their measured awareness of the brand (B_i), their willingness to directly contribute feedback (F_i) and the likelihood that they will engage in specific activities on the site designed to increase awareness and create a lasting impression (I_i)". In this research, we propose a behavioral – based modelling considering the most commonly identified

user navigation patterns across information networks, illustrated through the example of an electronic bookstore. One of our goals is to indicate that such behavioral – based model which provides profound knowledge on the processes of navigation could be used for user engagement metrics calculation. As another contribution, we illustrate a direct calculation of the user engagement components of Peterson's model (Peterson & Carrabis, 2008). Additionally, we compute conversion rate and abandonment rate metrics, thus obtaining a more detailed picture of the overall user experience.

Usually, user navigation behaviour modelling relies on predictive mathematical models built on log data collected in a period of time (Pass et al., 2006; Ruthven, 2011). This is massive data, usually with complex structure, and its analysis discovers important information about the users and their behaviour in different aspects. Statistical analyses and applications of data mining techniques are usually applied to study such kind of data in order to gain knowledge of the user's actions: discovering user search or navigation patterns (Awad & Khalil, 2012; Tadeusz et al., 2000), predicting and proposing future user actions and personalization based on user behaviour (Leino, 2014; Schafer et al., 2001), etc.

Riivo & Potisepp (2009) used a data mining approach to discover common user navigation behaviour patterns and sequences of transitions based on the website visit duration. An application of another data mining approach in combination with a semi-Markov process in discrete time is mentioned in Jenamani et al. (2003), once again, to understand and describe the user behaviour in order to improve the design of the website and the evaluation performance. A combination of log data and website structure analysis is used in the approach of Priyanka et al. (2010) for predicting user behaviour with the aim to improve the website performance. Both the discovering of the website structure and the prediction of the next user's action prediction are realized using the concept of Petri Nets. For the same purpose – predicting the next user's action, Bahadori et al. (2013) developed a model using Coloured Petri Nets, based on former user profiles and current user sessions. Modelling and analysing the structure of the website using another Petri Nets formalism for path completion is found in the work of Shih-Yang et al. (2007), where web pages are presented as states, and transitions as arcs. An evolution of this approach can be found in the research of Po-Zung et al. (2008), where the concept of Stochastic Time Petri Nets is implemented for website structure modelling and subsequent future user behaviour predicting. A behaviour model of an electronic store costumer exploiting the Stochastic Petri Nets (Ajmone et al., 1984) in order to improve the quality of web services, reliability, performance and availability is suggested by Mitrevski et al. (2014). Detailed investigation of user online search and navigation behaviour in enclosed spaces is done by Yongli et al. (2014) in order to improve user satisfaction when using the Internet and doing online shopping.

Another interesting aspect of user navigation behaviour is the clustering process. For example, the research of Su & Chen (2014) presents a new approach for clustering patterns revealed of navigational data from Chinese electronic store, which combines navigational paths, visit frequency and retention on a web page or category. Sequence clustering is specific type of clustering which utilizes not only the sequences of transitions but also their order as a way to discover useful information about the user behaviour. One such method used to group user behaviour according to the similarity of the user actions order is presented by (MacLennan & Tang, 2005), where combination of standard clustering methods and techniques to analyse sequences based on Markov Chain is used.

Our paper adds to this work, exploring the concept of Generalized Stochastic Petri Nets for modelling and analysis of navigation patterns across information networks. The model complex solution relies on Continuous Time Markov Chain and it is applied on two dynamically detected types of users, based on their traverse paths (Gega & Mitrevski, 2015a; Gega & Mitrevski, 2015b). User navigation behavior can be used to analyze and measure the user engagement. Thus, as another contribution, we propose application of the developed behavioral – based model for measuring user engagement in information network.

The rest of the paper is organized as follows. Section 2 represents an overview of the Petri Net formalism. Also, details about the model, the chosen scenario and the clustering methodology are provided. Section 3 focuses on the Peterson's user engagement model, thus direct application of this model is shown and described. Additionally, this section presents some basic ideas and initial directions towards user engagement metrics calculation using the proposed behavioural – based model. Finally, we conclude the paper and give future research direction in Section 4.

Using petri nets to capture navigation behaviour patterns

The basic elements of PN are: places, transitions, tokens and arcs. Usually, places are presented as circles, transitions as rectangular boxes and tokens as black dots. Places are related to states, transitions are related to actions that can change the states and arcs determine directed relation between places and transitions (Molloy, 1982; Murata, 1989). The marking of PN is closely related to tokens and it is used to describe the dynamic behaviour of the system. In that context, a transition is enabled if all its input places contain at least one token. An enabled transition can fire by removing one (more) token from all its input places, and adding one (more) token in all its output places, following the arcs.

The behaviour in a standard PN is discrete only. It means all the transitions are instantaneous or fire instantly. An extension of this concept is GSPN (Ajmone et al., 1984; Dingle et al., 2009), where immediate and timed transitions are introduced. Here, the firing delays of timed transition are stochastic, usually exponentially distributed random variables. It means a timer is associated to each enabled timed transition, in such way that the timer value is sampled from (negative) exponential distribution with appropriate rate parameter. The timer constantly decreases, and when its value will become zero, the timed transition will fire. Usually, immediate transitions are presented as black rectangular boxes or bars, and timed transitions as white rectangular boxes. The immediate transition fire with priority over timed transitions. The markings could be tangible and vanishing. A tangible marking is a marking where only timed transitions are enabled. Contrary to this, a vanishing marking is a marking where immediate transitions are enabled (or combination of immediate and timed transitions are enabled).

The dynamic nature of PN, especially GSPN, indicates that they can be accommodated for modelling real user navigation behaviour. The focus in our research is placed on discovering and modelling navigation behaviour of bookstore users, specifically examined in the case of the first Macedonian electronic bookstore[2] (Antovski & Armenski, 2012;), but easily generalized and applicable on other much known services, which structure is similar to the selected scenario.

Case study model

The navigation data used for this study is collected server side in a certain time frame and it is in a standard W3C format. The log file contains about 415000 records so that each record contains data belonging to several categories (attributes), but for this research particularly important categories are: userID, Date and Time and URL visited. In order to transform the data into easily interpretable format for further usage, it was necessary to do pre-processing and data cleaning tasks. It means removal of all incomplete records or records that lack any of the key attributes or they are inconsistent. We dynamically identify users, visits, and page views per user and per visit, as shown in table 1.

Table 1. Information about the log file

The total number of records analyzed	322000		
Unique users	1984		
Unique visits	15433		

[2] www.kupikniga.mk

	Mean	Min/Max	SD
Time between two successive visits	3d:20h:06m:59s	(0d:0h:0m:0s, 332d:23h:52m:57s)	
Visits per user	7.78	(1, 229)	13.89
Views per visit	20.85	(1, 2929)	52.60

Prior to model construction, there was a need to understand the web site topology, which means creating representation of the web site from the log data, as it is shown in figure 1. All web pages revealed are classified into four categories, based on their use. The 'A' category web pages are related to the products sold, in our case books. Web pages referred to general store information and help are categorized as 'B'. Next, the 'C' category web pages are only available for logged users, and they are associated with user's profile information. At the end, the 'D' category web pages are directly connected with the purchasing process.

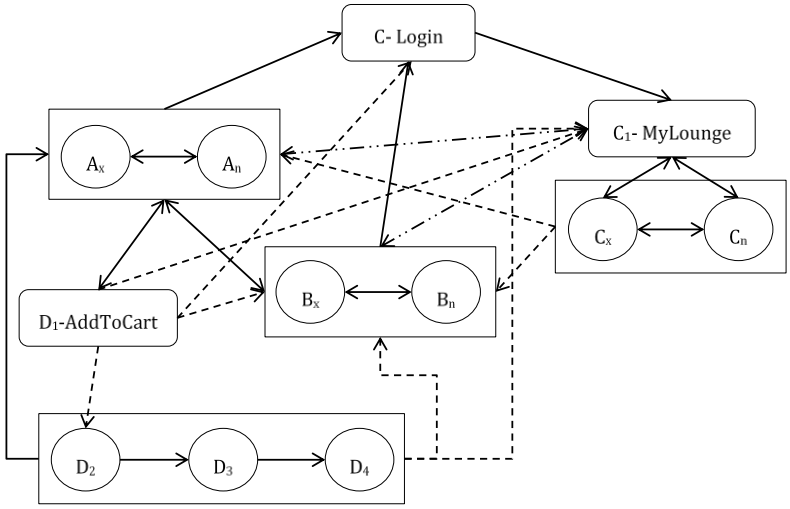

Figure 1. Web site representation

Based on the web site topology, using GSPN notation, our model comprises 10 places, and 44 timed transitions, as shown in table 2 and table 3, respectively.

Table 2. Places in the GSPN model

Place name	Place description	Initial marking
A	A category page	1
B	B category page	0
E	Ended user session	0
L	Login	0
ML	MyLounge (MyBookStore)	0
C	C category page	0
D1	AddToCart	0
D2	AddressEntry	0
D3	LastCartPreview	0
D4	PayingCasys	0

Table 3. Transitions in the GSPN model

Transition name	Transition description	Rate
tA_cont, tA1, tA2, tA3, tA4, tA5, tA6	Visit an A category page	α
tB, tB_cont, tB1, tB2, tB3, tB4, tB5	Visit a B category page	λ
tE_A, tE_B, tE_L, tE_ML, tE_C, tE_D1, tE_D2, tE_D3, tE_D4	End the session	μ

tL, tL1, tL2, tL_cont	Login	κ
tML, tML1, tML2, tML3, tML4, tML5, tML_cont	Visit MyLounge page	ν
tC, tC_cont	Visit a C category page	θ
tD1, tD1_cont	Visit AddToCart page	ε
tD2, tD2_cont	Visit AddressEntry page	γ
tD3, tD3_cont	Visit LastCartPreview page	δ
tD4, tD4_cont	Visit PayingCasys page	β

The GSPN dynamics of the model is graphicaly depicted in figure 2, and can be described using the terminology of CTMC (Ajmone et al., 1984; Murata, 1989).

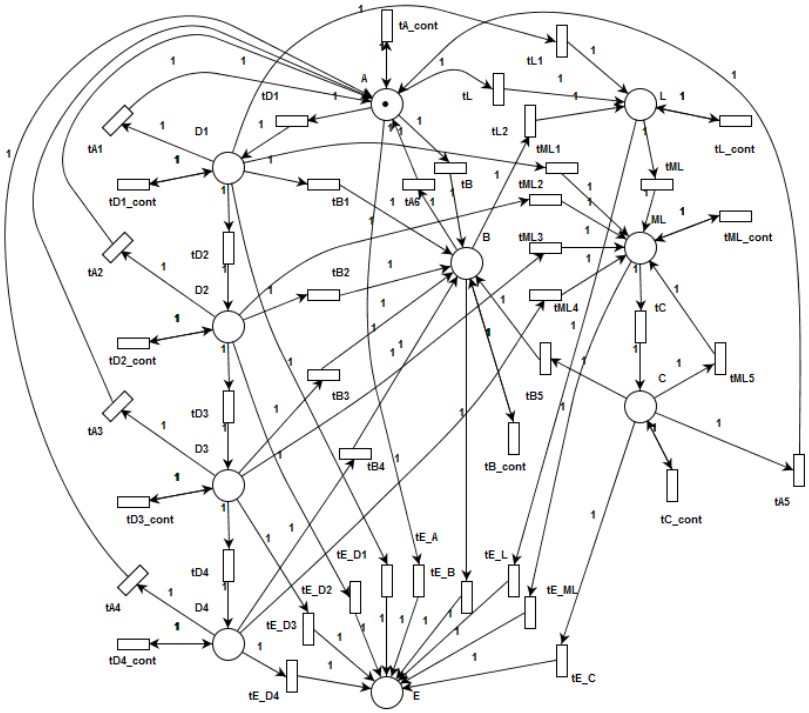

Figure 2. Graphical model presentation

Clustering methodology

Detected user behavior from navigation patterns indicates that there are different types of users. Therefore, we group the users based on their navigation patterns, applying the sequence clustering algorithm proposed by (MacLenann, 2005). According to the set of input parameters and the data set, the users are automatically grouped into two clusters, as an optimal solution for this problem. As illustrated in figure 3, the number of users in each cluster is approximately equal, 916 in the first cluster and 1068 in the second cluster.

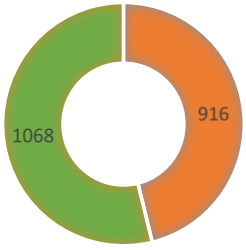

Figure 3. Number of users in the two clusters

The visual cluster profiling is given in figure 4, thus each column represents a cluster, each row represents a sequential attribute and each cell contains a histogram of user actions sequence.

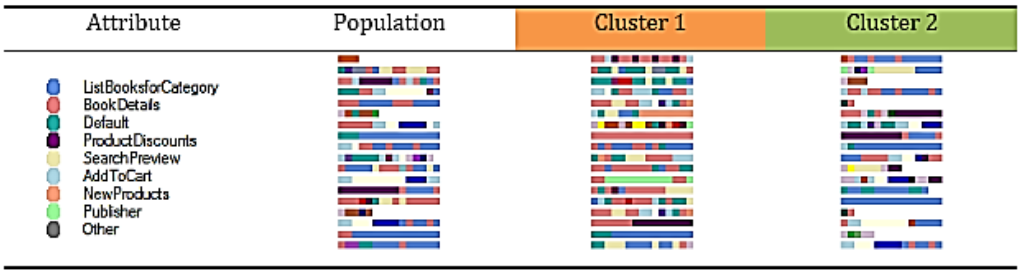

Figure 4. Cluster profiles

From the visual inspection of figure 4, we can notice that most prevalent actions differ in the both clusters. Such a difference is the following: in the first cluster's sequences, seeing a book details as an A category page, is most common, versus the second cluster where users prefer filtering books by category, also as an A category page. Deeper insight is given in figure 5 and figure 6, where characteristic transitions for the both clusters are visualized. Actually, these diagrams represent Markov chains where each node is a sequential state and each arc is a transition from one state to another, which means it is directed. Also, each arc contains a weight which is related to the transition probability. The diverse colour of the node is linked to the node popularity, which means more popular states are darker and vice versa. In this direction, we can discuss some of the most frequent navigation behavior patterns for the first cluster, demonstrated in figure 5:

- 40% of the users start their visit seeing details for a book ('A' category). Around half of them (47%) choose to visit the same page again later.
- 22% of the users start their visit with the default page ('A' category). Also, around 40% of these users visit the same page again.
- Similar, around 20% of the users start their visit previewing their lounge ('C' category), and only 22% of them decide to review their orders.
- One third of the users that have bought a book ('D' category) choose to visit the default page ('A' category).
- Similar, 32% of the users that have changed the delivery address ('C' category), choose to see their cart ('D' category), but almost 62% of them decide to see their cart again later.
- etc.

Similar, some of the most frequent navigation behavior patterns for the second cluster, demonstrated in figure 6 are:

- 35% of the users start their visit seeing details for a book ('A' category). One third of them (36%) choose to visit the same page again later.
- 22% of the users start their visit previewing their lounge ('C' category).
- 66% of the users which have invited friends ('C' category) choose to ask for help ('A' category) and 37% of those users end the visit.
- 27% of the users that have changed the delivery address ('C' category), choose to see their cart for the last time ('D' category), but almost 64% of this users decide to see their cart again later.
- etc.

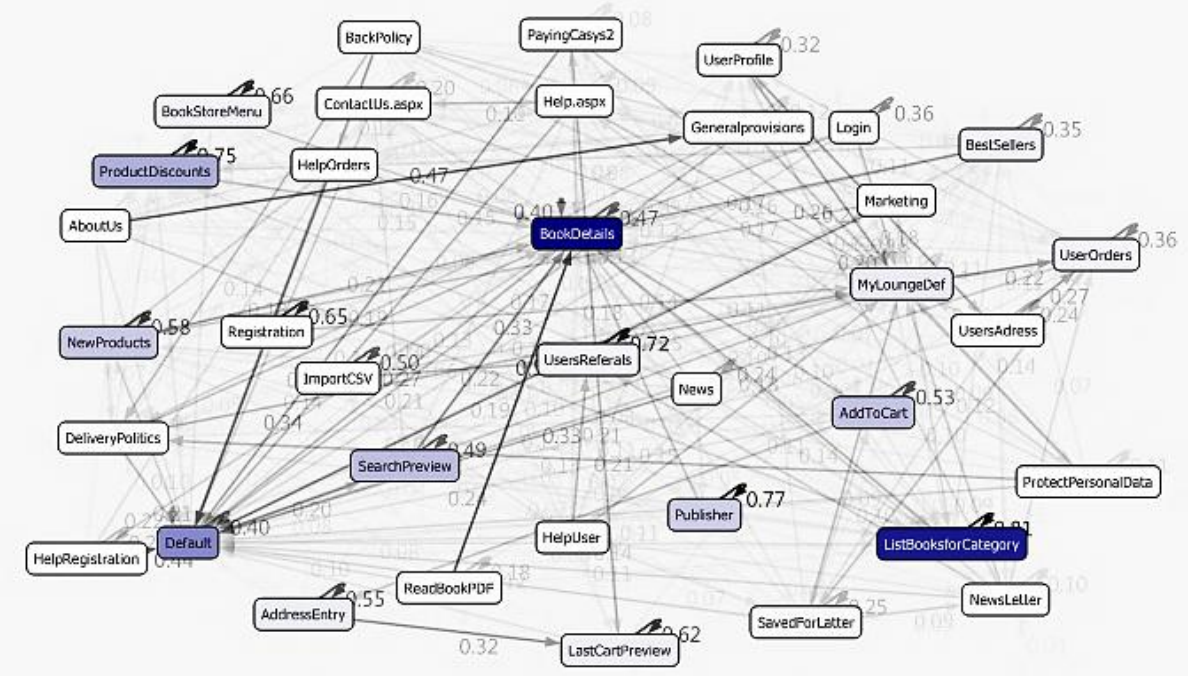

Figure 5. State transition diagram for Cluster 1

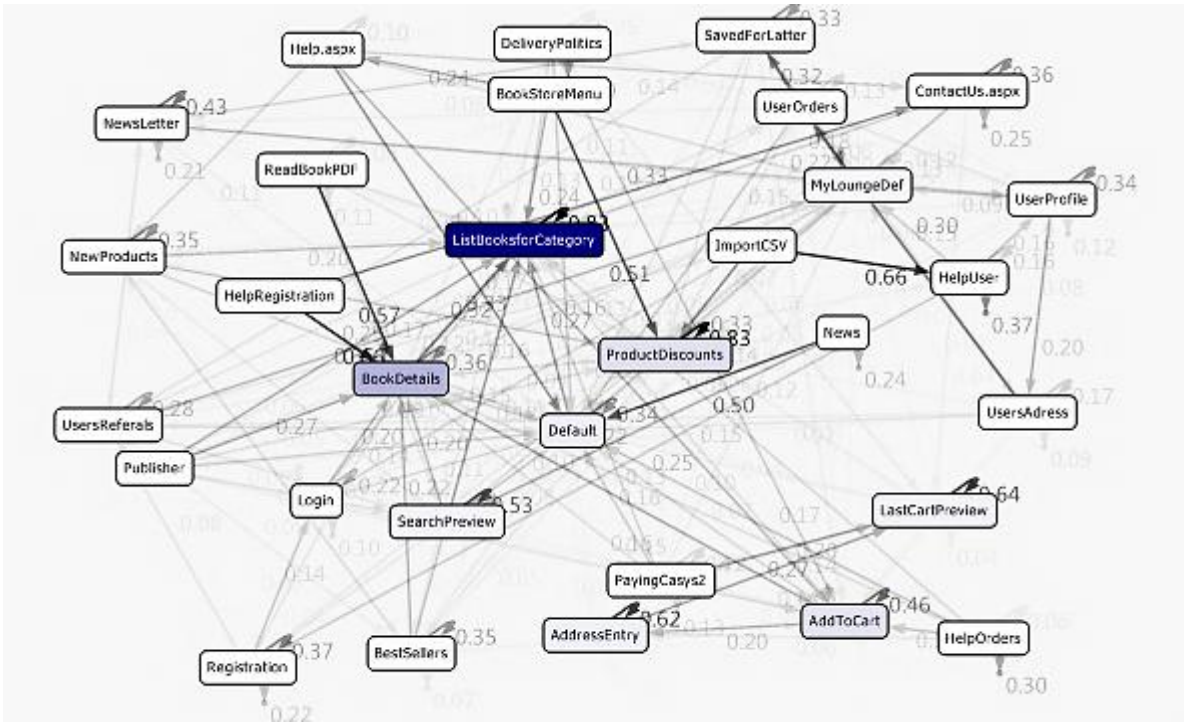

Figure 6. State transition diagram for Cluster 2

In this regard, the understanding of the stochastic process that underlies this GSPN model and evaluation of appropriate performance measures for both clusters, is subject of our other study (Gega & Mitrevski, 2015a; Gega & Mitrevski, 2015b), where we use efficient time and space algorithm for computing steady state solutions of deterministic (DSPN) and stochastic (SPN) Petri Nets, proposed by (Bause & Kritzinger, 1996; Ciardo et al., 1989). The obtained results include the average sojourn time in each of the transient states, the total time spent in these transient states, the average number of visits, as well as the cumulative sojourn time. This is not the scope of this paper.

'Measuring the immeasurable' in practice

In the Peterson's model (Peterson & Carrabis, 2008), as we mentioned in the introductory section, the user engagement is defined as a composition of seven different components (indexes). In this section we use the quantitative data obtained from the log analysis in order to compute some of these components.

$$UE = \sum(C_i + D_i + R_i + L_i + B_i + F_i + I_i) \quad (1)$$

The first component (C_i) is ClickDept index and it is related to the percentage of users who have number of page views per session greater than or equal to some threshold. On user level, this index is calculated as:

$$C_i = Sessions\ having\ at\ least\ \text{"threshold"}\ page\ views\ /\ All\ sessions \quad (2)$$

The second component (D_i) is Duration index and it refers to the percentage of users whose session duration is greater than or equal to some threshold. Once again, on user level, it is computed as:

$$D_i = Sessions\ having\ duration\ over\ \text{"threshold"}\ \text{time}/\ All\ sessions \quad (3)$$

The percentage of users who come and leave the web site within some time frame, usually expressed in days, is represented by the Recency index (R_i). With other words, it is web site visiting frequency, thus the index value is directly proportional to the sessions frequency. For each user separately, the index is computed as:

$$R_i = 1/Number\ of\ days\ elapsed\ since\ the\ most\ recent\ session \quad (4)$$

The Loyalty index (L_i) is the fourth component, concerning the percentage of users whose number of reoccurring sessions is greater than or equal to some threshold. The basic formula on user level is:

$$L_i = 1 - (1/Number\ of\ visitor\ sessions\ during\ the\ time\ frame) \quad (5)$$

The percentage of users who come to the site directly or by searching branded keywords is presented by the Brand index (B_i). It means the level of attention the visitor is paying to the brand prior to visiting the web site. Next, the sixth component (F_i) is Feedback index and it concerns the percentage of users who leave feedback, comment or rating, on the web site. For each user it is calculated as:

$$F_i = Sessions\ where\ visitor\ submits\ feedback/All\ sessions \quad (6)$$

The last component is (I_i) is Interaction index and it is related to the percentage of users who visit or have interaction with a specific content on the web site, for example, adding an item to a cart or placing an order. The corresponding formula on user basis is:

$$I_i = Sessions\ where\ visitor\ completes\ an\ action/All\ sessions \quad (7)$$

If one of the indexes in formula 3.1 could not be calculated, it should be replaced with zero, which will not affect the result at all.

Determining user engagement components from log data

Instead of sessions used for the calculations in the original work (Peterson & Carrabis, 2008), this research work illustrates application of the time between two subsequent visits. In this section we discuss the threshold values and results determined from the log data for the both clusters respectively.

For the first cluster it is found that the largest percentage of users (around 48%) have visits with average page views per visit between 1 and 10. Similar, in the second cluster, more than a half of the users (around 59%) also have visits with average page views per visit between 1 and 10. It is shown in figure 7, for the both clusters respectively. This threshold is used for ClickDepth index calculation, thus the distributions of C_i, for all 916 users over 7570 visits in the first cluster and for all 1068 users over 7863 visits in the second cluster, are depicted in figure 8.

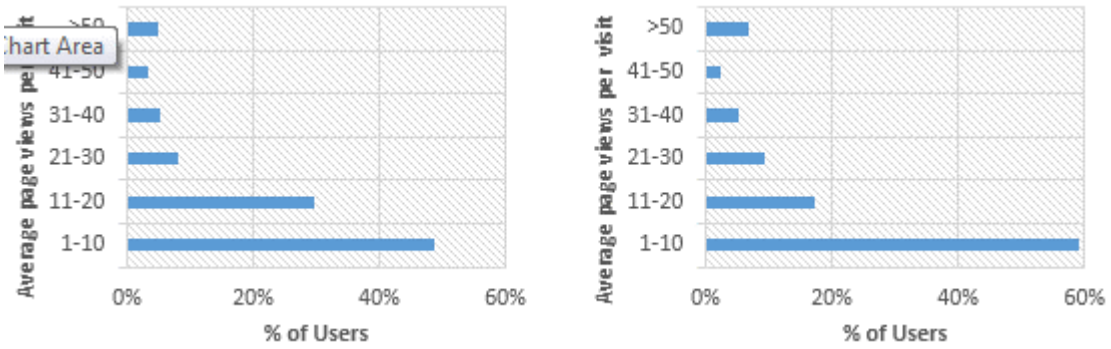

Figure 7. Average page views per visit in Cluster 1 (left) and Cluster 2 (right)

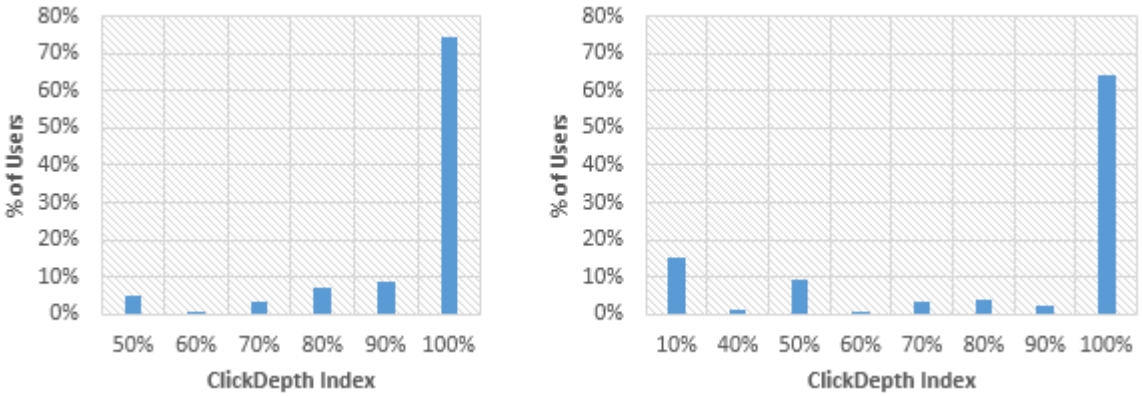

Figure 8. Distributions of C_i for all 916 users over 7570 visits in Cluster 1 (left) and for all 1068 users over 7863 visits in Cluster 2 (right)

For the first cluster it is found that the largest percentage of users (around 17%) have visits with average length between 1 and 10 minutes. In the second cluster, around one third of the users (27%) have visits with average length less than 1 minute and between 1 and 10 minutes. It is graphically depicted in figure 9, for the both clusters respectively. This threshold is used for Duration index calculation, thus the distributions of D_i, for all 916 users over 7570 visits in the first cluster and for all 1068 users over 7863 visits in the second cluster, are depicted in figure 10.

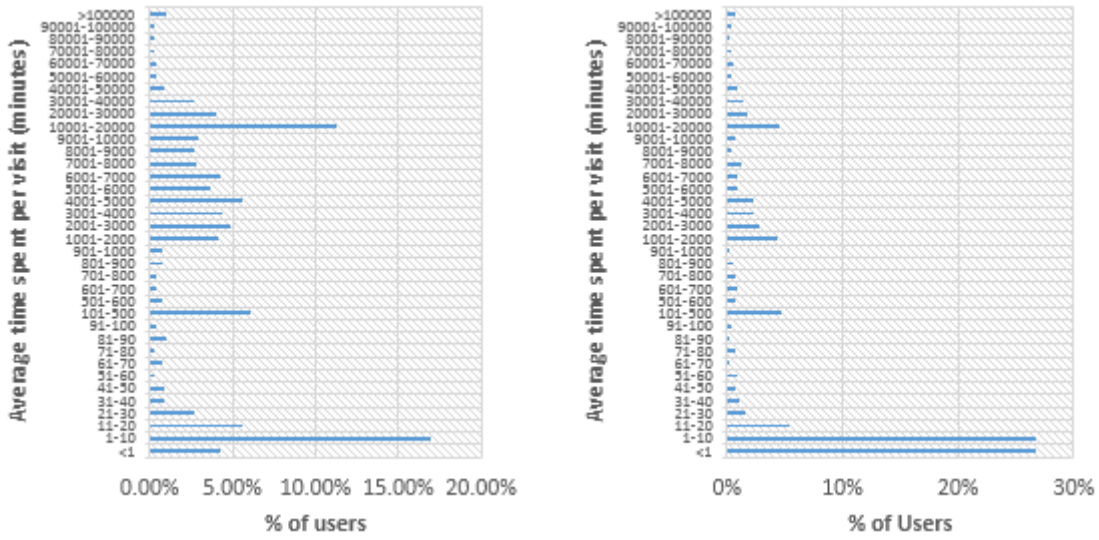

Figure 9. Average visit length in Cluster 1 (left) and Cluster 2 (right)

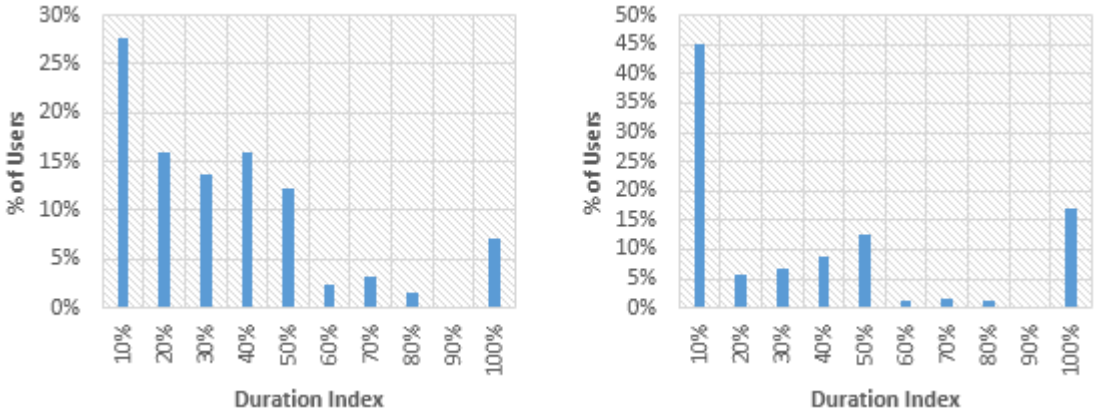

Figure 10. Distributions of D_i for all 916 users over 7570 visits in Cluster 1 (left) and for all 1068 users over 7863 visits in Cluster 2 (right)

Our study discovers that most of the users return to the web site in the rage of 1 to 10 days. Specifically, 31% of the users in the first cluster and 58% of the users in the second cluster return to the site in that range. Graphical representation is given in figure 11. This range is considered as a threshold used for Recency index calculation, thus the distributions of R_i, for all 916 users over 7570 visits in the first cluster and for all 1068 users over 7863 visits in the second cluster, are shown in figure 12.

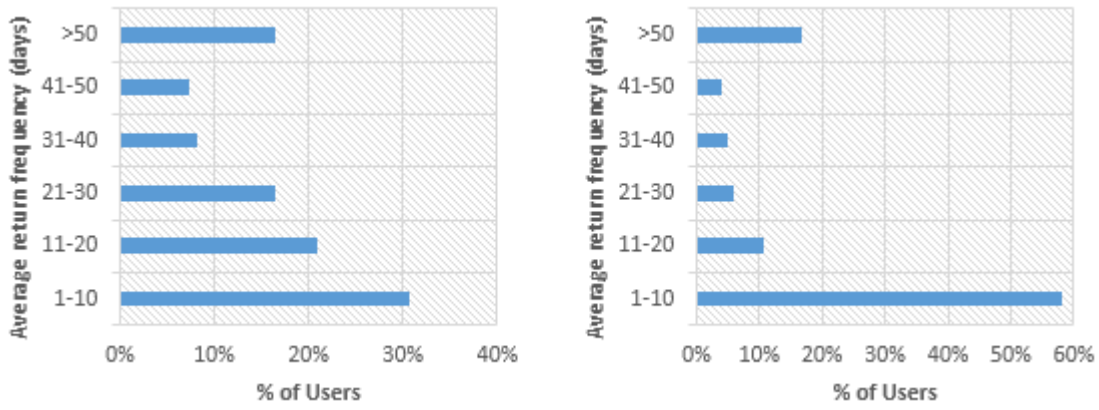

Figure 11. Average return frequency in Cluster 1 (left) and Cluster 2 (right)

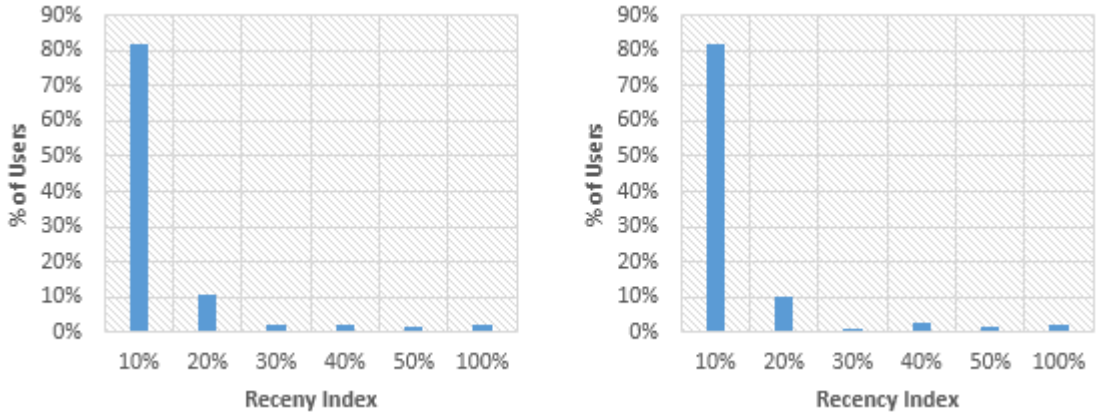

Figure 12. Distributions of R_i for all 916 users over 7570 visits in Cluster 1 (left) and for all 1068 users over 7863 visits in Cluster 2 (right)

Also, we reveal that very large portion of the users, more accurately, 74% of the first cluster and 86% of the second cluster, have between 1 and 10 reoccurring visits, as it is shown in figure 13. This threshold is used for Loyalty index calculation, thus the distributions of L_i, for all 916 users over 7570 visits in the first cluster and for all 1068 users over 7863 visits in the second cluster, are shown in figure 14.

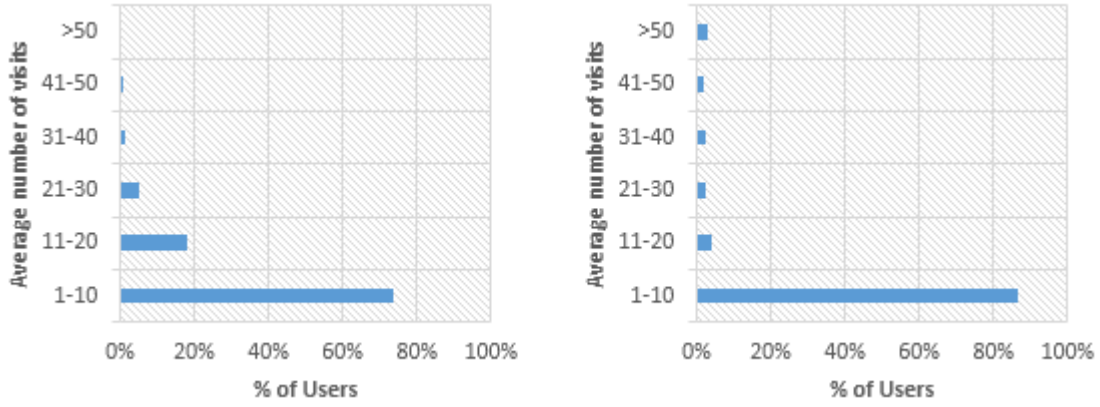

Figure 13. Average number of visits in Cluster 1 (left) and Cluster 2 (right)

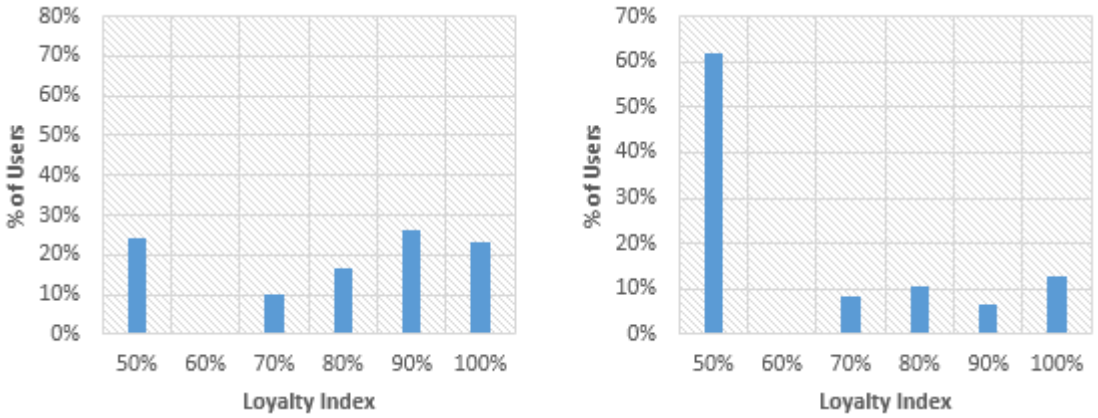

Figure 14. Distributions of L_i for all 916 users over 7570 visits in Cluster 1 (left) and for all 1068 users over 7863 visits in Cluster 2 (right)

For I_i calculation, we decide to use adding a book to the cart and buying a book as content of interest. The distributions of I_i, for all 916 users over 7570 visits in the first cluster and for all 1068 users over 7863 visits in the second cluster, are shown in figure 15.



Figure 15. Distributions of I_i for all 916 users over 7570 visits in Cluster 1 (left) and for all 1068 users over 7863 visits in Cluster 2 (right)

Our log doesn't contain data that can be used for B_i and F_i indexes calculation, so they are replaced with 0 in the final formula. The average user engagement in the first and the second cluster respectively is:

$$UE_cluster1 = (\mathbf{89.80} + 69.24 + 27.93 + \mathbf{62.62} + 0 + 0 + 38.88)/7 = 41.21\%$$
$$UE_cluster2 = (86.05 + \mathbf{81.00} + \mathbf{30.44} + 34.78 + 0 + 0 + \mathbf{57.24})/7 = 41.36\% \quad (8)$$

According to the results, we see that some components in the first cluster have higher values than in the second cluster and vice versa, but the final outcome shows that users in the second cluster are a little bit more engaged than the users in the first cluster.

Additional to this components we compute conversion rate and abandonment rate, thus we get more detailed picture of the whole process in relation to the set of unique users actions. According to the log data, the web site is visited by 1984 unique users but only 362 of them bought a book, which leads to 18.25% conversion rate for buying a book. Also, one third of 961 users who added something to the cart become buyers and two thirds of them left the cart without a purchase, which causes 62.33% abandonment rate. Specifically, in the first cluster only small portion of 916 users successfully bought a book, resulting in 20.21% conversion rate. This is slightly better than

the second cluster where only 16.57% from 1068 users have buying transactions. Similar to the global scenario, one third of 514 users in the first cluster and 449 users in the second cluster who added a book in the cart, become buyers. Also, two thirds of the users left the cart without a purchase, which results in abandonment rate of 64.01% in the first cluster and 60.58% in the second cluster. In the three cases, abandonment rate is within commonly accepted boundaries and near to the average[3].

Behavioral – based approach for determining user engagement components

Assuming that several user sessions/visits are initialized in a certain time frame, following the "token game" (Murata, 1989), the proposed behavioral – based model could be used for user engagement metrics calculation. In this section we discuss some basic ideas and give initial directions.

The C_i component could be calculated as average number of transition firings per unit of time, as given in formula (3.10). It is actually the frequency of firing a transition (Ajmone et al., 1984) calculated as given in formula 3.9. The transitions in our model are related to clicks, so actually we obtain the number of clicks during a visit and each visit has different number of clicks.

$$f_j = \sum_{M_i \in A_j} \lambda_j * \pi_i \quad (9)$$

$$C_i = \sum f_j * session\ or\ visit\ duration \quad (10)$$

Above, M_i is a marking in the set of markings A_j where the transition is enabled. The average sojourn time in a given marking can be calculated as a reciprocal value of the sum of the rates of the transitions enabled in that marking which lead the GSPN from that marking to other marking (Ajmone et al., 1984):

$$ST_i = \frac{1}{\sum_{t_j \in E(M_i)} rate_j} \quad (11)$$

Based on this, if we adjust the model to contain information about the distribution of the times between two successive visits, the component D_i could be easily calculated. Additionally, if this kind of model is generated on user level, the components R_i and L_i also could be determined. As we said in the previous section, the log doesn't contain data related to the B_i and F_i components calculation. From this fact we can also say that the model is not developed in direction for B_i и F_i components calculation. The I_i component could be calculated as total probability of taking some desired action. As we mentioned earlier, adding a book in the cart could be considered as an action of interest, thus following the model it is presented by the state D_4, which is reachable only by sequential firing of several transitions in the following order: $tD_1 \rightarrow tD_2 \rightarrow tD_3 \rightarrow tD_4$. According to this, I_i could be computed as:

$$I_i = P(t_{D1}) * P(t_{D2}) * P(t_{D3}) * P(t_{D4}) \quad (12)$$

Once again, the model doesn't include information about the number of users who took the action of interest, thus we equalize the conversion rate calculation with I_i component calculation, which means equalization of the term 'user' with the term 'time between two successive visits'. Abandonment rate could be calculated as:

$$abandonment = \left(1 - P(tD_4)\right) * P(tD_1) \quad (13)$$

[3] http://baymard.com/lists/cart-abandonment-rate

It expresses the probability of adding something to the cart (tD_1), but not making a purchase $\neg(tD_4)$.

Conclusion

This paper proposes a novel approach for calculating several user engagement components, based on a GSPN behavioral – based navigation model. We identified and analyzed the most common user navigation patterns across information networks, illustrated through the example of an electronic bookstore. The navigation data used in this study is collected server side in a certain time frame and contains about 415000 records.

First, the main contribution of this research work is the suggested behavioural-based model that provides profound knowledge about the processes of navigation, specifically examined for different types of users, automatically identified and clustered into two clusters according to their navigational behaviour. The developed model is based on stochastic modelling using the concept of Generalized Stochastic Petri Nets which complex solution relies on Continuous Time Markov Chain.

Second, we described several approaches for measuring the user engagement, but we gave special prominence to the Peterson's model (Peterson & Carrabis, 2008). After the observation of user navigation behavior from the log, we illustrated direct calculation of the Peterson's model user engagement components, for the both clusters of users separately. Additionally, we computed conversion rate and abandonment rate metrics, thus more detailed picture of the overall user experience is obtained.

Third, assuming that several user sessions/visits are initialized in a certain time frame, following the dynamic "token game" (Murata, 1989), we indicated that the proposed behavioral – based model could be used for user engagement metrics calculation, thus we discussed some basic ideas and gave initial directions. However, this suggested approach requires a lot of improvements and inclusions in order to enable calculation of all presented user engagement components, and that could be considered as a future direction.

References

Ajmone, M., Conte, G. & Balbo, G. (1984) 'A Class of Generalized Stochastic Petri Nets for the Performance Evaluation of Multiprocessor Systems', ACM Transactions on Computer Systems, Vol. 2, No. 2, pp. 93-122.

Antovski, L., & Armenski, G. (2012) 'KupiKniga.mk: transforming a website into a profitable e-commerce system using assisted conversions funnel', ICT Innovations 2012 Web Proceedings, ICT ACT, Skopje, pp. 321-329.

Arapakis, I., Lalmas, M., Cambazoglu, B. B., Marcos, M. C., & Jose, J. M. (2014) 'User engagement in online News: Under the scope of sentiment, interest, affect, and gaze', Journal of the Association for Information Science and Technology, 65(10), 1988-2005.

Awad, M., & Khalil, I. (2012) 'Prediction of user's web-browsing behavior: Application of markov model', Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 42(4), 1131-1142.

Bahadori, M., Harounabadi, A., & Sadeghzadeh, M. (2013) 'Modeling user navigation behavior in web by colored Petri nets to determine the user's interest in recommending web pages', Management Science Letters, Volume 3 Issue 1, pp. 359-366.

Bause, F., & Kritzinger, P. S. (1996) 'Stochastic Petri Nets' , Vieweg+ Teubner Verlag, pp: 163-171.

Chiola, G., Franceschinis, G., Gaeta, R., & Ribaudo, M. (1995), 'GreatSPN 1.7: graphical editor and analyzer for timed and stochastic Petri nets', Performance evaluation, 24(1), pp: 47-68.

Ciardo, G., Muppala, J. & Trivedi, T. (1989) 'SPNP: stochastic Petri net package', Petri Nets and Performance Models. PNPM89., Proceedings of the Third International Workshop on Digital Object Identifier, 10.1109/PNPM.1989.68548, pp. 142-151.

Dingle, N. J., Knottenbelt, W. J., & Suto, T. (2009) 'PIPE2: a tool for the performance evaluation of generalised stochastic Petri Nets', ACM SIGMETRICS Performance Evaluation Review, 36(4), pp. 34-39.

Gega, V., & Mitrevski, P. J. (2015a), Towards Modelling and Analyzing Navigation Behaviour Patterns Using Generalized Stochastic Petri Nets, International Conference on Applied Internet and Information Technologies (AIIT 2015), pp. XX-XX, Zrenjanin, Serbia.

Gega, V., & Mitrevski, P. J. (2015b) 'Using generalised stochastic Petri nets to model and analyse search behaviour patterns', International Journal of Reasoning-based Intelligent Systems, 7(1-2), pp. 26-34.

Jenamani, M., Mohapatra, P. K., & Ghose, S. (2003) 'A stochastic model of e-customer behavior', Electronic Commerce Research and Applications, 2(1), pp. 81-94.

Jensen, K. (2013) 'Coloured Petri nets: basic concepts, analysis methods and practical use', (Vol. 1), Springer Science & Business Media.

Leino, J. (2014) 'User Factors in Recommender Systems: Case Studies in e-Commerce', News Recommending, and e-Learning.

Lehmann, J., Lalmas, M., Yom-Tov, E., & Dupret, G. (2012) 'Models of user engagement', In User Modeling, Adaptation, and Personalization, pp. 164-175, Springer Berlin Heidelberg.

MacLennan, J., & Tang, Z. (2005) 'Data Mining with SQL Server 2005', Wiley.

Mitrevski, P.J. & Hristoski, I.S., (2014) 'Behavioral-based performability modeling and evaluation of e-commerce systems', Electronic Commerce Research and Applications 13(5), pp. 320-340.

Molloy, M. K. (1982) 'Performance analysis using stochastic Petri nets', Computers, IEEE Transactions on, 100(9), pp. 913-917.

Murata, T. (1989) 'Petri Nets: Properties, Analysis and Applications', Proceedings of the IEEE, Vol.77, No.4, pp. 541-580.

O'Brien, H. L., & Toms, E. G. (2008) 'What is user engagement? A conceptual framework for defining user engagement with technology', Journal of the American Society for Information Science and Technology, 59(6), pp. 938-955.

Pass, G., Chowdhury, A., Torgeson, C., (2006) 'A picture of search', Proceeding InfoScale '06 Proceedings of the 1st international conference on Scalable information systems, Article no.1.

Peterson, E. T., & Carrabis, J. (2008) 'Measuring the immeasurable: Visitor engagement', In Web Analytics Demystified, 14, 16.

Po-Zung, C., Sun, C., & Shih-Yang, Y. (2008) 'Modeling and analysis the web structure using stochastic timed Petri nets' Journal of Software 3.8, pp.19-26.

Priyanka, M., Gulati, P., & Sharma, A., (2010) 'A novel approach for predicting user behavior for improving web performance', International Journal on Computer Science and Engineering 2.04.

Riivo, K., & Potisepp, K. (2009) 'Click-stream Data Analysis of Web Traffic'.

Ruthven, I. (2011) 'Information Retrieval in Context', Advanced Topics in Information Retrieval, The Information Retrieval Series, vol. 33, pp. 187-207.

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.

Shih-Yang, Y., Po-Zung, C., & Sun, C. (2007) 'Using Petri Nets to enhance web usage mining.' Acta Polytechnica Hungarica 4.3.

Su, Q., & Chen, L. (2014). 'A method for discovering clusters of e-commerce interest patterns using click-stream data'. Electronic Commerce Research and Applications.

Tadeusz, M., Wojciechowski, M., & Zakrzewicz, M., (2000). 'Web users clustering', Proc. of the 15th International Symposium on Computer and Information Sciences.

Yongli, R., Tomko, M., Ong K., & Sanderson M. (2014) 'How People Use the Web in Large Indoor Spaces', Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management. ACM.