The Emergence of Crowdsourcing among Pokémon Go Players

Priscila Martins
Federal University of Minas Gerais
Belo Horizonte, Minas Gerais
primsouza@dcc.ufmg.br

Manoel Júnior
Federal University of Minas Gerais
Belo Horizonte, Minas Gerais
manoelrmj@dcc.ufmg.br

Fabrício Benevenuto
Federal University of Minas Gerais
Belo Horizonte, Minas Gerais
fabricio@dcc.ufmg.br

Jussara Almeida
Federal University of Minas Gerais
Belo Horizonte, Minas Gerais
jussara@dcc.ufmg.br

ABSTRACT

Since its launching, Pokémon Go has been pointed as the largest gaming phenomenon of the smartphone age. As the game requires the user to walk in the real world to see and capture Pokémons, a new wave of crowdsourcing apps have emerged to allow users to collaborate with each other, sharing where and when Pokémons were found. In this paper we characterize one of such initiatives, called PokeCrew. Our analyses uncover a set of aspects of user behavior and system usage in such emerging crowdsourcing task, helping unveil some problems and benefits. We hope our effort can inspire the design of new crowdsourcing systems.

KEYWORDS

crowdsourcing, Pokémon Go, colaborative systems

ACM Reference format:
Priscila Martins, Manoel Júnior, Fabrício Benevenuto, and Jussara Almeida. 2017. The Emergence of Crowdsourcing among Pokémon Go Players. In *Proceedings of SIDEWAYS'17, ACM Hypertext Workshop, Prague, Czech Republic, July 2017,* 4 pages, CEUR-WS.org.

1 INTRODUCTION

The mobile games industry experienced an exponential growth in the past decade, motivated mainly by (i) an ever increasing worldwide penetration of smartphones and mobile devices, (ii) the ability of such devices to deliver quality audio and video; and (iii) the increasing capacity of network transmissions of these devices, allowing users to download larger and more complex games [7].

The largest gaming phenomenon of the smartphone age so far has been the augmented reality game *Pokémon Go*. It was launched in July 2016, firstly in Australia, New Zeland, and USA. Yet, in one week after launching, it had already reached seven million users, accounting for three to six times more downloads of the most popular games in history at that time [5]. The game makes use of GPS, camera, and position sensors of smartphones which allow its users to capture, battle and train virtual creatures called Pokémons. These creatures appear on the phone screen as if they were in the real world. The set of technologies that allow this kind of experience support the so-called augmented reality, a field that has received a lot of attention after the game success.

There has been a number of recent studies exploiting behavioral changes among Pokémon Go players. Nigg et al. [6] suggest that Pokémon Go may increase physical activity and decrease sedentary behaviors. Other efforts [2, 3, 8] argue that the game may represent a new shift in perspective: players tend to socialize more while playing as they tend to concentrate in popular areas of the game, often called PokéStops.

Since the game requires the user to walk in the real world to see and capture the Pokémons nearby, a new wave of supporting apps has emerged. In these apps, players can collaborate with each other, sharing where and when Pokémons were found. They represent the emergence of a crowdsourcing effort of the game players to find rare and valuable Pokémons. PokeCrew[1], one such app of great popularity, is a crowdsourced Pokémon Go map. It shows reports of locations of Pokémon posted by players in real time in a map and it became quite popular among the most active users. For example, this website was ranked among the top 15000 domains in the Web, according to Alexa.com [1] and its IOS and android versions had hundreds of thousands downloads.

In this paper, we characterize the crowdsourcing effort of Pokémon Players through PokeCrew. Crowdsourcing systems enlist a multitude of humans to help solve a wide variety of problems. Over the past decade, numerous such systems have appeared on the Web. Prime examples include Wikipedia, Yahoo! Answers, Mechanical Turk-based systems, and many more [4]. Our effort consists in characterizing an emerging type of crowdsourcing system, identifying many interesting technical and social challenges. To that end, we obtained a near two-month log of reports from the game players, containing 39,895,181 reports of Pokémon locations. Our analyses uncover a set of aspects of user behavior and system usage in an emerging crowdsourcing task. We hope our effort can inspire the design of emerging crowdsourcing systems.

In the following, we first describe the data used in our study and then analyze how users collaboratively help each other within the Pokecrew platform. We finish this paper with our conclusions and possible directions for future work.

SIDEWAYS'17, Prague, Czech Republic

[1]https://www.pokecrew.com/

2 DATASET

With the increasing popularity of Pokémon Go in the whole world, many applications emerged with the purpose of enhancing the players experience with the game. Among the most popular ones are Pokecrew[2], PokeRadar[3], and PokeVision[4]. The idea is to share Pokémon maps and their locations in a crowdsourced way: after finding Pokémons in the game itself users may report them in the supporting app, making the creatures visible to other players that are not in that specific location and time. This can be very useful to players, since, in the game, it is not possible to see Pokémons far from where the player is currently physically located.

We have obtained data from Pokecrew, a popular app that offers to users a map containing the location of Pokémons reported by other users. The application can be found in the PlayStore, AppStore, as well as on the Web. Our dataset contains 39,895,181 reports, from July 12^{th} to August 24^{th} 2016. Each report contains several information fields, including: a report id, reported Pokémon id, geographic coordinates, time when the report was created and, in some registers, a username.

Our results show that most reports in our dataset (98.7%) do not include a valid username, since the app does not require the user to identify itself in order to create reports. Thus, although we report general statistics computed over the whole dataset in the next section, we focus on the subset of reports with valid usernames to study user behavior. We note that, despite the small percentage, there is still a considerable amount of identifiable reports (over 500k) on which we can perform such analysis. Finally, we also note that the spatial information in our dataset refers to geographic coordinates of the reports.

3 REPORTED POKÉMONS AND THEIR LOCATIONS

We start our characterization by analyzing which Pokémons are the most reported ones and where they were reported. We then discuss the application adoption on specific locations.

3.1 Most Reported Pokémons

Table 1 shows the top-10 most reported Pokémons in our dataset. We note that these Pokémons are evolutions or difficult to find in the game. An evolution of a given Pokémon consists of a similar monster but with a higher power, which is very important to battle with other Pokémons in the so called 'gyms', popular places across the world where a user (a Pokémon master) battles with other users in order to take control of that gym. Besides that, having these evolutions contributes to the user's Pokedex, which is a list containing detailed stats for every creature from the Pokémon games. The more Pokémons a user has in her list, more experience she has on the game, which is also important to battle with other players at the gyms. Capturing an evolution is attractive to players because the only alternative way to obtain them is to use an egg, which is earned as the player progress in the game. With this egg, the Pokémon master can put it to crash, which is achieved by walking. The distance required to crash an egg may vary (2, 5 or 10 kilometers). The higher the distance, the more valuable the Pokémon that comes out of the egg is. Thus, it is much more convenient to catch the evolved Pokémon right away than it is to walk waiting for the egg to crack and, luckily, be rewarded with a powerful Pokémon.

This observation shows the great contribution of Pokecrew to Pokémon players: the interest in rare Pokémons or evolutions is what drive players to appeal to these crowdsourced apps. These Pokémons are more valuable than the most commonly found, which normally have less power to battle.

Pokémon's Name	Number of Reports
Fearow	4,957,913
Raichu	3,910,515
Slowbro	2,395,383
Golduck	2,150,770
Pidgeot	2,051,487
Nidoran	1,507,741
Tentacool	1,199,451
Nidoqueen	1,044,395
Magnemite	938,829
Clefable	928,105

Table 1: Top 10 reported Pokémons

3.2 Pokémon Location

We note that our dataset contains only geographic coordinates of the reports. In order to characterize the location where these reports were made, we first converted the coordinates to the cities and countries where the reports are made. Our approach to do that consisted of using a reliable Python library, namely geopy[5], which allows us to retrieve the nearest town/city for a given latitude/longitude coordinate.

City	Number of reports
New York City	2,722,902
San Francisco	2,678,151
Taipei	747,773
Paris	738,837
Santiago	711,977
Johor Bahru	618,659
Kuala Lumpur	538,603
Long Island City	528,104
Tokyo	526,416
Kampung Pasir	430,890

Table 2: Top 10 popular cities

3.3 Crowdsourcing Adoption

In this kind of application, the initial stage of its lifespan can be unattractive to the users due to the lack of data in the system. In PokeCrew and other competing apps, this aspect is even more

[2] www.pokecrew.com
[3] https://www.Pokemonradargo.com/
[4] www.pokevision.com
[5] https://pypi.python.org/pypi/geopy/

critical, since the adoption of Pokémon GO was fast and user engagement was very strong. In an application like Pokecrew, if the user does not encounter Pokémon reports, it is very likely she will not be motivated to use the system and therefore won't be encouraged to create new reports.

Figure 1: Geographical Location of Top 20 regions from NY

To assess whether the previous presence of Pokémons in the application influences the user to create new reports, we compared the amount of reports in a popular area among the days. To that end, we focused on the reports in the city of New York, which concentrates most part of reports. The city geographic area was first divided into 800 regions of approximately $500m^2$, and for each region we counted the total number of reports created on each day. Besides that, we considered the period from August 14^{th} to 26^{th}, which concentrates a larger number of reports. It is possible to see in Figure 1, most of the reports made in New York City, were created in the Central Park area, a very popular place in the game itself. Figure 2 shows a heat map correlating the number of reports in each region showed in Figure 1 in each date.

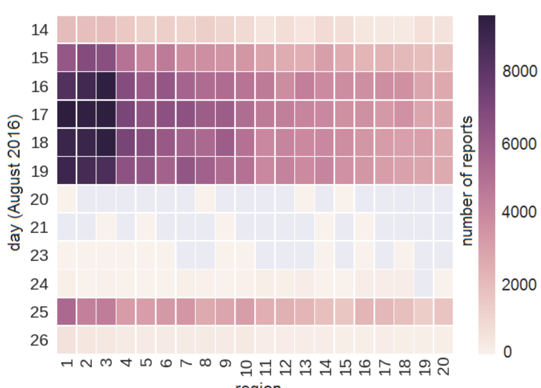

Figure 2: Heat Map of Top 20 regions from NY

4 USER BEHAVIOR

Next, we provide a characterization of the crowdsourced Pokecrew data, exploring aspects of user behavior and their engagement within this crowdsourced system. The dataset collected presents 39,895,181 reports, but only 452,359 (1.3%) are registered users (non-anonymous). We focus the next analysis on the behavior of this identified group of users.

4.1 User Engagement

Analyzing the reports made by registered users, we can notice that some of them, mostly users who contributed with the largest amounts of reports, reported a large number of sightings in just one day. For example, the top 1 user reported 184,615 times, being her reports concentrated between July 21^{st} and 24^{th}. Specifically, reported 13,007 times on July 21^{st} and 75,574 on July 22^{nd}, which are very expressive numbers. Similarly, 31,656 reports made by the second most active user, which corresponds to almost 99.7% of his contribution to the Pokecrew system, were concentrated on a single day, August 12^{th}. The 10 most active users in the Pokecrew system, in terms of reports of Pokémon sightings, are shown in Table 3. We note that some users have reported far many sightings than others, especially the 4 most active ones. Given the large amount of reports associated with these users, and the short time interval during which they were made, we speculate that these reported sightings may not have been made by "legitimate" Pokecrew users.

Based on the amount of reports that the top 5 have made, and the fact that these reports are, in general, concentrated in a few set of days, we've disregarded this data for some analysis. Just a few number of users have reported from 30 to 261 times. Most of users have reported from 1 to 20 times. We have categorized users who have reported from 1 to 5 times as less active users, representing 80% of the database.

Given that for each report we have the information about the time when it was created and its location, we can calculate for each pair of reports of a single user the speed in which he would have to dislocate in order to make both reports. Calculating the speed for each pair of user's report, we obtained a set of speeds from the identified user's reports. The chart bellow shows the speed distribution across the total identifiable users:

As we can see in the plot above, a considerable number of reports made by the users in the system reveal that they would have to

Counting reports	User name
184615	yay1199
31766	hi
609	mongrelo
413	luchocadaingles
270	maestroPokémonbelloto
261	rooty
255	srdandrea
239	crescenttough
236	ceryatec
235	guantoresp

Table 3: Top 10 users and their counting reports

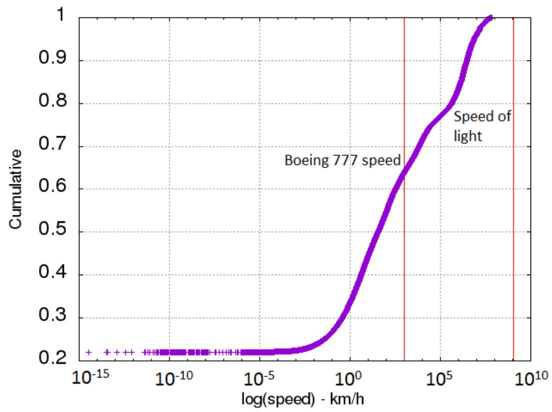

Figure 3: Speed distribution across identified users

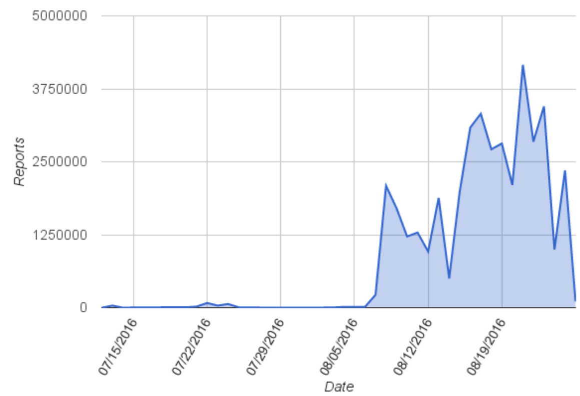

Figure 4: Amount of reports per date

move at abnormally high speeds for this kind of game. Even if we consider the case where a user report a Pokémon in one place, go on an airplane trip and then report another in the destination, which although maybe uncommon, is possible, some speeds are not feasible to achieve. The problem is that, even though the PokeCrew app uses the GPS data from the smartphone and places the map in the user current location, the player has the ability to move the map to any place in the world and therefore report a sighting from anywhere. There is no validation in the app if the report being created is trustful. This opens a serious flaw in the system, because malicious users and even bots could create fake sightings to spoof the system and degrade the legit user experience. Even if the user could not change the location in the map, it would still be possible to report a Pokémon that does not exist in the game itself at that given time and location. This kind of validation is a key problem in collaborative systems, and can be very challenging due to the lack of mechanisms to control whether the information being supplied to the database is true or not.

4.2 Temporal Analysis

The created at field represents the date when the the sighting was reported by a user. The chart below shows an analyse about how many reports were made in each day. This scenario changes after August 8^{th}, with a peak of reports on August 21^{st}.

5 CONCLUDING DISCUSSION

This paper explores one out of a new wave of applications that rely on collaborative databases in the mobile gaming environment. Pokémon Go was a huge phenomenon in this area and motivated the creation of innumerable initiatives to enhance the gamers' experience. Since the game requires the user to walk and visit places in order to succeed at being a Pokémon master, this kind of support showed to be extremely useful to players, because they could go straight to the exact location where a desired Pokémon is located instead of randomly walk hoping to find some valuable monster, as we could see with the most reported Pokémons. One important aspect in this work was to show how vulnerable such systems are to spoofed data. Beyond the gaming environment, we can cite other collaborative systems such as Wikipedia, where it's possible to see articles with dubious content, although they have more strict controls to prevent these kind of activity. In the Pokecrew application, inconsistent locations over time for some users and the noticeably high amount of reports these users pushed to the system certainly had a negative impact over the legit final user experience, who could encounter fake reports on the map. The recurrent presence of fake data on crowdsourced applications can be very frustrating to the final user, who can be discouraged to continue using the system and therefore stop providing useful information to the database. One possible solution for the considered scenario would be add an authentication layer to the system to detect and possibly ban users spamming the system. Besides that, some metrics can be useful to track abnormal activity, as we showed with the speed chart on section 4. Thus, it's highly recommended for collaborative systems such as the application studied in this work to implement these mechanisms to keep the system safe from unwanted data.

ACKNOWLEDGMENTS

We would like to thank Pokecrew for kindly sharing its data with our research group and also the Sideways organization for the space to spread this work. This work is supported by author's individual grants from Capes, Fapemig, and CNPq. F. Benevenuto is also supported by Humboldt Foundation.

REFERENCES

[1] Pokecrew Traffic Statistics. (????). http://www.alexa.com/siteinfo/pokecrew.com Available at http://www.alexa.com/siteinfo/pokecrew.com.
[2] Fabio de Oliveira Roque. 2016. Field studies: Could Pokemon Go boost birding? *Nature* 537, 7618 (2016), 34–34.
[3] Leejiah J Dorward, John C Mittermeier, Chris Sandbrook, and Fiona Spooner. 2016. Pokémon Go: Benefits, Costs, and Lessons for the Conservation Movement. *Conservation Letters* (2016).
[4] Jeff Howe. 2006. The rise of crowdsourcing. *Wired magazine* 14, 6 (2006), 1–4.
[5] Eddie Makuch. Pokemon Go Reaches $600 Million, Faster Than Any Mobile Game in History - Report. (????). http://www.gamespot.com/articles/pokemon-go-reaches-600-million-faster-than-any-mob/1100-6444687/ Available at http://www.gamespot.com/articles/pokemon-go-reaches-600-million-faster-than-any-mob/1100-6444687.
[6] Claudio R Nigg, Desiree Joi Mateo, and Jiyoung An. 2017. Pokémon Go may increase physical activity and decrease sedentary behaviors. (2017).
[7] Jason O. B. Soh and Bernard C. Y. Tan. 2008. Mobile Gaming. *Commun. ACM* 51, 3 (March 2008), 35–39. DOI: https://doi.org/10.1145/1325555.1325563
[8] Masaru Tateno, Norbert Skokauskas, Takahiro A Kato, Alan R Teo, and Anthony PS Guerrero. 2016. New game software (Pokémon Go) may help youth with severe social withdrawal, hikikomori. *Psychiatry Research* (2016).