Acta Universitatis Sapientiae
Electrical and Mechanical Engineering, 5 (2013) 21-43

New Approaches Analyzing Usage of the Telecommunication Services

László MOLNÁR[1], Szilárd ANDRÁS[2],
Ibolya KOTTA[3], Iuliu SZÉKELY[4]

[1]Department of Electrical Engineering, Faculty of Technical and Human Sciences,
Sapientia Hungarian University of Transylvania, Tg. Mureş,
e-mail: laszlo.molnar@ms.sapientia.ro
[2]Faculty of Mathematics and Computer Science, Babeş-Bolyai University, Cluj,
e-mail: andraszk@yahoo.com
[3]Faculty of Psychology and Education Sciences, Babeş-Bolyai University, Cluj,
e-mail: kottabambi@yahoo.com
[4]Department of Electrical Engineering, Faculty of Technical and Human Sciences,
Sapientia Hungarian University of Transylvania, Tg. Mureş,
e-mail: gszekely@ms.sapientia.ro

Manuscript received September 04, 2014; revised December 18, 2014.

Abstract: In this paper we analyze the media usage of students and the relation between the media usage and known psychological factors. Using hierarchical clustering, we identified three main categories: the socializing, the pragmatic and the all-consuming categories. Our results show that the media usage is also highly age and gender dependent even inside a very tight age group (19-22 years old students). Using PCA, discriminant analysis, and a special transformation that highlights the consentaneous reactions to different questions we also obtained some positive results that can be used to create an artificial intelligent agent for a posteriori clustering and some negative results related to the relation between most of the studied psychological factors and the media usage.

Keywords: Media usage, psychological factors, PCA, discriminant analysis, simultaneous bursts.

1. Introduction

The present telecommunication activities are converging in at least three ways: telecommunication and computer networks convergence; converging towards the Next Generation Network (NGN) architectures (all IP) [13]; convergence of electronic media and telecommunication services. Radio and television changed deeply the psycho-social background of life. They are

accessible by Internet today, but there are much more services accessible by Internet, which are deeply transforming usual life relations [5], [6], [7].

We study the development of the telecommunication services taking into account their psycho-social effects. The research has been intended to have two main stages. At the first stage, a statistical study was achieved concerning the correlation between the usage of the telecommunication services and the user's personality traits. At the second stage, methods of selective telecommunication services and contents delivery according to the individual and group users' profile are to be defined and developed.

The present paper relates the first stage. Selectively delivering of telecom content according to the user's personality traits is our final target.

The phenomena resulting from the usage of telecommunication services are studied mainly by psychology and/or telecommunication marketing.

The first approach studies the effects of using the different telecommunication services, which are considered mainly negative and undesirable (dependence, addiction, anxiety, fear, depression, etc.). The resulted recommendations are mainly targeted to avoid using or to reduce the time to be connected to telecommunication services.

The second approach aims to maximize the usage time of offered services, maximizing the profit which is realized.

These two approaches are rather contradictory.

Studies made after year 2000, show that young people aged between 15 and 24 years spend daily 1-3 hours on Internet. Main purposes are music listening, entertainment and searching for informational contents.

There are several models to evaluate human personality traits. All these models are grounded on the (lexical) hypothesis, which states that building and maintaining inter-personal relations are coded in language, through the human habits describing concepts.

After Howard (2004), the factors of the Big-Five Model are:

1. Neuroticism - N
2. Extraversion - E
3. Openness to experience - O
4. Agreeableness - A
5. Conscientiousness - C

A very interesting direction of the research is in which the main character traits are described by studying the preferences in accessing Internet contents. It is confirmed in this way too, the existence of 5 main character groups (clusters) [9].

The effects of the usage of the telecommunication services are very important to be studied. In order to develop telecommunication services according to their psycho-social effects, deeper knowledge of relationship is

required. A questionnaire to measure this relationship with the youth has been elaborated using an adapted version of the Delphi consultation ([16]). Instead of asking to estimate the exact number of calls, or exact duration of service usage, 6 quantified possible answers were introduced:

0 = not at all, 1= very little, 2 = little, 3 = average, 4 = much, 6 = very much.

This combined questionnaire [17] has been applied on a sample of students. Having been suggested by psychologists, we carried out the survey through handouts, not online. From the total of 1100 students of Tg. Mures Faculty of the Sapientia University about 300 students have been randomly chosen to respond to the questionnaire. For processing the results, we used multi-linear regression [3], [8], [18], cluster analysis [1], discriminant analysis for the cluster classification [2], Principal Component Analysis (PCA) [10], and a newly introduced model based on simultaneous bursts [14], [15].

To identify the different types of behaviour in using the telecommunication services we have applied the clustering method and thus succeeded in identifying three significantly different types:

1. The socializing;
2. The pragmatic;
3. The all-consuming.

Having used the discriminant analysis we have studied the relationship between the psychological factors obtained from the standard psychological tests and the classification obtained from the clustering technique as well as from identifying the different age and gender categories from the information related the usage of the telecommunication services. We have noticed that:

- Psychological factors explain a minor part of the classification;
- There is a significant difference between the very close age groups like 19-20 years and 21-22 years of age;
- There are significant differences between men and women what media usage concerns;

Using the factorial analysis we have chosen 6 principal components (with the varimax method) out of the data related to the media usage. These factors could be used to conduct a new poll with much fewer questions and also with relevant results.

To strengthen the explicit power of the information of the survey we have created a series of transformed variables which expresses the existence of coincidences in the responses given to the pairs of questions. Using these transformed variables we managed to build a model which can explain about 85% of the classification with the help of 19 such coincidences.

2. Cluster analysis of the data related to the usage of media

Cluster classification was made on the basis of media usage and of the media contents, 61 possibilities in 5 groups. For the obtained data through the questionnaire on media usage we adopted different clustering techniques (hierarchical, K-mean) and different types of distances (Euclidian, s-Euclidian, etc.) to compare the obtained classifications. These numeric experiences were intended to identify the number of the categories adopted at classification and to study (at least on intuitive level) the stability of the created categories with respect to the parameters of the method we have used of the created categories. We have found that the number of the significantly different categories and which are not reduced to very small groups (with changed classification techniques) is between 2 and 4.

We calculated the mean (m) and dispersion (s) within each category related to each question for two classifications: one with three (*Table 1*) and one with four categories. Owing to the hierarchical categorizing method the three categories in the first group may be obtained from those four in the second group by merging two groups. Consequently the second category in the first group (where we have 3 categories) is obtained by merging categories 2 and 3 in the second group (where we have 4 categories).

Table 1: Mean and dispersion in the structures with 3 categories (clusters).

	cl1	cl2	cl3	cl1	cl2	cl3
	Soc (97)	Cons (102)	Prag (63)	Soc (97)	Cons (102)	Prag (63)
Mobile calls	*m*	*m*	*m*	*s*	*s*	*s*
Solving personal problems	2.60	2.45	2.21	1.24	1.32	1.14
Brief personal communication	2.88	2.54	1.97	1.06	1.31	1.27
Long personal communication	1.94	2.16	1.24	1.27	1.22	1.10
Long communication	0.93	1.69	0.41	1.14	1.23	0.69
SMS	3.05	2.83	1.56	1.46	1.62	1.54
Mobile	m	m	m	s	s	s
Music	1.23	1.81	0.75	1.26	1.40	1.22
Internet browsing	1.05	1.88	0.86	1.36	1.52	1.19
Internet chat	0.46	1.50	0.40	0.99	1.49	0.91
Internet e-mail	0.56	1.25	0.41	1.00	1.24	0.85
Games	0.73	1.34	0.48	0.99	1.29	0.91

	cl1	cl2	cl3	cl1	cl2	cl3
	Soc (97)	Cons (102)	Prag (63)	Soc (97)	Cons (102)	Prag (63)
Radio	*m*	*m*	*m*	*s*	*s*	*s*
Information	1.09	1.77	0.32	1.11	1.24	0.62
News	1.22	1.78	0.38	1.14	1.27	0.61
Politics	0.37	1.14	0.13	0.86	1.15	0.42
Culture	0.90	1.69	0.17	1.16	1.28	0.52
Science	0.73	1.70	0.19	1.13	1.33	0.47
Music	2.56	2.94	0.78	1.68	1.60	1.05
Sports	0.52	1.45	0.16	0.93	1.38	0.54
Quiz	0.20	0.99	0.13	0.51	1.09	0.42
TV	*m*	*m*	*m*	*s*	*s*	*s*
News	0.81	2.19	0.40	1.03	2.36	0.77
Politics	0.28	1.44	0.08	0.64	1.10	0.33
Culture	0.66	1.96	0.17	0.98	1.15	0.58
Science	0.99	2.32	0.68	1.31	1.22	1.16
Music	1.06	2.52	0.16	1.31	1.38	0.48
Sport	0.67	2.03	0.48	1.21	1.45	1.22
Quiz	0.26	1.24	0.03	0.77	1.24	0.25
Interactive programs	0.44	1.84	0.29	0.92	1.31	0.83
Video clips	0.53	2.14	0.02	0.97	1.27	0.13
Feature film	0.40	1.70	0.08	0.95	1.34	0.41
Movie serials	0.64	2.40	0.30	1.16	1.53	0.75
Documentaries	1.11	2.15	0.40	1.38	1.41	0.93
Action films	0.57	2.12	0.29	1.06	1.40	0.71
Comedies	1.04	2.82	0.27	1.36	1.28	0.70
Plays	0.25	1.37	0.00	0.69	1.22	0.00
Romance films	0.48	2.05	0.11	1.06	1.35	0.44
Reality show	0.38	1.73	0.11	0.89	1.53	0.44
Horror	0.32	1.68	0.03	0.90	1.54	0.25
Erotic	0.16	1.31	0.02	0.55	1.47	0.13
Internet	*m*	*m*	*m*	*s*	*s*	*s*
Professional resource	2.58	2.61	2.16	1.25	1.20	1.46

	cl1	cl2	cl3	cl1	cl2	cl3
	Soc (97)	Cons (102)	Prag (63)	Soc (97)	Cons (102)	Prag (63)
E-mail	2.30	2.42	1.54	1.09	1.21	1.19
Games	1.06	1.88	0.89	1.43	1.54	1.43
Messenger (e.g. Yahoo)	3.13	2.82	1.57	1.44	1.42	1.15
Socialization sites (e.g. Facebook)	3.18	2.97	1.63	1.38	1.45	1.39
Information	2.57	2.82	1.79	1.26	0.98	1.35
E-books, e-textbooks	1.84	2.18	1.06	1.35	1.30	1.26
Encyclopedia (e.g. Wikipedia)	1.71	2.25	1.08	1.32	1.15	1.04
E-shopping	0.53	1.22	0.27	0.98	1.18	0.65
News	1.20	2.05	0.48	1.28	1.20	0.82
Politics	0.27	1.21	0.27	0.65	1.21	0.77
Culture	1.26	1.77	0.37	1.37	1.24	0.77
Science	1.52	2.29	1.19	1.46	1.22	1.33
Music	2.98	3.43	1.83	1.40	1.16	1.67
Sport	0.75	1.88	0.46	1.16	1.51	1.06
Matrimonial ads	0.26	0.92	0.02	0.79	1.20	0.13
Video clips	1.61	2.48	1.03	1.48	1.36	1.49
Feature film	0.94	1.46	0.16	1.38	1.45	0.63
Documentaries	1.41	1.80	0.59	1.53	1.46	1.10
Action films	1.35	2.01	0.49	1.59	1.47	1.05
Comedies	1.99	2.75	0.78	1.64	1.32	1.18
Romance films	0.98	2.16	0.16	1.43	1.45	0.60
Horror	0.87	1.71	0.24	1.40	1.48	0.76
Erotic	0.34	1.23	0.17	0.78	1.41	0.55

The table shows that those belonging to the first category use the services mainly for socialization and for this reason we called this type socializing behaviour. Those in the last category use these services mainly to solve current problems and for professional activities. For this we called the type pragmatic behaviour. Second category can be called all-consumers.

These categories can be differentiated at each type of services, moreover, at the majority of the questions of the questionnaire. The dispersion does not differ significantly (it has been tested by F test), and the means are significantly different (Students's t test): $p < 0.05$.

For a suggestive visualization we have calculated the centroids, the equation of the plane defined by these three centroids, and we projected each individual on this plane. This step has been realized with MATLAB because the adopted application for the rest of the calculations (SPSS ver.21) has not got this function. Thus the data from SPSS has been exported to EXCEL and from there imported to MATLAB ([4], [12]) for processing.

Thus we obtained *Fig. 1* where the green spots correspond to the socializing category, the blue ones to the consumers and the yellow spots to the pragmatists. The red circles mark the centroids.

These categories generate an avalanche of problems related to future research: how can these individuals' opinions be used to create a psychological index of the contents based on "face recognition" and "voice recognition" concerning the extraction of the users' reactions? Is it true that the pragmatists first choose the content and only afterwards the form of the service?

Concerning the possibility of identifying the age and sex categories we analyzed the cluster categorization, too, on the following subcategories: women aged 19-20 (21), women aged 21-22 (22), men aged 19-20 (11) and men 21-22 (12). For the first two categories we had to eliminate some data in order to obtain substantial categories. So for category 21 we excluded 11 persons, from category 22 we excluded 6 persons.

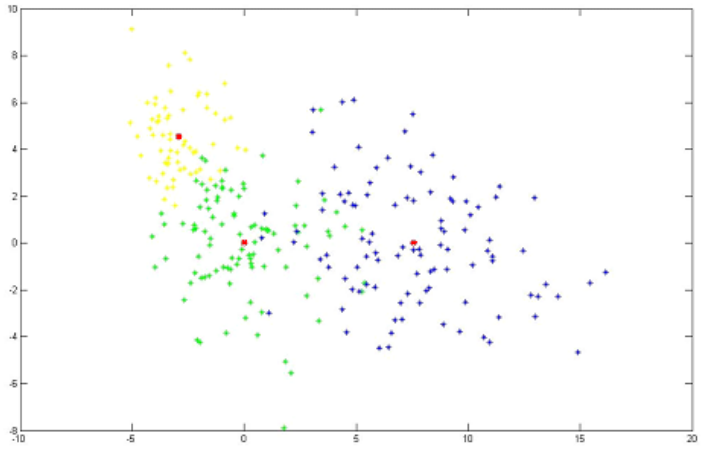

Figure 1: Projection of each individual on the plane defined by the three centroids. Values are coordinates in this plane.

Both classifications (with two classes) show that on subcategory level two behaviours can be identified (compared with the three behaviors identified for the whole population). At subcategory 21 a more emphatic socializing behavior can be identified for a part of the population and the rest of the population (including a part of the socializing category, all pragmatics and all consumers from this subcategory) can be considered as a separate category. At subcategory 22 the emphatic socializing group becomes more significant and includes a part (small) of the consumers and the other part is formed by a significant nucleus of pragmatists, a part of consumers and a negligible part of the socializing category.

It can be noticed with the women a radical change as regards the structure of the population even at this slight age difference.

Analyzing from the same point of view the subpopulation composed of men we have noticed that it is not necessary to exclude any data (without individuals with exaggerated "individualism") and it cannot be noticed the change of the structure either. Certainly these results raise a series of questions: how this structure of usage of services evolves in time, which are the external parameters the evolution depends upon, what kind of relations exist between this structure and other social characteristics (for example whether one is involved in a relationship or not) etc.

3. Discriminant analysis of the classification obtained from cluster analysis using psychological factors

The discriminant analysis between the factors of the psychological tests and of the results of the cluster analysis obtained in the previous paragraph offers a model which shows the explicative power of the psychological factors in the classification obtained from the usage of the media. From *Table 2* results that two functions are sufficient and these have a relatively big explanatory power in case of the target variable (59% respectively 41%).

Table 2: Eigenvalues.

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.370a	59.0	59.0	.520
2	.257a	41.0	100.0	.452

Note: First 2 canonical discriminant functions were used in the analysis.

The values obtained with the Wilks' Lambda test (*Table 3*) show that there are significant differences between the constructed functions. We applied the Box test also for the equality of the covariance matrices and the results show that the discriminant analysis can be applied.

Table 3: Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.580	86.492	54	.003
2	.795	36.400	26	.085

The following two tables contain discriminant function coefficients and correlated coefficients of the initial variables (psychological factors) and values created through discriminant functions.

Table 4: Standardized Canonical Discriminant Function Coefficients.

	Coefficients of the functions	
	1	2
STAI_S	-.114	.188
STAI_T	-.099	.504
BECK_DEPR	-.150	.111
SBQ	.083	-.183
COP_PROBLEM	-.343	.051
COP_EMOTION	.680	.284
COP_AVOID	.474	-.301
PISI_01	-.413	.478
PISI_02	.048	-.358
PISI_03	.339	-.182
PISI_04	-.403	-.022
PISI_05	-.116	-.182
PISI_06	.173	.010
PISI_07	.198	-.536
PISI_08	-.240	.673
PISI_09	.287	-.366
PISI_10	.027	-.068
PISI_11	.670	.321
PISI_12	.003	-.104
PISI_13	-.008	.107
PISI_14	-.170	-.628
PISI_15	.540	1.014
PISI_16	.047	-.132
EPQ_P	-.196	.147
EPQ_E	.134	-.151
EPQ_N	.226	-.084
EPQ_L	-.096	.463

The functions are in fact the linear combinations from the explicative variables values, and the values in the tables are the coefficients of these

combinations. The values of these functions will give us a categorization based on the explicative variables and this categorization will be compared with the initial one defined through the target variable. *Table 4* also shows the impact of the initial variables on the two discriminant functions. The green color marks the factors with raised share in the discriminant function 1, and the yellow marks the high share in the discriminant function 2.

Table 5 contains the correlated coefficients between the psychological factors and the discriminant functions.

Table 5: Structure Matrix.

	Function	
	1	2
COP_AVOID	.558*	-.018
PISI_11	.490*	.140
PISI_14	-.390*	-.341
COP_EMOTION	.361*	.111
PISI_16	-.336*	-.114
EPQ_E	.319*	-.148
PISI_07	.282*	-.241
EPQ_N	.239*	.090
PISI_12	-.237*	-.190
EPQ_L	-.230*	.149
PISI_10	.216*	-.092
PISI_15	-.216*	.001
PISI_13	-.181*	-.085
PISI_06	.179*	.135
SBQ	.143*	.003
PISI_04	-.133*	.031
PISI_09	.066*	-.056
BECK_DEPR	.108	.227*
STAI_S	-.003	.208*
PISI_05	-.173	-.205*
PISI_08	.199	.203*
STAI_T	.092	.196*
COP_PROBLEM	-.017	-.174*
PISI_03	-.084	-.140*
EPQ_P	.027	.122*
PISI_02	.028	-.098*
PISI_01	-.028	.065*

The calculated discriminant function values on the mean of the three categories (*Table 6*) show that these have different comportment (these types of comportment can be identified with plane vectors having different directions,

the first being in the first quadrant, the second in the fourth one and the third in the second one).

Table 6: Functions at Group Centroids.

Cluster	Function	
	1	2
1.00	.473	.454
2.00	.124	-.646
3.00	-1.107	.244

Fig. 2 shows the values of the discriminant functions and the classification obtained based on them. The discriminant function values for each member of the population: each circle represents an individual of the population, the two coordinates are the values of the two discriminant functions, and the color is obtained by categorizing based on these functions.

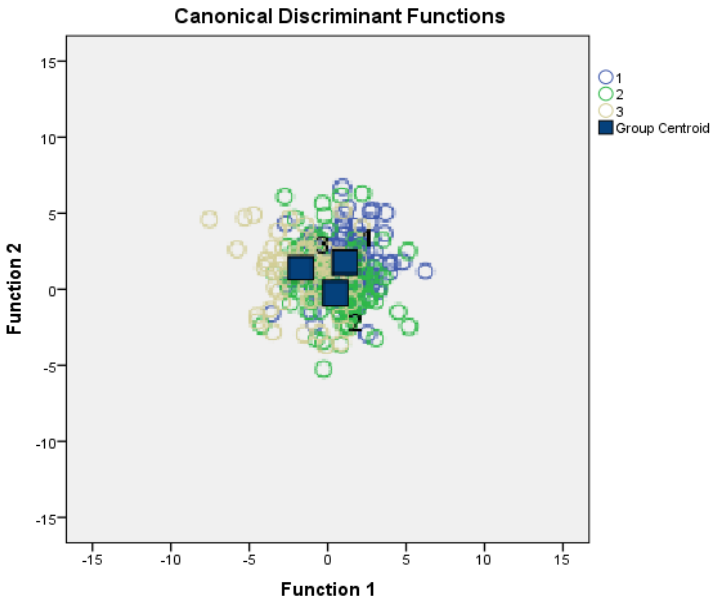

Figure 2: Discriminant function values for each member of the population.

Table 7 shows the fact that on the basis of the calculated factors of the psychological tests we can correctly classify 55.7 % of the whole population.

Studying the structure of the discriminant functions we can see that not all the psychological factors are relevant in this analysis.

For this we construct two more models, the first based on the PISI test and COPING and the second on just some factors selected from these.

Certainly through selecting these variables we will lose from the explanatory power but we will get a correct classification for 48.5% of the population based on only 4 psychological factors instead of 27.

Table 7: Classification Results.

		Cluster	Predicted Group Membership			Total
			1.00	2.00	3.00	
Original	Count	1.00	70	24	8	102
		2.00	35	47	15	97
		3.00	18	16	29	63
	%	1.00	68.6	23.5	7.8	100.0
		2.00	36.1	48.5	15.5	100.0
		3.00	28.6	25.4	46.0	100.0

Note: 55.7% of original grouped cases correctly classified

4. Discriminant analysis, Model based on PISI and COPING factors

The Wilks' Lambda, Chi-square and Box gave positive results. Coefficients calculated (*Table 8*):

Table 8: Standardized Canonical Discriminant Function Coefficients.

	Function	
	1	2
COP_PROBLEM	-.299	.316
COP_EMOTION	.504	.014
COP_AVOID	.458	.714
PISI_01	-.249	-.410
PISI_02	.028	.278
PISI_03	.304	.179
PISI_04	-.346	.062
PISI_05	-.237	.262
PISI_06	.169	-.048
PISI_07	-.016	.612
PISI_08	.031	-.578
PISI_09	.293	-.018
PISI_10	.254	-.114
PISI_11	.479	-.037
PISI_12	-.021	.347
PISI_13	.032	.038

	Function	
	1	2
PISI_14	-.330	.342
PISI_15	.558	-.311
PISI_16	-.156	.203

From this table we can extract the factors with relatively high contribution (green colour). The next model will be constructed on these four factors.

Also in this case the calculated discriminant functions in the centroids of the three categories separated out into 3 quadrants (*Table 9*):

Table 9: Functions at Group Centroids.

Cluster	Function	
	1	2
1	.527	-.288
2	-.033	.471
3	-.821	-.341

Each circle in *Fig. 3* represents an individual of the population.

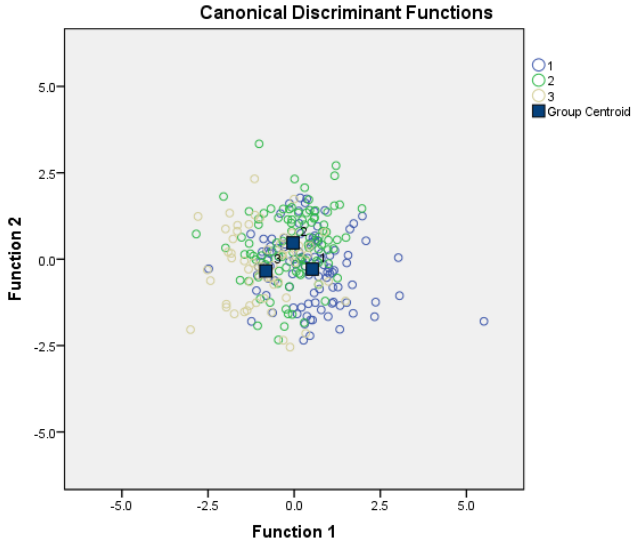

Figure 3: Discriminant function values, COPING and PISI models.

The 2 coordinates are the values of the 2 discriminant functions, and the colors are obtained by categorizing based on these functions.

In the previous model where 55.7% could be classified, in this situation still 54.6% of the cases can be classified (*Table 10*).

Table 10: Classification results.

		Cluster	Predicted Group Membership			Total
			1.00	2.00	3.00	
Original	Count	1.00	57	36	9	102
		2.00	27	57	13	97
		3.00	10	24	29	63
	%	1.00	55.9	35.3	8.8	100.0
		2.00	27.8	58.8	13.4	100.0
		3.00	15.9	38.1	46.0	100.0

5. Discriminant analysis, Model based on the 4 most relevant factors

- COP_EMOTION- Coping centered on emotions (Coping is an active process through which an individual can deal successfully with a stressful situation and succeeds in controlling it [19]).
- COP_AVOID-coping centered on avoiding
- PISI_11 - The power of social creation
- PISI_15 - Controlling sentiments

Similarly the previous steps were remade starting just from the four most relevant factors.

In this case too, the discriminant function values calculated in the centre of centroids of those three categories separated into 3 quadrants (*Table 11*).

Table 11: Functions at Group Centroids

Cluster	Function	
	1	2
1.00	.350	.136
2.00	.100	-.172
3.00	-.722	.066

Each circle in *Fig. 4* represents an individual of the population, the 2 coordinates are the values of the 2 discriminant functions, and the color is obtained by categorizing based on these functions.

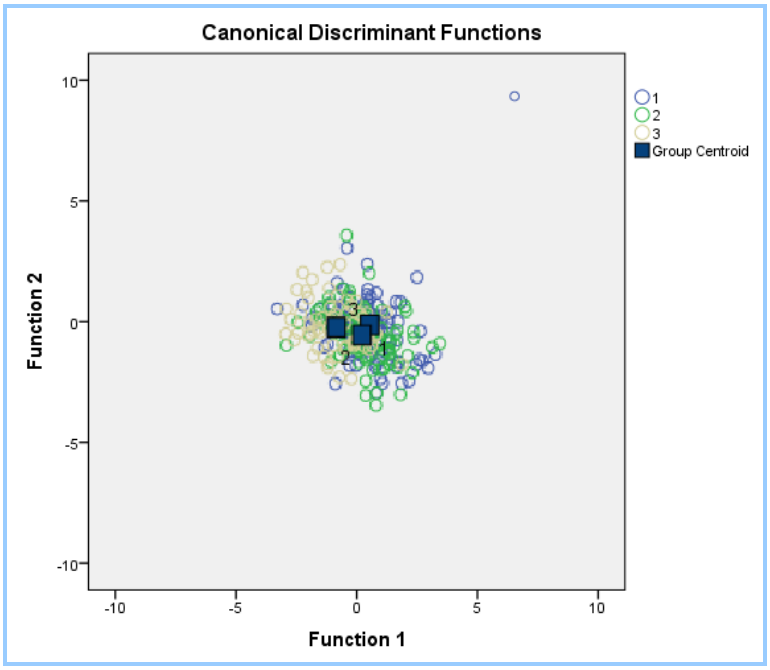

Figure 4: Discriminant function values, model with 4 factors.

In the previous model 54.6% could be classified. In this situation 48.5% of the cases can be classified (*Table 12*):

Table 12: Classification results.

		Cluster	Predicted Group Membership			Total
			1.00	2.00	3.00	
Original	Count	1.00	47	42	13	102
		2.00	32	52	13	97
		3.00	12	23	28	63
	%	1.00	46.1	41.2	12.7	100.0
		2.00	33.0	53.6	13.4	100.0
		3.00	19.0	36.5	44.4	100.0

Note: 48.5% of original grouped cases correctly classified

Consequently by reducing the psychological factors from 27 to 4 we lose only 7.2% from the explanatory power.

6. Principal component analysis

The value for the Bartlett test and Kaiser-Meyer-Olkin measure show that the correlation matrix is significantly different from the identity matrix and that the correlative coefficients can be explained relatively well through partial correlates, so there is no need for Anti-Image Correlation Matrix and the factorial analysis can be used.

The table of the communalities (not included) shows how much per cent of each variant of the initial variable can be explained by using the 6 factors drawn out through the PCA method (and the rotation of these using the Varimax process).

Table 13, contains the global explanation of the factors and shows the factors extracted on the bases of Cattell criterion (Scree test).

The graph obtained from eigenvalues (Scree plot) explains 50.65% of the total variants.

Table 13: Eigenvalues for extracted components (fragment).

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	15.706	25.748	25.748
2	4.170	6.837	32.584
3	3.736	6.124	38.709
4	2.722	4.463	43.171
5	2.672	4.381	47.552
6	1.892	3.101	**50.653**
7	1.766	2.895	53.548
8	1.504	2.466	56.014
9	1.434	2.351	58.365
10	1.376	2.256	60.621
11 ...			

Certainly the analysis can be repeated to extract more factors, but in this case we will have two problems: we are either notable to explain the factors or the methods of calculating the optimal rotations are not convergent.

For this reason we kept just 6 factors and in the annexed table we included just the first 36 values (which are necessary to be able to explain 90% of the total variant).

The Cattell (scree test) is a pseudo-empirical test to establish the number of relevant factors. Consequently the eigenvalues, which line up along a straight line, will not be taken into consideration.

In this case, after the 6th eigenvalue the others are practically on a straight line, so according to the Cattell criterion the extraction of 6 values is recommended (*Fig. 5*).

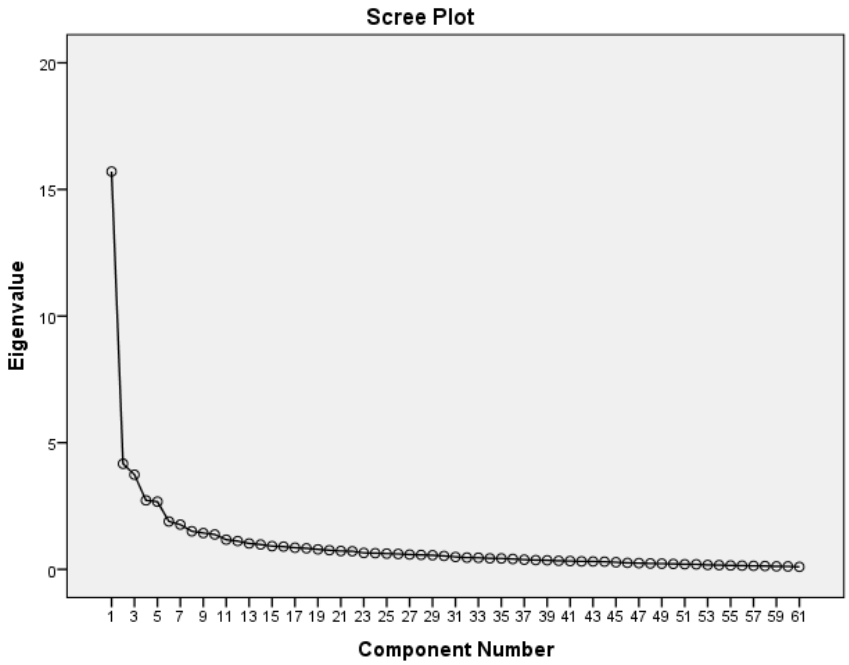

Figure 5: Eigenvalues for Cattel test.

To obtain a more suggestive correlation between the initial variables and obtained components, and to point out the possible interpretations of gained factors we rotated them using the Varimax process (we took just the relevant parts where the coefficient of correlation is at least 0.6).

We obtained 6 new components (*Table 14*).

The first factor includes the usage of cultural services on a large scale (culture, plays, films, music etc.).

The second factor is mainly linked to information (information, politics, science, sport, news).

The third factor refers to the information accessed electronically.

The fourth factor refers to Internet usage for mail, chatting, searching, the fifth one is related to action and horror films and the sixth to socializing.

Table 14: Rotated Component Matrix, relations between components and initial variables.

	Component					
	1	2	3	4	5	6
Mobile						
Internet browsing				.778		
Internet chat				.758		
Internet e-mail				.720		
Radio						
Information		.719				
News		.697				
Politics		.621				
Culture		.711				
Science		.699				
Sports		.633				
TV						
Culture	.602					
Music	.693					
Video clips	.692					
Movie serials	.708					
Comedies	.746					
Plays	.604					
Romance films	.747					
Reality show	.661					
Internet						
E-mail			.654			
Socialization sites (e.g. Facebook)						.635
E-books, e-textbooks			.720			
Encyclopedia (e.g. Wikipedia)			.686			
Culture			.625			
Action films					.725	
Horror					.689	

Notes: Extraction Method: Principal Component Analysis
Rotation Method: Varimax with Kaiser Normalization
Rotation converged in 8 iterations

7. Model based on coincidences (simultaneous bursts)

To increase the explanatory power of the initiative classification we constructed for each pair of initial variables (Kx.y in *Table 15*) a new variable which measures the simultaneous bursts of the two initial variables.

So the new variable can take the values of -2, -1, 0, 1 or 2 depending on the values of the initial variables according to the following rules:

- The value is 2 if the initial variables have the same value;
- The value is 1 if the modulus of the difference between the initial variables is 1;
- The value is -1 if the modulus of the difference between the initial variables is one unit smaller than the possible maximal value;
- The value is -2 if the modulus of the difference between the initial variables is maximum possible;
- The value is 0 in all the other cases.

So of the 61 initial variables we constructed 61×60/2 = 1830 new variables (using the syntactic programming of SPSS - [11]) to which we attached the factors too, obtained from evaluating the psychological tests.

After this by constructing some successive regressive models we eliminated the variables which did not present a significant contribution to the explanatory power of the module.

As a consequence of this interactive process (after 9 steps) we obtained a set of 21 significant variables which together have an explanatory power of 85%. These variables being numbered in order of their automatic generation we decoded them and we obtained the decoding table (*Table 15*):

To select from these variables those which have a significant contribution in the explanation of categorization it was calculated the coefficient of correlation among these variables and the target variable and the significance of each test was tested.

Thus 9 variables were eliminated from the model and it was obtained a model of linear regression which has an explanatory power of 74% based on 2 psychological factors (PISI_11, COP_EMOTION) and 10 coincidences to the questions referring to the usage of telecommunication services.

To select from these variables those which have a significant contribution in the explanation of categorization it was calculated the coefficient of correlation among these variables and the target variable and the significance of each test was tested.

Table 15: Decoding table for the new variables.

NewVar	i			j
384	7	K.2.2	K.5.15	52
422	8	K.2.3	K.5.1	38
423	8	K.2.3	K.5.2	39
429	8	K.2.3	K.5.8	45
579	11	K.3.1	K.5.8	45
646	13	K.3.3	K.3.7	17
739	15	K.3.5	K.4.1	19
757	15	K.3.5	K.4.19	37
767	15	K.3.5	K.5.10	47
806	16	K.3.6	K.5.5	42
1463	35	K.4.17	K.5.16	53
1492	36	K.4.18	K.5.21	58
1502	37	K.4.19	K.5.8	45
1508	37	K.4.19	K.5.14	51
957	20	K.4.2	K.4.9	27
961	20	K.4.2	K.4.13	31
978	20	K.4.2	K.5.11	48
993	21	K.4.3	K.4.6	24
1628	43	K.5.6	K.5.17	54

After this analysis *Table 16* was obtained and thus in the decoding table the significant variables were marked with green.

Thus 9 variables were eliminated from the model and it was obtained a model of linear regression which has an explanatory power of 74% based on 2 psychological factors (PISI_11, COP_EMOTION) and 10 coincidences to the questions referring to the usage of telecommunication services.

For the rest of the variables the significance level (Sig) is bigger than 0.05, so we omitted these variables.

These results can be used for both designing a new questionnaire and for identifying the belonging of some individuals to the identified categories.

Beside these practical aspects the above described process raises a series of theoretical questions related to the possibility of generalizing this process to obtain factors based on creating variables of coincidences in general,

respectively the a priori testing of the possibility of some improvement by creating variables of coincidence type.

Table 16: Correlation among significant variables and the target variable.

Model	Unstandardized Coefficients		Stand-ized Coeffic	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
(Constant)	4.108	.435		9.435	.000	3.248	4.967
NewVar423	-.291	.068	-.273	-4.279	.000	-.425	-.157
PISI_11	-.036	.009	-.187	-3.798	.000	-.055	-.017
NewVar422	.198	.055	.191	3.621	.000	.090	.307
NewVar757	-.194	.056	-.212	-3.490	.001	-.303	-.084
NewVar1463	-.169	.054	-.173	-3.119	.002	-.276	-.062
NewVar993	-.151	.054	-.162	-2.778	.006	-.258	-.044
NewVar429	.158	.060	.158	2.651	.009	.040	.276
NewVar384	-.115	.044	-.119	-2.581	.011	-.203	-.027
NewVar961	-.172	.068	-.185	-2.516	.013	-.306	-.037
NewVar767	-.111	.044	-.119	-2.507	.013	-.199	-.024
NewVar978	-.139	.056	-.119	-2.493	.014	-.250	-.029
NewVar579	-.122	.053	-.123	-2.289	.023	-.228	-.017
NewVar1492	-.135	.063	-.154	-2.150	.033	-.260	-.011
NewVar1628	.098	.046	.098	2.124	.035	.007	.189
NewVar739	.122	.059	.127	2.056	.041	.005	.238
COP_EMOTION	-.009	.005	-.135	-2.043	.043	-.019	.000
NewVar957	-.112	.057	-.114	-1.971	.050	-.224	.000

3. Conclusion

The majority of the psychological factors (from clinical tests) do not have a significant relevance for the usage of telecommunication services apart from COP_EMOTION, PISI_11 factors. For further research the identification of other psychological factors is needed.

In the usage of telecommunication services three big categories of users can be identified: the socializing, the all-consuming and the pragmatic. These

categories represent a quite interesting dynamics if we analyze the different subcategories (women-men, age categories). A more detailed study of this phenomenon would deserve a supplemental attention.

To construct content filters it is advisable to exclude the pragmatists (or to start working with them in an alternative, more conscientious way).

References

[1] András, Sz., and Baricz, Á., "Statisztika közgazdászoknak (Statistics for Economists)", Editura Státus, Miercurea Ciuc, 2010.

[2] Andrews, N. O., and Fox, E. A., "Recent Developments in Document Clustering", *Technical Report TR507535, Computer Science*, Virginia Tech., 2007.

[3] Bilodeau, M., and Brenner, D., "Theory of Multivariate Statistics", Ed. Springer, 1999.

[4] Blaga, P., "Statistică prin Matlab", Ed. Presa Universitară Clujeană, 2002.

[5] Bromley, R., V., and Bowles, D., "Impact of Internet on use of traditional news media", *Newspaper Research Journal*, Department of Journalism, University of Memphis 16(2), 1995.

[6] Chan, K.; and Fang, W., "Use of the internet and traditional media among young people, Young Consumers: Insight and Ideas for Responsible Marketers", *Emerald* 8(4) pp. 244-256, 2007.

[7] Chen, R, Su, E., Shen, V., Wang, Y., "Introduction to IP Multimedia Subsystem (IMS), Part 1: SOA Parlay X Web services", IBM, http://www.ibm.com/developerworks/ webservices/library/ws-soa-ipmultisub1/, 2006.

[8] Ezekiel, M., Fox, K. A., "Korreláció és regresszióanalízis (Correlation and Regression Analysis)", Közgazdasági és Jogi Kiadó, Budapest, 1970.

[9] Faber, M. A., Mayer, J. D., "Resonance to archetypes in media: There's some accounting for taste", *Journal of Research in Personality*, 43(3): pp. 307-322, 2009.

[10] Jackson, J. E., "A User's Guide to Principal Components (Wiley Series in Probability and Statistics)", Ed. Wiley-Interscience, 2003.

[11] Landau, S., Everitt, B. S., "Handbook of statistical analysis using SPSS", Chapman&Hall/ CRC, 2004.

[12] Martínez, W. L., Martínez, A. R., "Computational Statistics Handbook With MATLAB", Chapman and Hall/CRC Computer Science and Data Analysis Series, Taylor and Francis, 2008.

[13] Milne, R., "ICT quality of service regulation: practices and proposals", ITU website, 2006, [Online]. Available on 1st March 2013, http://www.itu.int/ITU-D/treg/Events/Seminars/ 2006/QoS-consumer/documents/SI-2_QOS-bkgpaper-presentation.pdf. 2006.

[14] Molnár, L., "Development of the Telecommunication Services According to their Psycho-Social Effects", PhD thesis, Transilvania University, Brasov, Romania, 2013.

[15] Molnár, L., András, Sz., Székely, I., "Relation between media content usage and known psychological factors", *4th International Conference on Recent Achievements in Mechatronics, Automation, Computer Science and Robotics (MACRo2013)*, Sapientia University, Tg. Mureş, 2013.

[16] Molnár, L., Székely, I., "Adapted Delphi Consultation to Elaborate a Questionnaire for Measuring the Usage of Different Telecommunication Services", in *The 6th edition of the Interdisciplinarity in Engineering International Conference* (*InterEng*), "Petru Maior" University of Tîrgu Mureş, pp. 388-393, 2012.

[17] Molnár, L., Székely, I., "Simultaneously Measurement of the Usage of some Telecommunication Services and some Psycho Factors", *XXII*nd *Conference on Computers and Education - SzámOkt*, Alba Iulia, pp. 319-324, 2012.
[18] Mundruczó, Gy., "Alkalmazott regressziószámítás (Applied Regression Analysis)", Akadémiai Kiadó, Budapest, 1981.
[19] MUT_VITCU&Asociatii, "Index de termeni", http://www.mutvitcu.ro/index-de-termeni/coping/20.