Library User Behavior Analysis – Use in Economics and Management

P. Hájek, J. Stejskal
Institute of System Engineering and Informatics
Institute of Economic sciences
University of Pardubice
84, Studentská, 532 01 Pardubice
CZECH REPUBLIC
petr.hajek@upce.cz, Jan.stejskal@upce.cz

Abstract: - This paper develops a method of bibliomining, including the characteristics of the various stages of the process. Furthermore, the whole process is applied to research conducted in 2012 in the largest public library in the Czech Republic – the Municipal Library of Prague. The results are interpreted and a proposal for continuation of research is also included.

Key-Words: - bibliomining, cluster analysis, economics, library user behavior, management, public library.

1 Introduction

DATA mining has recently gained attention through the realization of a wide range of technical and non-technical issues. In management and marketing, these issues are primarily the forecasting of supply and demand, customer classification or identification of patterns of customer behavior. Data mining can be defined as the process of extracting meaningful and useful patterns and rules from large data sets or databases. Data mining methods mainly include regression, classification, clustering methods, and association rules. These methods are used to obtain knowledge in a form which is understandable to the user, and usually in the form of conditional rules.

Bibliomining is defined in the context of the provision of library services as the process of applying data mining techniques to extract patterns of behavior from library databases [17,16]. This process consists of several steps, namely, target identification, collection and preprocessing of data sets, discovery of knowledge in data, result evaluation, and application of the acquired knowledge in practice. The aim of this process is to obtain behavioral patterns of library users that enable more efficient management of customer relations and continued improvement in the quality of services offered by libraries. Bibliomining is becoming an important tool for supporting library management's decision making processes.

In the past, research in the area of bibliomining has only focused on defining the concept [16] or the technological background of the application, which involves the process of designing a suitable database [8,24]. Bibliomining applications have hitherto been focused on the optimal allocation of budgetary resources based on the past circulation of documents [11]. What is lacking is a specific application of bibliomining methods based on the recognition of patterns of user behavior. This work aims to bridge this gap and presents the application of cluster methods and methods of extracting attributes of real data of a public library in order to find similarities in the services provided by public libraries based on a representative set of behavior patterns of users. The results of this work can be implemented to achieve more efficient management of the public library.

The rest of the paper is structured as follows. Section 2 is devoted to the characteristics of library services. There is then a description of data collection and processing. Section 4 characterizes the use of bibliomining methods and the results are interpreted. The paper ends with a summary of the conclusions.

2 Bibliomining

A conceptual framework for bibliomining was defined by Nicholson [16]. It includes five basic elements: operation of the library, bibliographic

records, bibliometric data, library services and demographic structure of users.

Operation of the library is represented by the library staff, their patterns of behavior, time allocation, etc. Bibliographic records usually include the name of the item (document), its author, classification, abstract, keywords and data on its availability. Bibliometric data includes citations and cross-references (for more about bibliometrics see [6]). Library services can be divided into searching, circulation of documents (use), librarian assistance, educational and cultural programs, etc. Important information for the demographic structure of users includes place of residence, interests and membership in interest groups, employment and specific job position, education, age, etc.

Bibliomining has evolved from knowledge discovery in databases and data mining, respectively. Knowledge discovery in databases is defined as the (iterative and interactive) nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [5]. Data mining is closely related to this term and it is defined as the step in the process of knowledge discovery in databases, that inputs predominantly cleaned, transformed data, searches the data using algorithms, and outputs patterns and relationships to the interpretation / evaluation step of the whole knowledge discovery in databases process. The outline of these concepts is depicted in Fig. 1 and Fig. 2. Fig. 2 presents the tasks and applications of data mining. Based on the tasks of data mining we distinguish classification methods (such as decision trees [3], neural networks [9], etc.), methods for the generation of association rules (apriori algorithm, tertius-type algorithm, etc., see [1]) and clustering methods (K-means algorithm, hierarchical clustering methods, self-organizing maps, etc., see [10]).

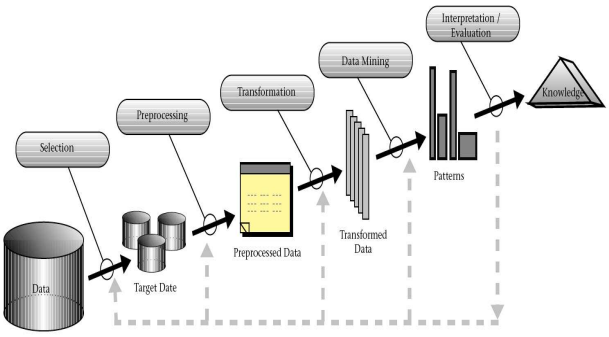

Fig. 1: Knowledge discovery in databases
Source: [5]

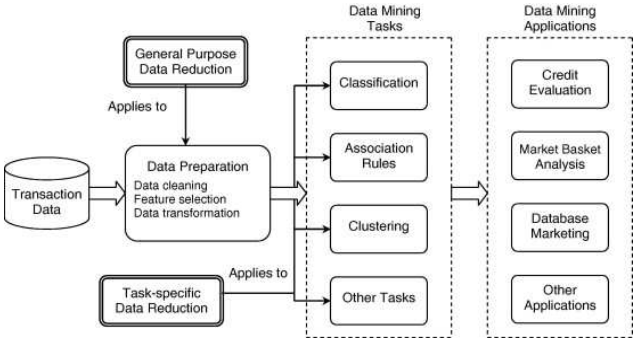

Fig. 2: Data mining tasks and applications
Source: [13]

Several recently proposed groups of methods belong to the area of bibliomining. For example web usage mining, where a dataset is obtained from the behavior of users on the Internet or their IP addresses and records in log files [20]. This can track user behavior, how users reach a website, where they are directed next, in what sequence, and how long they send browsing the individual items, etc. Web usage mining has been applied for a research library in the work of Bollen et al. [2]. In [19] log files were analyzed by clusters based on graph structures in order to find communities of common interest.

Text mining is a suitable method for the analysis of documents. Based on the textual content of documents it is possible to find a similar of set documents, extract keywords from documents, etc. Text mining is not focused directly on user behavior but on the analysis of text documents. The link to user behavior is achieved through the process of the user searching through a set of text documents. Text mining can therefore contribute to more effective searching and more accurate search results. Combining text mining methods with bibliometrics was performed in [12], where it was possible to remove unrelated links from the summary of document citations.

Bibliomining can be used to realize a wide range of issues associated with the behavior of library users. It can help libraries to reveal the future behavior of users by predicting the circulation of documents. These predictions can then be used to made decisions on the acquisition and allocation of funds. Furthermore, this knowledge can be used to make decisions about future between-borrowing cooperation. The knowledge gained by the circulation of documents can be further used to decide on the location of different types of items in the library, to change the opening hours and the

number of librarians, etc. On the other hand, these results can be used for more efficient marketing of the library and personalization of services, such as recommendations based on the analysis of the behavior of similar users [7].

3 Data Collection and Pre-processing

The survey was carried out in the pilot phase of the project Methodology for measuring the value of library services. It was a purely quantitative questionnaire survey conducted to verify the proposed method of determining empirical data among its respondents. The partial aim was to provide enough empirical data for verification of method of questioning the respondents on their individual preferences of the public services provided by the library, including questions on the valuation of their importance or value.

The basic set of respondents was made up of readers at the Municipal Library of Prague (MLP). The total number of members of the questionnaire panel was 1 061, a total of 620 responses were processed. Among the panel members were randomly selected MKP readers who meet the following criteria:

- age 15 + (1994 and older)
- loaned at least one item in 2010,
- gave an email address,
- responded in the first round of questioning, thus agreed to be included in the panel.

The panel members were surveyed using an online questionnaire (CAWI type) during October and November 2011. In the pilot survey the panel of respondents was divided into 7 groups, who provided with different variants of the same questions (different concepts of the questions or different questions focusing on the same area of preference, etc.). The method of questioning comes from foreign studies [18-14], which use WTA and WTP procedures to determine the respondents' opinions on the value of the services (for an explanation see [18]). Part of the questions was designed separately by the authors of this paper.

The return rate of the questionnaires was 35% on average. For more detailed information see Table 1. Input attributes were related to the basic socio-economic and demographic characteristics (gender, age, economic activity and education) of the reader, frequency of visits and loans and finally customer behavior, i.e. whether the individual uses the individual services offered by the library.

Table 1: Return rate of 7 versions of the questionnaire

Name of questionnaire	Answers	Sent	Return rate
Ask the readers 1/2011	48	150	32 %
Ask the readers 2/2011	50	150	33 %
Ask the readers 3/2011	61	152	40 %
Ask the readers 4/2011	49	152	32 %
Ask the readers 5/2011	60	153	39 %
Ask the readers 6/2011	62	151	41 %
Ask the readers 7/2011	44	153	29 %
Total	374	1061	35%

An overview of the attributes relating to the readers' user behavior is shown in Table 2. This table shows an overview of library services and the frequency of their use by the reference sample of readers. More than 88% of readers use the service loan of books and other media without the assistance of a librarian. This assistance, however, is utilized by 52% of users. Other services are used by less than half of the users. Approximately 43% of users spend time in the library reading documents retrieved without the assistance of a librarian. Less than 10% of users utilize the service of copying and printing of documents. A similar situation is observed in the case of reading and downloading of electronic documents. Half of the readers use the library's website to search for documents in the catalogue. About 38% need the services of a librarian to search. None of the users utilized the possibility of processing research documents and only 2% used the services "Ask a Librarian". Assistance of a librarian is mainly used for the purpose of obtaining information on the availability of books in another branch or library (32%) and instruction on library services (20%). Approximately 17% of users participated in an educational or cultural program in the library, and 7% outside the library. About 41% of users perceive their stay in the library as a form of relaxation and 12% as a place to meet friends. In order to access the Internet, users utilize the library's computers (16%) in approximately the same extent as they use their own computers through Wi-Fi (17%).

The socio-economic and demographic characteristics of the users are included in Table 3. About 80% of the users were women. The sample of users was dominated by persons aged 40 years and over (17% aged 40-49 and 30% aged 50 and over). Regarding education, about 37% had graduated from secondary school and 47% had a university degree. Most users were economically active (55%). In addition, students are represented (21%) and pensioners (16%), housewives (7%) and the unemployed (less than 1%).

Table 2: Reader attributes

Abbr.	Characteristic	Frequency of usage (yes/no)
x_1	Loan of books, magazines, CD retrieved without the help of a librarian	547/73
x_2	Loan of books, magazines, CDs from stock, retrieved with the help of a librarian	320/300
x_3	Reading books, newspapers or magazines, etc. studying documents etc. in the library - retrieved without the help of a librarian	269/351
x_4	Reading of books, newspapers or magazines, etc. study of documents etc. in the library – retrieved with the help of a librarian	104/516
x_5	Copying of documents (without the assistance of a librarian)	60/560
x_6	Copying of documents (with the assistance of a librarian)	27/593
x_7	Printing of documents (without the assistance of a librarian)	14/606
x_8	Printing of documents (with the help of a librarian)	22/598
x_9	Use of electronic databases	25/595
x_{10}	Reading e-books	5/615
x_{11}	Downloading e-books from the website	59/561
x_{12}	Downloading other electronic documents from the website	20/600
x_{13}	Search in the paper catalogue	21/599
x_{14}	Search in the electronic catalogue on a computer	285/335
x_{15}	Search in the electronic catalogue on the website	308/312
x_{16}	Search with assistance of a librarian	238/382
x_{17}	Processing of written research	0/620
x_{18}	Information on the availability of a book at another branch or library	201/419
x_{19}	Asking a librarian for instruction on library services	124/496
x_{20}	Use of the service “Ask a librarian”	12/608
x_{21}	Participation in an educational or cultural program in the library	107/513
x_{22}	Participation in an educational or cultural program outside the library	44/576
x_{23}	Spending time in the library for personal relaxation	254/366
x_{24}	Spending time in the library for studying your own documents	93/527
x_{25}	Meeting with friends	72/548
x_{26}	Use of electrical plug sockets for your own computer	56/564
x_{27}	Use of library computers to access the Internet	102/518
x_{28}	Use of Wi-Fi connection in the library	104/516
x_{29}	Use of technical equipment of the library	55/565
x_{30}	Use of library for Opencard related services	104/516

Table 3: Socio-economic and demographic characteristics of the users

Abbrev.	Characteristic	Frequency
x_{31}	Gender	Male: 126
		Female: 494
x_{32}	Age	15-17: 17
		18-19: 28
		20-24: 82
		25-29: 73
		30-34: 56
		35-39: 72
		40-49: 104
		50 and over: 188
x_{33}	Education	Basic or apprenticeship: 46
		Secondary: 230
		Post-secondary: 51
		University: 293
x_{34}	Socio-economic status	Student: 133
		Housewife, maternity leave/parental leave: 42
		Pensioner: 99
		Unemployed: 2
		Economically active: 344

The frequency of visits by users and their loan activity in the library are presented in Table 4. District branches of the library are used more often. Approximately 60% of readers use only or mostly district branches of the library. In contrast, less than 7% of readers only use the central library. The frequency of visits is dominated by the readers who visit the library about once a month (68%). Most users loan 3-6 books per month (58%).

Table 4: Frequency of visits of the users

Abbr.	Characteristic	Frequency
x_{35}	Library branch	Only district: 155 Mostly district: 219 Both the same: 115 Mostly central: 88 Only central: 43
x_{36}	Visit frequency	Less than once per month: 96 About once per month: 422 One per week and more: 102
x_{37}	No. of loaned items per month	0: 36 1-2: 131 3-4: 193 5-6: 166 7-10: 66 11 and more: 28

4 Bibliomining Methods Used

The K-means algorithm was used for the identification of services of the public library. This method belongs to the group of unsupervised learning methods as well as clustering methods. Clustering methods are used in data mining to identify certain groups of objects with similar characteristics.

The K-means method is a non-hierarchical clustering method. These methods are preferable for discovering knowledge in databases because they allow more efficient processing and interpretation of large data sets. In contrast, hierarchical clustering methods are used to create a tree structure of clusters, called a dendrogram. Clustering results are affected by outlying values and unwanted previous combinations of samples remain in the analysis.

In non-hierarchical algorithms (such as K-mode or K-means algorithms) the samples are a predetermined number of clusters. In the case of K-means algorithms initial cluster centers are set first and then the samples, located within a given distance from the center of the cluster, are assigned to the cluster. The key step here is the initial setup of the centers of the clusters (here carried out using a hierarchical single linkage algorithm), which enables the efficient functioning of the K-means algorithm and reduces the possibility of the error (utility) function being stuck at the local minimum. The aim of K-means algorithms is to minimize the utility function J, which has the form:

$$J = \sum_{r=1}^{m} \sum_{i=1}^{n} \left\| \mathbf{p}_i^r - \mathbf{c}_r \right\|^2 \quad (1)$$

where $\mathbf{p}_i^r$ is the sample i belonging to the cluster r and $\mathbf{c}_r$ is the center of the cluster r.

For clustering a K-means algorithm is used for which a fixed number of clusters m=11 is chosen. This number was determined based on the shape of utility function J for a varying number of clusters m. In the area behind m=11 there was no further significant decrease in the utility function J.

The bibliomining process for the analysis of user behavior in a public library is presented in Fig. 1.

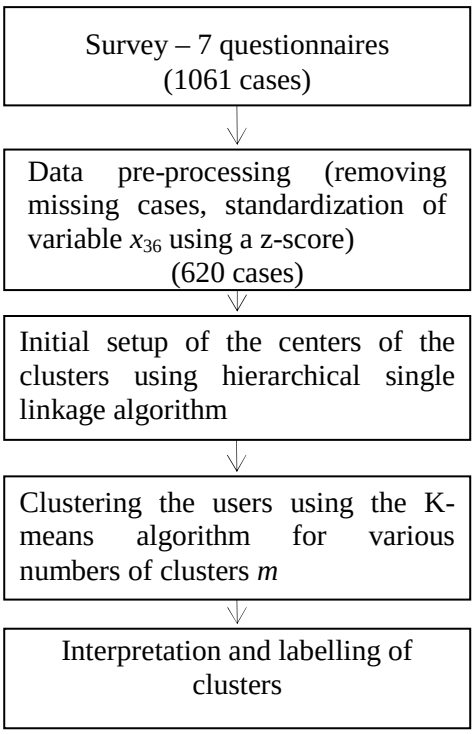

Fig. 3 Bibliomining process for the analysis of user behavior in a public library

5 Results and their Analysis

The results of clustering are presented in Tables 5 to 8. The following are typical cases (center of clusters), where cluster c_1 corresponds to 105 readers in the test sample, cluster c_2 corresponds to 91 readers, etc., see Table 3. The largest number of visitors is represented by cluster c_5, namely 20% of the total number of 620 in the test sample (see Table

7). Table 4 shows the user behavior of typical readers in a particular cluster. This typical reader is represented by the center of the cluster. The first column shows the values with the highest frequency for the entire dataset in order to identify differences between typical readers of individual clusters from that of a single typical reader for the entire data set. A typical reader for the entire data set uses both loan services (home lending) without the assistance of a librarian with searching and with this assistance. Furthermore, she is a woman older than 50 years, economically active with a university education, using mostly a district library, visiting the library once a month with the monthly number of loans in the range of 3 to 4 items. In this way it is also possible to characterize the individual readers in the individual clusters.

Readers in clusters c_1 and c_5 are the least demanding on the library services. These users only use distance loans of book, magazines or CD retrieved either with or without the assistance of a librarian. They are economically active women with a secondary school education between 40 and 49 years (cluster c_1) or university education and age over 50 years (cluster c_5). They only or mostly use the services of a district branch of the library, and once a month to borrow 3-6 books. Cluster c_1 can be characterized as "*borrowers requiring assistance in searching*" and a cluster c_5 as "*independent borrower*s."

Typical readers in other clusters also use other services offered by the library. Cluster c_2 can be characterized as "*users searching for a specific title.*" They use both the electronic catalogue and the assistance of a librarian. They are economically active women over 50 years old with a university education who visit the library once a month and use both the district and central libraries to the same degree to borrow 1-2 books. A typical reader in cluster c_3 searches for items without the assistance of a librarian and reads in the library for personal relaxation. He or she is retired and uses the library once a month to borrow 3-4 books. They are therefore people who do not come with a specific idea of what to borrow i.e. "*users looking for rest and relaxation*".

Cluster c_4 is another group of readers who can be described as "*users demanding on the assistance of a librarian*". They are economically active women aged 40-49 years who mostly visit a district library once a month borrowing 5 to 6 items.

Clusters c_6 and c_7 include students who differ in their use of services based on their hitherto level of education. Secondary school students are in cluster c_6, while in cluster c_7 there are university students. Students in cluster c_6 search in the catalogue and on the Internet looking for specific information related to Opencard (city card) services. They mostly use district libraries once a month with a lower frequency of borrowing i.e. 1 to 2 items. They can be described as "*students looking for specific information*". In contrast, students in cluster c_7 use the library for borrowing as well as in-house reading of documents that are retrieved both with and without the assistance of a librarian. They also search in electronic catalogues with and without the assistance of a librarian. They also ask about the availability of books outside of the library and instruction on library services. They use the library both for personal relaxation and to study their own documents. Furthermore, they use the library computers or their own computers connected to the Internet via Wi-Fi for their studies and other activities. From the point of view of the library they are the most demanding customers and they can be characterized as "*customers demanding on modern technology*". Users in cluster c_9 also use the library for their studies. This is a less demanding user who mainly uses the central branch of the library borrowing to 5-6 items per month. They use only printed documents for their studies, so they can be described as "*studiers of printed documents*".

Users in cluster c_8 are university educated, aged 30-34 years, who search for books and other documents only with the assistance of a librarian, both for distance and for in-house reading. This corresponds to the low frequency of borrowed items. This cluster can be described as "*users needing the assistance of a librarian*."

User behavior in clusters c_{10} and c_{11} is similar; the difference lies only in the use of a librarian to search for the required books. While readers in cluster c_{10} (women with a university education, 30-34 years) make use of this service, 40 to 49 year old men with a secondary school education in cluster c_{11} do not. These clusters also differ in the frequency of use of library services. While the typical user in cluster c_{10} uses the library once a month and has a higher number of loans (5-6), a user in cluster c_{11} borrows fewer books (3-4), but the frequency of visits is the highest of all the clusters (more than once per week). These clusters can be described as "*young employed reader*s" (cluster c_{10}) and "*readers with a high frequency of visits*" (cluster c_{11}).

Table 5: Typical readers in the individual clusters

Service	Total	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}
x_1	yes	yes	yes	yes	yes	yes	yes	yes	no	yes	yes	yes
x_2	yes	yes	yes	no	yes	no	no	yes	yes	no	yes	no
x_3	no	no	no	yes	no	no	no	yes	no	yes	no	no
x_4	no	no	no	no	no	no	no	yes	yes	no	no	no
x_5	no	no	no	no	no	no	no	no	no	no	no	no
x_6	no	no	no	no	no	no	no	no	no	no	no	no
x_7	no	no	no	no	no	no	no	no	no	no	no	no
x_8	no	no	no	no	no	no	no	no	no	no	no	no
x_9	no	no	no	no	no	no	no	no	no	no	no	no
x_{10}	no	no	no	no	no	no	no	no	no	no	no	no
x_{11}	no	no	no	no	no	no	no	no	no	no	no	no
x_{12}	no	no	no	no	no	no	no	no	no	no	no	no
x_{13}	no	no	no	no	no	no	no	no	no	no	no	no
x_{14}	no	no	yes	no	no	no	no	yes	no	yes	yes	yes
x_{15}	no	no	yes	no	no	no	yes	yes	yes	yes	yes	yes
x_{16}	no	no	yes	yes	yes	no	no	yes	no	no	no	no
x_{17}	no	no	no	no	no	no	no	no	no	no	no	no
x_{18}	no	no	no	no	yes	no	yes	yes	no	no	no	no
x_{19}	no	no	no	no	yes	no	no	yes	no	no	no	no
x_{20}	no	no	no	no	no	no	no	no	no	no	no	no
x_{21}	no	no	no	no	no	no	no	no	no	no	no	no
x_{22}	no	no	no	no	no	no	no	no	no	no	no	no
x_{23}	no	no	no	yes	no	no	no	yes	no	no	no	no
x_{24}	no	no	no	no	no	no	no	yes	no	yes	no	no
x_{25}	no	no	no	no	no	no	no	no	no	no	no	no
x_{26}	no	no	no	no	no	no	no	yes	no	no	no	no
x_{27}	no	no	no	no	no	no	no	yes	no	no	no	no
x_{28}	no	no	no	no	no	no	no	yes	no	no	no	no
x_{29}	no	no	no	no	no	no	no	no	no	no	no	no
x_{30}	no	no	no	no	no	no	yes	no	no	no	no	no

Table 6: Socio-economic and demographic characteristics of the typical readers

Service	Total	c_1	c_2	c_3	c_4	c_5
x_{31}	female	female	female	female	female	female
x_{32}	50+	40-49	50+	50+	40-49	50+
x_{33}	Uni	SS	Uni	SS	SS	Uni
x_{34}	Eco	Eco	Eco	Pen	Eco	Eco
Service	c_6	c_7	c_8	c_9	c_{10}	c_{11}
x_{31}	female	female	female	female	female	male
x_{32}	20-24	20-24	30-34	20-24	30-34	40-49
x_{33}	SS	Uni	Uni	SS	Uni	SS
x_{34}	Stu	Stu	Eco	Stu	Eco	Eco

Legend: SS is secondary school education with graduation exam, Uni is university education, Eco is economically active, Pen is pensioner and Stu is student.

Table 7: Characteristics of the typical readers based on the frequency of visits to the library

Service	Total	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}
x_{35}	MB	MB	BC	MB	MB	B	MB	BC	MB	MC	MB	MC
x_{36}	Mon	Mon	Mon	Mon	Mon	Mon	Mon	Mon	Mon	Mon	Mon	Wk
x_{37}	3-4	3-4	1-2	3-4	5-6	5-6	1-2	3-4	1-2	5-6	5-6	3-4

Legend: MB is mostly branch library, BC is both branch and central library the same, B is branch library only, MC is mostly central library, Mon is once per month and Wk is once per week.

In terms of frequency, Table 8 shows clear that the highest frequency (37% of users in total) is characterized by less demanding readers in clusters c_1 and c_5 i.e. "*borrowers requiring assistance in*

searching" and "*independent borrowers*". The following are the clusters c_2 and c_3 i.e. "*users searching for a specific title*" and "*users looking for rest and relaxation*". These clusters are made up of users who are less demanding on additional library services. Other clusters contain less than 10% of users. On the other hand, these clusters use specific library services, such as increased assistance of a librarian (cluster c_4) or technology (cluster c_7). We can therefore say that in addition to the basic services – loans at home, other services are used at high levels only by selected groups of users. In the clusters there are groups that use a large amount of services but the number of users is relatively low.

Table 8: Frequency of readers in clusters

Cluster	No. of readers (absolute)	No. of readers (relative in [%])
c_1	105	17
c_2	91	15
c_3	68	11
c_4	33	5
c_5	126	20
c_6	23	4
c_7	48	8
c_8	12	2
c_9	47	8
c_{10}	51	8
c_{11}	16	3

6 Conclusion

Public institutions currently provide public services based on a commission or public order from the central authorities of the state administration. These are mainly decided based on the amount of public funds allocated. Another consideration can be the fulfillment of the electoral program or vested interests. However, neither of these latter aspects is applicable in the case of public libraries. These public services depend primarily on education and further contribute to the elimination of information asymmetry. However, there is no reliable method to determine the appropriate amount of public resources to be allocated and to ensure the effective level of public services.

This paper aims to contribute to libraries determining the extent of public services by determining the typical consumer and not based on financial allocation or evaluation of public services using ROI [15,23]. If it is possible to determine the typical consumer of library services, then it will be possible for the library management to ensure the provision of the appropriate level of public services. This can be used to determine the necessary level of financial resources needed to ensure the standard of public service of a certain kind. This will help donors to better target funds or control their use.

The empirical data and bibliomining methods used in this paper determined the typical features of user groups, including their characteristics. The most frequent clusters in this case were "*borrowers requiring assistance in searching*" and "*independent borrowers*". In terms of frequency then followed "*users searching for a specific title*" and "*users looking for rest and relaxation*". These four user groups included about 63% of all users in the studied sample of respondents.

The sample results can be used in practice to change the attitudes of library management. Specifically, it is possible to assume from the result the following:

- who the typical reader is, i.e. the customer of library services, from which it is possible to deduce what services, in what quantity, quality and time will be required,
- it is possible to use an inverse approach. It will be possible to determine who is not a typical customer and if there is a public interest the library management will change the marketing of the library and try to attract the customer,
- it is possible to determine the customer's anticipation of the public services, thus helping to determine the standard of public services,
- this allows an analysis of the services provided or a specific set of public services (an analysis of services has never been performed in public libraries - what it should involve, there is no taxonomy, etc.)
- the results will affect the marketing of public institutions. It will no longer be possible to blame the immateriality of the services and hence the elusiveness of the characteristics or features of the marketing mix
- the results also have an economic significance. The employer can analyze manpower requirements in public services. Furthermore, it provides an analysis of the facilities background and its components, such as buildings and equipment. It has been clearly established that some services can be provided without the assistance of a librarian.
- for the libraries an analysis of the book collections can be performed and their range can be reviewed depending on popularity/need or usefulness to the user.
- development of user needs can be monitored and analyzed over the long-term. Based on the results the library can update the range of public services offered.

The method has many obstacles to its use. One of the fundamental obstacles is the absence of an analysis of the services public libraries provide. The range is too broad and varies from library to library. The results of an analysis of library services would enable the creation of standards in this sector of public services.

The use of the above procedures in practice prevents the need for a large amount of empirical information, the need for its explanatory capability and representativeness. On the other hand, it represents a unique approach to a long-term unresolved issue and provides proof that it is possible to head in this direction. Collecting data on the behavior of users of public libraries in the future will be automated e.g. by the implementation of radio frequency identification [22]. It would also be fruitful to explore the dependence of the use of the services on the information literacy of the readers [21].

Acknowledgements:
This paper was prepared as part of the research project No. 3052/2012 SOKL entitled "Methodology for measuring the value of library services" funded by the Ministry of Culture of the Czech Republic.

References:

[1] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, I. Verkamo, Fast Discovery of Association Rules. In *Advances in Knowledge Discovery and Data Mining*, U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy (Eds.), Menlo Park, California: AAAI Press, 1996, pp.307-328.

[2] J. Bollen, R. Luce, S. Vemulapalli, W. Xu, Usage Analysis for the Identification of Research Trends in Digital Libraries. *D-Lib Magazine*, Vol.9, No.5, 2003, available online at: http://www.dlib.org/dlib/may03/bollen/05bollen.html.

[3] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, *Classification and Regression Trees*, Belmont, California, Wadsworth, 1984.

[4] British Library, *Measuring our Value*, British Library Annual Report 2003/2004, 2004, available online at: http://www.bl.uk/pdf/measuring.pdf.

[5] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, From Data Mining to Knowledge Discovery: An Overview. Advances in Knowledge Discovery and Data Mining, *AI Magazine*, Vol.17, No.3, 1996, pp.37-54.

[6] E. Garfield, A.I. Pudovkin, V.S. Istomin, Mapping the Output of Topical Searches in the Web of Knowledge and the Case of Watson-Crick, *Information Technology and Libraries*, Vol.22, No.4, 2003, pp.183-187.

[7] A. Geyer-Schulz, A. Neumann, A. Thede, An Architecture for Behavior-Based Library Recommender Systems, *Information Technology and Libraries*, Vol.22, No.4, 2003, pp.165-174.

[8] K. Guenther, Applying Data Mining Principles to Library Data Collection, *Computers in Libraries*, Vol.20, No.4, 2000, pp.60-63.

[9] P. Hajek, Municipal Credit Rating Modelling by Neural Networks, *Decision Support Systems*, 2011, Vol.51, No.1, pp.108-118.

[10] P. Hajek, V. Olej, Municipal Creditworthiness Modelling by Clustering Methods. In *10th International Conference on Engineering Applications of Neural Network*, EANN07, Thessaloniki, Greece, 2007, pp.168-177.

[11] S.C. Kao, H.C. Hang, C.H. Lin, Decision Support for the Academic Library Acquisition Budget Allocation via Circulation Database Mining, *Information Processing and Management: an International Journal*, Vol.39, No.1, 2003, pp.133-147.

[12] R. Kostoff, J. del Rio, J. Humenik, E. Garcia, A. Ramirez, Citation Mining: Integrating Text Mining and Bibliometrics for Research User Profiling, *Journal of the American Society for Information Science and Technology*, Vol.52, No.13, 2001, pp.1148-1156.

[13] X.B. Li, V.S. Jacob, Adaptive Data Reduction for Large-scale Transaction Data, *European Journal of Operational Research*, Vol.188, No.3, 2008, pp.910-924.

[14] T. Lynch, *A Study of Taxpayer Return on Investment (ROI) in Florida Public Libraries: Part III - REMI Details*. Tallahassee, FL: State Library and Archives of Florida, 2004.

[15] J.R. Matthews, *Measuring for Results – The Dimensions of Public Library Effectiveness*, Libraries Unlimited, London, 2004.

[16] S. Nicholson, The Basis for Bibliomining: Frameworks for Bringing together Usage-based Data Mining and Bibliometrics through Data Warehousing in Digital Library Services, *Information Processing and Management*, Vol.42, 2006, pp.785-804.

[17] S. Nicholson, The Bibliomining Process: Data Warehousing and Data Mining for Library

Decision-Making, *Information Technology and Libraries*, Vol.22, No.4, 2003, pp.146-151.

[18] Ohio, *Value for Money: Southwestern Ohio's Return from Investment in Public Library*, Driscoll & Fleeter, Columbus, OH, 2006.

[19] C. Papatheodorou, S. Kapidakis, M. Sfakakis, A. Vassiliou, Mining User Communities in Digital Libraries, *Information Technology and Libraries*, Vol.22, No.4, 2003, pp.152-157.

[20] D. Pierrakos, G. Paliouras, C. Papatheodorou, C.D. Spyropoulos, Web Usage Mining as a Tool for Personalization: A Survey, *User Modeling and User-Adapted Interaction Journal*, Vol.13, No.4, 2003, pp.311-372.

[21] A. Repanovici, Information Technology Implication in Student Behaviour for Information Literacy Skills. In *Proceedings of the 4th WSEAS/IASME International Conference on Educational Technologies* (EDUTE'08), Corfu, Greece, October 26-28, 2008, pp.81-86.

[22] A. Repanovici, M. Turcanu, L. Cristea, M. Baritz, I. Moisil, Smart Library: RFID Implementation in Libraries. In *Proceedings of the 8th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering & Data Bases* (AIKED '09), 2009, pp.523-526.

[23] J. Stejskal, P. Kotatkova Stranska, K. Matatkova, P. Hajek, Public Services Value Determining – Case of Public Libraries. In *Proceedings from 4th WSEAS World Multiconference on Applied Economics, Business and Development* (AEBD 12), Porto, 2012, pp.140-145.

[24] J. Stejskal, K. Matatkova, Descriptive analysis of the regional innovation system - novel method for public administration authorities. In *Transylvanian Review of Administrative Sciences* (39), pp. 91-107.