Group-based Communication in WhatsApp

Michael Seufert*,Tobias Hoßfeld†, Anika Schwind*, Valentin Burger*, Phuoc Tran-Gia*
*Institute of Computer Science, University of Würzburg, Würzburg, Germany
{seufert | anika.schwind | valentin.burger | trangia}@informatik.uni-wuerzburg.de
†Chair of Modeling of Adaptive Systems, University of Duisburg-Essen, Essen, Germany
tobias.hossfeld@uni-due.de

***Abstract*—WhatsApp is a very popular mobile messaging application, which dominates todays mobile communication. Especially the feature of group chats contributes to its success and changes the way people communicate. The group-based communication paradigm is investigated in this work, particularly focusing on the usage of WhatsApp, communication in group chats, and implications on mobile network traffic.**

***Index Terms*—Group-based communication, WhatsApp, mobile messaging application**

I. Introduction

As the Internet has become omnipresent, nowadays a complex interplay of Internet technology and human behavior can be observed. On the one hand, the Internet is changing our daily lives, the society as well as industry and business. On the other hand, the Internet technology is driven by the adoption of the end users and stakeholders in the ecosystem.

In particular, we have a closer look at the de facto communication applications in the Internet. We notice that the applications changed the usage behavior of users. YouTube, being a prominent example of a video streaming service, made it possible to upload and stream user-generated videos, which lead to an unprecedented increase of global Internet traffic. Another change of user behavior can be seen with mobile messaging applications recently. With these applications, such as WhatsApp, users are now communicating asynchronously in groups, which are created spontaneously or which exist over a longer period. The users activity is triggered by events in these groups, i.e., posted messages, which can be enriched by user-generated images and videos. Thus, as users are always online and interacting due to smartphones and network connectivity everywhere, their activity patterns are changing.

Then again, the emerging user behavior may also change the underlying Internet technology. This happened in the past and may also happen in the future. Coming back to the same examples mentioned above, we observed the need for content delivery networks (CDNs) due to the increased video demand on YouTube. CDNs allowed to place and cache the content closer to the users and to take into account regional or social interests. This new Internet technology subsequently spread, and now many different types of applications rely on CDNs. Similarly, the changing usage behavior of group-based communication in WhatsApp might have possible disruptive implications for the future Internet.

ISBN 978-3-901882-83-8 © 2016 IFIP

The mobile messaging applications establish a publish-subscribe paradigm on application layer, which may be efficiently implemented on the network layer. Moreover, the exchange of user-generated content in groups fosters caching approaches close to the edge and the social groups. Thus, research proposals like information-centric networking (ICN) could introduce benefits. However, the increasing privacy awareness of users might lead to encrypted data communication hindering network management for ISPs. It might not yet be obvious, which technology will be employed to cope with the new challenges and demands. Nevertheless, it is the user behavior that dictates the path of technology through service acceptance and adoption.

In this paper, we show researchers how group-based communication changes the activity patterns of multiple users. This should be taken into account, e.g., when evaluating communication technologies. Models from the past (e.g., Poisson arrival process of end-to-end voice calls) cannot be directly used for nowadays applications (e.g., WhatsApp messages are exchanged in groups of users) and need to be adapted to integrate interaction of users on a smaller time scale. Therefore, we present measurement results of WhatsApp and the group communication behavior and discuss possible implications of this emerging communication paradigm on networking technology.

Section II describes the evolution of communication paradigms from one-to-one communication towards group-based communication. Section III introduces WhatsApp, the most popular application for group communication. The implications of the usage of WhatsApp on user activity and the network traffic are discussed in Section IV, and Section V concludes.

II. Evolution of Communication Paradigms

During the last two decades, the evolution of technology and especially of the Internet changed our message telecommunication. Figure 1 shows this evolution by presenting some important messaging services in chronological order of their release. It furthermore assigns them to three different ways of communication: one-to-one, one-to-many and group communication.

The 1990s were characterized by private, immediate exchange of messages between two equal partners of communications. This way of communication is called one-to-one communication. In the middle of the 1990s, for the first time

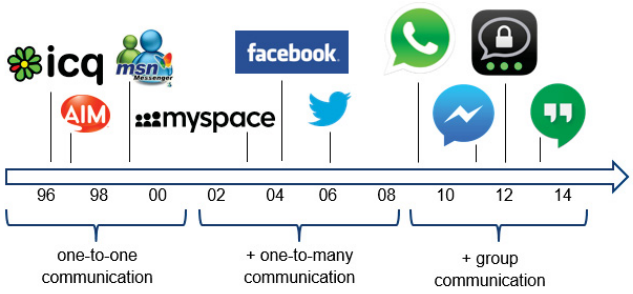


Fig. 1. Evolution of communication paradigms and notable messaging services

the concept of personal instant messaging over the Internet became available in the form of so called instant messengers. Instant messengers are chat programs, which allow users to instantly exchange messages of short size with each other. In that, they are similar to the Short Message Service (SMS), which already became available to mobile phone users in the early 1990s.

With the release of the instant messenger ICQ[1] in 1996, an alternative to the traditional email communication became popular on personal computers. In comparison to emails, which are mostly written formal, containing long and fully formulated sentences and usually have a longer response time, instant messengers are made to communicate with others similar to a real life oral conversation. Here, the content is informal with mainly short sentences. Also, a quick response is expected, so that a fluent conversation results [1]. Other well-known instant messengers are, for example, Microsoft's discontinued MSN Messenger and AIM[2], the instant messenger of AOL.

After the turn of the millennium, a new trend in online communication emerged. For the next couple of years, one-to-many communication became more and more popular. One-to-many communication is a way of broadcasting messages to a receiving group of users. Messages are simply published and can be read by everyone or by a restricted set of users, which were given rights to read the contents, or which subscribed to the feed. The messages are not necessarily answered by the recipients directly, as the main idea is to make messages available to many people at the same time. Nevertheless, it is possible to reply to messages by broadcasting an answer, so the communication is not unidirectional. This way of communication is mainly used in online social networks. These networks also support one-to-one communication, but the main usage is one-to-many communication.

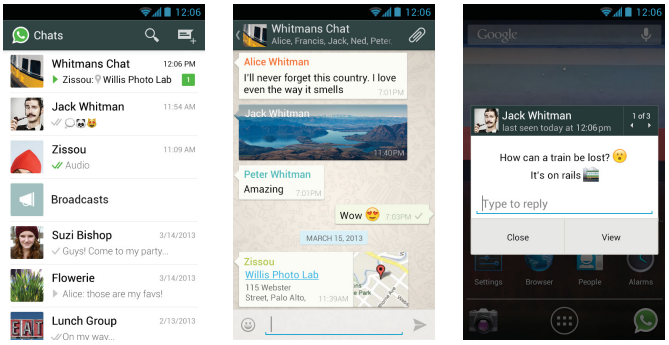
In 2003 and 2004, two of the most popular online social networks, Myspace[3] and Facebook[4], were launched. This development increased the desire of users to publish messages within their social environment. A prime example for this kind of communication is also Twitter[5], which is an online service started in 2006. Twitter allows users to broadcast short messages with up to 140 characters to their so called followers, which are passive recipients.

Since the middle of the 2000s, online communication has been supplemented by group communication. In this context, group communication means a conversation of a fixed group of users, which can equally participate. This process started in 2007, when the first iPhone was introduced and changed mobile communication significantly. Since this time, smartphones have become more and more popular, whereas SMS has been pushed into the background by the increasing usage of Mobile Messaging Applications (MMA). These applications are a form of instant messengers for smartphones. In contrast to emails and online social networks, MMAs are designed to allow immediate responses in real time similar to instant messengers. Additionally, communication in MMAs is not limited to one-to-one or one-to-many conversations, as many MMAs provide group conversation features. Older technologies like mailing lists and IRC also provided the ability to communicate in groups, but only with the introduction of MMAs this type of communication became widely popular. A further advantage of MMAs is the mobility, which allows to easily communicate with others from anywhere and at any time. One of the most popular MMAs is WhatsApp[6], which will be covered in more detail later [2].

Online social networks like Facebook also follow the trend of mobile communication and offer their own MMAs like the Facebook Messenger[7]. Moreover, they also added group communication explicitly, e.g., Facebook, or implicitly, e.g., Twitter hashtags, which are implicitly forming theme-based groups. Thus, the conversation in groups, the possibility of immediate responses, and the omnipresence of smartphones move written online communication further into the direction of real-life conversations.

In the future, data security and privacy, integration into the cloud, and the availability of the same services on all devices will be main points of interest. First steps in this direction have been taken, for instance, by Telegram[8], an MMA founded in 2013. This application offers end-to-end encryption of chats and cloud features. A further trend is unified communication, especially in the professional area. Unified communication combines real-time communication services, such as instant messaging or IP telephony, with non-real-time communication services, like email or fax. Finally, the support of the group communication paradigm by communication networking technology is expected. This means that traffic management solutions take into account group communications, especially in mobile networks, for example, by implementing the publish-subscribe pattern.

[1]ICQ. http://www.icq.com/ – Accessed: February 15, 2016
[2]AIM. http://www.aim.com/ – Accessed: February 15, 2016
[3]Myspace. http://www.myspace.com/ – Accessed: February 15, 2016
[4]Facebook. http://www.facebook.com/ – Accessed: February 15, 2016
[5]Twitter. http://www.twitter.com/ – Accessed: February 15, 2016
[6]WhatsApp. http://www.whatsapp.com/ – Accessed: February 15, 2016
[7]Messenger. http://www.messenger.com/ – Accessed: February 15, 2016
[8]Telegram. http://www.telegram.org/ – Accessed: February 15, 2016



(a) Chat overview. (b) Conversation in a group. (c) Pop-up notification.

Fig. 2. Screenshots of WhatsApp, retrieved from the official website of WhatsApp.

III. Background on WhatsApp

WhatsApp Inc. was founded by Jan Koum and Brian Acton in Santa Clara, California, in 2009. Starting as an iPhone application, WhatsApp soon became more popular and also available for Android, Windows Phone, BlackBerry, and Nokia. In February 2014, Facebook Inc. bought WhatsApp for USD 19 billion. In September 2015, it had more than 900 million monthly active users worldwide [3], being especially popular among young people [4] and reaching a usage share of up to 77% of mobile Internet users in some countries [5].

It is very simple to start using WhatsApp because it is free to download and no complex registration is required. The application automatically identifies users by their phone numbers. Those contacts saved on the smartphone that are also users of WhatsApp will be automatically added to the application's contact list.

WhatsApp combines one-to-one, one-to-many, and group communication by offering private chats, broadcasts, and group chats. In the beginning of 2015, a desktop client and a feature for voice calls via VoIP was added [6].

Figure 2 shows various features of WhatsApp (in this example on an Android device). In Figure 2a, there is an overview over all chats and broadcasts. On top of this list is a toolbar. A private chat can, for example, be started by touching the message icon on the toolbar and choosing a contact of your list.

To start a new group, one has to open the menu and choose new group. Then, the subject of the group (a free text) must be defined and a group icon can be uploaded. Afterwards, up to 100 contacts from the contact list can be invited to join the group. The creator of a group has administrative privileges and can add and remove people from the group at any time and also promote other group members to group administrators.

An example for a group chat can be seen in Figure 2b. Each posts of a member of the group is represented by a speech bubble. Apart from the exchange of text messages, WhatsApp also allows to send photos, videos, and audio files, contact data, as well as the current location of the user. In a conversation, every type of message is seamlessly integrated into a single view, as Figure 2b shows. Every user of WhatsApp will be notified as soon as a new post arrives, whether in a group or in a private chat. This notification can be a sound, an icon, or a pop-up window, which is depicted in Figure 2c.

In contrast to SMS, WhatsApp needs an Internet connection to send and receive messages. For this purpose, it uses the Extensible Messaging and Presence Protocol (XMPP) [9]. WhatsApp is a fully centralized service, i.e., it is a service, which is operated exclusively by the US based cloud provider SoftLayer [7] This work will investigate the way users communicate using WhatsApp. Particularly, the focus will lie on group-based communication.

IV. Implications of Group-based Communication

To analyze the usage of WhatsApp and the implications of group-based communications, we conducted a measurement study on group communications, as well as a survey on the campus of the University of Würzburg, Germany in November 2014 [8]. The survey was divided into three different parts: demographic questions, group communications, and network usage statistics. The participants answered the questions of the survey in a dedicated room using personal or laptop computers, which took around 15 minutes. Questions had to be answered using text fields, single choice, or multiple choice options.

In total, 243 persons participated in the survey, which all had WhatsApp installed on their smartphones. After filtering out invalid or inconsistent answers, 209 participants remained – 106 female and 103 male. The ages of the participants ranged from 17 to 29. The average age was 21.4, which is because mostly students took part in the study. After taking part in the survey, the participants were asked to send us some of their messaging histories from WhatsApp groups by email. In that way, 402 messaging histories have been collected.

A. Usage of WhatsApp

Comparing the usage of WhatsApp to SMS, the survey showed that WhatsApp is used significantly more often than SMS. 85.17% of the participants use WhatsApp at least once or twice a day, whereas only 6.69% use SMS so frequently. This leads to the conclusion that WhatsApp communication was preferred considerably to SMS communication by the participants. The participants also had to indicate if and which other mobile messaging applications they use besides WhatsApp. Most participants (81.82%) also use other MMAs showing that WhatsApp is not the only well-established MMA. The reason is that competitive mobile messaging applications (e.g., Facebook Messenger, Skype, Threema) provide additional or different features than WhatsApp.

The participants were also asked for which purposes they used WhatsApp. 98.09% of the participants answered that they used WhatsApp for private purposes, 92.34% for organizational purposes, 77.51% for fun, 50.24% for important issues, and 33.01% for professional purposes. It follows that,

[9]The XMPP Standards Foundation. http://xmpp.org/ – Accessed: February 15, 2016

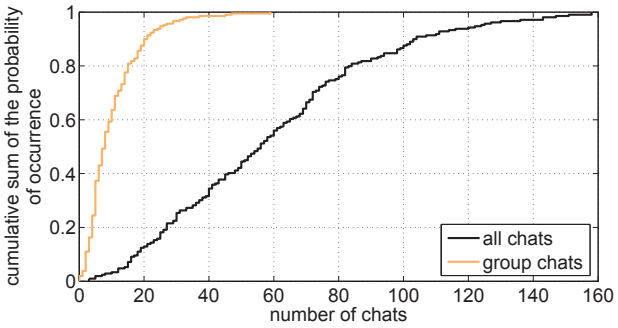


Fig. 3. CDF of the number of all chats compared to the number of group chats

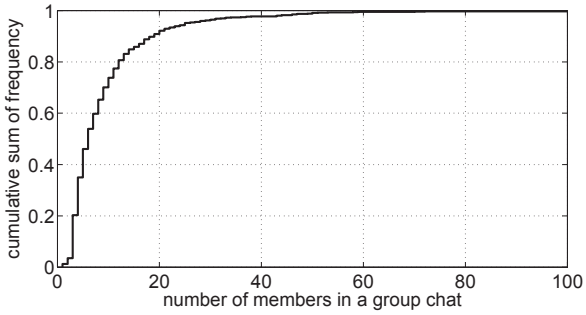
for many people, WhatsApp has become an important means of communication in many conditions of life. Moreover, WhatsApp and group communication seems to be so useful for organizing things that it also qualifies for important and professional purposes.

B. Group Chatting in WhatsApp

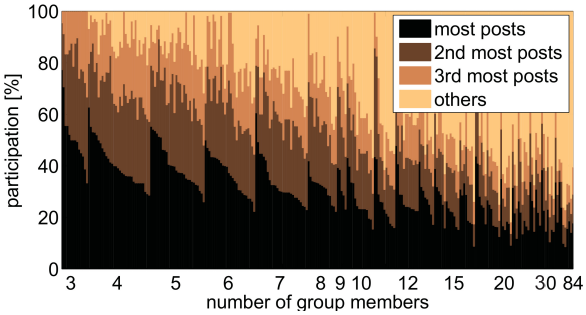
The high popularity of WhatsApp can be seen in the number of different communication chats per user. A chat in WhatsApp can be either a one-to-one chat with only one communication partner or a group chat with group sizes of 3 to 100. Thus, the participants had to count how many chats in total they had on their devices and how many of these chats were group chats.

Figure 3 depicts these distributions by comparing the number of all chats (including group chats) to the number of group chats. The average number of chats is 59. 70.32% of the participants exchanged messages with at least 30 different partners. 12.44% of the participants had even more than 100 chats, the maximum being 158. As can be seen in the brown curve, the number of groups ranged up to 59, the average being 10. Only 1.91% of the participants did not participate in any group chat. Please note that it is possible to delete group chats and the participants estimated during the study that they had already deleted on average 7 groups. Still the share of group chats among all chats is fairly high having an average of 17.94%. For most of the participants (83.28%), this share ranged between 5% and 30%. All in all, this supports the assumption that the group chat feature is used frequently by almost every WhatsApp user, which makes it a key feature of WhatsApp.

We further asked about each group chats and the participants had to specify the number of members of each group, the number of members they did not know, and their personal participation in the group chat. Figure 4a shows the distribution of group chat sizes. The average number of group members is 9, and on average one of them is unknown to the user, i.e., not in his contact list. The average group size is considerable, but low considering that WhatsApp allows creating group chats with a maximum of 100 members. Only few group members are unknown, which leads to the assumption that group chats are mainly used for communicating with specifically selected



(a) Distribution of group chat members.



(b) Active participation of group chat members.

Fig. 4. Distribution of the number of members and their active participation.

people who know each other. Nevertheless, we see that WhatsApp, like online social networks, is able to link people who do not yet know each other (triadic closure).

Next, each collected messaging history was analyzed with respect to how much posts each member actually sent in relation to all sent posts in a group chat, which is visualized in Figure 4b. It shows a bar for each group that indicates the shares of the top-3-contributors and the share of the remaining members. The bars are sorted by group size and the share of the user who sent the most posts. It can be seen that group conversations are rarely balanced. In 8.1% of the group chats, there was one user who dominated the group and sent more than 50% of all posts in this group chat. Furthermore, in 19.2% of the group chats, there were two or three members who dominated, each sending 30% or more of all posts. Also for many large groups, it can be observed that the top-3-contributors account for a quite big share of messages. This leads to the conclusion that most group chats consist of few active, dominating and several passive members. The active members send most posts in a group chat while the others in most cases only read the messages.

C. Impact on Network Traffic

Investigating the impact of group communication on the network, it is important to understand how often messages have to be transmitted. The black line in Figure 5 shows the distribution of the inter-arrival times of posts in group chats. Here the x-axis represents the inter-arrival time in minutes and the y-axis the cumulative distribution. It can be seen that many posts are replied very fast, but some messages have a very

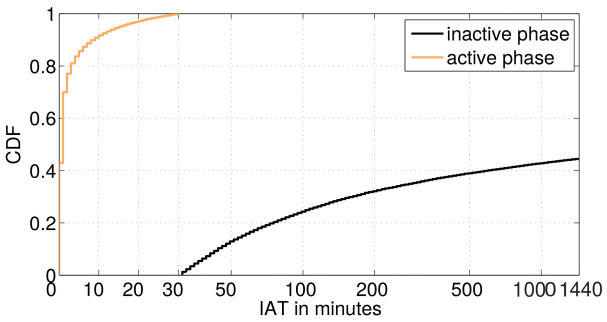


Fig. 5. Inter-arrival times (IAT) of messages in active and inactive phases

large inter-arrival time. 59.60% of the posts were immediate responses and had an inter-arrival time of one minute or less. Considering the range from 0 to 30 minutes, it can be seen that 84.90% of all inter-arrival times are included. Only 15.10% of all inter-arrival times are 30 minutes or longer. This suggests that there are different communication phases in group communications, which resembles older on-off-models (e.g., Markov-modulated Poisson processes) in telephony or networking [9].

Thus, in Figure 5, we also separate the group communication into active and inactive phases. An inter-arrival time equal or lower than 30 minutes, which is the default session timeout for web servers such as Apache Tomcat, is assigned to an active phase, higher than 30 minutes to an inactive phase. The active phase is plotted in orange on a linear scale from 0 to 30 minutes, and the inactive phase is shown in brown on a logarithmic x-axis from 30 minutes to almost two years (~106 minutes). More than two thirds (69.92%) of the inter-arrival times in active phases were immediate answers in one minute or less. In inactive phases, 44.42% of the inter-arrival times are one day or less and almost all messages are replied within one week (~104 minutes). The probability of changing from an active to an inactive phase is 12.24%, from inactive to active it is 69.99%. This distributions support the statement that WhatsApp constitutes a very fast communication.

Over the course of a day, the typical diurnal pattern can be observed with most posts being sent in the evening from 6pm to 8pm (15.10%), fewest between 5am and 7am (0.44%). This also confirms the statement of the participants that only few use WhatsApp for professional purposes. The participants were asked to enter the statistics of WhatsApp's network usage, which are collected by the application on every device. Generally, the communication of each user in WhatsApp is balanced, however, an average user receives roughly 21% more messages than he actually sends. The slightly higher rate of received messages is likely to be caused by group chats. Recall that in these chats, every message sent by one user is potentially received by a multitude of other users. We observed in the messaging histories that media posts, e.g., photos, videos, or voice messages, are sent very rarely. On average only 6.53% of all posts in a group were media posts. However, considering the relation between received media

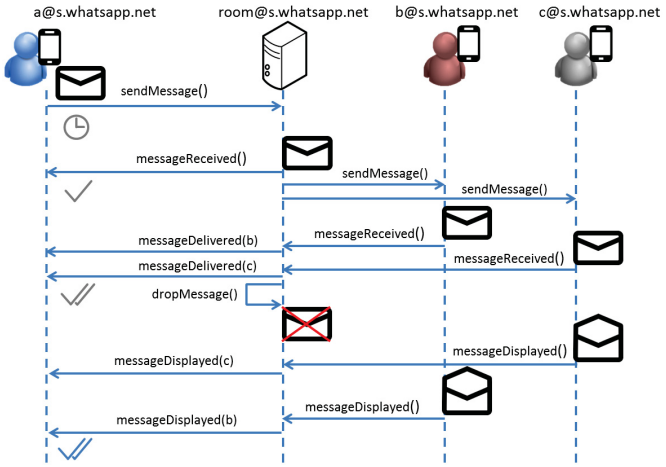


Fig. 6. Process of sending a message in a group chat with three users

bytes and received bytes in total, nearly 86% of the total bytes can be attributed to media posts. This ratio has a linear behavior among our participants with high Pearson correlation coefficient of 0.92. This indicates that media posts generally cause the largest part of WhatsApp's network traffic. Similar findings obtained from passive measurements in a cellular network are presented in [7]. With respect to message sizes, a simple measurement showed that, as expected, the size of a text message increases linearly with the number of characters. During our study, we observed an average message length of 36 characters, which resulted in an average text message size of 317B. For media messages, [7] reports an average size of 225KB. However, a more thorough investigation is needed, as we noticed during our measurement that WhatsApp applies transcoding and scaling to transmitted images and videos.

Not only the data volumes of the messages themselves, but also a lot of application-layer signaling traffic puts load on the mobile network. Figure 6 illustrates the process of sending a message in a minimal group chat with only three users following the specifications of the XMPP protocol [10]. User A sends a message to the server addressing it to the room of the group. The message can be a text post or media post and its content can potentially have a high data volume. As soon as the server receives the message, it sends an acknowledgement to user A. With a simple setup, we measured that sizes of acknowledgements are around 60B. Upon arrival of the acknowledgement, the clock symbol turns into a single grey checkmark on user As phone. The server forwards the message with the content to all members of the group, which are user B and user C. The message receptions are acknowledged to the server. The server forwards the acknowledgements to user A, signaling that the message was delivered. As soon as the message is received and acknowledged by all members of the group, the server drops the message and the symbol in the chat turns into a grey double checkmark. If a user in the group displays the message to eventually read it, it is reported to the server, which forwards it to the sender.

After all group members were reported to have displayed the message, the double checkmark turns blue. Hence, sending a message within a group implies a number of subsequent messages signaling the reception and processing state of the message of each user. Additionally, each application signaling induces a lot of signaling in the mobile network [11]. In groups with many members this results in a significant traffic volume and high number of signaling messages, which have to be processed by the network. In peak hours or in case of flash crowd events, this may lead to problems in the network and requires management of the traffic.

V. Conclusion

In this paper, we investigated group-based communication in WhatsApp. Communication in groups constitutes an emerging communication paradigm, which has a huge impact on todays mobile networks. We conducted a survey on WhatsApp usage and analyzed collected messaging histories to better understand group communication and its impact on network traffic.

All in all, this work provided a first investigation of group-based communication in WhatsApp, which changes the way how people communicate. The analyses presented in this paper allow for modelling and simulating communication in groups. Thereby, novel traffic management mechanisms can be designed and evaluated in order to better cope with the network demands. These might include the ICN proposal and caching of content close to the end users, publish-subscribe mechanisms, or multicast transmissions instead of transmitting content naively to each individual group member. Moreover, other approaches like mobile ad hoc transmissions could become relevant, especially in cases in which (parts of) the virtual WhatsApp group physically meet. Content could then be exchanged directly when the members are in the same WiFi network or by short-range device-to-device communication as currently discussed for 5G networks. It remains for future work to analyze the benefits and to study the applicability of each approach in order to adapt the current Internet technologies to the changing user behavior with group communication being well underway.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under grants HO4770/1-2 and TR257/31-2 (DFG project OekoNet). The authors alone are responsible for the content.

References

[1] A. F. Cameron and J. Webster, "Unintended Consequences of Emerging Communication Technologies: Instant Messaging in the Workplace," *Computers in Human Behavior*, pp. 85–103, 2005.

[2] C. Montag, K. Błaszkiewicz, R. Sariyska, B. Lachmann, I. Andone, B. Trendafilov, M. Eibes, and A. Markowetz, "Smartphone Usage in the 21st Century: Who is Active on WhatsApp?" BMC Research Notes 8.1: 331, Tech. Rep., 2015.

[3] L. Rao, "WhatsApp hits 900 million users," 2015. [Online]. Available: http://fortune.com/2015/09/04/whatsapp-900-million-users/

[4] J. Fetto, "The $19 billion question: Who uses WhatsApp and why are they so important to Facebook?" 2014. [Online]. Available: http://www.experian.com/blogs/marketing-forward/2014/02/21/the-19-billion-question-who-uses-whatsapp-and-why-are-they-so-important-to-facebook/

[5] J. Mander, "WhatsApp Usage Highest in LatAm and MENA," 2014. [Online]. Available: https://www.globalwebindex.net/blog/whatsapp-latam-mena

[6] A. Chowdhry, "WhatsApp Android App Now Has Free Voice Calling For Everyone," 2015. [Online]. Available: http://www.forbes.com/sites/amitchowdhry/2015/03/31/whatsapp-calls-android/

[7] P. Fiadino, M. Schiavone, and P. Casas, "Vivisecting Whatsapp Through Large-scale Measurements in Mobile Networks," in *Proceedings of the ACM SIGCOMM*, Chicago, IL, USA, 2014.

[8] M. Seufert, A. Schwind, T. Hoßfeld, and P. Tran-Gia, "Analysis of Group-based Communication in WhatsApp," in *Proceedings of the 7th EAI International Conference on Mobile Networks and Management (MONAMI)*, Santander, Spain, 2015.

[9] D. L. Jagerman, B. Melamed, and W. Willinger, "Stochastic Modeling of Traffic Processes," *Frontiers in Queueing: Models and Applications in Science and Engineering*, pp. 271–320, 1997.

[10] P. Saint-Andre, "XEP-0045: Multi-User Chat," 2008. [Online]. Available: http://xmpp.org/extensions/xep-0045.html

[11] C. Schwartz, T. Hoßfeld, F. Lehrieder, and P. Tran-Gia, "Angry Apps: The Impact of Network Timer Selection on Power Consumption, Signalling Load, and Web QoE," *Journal of Computer Networks and Communications*, 2013.