Harnessing Twitter and Crowdsourcing to Augment Aurora Forecasting

Nicolas LaLone
Penn State University
nlalone@psu.edu

Andrea Tapia
Penn State University
atapia@ist.psu.edu

Elizabeth MacDonald
NASA Goddard
New Mexico Consortium
elizabeth.a.macdonald@nasa.gov

Nathan Case
NASA Goddard
New Mexico Consortium
nathan.a.case@nasa.gov

Michelle Hall
Science Education Solutions
hall@scieds.com

Jessica Clayton
Science Education Solutions
jclayton@scieds.com

Matthew Heavner
US White House OSTP
Los Alamos National Laboratory
New Mexico Consortium
Matthew_J_Heavner@ostp.eop.gov

CSCW '15 Companion, Mar 14-18, 2015, Vancouver, BC, Canada
ACM 978-1-4503-2946-0/15/03.
http://dx.doi.org/10.1145/2685553.2702671.

Abstract
The aurora borealis and aurora australis are beautiful space weather driven events whose sighting is typically based on luck given that forecasting is not spatially or temporally precise. To help increase the accuracy and timeliness of auroral forecasting, we have designed a multi-faceted system called Aurorasaurus. This system allows crisis management specialists to test reactions to rare event notifications, space weather scientists to get direct sighting information of auroras (complete with pictures), and science education researchers to evaluate the impact of educational materials about the aurora and the physics surrounding this unique phenomenon. Through manual tweet verification and directly reported aurora borealis or aurora australis sightings, everyday users help make space weather and aurora forecasting more accurate.

Author Keywords
Early warning systems; Twitter; Social Media; Disaster; Crisis; aurora borealis.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Figure 1 - Aurorasaurus Branding

http://www.aurorasaurus.org

Figure 2 - Mobile Application Homescreen.

Introduction & Background

Citizen Science developers have long struggled to solve two simple problems: recruitment and training [2]. In an evaluation of over 340 citizen science products, researchers found that approximately 11% of those used mobile devices [2]. This means that over 80% of these products, rely on desktop or laptop computers for their ability to run an internet browser. Considering the proliferation of mobile devices over the past decade, HCI practitioners employed by citizen science initiatives should address how to encourage users to help them with the device they have in their pocket.

In addition to the lack of mobile device adaptation, another issue is becoming more apparent: social media data. To wit, millions of users are broadcasting vast amounts of data through social media. These data are often used by researchers to predict everything from movies sales to the spread of the seasonal flu virus [1]. Crisis management researchers have begun to use twitter to augment rescue efforts [7].

As such, citizen science initiatives may find that it is easier to simply engage the Twitter API than it is to design a new data collection vehicle that requires the training of new users or tools. Depending on the granularity, or effort required to complete a citizen science-based task, recruitment and motivation of users can be problematic [4]. To that end, we have created Aurorasaurus which includes a hybrid approach to data collection using both purposeful citizen scientists and ad hoc tweets

Aurorasaurus is a website, a mobile application, and a citizen science initiative that harnesses the Twitter API and a simple voting system in order to comb through geocoded and time-stamped tweets that reference the aurora borealis, the northern lights, or the aurora australis. These data are evaluated by users who are interested in the auroras or the science behind them. Once evaluated, tweets that contain a verified aurora sighting augment the coarse means through which space weather scientists currently use to predict where the aurora can be seen.

The aurora borealis and aurora australis are the result of charged particles from the sun colliding with the atmosphere near the magnetic poles after being accelerated in the Earth's magnetosphere. The strength of an aurora, and thus its location, is driven by the sun's activity and the conditions in interplanetary space. Predicting when and from where the aurora can be seen is a non-trivial task and so space weather scientists have to use a suite of methods to help them do so (e.g. [3, 6]).

The Aurorasaurus System

An alpha version of Aurorasaurus (Aurorasaurus 1.0) has been live on the web for two years. This November, we re-launched with new branding and a new mobile application on iOS and Android. To date, Aurorasaurus has shown promising results [6]. At its core, Aurorasaurus offers an estimate of the location and character of the aurora borealis.

The system estimates these aspects of the aurora through two different types of participation: active and passive. First, active participation is sought through form-submitted direct observations of the aurora. These are entered through a form in the mobile application or browser window. In addition, we also aggregate tweets about the auroras (passive) and

present them to our users for verification (active). The addition of passive participation in the form of collecting tweets for active participants to categorize is key to our tool.

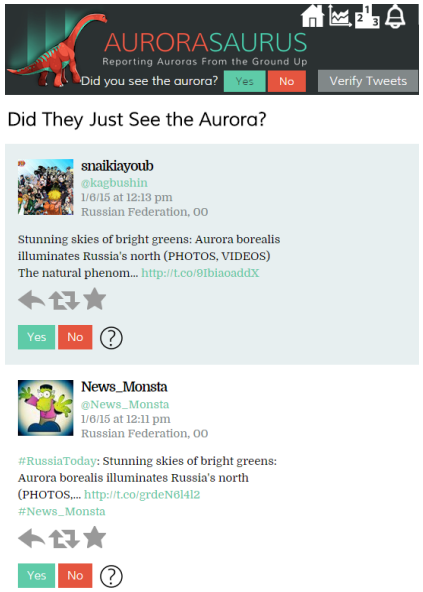

Figure 3 - Twitter Verification Screen. In addition to buttons for Twitter commands, users can also gain access to more information about what our team is looking for when we ask, "Did they just see the Aurora?"

Aurorasaurus gathers tweets in real time via keyword searching through the Twitter streaming API. These tweets are added to a database that users are encouraged to help clean with some automated filtering we hope to augment in the future. Anyone with the application or website open, not just registered users, can read tweets and answer "Yes" or "No" to the question, "Did They Just See the Aurora?" This screen can be seen in Figure 3. Once a certain score has been recorded, that tweet becomes a verified sighting or is rejected as noise.

Verified sightings are then tested against predictive data from space weather observatories from around the world. When applicable, these predictions are displayed on our application's map (Figure 4). Since the mobile launch in late 2014, we have recruited over 500 users even without particularly active recruitment.

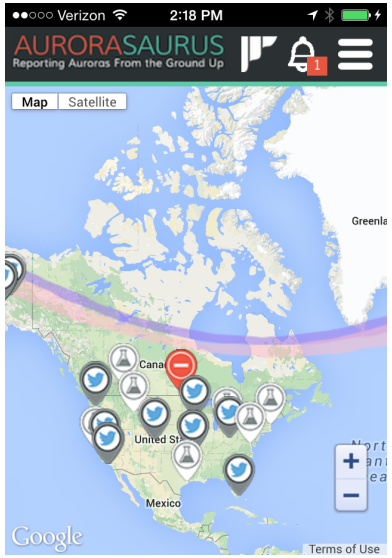

Figure 4 - The two color bands in northern Canada are estimates of Auroral activity.

Current & Future Development

We have begun to see more activity as we add additional layers to our Aurorasaurus. First, a community of hundreds of registered users and thousands of social media followers has begun to

develop. Second, we are beginning to lay the groundwork for the incorporation of data mining and informational retrieval experts to change the way we interact with our corpus of tweets. We are doing this in three distinct ways: interaction, information, and automation.

First, we are interacting with our participants through social media and the Google+ Hangout system. Here, we hold monthly meetings whereupon any interested user can ask the team aurora-related questions. In addition, users can ask for information or simply observe as we talk about future Aurorasaurus developments. Our hope is that by creating a means through which users can engage the Aurorasaurus team that participation of our users will increase.

The goal of this initiative is meant better inform space weather scientists and everyday citizens about auroral activity. To achieve this goal, we are developing educational materials for the classroom and gamification-inspired extrinsic rewards of the site (badges, icons, etc). In this way, gain a means through which to gauge public understanding of the physics surrounding space weather.

Finally, as Aurorasaurus grows and becomes more active, we hope to add data cleaning algorithms to help automate the cleaning of tweets. Because engaging large amounts of tweets is so complex, the addition of a limited context should prove to be an excellent testing ground for such algorithms.

In these three ways, we are engaging the concept of citizen science and incorporating the interdisciplinary nature of this movement to assemble an initiative that is focused not solely on citizen participation, but on interdisciplinary cooperation as well.

ACKNOWLEDGMENTS

This work is partially supported by the National Science Foundation #1344296, INSPIRE: Aurorasaurus Citizen Scientists Experiencing the Extremes of Space Weather.

REFERENCES

[1] Chen, H. Achrekar, B. Liu, and R. Lazarus, "Vision: towards real-time epidemic vigilance through online social networks: introducing sneft – social network enabled flu trends." ACM Mobile Cloud Computing and Services, San Francisco, California, June 2010.

[2] Kim, S., Mankoff, J., & Paulos, E. (2013, February). Sensr: evaluating a flexible framework for authoring mobile data-collection tools for citizen science. In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 1453-1462). ACM.

[3] Newell, P. T., Sotirelis, T., & Wing, S. (2009). Diffuse, monoenergetic, and broadband aurora: The global precipitation budget. *Journal of Geophysical Research: Space Physics* (1978–2012), 114, A9.

[4] Nov, O., Arazy, O., & Anderson, D. (2011, July). Technology-Mediated Citizen Science Participation: A Motivational Model. In ICWSM.

[5] Sigernes, F., Dyrland, M., Brekke, P., Gjengedal, E. K., Chernouss, S., Lorentzen, D. A., Oksavik, K., & Center, S. W. (2011). Real time aurora oval forecasting-SvalTrackII. In *other words*, 3, 3.

[6] Tapia, A.; Lalone, Nicolas; MacDonald, Elizabeth; Hall, Michelle (2014) Crowdsourcing Rare Events: Using Curiosity to Draw Participants into Science and Early Warning Systems. *Proceedings of the 11th ISCRAM*.

[7] Tapia, A. H., LaLone, N., & Kim, H. W. (2014). Run Amok: Group Crowd Participation in Identifying the Bomb and Bomber from the Boston Marathon Bombing. *Proceedings of the 11th ISCRAM.*