

Visual decisions in the analysis of customers online shopping behavior

Julija Pragarauskaitė, Gintautas Dzemyda

Institute of Mathematics and Informatics, Vilnius University
Akademijos str. 4, LT-08663 Vilnius
julija.pragarauskaite@gmail.com; gintautas.dzemyda@mii.vu.lt

Received: 5 November 2011 / **Revised:** 2 February 2012 / **Published online:** 5 July 2012

Abstract. The analysis of the online customer shopping behavior is an important task nowadays, which allows maximizing the efficiency of advertising campaigns and increasing the return of investment for advertisers. The analysis results of online customer shopping behavior are usually reviewed and understood by a non-technical person; therefore the results must be displayed in the easiest possible way. The online shopping data is multidimensional and consists of both numerical and categorical data. In this paper, an approach has been proposed for the visual analysis of the online shopping data and their relevance. It integrates several multidimensional data visualization methods of different nature. The results of the visual analysis of numerical data are combined with the categorical data values. Based on the visualization results, the decisions on the advertising campaign could be taken in order to increase the return of investment and attract more customers to buy in the online e-shop.

Keywords: analysis of online shopping behaviour, visual analysis, advertising campaigns, multidimensional data.

1 Introduction

The number of Internet users has increased exponentially around the world. Web applications have been growing explosively as well; people enjoy a wide variety of online services, from e-mail and browsing to information services and search, or collaborative services such as wikis, blogs, and social networks. Online purchasing of goods, both expensive and cheap, has been more and more preferable to a much larger extent in recent years due to convenience, speedy transactions and enhanced shopping experience. Customers use the Internet to buy the product online, to compare prices and products. Europeans spend an average of 24 hours on the Internet each month [1]. Almost all of the services that they use regularly – from the e-mail, instant messaging, maps to social networks, games, music, video sites to search and price comparison – are free, funded largely by online advertising.

All e-shop website customers could be classified into either coming directly to this website or through an advertising campaign. The online advertisements in various Internet channels significantly increase the number of customers of the e-commerce websites, especially if the advertisements announce special sales or discounts. According to “The 2010 Europe Digital Year in Review” [2] published by comScore Inc. (a company measuring the digital world performance and newest trends) even 97% of all Internet users were reached by display advertisements in Germany, UK and France. The increased level of engagement of Internet users was shown by rich media advertisements and especially by online video advertisements, where a regular EU Internet user spent approximately 15 hours in 2010 watching online videos on the Internet.

The Interactive Advertising Bureau (IAB) Europe reports that Internet advertising is approaching a 20% share of total advertising spending [1], almost doubling its share over the last two years. While the previous years saw a contraction in marketing budgets, the digital marketplace in Europe continues growing. With its growth there comes innovation – from advanced planning tools to sophisticated targeting techniques. Social media grew immensely in 2010 and saturated at least three quarters of individual European markets, whereas the display advertising market continued to grow in 2010, with the social networking category accounting for a rapidly increasing share of impressions. The online advertising market continues evolving, with more advanced advertisements, improved targeting capabilities, and higher quality creativity.

The analysis of online customers behavioral shopping data is very subjective, because such data usually is both numerical and categorical and the priorities of each feature describing the data are not clearly defined. Our idea is to analyse the data using a number of different visualization methods, to visualize customer behavioral data and to present the visualization results for the decisions on the strategy of the advertising campaign to a non-technical person, usually the media planner. Data on the online customer shopping behavior is multidimensional and consists of two types: numerical and categorical. Several multidimensional data visualization methods are applied in displaying online shopping data and their relevance is analyzed. The results of the visual analysis of numerical data are combined with the categorical data. The visualization would be very helpful for assessing and improving clustering and making decisions on the advertising campaign strategy, efficiency and usability of e-shop, because the decisions are carried out by a non-technical persons, so the analysis results should be displayed in the easy and understandable way. Correctly defined customer clusters can be used to identify up-selling and cross-selling opportunities with existing customers. One can also cluster products that tend to sell together. Clustering and visualization of transactional data is a vital component in the retail industry. The proper strategy of the advertising campaign helps advertisers to stay ahead in highly competitive markets.

2 Customer online shopping behaviour data

Each advertising campaign captures a huge number of various online data, e.g. referrals, impressions and clicks on advertisements in media websites, page views, products

viewed, products purchased, the price paid, etc. All this data is usually measured using browser cookies which are text files stored on the user's computer by the web browser. The online statistics show that 87% of users have accepted the 3rd party cookies [3] and this allows analyzing the performance of advertising campaigns and shopping behavior of different customer segments quite accurately. The customers can come to the e-shop via different channels (campaign traffic and non-campaign related traffic). In campaign traffic, the customers interacted with the advertising campaign material before they came to the advertiser website, e.g. they saw or clicked on the campaign banners or the adwords in the search engines, etc.; whereas in the non-campaign related traffic, the customers did not interact with the campaign material and went straight to the advertiser website. Note that for the non-campaign traffic, it is impossible to capture the same features as for the campaign traffic, e.g. campaign name, advertisement type, advertisement size, etc.

There is a number of different analysis methods performed recently with online data. The online data may be obtained from a clickstream. The clickstream can be defined as a path a customer takes through one or more websites and it can include within-site information such as the pages visited, time spent, etc. Researchers have investigated customer behaviors across websites [4–6] and within a particular website [7, 8]. Another study aimed at investigating the complexity of online shopping behavior [9] from many different online decision-making processes.

In this paper, the online shopping data has both numerical and categorical values. There were two advertising campaigns running for a particular advertiser in August 2010 (5361 purchases in total), with various creative material (flash and rich media) shown in the most visited media websites. Also search adwords were shown in search engines (Google, Yahoo, MSN) when customers were searching for specific queries related with the advertiser. The advertiser website was entirely tracked including the purchase amount. Each purchase in the advertiser online shop was defined by 9 different features x_1, x_2, \dots, x_9 . Five numerical features are detailed below:

- x_1 – campaign channel, which leads customers to the advertiser website: 1 – campaign traffic (customers who came through campaign material), 0 – non-campaign traffic (customers who came directly to the advertiser website);
- x_2 – the advertisement size in cm^2 ;
- x_3 – minutes from the last interaction: for campaign traffic – it shows the minutes from last interaction with campaign material to the purchase made in the advertiser website (the time from the last moment when the user interacted with the campaign material and the moment of purchasing; not necessary the purchase is performed in the same session as the interaction with campaign material); for non-campaign traffic – it shows the minutes since the customer opened any page in the advertiser website to the purchase without interacting with any campaign material;
- x_4 – the number of interactions before the purchase was made;
- x_5 – the sale amount of the purchase.

Four categorical features are detailed below:

- x_6 – the name of campaign, which leads customers to the advertiser website (non-campaign, Campaign 1, Campaign 2);
- x_7 – the interaction type with the advertising material in media site (non-campaign, impression, click, direct link);
- x_8 – the referrer, e.g. the media name where advertising material was shown for campaign traffic or the external website from where the customer went to the advertiser website (advertiser:frontpage, advertising.com, google.com:AdWords, e-travel.com, msn.com, tradedoubler.com, others);
- x_9 – the advertisement type (non-campaign, flash, search, link, rich media).

Due to large differences in the scales of each numerical feature, the values were normalized so that the values belong to the interval $[0, 100]$.

The numerical data analysis methods were applied to the matrix $\{x_{ji}, j = 1, \dots, m, i = 1, \dots, n\}$ consisting of $m = 2644$ rows and $n = 5$ columns, because this number of rows from 5361 possible ones has no empty cells.

3 Methods for analysis of online shopping behavior

The multidimensional data analysis is a very important task, which could be performed using classification, clustering, or visualizing data. The analysis results of customer online shopping are usually reviewed and understood by a non-technical person, therefore the results must be displayed in the easiest possible way – visualization of data. Based on the results, different strategies in the advertising campaign could be taken. In this paper, several direct visualization methods, where each feature is displayed in the visualization, and dimensionality reduction methods, which transform the initial data set from R^n to a lower-dimensional space R^d ($d < n$, usually $d = 2$ or 3), are combined in the integrated approach for the visual analysis of customer online shopping data.

The idea of visualization of the multidimensional data will be introduced below in brief. Let the purchase be described by an array of features x_1, x_2, \dots, x_n . Any feature may take some numerical values. A combination of values of all the features characterizes a particular purchase $X_j = (x_{j1}, x_{j2}, \dots, x_{jn})$ from the set $\{X_1, X_2, \dots, X_m\}$, where n is the number of features and m is the number of analysed purchases. X_1, X_2, \dots, X_m can be interpreted as the points in the n -dimensional space R^n . In fact, we have a table of numerical data for the analysis: $\{x_{ji}, j = 1, \dots, m, i = 1, \dots, n\}$.

A lot of methods have been developed for the direct data visualization. It is a graphical presentation of the data set providing the quality understanding of the information contents in a natural and direct way: parallel coordinates, scatter plots, survey plots, Chernoff faces, dimensional stacking, etc. (see [10–12]). There are a lot of so-called projection methods that can be used for reducing the dimensionality, and, particularly, for visualizing the n -dimensional points $X_1, X_2, \dots, X_m \in R^n$. A deep review of the methods is performed e.g. in [10, 13–15].

The goal of projection methods is to represent the multidimensional data points in a lower-dimensional space so that certain properties of the structure of the data set were preserved as faithfully as possible. Suppose that we have m data points, $X_j = (x_{j1}, x_{j2}, \dots, x_{jn})$, $j = 1, \dots, m$ in the R^n space and, respectively, the goal of dimensionality reduction is to define m points $Y_j = (y_{j1}, y_{j2}, \dots, y_{jd})$, $j = 1, \dots, m$ in the R^d space ($d < n$). The projection can be used to visualize the data set if a sufficiently small output dimensionality is chosen. One of these methods is the principal component analysis (PCA). The well-known principal component analysis [10, 16] can be used to display the data as a linear projection on a subspace of the original data space such that best preserves the variance in the data. Several approaches have been proposed for reproducing nonlinear multidimensional structures on a lower-dimensional display. The most common methods allocate a representation for each data point in a lower-dimensional space and try to optimize these representations so that the distances between them were as similar as possible to the original distances of the corresponding data items (points). The methods differ in that how the different distances are weighted and how the representations are optimized. Multidimensional scaling (MDS) refers to a group of methods that is widely used [10, 17]. The starting data of MDS is a matrix consisting of pairwise dissimilarities of the data items. In general, the dissimilarities need not be distances in the mathematically strict sense. There is a multitude of variants of MDS with slightly different cost functions and optimization algorithms. The MDS algorithms can be roughly divided into two basic types: metric and nonmetric MDS. The goal of projection in the metric MDS is to optimize the representations so that the distances between the items in the lower-dimensional space were as close to the original distances as possible. Denote the distance between the points X_i and X_j in R^n by d_{ij}^* , and the distance between the corresponding points Y_i and Y_j in R^d by d_{ij} . In our case, the initial dimensionality is n , and the resulting one is $d = 2$. The metric MDS approximates d_{ij}^* by d_{ij} . If a square-error cost is used, the objective function to be minimized can be written as

$$E_{\text{MDS}} = \sum_{\substack{i,j=1 \\ i < j}}^m w_{ij} (d_{ij}^* - d_{ij})^2,$$

where w_{ij} are some positive weights.

The following weights are frequently used:

$$w_{ij} = \left[\sum_{\substack{k,l=1 \\ k < l}}^m (d_{kl}^*)^2 \right]^{-1},$$

$$w_{ij} = \left[d_{ij}^* \sum_{\substack{k,l=1 \\ k < l}}^m (d_{kl}^*) \right]^{-1},$$

$$w_{ij} = \frac{1}{m d_{ij}^*}.$$

Most often, the Euclidean distances are used for d_{ij} and d_{ij}^* . A particular case of the metric MDS is Sammon's mapping [10, 18]. It tries to optimize a cost function called Sammon's stress that describes how well the pairwise distances in a data set are preserved:

$$E_S = \frac{1}{\sum_{\substack{k,l=1 \\ k < l}}^m d_{kl}^*} \sum_{\substack{i,j=1 \\ i < j}}^m \frac{(d_{ij}^* - d_{ij})^2}{d_{ij}^*}. \quad (1)$$

Sammon's stress function E_S (1) is more sensitive to small distances than to large ones, i.e. the Sammon's mapping gives more weight to small distances than to large ones.

Artificial neural networks also became as a mean for the dimensionality reduction and data visualization [19]. The self-organizing map (SOM) [13, 14, 20] is a class of neural networks that are trained in an unsupervised manner using competitive learning. It is a well-known method for mapping a multidimensional space onto a low-dimensional one. We consider here the mapping onto a two-dimensional grid of neurons. Let $X_1, X_2, \dots, X_m \in R^n$ be a set of n -dimensional points for mapping. Usually, the neurons are connected to each other via a rectangular or hexagonal topology. When the training of SOM is completed, the winning neurons are determined for all the multidimensional points. Usually, the grid of neurons is interpreted as a plane and the position of the multidimensional point is completely defined by the position of corresponding winning neuron on the grid.

In this paper, the integrated approach for the visual analysis of customer online shopping data covers:

- visual analysis by geometric method (scatter plots);
- visual analysis by the multidimensional scaling (Sammon's mapping);
- visual analysis by the artificial neural networks (self organizing map).

4 Visual analysis by scatter plots

Due to a large number of parameters and their values of the customer online shopping data, only geometric visualization of scatter plots were used, because the iconographic and hierarchical displays would not be informative for the non-technical person. Scatter plots visualize dependencies between the selected features and allow finding clusters, outliers, and trends within the data. The scatter plot matrix can be used in order to extend the scatter plot to higher dimensions, which is useful for looking at all possible two-way interactions or correlations between features. In Fig. 1, the numerical online shopping data is displayed in the scatter plot matrix using RapidMiner software [21] with colored points by the campaign traffic type (light – campaign, dark – non-campaign).

The advantage of scatter plots is that the visualized features in pairs can be easily interpreted, whereas the main drawback of the scatter plot matrix is that with an increase of the dimensionality, less screen space is available for each projection. In [22] the technique to lighten this problem by means of the use of color is presented, which was also used in Fig. 1.

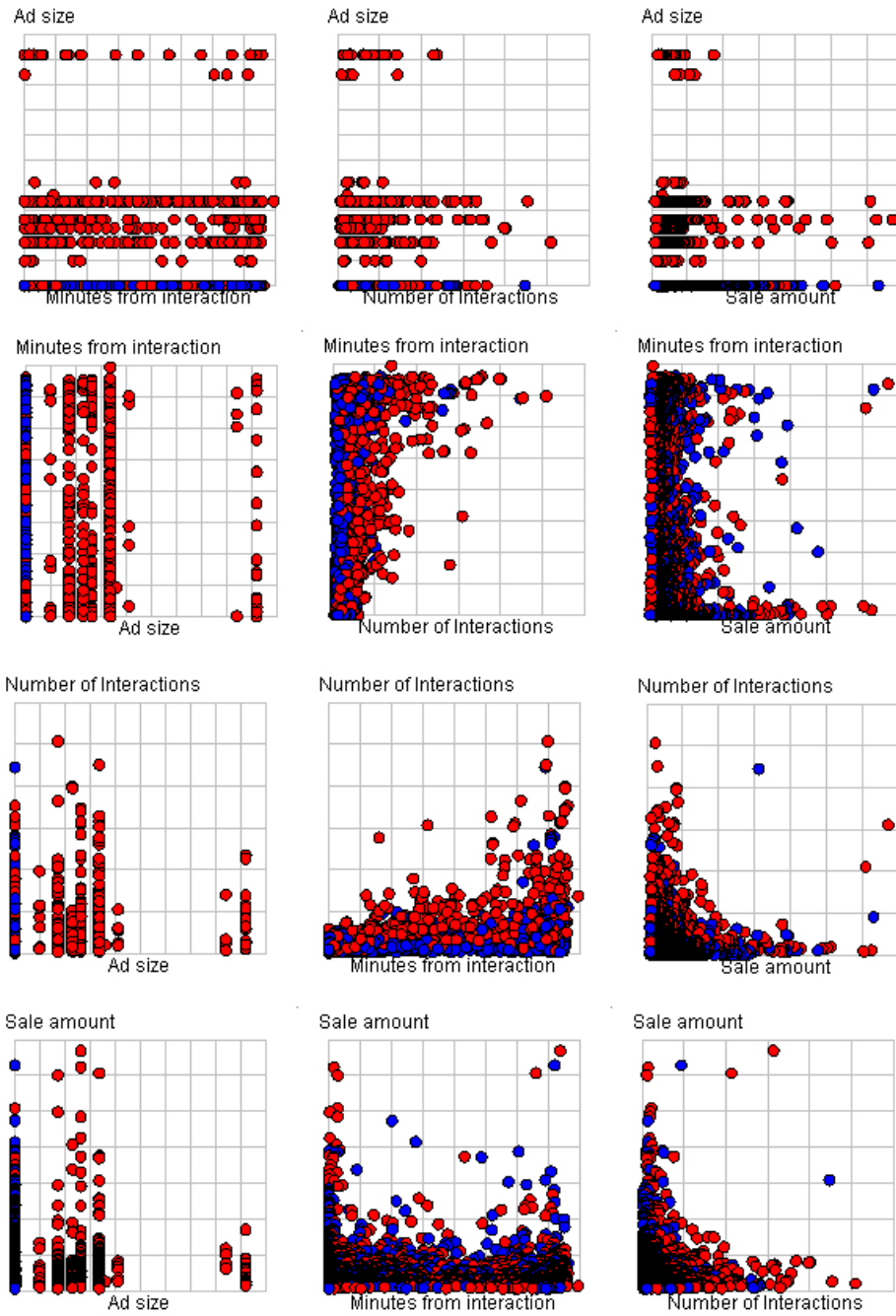


Fig. 1. The numerical online shopping data visualization using scatter plots.

Based on the scatter plot visualization, the following conclusions could be drawn:

- Most customers had just few interactions with the campaign advertisements before purchasing anything in the advertiser website.
- Most customers purchased goods in the advertiser website after quite a short period of time since the first interaction.
- The middle size advertisements and non-campaign interactions were the most popular ones in terms of return of investment.
- The customers who bought more expensive goods in the advertiser website belong to both channels: campaign and non-campaign traffic.
- The customers who bought more expensive goods in the advertiser website usually spent either a very short time period or a very long time period since the last interaction compared to the regular customer.
- The customers who bought more expensive goods in the advertiser website had very few interactions before purchasing anything or significantly smaller number of interactions with the campaign material compared to the standard customer.
- There are no separated clusters of customers.

5 Visual analysis by the multidimensional scaling

For the visualization using multidimensional scaling (MDS), only numerical data was used and it was mapped on the plain ($d = 2$). For the visualization of customer shopping behavior, Sammon's mapping was used. The visualization results are presented in Fig. 2. We do not give legends and units for both axes in Fig. 2 with the visualization results, because we are interested in observing the interlocation of the purchases $X_j = (x_{j1}, x_{j2}, \dots, x_{jn})$, $j = 1, \dots, m$ on a plane only.

The advantage of MDS visualization is that it shows the relationships between the features of data: multidimensional data points that are similar will appear close together while the points that are different will appear far away from one another. MDS is best used in situations where there is a large amount of data organized in table form. However, particularly for large data sets MDS is slow and time consuming task. Also, since MDS is a numerical optimization technique, it can fail to find the true best solution because it can become stuck on the local minima, solutions that are not the best solution but that are better than all nearby solutions. The increased computational speed now allows MDS ordinations even of large data sets and multiple features to be run, so that the chance of being stuck on a local minimum is significantly decreased.

The experiments were performed with various realizations of MDS in Orange 2.0 tool [23]. Only Sammon's mapping gave clearly separated clusters, see Fig. 2.

Based on the MDS visualization, useful conclusions could be drawn. It is easy to notice separate clusters, which represent different customer groups. If clusters are displayed close to each other in the map, then these clusters have more similarities. The cluster in the

bottom of Fig. 2 (colored in black) represents the customers, who came directly to the advertiser website and purchased goods without any interaction with the campaign material. The group of 5 different clusters in the top of Fig. 2 represents all the customers who came to the advertiser website through the campaign material. The clusters within this group are divided according to the advertisement size: the top cluster represents the customers who were attracted to the advertiser website via largest advertisements, whereas other clusters are ordered from larger advertisements at the top to smallest advertisements at the bottom. Having the results in Fig. 2, the manager can make a decision on the efficiency of the campaign – how much the size of advertisement influences purchasing.

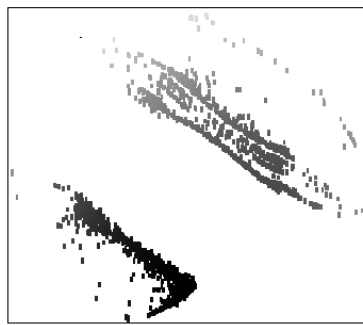


Fig. 2. The visualization of numerical online shopping data by Sammon's mapping.

6 Visual analysis by the self organizing map (SOM)

In this paper, the Viscovery SOMine software [24] was applied to visualize the advertising campaign data using the hexagonal SOM network topology and the Ward neighborhood function. The classical clustering method of Ward belongs to the hierarchical agglomerative clustering algorithms. Ward's algorithm uses the following distance between two clusters, A and B :

$$\begin{aligned} d(A, B) &= \sum_{i \in A \cup B} \|X_i - \vec{m}_{A \cup B}\|^2 - \sum_{i \in A} \|X_i - \vec{m}_A\|^2 - \sum_{i \in B} \|X_i - \vec{m}_B\|^2 \\ &= \frac{n_A n_B}{n_A + n_B} \|\vec{m}_A - \vec{m}_B\|^2, \end{aligned}$$

where \vec{m}_j is the center of cluster j , and n_j is the number of points in it. $d(A, B)$ is called the merging cost of combining the clusters A and B . With hierarchical clustering, the sum of squares starts at zero (because each point is in its own cluster) and then grows as we merge clusters. Ward's algorithm keeps this growth as small as possible.

The SOM network allows analyzing both numerical and categorical data in the easiest and most understandable way for a non-technical person. The disadvantage is that, in order to create a proper SOM visualization, a number of different additional parameters are needed and they should be defined by a person, who already has knowledge about the initial data and can prioritize the data in the proper way.

The results on application of the SOM to the online shopping data are discussed below.

Eight clusters of SOM neurons were discovered. The 1.5 priority was given to the feature ‘sale amount’, whereas all the rest features had 1.0 priority in the analysis. The information on each cluster is displayed in Table 1, where the most beneficial customer cluster is C7 and the customers, who belong to this cluster, bought the most expensive goods in the advertiser website. The customers of cluster C7 made around 5 interactions in average with the campaign material before purchasing and the time since the last interaction with the campaign material until the purchase, is around 20 minutes. The customers, who came directly to the advertiser website without interacting with any campaign material (cluster C1), also spent around 20 minutes in the website before purchasing and make 3 interactions in the advertiser website in average before purchasing. This shows that customers, belonging to both C1 and C7 clusters, have similar behaviour in terms of time spent and interactions made.

Table 1. Cluster information.

Cluster	C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8
Frequency	44.89%	16.15%	13.09%	13.05%	4.88%	1.89%	4.05%	2.00%
Campaign type	0	98.1	100	0	100	100	45.8	73.5
Advertisement size	0	27.3	0	0	35.3	7.9	7.4	58.8
Minutes from last interaction	20.6	36.9	30.1	7.5	43.5	38.2	19.5	32.9
Number of interactions	3.2	12.4	7.3	0.4	15.7	8.8	4.8	10.7
Sale amount	5.6	5.8	6.5	6.9	5.3	8.5	40.7	5.5
Campaign name:								
Non-campaign	100%	2%	0%	100%	0%	0%	54%	27%
Campaign 1	0%	98%	100%	0%	100%	100%	46%	8%
Campaign 2	0%	0%	0%	0%	0%	0%	0%	65%
Interaction type:								
Non-campaign	100%	1%	0%	0%	0%	0%	44%	0%
Impression	0%	95%	0%	0%	83%	28%	23%	55%
Click	0%	3%	100%	0%	17%	72%	22%	18%
Direct Link	0%	1%	0%	100%	0%	0%	10%	27%
Referrer:								
Advertiser:Frontpage	100%	1%	0%	0%	0%	0%	44%	0%
Advertising.com	0%	98%	0%	0%	0%	0%	24%	0%
Google.com:AdWords	0%	0%	100%	0%	3%	0%	21%	0%
E-travel.com	0%	1%	0%	100%	0%	0%	9%	0%
Msn.com	0%	0%	0%	0%	96%	0%	0%	0%
Tradedoubler.com	0%	0%	0%	0%	0%	100%	0%	0%
Others	0%	0%	0%	0%	1%	0%	2%	100%
Advertisement type:								
Non-campaign	100%	2%	0%	100%	0%	0%	54%	27%
Flash	0%	98%	0%	0%	97%	28%	25%	74%
Search	0%	0%	100%	0%	3%	0%	21%	0%
Link	0%	0%	0%	0%	0%	72%	0%	0%
Rich Media	0%	0%	0%	0%	0%	0%	0%	0%

The component planes representing numerical features on the SOM are given in Fig. 4. The customers that belong to the most beneficial cluster C7 spent the average time period before purchasing, however were not coming back to the advertiser more times than a regular customer. Approximately half of the customers belonging to the C7 cluster came to the advertiser website from the campaign and were interacting with the middle size campaign advertisements.

The component planes representing categorical features on the SOM are given in Figs. 5–8.

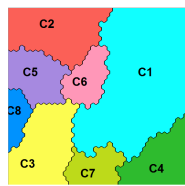


Fig. 3. Clusters in SOM obtained by analysing the numerical online shopping data.

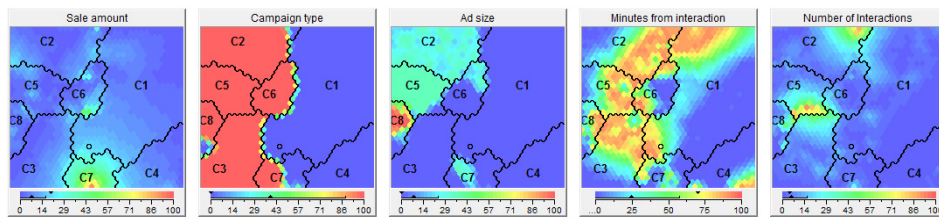


Fig. 4. Component planes representing numerical features on the SOM.

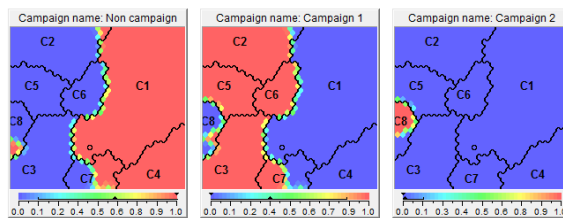


Fig. 5. Component planes representing campaign names on the SOM.

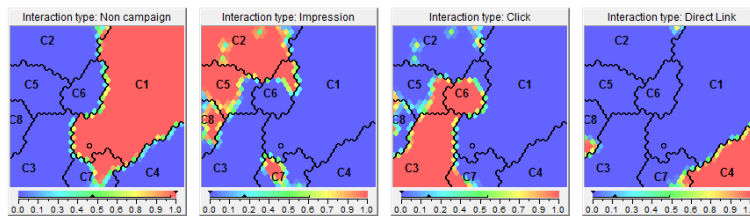


Fig. 6. Component planes representing interaction types on the SOM.

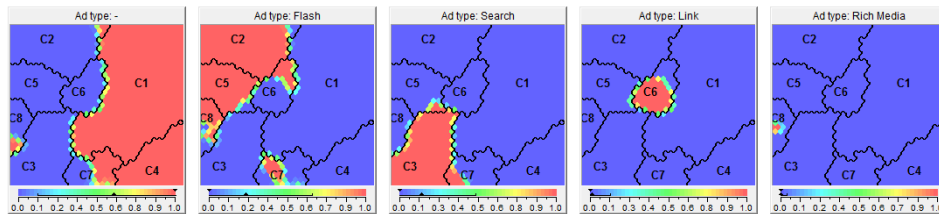


Fig. 7. Component planes representing advertisement types on the SOM.

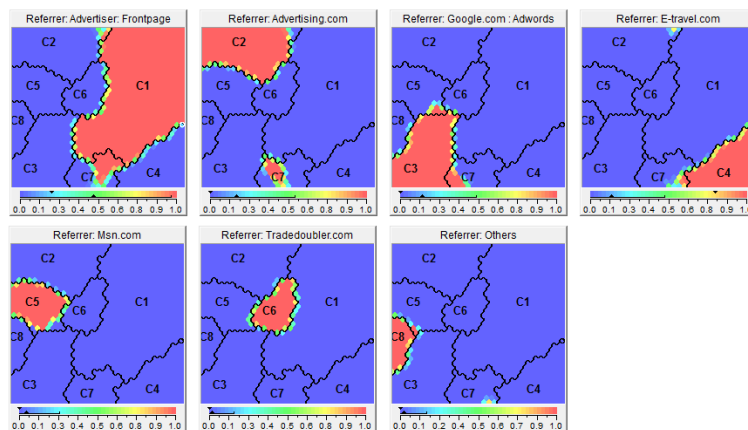


Fig. 8. Component planes representing the referrers on the SOM.

The following conclusions could be drawn using the SOM visualization:

- A half of the most beneficial cluster *C7* customers came to the advertiser website through Campaign 1 and the other half of *C7* customers did not interact with the campaign material.
- The *C7* customers who came through the advertising campaign, a half of them saw campaign material and the other half clicked on the campaign advertisements.
- The most popular advertisement type of *C7* cluster was flash or search in search engines, whereas the rich media was the least popular advertisement type and does not belong to cluster *C7*.
- The most beneficial customers from *C7* cluster were interacting mostly with advertisements on advertising.com and Google AdWords.

7 Conclusions

In this paper, the approach has been proposed for the visual analysis of the online shopping data and their relevance. It integrates several multidimensional data visualization methods

of a different nature. The results of the visual analysis of numerical data are combined with the categorical data values.

There is no single visualization that is best for multidimensional data exploration or data mining. Some visualization methods are better for showing clusters or outliers, e.g. self organizing maps or multidimensional scaling, other visualizations can show two way relations, e.g. scatter plots. However, direct visualization methods cannot properly visualize a large number of data records, since the plots become non-informative and difficult to understand. In this paper, the SOM visualization has showed the best results for the analysis and visualization of the online shopping behavior, however it requires some training and experience by a user, who will analyze the data. Based on the visualization, the decisions could be taken by a non-technical person who is managing the advertising campaigns. The visualization helps to determine the behavior of the interesting clusters and their interactions with the campaign material. Thus the decisions on the advertising campaign strategy could be made, e.g. which type of advertisement, on which media sites to put advertisements, etc. Moreover, the decisions on the advertising campaign could be made with a view to increase the return of investment and attract more customers to buy in the online e-shop.

References

1. iab: Consumers Driving the Digital Uptake: The Economic Value of Online Advertising-based Services for Consumers, 2010, http://www.iab.net/insights_research/947883/consumers_driving_digital_uptake.
2. comScore: The 2010 Europe Digital Year in Review, 2011, http://www.comscore.com/Press_Events/Presentations_Whitepapers/2011/2010_Europe_Digital_Year_in_Review.
3. GRC Visitor Browser & Cookie Demographics, <http://www.grc.com/cookies/stats.htm>.
4. R.E. Bucklin, J.M. Lattin, A. Ansari, S. Gupta, D. Bell, E. Coupey, J.D.C. Little, C. Mela, A. Montgomery, J. Steckel, Choice and the internet: From clickstream to research stream, *Marketing Letters*, **13**(3), pp. 245–258, 2002.
5. R.E. Bucklin, C. Sismeiro, A model of web site browsing behavior estimated on clickstream data, *Journal of Marketing Research*, **40**(3), pp. 249–267, 2003.
6. Y. Park, P.S. Fader, Modeling browsing behavior at multiple websites, *Marketing Science*, **23**(3), pp. 280–303, 2004.
7. A. Goldfarb, Analyzing website choice using clickstream data, in: M.R. Baye (Ed.) *The Economics of the Internet and E-Commerce*, Advances in Applied Microeconomics, Vol. 11, Elsevier, 2002, pp. 209–230.
8. W.W. Moe, Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream, *Journal of Consumer Psychology*, **13**, pp. 29–40, 2003.

9. S. Senecal, P.J. Kalczynski, J. Nantel, Consumers' decision-making process and their online shopping behavior: A clickstream analysis, *Journal of Business Research*, **58**, pp. 1599–1608, 2005.
10. G. Dzemyda, O. Kurasova, J. Žilinskas, *Multidimensional Data Visualization: Methods and Applications*, Springer Optimization and Its Applications, Vol. 75, Springer, 2012.
11. G. Grinstein, M. Trutschl, U. Cvek, High-dimensional visualizations, in: *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 26–29, 2001, San Francisco, California, USA*, 2001, pp. 1–14.
12. P.E. Hoffman, G.G. Grinstein, A survey of visualizations for high-dimensional data mining, in: U. Fayyad, G.G. Grinstein, A. Wierse (Eds.), *Information Visualization in Data Mining and Knowledge Discovery*, Morgan Kaufmann Publishers, San Francisco, CA, 2002, pp. 47–82.
13. S. Kaski, *Data Exploration Using Self-Organizing Maps*, Acta Polytechnica Scandinavica, Mathematics, Computing and Management in Engineering Series, Vol. 82, Espoo, 1997.
14. T. Kohonen, *Self-Organizing Maps*, 3rd edition, Springer Series in Information Sciences, Vol. 30, Springer, 2001.
15. O. Kurasova, *Visual Analysis of Multidimensional Data Using Self-Organizing Maps (SOM)*, Ph.D. dissertation, Institute of Mathematics and Informatics, Vilnius, 2005.
16. D.N. Lawley, A.E. Maxwell, *Factor Analysis as a Statistical Method*, 2nd edition, Butterworths, London, 1971.
17. I. Borg, P. Groenen, *Modern Multidimensional Scaling: Theory and Applications*, Springer, New York, 1997.
18. J.W. Sammon, A nonlinear mapping for data structure analysis, *IEEE Transactions on Computers*, **18**, pp. 401–409, 1969.
19. G. Dzemyda, O. Kurasova, V. Medvedev, Dimension reduction and data visualization using neural networks, in: I. Maglogiannis, K. Karpouzis, M. Wallace, J. Soldatos (Eds.), *Emerging Artificial Intelligence Applications in Computer Engineering – Real World AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*, Frontiers in Artificial Intelligence and Applications, Vol. 160, IOS Press, 2007, pp. 25–49.
20. G. Dzemyda, Visualization of a set of parameters characterized by their correlation matrix, *Computational Statistics and Data Analysis*, **36**(1), pp. 15–30, 2001.
21. Rapid Miner 5.0, <http://rapid-i.com/>.
22. W. Colin, J.C. Beatty, Using color dimensions to display data dimensions, *Human Factors*, **30**(2), pp. 27–142, 1988.
23. Orange 2.0, <http://orange.biolab.si/>.
24. Viscovery SOMine 5.2, <http://www.viscovery.net/somine/>.