

# **PlanetSense: A Platform for Gathering Real-time Geo-spatial Intelligence from Crowd-sourced and Social-media Data**

**PRINCIPAL INVESTIGATOR:** Gautam S. Thakur

**PROJECT START DATE:** 2015

**PARTICIPATING STAFF:** Kevin Sparks, Robert Stewart, Marie Urban

**SPONSOR:** US Department of Defense

## **PROJECT DESCRIPTION**

Authoritative sources of information and archived data are used to perform forensics on incidences happened in the past. Monitoring continuously, acting in real-time, and developing preemptive actions is the key to avoid them. Currently, no platform exists that supports end-to-end mechanism of large-scale harvesting of real-time data, perform necessary data fusion operations, and generate context-aware intelligence for thwarting events. Crowd-sourced and volunteered information, social media, and participatory sensors are capable of providing real-time activity data. Monitoring these sources in time of relevance and then using them to gather operational intelligence is an important approach for responders and proactive threat management. However, current techniques are limited to web-scraping, human copy-and-paste, semantic annotation recognizing, and vertical aggregation, which are illegal, slow, and non-scalable. In this work, we propose PlanetSense, a platform that is built to harness the existing power of archived data and add to that, the dynamics of heterogeneous real-time streaming data - seamlessly integrated with sophisticated deep machine algorithms and visualization tools for serving real-time intelligence, anticipatory threat discovery, mitigation and impact forecasting. This distributed and scalable platform strictly follows legit ways of harvesting very large amounts of data, while the actual deployment will be tailored for intelligence work. It has four main components – i) Hardware design, methods and algorithms to collect and process high-volumes of streaming data in a Geo-data Cloud; ii) Data fusion methods and algorithms for integrating disparate data sources with existing legacy products; iii) Design and development of data analytics and machine learning methods as a service; iv) Presentation and visualization through immersive web interface and REST data services. The information generated through intelligence has application in several important areas including national security, emergency management, cyber security, and topographical mapping

## **SIGNIFICANCE**

While acting based on historic data is effective, it entirely misses the opportunity to harvest real-time data and instantly act on it. We cannot deny the importance of archived data with the pursuit of gathering intelligence, the addition of real-time data will allow for a level of awareness and responsiveness to developing events that old data alone cannot provide. The proposed platform will create a system for continuous, real-time monitoring, anticipatory threat reduction in a more structured and scalable fashion. The proposed system will avoid illicit data collection approaches and respect privacy laws. It will follow the constraints imposed by data owners, while still enabling accurate, on-time intelligence. This will be the first demonstration of such a platform that is highly scalable, easy to use, and provide end-to-end support. This work is a direct response to LDRD's cross-cutting initiative of Anticipatory Threats & Impact Forecasting.