Game Principles for Enhancing the Quality of User-generated Data Collections

Olga Yanenko Christoph Schlieder
University of Bamberg
Chair of Computing in the Cultural Sciences
96047 Bamberg, Germany
{olga.yanenko, christoph.schlieder}@uni-bamberg.de

Abstract

In this paper we discuss gamification principles in the context of Volunteered Geographic Information (VGI). Our main focus is the quality of the data collected in a VGI game. Two methods for improving the quality of crowdsourced data are compared: the widely used approach of mutual confirmation of data sources which is known to work well for objective data (e.g. mapping bus stops) and the retest mechanism, a technique adapted from the cognitive sciences for validating subjective data with high inter-individual variation (e.g. judgements about the safety of a place). We formulate design criteria for a game based on retesting and present the VGI game Alien GeoSpy as a proof of concept which applies both principles as in-game elements.

Keywords: VGI, Gamification, Data Quality, Confirmation, Retest

1 Introduction

The research areas of mobile location-based games as well as Volunteered Geographic Information (VGI, [7]) have emerged within the last decade based on the possibilities provided by the proliferation of technologies such as GPS and the mobile internet. Prior research has identified four main challenges to the success of VGI: (1) contributor motivation [3, 6], (2) data quality [1, 6, 7, 13, 16] and (3) the spatial and (4) temporal coverage of the collected data [7]. In this work we mainly concentrate on the second challenge and propose to study data quality in the context of social reporting scenarios in which volunteers submit eye witness reports about geographic objects, locations or events. Data validation in social reporting requires (semi-) automated mechanisms that can be directly integrated in the data collection process because human processing of reports does not scale up with the number of contributors. Gamification is a proven method for increasing user motivation in a wide variety of applications [5, 9].

In this article we present two approaches for data validation that can be included in mobile data collecting games. The approaches are compared with respect to their applicability to the VGI scenario, the expected effect on the quality of the dataset collected in the game as well as on the player's in-game behaviour. This comparison leads to the formulation of three general design requirements. Finally, a prototype implementation of the VGI game Alien GeoSpy is presented as a proof of concept which demonstrates that it is possible to satisfy the design requirements.

The remainder of this article is organized as follows: Section 2 discusses the background and introduces related approaches. In section 3 we introduce a new game pattern—the retesting mechanism—and contrast it with the widely used confirmation mechanism for data validation. Section 4 presents the VGI game Alien GeoSpy. Section 5 draws conclusions, points to open research questions and gives a brief outlook on future work.

2 Background and Related Work

Data produced by volunteers is often considered as being of lesser quality than data produced by experts [3, 8]. On the other hand, VGI provides a rich source of user-generated content which is difficult to substitute. Crowdsourced datasets are often more detailed than professional datasets due to the fact that a lot more people are contributing. For spatial datasets (e.g. OpenStreetMap) this is especially the case in urban areas [8]. At the same time the data is more heterogeneous and less consistent [3, 7, 8]. To understand the outcome of user-generated data collection one has to understand the motivation of the contributors. Several reasons to contribute to VGI were identified—positive ones like altruism, personal interest or social reward and negative ones like vandalism and malice or even criminal intent [3]. Negative motivation usually results in erroneous data being submitted by the data producer, but errors can also occur unintentionally as a result of a lack of expertise in a specific technical subject [1, 3].

Gamification is known to enhance the motivation of users for different human computation tasks [6], e.g. in collecting and tagging geospatial data [9]. But since the main goal of the players is to win the game, the task itself is less important to them. If this is true, gamification will increase participation but it might have a negative effect on data quality. In most of the data collection games, the players have to complete some game actions within the shortest possible time or to complete as many actions as possible in a delimited period. In both cases, time becomes a critical factor for winning the game or obtaining a higher score. From a data collection perspective, this affects the resulting dataset since the contributors are less accurate if they are acting in a hurry. Even fake data is likely to be produced by cheating players which try to save time.

Methods for ensuring a certain level of data quality are indispensable for gamified VGI approaches. Different quality assurance methods are known to work for user-generated content. The most reliable procedure consists in verifying the data by actually inspecting it on-site. Since this is the most costly and time-consuming method, it is not practicable in many scenarios. Another approach bases data validation on measuring the contributor's trust and reputation [1]. Typically, the measurement evaluates reviews of other data produced by the contributor and is expressed in form of a trust-value. These trust-values are used to predict the reliability of new pieces of information. Somewhat related to the authority approach is the confirmation method [14]. Pieces of data submitted by several independent contributors are considered more reliable than data from a single source.

Several non-spatial games include some kinds of review and confirmation method to improve data quality. The Google Image Labeler (previously known as ESP game and developed by Luis von Ahn in 2004 [14]), a human computation tagging game, was one of the first attempts to include the confirmation task directly in the game mechanics. Two players of the game were only rewarded with game points if they both agreed on their tag suggestions for an image without knowing the tags proposed by the other player.

The location-based mapping game Urbanopoly based on the board game Monopoly where people map real-world venues by buying them includes hidden human computation tasks in the form of mini-games [2]. Two different methods for data validation are implemented: a quiz where players are asked to choose an attribute value for an existing venue (e.g. the cuisine type of a restaurant), and a rating task where the player is asked to rate an artefact about a venue (e.g. a photo) on a five star scale.

Another example is the chasing game MapSigns [4]. The players in this game are divided into thieves and policemen. The goal of the game consists in collecting traffic signs on a route between a predefined start and end location while the traffic signs are regarded as heists from the thieves' point of view and depict heist evidence for the policemen. At the end of the game, all players together decide for every collected traffic sign if the item is correctly mapped or not.

However, all of these approaches only take into account objective data like the categorization of traffic signs. An emerging field of VGI, the collection of subjective data, is not considered so far. By subjective data, we understand data with high inter-individual variation. Typical examples are aesthetic preferences ("Do you think this place is beautiful?") or judgements about emotional qualities ("Do you feel safe at this place?"). Subjective data is especially of interest for Geodesign and E-Government scenarios [3, 6, 12].

We address the following research question: How can a VGI game mechanism validate subjective data during the collection process? The mechanism should at the same time prevent the players from producing fake data, support the game flow and integrate into the game narrative in such a way that the playing experience does not suffer.

3 Confirmation vs. Retesting Methods

With data confirmation approaches, the more contributors agree on a piece of information, the more reliable it is. The principle of confirmation can be used in an *implicit* and in an *explicit* way. In this work, we concentrate on tasks in which objects from the geographical environment (e.g. a bus stop) are categorized, that is, reported to belong to a certain object category. Within this context implicit confirmation is given if at least two people map an item of the same category at the same location. Implicit validation is usually applied when evaluating an already existing dataset and works best with large numbers of observations [11]. Explicit confirmation, in contrast, is based on a review mechanism where the data produced by a contributor is presented to others in order to obtain a second categorization. In the Urbanopoly game, two types of reviews are included in the game mechanics, the validation of item attributes and the rating of item artefacts. The MapSigns game also implements a reviewing mechanism by letting all players explicitly validate the collected data after the main game. Another possibility of using confirmation consists in guiding a person to a location in order to confirm that an item mapped by someone else is actually there. However, this mechanism is usually too time-consuming. We propose an explicit confirmation mechanism that is based on the review of photos made of the item. A picture of a categorized item (e.g. a bus stop) that was taken by one contributor is shown to another one who is asked to re-categorize the item. The assignment is usually made by choosing an item from a category list as opposed to a free text description. This avoids dealing with unstructured data which is more difficult to process automatically.

Both, implicit and explicit confirmation work well with objective data, when no individual bias is expected to affect the outcome of the data collection task, such as it is the case when identifying and mapping bus stops. The validation of subjective data, however, is not straightforward. If players are asked to map subjective items like "a place for a romantic picnic" confirmation mechanisms fail because of the high inter-individual variation of the categorization. This problem is well known to research in cognitive psychology. A standard soltion consists in measuring the reliability of test data. Formally, *test-retest reliability* measures a correlation coefficient of two identical tests taken by a subject twice with a time interval in between [10, 15]. We propose to integrate a similar type of measurement into the game mechanics. Similar to the procedure used with the confirmation method, the player is showed a photo of an item which she or he had mapped at an earlier stage of the game. In the retest, the player has to re-categorize the item based on the photo.

Generally speaking, the principles of confirmation and retesting are based on the same idea—the correlation of two different statements; the main difference is that confirmation relies on *inter-individual agreement* while retesting accounts for *intra-individual agreement*. Both approaches can be applied for *item-based* validation as well as *user-based* validation. In item-based confirmation, every piece of information is reviewed on its own, i.e. for each observed item, the number of confirming or disconfirming statements is counted regardless of the reporters' trust-values and a reliability score is computed for the report itself. Given a set of independent (non- contradicting) item categories, the correlation coefficient for an item-based confirmation ranges between 0 and 1 depending on the number of statements about the item. The same applies to user-based confirmation since the reliability score for an individual user is derived from the correlation coefficients of all items she or he mapped. When performing an item-based retesting, each mapped item can be regarded as a separate test. Since only the initial and the retest categorization of one user is taken into account, the correlation coefficient is either 0 (if the categorizations do not match) or 1 (if both categorizations are consistent). Again, these measures can be used for computing a user-based reliability score which ranges between 0 and 1.

In psychological retest situations, e.g. when measuring intelligence, generally, the person tested takes the same test twice. In gaming this procedure forces the player to repeat every task which would be annoying and detrimental to the playing experience. We argue that for measuring user-based retest reliability it may be sufficient to present retests only for a part of the categorization tasks. If the tasks are selected appropriately they should permit to reveal non-consistent statements and to recognize cheaters. A combination of both approaches can be useful as well: the aggregation of a larger dataset where reports can be used for computing individual reliability scores for the data producers themselves and in return these values can be taken into account when computing the reliabilities for individual reports. Table 1 summarizes the characteristics of the confirmation and retesting principles.

Table 1: Confirmation vs. retesting

	confirmation	**retesting**
basic principle	inter-individual agreement	intra-individual agreement
data	objective data	subjective data
item-based validation score	[0, 1]	0, 1
user-based validation score	[0, 1]	[0, 1]
player amount	at least 2	at least 1

As proposed in our prior work [11], the principle of spatio-temporal proximity and social distance can be used for implicit confirmation when evaluating an existing dataset produced in a gaming scenario. The explicit confirmation as well as the retesting mechanism in contrast can be included in the gaming mechanics directly.

4 Alien GeoSpy: a VGI Game

We developed a prototype of a VGI game that can be used for different mapping scenarios. For the implementation of the validation techniques as part of the game mechanics we identified the following requirements:

(1) Both, item-based as well as a user-based validation have to be supported.

(2) The validation mechanism should seamlessly integrate into the game flow and not be recognizable to the user.

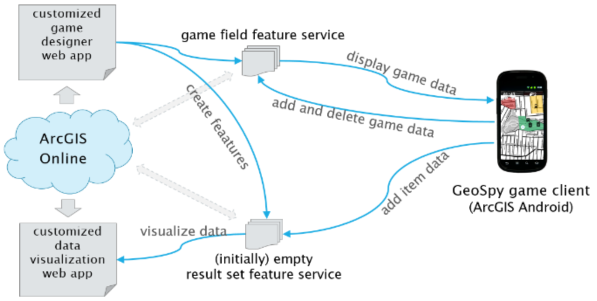
(3) The balancing of the game should not be affected by the introduction of the validation mechanisms.

The narrative of the game is based on an alien invasion scenario in which the player's role consists in spying on the planet Earth by gathering information about given regions. Items belonging to several predefined categories have to be found by inspecting the regions and are documented by mapping and photographing them within the game. The goal of the game consists in obtaining the highest score by collecting more items than the opponents. To ensure that every item is mapped only once, a predefined distance between two items of the same category is specified in the game configuration. Additional points can be obtained by completing missions or conducting some in-game tasks. A point loss is also possible by stepping on a bomb.

Figure 1: Alien GeoSpy user interface

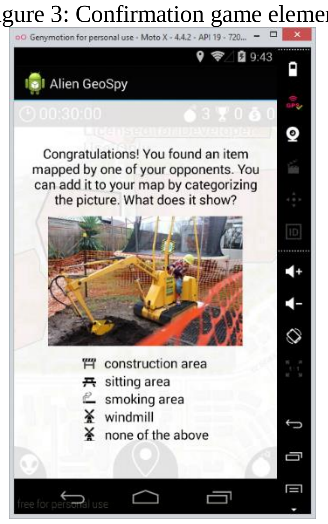
Figure 1 shows the user interface of the game with four visible regions, each in a different state denoted by the region's fill color. Regions with no items having been mapped by the player are displayed in a transparent red color. As soon as an item is mapped within a region, the region turns yellow. Finally, if at least one item of each category is mapped within a region it turns green. Figure 2 shows the overall game architecture. The game is implemented for the Android platform and uses GI technologies for all three parts of a gaming event: the creation of game instances (ArcGIS API for JavaScript, ArcGIS Online), the gameplay (ArcGIS Runtime SDK for Android) and the post-game visualization and analysis (ArcGIS API for JavaScript, ArcGIS Desktop).

Figure 2: Alien GeoSpy Game Architecture



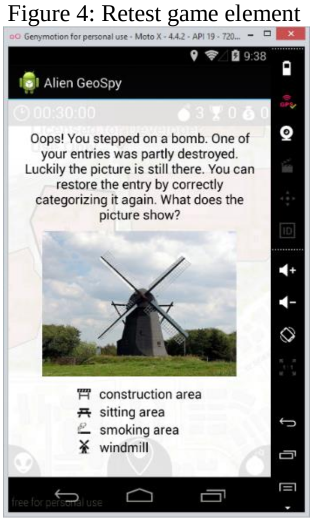
The data quality issue is addressed by two different game elements that implement the confirmation and the retesting mechanisms. From time to time a player gets the chance to earn some extra points by reviewing an item collected by one of her or his opponents. This task is realized by showing a blinking treasure symbol on the map. The player can collect this treasure by moving to its location. She or he is then told that an element mapped by one of the other spies has been found. The player is asked to categorize the item by viewing the photo taken by the opponent player. If both of the players agree in their categorization, the item and the according points are added to the player's account. The players are not penalized for a disagreement because it is not clear which of the two is wrong. The game organizer defines the frequency with which the element appears. With a higher frequency, more items are reviewed by different users resulting in more reliable confirmation values for the individual items. Note that it is important to relate the frequency to the number of items already collected by the users, otherwise the players will be too engaged in collecting other's items instead of mapping their own. The confirmation game element is illustrated in Figure 3.

Figure 3: Confirmation game element



The retesting mechanism is realized by a bomb-element. All players have a specified number of bombs at the beginning of a game. These bombs can be used for hampering the other players by placing them on the geographical game board. A bomb is harmless for the player who placed it. However, if another player moves close to a bomb, it will explode and fully or partly destroy all items mapped by this player within a certain radius of the explosion. The player will be asked to restore the partly destroyed elements by viewing the corresponding picture and by categorizing them again. If the categorization matches the initial one, the item will be restored to the player's map. Otherwise she or he loses the entry and the according points. Note that there will always be more partly destroyed items than fully destroyed items in order to guarantee a less frustrating game experience. Moreover, the fully destroyed items can easily be restored without big effort by mapping them again since the player is still within a small distance to them due to the peripheral impact of the bomb. The bomb element is illustrated in figure 4.

Figure 4: Retest game element



The possibility to destroy the opponents' mappings is limited by the number of bombs to maintain the balance of the game. The players will have to deliberate about where and when to place the bombs.

The game elements we presented do not only provide a mechanism for validating the collected data but also have the effect of keeping the players from cheating since the users are motivated to make correct statements and to take meaningful pictures in order to be able to resume their contributions in the case of a damage of their collection.

5 Conclusion and Outlook

In this article we discussed the principles of confirmation and retesting as mechanisms for data quality improvement. We presented a VGI game implementing both patterns as in-game elements without disrupting the game flow.

There is still need for further research. One of the main requirements for a retesting mechanism is that a player in the retest does not remember the answer given in the test. Since in spatial scenarios, episodic memory is aided by environmental cues, a better understanding of the memory effects involved would be helpful.

Our future work will concentrate on the evaluation of the game principles from different perspectives: improvement of the data quality, playability of the game and game balancing. Furthermore, the determination of adequate reliability scores for different scenarios is of interest in order to identify cheating players. Research on the smallest number of item-retests sufficient for computing a user-based reliability score is also fundamental for a pleasant user experience in a game.

Acknowledgements

Research was funded in part by Esri Inc., Redlands, CA. Special thanks are extended to the Esri Education Team, the Applications Prototype Lab and the Android Development Team for providing great support.

References

[1] Mohamed Bishr and Krzysztof Janowicz. Can we Trust Information? – The Case of Volunteered Geographic Information. In *Towards Digital Earth Search Discover and Share Geospatial Data Workshop at Future Internet Symposium*, vol. 640, 2010.

[2] Irene Celino, Dario Cerizza, Simone Contessa, Marta Corubolo, Daniele Dell'Aglio, Emanuele Della Valle and Stefano Fumeo. Urbanopoly—A Social and Location-Based Game with a Purpose to Crowdsource Your Urban Data. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Conference on Social Computing (SocialCom)*: 910-913. IEEE, 2012.

[3] David J. Coleman, Yola Georgiadou and Jeff Labonte. Volunteered geographic information: The nature and motivation of produsers. In *International Journal of Spatial Data Infrastructures Research* 4.1: 332-358, 2009.

[4] Nicola Davidovic and Leonid Stoimenov. Using Location Based Game MapSigns to motivate VGI data collection related to traffic signs. In *Online proceedings of the International Workshop on Action and Interaction in Volunteered Geographic Information (ACTIVITY)* at the 16th AGILE Conference on Geographic Information Science. Leuven, Belgium, May 2013.

[5] Sebastian Deterding, Dan Dixon, Rilla Khaled and Lennart Nacke. From game design elements to gamefulness: defining "gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (MindTrek '11): 9-15, ACM, New York, NY, 2011.

[6] Andrew J. Flanagin and Miriam J. Metzger. The credibility of volunteered geographic information. In *GeoJournal* 72: 137-148, 2008.

[7] Michael F. Goodchild. Citizens as sensors: the world of volunteered geography. In *GeoJournal* 69: 211-221, 2007.

[8] Christian Heipke. Crowdsourcing geospatial data. In *ISPRS Journal of Photogrammetry and Remote Sensing* 65(6): 550-557, November 2010.

[9] Sebastian Matyas, Peter Kiefer, Christoph Schlieder and Sara Kleyer. Wisdom about the crowd: assuring geospatial data quality collected in location-based games. In *Entertainment Computing—ICEC 2011*: 331-336, Springer Berlin Heidelberg, 2011.

[10] John Rust and Susan Golombok. *Modern Psychometrics: The Science of Psychological Assesment*. Taylor & Frances/Routledge, Florence, KY, US, 1989.

[11] Christoph Schlieder and Olga Yanenko. Spatio-temporal proximity and social distance: a confirmation framework for social reporting. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*: 60-67, ACM, November 2010.

[12] Weidong Song and Guibo Sun. The role of mobile volunteered geographic information in urban management. In *Proceedings of the 18th Conference on Geoinformatics*: 1-5, June 2010.

[13] David L. Tulloch. Is VGI participation? From vernal pools to video games. In *GeoJournal* 72: 161-171, 2008.

[14] Luis von Ahn and Laura Dabbish. Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems*: 319-326, ACM, April 2004.

[15] Mark Wilson. *Constructing measures: An item response modeling approach.* Routledge Academic, 2013.

[16] Olga Yanenko and Christoph Schlieder. Enhancing the Quality of Volunteered Geographic Information: A Constraint-Based Approach. In *Bridging the Geographic Information Sciences*: 429-446, Springer Berlin Heidelberg, 2012.