Using Natural Language Processing in an Instant Messaging Environment for User Analysis

Lukas Stasytis
Faculty of Informatics
Kaunas Technology University
Student g. 50, Kaunas 51368
Email: lukas.stasytis@ktu.edu

Abstract—**With recent advances in machine learning technology and a resurgence of Instant Messaging (IM) software, a possibility to incorporate natural language processing (NLP) solutions into IM servers for user personality profiling and monitoring has presented itself. This paper presents a novel use-case for NLP in a rapidly expanding data-generating environment - instant messaging application servers to gauge emotional profiles of internet users over time and to appropriately respond without the need for any human interaction from the side of the monitor. IBM Watson's Personality Insights API is looked at as a case-study NLP system for analyzing user data and the IM software Discord as a user-data-generating and user-monitoring environment using a 300.000 message sample. Results show clear and consistent differences in user personality profiles, suggesting that the IM space is a promising environment for further user analytics based on NLP.**

February 25, 2018

I. Introduction

Over the last few years, natural language processing techniques have greatly improved and expanded [2] [3] to the point where robust engines are now being offered as a service by companies like IBM [6]. These cloud-based engines provide the ability to implement NLP-based solutions for many use-cases without the need for training and prior-datasets. Additionally, a field that has rapidly risen in the recent years is instant messaging applications - an environment dominated by natural language. Thus, the use of NLP engines in an instant messaging environment becomes a natural next step. This could lead to a number of use-cases, such as emotional monitoring for suicide prevention, community-wide emotional range assessment and more. This paper overviews the effectiveness of instant messaging applications as a user data environment, the primary method for extracting said user data by using bots, proposes a methodology for extracting and analyzing the data, furthermore, provides potential use-cases for the application and looks at a case-study done on a 3.000-member server of a popular instant messaging application - Discord, using a dataset of 300.000 messages collected over a months time. Lastly, given the proposed use-cases and results of the case-study, conclusions are drawn.

Copyright held by the author(s).

II. User data environment

In recent years, as shown in the INTSIGHTS report [8], internet messaging applications have made a significant resurgence, with skype having 300 million users, WhatsApp - 1.2 billion, ICQ - 11 million and Discord at 45 million, at the time of the report. Slack has also announced last year to having reached 6 million daily uses on their application [13].

Looking closer at Discord, since the aforementioned report, a more recent one published in Dec, 2017 by the Discord team themselves mentions a figure of 90 million users and a growth rate of 1.5 million users / week [1]. Showing an upwards trajectory in user engagement in IM services. Additionally, the current culture of topic-based IM servers, such as fandoms, gaming communities, corporation servers can all provide highly diverse environments, encompassing people of varying personalities and age groups.

There is great potential for data analytics, given abundance of user messages that can be used as primary datasets. Each message written by a user contains, usually, the message authors identification number, the date of the message, the specific channel or room within the server where the message was written and the message itself. This much data is enough to start grouping messages based on servers, channels, users and time to create personalized and group-wide statistics, such as personality and emotional assertion.

Given that the environment for data is there, the next step is extraction. Which leads to bots.

III. Bots

A key component of current IM applications like Discord and Slack, as well as IRCs, is the inclusion of bots. Bots are non-human users within IM servers that are running client scripts while interacting with the IM server application interface (API). These bots are highly modular, require no additional fees to be in the servers and are, above all, user friendly and universal. Using Discord as a case-study, a simple bot can be started by just installing the appropriate

CEUR-WS.org/Vol-2145/p25.pdf

programming language library (of which there are options for every popular programming language), registering the bot client identification number on the Discord official websites developer interface and finally using the provided bot token to communicate with the Discord API. The bot can then be added by a server owner using a simple invitation link and the developer simply has to run the script to have the bot connect to all the servers it has been added to. At this point, a bot can parse every message written in the server it is in and depending on the permissions given by the server administrators respond appropriately. Generally, these bots are used for simple services, like server administration using the chat interface, acting as chatterbots or allowing to interface with the world wide web using API calls and just the server text interface from the users end. These bots, while seeming like a novelty, are highly popular, which can be seen by looking at some of the more popular bots official pages and the number of servers they are present in. An example being the music bot Himebot, currently running on over 66.000 different Discord servers [12].

The most popular bots are following hundreds of thousands of servers at once, which can house thousands of users each, each writing messages that can be used as datasets for analysis. The case-study that will be explored further in this paper explores the use of a bot on a single Discord server with 3.000 users that generated 300.000 messages in just under a month.

An important consideration is the legality of using these bots for logging messages. Using the Discord example, their current terms of service state as following: Developers: ”Developers using our SDK or API will have access to their end users information, including message content, message metadata, and voice metadata. Developers must use such information only to provide the SDK/API functionality within their applications and/or services.” [5] This falls in line with the proposal of using the bots for user profiling, given the option for users to view their own profiles and keeping the process transparent. Additionally, according to Section 2.4 of the Developer Terms of Service, the bot functionality would have to be limited to non-commercial, non-advertisement-based use-cases, which fall in line with the proposals presented later in this paper.

Lastly, given the bots modularity, the bots can act as simple interfaces between IM applications and analytics servers, allowing for bots written in many programming languages, housed in many different servers across a wide array of IMs and IRCs to all interface with the same central server that would analyze all user messages, allowing for vast scalability of the service, whilst being easily incorporated into current bots as a new feature.

IV. User message analysis model

With the increasingly growing userbase for IM services, the potential for Big Data analytics projects develops. Abundant public messages that are exchanged every day on these services can be used in natural language processing-based solutions for both training and inference of machine learning implementations. The developed models can be used to provide powerful user emotional and personality trait-based modelling. A proposed methodology for user message analysis is to:

(a) Group the messages sent on the monitored servers based on time, servers, channels or rooms within those servers and message authors.
(b) Input the group datasets to NLP tools, such as IBM Watsons Personality insights API to obtain personality insight tags for the datasets.
(c) Use a different, unused metric in step 1 to plot the tagged messages and analyze personality trends.
(d) Look for abnormalities within the plotted datasets or noteworthy trends and react appropriately by either contacting the users in question directly via the bot or by sending warning signals to social services that might be able to respond appropriately themselves.
(e) Provide open access to the generated datasets for users within their own message-scope or wider if the message authors comply to fall in line with all terms of service for the applications and provide transparency for the users.

This proposed model allows to look for trends in user or community personalities and emotions in a time space or across different environments (servers and channels). Thus, different applications can be created, making use of the varying combinations of results achievable.

Much work has already been done in the field of social network analysis, some with practical case-studies that proved highly encouraging results [10] [7] [11]. IM applications can be a potential new input environment to add to the field of research.

V. Use-cases

The proposed user analysis model, given data reliability for analysis engine inference, can be a promising tool for user personality profiling, monitoring and implementation of autonomous systems that respond to shifts in user personalities and emotional ranges. Some of the proposed use-cases for the model are:

(a) Emotional monitoring for early spotting of depression, stress and suicidal tendencies and automated responses using the same bot system.
(b) General community personality analysis for detection of groups of individuals falling into set personality groups that a corporate or individual entity might be seeking.
(c) User retention in a text message communication environment analysis.
(d) User personality assessment without the need for personalized testing in the form of quizzes, allowing for a more unbiased result.

VI. Results

For the purpose of testing the capabilities of current NLP systems in an IM environment, a case-study Discord server, which houses 3.000 users was monitored from January 25th, 2018 to February 25th of the same year. Over the course of the month, a bot was tracking every message written within the server and saving it in a logging file alongside extra message metadata.

We start by extracting every message of every user which was posted within the server and saving the message contents, author's identification tag, time and the name of the channel in which the message was written to a logging file for further processing.

In our case study, after obtaining sufficient amounts of user information, the messages are grouped up based on authors and a few of the most active authors are taken and their messages grouped by date - a parameter not used in the initial message grouping. For this example, we will be taking 2 of the most active members on the server and the server average. Comparing their Big 5 personality assessment and 3 'needs' as output by the IBM NLP engine - Watson's Personality Insights.

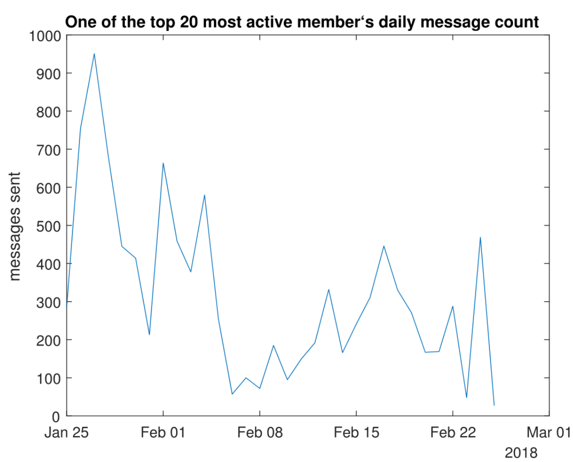

Fig. 1. The message count of one of the most active users of the server for whom we will use a nickname "Bob". This shows the amount of data possible to generate from a single user in an IM environment.

For this particular case-study, we take 2 of the most active 50 members on the server, each providing, on average, a few hundred messages per day and sending these grouped messages to IBMs Personality Insights engine - Watson. We also look at the server's average results.

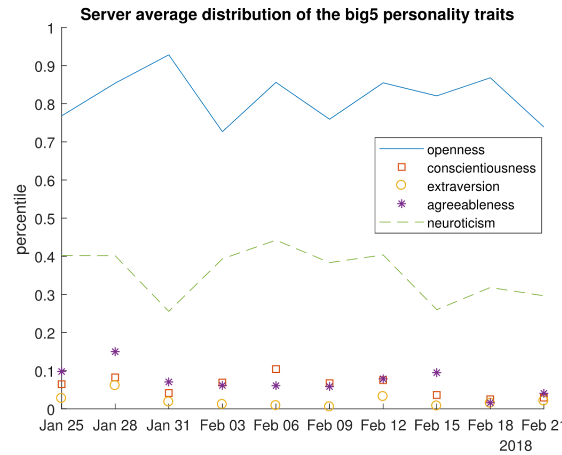

Fig. 2. The server's average distribution of the big 5 personality index in percentiles

A trend in user personality traits analyzed by Watson emerges. A high degree of openness and neuroticism is present in all users, whilst traits like agreeableness, conscientiousness and extraversion are vastly underrepresented.

Next we take our two example users and compare their results in relation to the server average. This allows us to find users that are anomalous within the target group and prevent biased results that might emerge from the format in which the messages are sent. Figures 3 and 4 represent these users.

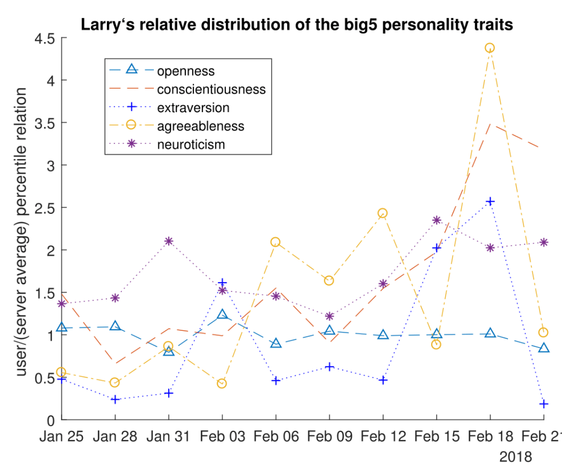

Fig. 3. Larry's Big 5 personality insight output

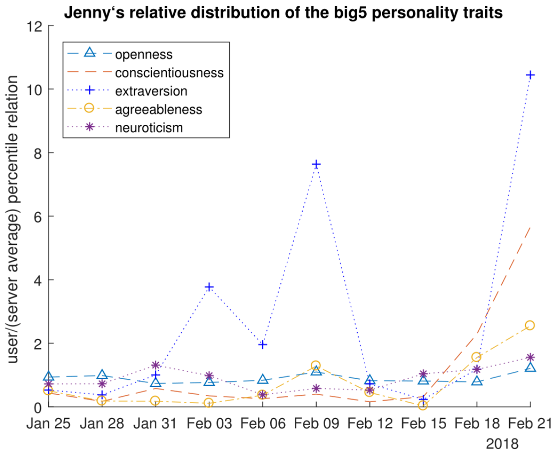

Fig. 4. Jenny's Big 5 personality insight output

A focal point is the fact that, whilst two users can have considerably different representations of the common five traits, the average over the course of the month remains relatively the same with specific spikes happening that are the precise type of anomaly that we would be looking for with such a system. This can, however, be a product of overfitted datasets. Careful message analysis would be required to conclude how accurate Watson's estimation really is in this environment.

Lastly, we look at some of the "needs" of the same users presented by the Watson engine. For the sake of focusing on potential cases of depression, we will focus on only these 3 parameters - 'love', 'closeness' and 'stability'.

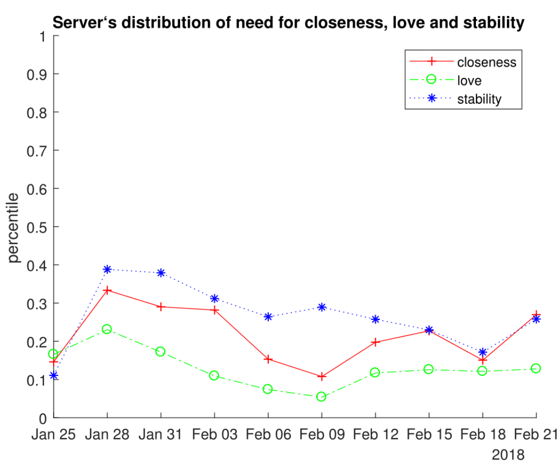

Fig. 5. Server's average closeness, love and stability time graphs

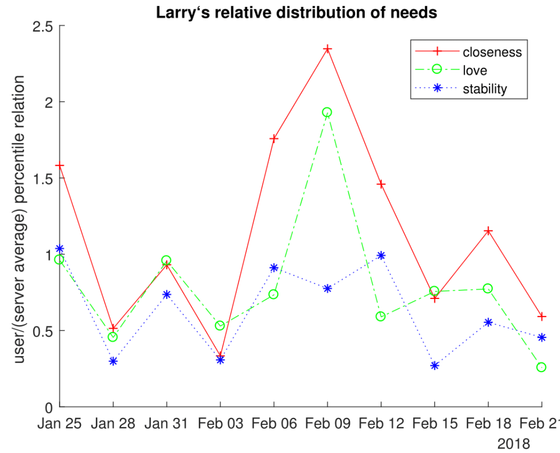

Fig. 6. Jenny's closeness, love and stability time graphs

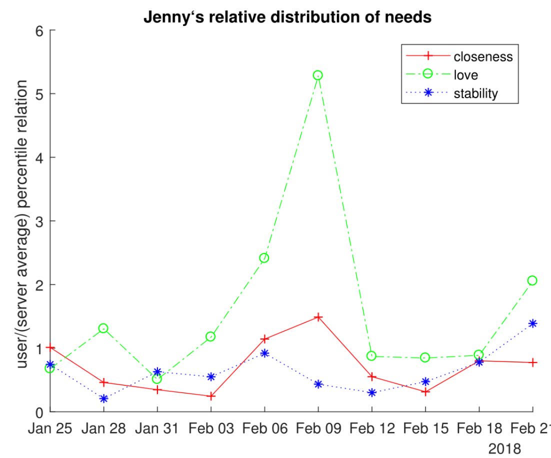

Fig. 7. Lenny's closeness, love and stability time graphs

The results are a lot less spread out than with the Big 5 analysis, however we can still see outliers once each user's analysis is presented relative to the average. Most notably - the spike around February 9th, where the need for love spiked for both users, as far as 6 times the norm for Jenny, and the increase in need for closeness from Larry's side.

The results all have a common trend of individual users having consistently diverse personality assessments by Watson, while maintaining a level of variation over time, which can closely match their mood and personality changes over time. This aligns with recent studies which have been done, showing that there is valuable insight to be gained from analyzing and modelling emotions, sentiments and personality traits. [9]

Lastly, at the end of the manuscript, Table 1 shows the entirety of the server's average needs output by the Watson engine. Some additional trends that we can find are that the need for curiosity and ideal are much more represented than the rest and the percentile of each need remains consistant.

TABLE I
THE SERVER'S ENTIRE 'NEEDS' OUTPUT FROM THE WATSON ENGINE IN PERCENTILES

date	closeness	curiosity	excitement	harmony	ideal	liberty	love	practicality	self-expression	stability	structure
1/25/2018	0.15	0.60	0.15	0.14	0.46	0.25	0.17	0.46	0.28	0.11	0.27
1/28/2018	0.33	0.72	0.28	0.23	0.64	0.39	0.23	0.43	0.35	0.39	0.15
1/31/2018	0.29	0.72	0.07	0.29	0.40	0.18	0.17	0.27	0.30	0.38	0.26
2/3/2018	0.28	0.69	0.34	0.19	0.58	0.39	0.11	0.45	0.25	0.31	0.21
2/6/2018	0.15	0.72	0.17	0.20	0.47	0.28	0.07	0.39	0.24	0.26	0.28
2/9/2018	0.11	0.73	0.26	0.16	0.56	0.36	0.05	0.49	0.28	0.29	0.36
2/12/2018	0.20	0.70	0.19	0.20	0.50	0.26	0.12	0.34	0.28	0.26	0.27
2/15/2018	0.23	0.74	0.16	0.19	0.48	0.29	0.13	0.32	0.33	0.23	0.16
2/18/2018	0.15	0.83	0.24	0.18	0.50	0.37	0.12	0.49	0.35	0.17	0.23
2/21/2018	0.27	0.65	0.25	0.21	0.48	0.32	0.13	0.37	0.29	0.26	0.13
Average	0.22	0.71	0.21	0.20	0.51	0.31	0.13	0.40	0.29	0.27	0.23

VII. CONCLUSION

In this paper, the motivation for implementing IM server tracking bots for user personality and emotional analysis was overviewed and a methodology for implementing such a system looked at. Additionally, a case study was explored of using the proposed methodology on a 3.000 user server over a period of one month and example datasets of some of the users presented while using the IBM Watson Personality Insights engine for the inference of the messages.

The results show that active members on IM servers can provide plenty of data to extract and analyze, with over 300.000 messages being generated per month on the case study server and at least 100 messages being sent by active users every day. Furthermore, there is a clear and even distribution in the average engine output, with a very high representation of openness (80th percentile), a moderate level of concientiousness (30th to 40th percentile) and a very low representation of extraversion, agreebleness and neuroticism (all under the 10th percentile). The same underrepresentation could be noticed with the 'need for consideration, love and stability' parameters.

Additionally, relative graphs of users, compared to the server can be used to find anomalous behaviour. The results, showing clear distinctions between individual users, suggest the engine having an analytical accuracy in the IM space worthy of further consideration and scientific analysis, giving weight to the proposed use-cases in section V.

For future work, using the generated emotional range results as parameters for a recurrent neural networks, an automated system to predict future changes in user emotional range will be looked at.

ACKNOWLEDGEMENTS

I would like to express my deepest thanks to my mentor Kazimieras Bagdonas for suggestions in implementing the system and providing guidance in the writing of the paper itself.

REFERENCES

[1] J. Alexander. As discord nears 100 million users, safety concerns are heard. Technical report, 2017.

[2] Y. Bengio, R. Ducharme, P. Vincent, and C. Jauvin. A neural probabilistic language model. *Journal of machine learning research*, 3(Feb):1137–1155, 2003.

[3] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa. Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, 12(Aug):2493–2537, 2011.

[4] R. Damaševičius, C. Napoli, T. Sidekerskienė, and M. Woźniak. Imf mode demixing in emd for jitter analysis. *Journal of Computational Science*, 22:240–252, 2017.

[5] Discord. *Section 2.4 - End User Data.*, 2017.

[6] R. High. The era of cognitive systems: An inside look at ibm watson and how it works. *IBM Corporation, Redbooks*, 2012.

[7] R. Lambiotte and M. Kosinski. Tracking the digital footprints of personality. *Proceedings of the IEEE*, 102(12):1934–1939, 2014.

[8] I. C. I. Ltd. Threat report: Messaging applications: The new dark web. Technical report, 2017.

[9] M. Mostafa, T. Crick, A. C. Calderon, and G. Oatley. Incorporating emotion and personality-based analysis in user-centered modelling. In *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pages 383–389. Springer, 2016.

[10] G. Oatley and T. Crick. Measuring uk crime gangs: a social network problem. *Social Network Analysis and Mining*, 5(1):33, 2015.

[11] G. Oatley, T. Crick, and M. Mostafa. Digital footprints: envisaging and analysing online behaviour. In *Proceedings of AISB Symposium*, 2015.

[12] H. official website.

[13] Slack. Where work happens around the world. Technical report, 2017.

[14] A. Venckauskas, A. Karpavicius, R. Damaševičius, R. Marcinkevičius, J. Kapočiūte-Dzikienė, and C. Napoli. Open class authorship attribution of lithuanian internet comments using one-class classifier. In *Computer Science and Information Systems (FedCSIS), 2017 Federated Conference on*, pages 373–382. IEEE, 2017.

[15] M. Wróbel, J. T. Starczewski, and C. Napoli. Handwriting recognition with extraction of letter fragments. In *International Conference on Artificial Intelligence and Soft Computing*, pages 183–192. Springer, 2017.