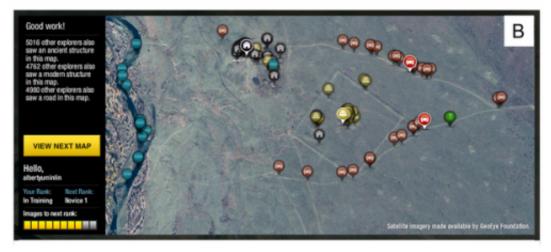# Crowdsourcing used in Genghis Khan tomb search

January 9 2015, by Nancy Owano

Tagging Interface: (a) Example of tags being made. Tags are color coded with roads (red), rivers (blue), ancient (yellow), modern (grey), and other (green) structures; (b) Example of peer feedback after a participant completes their annotation task. Results of all previous observers of that image tile are shown. Satellite imagery provided courtesy of the GeoEye Foundation. Credit: doi:10.1371/journal.pone.0114046.g001

Researchers from the University of California San Diego have written "Crowdsourcing the Unknown: The Satellite Search for Genghis Khan," published last month on *PLOS ONE*, the peer-reviewed, open-access, online publication. Briefly, they charged an online crowd of volunteer participants with the challenge of finding the tomb of Genghis Khan. Their field expedition to look for this tomb was designed to engage tens of thousands of public volunteers and generate contributions towards an archaeological satellite imagery survey. The crowdsourcing figures attributed to the study are impressive: Over 10,000 online volunteers contributed a combined total of 30,000 hours (3.4 years) of human visual analytics, calculated from user interaction time logs, and generated 2.3 million feature categorizations.

This was a virtual exploration system that launched on June 10, 2010. Participants generated inputs (tags), creating a geospatial map that highlighted regions of crowd consensus among inherently noisy data. Motivation? One can answer the obvious, to find the tomb of Genghis Khan. Location of the tomb was, after all, an archaeological enigma lacking historical descriptions of its visual appearance. The authors said that not a single burial of the Mongolian imperial family has been identified. What is more, there are largely undocumented cultural heritage sites across a sparsely populated and undeveloped landscape. So while looking for the tomb was a motivation, the effort was also to leverage the power of human perception in a search for the unexpected. This was a challenge. The authors said that without a pre-existing reference for validation they turned towards consensus, defined by kernel density estimation, to pool human perception for "out of the ordinary" features across a vast landscape, developing a collective reasoning engine for anomaly detection.

Leave alone looking for a needle in a haystack, in the scenario here, the

appearance of the needle was undefined. Speaking about the enigma, an article in Smithsonian.com said, "The tomb of Genghis Khan—Mongolian ruler, warrior, and ancestor to an estimated one out of every 200 humans alive today—has been a mystery for almost as long as the man has been buried. Legend has it that when he died in 1227, soldiers killed the tomb builders as well as every person the funeral procession passed. Then, it is said, the soldiers themselves were killed so that no one who knew the tomb's location would live to share it."

Dr. Albert Yu-Min Lin, co-author of the paper, said in National Geographic, "Using traditional archeological methods would be disrespectful to believers." The goal would be to identify archaeological sites without disturbing them–in the area of Mongolia's most sacred heritage–Genghis Khan's homeland, where the tomb would be an extremely sacred place and not to be desecrated. In turn, he and his team chose a noninvasive way without overstepping cultural barriers. "It also allows us to empower Mongolian researchers with tools they might not have access to otherwise," he said.

Finding the tomb, however, was not the only goal. The study is important as a test of the strengths of crowdsourcing under difficult conditions. The authors said that "Turning to the crowd as the 'partner of choice' for scalable problem solving is becoming increasingly attractive across broad domains both in science and industry."

Their approach involved loosely guided, online volunteer participation, they said, to generate human identifications of unknown anomalies within massive volumes of geospatial remote sensing data. They utilized sub-meter resolution satellite imagery of the Mongolian steppe. They described satellite imagery (0.5 meters/pixel), tiled into 84,183 small, semi-overlapping image tiles (1236×630 pixels) and presented to the public via a National Geographic website.

The research resulted in a map that confirmed 55 sites with archaeological significance. They said the 55 potential archaeological anomalies verified by the field team ranged from bronze age to Mongol period in origin

As important, the authors had this to say about the power of crowdsourcing: the power "lies not only in harnessing parallel networks for scalable analytics, but in forming the collaborative frameworks necessary to cultivate collective reasoning."

Abstract

Massively parallel collaboration and emergent knowledge generation is described through a large scale survey for archaeological anomalies within ultra-high resolution earth-sensing satellite imagery. Over 10K online volunteers contributed 30K hours (3.4 years), examined 6,000 km2, and generated 2.3 million feature categorizations. Motivated by the search for Genghis Khan's tomb, participants were tasked with finding an archaeological enigma that lacks any historical description of its potential visual appearance. Without a pre-existing reference for validation we turn towards consensus, defined by kernel density estimation, to pool human perception for "out of the ordinary" features across a vast landscape. This consensus served as the training mechanism within a self-evolving feedback loop between a participant and the crowd, essential driving a collective reasoning engine for anomaly detection. The resulting map led a National Geographic expedition to confirm 55 archaeological sites across a vast landscape. A increased ground-truthed accuracy was observed in those participants exposed to the peer feedback loop over those whom worked in isolation, suggesting collective reasoning can emerge within networked groups to outperform