Noname manuscript No.
(will be inserted by the editor)

User Interactions and Behaviors in a Large-Scale Online Emotional Support Service

Derek Doran · Luisa Massari · Maria Carla Calzarossa · Yu Li · Daniele Tessera · Junjie Zhang

Received: date / Accepted: date

Abstract Among the most important reasons why people communicate with each other is to share and support each other through emotional problems, yet most online social systems are uncomfortable or unsafe spaces for this purpose. This has led to the development of online emotional support systems, where users needing to speak to someone can anonymously connect to a crowd of trained listeners for a one-on-one conversation. Towards understanding the qualities of this emerging type of online social system, this article examines the users, conversations, and activities performed across *7 Cups*, a massive, vibrant emotional support system with a community of listeners ready to help those with any number of emotional issues. The study makes intriguing insights along the world-wide adoption of the service, the need of its users to seek support from many others, a power-law effect of listener popularity, that users have a penchant to connecting to others along common interests and that a core-periphery like structure emerges among conversation networks, and identifies qualities of the system that drive user engagement and retention. We further study the words and actions of misbehaving users who have been reported on or blocked, and build a machine learning classifier able to anticipate their undesirable actions with reasonable accuracy and very low false positive rate. The qualities

Derek Doran
Dept. of Computer Science and Engineering, Wright State University, Dayton, OH, USA
E-mail: derek.doran@wright.edu

Luisa Massari
Dept. of Electrical, Computer, and Biomedical Engineering, University of Pavia, Pavia, Italy

Maria Carla Calzarossa
Dept. of Electrical, Computer, and Biomedical Engineering, University of Pavia, Pavia, Italy

Yu Li
Dept. of Computer Science and Engineering, Wright State University, Dayton, OH, USA

Daniele Tessera
Dept. of Mathematics and Physics, Catholic University of the Sacred Heart, Brescia, Italy

Junjie Zhang
Dept. of Computer Science and Engineering, Wright State University, Dayton, OH, USA

recovered shed insight into the user dynamics and communication structure of an online emotional support service, the features that drive engagement, and a means of identifying misbehaving users automatically.

Keywords emotional support systems · workload characterization · time series analysis and forecasting · network analysis · behavior prediction · harassment detection

1 Introduction

When people are faced with an emotional problem, it is often their friends, colleagues, and family to whom they turn to seek relief [41]. A common mode of communication to these parties are through online-based social platforms, that include Facebook, LinkedIn, Twitter, Snapchat and Kik. Yet the public or semi-public nature and functions of the platforms and the permanency of their communication records [28] make them an, at best, uncomfortable way to seek out and obtain support for emotional, private, and psychological problems. This discomfort is further exasperated by the social pressures of users to display an 'idealistic' or particular impression about themselves [39] on social media websites. On Facebook, for example, users may curate their profiles and photos to express an image compatible with their wide network of friends, relatives and professional colleagues, and a user may choose not to display any personal information on LinkedIn that would not be well received by fellow business professionals. It is thus no surprise that users do not often use commonplace online social systems to express emotional despair or as a platform to call for help.Yet online social systems are continuing to rise as a dominant form of communication in society [36], and without an outlet for Internet communication suitable for emotional support, there is concern that people will be further discouraged to seek help [48].

Existing emotional support systems on the Internet that help people in distress vary in regards to the type of support offered. Some sites offer generic, written advice for common emotional and mental ailments[1], while some others offer tools for visitors to perform a self-assessment of mental conditions [53]. Other systems take the additional step of providing access to a live therapist, but only in cases when a user is suffering from specific and critical conditions (e.g., suicide contemplation[2] or upon receiving a life-threatening health prognosis[3] [2,22,42]). Yet general emotional conditions that are perceived as "less critical" in nature, such as minor bout of depressions, feelings of anxiety about upcoming events, hurt feelings after losing a loved one, or consistent levels of stress caused by work and family, can be most detrimental to the long-term health of a person (including loss of lifespan [33] and a greater propensity for becoming irritable [45]). The countless types of 'minor' emotional problems that plague people make forming online systems to support them intractable. Another shortfall of many existing systems is that users are helped in *isolation* through private evaluations or by one-on-one engagements with a professional.

[1] `http://www.stress.org/emotional-and-social-support`
[2] `http://www.crisischat.org`
[3] `http://www.cancersupportcommunity.org`

But evidence suggests that people who are able to integrate into a network of peers supporting each other find greater improvements in their mental health compared to those who seek isolated support [44].

These observations suggest the need for a new kind of online social system that offers private, anonymous, quick, and live emotional support for those who prefer to communicate online and need either immediate help or simply someone to talk to about major and minor emotional concerns [4,25]. The social elements of such a system (e.g., user profiles, group communication and messaging mechanisms) would offer group support, and could have a critical mass of participants where almost any type of emotional problem can be discussed. However, it is unclear exactly *how* such systems should be designed to maximize the help they can provide to their users. For example, how should a system identify ideal volunteers to communicate with? How can a system engage volunteer listeners and members needing help to guarantee that they return and receive help often? What are the best mechanisms for facilitating meaningful interactions, in ways that minimize the possibility of harassment or 'trolling' behaviors? And is there any evidence to suggest that such a system can obtain a critical mass of users so that a live volunteer is always available to help someone in need?

Unfortunately, a meta-analysis of medical, social science, and other online publication databases found just a handful (513) on the topic of online emotional support systems [11]. Of those, 52% are papers simply reviewing and commenting on a site, 41% are studies of their use involving simple, descriptive statistics, and only 6.6% evaluate the services through a careful intervention study. We assert that the best way of answering the questions listed above is to study measurements that can be distilled into actionable insights about the users, conversations, actions, and behaviors of users on an existing online emotional support service. Towards this end, in this paper we characterize the users of, and their interactions and actions within, a large and successful example of an online social system called *7 Cups*[4]. This work offers an organized systematic multi-faceted and complementary analysis of this emotional support system by integrating and extending previous conference papers [8], [15], and [32]. In particular, the extensions refer to the application of time series analysis and forecasting for modeling and predicting the temporal patterns of user conversations. In addition, we extend the analysis of the structural properties of bipartite networks to different categories of users (e.g., registered and occasional users). This analysis allows us to capture in more details the structure of the interaction co-occurrences between users and to asses to what extent the registration process represents a barrier for users who seek support from the service.

7 Cups is a platform that facilitates connections between a crowd of live "listeners" (i.e., para-professionals wanting and willing to support people facing a variety of emotional problems) and users who need support. "Members" and "guests" can have a live, one-on-one conversation with a listener, can participate with other members and listeners on group message boards and live chats, and can browse the profiles of other members. Both members and listeners are able to show off their account level and badges, which are gamification mechanisms that encourage users to remain

[4] `http://www.7cups.com`

engaged on the website. 7 Cups attracted about 90,000 listeners who have helped over 1M people in about 3M *one-on-one conversations* (i.e., private asynchronous or real-time message exchanges) during its first two years and has experienced an exponential grown in popularity since. Its success and growth demonstrate the demand for safe online spaces for people needing emotional support.

The layout of this paper is as follows: Section 2 presents some related work on studying online systems offering emotional support. Section 3 describes the main characteristics of the platform and of the dataset analyzed in this study. Section 4 focuses on some activity analytics, while the detailed evaluation of one-on-one conversations and the models of their temporal patterns are described in Section 5. The interactions of users and listeners in one-on-one conversations and the various types of relationships induced by these interactions are discussed in Section 6. Section 7 then examines engagement on the platform, and uses predictive analytics to identify user behaviors and platform features that drive engagement. Finally, Section 8 examines the kind of user harassment that could happen on the platform through topic and machine learning models. We summarize and conclude the study in Section 9.

2 Related Work

Owing to the fact that online emotional support systems are an emerging communication medium, the literature on their characterization, analysis, and evaluation is limited. Past work examines particularly small online support communities. For example, Maloney-Krichmar *et al.* investigated the dynamics of group interactions among an online self-help group for knee injuries [35]. Barak *et al.* established a positive relationship between the amount of activity of adolescents in an online support group and the emotional relief they felt [3], underscoring the importance of building online systems that facilitate user interactions. Yuen *et al.* highlighted how remote assessment, treatment and consultation provided over the Internet, and video conferencing have the potential to increase access to quality psychological services [52]. They also discussed the clinical, ethical, logistical challenges involving security, competence, usability and technical difficulties on such platforms. Wang *et al.* analyzed the relationship between the support a user receives in an online social group and their engagement with others online [46]. Ploderer *et al.* delved into the discussion topics on a Facebook group of people trying to quit smoking, and found that most supportive responses come from those who just began trying to quit, rather than long-term quitters [40]. In this effort, we study a unique emotional support system designed from the ground up to facilitate online anonymous emotional support, rather than simply examining message boards or conversations on top of existing online social networks. The hundreds of thousands of users present on 7 Cups also enable a more through data-driven approach to understand its population of users and the conversations they hold.

An important part of our study involves measuring the *engagement* of a user on an emotional support platform. One body of work considers physiological measures of engagement [1,26,29,43], that may come from sensors like eye trackers, mouse movements, blood pressure, heart pulses, and cameras and measure features such as eye movements, heart rate, and mouse clicks. Such analysis reveals bodily reactions that are cognitively linked to mental engagement in a task, but physiological readings are not easily relatable in the context of an emotional support system although they would be very beneficial.

Previous research has proposed a variety of measures for quantifying or measuring user interactions. The most popular approaches involve self-reporting, where questionnaires and interviews are given to users that must respond honestly [30]. For example, Webster *et al.* [47] developed a seven item questionnaire with items that include the degree of attention, challenge, intrinsic interest, and variety in the context of presentation software, while Jacques [27] developed a 13-item survey to evaluate user engagement in e-commerce environments. More modern approaches measure user behaviors from the software system itself to determine engagement. These behaviors, often called Key Behavior Indicators (KBIs), may be based on average page views, bounce rate (people arriving at a website and leaving immediately), and user loyalty metrics [5]. By themselves, KBIs reveal the engagement characteristics of an individual user. Generalizing KBI-based approaches to any online platform, however, is not realistic since the KBIs used in each study vary widely and are platform specific. Instead, this paper proposes a framework for engagement analysis that is applicable to any type of online social system. To ensure broad applicability, the approach generalizes notions of engagement through multiple abstract dimensions into application-specific KBIs.

Another dimension of our analysis is the study of users who are *suspicious* and may be displaying behavior discouraging others from participating. There are some studies reporting on the negative impact such users may have on an entire online social system, and especially when users are asked to share personal information. Newman *et al.* showed that people are cautious while sharing health related information on online social networks like Facebook [39]. They also demonstrated that users are hesitant to share certain types of information, especially personal or information that may be used as fodder for cyberbullies, to protect themselves and to manage their online impression.

Detecting suspicious users is also an important task for an online emotional support system. Many kinds of detection systems for this task have been proposed [10, 12, 16, 23] but define what a "suspicious activity" is in different ways. Most methods focus on detecting accounts that send spam messages initialized by an attacker, flows through a series of suspicious accounts, and reaches a victim account. Gao *et al.* [18] designed a method to reveal campaigns of suspicious accounts by clustering accounts sending messages with similar content. Yang *et al.* [50] extracted a graph from the follower relationship of Twitter accounts and propagated a maliciousness score using the derived graph. Wu *et al.* [49] proposed a method to detect both social spammers and their messages by learning from relationships between users and messages. The challenges to detect suspicious users on an emotional support system, however, is unique. For example, elements important for detecting spam attacks such as URLs and binaries do not have to be present to enable successful user harassment attacks. Moreover, online emotional support systems define different kinds of associations between users beyond the traditional "friendships" and "follower/followee" relations.

3 Dataset and Platform Characteristics

Our study is based on a database graciously provided by 7 Cups that captures the attributes of interactions, users, and their activities between the twenty months of December 5th, 2013 and August 14th, 2015. For privacy reasons, the database does not include any feature related to true identity and contact information of the users. Instead it includes the identifiers of the users involved in each interaction together with the specific features of the individual activities (e.g., timestamps, type of user, number of messages exchanged/posted). Although this database is dated, it represents the most comprehensive and detailed dataset on the users and interactions in an online emotional support system that, to the best of the authors' knowledge, have been studied.

The platform is centered around users who seek emotional support services for everyday issues and users who provide these services. Depending on their role, 7 Cups users are classified in three groups, namely:

- **Listeners** who have been trained in active listening to understand and support others;
- **Members** who register an account on the site to speak with someone about emotional distress;
- **Guests** who wish to speak to someone about emotional distress without registering on the site.

Any guest can eventually transition to become a member. 7 Cups allows multiple types of interactions with others, including by:

- *One-on-one conversations*, where a member or a guest privately speaks to a listener;
- *Topical group chat rooms*, where users hold live chats with people who have overcome similar challenges;
- *Thematic support forums*, where users engage in asynchronous discussions about particular topics in discussion threads;
- *Listener blogs* where listeners publish articles about emotional support topics.

To obtain emotional support services tailored to their age, members and guests have the option of specifying if they are under or at least 18 years of age. 7 Cups then provides chatrooms dedicated to 'younger' and 'older' users as well as listeners specialized in the support of each age group. In what follows, we denote these groups of users as *teens* and *adults*, respectively.

It is worth noting that listeners are not allowed to give any advice in their interactions with members and guests. The main task of listeners is to validate – thanks to their active listening skills – the concerns of the users and provide support by helping them explore what is going on in their life. Moreover, a listener can jump in any chat room or forum at any time to help the discussion stay productive and supportive.

4 Activity Analytics

We begin the study with basic measurements on the activities within 7 Cups. Summarized in Table 1, we find that 7 Cups users take significant advantage of the emotional support services of the platform. In the first twenty months of operation, nearly 132M messages were exchanged in one-on-one conversations between a member or a guest and a listener, and some 25M messages posted in group chat rooms by members and listeners. Moreover, users participated in thematic forums – with about 350K messages distributed across more than 20K threads – and received tips and support from 74 different blogs. As can be seen, one-on-one conversations, where users engage in

Conversations		Chat rooms		Forums	
Volume	Exchanges	Volume	Exchanges	Volume	Exchanges
3,078,158	131,659,409	52	24,872,560	453	353,521

Table 1: Summary of the volume of interactions and exchanges of 7 Cups users over twenty months.

an anonymous private exchange with listeners about their emotional distress, are the most popular service on 7 Cups. We evaluate this service further in Section 5.

Users are also active in group chats and support forums where they discuss a large variety of subjects, including, among the others, depression, work stress, eating disorder, loneliness, parenting support, anger management as well as their favorite hobbies. The list of subjects of topic forums and group chats are constantly evolving on 7 Cups.

Members posted nearly 17M messages in group chats, which is about 13% of the overall volume of one-on-one messages exchanged, with listeners accounting for approximately 30% of group chat messages. Listeners were particularly active in the group chats specifically dedicated to discussions on issues related to their listening activity and on general issues about the platform (e.g., Listener Community Room, Teen Listeners).

The support forums on the platform involved nearly 328K users whose discussions were grouped across 23K threads. These users, who subscribed on average to 11.2 threads each, posted messages, ‘upvoted’ the posts they found most helpful or simply followed the ongoing discussions. The largest number of upvotes was received by the posts of the threads on motivational quotes and appreciation (nearly 20,000 upvotes each).

In general, the majority of 7 Cups users strictly obeyed the terms of service stated by the platform: they did not post any inappropriate content and they did not start any hostile interaction. In fact, a very small fraction of the messages posted by the users, (e.g., 0.15% of the chat messages) was removed from the platform because of harassing, privacy invading or threatening content. Similarly, few one-on-one conversations have been blocked due to inappropriate behaviors. All these measures have been implemented into the platform to protect users from unwanted interactions.

5 Conversation Analysis

The one-on-one conversations is the core support service offered by 7 Cups. A detailed evaluation of these private message exchanges between a member (or a guest) and a listener is fundamental to understand user behavior and interactions being developed in the platform between users and listeners. In what follows we characterize the conversations commenced by 7 Cups users and we identify models for describing their temporal patterns. We note that the content of these one-on-one conversations is not considered in our analysis because it is never recorded by 7 Cups.

5.1 Listener matchmaking

7 Cups allows members and guests to initiate a conversation with a listener they select from profiles or with a listener the system automatically selects from among those currently available. Specific listeners may be selected because their stated interests and expertise align with a particular topic a user wants to discuss, while random listeners may be selected by those who are testing the service, or are facing a general anxiety and simply wish to speak to anyone. Only one third of the conversations had a member or guest select the listener they wished to talk to, suggesting a need for an emotional support service that provides someone with general, rather than specific expertise.

Of the 3M one-on-one conversations commenced by 7 Cups users, we found 44% initiated by members and the remaining 56% by guests. These conversations involved over 1.3M users and nearly 90,000 listeners as detailed in Table 2.

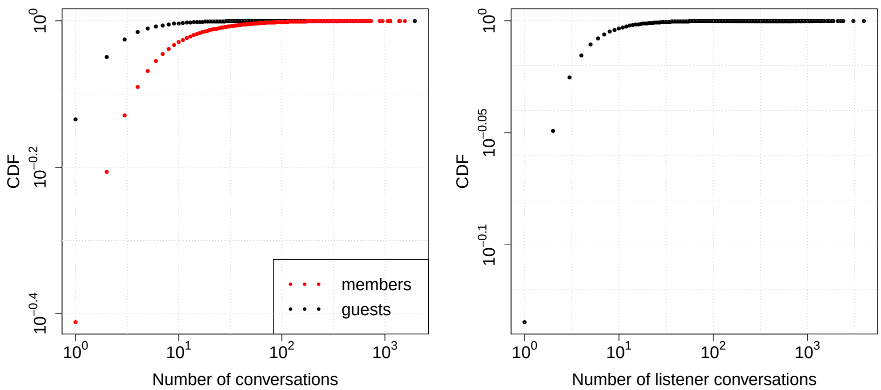
Listeners	Members	Guests
89,347	297,151	1,043,821

Table 2: Number user types in one-on-one conversations.

It is interesting to note how most conversations were created by guests; this may indicate that such guest users of 7 Cups are simply ‘trial’ users who want to engage in a conversation to determine if the service will be beneficial to them. It may thus be important for an online emotional support service to offer a guest account or some frictionless mechanism for people to immediately begin a conversation, without needing to create “ a member account or provide personal information to the service.

5.2 Conversational diversity

The literature reports a positive effect of receiving emotional support and guidance from multiple others when going through crises such as depression [6] and poor health prognoses [14]. An online emotional support service’s ability to let a user communicate with many others, rather than a single professional, is a key differentiator between seeking online and offline help. Conversing with many listeners may thus be a key advantage of using an online support service, especially when people may feel uncomfortable in sharing with friends and family or if it is too costly to see multiple professionals. To investigate this phenomenon on 7 Cups, Figure 1 plots on a log-log scale the cumulative distribution functions (CDFs) of the number of conversations held by members/guests and by listeners, respectively. We positively note that while



(a) Member/guest conversations

(b) Listener conversations

Fig. 1: CDF on log-log scale of the number of conversations held by members and guests (a) and by listeners (b).

a high percentage of guests engage in a single one-on-one conversation (approximately 73%), this is true only for about 38.7% of all members. We also measure a quarter of members with at least four one-on-one conversations each, enabling them to obtain the support of different listeners. On the contrary, guests are sort of occasional 'visitors' who tend to converse with one listener only. The ability for members to engage with multiple listeners, and their willingness to do so, indicate that the online emotional support service enables guidance and advice from many others that may otherwise be difficult to achieve.

Interestingly, listeners are keen to provide support to others despite the fact that they are volunteers: half of the listeners supported up to 13 users in one-on-one conversations and one quarter supported 35 users or more. These high values may be promoted by the fact that listeners are trained by 7 Cups to be able to help most kinds of general emotional problems for users of different age groups. The CDF (see Fig.1(b)) identifies a substantial group of listeners (about 10% or $8,935$) who were very active and engaged in conversations with at least 80 different users. Moreover, a small proportion of listeners (7%) held only one conversation during the observation interval. These statistics are very encouraging: most listeners decide to engage with more than one user, and there are many listeners who devote substantial energy and time to engage with a large number of others. This phenomenon could be driven by the site's gamification mechanisms to encourage engagement, but nevertheless, it is possible for an online social system to cultivate an active community of volunteers to help others in need.

About three quarters of listeners specialized in adult's support and one quarter in teen's support, while a very small fraction of listeners (i.e., 1.9%) offered their support to users of both age groups.

The volume of conversations held by members and guests, subdivided according to these age groups, is summarized in Table 3. In our dataset we identified some users (i.e., about 41,000) who changed their age group during the first twenty months of 7 Cups. A key takeaway from the table is the proportion of conversations of members and guests who are 18 years of age or younger (29.7% of members and 28.5% of guests). On the one hand, any kind of online social media may attract a younger set of users causing this proportion to be large, but on the other hand we find a very significant number of adult users willing to use this non-traditional means of finding emotional support. Table 3 also reports the volume of messages exchanged in the conversations. The largest number of messages is exchanged in the conversations of members, namely, 71 and 45 messages per conversation of adults and teens, respectively.

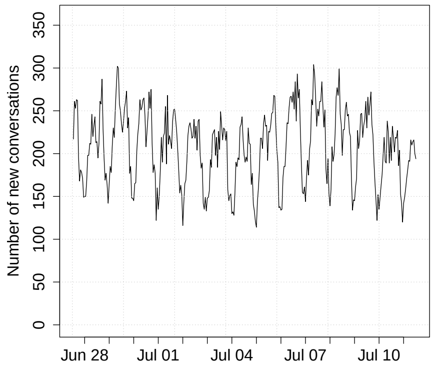
	Members		Guests	
	Adult	Teen	Adult	Teen
Num. conversations	951,701	403,903	1,231,414	491,140
Num. messages	67,395,434	18,260,874	36,134,943	9,868,158

Table 3: Volume of conversations of members and guests controlling for age group.

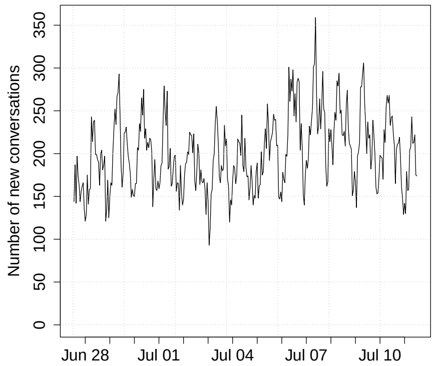
We remark that one-on-one conversations are permanent, in the sense that they last forever and either participant can resume them at any time. However, the platform allows users to "hide" conversations from their user interface in the event a conversation has become inactive or if the member or listener has decided to ignore it. Nearly 25% of all conversations are hidden by members, which could suggest that most of the conversations they engage in are worthwhile enough to remain active. Listeners hid 49% of their conversations, which may be caused by engaging with guests or members who ignore their initial greeting or leave the site after a brief trial.

5.3 Temporal conversation patterns

7 Cups timestamps the initiation of every conversation in their database. We use this information to study temporal patterns in the intensity of conversations. Figure 2 plots the number of new conversations created per hour by members and guests on 7 Cups between June 28th and July 12th 2015. Although 7 Cups can be accessed by users at any time from anywhere in the world by means of any type of device, we identify clear diurnal patterns. Both members and guests prefer to start conversations in the evening and overnight, while they start very few conversations in the middle of the



(a) Member initiated conversations



(b) Guest initiated conversations

Fig. 2: Temporal patterns of the volume of conversations commenced by members (a) and guests (b) over a two week period. x-axis labels are centered to 12pm UTC.

day. For example, the maximum number of conversations (i.e., 353) is detected in a one hour interval starting at 7pm (see Fig. 2 (b)). We also observe some fluctuations in the temporal patterns that are particularly noticeable in guest conversations, and that new conversations are not evenly distributed across days and hours.

Despite this, we are able to build a model of conversation intensity over time that fits this data well. The modeling methodology follows the approach outlined in [7, 9], namely:

1. analysis of time series properties to find possible outliers;
2. estimation of the periodicity of the time series by spectral analysis of its autocorrelation function;
3. time series decomposition into trend, seasonal and irregular (noise) components;
4. identification of component models by numerical fitting and moving average auto regressive techniques; and finally
5. forecasting future dynamics by extrapolation over a moving horizon.

We apply this approach to build a time series model representing the hourly patterns of the 530,359 conversations commenced by both members and guests over eight weeks (until July 18th, 2015). The spectral analysis of its autocorrelation function leads to the identification of patterns that repeat in time with a weekly and a daily periodicity. The decomposition of the time series into its trend, seasonal and irregular components is shown in Figure 3. The seasonal component is paired with an increasing trend over the entire two week period, indicating that the volume of conversations steadily grows with time. This is due to an increased number of users taking advantage of the emotional support services provided by 7 Cups as well as to the tendency of members to seek support from multiple listeners, thus commencing multiple conversations.

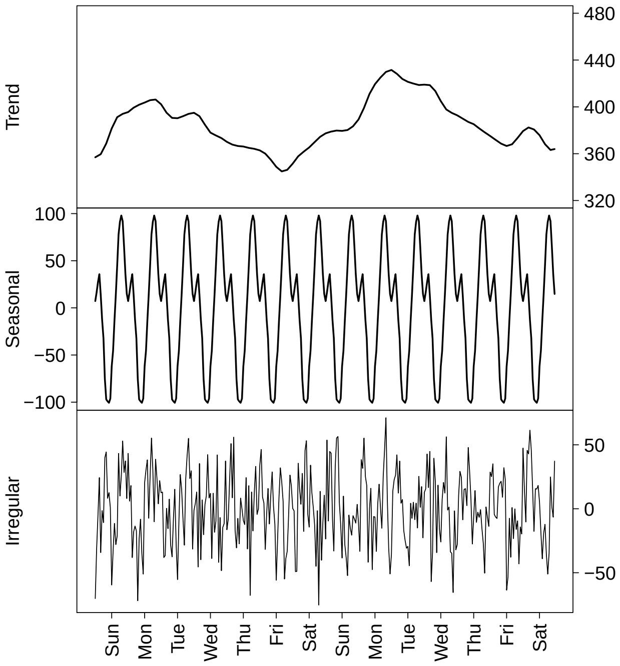


Fig. 3: Decomposition of the time series representing the hourly patterns of the volume of conversations commenced by members and guests over two of the eight weeks of the training set. x-axis labels are centered to 12pm UTC.

Trigonometric polynomials of degree six and three are identified by numerical fitting techniques to model the trend and seasonal components, respectively. Note that the trend model also includes a linear component that accounts for the growth of the volume of conversations over time. A Seasonal Auto Regressive Moving Average model $(1, 0, 1) \times (1, 0, 0)_{24h}$ is identified for modeling the irregular component. The final model is then obtained as the superposition of the individual models. The fitted model over a three week snapshot of the data in Figure 4 shows a very nice fit with mean absolute percentage error at just 6.7%. The figure also shows accurate one-step ahead forecasting after July 19th, with overall mean absolute percentage error of 6.86%.

6 Interaction Analysis

We now move from studying user activities to interactions between users, and more specifically, how these interactions are structured across the online social system. For this purpose, we build a social network where users and listeners are related if they hold a conversation between each other. The social network enables a structural analysis of interactions that can inform how users choose to engage with listeners on

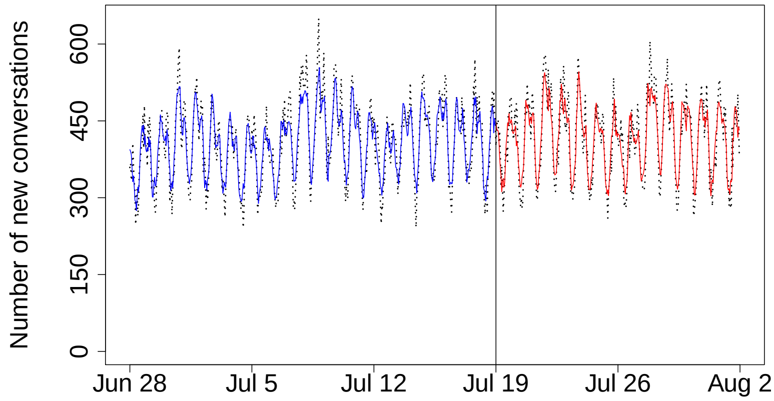


Fig. 4: Hourly patterns of the volume of conversations (dots) together with their overall model over three weeks (blue curve) and one-step ahead prediction over two weeks (red curve).

7 Cups, if some subsets of listeners are more popular than others, and if a pattern of users selectively choosing listeners can be seen.

6.1 Structural analysis

We represent conversations across the system as a bipartite network whose nodes are partitioned by whether they correspond to a user or a listener. We visualize a random subsampling of 30,000 relations in the bipartite graph and their incident users in Figure 5. Member and listener nodes are colored in red and blue, respectively, under a Yifan Hu layout [24] with node size proportional to its degree. The layout reveals two dense, distinct subnetworks of members and listeners, which may correspond to teen and adult members and the listeners they interact with (note that listeners on 7 Cups are trained to support only adults, teens, or both). We further observe that the two dense communities are anchored by a small number of members whose degrees are strikingly high, e.g., by a small proportion of members who connect with a large number of listeners in the network. We hypothesize, but cannot confirm, that some of these very high degree members may be a bot or some malicious users: the members from this sample with highest degree connected to 1,065, 681, and 352 distinct listeners, respectively (note that the average degree of the bipartite graph is 7.13, two to three orders of magnitude smaller than those samples with highest degree). It is difficult to have sampled so many relations to such members by chance, particularly because over 1M edges exist in the bipartite network. It would be prudent for an emotional support service to detect, track, and expel such outlying accounts that are, perhaps needlessly, draining the time and energy of a significant number of volunteer listeners. Such listeners may become dissuaded from using the site further should they have a negative experience with these ‘spamming’ members.

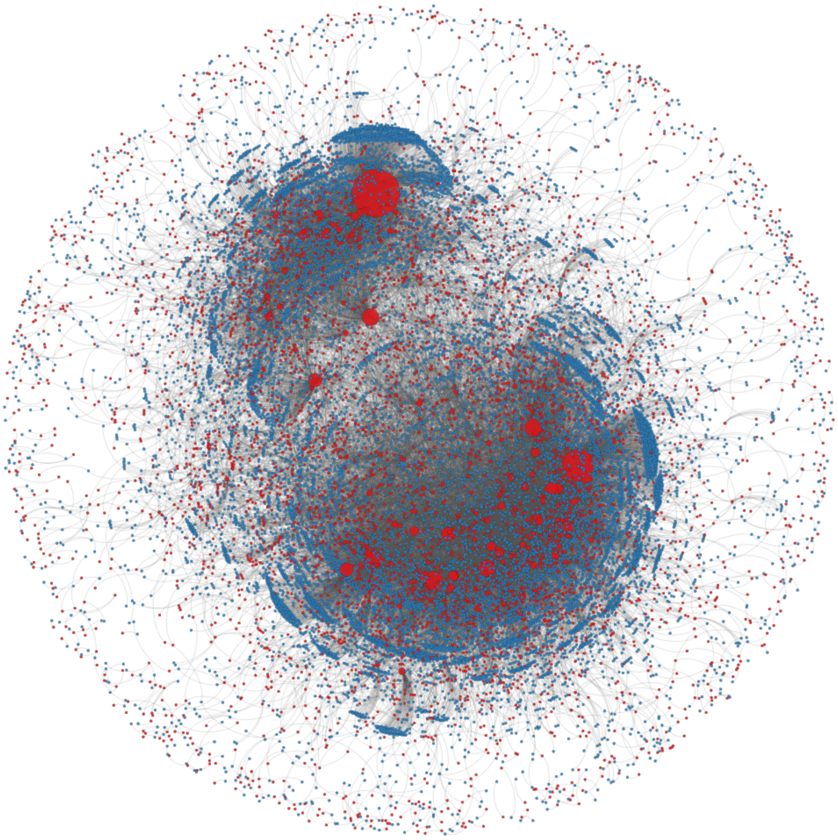


Fig. 5: Bipartite network of member conversations (best viewed digitally).

Table 4 list some common structural properties and statistics of the bipartite network [17, 34, 37]. The statistics reveal over 2,500 disconnected components, the largest of which (GCC) includes almost every user (99.55%) on the platform. That almost all members, guests and listeners emerge in a single GCC is positive as it suggests that isolated communities of exclusive users and listeners may not exist on the site. This is crucial for users to be exposed to a wide mix of listeners and their suggestions.

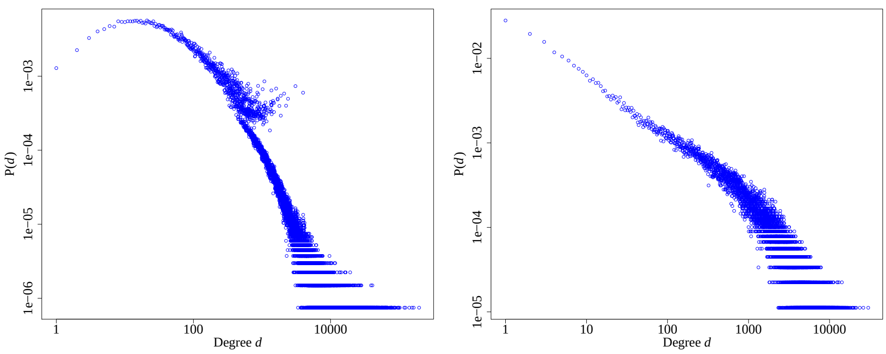
To capture the structure of interaction co–occurrences between listeners and users, we also study the one-mode projections of these bipartite networks. We recall that a one-mode projection of a bipartite network of m listeners and g users is defined by a matrix $\mathbf{B} \in \mathbb{R}^{m \times g}$ where $\mathbf{B}_{ij} = 1$ if listener i ($i = 1, 2, \ldots, m$) has a conversation with user j ($j = 1, 2, \ldots, g$). Then we define $\mathbf{P}^{(l)} = \mathbf{B}\mathbf{B}^T \in \mathbb{R}^{m \times m}$ and $\mathbf{P}^{(u)} = \mathbf{B}^T\mathbf{B} \in \mathbb{R}^{g \times g}$ as the adjacency matrices of the listener and user projection

	Bipartite network	Users (Members and Guests)	Listeners
$\mid V \mid$	1,430,319	1,340,972	89,347
$\mid E \mid$	3,078,158	294,057,991	40,600,172
Average degree	4.30	438.57	908.82
Clust. coeff.	NA	0.85	0.57
Assortativity	NA	0.62	0.47
Density	NA	3×10^{-4}	1.02×10^{-2}
Components	2,518	2,518	2,518
GCC size	1,423,965 (99.55%)	1,337,145 (99.71%)	86,820 (97.17%)

Table 4: Structural properties of the overall bipartite network and of the user and listener projections.

networks, respectively. In these matrices, $\mathbf{P}_{ij}^{(l)} = c$ ($\mathbf{P}_{ij}^{(u)} = c$) if listeners (users) i and j hold a conversation with c common users (listeners).

Structural properties and statistics of the projections are summarized in Table 4. User projection is characterized by a lower degree (i.e., 438.57) and larger clustering coefficient (i.e., 0.85) with respect to listener projection, implying a weak penchant for users to form clusters by the common listeners they connect to. We also examine the degree distributions of the user and listener projection networks on log-log scale in Figure 6 (note that Figure 1 also represents the degree distribution of the bipartite network).



(a) User projection

(b) Listener projection

Fig. 6: Degree distributions of projection networks.

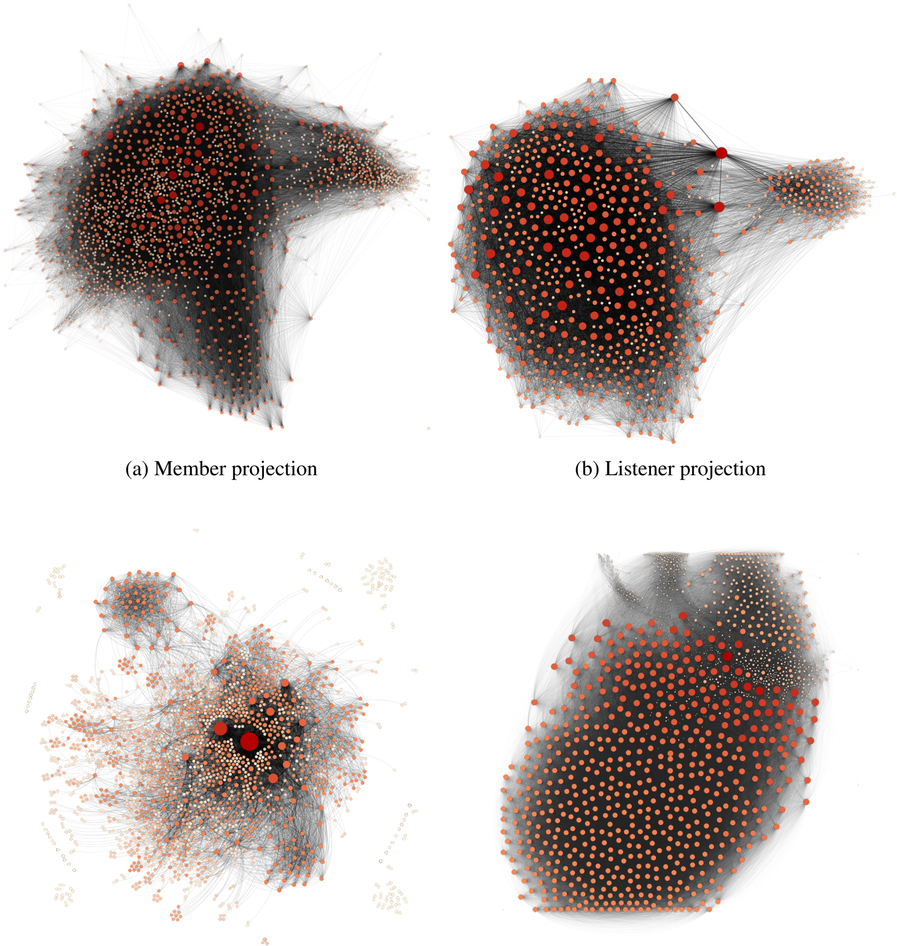
The listener degree distribution matches something like a double-Pareto log-normal distribution, while user projection degree distribution exhibits power-law behavior except for nodes with degree of order 10^3 that bucks the trend. This phenomenon may be a product of groups of users who are almost fully connected to each other in the projection network. Such components would represent a subset of users that have all connected to nearly the same set of listeners on 7 Cups. These listeners may indicate in their profiles their specialization for help with specific kinds of problems. The high degree assortativity for both projections also indicates an inclination for users (listeners) who share many common listeners (users) with others, share them with those who have a large number of listeners or users in common with others.

To highlight peculiarities differentiating member and guest behavior, we study separately their interactions with listeners. Table 5 gives the structural properties of the member and guest bipartite networks and their corresponding one-mode projections. Density values indicate that both bipartite networks and their projections are sparse. As for the entire network (see Table 4), the high mean degree and large GCC size of member projection networks support the hypothesis that members and listeners do not limit themselves to interact with a small subset of listeners or members, respectively. On the contrary, the small value of the average degree of the guest bipartite network may be due to the presence of a high number of guests holding one conversation, that is, communicating with one listener. For the same reason, we identify in listener projection network more than 6,200 nodes with degree 0 corresponding to listeners who conversed with very few different guests, on average 1.6 each. In addition, a very high number of disconnected components (6,327) characterizes guest interactions, the largest of which includes 98.42% of the guests and listeners.

	Members			Guests		
	Bipartite network	Member projection	Listener projection	Bipartite network	Guest projection	Listener projection
$\mid V \mid$	380,037	297,151	82,886	1,126,206	1,043,821	82,385
$\mid E \mid$	1,355,604	55,051,363	34,698,317	1,722,554	100,343,247	7,808,959
Avg. degree	7.13	370.53	837.25	3.06	192.26	189.57
Clust coeff.	NA	0.73	0.60	NA	0.89	0.53
Assortativity	NA	0.47	0.46	NA	0.75	0.61
Density	NA	1×10^{-3}	1×10^{-2}	NA	1.8×10^{-4}	2.3×10^{-3}
Components	1,413	1,413	1,413	6,327	6,327	6,327
GCC size	376,991 (99.1%)	295,530 (99.45%)	81,461 (98.28%)	1,108,426 (98.42%)	1,032,416 (98.90%)	76,010 (92.26%)

Table 5: Structural properties of member and guest bipartite networks and of the corresponding one-mode projections.

Figures 7(a) and 7(b) show the projection networks of the member bipartite network. Nodes are drawn under the Fruchterman Reingold layout and sized proportionally to their degree. Note that due to the network size, the visualizations represent a 10% random edge sampling of the projection networks. The member projection network exhibits two distinct groups of members not unlike the bipartite network of Figure 5. The small group of members who are weakly connected to the larger group generally corresponds to teen users who primarily connect to the subset of listeners for teens. The small number of members realized by the nodes between the two large groups may thus signify those who connect to listeners for both teens and adults, implying that most members do seek out a listener appropriate to their age group.



(a) Member projection

(b) Listener projection

(c) Guest projection

(d) Listener projection

Fig. 7: One-mode projections of member-exclusive [(a), (b)] and guest-exclusive [(c), (d)] bipartite networks. Thicker edges show more common listeners (members/guests) between members/guests (listeners).

Members seen in the periphery of the network visualization are those who connected to a very small number of listeners, thus limiting the number of other members they may share a listener with. A similar structure, further reflecting the difference between teen and adult communities on 7 Cups, is seen in the listener projection network. A key difference are listeners with particularly high degree in the projection network that lie between the communities of adult and teen listeners. Such listeners may support a mixture of adults and teens, or are popular listeners that members seek out (perhaps due to a very high account level or numerous accomplishment badges).

Figures 7(c) and 7(d) plot edge sampled projection networks of guests conversations under a Force Atlas layout so that nodes with higher connectivity to each other are positioned closer together.

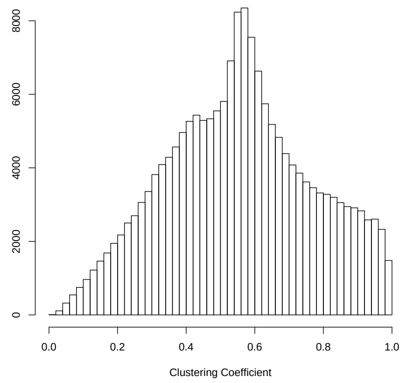
The node size and the color of the node are proportional to their degree, while edge width is proportional to its weight in the projection matrix. The numerous nodes on the periphery of the projection network represent guests who had few common listeners with others, which is in alignment with the small and zero degree nodes observed in the statistics of the listener projection network in Table 5. Two guest nodes with exceptionally high degree also stand out. We hypothesize, but are unable to verify with the data we have available, that these two anomalous guests may have been some kind of spam account or penetration test carried out by 7 Cups. The listener projection network is well connected where most listeners have connected to a similar number of other guests. This is a sign that most guests select an arbitrary listener to connect to or that 7 Cups automatically pairs guests with a listener in a random fashion.

We finally observe that the large global clustering coefficient (see Table 5), characterizing projection networks for both members and guests, denote the existence of numerous transitive relationships. Looking at local clustering coefficient in Figure 8, we note that only 10.3% of listeners but more than 72% of guests have a clustering coefficient equal to one. On the contrary, only 12.5% of listeners and 38.2% of members have clustering coefficient equal to one, meaning that the majority of guests but only a small portion of members have fully connected neighbors. We further note that the clustering coefficient distributions for the member projection networks are shaped somewhat like a normal distribution. Normally distributed cluster coefficient distributions are usually seen in co-occurrence networks [19,31,51], but the irregularity of the distribution's shape and the spike of listeners in the member projection network where the clustering coefficient is 1 is unique to 7 Cups. This may be a sign of some innate process on 7 Cups that is not related to the fact that these projection networks are simply another type of co-occurrence network. For example, members may selectively connect to the same pool of listeners who may have similar ratings, experiences, or bios that indicate a common expertise. This may lead to a large set of members connecting to this pool of listeners leading to a spike in the number of listeners with maximum clustering coefficient.

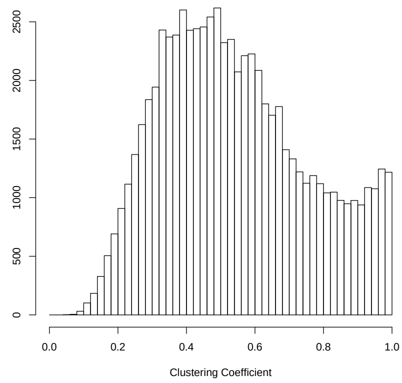
These results suggest interesting future developments aimed at better understanding the effects of network topologies and the many ways community structures influence user behavior.

6.2 Evolutionary analysis

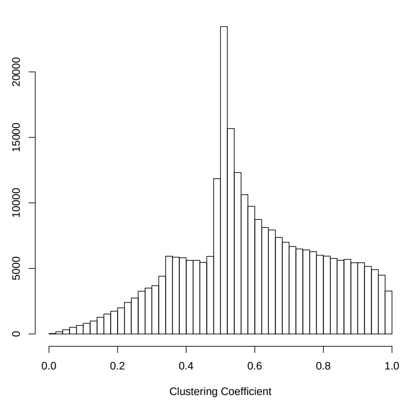
We further examine how conversations develop across 7 Cups and whether the growth of the network represents a preferential attachment process [38] where connections are more likely to be established with a node that already exhibits relatively high degree. For this evaluation, we empirically compute the probability $p_e(d)$ that a member



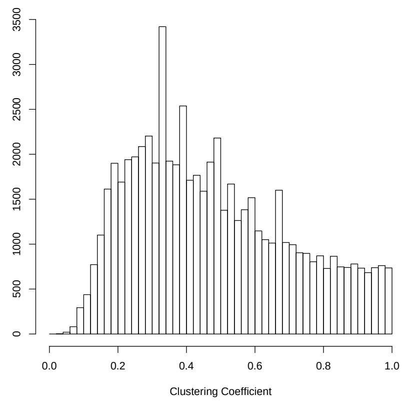
(a) Member projection



(b) Listener projection



(c) Guest projection



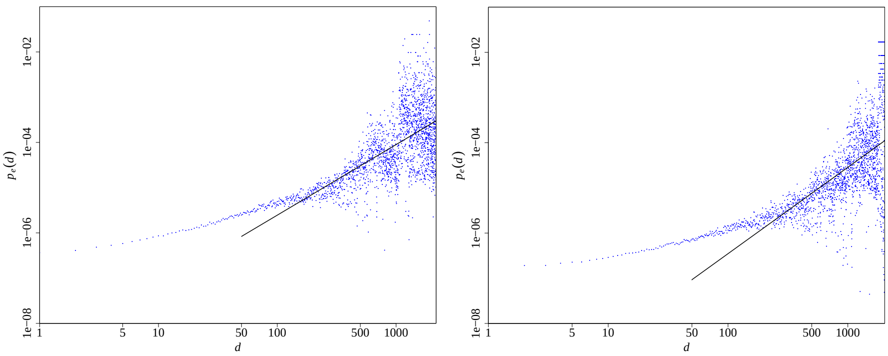
(d) Listener projection

Fig. 8: Clustering coefficient distributions of member and guest one-mode projection networks. (b): projection for member network. (d): projection for guest network.

or guest u at time t will hold a conversation with listener l having degree d as:

$$p_e(d) = \frac{\sum_t \mathbf{1}(e_t = (u,l))\mathbf{1}(d_{t-1}(l) = d)}{\sum_t \mid \{l : d_{t-1}(l) = d\} \mid}$$

where e_t denotes a conversation created at time t, $d_{t-1}(l)$ the degree of listener l at time $t-1$, and $\mathbf{1}(\cdot)$ is the indicator function. Figure 9 shows, on log-log scale, the probability $p_e(d)$ for members and guests to connect to a listener characterized by a degree equal to d. It illustrates the emergence of a rich-get-richer phenomenon where the probability that an edge (conversation) is added to a node (listener) of



(a) Members

(b) Guests

Fig. 9: Probability a member or guest will start a conversation with a listener having degree d.

degree d is proportional to d^α. Specifically, for all degrees smaller than 100, we measure $\alpha = 0.66$ for members and $\alpha = 0.52$ for guests. For larger degree nodes we find $\alpha = 1.27$ and 1.91 for members and guests, respectively, suggesting a super-linear preferential attachment. Hence, preferential attachment is somehow weaker for low degree nodes, whereas edges attach preferentially to higher degree nodes. Once listeners connect with a number of users they asymptotically become connected to all users on the platform as their time t on the site goes to infinity. This asymptotic result may not be that surprising: a listener who decides to connect to a large number of members and guests over time may be expected to eventually connect to all members and guests given enough time to do so.

It is also interesting to find preferential attachment to hold for both members and guests, although the phenomenon is more evident for guest conversations. This phenomenon may emerge because, in personal conversations, users select a listener based on a profile, which includes information about their experience and amount of activity on the site. Thus, it may be expected that a user will always choose a listener with more experience rather than one who has only helped a small number of others. However, as stated in Section 5, only one third of conversations are classified as personal, and this percentage is even lower (i.e. 25%) for guests conversations. Preferential attachment may be hence explained by the presence of a number of listeners who are often online and available to provide their support. This phenomenon could also be indicative of an underlying mechanism on 7 Cups that prefers to involve more experienced listeners when many are available at the same time.

6.3 Communication densification

We also examine whether the network of conversations among users *densifies* over time. Densification of the conversation network defines the degree to which new conversations are established compared to new users added over time. Figure 10 plots on log-log scale the number of nodes (users and listeners) against the number of edges per month in the conversation network. The behavior indicates that densification is occurring, as the ratio of the number of edges to nodes grows as $e(t) \propto n(t)^{\alpha}$ where $e(t)$ and $n(t)$ represent the number of edges and nodes of the graph at time t. We measure the densification exponent α to be equal to 1.1. This observation further underscores the users' desire to communicate with a number of other listeners on an emotional support system. It also indicates that the scalability of the system hinges not on the number of users it can support, but on the number of conversations it fosters.

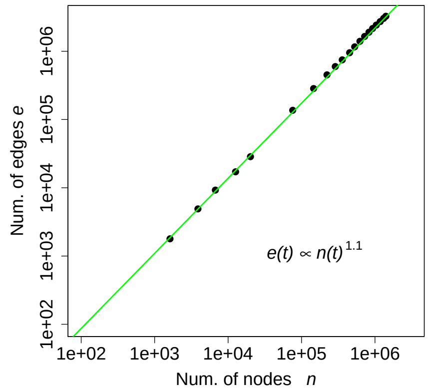


Fig. 10: Number of nodes and edges in the conversation network over time.

7 User Engagement

We next evaluate how members signed up for 7 Cups engage with the service. An engagement analysis tries to provide mechanisms and properties of the platform that encourage members to return, listeners to stay active, and for members to have multiple, fruitful conversations. Engagement analysis is practically important to a service for member retention and community growth, but also identifies qualities that encourage people to seek follow-up emotional support. Due to the proprietary nature and availability of the dataset this engagement analysis was only carried out for records between December 5th 2013 and November 18th 2014.

Coins	0.247	Growth Points	0.977
Compassion Hearts	0.243	Signup Date	-0.009
Last Login Date	0.133	Distress Level	0.004
Group Chat Msgs	0.120	Page Views (Web)	-0.002
Page View (iOS)	-0.001	Login Count	-0.001
Conversation Requests	0.001	Self Help Views	0.005
Forum Posts	-0.001	Forum Views	-0.001
Forum Upvotes	0.201		

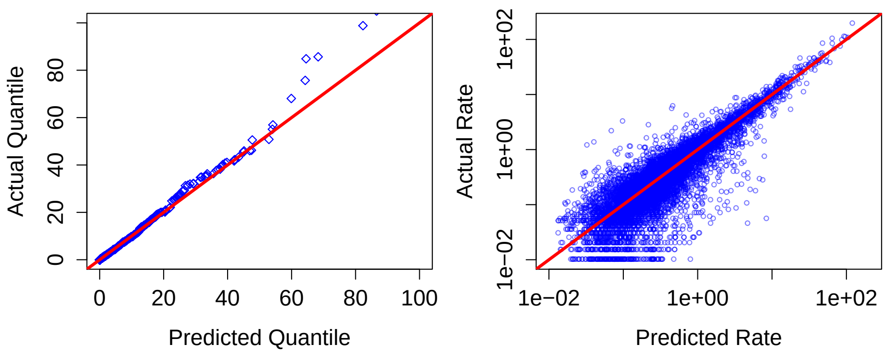
Table 6: Pearson correlation between message rate and user or behavior features

7.1 Factors driving engagement

We first study the features and behaviors of members and their relationship to measure of site engagement. Members sending messages to listeners is the principle purpose of 7 Cups, so we we quantify engagement as the *daily average message rate* of members in conversations. We examine the following behavioral and user features that could be related to engagement: (i) the number of coins, growth points, and compassion hearts, which are gaming and progress measures related to a members reputation and experience; (ii) the signup and last login date; (iii) the reported distress level when members register; (iv) the number of group chat messages; (v) the number of page views from the 7 Cups Web and iOS applications; (vi) the number of logins; (vii) the number of conversation requests sent; (viii) the number of self-help page views; (ix) the number of forum posts, views, and upvotes. Table 6 lists the Pearson correlation coefficient between each of the features listed above and the daily average message rate of members. The correlations show that gamification "rewards" (accumulated coins, hearts, and growth points) have the strongest relationship with user engagement. The volume of messages sent by members automatically increases their growth points, and so their 0.977 Pearson correlation is noteworthy. Behavior features unrelated to communication (signup and last login date, distress level, page views, and help article views) have almost no correlation.

Interactions among features may also be positively correlated with engagement. For example, users who exhibit a high distress level could initiate many conversation requests and thus have a high level of engagement even though the features are individually not correlated. To consider such feature interactions, we built a random forest model that predicts user engagement by a regression over all features. A random forest is an ensemble of decision trees trained over different bootstraps of the data. During training, each tree is limited to the use of distinct small subsets of features to make splitting decisions. If X_u is a vector of member u's features, the random forest predicts the engagement of u as $\hat{f}(X_u) = N^{-1} \sum_{i=1}^{N} \hat{f}_i(X_u)$ where $\hat{f}_i$ is the predicted engagement value from the i^{th} of N decision trees in the random forest. The bootstraps, limited choice of features for tree splitting, and result averaging across the tree ensemble ensure the forest does not overfit the data even for large N [21]. We compute the importance of each feature to the random forest regression model as follows: let $C = \sum_{i=1}^{m}(y_i - \hat{f}(X_i))^2$ be the mean square error (MSE) of the random forest predictions against the actual engagement y_i of every member i. The importance of feature ℓ may be found by randomly perturbing the values of ℓ across every member's feature vector. Let $X_i^{(\ell)}$ be the feature vector of member i whose ℓ^{th} element is perturbed and $C_\ell = \sum_{i=1}^{m}(y_i - \hat{f}(X_i^{(\ell)}))^2$ the MSE of the model using the perturbed vectors. Then the importance of ℓ may ranked by the percent increase in MSE between C and C_ℓ. For example, if feature ℓ is not important, the errors of the model will be less sensitive to a reshuffling of its values across all users.

We trained a random forest using 75% of the members with $N = 1000$ trees and randomly chose one third of the features for every tree splitting decision. Figure 11 presents the quantile and prediction scatter plots of the predicted and actual message rates for the 25% of users not used to train the random forest. The figure illustrates how well the random forest models engagement ($R^2 > 89\%$), with an especially tight linear relationship between the predicted and actual engagement distributions up to the 60^{th} percentile. The predicted vs. actual engagement rates in Figure 11 (b) exhibit small gaussian errors while predicting the engagement level of members with low engagement.



(a) Quantiles (b) Rates

Fig. 11: Random forest predicted vs. actual engagement

We also run a feature importance analysis using the random forest model. Figure 12 shows the percent increase in MSE of a random forest where each factor was individually perturbed across the training data. Like Table 6, the number of growth points of a member is considered to be the most important predictor for engagement, but message rate and user growth points accumulate together on the site. The signup date and last login date of a member are seen to be the next most important features that predict engagement, as each feature increases the model's MSE by over 20% when they are perturbed. This is followed by the number of messages sent in group chat s as an important feature not related to gamification mechanisms on the site. Participating in group chats may thus encourage users to progress to one-on-one conversations, or to become more active in the one-on-one conversations they have already started. We also see how the volume of upvotes a member has on the forum actually introduces noise in the model, since perturbing this factor decreases the MSE of the random forest. One explanation may be that members who gain recognition for forum posts become disinclined to invest time in participating in other communication modes.

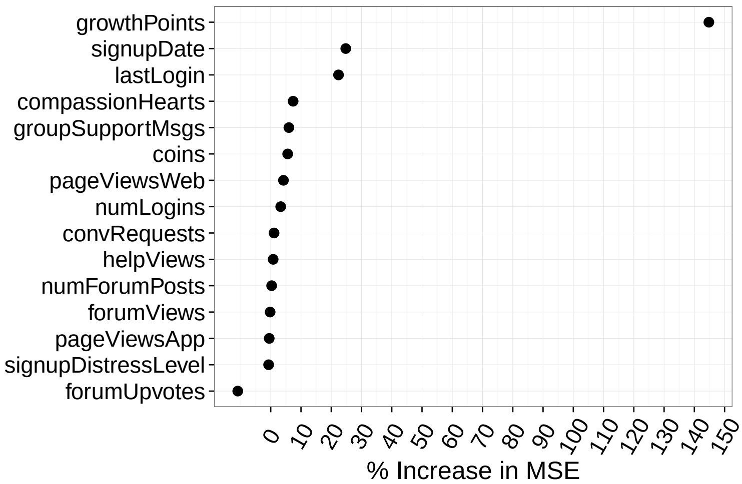
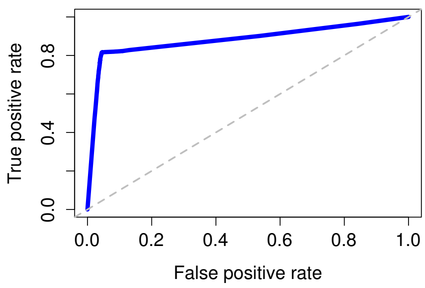


Fig. 12: Feature importance in the random forest model

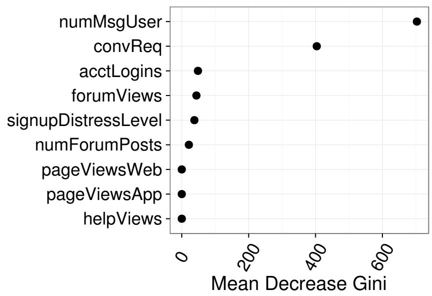
7.2 New user engagement prediction

New and unengaged members may have a propensity to become 'inactive' or never return to an emotional support system. Early identification of such users is helpful to a service operator to identify those users who need further encouragement, incentives, or other reasons to return to the site to continue receiving emotional support. We thus built another random forest classifier that identifies if members, based on actions during their first two weeks on 7 Cups, will become active users. We consulted with 7 Cups staff to define an active user as one who: (i) has been registered for at least six weeks; and (ii) has performed at least two actions on the service over the past month. A 'new user' is one who has registered within the last two weeks. We identified all members who registered between May 7th 2014 (the first date user action data was recorded) and November 18th 2014 and mark them as 'active' or 'inactive'. We then collect the following actions from their first two weeks on the site: (i) number of conversation requests; (ii) number of messages sent; (iii) number of forum posts made and viewed; (iv) number of logins performed; (v) number of help page views; and (vi) number of site pages accessed via 7 Cups's Website and iOS app.

52,803 members registered on 7 Cups during the time period considered, of which 11,117 (21%) became active and 41,686 (79%) became inactive. We built a training set from a random sample of 66% of members to predict if they are active. Because



(a) ROC curve



(b) Gini Coefficient

Fig. 13: Active user classifier evaluation

of the nearly 20/80 imbalance in the number of inactive and active members in the data the training set was sampled from, the minority class is randomly oversampled so that equal numbers of inactive and active cases are provided for training [21]. The test set is composed of the remaining 33% of users. The classifier achieves a very promising accuracy of 92.5% and the ROC curve in Figure 13 (a) suggests a small to moderate false positive rate.

As before, we assess the importance of the factors used for predicting active users. MSE is incompatible with the notion of a binary classification decision, so we consider the Gini index [21] instead. The Gini index of a decision tree node t is defined by $G_t = p_{ta}(1 - p_{ta}) + p_{ti}(1 - p_{ti})$ where p_{ta} (p_{ti}) is the proportion of members marked active (inactive) that fall into node t based on the splitting criteria of its parent node. A near zero Gini index suggests that the splitting rule at the parent of a tree node divides the data into separate classes almost perfectly, which is a property of strong decision tree classifiers. We thus rank the importance of a factor by the average decrease of the Gini coefficient across all splits in all trees of the forest in Figure 13 (b). The figure shows how the volume of messages sent and conversation requests submitted during the first two weeks best predict whether a user will become active. Figure 13 (b) further illustrates that the number of account logins performed, the member's self-reported distress level, and any activity on the online forums are not significant predictors of becoming an active user.

8 Suspicious Behaviors

Finally, we study the degree to which 7 Cups users behave *suspiciously*, which means they demonstrate malicious, rude, or behaviors that go against the purpose of the site and hence are suspicious.

Users	Count	Percentage
Suspicious Members	19,281	6.49%
Active Suspicious Members	15,305	5.15%
Affected Listeners	37,262	41.7%

Table 7: Number of suspicious members found and exposure to listeners

8.1 Finding suspicious users

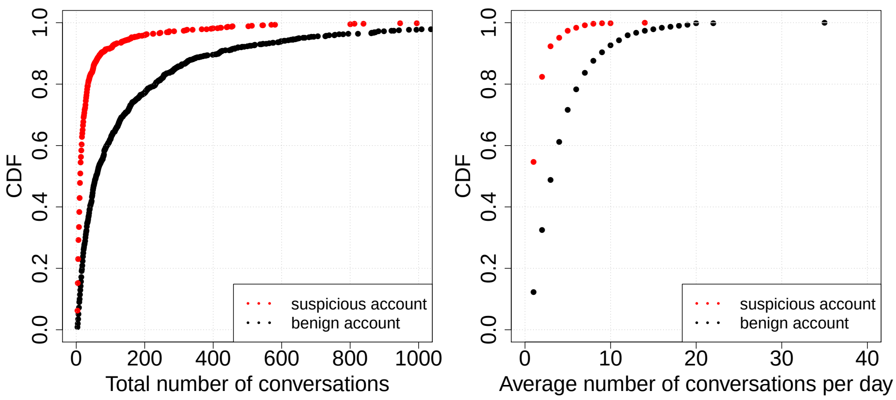
7 Cups provides a mechanism for users to report another if they witness any form of harassment, belittling, or other text that causes discomfort. Two types of actions either a *block* or a *ban*, are taken against misbehaving members by the administrators of 7 Cups depending on the frequency and the severity of reports against an individual. A block means the member can no longer have a one-on-one conversation with the listener who submitted the block action. Blocks are applied automatically, so that an offended user can immediately cut off contact with the offender. A ban is a more severe restriction imposed by 7 Cups administrators who stop the banned member from using the service. We say that any member who has been blocked or banned is *suspicious*, in the sense that there is evidence of demonstrating malicious or rude behaviors. In this section, we quantitatively explore the features of suspicious members, and then develop a classifier able to reliably detect if an unsuspicious user will become one in the future. The ability to detect such users may be important for 7 Cups and other online emotional support systems to be a safe place for listeners to volunteer their time and effort.

15,305 active members have been blocked from at least one conversation or banned. We further counted the number of listeners who hold a conversation with these members and find that 41.7% of them were exposed in a one-on-one conversation. This statistic underscores the need to be able to identify and halt suspicious behavior by users to lower their exposure to listeners. This information is summarized in Table 7.

8.1.1 Feature evaluation

We compare the behavior of suspicious members against benign ones to better understand their contrasting qualities. Since benign members seek emotional relief, they may focus on establishing a small number of listeners who they trust to express their concerns with. On the other hand, suspicious members may solicit a large number of listeners to create conversations without a specific purpose. Some listeners may also realize the actual intentions of suspicious members and consequently refuse to respond or terminate any ongoing conversations. We therefore hypothesize that benign members have high-quality conversations with a limited set of listeners, while suspicious members may create a large number of conversations to promote their exposure and to harass others without regard. We thus contrast the total number of conversations and average number of conversations per day between suspicious and benign members to test this hypothesis in Figure 14. As shown in the distribution of Figure 14 (a), more than 60% suspicious members have total conversations higher than

50 while more than 96% benign members have less than 50. Figure 14 (b) shows that nearly 90% of benign members participate in less than four conversations per day. In contrast, more than 50% of suspicious members have more than four conversations per day.



(a) Total conversations (b) Average conversations

Fig. 14: CDF of the total number of conversations held by users (a) and of the average number of conversations held per day (b)

We also examine the number of personal conversations of a member. In these conversations, members can search for specific listeners based on criteria like age, country, and keywords contained in a listener profile. While both benign and suspicious members have the freedom to identify their preferred listeners, suspicious members may seek to identify particular targets for their harassment attacks. For example, a member who plans to conduct sexual harassment may infer the gender of a potential target listener based on the portraits (or avatar icons) uploaded by listeners. Figure 15 presents the distribution of personal conversations for suspicious and benign members, showing that suspicious members tend to generate more conversations compared to benign members. More than 65% of suspicious members held more than 20 personal conversations, while only about 10% of benign accounts held more than 20 personal conversations.

Conversational behaviors are another possible distinction between benign and malicious members. Different from typical online chatting systems, the state of a conversation of 7 Cups is maintained even if the member or listener in this conversation logs out or disconnects. We thus consider the possibility that a suspicious member may be inclined to be more active in sending messages to attract listeners' attention or to disturb their normal consulting activities. We also consider the reasons for the termination of a conversation. Both the member and listener can re-engage themselves in previously-established conversations to continue discussion, thus fos-

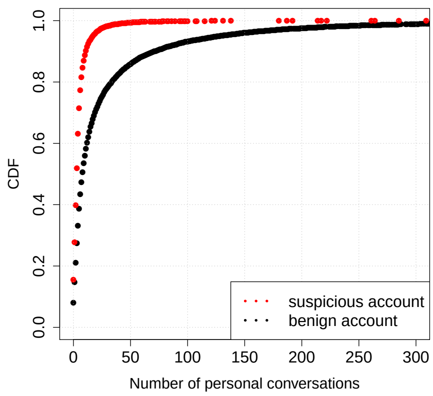
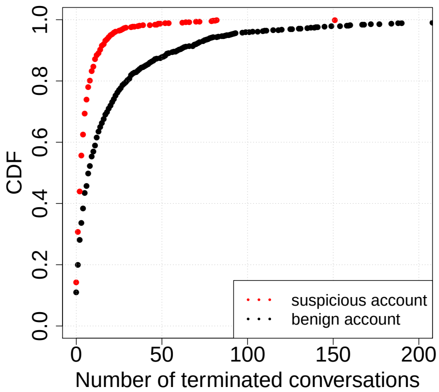


Fig. 15: CDF of the number of personal conversations

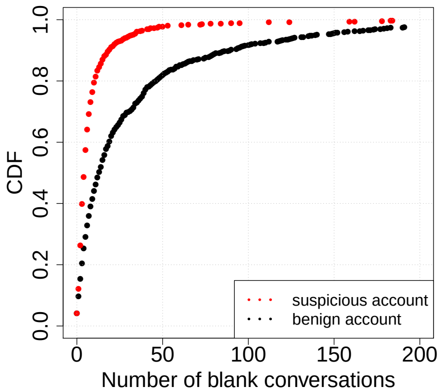
tering long-term, trustworthy, and high-quality interaction between the listener and member. A conversation between a listener and a benign member may be terminated if, for example, the member has obtained sufficient mental support from the listener, or if the present listener cannot satisfy the needs of the member. Conversation terminations are thus presumed to be a rare event as it takes a significant amount of interactions (and thus time) for a benign member to be adequately relieved. In contrast, conversations between a listener and a suspicious member may be terminated more frequently. To check this, Figure 16 (a) presents the distribution of this feature for both benign and suspicious members.

We also consider the conversational behavior of *silence*. When a benign member initiates a conversation, the dialog usually starts with greetings from the member and a response from the listener. A suspicious member, however, may start conversations with no message exchanged. Blank conversations may be an indication of a type of denial of service (DoS) attacks by spamming listeners with the hope of wasting their time and energy, and to leave the impression that most members on 7 Cups are not serious about obtaining help. A suspicious member may also start many conversations simultaneously with a collection of listeners. When suspicious members get a response from a listener, they may stop greeting the following ones, thus introducing blank conversations. To investigate this, Figure 16 (b) gives the distribution of blank conversations for benign and suspicious members. Approximately 80% benign members held less than 20 blank conversations, while over 60% suspicious members initiated more than 20 blank conversations and about 10% of suspicious members in fact generated over 100 blank conversations.

Finally, we examine features of benign and suspicious members that correspond to reputation: the volume of blocks submitted and gamification points for participating on the website. Users are provided with the ability to block listeners should they behave offensively. Members are much more likely to be blocked compared to listeners since members are those seeking help on the service whereas listeners are



(a) Terminated conversations



(b) Blank conversations

Fig. 16: CDFs of the number of terminated (a) and blank (b) conversations

volunteering their time and energy for sake of supporting others, and hence is less likely to share harassing or offensive messages to members. Still, the act of intentionally blocking a listener can be seen as an abusive action. For example, when suspicious members find that their dubious behavior is identified by a listener, they could proactively block the listener to either disrupt the reporting process or to confuse an administrator and avoid a ban. We thus examine the distribution of the number of listeners blocked by members in Figure 17. Indeed, we see that 98% of benign members never blocked any listener, while almost 70% of suspicious members blocked at least one listener.

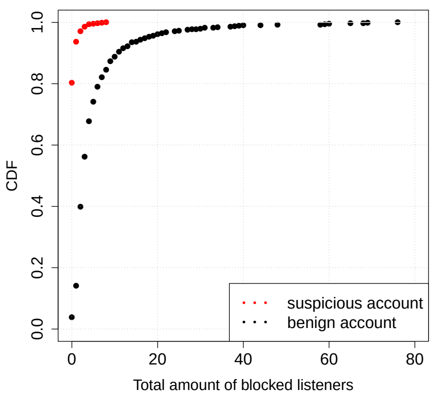


Fig. 17: CDF of the number of listeners blocked by members

Another reputation-based feature is the accumulation of gamification points. Suspicious members who are more likely to frequently establish new conversations and perform specific site actions can easily acquire larger levels of gamification points, while benign members may have a limited point accumulation by focusing on activities that supporting their health (e.g., by only posting on a forum or speaking to the same small set of listeners). We thus expect that suspicious members will tend to have a higher volume of gamification points compared to benign members. This is confirmed in Figure 18. where approximately 90% of benign users have growth points less than 2,000, 50% of suspicious accounts accumulate more than 2,000 growth points.

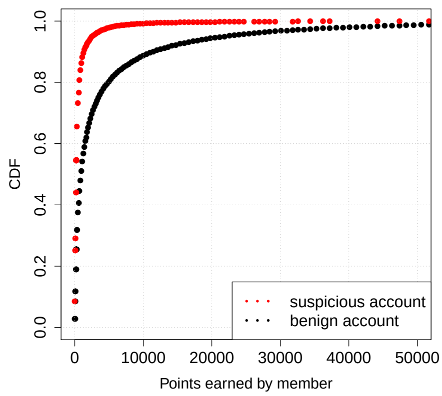


Fig. 18: CDF of member level

8.1.2 Detecting suspicious members

We trained multiple classifiers to identify suspicious users from a dataset of 15,305 suspicious members and a random sampling of 15,305 benign members. Our evaluation focuses on detection accuracy and any correlations between features and accuracy. We consider three different classifiers including a Random Forest, Support Vector Machine [13], and Gradient-Boosted Tree [20]. A 10-fold cross-validation is applied with one fold held out to test the detector. Figure 19 visualizes the receiver operating characteristic (ROC) for a trained Random Forest to illustrate the trade-off between the detection rate and the false positive rate. Encouragingly, the the detector can achieve a suspicious user detection rate of 77.8% if a small tolerance of a 1% false positive rate is allowed.

We further investigate the relative importance of the proposed features in the context of Random Forest classifier with the same feature permutation test discussed in Section 7.1. The rank of features based on the variable importance is shown in Table 8, where we can find that the number of blocked listeners, the member's level, and

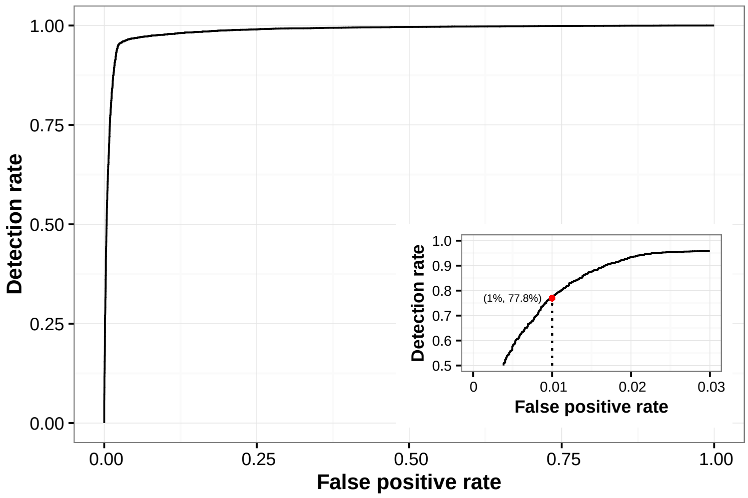


Fig. 19: ROC curve on eight features

Feature	Importance
Number of Blocked Listeners	815.4
Member Level	131.8
Average Number of Conversations per Day	80.6
Total Number of Conversations	49.4
Number of Black Conversations	46.9
Average Daily Messages Sent	43.1
Number of Terminated Connections	36.6
Number of Personal Conversations	33.5

Table 8: Feature importance using random forest to identify suspicious members

the average number of conversations per day most strongly drive suspicious member detection. It is interesting to note that the two reputation features are most important in making a classification decision. In fact, the number of listeners a member has chosen to block proves to the strongest indicator of a suspicious member by far. This lends support to our hypothesis that suspicious users can obtain a quick sense of when a punitive action will be taken against them, blocking the connection prematurely. It also suggests a trolling attack where members begin a conversation, launch a verbal attack or harassment, and then quickly block the listener to terminate the conversation and prevent feedback. Moreover, benign users with good intentions may never find a need to block a listener since conversations persist on 7 Cups. The importance of the level of a member also matches the intuition that longstanding members who contribute to the community positively are very low risk for becoming a future suspicious member.

9 Conclusions

This paper documents a comprehensive analysis of the behaviors, interactions, and engagement of users on a large-scale online emotional support service. We conclude by summarizing key insights identified throughout:

- That one-on-one conversations between members and listeners dominate, although they also take advantage of group chats and forums to communicate broadly;
- That members more often need to speak with *anyone*, rather than *someone* with a particular topic expertise;
- That a "guest" membership where users do not need to register or sign-up to receive emotional support is an important way to ease new users onto the site;
- That it is possible to cultivate a large and active community of volunteers to provide on-demand support for those in need;
- That members have a tendency to seek out listeners who are appropriate for their age group;
- That members desire communication with a large number of other listeners, and scalability hinges on a system that can support high volumes of conversations (rather than users);
- That gamification mechanisms and group chat participation are correlated with user engagement;
- That behavioral features unrelated to communication with others have almost no correlation with engagement;
- That a large proportion of volunteer listeners are exposed to a relatively small proportion of members who claim they seek help;
- That the number of blocked listeners, the level of a member, and the average number of conversations a member holds are important features for determining if a member is abusive to listeners.

Our insights are based on mechanisms common to most online social services, including gamification aspects to encourage engagement and the use of a crowd of volunteer listeners that enable the service to scale to large user bases. We believe this makes our findings generalizable to other types of online emotional support services exhibiting similar characteristics.

References

1. Attfield, S., Kazai, G., Lalmas, M., Piwowarski, B.: Towards a science of user engagement (position paper). In: Proc. WSDM workshop on user modelling for Web applications, pp. 9–12 (2011)
2. Bar-Lev, S.: "We are here to give you emotional support": Performing emotions in an online HIV/AIDS support group. Qualitative health research **18**(4), 509–521 (2008)
3. Barak, A.: Emotional support and suicide prevention through the internet: A field project report. Computers in Human Behavior **23**(2), 971–984 (2007)
4. Binik, Y.M., Cantor, J., Ochs, E., Meana, M.: From the couch to the keyboard: Psychotherapy in cyberspace. Culture of the Internet pp. 71–100 (1997)
5. Booth, D., Jansen, B.J.: A review of methodologies for analyzing websites. In: Web technologies: Concepts, methodologies, tools, and applications, pp. 145–166. IGI Global (2010)
6. Bright, J.I., Baker, K.D., Neimeyer, R.A.: Professional and paraprofessional group treatments for depression: A comparison of cognitive-behavioral and mutual support interventions. Journal of consulting and clinical psychology **67**(4), 491 (1999)

7. Calzarossa, M., Della Vedova, M., Massari, L., Nebbione, G., Tessera, D.: A methodological approach for time series analysis and forecasting of web dynamics. LNCS Transactions on Computational Collective Intelligence XXXIII **11610**, 128–143 (2019)
8. Calzarossa, M.C., Massari, L., Doran, D., Yelne, S., Trivedi, N., Moriarty, G.: Measuring the Users and Conversations of a Vibrant Online Emotional Support System. In: Proc. of the IEEE Symposium on Computers and Communication (ISCC), pp. 1193–1199 (2016)
9. Calzarossa, M.C., Tessera, D.: Modeling and Predicting Temporal Patterns of Web Content Changes. Journal of Network and Computer Applications **56**, 115–123 (2015)
10. Chen, Y.R., Chen, H.H.: Opinion spammer detection in web forum. In: Proc. of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 759–762. ACM (2015)
11. Chou, W.y.S., Prestin, A., Lyons, C., Wen, K.y.: Web 2.0 for health promotion: reviewing the current evidence. American journal of public health **103**(1), e9–e18 (2013)
12. Chu, Z., Gianvecchio, S., Koehl, A., Wang, H., Jajodia, S.: Blog or block: Detecting blog bots through behavioral biometrics. Computer Networks **57**(3), 634–646 (2013)
13. Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines and other Kernel-based Learning Methods. Cambridge University Press (2000)
14. Davison, K.P., Pennebaker, J.W., Dickerson, S.S.: Who talks? The social psychology of illness support groups. American Psychologist **55**(2), 205 (2000)
15. Doran, D., Yelne, S., Massari, L., Calzarossa, M., Jackson, L., Moriarty, G.: Stay Awhile and Listen: User Interactions in a Crowdsourced Platform Offering Emotional Support. In: Proc. of the IEEE/ACM Int. Conf. on Advances in Social Networks Analysis and Mining (ASONAM), pp. 667–674. ACM (2015)
16. Fakhraei, S., Foulds, J., Shashanka, M., Getoor, L.: Collective spammer detection in evolving multi-relational social networks. In: Proc. of the 21st ACM International Conference on Knowledge Discovery and Data Mining, pp. 1769–1778 (2015)
17. Fortunato, S.: Community detection in graphs. Physics Reports **486**(3), 75–174 (2010)
18. Gao, H., Hu, J., Wilson, C., Li, Z., Chen, Y., Zhao, B.Y.: Detecting and characterizing social spam campaigns. In: Proc. of the 10th ACM SIGCOMM Internet Measurement Conference (IMC), pp. 35–47 (2010)
19. Grando, F., Granville, L., Lamb, L.: Machine learning in network centrality measures: Tutorial and outlook. ACM Computing Surveys **51**(5), 102:1–102:32 (2018)
20. Han, J., Kamber, M., Pei, J.: Data mining: Concepts and techniques. Morgan Kaufmann (2006)
21. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer (2009)
22. Hemmati, A., Chung, K.S.K.: Associations between personal social network properties and mental health in cancer care. In: Proc. of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 828–835 (2014)
23. Hu, X., Tang, J., Liu, H.: Online social spammer detection. In: Proc. of the 28th AAAI Conference on Artificial Intelligence, pp. 59–65 (2014)
24. Hu, Y.: Efficient, high-quality force-directed graph drawing. Mathematica Journal **10**(1), 37–71 (2005)
25. Huang, M.P., Alessi, N.E.: The internet and the future of psychiatry. American Journal of Psychiatry **153**(7), 861–869 (1996)
26. Ikehara, C.S., Crosby, M.E.: Assessing cognitive load with physiological sensors. In: Proc of 38th Annual Hawaii International Conference on System Sciences. IEEE (2005)
27. Jacques, R.D.: The nature of engagement and its role in hypermedia evaluation and design. Ph.D. thesis, South Bank University (1996)
28. Kayes, I., Iamnitchi, A.: Privacy and security in online social networks: A survey. Online Social Networks and Media **3-4**, 1 – 21 (2017)
29. Konradt, U., Sulz, K.: The experience of flow in interacting with a hypermedia learning environment. Journal of educational multimedia and hypermedia **10**(1), 69–84 (2001)
30. Lalmas, M., O'Brien, H., Yom-Tov, E.: Measuring user engagement. Synthesis Lectures on Information Concepts, Retrieval, and Services **6**(4), 1–132 (2014)
31. Leskovec, J., Kleinberg, J., Faloutsos, C.: Graph evolution: Densification and shrinking diameters. ACM Transactions on Knowledge Discovery from Data **1**(1) (2007)
32. Li, Y., Kim, D.W., Zhang, J., Doran, D.: Teafilter: Detecting suspicious members in an online emotional support service. In: Proc. EAI International Conference on Security and Privacy in Communication Networks (2018)

33. Lupien, S.J., McEwen, B.S., Gunnar, M.R., Heim, C.: Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nature reviews neuroscience **10**(6), 434 (2009)
34. Malliaros, F., Vazirgiannis, M.: Clustering and community detection in directed networks: A survey. Physics Reports **533**(4), 95–142 (2013)
35. Maloney-Krichmar, D., Preece, J.: A multilevel analysis of sociability, usability, and community dynamics in an online health community. ACM Transactions on Computer-Human Interaction **12**(2), 201–232 (2005)
36. Mislove, A., Marcon, M., Gummadi, K.P., Druschel, P., Bhattacharjee, B.: Measurement and analysis of online social networks. In: Proc. of the 7th ACM SIGCOMM conference on Internet measurement, pp. 29–42. ACM (2007)
37. Newman, M.E.J.: The structure and function of complex networks. SIAM Review **45**(2), 167–256 (2003)
38. Newman, M.E.J.: Networks: An Introduction. Oxford University Press (2010)
39. Newman, M.W., Lauterbach, D., Munson, S.A., Resnick, P., Morris, M.E.: It's not that I don't have problems, I'm just not putting them on Facebook: Challenges and Opportunities in Using Online Social Networks for Health. In: Proc. of the ACM 2011 Conference on Computer Supported Cooperative Work, pp. 341–350. ACM (2011)
40. Ploderer, B., Smith, W., Howard, S., Pearce, J., Borland, R.: Patterns of support in an online community for smoking cessation. In: Proc. of the 6th International Conference on Communities and Technologies, pp. 26–35. ACM (2013)
41. Procidano, M.E., Heller, K.: Measures of Perceived Social Support From Friends and From Family: Three Validation Studies. American Journal of Community Psychology **11**(1), 1–24 (1983)
42. Rochlen, A.B., Zack, J.S., Speyer, C.: Online therapy: Review of relevant definitions, debates, and current empirical support. Journal of clinical psychology **60**(3), 269–283 (2004)
43. Seah, M.l., Cairns, P.: From immersion to addiction in videogames. In: Proc. of British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction, pp. 55–63. British Computer Society (2008)
44. Stanton-Salazar, R.D., Spina, S.U.: Adolescent peer networks as a context for social and emotional support. Youth & society **36**(4), 379–417 (2005)
45. Tsai, Y.C., Liu, C.H.: Factors and symptoms associated with work stress and health-promoting lifestyles among hospital staff: a pilot study in taiwan. BMC health services research **12**(1), 199 (2012)
46. Wang, Y.C., Kraut, R., Levine, J.M.: To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups. In: Proc. of the ACM 2012 Conference on Computer Supported Cooperative Work, pp. 833–842. ACM (2012)
47. Webster, J., Ho, H.: Audience engagement in multimedia presentations. ACM SIGMIS Database: the DATABASE for Advances in Information Systems **28**(2), 63–77 (1997)
48. White, M., Dorman, S.M.: Receiving social support online: implications for health education. Health education research **16**(6), 693–707 (2001)
49. Wu, F., Shu, J., Huang, Y., Yuan, Z.: Social spammer and spam message co-detection in microblogging with social context regularization. In: Proc. of the 24th ACM International on Conference on Information and Knowledge Management, pp. 1601–1610. ACM (2015)
50. Yang, C., Harkreader, R.C., Gu, G.: Die free or live hard? Empirical evaluation and new design for fighting evolving Twitter spammers. In: Proc. of the International Workshop on Recent Advances in Intrusion Detection, pp. 318–337. Springer (2011)
51. Yang, J., Leskovec, J.: Defining and evaluating network communities based on ground-truth. Knowledge and Information Systems **42**(1), 181–213 (2015)
52. Yuen, E.K., Goetter, E.M., Herbert, J.D., Forman, E.M.: Challenges and opportunities in internet-mediated telemental health. Professional Psychology: Research and Practice **43**(1), 1–8 (2012)
53. Zuckerman, E.: Finding, evaluating, and incorporating internet self-help resources into psychotherapy practice. Journal of Clinical psychology **59**(2), 217–225 (2003)