HiggsHunters - a citizen science project for ATLAS

Andrew Haas, on behalf of the ATLAS collaboration
NYU Physics, 4 Washington Place, New York, NY 10003
andy.haas@nyu.edu

Abstract. Since the launch of HiggsHunters.org in November 2014, citizen science volunteers have classified more than a million points of interest in images from the ATLAS experiment at the LHC. Volunteers have been looking for displaced vertices and unusual features in images recorded during LHC Run-1. We discuss the design of the project, its impact on the public, and the surprising results of how the human volunteers performed relative to the computer algorithms in identifying displaced secondary vertices.

1. Introduction
Citizen science allows people (from the general public) to participate in scientific studies, thus allowing them to learn about the science more deeply, and benefitting the studies. Zooniverse.org is a series of web-based projects that makes it easy to join into citizen science. So far there have been 48 projects, with millions of participants, and hundreds of scientific publications resulting. The topics range from social science, where ancient texts are studied, to astronomy, where galaxy shapes are classified. With HiggsHunters, we attempted to make citizen science work for the high-energy collider experiment, ATLAS [1].

Figure 1: A typical event from the project.
To be successful, the project must be interesting, with real scientific goals, but also fun and easy to understand. We made it visual and interactive using clickable event displays. The idea was that people may be able to help us find new long-lived particles we may have missed. This is typically a hard problem for computers to reconstruct, given the large phase space of types of vertices in various parts of the detector.

Users click on the images they are shown, identifying locations with vertices, saying if there are 2, 3, 4, 5-10, or 10+ tracks. They can also identify “something weird”, which is followed up separately. We, the ATLAS team managing the project, receive monthly metadata with this info on each user’s clicks and classifications.

Figure 2: A typical image that a participant would classify.

2. Science behind HiggsHunters

Many events shown on HiggsHunters are real ATLAS data events from 8 TeV pp collisions in 2012, which have been selected to be enriched in Higgs boson decays. The premise is that Higgs bosons are more likely to have rare decays to long-lived particles, due to its naturally small width \[2\]. The events must pass a Z→mumu trigger, with two muons with transverse momenta \(p_T > 20 \text{ GeV}\) and \(|\text{eta}| < 2.5\) and \(|M(\text{mu}\mu) - 90| < 25 \text{ GeV}\), where M is invariant mass. The vector sum of the two muons’ momenta must also be >60 GeV (which is typical for a recoil off a Higgs). To enhance the possibility of new, rare Higgs decays, we select events with missing transverse momentum >40 GeV. Roughly 60 Zh associated production events are expected in this dataset, out of a total of ~60k.

In addition to the data events, a signal model was chosen and simulated events were created and shown to the users, to test their efficiency for identifying vertices. The simulated events were also used to “train” users in identifying events, but for final classification the users were not told whether the event was simulated, until after it was classified. The model had the Higgs decay to two new long-lived particles, with masses of either 8, 20, or 50 GeV, and decay lengths of 1, 10, or 100 mm. The particles could then decay to either a pair of tau leptons, or b-jets, depending on the particle mass. In
some sense, HiggsHunters is like a standard search for new physics, but using crowd-sourced human eyes and brains, rather than computer algorithms, for some reconstruction of the data.

Figure 3: The model of new physics, with the Higgs decaying to two long-lived particles.

3. HiggsHunters statistics
Since launching in Nov. 2014, there have been more than 30k participants, from 179 countries. About 85k images were made available, and 700k images have been classified, with over 1.2M clicks (vertices) identified. (Each event is viewed multiple times by many people, and results clustered.)

Figure 4: Statistics of HiggsHunters.
4. HiggsHunters supporting resources

In addition to the main image classification section of the website, higgshunters.org, there are supporting introductory tutorials and videos about the LHC, HEP, ATLAS, and the HiggsHunters science. These are essential for helping the users to understand what they are participating in. There are also more personal touches, such as a blog (blog.higgshunters.org) with posts from the ATLAS research team, further explaining the science as it relates to the project. There were also Twitter and Facebook postings to gather users to the site. The project gathered significant media attention from sites like Wired and Physics Today, helping to draw people to the site, and educate them about the science behind the project, even if they didn’t eventually participate.

In addition, a very successful element of the site was the “talk” site (talk.higgshunters.org), where users could discuss events that they saw with other users. They could #hashtag events with tags that would create “conversations” joining many events together. For instance, #bigvertex could be used to tag all events found with a very large vertex with many tracks. They could then discuss common explanations for these amongst themselves and with ATLAS experts on the talk forums. After some initial confusion, it was decided to clearly label events in chat forums as coming from simulation or not.

5. Results

The analysis of the users’ clicks was still ongoing at the time of this talk, Oct. 2016. Some preliminary findings are summarized below, in Fig. 6. More detailed findings can be seen in our recently submitted
Each event was also classified by the standard algorithms used to find displaced vertices at ATLAS. People do beat the algorithms in some cases, but it depended on the signal model parameters, as the algorithm was tuned for higher-mass vertices.

Figure 6: Some preliminary results comparing people to algorithms.

People were very good at detecting “weird stuff”, processes that made events which clearly looked different. ATLAS has no general algorithm for detecting such things, though the processes tend to be well-known over time, as the events show up in tails of distributions. For instance, #muonbundle was tagged in 397 events and #weird was tagged 963 times. Several cases were followed up in the talk forums by ATLAS experts and explained in detail, there and in blog posts. Users discovered many odd processes such as jet punch-through, and cosmic ray showers. It was encouraging to see that users would have spotted some striking new phenomena we hadn't designed an algorithm to see. And the users really enjoyed the excitement of possibly finding something new.

6. Demographics and survey
In addition to volunteered information from the users on the site, and geographic data from IP addresses, surveys were used to assess the user population. A large spread of ages was found, from middle-school students to retirees. One-third of the users were female. 75% had a college degree, but most had no physics training beyond high-school. 80% had participated in other Zooniverse projects.
7. Summary and future plans

The HiggsHunters project on Zooniverse was able to successfully attract and educate a large range of citizens about high-energy physics. New users are still joining daily and we are still analyzing clicks. HiggsHunters has also recently partnered with the Institute for Research in Schools (IRIS) to have students help analyze the user click data, an even more intricate pursuit. Finally, there is overwhelming support for a future LHC Zooniverse project. We are looking into a Run2 version which will continue to allow the public to engage in the hunt for new physics at the LHC.

References

