Studying human social environment and state with social network data

Anton Kolonin
Institute of Cytology and Genetics SB RAS
Aigents Group
Novosibirsk, Russia
akolonin@gmail.com

***Abstract*—The paper describes approach, practical solution and intermediate results for self-driven exploration of personal interests structure and nearest social environment for an individual in order to correct their own behavioral patterns and interactions with communication partners on social networks on the Internet.**

Keywords—collective consciousness; mental state; personal awareness; social network

I. Introduction

Tight connectivity of humans in modern world-wide social networks is rapidly increasing and now is getting close to connectivity of human brain. Importance of the latter transition for evolution and humanity could not be underestimated because of amount of nodes and links in modern computer networks is exponentially growing with addition of artificial agents being involved in "hybrid" human-computer networks. It becomes critically important to understand phenomena of social interactions in such over-connected societies, so that effects of social behavior based on collective consciousness could be well understood, predictable and manageable – from perspective of humans exposed to modern information networks and from perspective of entire society.

One of the key factors of modern social communities across the world is high volatility and scale of social dynamics, with social clusters being created, updated, merged, split and disappeared rapidly, which leads to qualitative changes in principals of social organization [1]. The impact of social network communications on each individual interacting with Internet on daily basis is normally underestimated by average human. At the same time, it can be significant enough to alter state of consciousness of a single person or entire community [2]. That means, having ability to measure impact of digital social environment of a human or society instrumentally would be beneficial for a person or entire society, so they could adapt to the social environment and/or correct their course of action in its regard.

II. Approach and practical solution

We consider approach for building computational model encompassing knowledge acquired by entire society by means of evidence supplied by each of it members, called "social evidence-based knowledge representation" [3]. In this work we focus on the nearest social surrounding of a person interacting with social networks – in attempt to build personal mental model identifying contextual clusters of private interest as well as capture structure if inter-personal relations in the nearest environment, building collective consciousness model of the community that person is belonging to. This would potentially let person identify their state within their social environment and increase their personal awareness.

To achieve this goal, we build software application which extracts social network data for specific user, given their agreement and then extracts patterns of "natural classification" [4] from it and evaluate different sorts of social relationship within the social cluster the user belongs to.

My interests

Topics
brain, brings, special, technopark
kite, shovel
entire, siberia
museum
adaptive, agents, personal
data, google, graph, more, query, search, us
ibm, watson
springer

Interests of my friends

Topics	Friends
поздравляем	Jgohqunu Zkuhcajtytonre, Qcaxpiqu Reqeninu
вчора, лавров	Purjip Kuzinegiyt, Сункйянйын Елгюнсу
днем, рождения	Huhcunu Xazuheqegu, Xpihkic Ejimrin, Шряцыпыл Фещын
good	Moi Funv, Kiytukx Niyteqj
aigents, could, human, internet, system, used	Unxkof Jtoqq, Gquxijqug Jekerin, Purjip Kuzinegiyt, Erjunu Reguqogu, Pijtu Reqenin, Jhomtun Gquxipik Zavus, Voekv Gokhimkuyteg, Xpihkc Mohanin, Кюлчют Нюзюцуш, Ярюскянцл Кощщуйын, Uivonhj, Пыжяыр Кожялюш, Iqcu Nonujtog, Xonij Exinereg, Шысйул Чурысуш, Unujhujiu Zqibniar, Сункйянйын Елгюнсу, Шряцыпыл Фещын, Kiytukx Niyteqj, Gquxipik Bugiuqeg, Ytupuku Woknunxe, Кюлчют Вопксыт, Ынюккя Гырысыня, Kucpenx xo Quyubo, Xpihkc Ekqeg, Gquxipik Rikiqonre

Fig. 1. Using "natural classification" technology to unravel "clusters of interests" of given user (upper table) and "clusters of fiends' interests" (bottom table) with Aigents Report for social network.

We use Aigents software application and web service [5] in order to collect information on personal environment of a human in social network. Each user of the application can connect Aigents service to number of social networks (such as Facebook, VKontakte or Google+) and let it collect public data from it. Most of social networks are substantially restricted in regard to scope of information that can be collected to endorsed application programmatically, so in most

cases only shallow layer of public personal profile can be used – gathering posts from the personal pages, comments on the posts and events that the posts and comments are being liked by other users.

For each user, given corpus of posts and comments made by them and liked by them and their friends or not, along with posts or comments made their friends and liked by user and friends or not, we can extract word feature vectors from the texts and perform clustering analysis [4] discovering categories of "natural classification" across the interests of the user themselves and across their friends associated with the texts of posts and comments.

Also, the similarity metric between each user and their friends can been calculated using feature vectors extracted from user texts and comments [6] with normalized overlap between the two vectors evaluated as mutual similarity measure between users.

Moreover, we can also evaluate number of relationships between the user and their friends, assuming L_{ij} can be used to denote number of "likes" user j gives to posts/comments made by user i and C_{ij} indicating number of comments user j makes in regard to posts/comments made by user i. For instance, "Best friends" metric can be used to indicate value of mutual likes between two users $L_{ij}*(L_{ji},+C_{ji})$, normalized my maximum number for given friend j of user i across all J users as follows.

$$B_j = L_{ij}*(L_{ji}+C_{ji}) / Max_{j=1,J} (L_{ij}*(L_{ji}+C_{ji}))$$

Further, "Fans" metric can be evaluated as amount of likes and comments the other user gives to posts of user of interest denominated by returned likes and comments, so that complete fan is one that pays more attention to given user while the latter one pays the least amount of attention to the fan.

$$F_j = (L_{ji}+C_{ji})/(1 + L_{ij}) / \mathrm{Max}_{j=1,J} ((L_{ji}+C_{ji})/(1 + L_{ij}))$$

Similar to me

Rank,%	Friend	Crosses	My likes	Likes	Comments	Words
100	Uivonhj	299	1	14	15	social information now has mo
37	Kiytukx Niyteqj	35	2	0	4	new anton good all thanks interne
16	Iqcu Nonujtog	22	2	0	5	можно новости поиска сами новы
14	Кюлчют Волксыт	15	1	3	1	learning cognitive human
14	Pijtu Reqenin	14	3	10	8	about us out time down aigent like hos
10	Кюлчют Нюзюцуш	13	1	0	2	агентов автоматическое например объе

Fig. 2. Displaying mutual similarity between social network user and their friends with histogram of most commonly used terms per each of friends.

Finally, the metric opposed to "Fans" is "Authorities" which corresponds to amount of attention (number of likes) paid by user of interest to other one denominated by amount of attention (likes and comments) returned by them.

$$A_j = L_{ij}/(1+L_{ji}+C_{ji}) / \mathrm{Max}_{j=1,J} (L_{ij}/(1+L_{ji}+C_{ji}))$$

In the following results, examples are given for personal account of author in one of the social networks having names of author's friends changed to meaningless labels for the purpose of privacy protection of these users.

III. Discussion of Results

For studied users, we have discovered patterns not obvious for conventional interactions with social networks, which has turned useful for a person to re-evaluate their state in social environment and possibly adjust their mental state and communication pattern.

Best friends

Rank,%	Friend	My likes	Likes	Comments
100	Gquxipik Rikiqonre	6	23	8
38	Qcaxpiqu Reqeninu	2	28	7
29	Pijtu Reqenin	3	10	8
24	Шысйул Чурысуш	2	18	4
16	Uivonhj	1	14	15
11	Xonij Exinereg	1	20	1
10	Unujhujiu Zqibniar	3	3	3

Fans

Rank,%	Friend	My likes	Likes	Comments
100	Ukhop Gihuq'ogiyt	0	23	0
83	Ytkijhewok Zaenhopmi	0	17	2
63	Uivonhj	1	14	15
52	Янцлют Йуркйыжын	0	12	0
51	Qcaxpiqu Reqeninu	2	28	7
46	Xonij Exinereg	1	20	1
43	Caqiu Xuniqegu	0	8	2
37	Pirtuiq Nireqjric	1	15	2
35	Eqov Zujreg	0	8	0
35	Eqvu Guqiogu	0	8	0
35	Uqodunxok Xazcnin	0	7	1

Authorities

Rank,%	Friend	My likes	Likes	Comments
100	Шряцыпыл Фещын	2	0	2
75	Xpihkc Mohanin	1	0	1
75	Ынюккя Гырысыня	2	1	2
64	Unujhujiu Zqibniar	3	3	3
60	Jhomtun Gquxipik Zavus	2	2	2
60	Kiytukx Niyteqj	2	0	4
50	Gquxipik Bugiuqeg	1	1	1
50	Iqcu Nonujtog	2	0	5
50	Кюлчют Нюзюцуш	1	0	2
50	Xpihkic Ejimrin	1	1	1

Fig. 3. Ranking friends by different metrics such as "Best friends", "Fans" and "Authorities" based on number of mutual "likes" ad cross-comments.

For instance, for given user account, for the year of study, five clusters of personal user interests have been discovered and confirmed as actually present (Fig.1): "brain, brings, special, technopark" (specific to brain sciences and technology), "kite, shovel" (specific to outdoor sports), "entire, siberia" (local geographical subject domain), "museum" (local museums), "adaptive, agents, personal" (major subject of the author), "data, google, graph, more, query, search, result" (most of Internet search technology discussions), "ibm, watson" (threads on respective technology from IBM) and "springer" (for Springer publications).

Respectively, the clusters of friends on Fig 1. split into few top groups of friends exchanging greetings and congratulations and the very last cluster (“aigents, could, human, internet, system, used”) to discuss major matters of the domain shared with author.

The top similar user friends discovered are shown on Fig.2 with topmost 100% overlap for primary professional English contact regarding “social” and “information” terms, then less formal positive communicating peer on Internet technology advances (“new, good, thanks, internet”) and similar Russian contact (“можно, новости, поиска”). Next two are contacts interacting regarding use of Aigents software for automatic time-critical information processing in English (“about us, time, aigent” and in Russian (“агентов, автоматическое”).

Rankings of relationships shown on Fig.3 gave unsurprising results with two closest relatives listed at positions 2 and 3 of the “Best friends” list. However it was more surprising to find a not very familiar person at row 1 of this sheet to make sense of tightness of mutually pleasant communications with the person. The “Fans” sheet gave the most surprise having second row occupied by complete stranger even not having direct friend relationship with account user, which has been considered as a suggestion to establish actual friend relationship. The “Authorities” ranking seemed not very salient at first glance but further consideration suggested that people at the top are either important persons or far-away friends, so for both categories attention to their comments to account user posts has been expressed with unconditional “likes”.

IV. Conclusion And Future Work

Described approach is considered practical for people who desire to increase their level of self-awareness while interacting with social networks and can help them to improve their social performance and level of comfort.

The described functionality is being made available as web service at https://aigents.com/. Future support for Android application “Aigents” for smartphones and tablets and standalone applications for Windows, Linux and Mac OS/X servers and desktops is expected in the future.

In the future work, we anticipate to extend depth and breadth of analysis to cover emotional and intentional aspects of social interactions of social network user as well as study dynamics of contextual, social, emotional and intentional changes over time.

References

[1] V. Makarov, A. Bakhtizin, “The New Form of Mixed Economy with Rationing: Agent - Based Approach”, Advances in Systems Science and Applications (2014) Vol.14 No.3. pp.278-285.

[2] L. Muchnik et al., “Social Influence Bias: A Randomized Experiment”, Science 341, 647, 2013, DOI: 10.1126/science.1240466.

[3] A. Kolonin, “Computable cognitive model based on social evidence and restricted by resources", 2015 SIBIRCON/SibMedInfo Conference Proceedings, ISBN 987-1-4673-9109-2, pp.41-45.

[4] E. Vityaev, “Unified formalization of «natural» classification, «natural» concepts, and consciousness as integrated information by Giulio Tononi”, The Sixth international conference on Biologically Inspired Cognitive Architectures (BICA 2015, November 6-8, Lyon, France), Procedia Computer Science, v.71, Elsevier, 2015. pp 169-177.

[5] A. Kolonin, “Aigents: Adaptive Personal Agents for Social Intelligence”, Knowledge Engineering and Semantic Web, 6th International Conference Proceedings, ISSN 1865-0929, Moscow, Russia, 2015, pp.283-290.

[6] A. Kolonin, "Automatic text classification and property extraction", 2015 SIBIRCON/SibMedInfo Conference Proceedings, ISBN 987-1-4673-9109-2, pp.27-31