Babeș-Bolyai University
Faculty of Sociology and Social Work
Sociology Doctoral School

The Structure and Dynamics of Popular Online Social Networking Systems

Ph.D. candidate: Radu-Mihai Meza

Supervisor: prof. Traian Rotariu, Ph.D.

February 2012

Table of Contents:

Introduction **7**

I The study of social networks **13**

I.1 Structural aspects of the study of human relationships 13

I.2 Forerunners of social network analysis 15

I.2.1 The preoccupation for structural approaches in the nineteenth century 15

I.2.2 Pareto and the ”80-20 rule” in the study of complex systems 16

I.2.3 Jacob Moreno and sociometry 18

I.2.4 The birth of social network analysis at Harvard 21

I.2.5 Merton, Lazarsfeld and structuralist approaches 22

I.3 Talcott Parsons, structural functionalism and the influences of cybernetics 24

I.4 Milgram and the six degrees of separation 26

I.5 Granovetter and the strength of weak ties 27

I.6 Niklas Luhmann and society as a system of communication 28

I.7 Social network analysis 29

I.7.1 Methods in social network analysis 31

I.7.2 The characteristics of networks 33

I.8 Conclusions 34

II Computer mediated communication (CMC) **36**

II.1 The personal computer as a commodity 36

II.1.1 Intel inside 36

II.1.2 IBM PC and the vision of ”a computer in every home” 37

II.1.3 Software and hardware 37

II.2 The Internet 39

II.2.1 Communication in local area networks 42

II.2.2 The transformation of the Internet in a network accessible to the broad public 43

II.3 World Wide Web 43

II.4 Theoretical perspectives in sociology applicable to the study of CMC 46

II.5 Theoretical perspectives in the study of CMC 49

II.5.1 New media theory 49

II.6 Computer mediated communication 54

II.6.1 The „cues-filtered-out” approach 55

II.6.2 Avatar research 56

II.7 The prehistory of computer mediated communication 56

II.8 Forms of asynchronous computer mediated communication 58

II.8.1 E-mail 58
II.8.2 From Bulletin Board Systems to forums and discussion groups 60
II.9 Forms synchronous computer mediated communication 65
II.9.1 Multi-User Dungeon (MUD) and MUD, Object Oriented (MOO) 65
II.9.2 Internet Relay Chat (IRC) 70
II.9.3 Instant Messaging (IM) 74
II.10 The study of social networks established through CMC in the 1990s 79
II.11 Conclusions 83
III Computer mediated communication in the Web 2.0 framework 85
III.1 Online systems - from Web 2.0 to „social media" 85
III.2 Services associate with Web 2.0 86
III.2.1 Blogs 87
III.2.2 Wikis 88
III.2.3 Multimedia sharing 89
III.2.4 Tagging and social bookmarking 91
III.2.5 Newsfeeds, RSS and content aggregation systems 91
III.2.6 Podcasting 92
III.2.7 Mash-ups 93
III.2.8 The reproduction of desktop applications as online services 93
III.2.9 Content Management Systems (CMS) 93
III.2.10 Social network sites and social networking systems 94
IV Social network sites (SNS) 97
IV.1 Definitions 97
IV.2 Classification 98
IV.2.1 Socializing SNS 99
IV.2.2 Networking SNS 100
IV.2.3 Social navigation SNS 100
IV.3 A classification of users and uses of social network sites 103
IV.4 A brief history of social network sites 105
IV.4.1 1985 - WELL 105
IV.4.2 1994 – Geocities 105
IV.4.3 1995 – Classmates.com 105
IV.4.4 1995 – Match.com and dating sites 106
IV.4.5 1997 – Sixdegrees.com 106
IV.4.6 1999 – Cyworld 107
IV.4.7 AsianAvenue, BlackPlanet and MiGente 108

IV.4.8 2002 – Friendster 108
IV.4.9 2003 – MySpace 109
IV.4.10 2003 – LinkedIn 113
IV.4.11 2003 – Gaia Online 116
IV.4.12 2003 – hi5 117
IV.4.13 2004 – Orkut, Google Buzz, Google Friend Connect, Google + ... 119
IV.4.14 2004 – Facebook 124
IV.4.15 Other popular social networking systems 131
IV.5 Approaches and methodologies in the study of social networking systems 131
IV.5.1 Quantitative studies on automatically extracted large datasets 131
IV.5.2 Qualitative studies of profiles and messages in social networking sites 134
IV.5.3 Studies centered on the use of social networking sites 135
IV.5.4 Studies using mixed methods 136
IV.6 Conclusions 140
V Media consumption preferences and personal networks on Facebook 142
V.1 Research design 142
V.2 The network of Facebook profiles 145
V.3 An analysis of media consumption preferences manifested in the studied network 151
V.4 Media consumption preferences with respect to network structure 155
V.4.1 A short introduction to formal concept analysis 158
V.5 Discussion 165
VI The collapse of contexts and public criticism of subcultural values and esthetics 167
VI.1 Satire 168
VI.2 Research design 169
VI.3 Cocalari 170
VI.3 Pițipoance 177
VI.4 Discussion 182
VII Conclusions 186
VII.1 The convergence of computer mediated communication in social networking sites 186
VII.2 Media preferences and the structure of social networks in social networking sites 187
VII.3 The collapse of contexts and value conflicts 189
VII.4 Conclusions referring to methodology and originality of research 190
VII.5 Limits, advantages and possibilities for further research 191
VII.6 Final considerations 192
Bibliography 194

Keywords: computer mediated communication, Internet, Web 2.0, online social networking systems, Facebook, social network analysis, media consumption preferences, self-presentation

Introduction

This thesis is based on theoretical aspects and uses methodologies from several different fields and thus constitutes an ample interdisciplinary effort.

One of its main aims is to explore interdisciplinary methodologies in the study of social networking systems, applying concepts from several fields, belonging to both hard sciences and social sciences.

I. The study of social networks

The first chapter of the thesis makes an inventory of the development of sociological interest for the study of social structures both at macro and micro levels and then describes the development of methods associated with social network analysis starting with the first sociometrists like Jacob Moreno, to Robert K. Merton, Paul Lazarsfeld and to the contributions of Stuart Milgram and Mark Granovetter with respect to small world phenomena and the role of weak ties in social networks.

From various relationships and hierarchies instituted in the pantheons of major religions to genealogical descendence used as a way of legitimizing status quo, the organizing and documenting of relations has always played an important role in human society.

The structural approach of the study of inter-human relations became central to the interests of the fathers of modern sociology especially under the influence of evolutionism and the enthusiasm of Darwinist theories. One can understand the preoccupation of the early sociologists like Comte, Spencer, Tönnies and Durkheim and political economy theorists like Marx and Weber for the study of vast social changes and the growing complexity of Western societies in the context of evolutionist theories which offered an apparently universally applicable paradigm and, in some cases, could be easily molded over older conceptualizations of society as an organism. The positivist inclination of the first sociologists greatly influenced structural approaches of the study of inter-human relations. At the beginning of the 20th century, we find the first attempts at mathematically modeling phenomena bearing characteristics of complex

systems – the principle launched by Italian economist Vilfredo Pareto under the influence of forerunners like Marx and Spencer who were primarily interested in theorizing social inequality.

Sociometry, one of the most important contributions to the development of social network analysis as a method in the first half of the 20th century belongs to psychiatrist Jacob Moreno who puts the failure of many solutions to social problems appearing throughout time and the failure of many doctrines into the account of the lack of preoccupation for the systematic study of the structure of society. As a consequence, his research focused on collecting data on positive and negative options and what individuals know each-other, observing models of interactions that connect individuals, discussing kinship, examining social roles. His most important innovation is introducing the sociogram as a method of representing the formal properties of social configurations through the use of diagrams where individuals were represented as points and social relations as lines.

Linton C. Freeman argues that in 1938 Moreno, with the help of Jennings and Lazarsfeld had already touched the four defining aspects of social network analysis as we know it today:

- Social network analysis is motivated by a structural intuition based on connections between social actors
- It is based on systematically collected empirical data
- It uses graphical representation methods
- It is based on mathematical or computational models

Also, research done by Mayo and Warner at Harvard University towards the end of the 1920s came very close to contemporary standards in social network analysis, meeting three of the four criteria mentioned above.

The long collaboration between Robert K. Merton and Paul Lazarsfeld at Columbia University had a strong influence on American sociology, giving it a structuralist direction. The success of Merton's essay "Social structure and anomie", published in 1938, revised in 1949 and extended eight years later, considered by many as being the most cited article in the history of sociology, is owed to a structural-functionalist approach. In explaining anomie and deviant behavior, Merton focuses on social order, rather than on the individual.

Paul Lazarsfeld and Elihu Katz launched the two-step flow mass communication model, thus trying to introduce elements regarding the study of inter-human relations, the roles played by individuals in the groups they belong to and the influence of some over the others – all elements associated with the study of social networks, into a new theory of mass communication.

After World War II, Talcott Parsons supported structural functionalism, inspired by the advent of cybernetics, and his conception of systems as structures saw them as hierarchical and decomposable into modules, each component functioning in accordance to intrinsic principles and each being influenced by other components at well-defined points through various inputs.

Stanley Milgram contributed to theories regarding social networks by introducing the concept we know today as “six degrees of separation”. It refers to the fact that any two randomly chosen individuals in a very large population are separated by a chain of relatively few social connections. This is also called the “small world phenomenon” generated by the fact that society is actually a very dense network if we consider each individual’s social relations.

Mark Granovetter is another researcher whose contribution to social network analysis in the 1960s cannot be omitted. While working on a study on how people find jobs, Granovetter focused his research on the way in which individuals use their social network or their social relations to find jobs. Surprisingly, the results showed that weak ties (i.e. acquaintances) and no strong ties (i.e. friends) are more important in identifying job opportunities or getting a job.

Albert Barabasi argues that Granovetter describes a new model for the structure of society – a society structured in very closely connected clusters, or very tight groups of friends, groups where everybody knows everybody.

In the 1980s and the beginning of the 1990s, German sociologist Niklas Luhmann focused on integrating sociology with research done in cognitive sciences an cybernetics by Humberto Maturana and Francisco Varela, two researchers that had become very influent in the emerging field of complexity science. In this context, Luhmann argues that society can be conceptualized as a self-generated, complex, emergent, self-organized system of communication, structurally open, organizationally closed, dynamic, capable of adapting, continuously evolving and autonomous.

In the last part of the first chapter we briefly covered some of the key methodological aspects of social network analysis, described the main types of data, widely used collection methods, and a series of measures defined for the study of network structure.

II. Computer mediated communication

In the second chapter, in order to explain the current popularity of social networking systems and to contextualize the research on virtual social networks, we discuss a series of contributing factors that led to the development of large-scale computer mediated communication and influenced the characteristics of this type of communication. We refer to a series of crucial moments in the history of information technology to explain computer mediated communication in its early stages and subsequent developments: how x86 architecture became a standard, the economic and socio-cultural impact of the IBM PC, the visions of Microsoft founders – "a computer in every home" and "software separated from hardware", the shift of hardware companies from vertical integration to horizontal integration production models (with the notable exception of Apple).

We describe the beginnings of the Internet, insisting on its architectural conception, the features that differentiate it from other communication networks, and the development of protocols that determined the types of computer mediated communication. Beyond the technical aspect, the development of the Internet and the World Wide Web are discussed in an economic context, stressing the preconditions of Internet business in 1995 (the growing number of online users, the popularity of the World Wide Web, the development of graphical browsers and the changes in legislation that allowed commercial Internet traffic), the "Internet bubble" phenomenon that lead to the crash of the stock market prices for "dot com" companies in 2000 and the advent of the term Web 2.0.

A section of the second chapter deals with the theoretical perspectives applicable to the study of computer mediated communication and covers some of the scientific approaches explored in the 1980s and early 1990s by the pioneers of computer mediated communication and virtual networks research: Henry Jenkins, George Gilder, Nicholas Negroponte, Howard Rheingold, Mark Poster, Sherry Turkle. One of the central ideas in the works of these theoreticians is media convergence – i.e. the convergence of several media technologies into one – the computer

connected to the Internet, sending and receiving any information in digital format. We also discuss authors like Brian Winston, Carolyn Marvin, Lisa Gitelman, Geoffrey Pingree, James Carey, Stephen Kern or Armand Mattelart who tried to temper the exagerated enthusiasm surrounding new media and their effects on society and communication by showing, usually through mainly historical accounts of innovations in media technology throughout the last 200 years, that many of the changes we attribute to the Internet and digital media had already appeared by the end of the nineteenth century, with the spread of the telegraph – the first technology that separated communication from transportation, thus producing an unprecedented compression of time and space, opening the path towards instant global communication and constituting the first instance of what came to be called cyberspace.

In order to cover the development of virtual social networks we discuss the development of various forms of computer mediated communication and the ways in which they offered the technological medium for the creation of social networks on the Internet.

We draw the distinction between asynchronous (e-mail, bulletin board systems) and synchronous (MUD, IRC, IM) computer mediated communication outline their features with respect to their potential of sustaining the creation and development of virtual social networks. In chapter III, we look at social networking sites such as Facebook as a corollary of several types of computer mediated communication.

The wide spread of social networking services (Friendster, MySpace, LinkedIn, Facebook, and recently Google+, just to name a few of the best-known), researchers in the fields of both sociology and computer science have considered profiting from the opportunity to study explicitly declared social networks, connected in gigantic, even global structures, with millions tens of millions or even hundreds of millions of nodes. This kind of research involves wide range of new problems that these researchers, regardless of their field have to deal with – from ethical or legal issues involved by the technical possibility of gathering data about users without their explicit approval to methodological problems raised by the difficulty of studying the macro structure of such complex networks.

III. Computer mediated communication in Web 2.0

The third chapter discusses a series of communication technologies on the World Wide Web commonly associated to the term Web 2.0. We emphasize the similarity between the principles of this new paradigm and the initial ideas behind the Web proposed by Tim Berners-Lee in 1990. We briefly define and discuss: blogs, wikis, multimedia-sharing systems, tagging and social bookmarking, newsfeeds, podcasting, mash-ups, content management systems and finally social networking services.

We provided a detailed description of online social networking systems that have currently become a convergence point for most of the other computer mediated communication technologies, exhibiting diverse characteristics, functionalities and uses.

A growing number of studies focus on social networking systems in the Web 2.0 paradigm (Hi5, Facebook, MySpace) using specific research methods. Given the popularity of these virtual systems, both in Romania and globally, researching their structure and dynamics and especially their potential of becoming databases containing information usable by marketers, sociologists, media distributors to define groups based on media preferences, based on self-applied tags used in such systems, on a large scale.

If media are the extensions of man, as McLuhan theorized over 40 years ago, we are now confronted with a situation where profiles and avatars interact with each-other in virtual spaces making these extensions explicit as a dramatic presence of individuals in public spaces, seen by the Canadian visionary as a bridge over space and time created by human beings, turning them into veritable projections (in the truest mathematical sense) of human beings into planes defined by Web 2.0 tags.

IV. Social network sites

In the first part of chapter IV we discuss definitions and classifications of social networking sites as well as a series of classifications of users and uses for these sites.

Next, we described (mainly chronologically) the development of social network systems, focusing on the most popular ones or on those that we considered to be turning points in the configuration of such online systems:

- 1985 – WELL;
- 1994 – Geocities;
- 1995 – Classmates.com;
- 1995 – Match.com și site-urile de anunțuri matrimoniale (”dating sites”)
- 1997 – Sixdegrees.com;
- 1999 – Cyworld, AsianAvenue, BlackPlanet și MiGente;
- 2002 – Friendster;
- 2003 – MySpace;
- 2003 – LinkedIn;
- 2003 – Gaia Online;
- 2003 – hi5;
- 2004 – Orkut, Google Buzz, Google Friend Connect, Google +;
- 2004 – Facebook.

Key functionalities were discussed while trying to explain the broad or limited success of each of the systems and finally the global spread of the use of the Facebook system. To justify our choice of Facebook as the focus of our research, we described the various functionalities offered by the service during its 7 years of activity and its corollary role with respect to computer mediated communication.

The last part of chapter IV is dedicated to reviewing some approaches and methodologies used in the study of online social networking systems:

- Quantitative studies on automatically extracted large datasets,
- Qualitative studies of profiles and messages in social networking sites,
- Studies centered on the use of social networking sites,
- Studies using mixed methods.

Defining and classifying social networking sites in the context of the last few years when their use has become globally spread among teenagers and youths especially, has become an increasingly important preoccupation for the members of the scientific community. Researchers in both hard sciences and social sciences seem to drive towards research methodologies that take advantage of the large quantities of data made public on these websites.

A significant portion of the research on large datasets explore different strategies of automatically processing information extracted from the profiles defined by users of social networking sites with respect to the network structure configured between them.

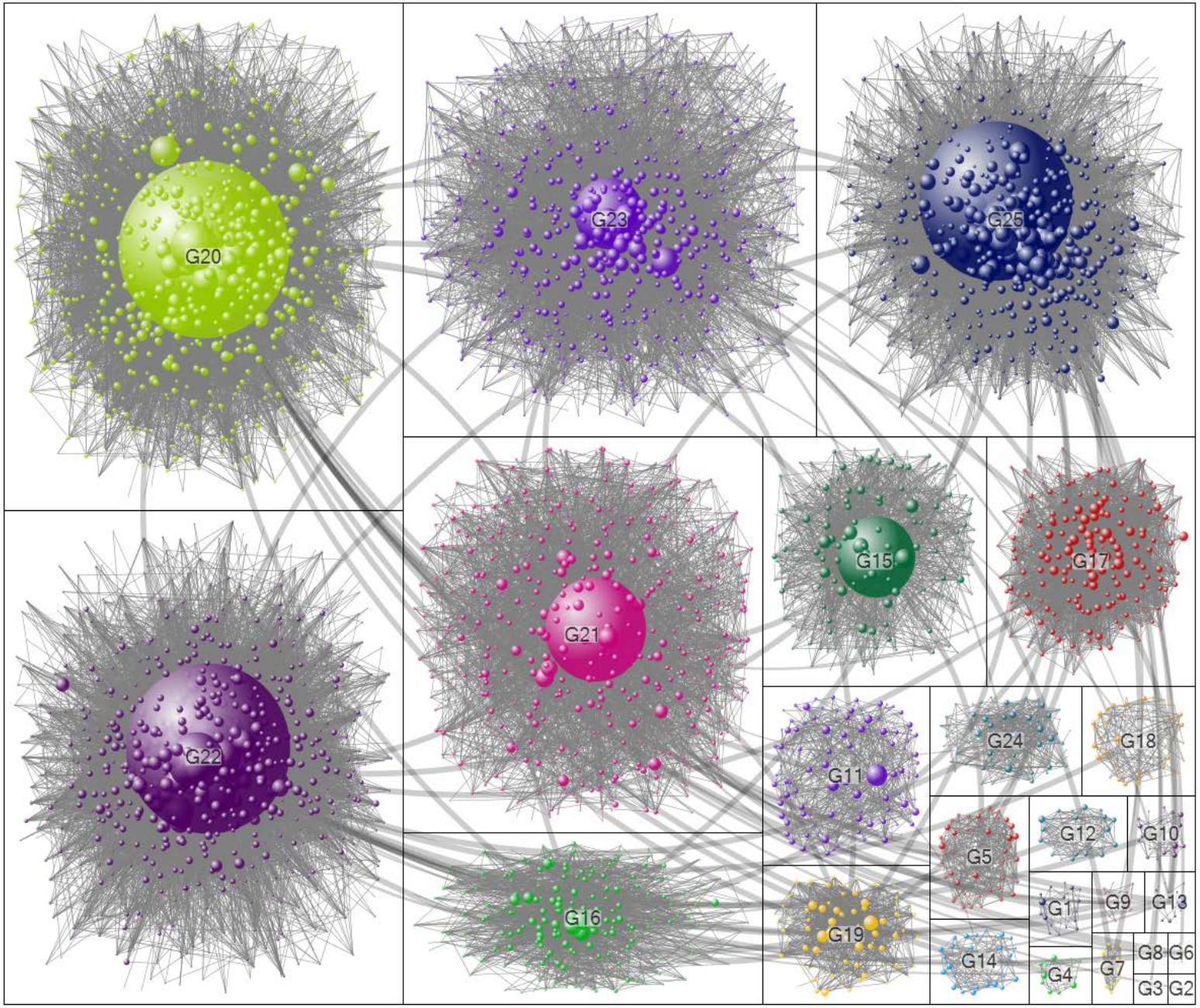
V. Media consumption preferences and personal networks on Facebook

Chapter V was configured around a study of online social networks in the popular Facebook system. We automatically collected data from user profile pages using specialized tools like NodeXL. This method allows the collection of egocentric networks associated to user accounts (with the account holder's permission). Some of the variables extracted for each profile are: sex, age, movie preferences, TV preferences, music preferences, book preferences and the list of friends. We analyzed information from over 4200 connected profiles.

We used social network analysis methods to identify structural clusters using relational data (the friendship link structure) as well as formal concept analysis, a method based on set theory, that allows the corroborated analysis of clusters and conceptual hierarchies of attributive data (both demographic and those that describe media preferences), thus revealing types of media consumption associated to demographic categories among the profiles identified as being most important (according to PageRank scores) in the network portion we studied.

Analyzing the community or cluster structure of the collected network portion, we discovered formations of structural clusters defined by large or small common social contexts (from cohorts of classmates, to smaller groups of relatives, childhood friends and co-participants to exclusive and well-defined (with respect to time and space) contexts – workshops, summer schools etc.). The large clusters observed in the network portion we studied indicate the existence of users who are very involved in this type of communication who link together several communities – these are nodes with high PageRank scores and are usually public communicators – bloggers or personal profiles built to represent local media institutions, thus breaking the Facebook rules. These users create a great number of connections, using the friendship link to build a network public that automatically becomes subscribed to all the messages they choose to make public.

Cazul 1 - Rețeaua de legături dintre profile structurată pe grupuri



Creat cu NodeXL

VI. The collapse of contexts and public criticism of subcultural values and esthetics

The sixth chapter presents a study of the satirical discourse constructed on certain sites based on photographs published on profile pages in popular online social networking sites. We analyze the self-presentation tactics associated to the subcultures that these sites name *piţipoance* and *cocalari*. Also, we discuss the construction of satirical discourse and the consolidation of stereotypes about these youth subcultures through the selection and taking out of the context of social networking sites of images exhibiting certain features.

We used qualitative analysis and the instruments of semiotics to study documents published on two websites: cocalari.com and pitzipoanca.org.

Popular online social networking sites facilitate explicit values conflict, allowing the meeting in the context of the same systems of people that come from different social contexts – the context collapse phenomenon. The egalitarian space of these systems themselves and enforcing the use of real identities within their boundaries controls aggressive practices, imposing a relatively civilized behavior. For this reason, social satire as a form of violence usually takes place in external contexts such as the blogs analyzed in this study, but also a great number of similar initiatives. Taking profile pictures out of contexts, separates the criticized cultural practices from the rest of the content in the profile, facilitating the dehumanization or rather the caricaturization of individuals, allowing for analogies to consolidated stereotypes.

VII. Conclusions

This thesis discussed the development of computer mediated communication in order to describe popular online social networking systems like Facebook and Google + as convergence points of several important tendencies in communication within virtual social networks.

The methodology used in the study of Facebook network portion combines social network analysis methods with data mining methods used to process the lists of media preferences, formal concept analysis as a method or hierarchical clustering of attributive data in Facebook profiles, and semiotic analysis. This type of methodological approach is innovative, having potential applications in sociology, and especially in market studies.

The interdisciplinary approach constitutes an important aspect of originality of this research. Furthermore, the use of recently developed and relatively easy-to-use (by social scientists)

software tools (like NodeXL, the Social Network Importer extension, the formal concept analysis extension or Google Refine) is an unique effort in social research being done in Romania at the moment.

With respect to the study of blog-type web-sites (pitzipoanca.org and cocalari.com) that use satire to criticize the self-presentation tactics of some users of popular social networking systems, the methodology used is not necessarily innovative, the topic of the research is a fairly new one for the social science community in Romania despite the considerable success of these websites.

On social network sites all relations, experiences and events are lived simultaneously, thus producing a ”personal present”, in the same way that, according to Marshall McLuhan, the electronic age – especially radio and television – have produced a ”global present”.

The macro level vision proposed by Castells can be used to explain the reconfiguration of personal or public space and time in these popular online social networking sites. Temporal sequences that defined experiences or events are super-imposed in a ”timeless present”, social contexts defined by different social interaction spaces (”at home”, ”at school”, ”at work”, ”going out”) are collapsing into each-other due to the fact that virtual environments defined by platforms like Facebook are always just a click away.

Understanding social networking sites as systems of communication, in the sense used by Niklas Luhmann, and defining all their elements, all profile elements, all posts (status updates, links, images) and threaded conversations, all relations between profiles or between profiles and other entities, al private messages both synchronous and asynchronous as instances of communication seems less counter-intuitive then understanding human society this same way. If we look at this form of computer mediated communication as an internal system that appeared in the environment of social communication in general by reducing its complexity (essentially defining simple types of relations), through the inherent boundaries between the online and ”offline” environments, the study of communication in these online systems (where instances of communication are persistent) can lead to a better understanding of society as a system of communication in the paradigm proposed by Luhmann, in a similar way that a simplified model of the solar system can lead to a better understanding of the interaction between planets.