

Observing Changes in Real-Time Sensor Observations

Alejandro Llaves
Institute for Geoinformatics, University of Muenster
Weselerstr. 253, 48151 Muenster, Germany
alejandro.llaves@uni-muenster.de

Chris S. Renschler
NCGIA, University at Buffalo
116 Wilkeson Quad, 14261 Buffalo (NY), USA
rensch@buffalo.edu

Abstract

The increasing occurrence of extreme events, like the tropical storm Irene which caused several floodings on the East of New York State on September 2011, calls for novel methodologies to analyze data and exchange updated geospatial information in a fast and accurate manner. Sensor networks have been out there measuring properties of environmental phenomena for years. The access to these massive data sets is essential to build analysis applications capable of detecting relevant changes as they occur. Event processing techniques allow defining event patterns which can be used to detect changes in geosensor data in near real-time. This research work explores the connection between change detection in time-series of observations and flood occurrences in the Schoharie Watershed.

Keywords: Environmental Monitoring, Semantic Sensor Web, Event Processing, Semantic Interoperability.

1 Introduction

The term *event* refers in this paper to anything that happens or is contemplated as happening at an instant or over an interval of time which is relevant for the observer¹. The purpose of applying Event Processing techniques to geosensor data is i) detecting changes or patterns of change in continuous data streams and ii) abstracting higher level entities from simple observations. We deal with these inferred entities again as *observations*, considering the change detection filters as a special type of *sensors*.

According to the National Oceanic and Atmospheric Administration (NOAA) in the United States of America, an event is called extreme if it is from the tails of the climatological distribution, occurring, for example, only 5% or less of the time [11]. Extreme events - in a more general and geographic context of environmental systems - are spatio-temporal occurrences that are rare, not expected, or significant in terms of changing a system's functionality, e.g. a severe flood that changes the flow pattern of the stream. In 2011, the number of extreme weather events reported by NOAA costed at least one billion dollars and a historical record of twelve significant events, including tornadoes, droughts, and more². Environmental experts predict an increase in the number of disasters in the next years as a consequence of climate changing conditions. To analyze, understand, and prevent extreme events environmental monitoring is essential. We focus on in-situ sensor observations as a starting point to detect changes in environment and infer new knowledge.

The concept of Sensor Web [3] envisioned eleven years ago a

¹This definition is an adaptation of the OGC definition emphasizing the role of the observer.

²From the keynote by Jane Lubchenco "Predicting and Managing Extreme Events", at the American Geophysical Union meeting on December 2011, San Francisco, California.

network of spatially-distributed and interconnected sensing devices able to monitor uncertain environments. A decade later, innovation in sensing technologies and standardization efforts like the Sensor Web Enablement (SWE) [1, 2], have pushed the concept of Sensor Web to a mature level. The Semantic Sensor Web (SSW) [9] goes further and aims at a Web where sensor data includes spatio-temporal and thematic annotations to enable interoperability, and improve data analysis and information discovery. Having access to these sensor data sets which are continuously updated, there is a need for novel analysis techniques that allow change detection in a spatio-temporal context.

Complex Event Processing (CEP) [6] provides computing capabilities to read, create, abstract, and discard events [7]. CEP allows defining event patterns that can be matched against message-like data flows in real-time. One approach to analyze massive amounts of geosensor data provided in real-time is to deal with every single observation as an event. This is discussed from an ontological point of view in [8]. Observation events are geospatial events because they are situated in a spatio-temporal context [10] implicit in the observation process.

Some communities use the same term to categorize situations happening under different conditions. For instance, see below three different definitions for *heavy rain* based on rainfall intensity and rainfall accumulation:

- American Meteorological Society: rainfall intensity over 7,6 mm per hour or more than 0,76 mm in six minutes³.
- British Meteorological Office: rainfall intensity over 4 mm per hour⁴.

³See definition at <http://amsglossary.allenpress.com/glossary/search?id=rain1>

⁴See definition at http://www.metoffice.gov.uk/media/pdf/4/1/No._03_-_Water_in_the_Atmosphere.pdf

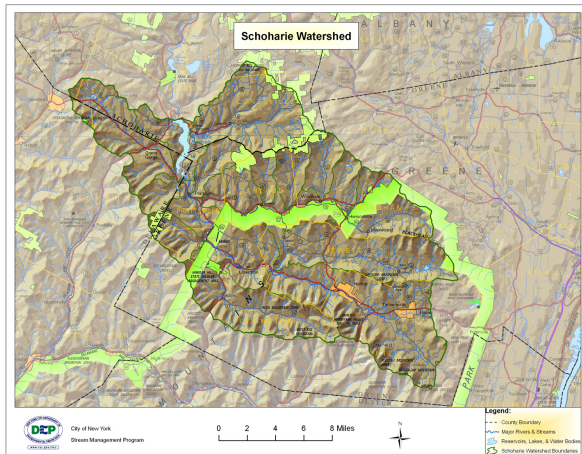


Figure 1: Map of the Schoharie Watershed

Source: <http://www.catskillstreams.org/>

- Central Weather Bureau of Taiwan: 24-hour accumulated rainfall exceeding 50 mm with at least one hour of 15 mm or above⁵.

Each organization uses different granularities and different thresholds, which can lead to misunderstandings at the time of comparing heavy rainfall records. Six minutes of precipitation can be enough to categorize a rainfall as *heavy* by the American Meteorological Society, whereas in Taiwan they need to measure the rainfall accumulated in one day. The knowledge about the categorization rules used to define environmental occurrences is essential to integrate data across different communities. Spatio-temporal models usually do not include the patterns of change used to identify relevant occurrences.

2 The Schoharie Watershed

On September 2011, after the pass of the Tropical Storm Irene, several populated areas were affected by floods at the Schoharie Watershed, New York (see figure 1). A post-disaster assessment of the watershed conditions and emergency management strategies reflected serious problems that could lead again to potential disasters in case of future heavy rains⁶.

USGS lists fourteen river monitoring stations which are measuring gage height and discharge at this area⁷. This data is provided in near real-time (normally, there is a difference of approximately one hour between the last observation and the current time) and available online through a USGS web service⁸. The National Weather Service of United States (NWS) defines the following flood categories for each registered forecast point⁹:

- *Minor flooding*: possible public threat, no property damage.

⁵See definition at Observations, Precipitations, Precipitation Classification at <http://www.cwb.gov.tw/eng/index.htm>

⁶Press release available at <http://www.buffalo.edu/news/12899>

⁷See *Stream Flow* link at http://cfpub.epa.gov/surf/huc.cfm?huc_code=02020005

⁸Available at <http://waterservices.usgs.gov/rest/IV-Service.html>

⁹Monitoring point within a watershed for which hydrological forecast and alert services are provided.

Table 1: Flood categories and gauge height thresholds (in feet) for the Prattsville site.

Flood Category	Threshold
Major flood stage	16 ft
Moderate flood stage	14 ft
Minor flood stage	12 ft
Action stage	9 ft

Source: Advanced Hydrologic Prediction Service (AHPS)

- *Moderate flooding*: inundations of road and structures close to the stream that could involve some evacuation.
- *Major flooding*: extensive inundation of road and structures, and massive evacuation of people.
- *Record flooding*: highest flood level recorded has been reached or exceeded.

Action stage is defined as the stage preceding the minor flooding category, and it is used by the Weather Forecast Offices (WFOs) to prepare potential responses in case of a flood occurrence. We consider that the river is on *normal stage* at a specific measurement point when the water level is below the action stage. The division between all these categories for each forecast point is described in terms of quantitative thresholds for water level, e.g. the minor flooding threshold for the forecast point “Schoharie Creek at Prattsville” is 12 feet (see table 1). For this study we considered the four forecast points providing continuous stream height observations within the Schoharie Watershed (from upstream to downstream): Prattsville, Gilboa Bridge, Breakabeen, and Burtonsville. Each monitoring station presents different thresholds for the flood categories.

3 Approach

To avoid misunderstandings when exchanging environmental information between different communities we have extended the existing W3C’s Semantic Sensor Network (SSN) ontology¹⁰ allowing to model spatio-temporal occurrences inferred from observations. A CEP engine will process the time-series of sensor observations and abstract higher level observations when the data matches the event descriptions.

3.1 Extending the SSN ontology

As we introduced in section 1, we consider the detection of a change in sensor data by means of Event Processing as a new observation. The namespace used for the new entities is *eo*, from *Event-Observation*¹¹. *dul* refers to concepts or relations from the DOLCE Ultralite ontology¹². *rdfs* relations are from the RDF Schema¹³. The extension consists of five new concepts and one relation which are depicted in figure 2. The concepts inherit the relations from their superclasses.

¹⁰See final report at <http://www.w3.org/2005/Incubator/ssn/XGR-ssn-20110628/>

¹¹Available at <http://purl.org/fifgi/event-observation>

¹²Available at <http://www.loa.istc.cnr.it/ontologies/DUL.owl>

¹³Available at <http://www.w3.org/TR/rdf-schema/>

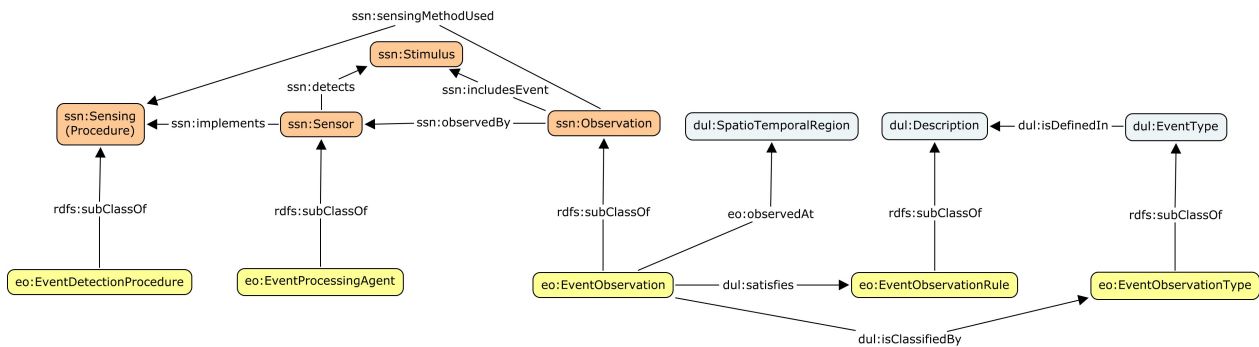


Figure 2: Concept map of SSN ontology extension.

An *eo:EventObservation* is an observed situation of change in the property of the geographical entity being observed. In the SSN ontology, an observation is a situation satisfying a description of the method that was used to observe it. We also describe the sensing method used in *eo:EventDetectionProcedure*, but additionally, we need a description of the situation itself and that is defined in the concept *eo:EventObservationRule*. This rule is used to trigger the change detection when the data analyzed meet certain conditions. An *eo:EventObservationType* can be related to different *eo:EventObservationRules*, hence this model helps solving the problem of different views on the same occurrence described at the end of section 1. The spatio-temporal location at which the *eo:EventObservation* was observed is inferred from the lower level observations processed by the *eo:EventProcessingAgent*. The event triggering the observation is represented as *ssn:Stimulus* and its identity is not affected by the different views. In section 4 we show an example of the representation of an observation of change using the extension presented here.

3.2 A CEP Engine for Observations

The Sensor Web allows online access to time-series of observations. To infer environmental occurrences using CEP we need to convert each individual observation into an *event*¹⁴ understandable by the CEP engine. This step is carried out by the *event producer* (following the terminology introduced in [4]). For this purpose we used Esper for Java¹⁵.

One of the main issues of most Web data services is that they usually follow a pull-based approach, i.e. synchronous request-response paradigm. Hence, to receive the sensor data continuously, “as it is observed”, it is necessary to simulate a push-based delivery by developing an application that requests the last observations regularly and sends them to the processing en-

¹⁴From [7]: “It is tempting to introduce two separate terms such as event and event object. However, in any discussion longer than a paragraph or two, this becomes intolerably clumsy and one finds the distinction being misused, forgotten or dropped altogether. For example, using the two separate terms would dictate that *event processing* (see below) should be *event object processing*. The best solution is to overload the word *event*. The context of each use becomes the indicator of which meaning is intended. This has been standard practice in the field of event-driven simulation for the past thirty years. It was the approach taken by the physicists of the early 20th century in discussions of relativity where *event* also has two meanings. We have chosen to follow their example in the knowledge which did not lead them into ambiguity problems.”

¹⁵Open source software available at <http://esper.codehaus.org/>

gine. The stream of events generated by the *event producer* pass through the *event processing agents* which are able to filter, transform, and/or detect patterns [4]. Each *event processing agent* may include various *event descriptions* (also called statements or event patterns) encoded in Event Processing Language (EPL). An *event consumer* is implemented as a listener of one or more *event channels*, where the events flow.

Our approach consists of “linking” each *event observation rule* to the URL of the event concept represented in the corresponding domain ontology. For instance, an *event processing agent* can have three different heavy rain descriptions defined by three organizations (see example in section 1), but each description will point to the same *heavy rain* concept in a domain ontology. When a group of incoming *events* (transformed observations) matches an *event observation rule*, a new instance of the domain ontology concept referred by the description is created and added to the ontology. Such instance contains the information inferred from the observation data (spatio-temporal location, observed property, etc.) and a description of the property parameters used to detect the change, as presented in the model of previous section.

4 Example of Use

The Schoharie Watershed lacks of a reliable and user-friendly floodwatch system to be used by citizens and domain experts. To test our research approach on change detection for environmental monitoring purposes, we decided to work on a mockup of the watershed. Next, we describe our first steps towards a floodwatch system for Schoharie.

The input of the floodwatch system consists of time-series of the gauge height observations corresponding to Prattsville, Gilboa Bridge, Breakabeen, and Burtonsville, which are published online through a Web service (see section 4). Our application retrieves this data hourly and converts each observation into an *event* to be processed by the CEP engine. Such *event* includes the gauge height value, information related to the monitoring station where the sensor is located (identifier, name, and spatial location), and the point in time at which the measurement was taken.

The stream of *events* pass through a pattern detection process to identify specific situations of interest. Eight event descriptions have been created for each one of the four forecast points within

the Schoharie Watershed. They represent the eight possible transitions between normal stage, action stage, minor flooding stage, moderate flooding stage, and major flooding stage¹⁶:

- Normal stage to Action stage
- Action stage to Minor flooding stage
- Minor flooding stage to Moderate flooding stage
- Moderate flooding stage to Major flooding stage
- Major flooding stage to Moderate flooding stage
- Moderate flooding stage to Minor flooding stage
- Minor flooding stage to Action stage
- Action stage to Normal stage

The following snippet shows how the “Action stage to Minor flooding stage” event description for the Prattsville site looks like in EPL:

```
SELECT obs1, obs2
FROM pattern [every (
  obs1=StreamGaugeObservations(obs1.value < 12)
  ->
  obs2=StreamGaugeObservations(obs2.value > 12)
)]
WHERE (obs1.sensor.id='01350000') and
(obs1.sensor.id=obs2.sensor.id)
```

In the SELECT clause we define the event attributes we want to use from the event stream of observations. In this case, we take all the data enclosed in the two *events* that matched the pattern. The FROM clause specifies the type of pattern we are looking for, i.e. every pair of events (*obs1*, *obs2*) from the event producer StreamGaugeObservations with the value of *obs1* below 12, and the value of *obs2* above 12, which is the threshold for the minor flooding stage at the Prattsville site (table 1). The arrow (->) means that *obs2* follows *obs1* in time. It is assumed that the observations are sent as they are observed, meaning that they are temporally ordered. In the WHERE clause we filter by the sensor identifier (*obs1.sensor.id*) of Prattsville and we specify that *obs1* and *obs2* must be produced by the same sensor.

All the event patterns are registered to the CEP engine, having a different event processing agent per forecast point (since thresholds are different for each site). When the incoming *events* match one of these patterns, an event instance is created and stored in a repository (see figure 3). A transition from action stage to minor flooding stage has been detected by the *EventProcessingAgent* for Prattsville. A new *EventObservation* instance *obs3* is created. The observed situation satisfies the event observation rule described above and belongs to the *ActionStageToMinorFloodingStage* type. The trigger was a change in the water level which increased above 12 feet. The spatio-temporal location of the observed situation is inferred from the timestamps and sensor locations of the observations analyzed. Some concepts in the graph belong to a simple flood ontology (namespace: *f*) we are developing for the mockup¹⁷.

Each forecast point has a website where the current state of the stage is shown (see figure 4). The chart is updated in near-real time with the same data used for the event processing. Our example of floodwatch system has been developed to detect the

¹⁶We do not consider the record flooding stage at the moment.
¹⁷Available at <http://purl.org/ifi/water/flood>

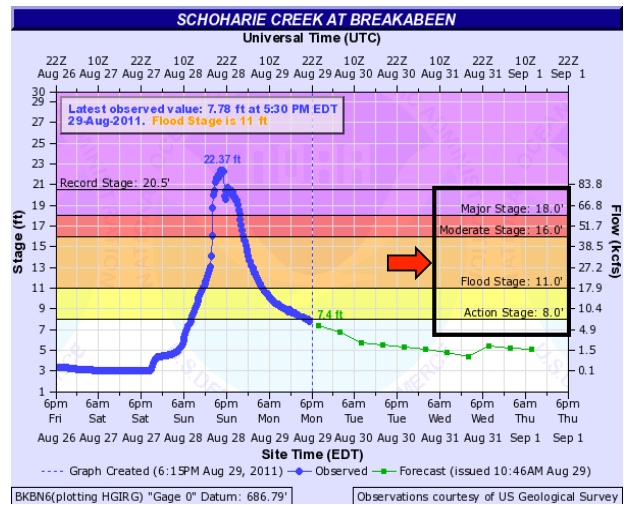


Figure 4: Chart showing real-time gauge height (in feet) observations for the Breakabeen monitoring station. In the background we can distinguish different areas representing the flood stages. The thresholds are inside the black box at the right side.

transitions between the flood stages in an integrated way. We can now detect and model in near real-time transitions between two flood stages (e.g. action stage to minor flood stage) pointing to the same “ontological concept of occurrence” but with different descriptions for the changes in the observed property.

5 Conclusion

This paper presented ongoing research work on a novel methodology to process and model spatio-temporal occurrences in the field of environmental monitoring. It introduces some variations in the way that observations are analyzed and integrated in the Semantic Sensor Web (see related work at [5]). The use of the SSN ontology as the foundation stone to build our extension is because of such model is the result of a thorough revision of several existing observation-centric and sensor-centric ontologies. We strongly believe that next research steps on the ontological representation of observations should be focused on the adoption and extension of the SSN ontology.

The approach introduced here solves the interoperability problems of different communities using the same term to refer to different occurrences. For that purpose, we find essential the inclusion of the situation’s description in the observation models and not only the description of the sensing method used. In terms of provenance, this new model can be useful to *trackback* inferences made over time-series of observations. The application of event processing techniques, like CEP, to geosensor data provided in near real-time looks promising. Yet, we think that more research needs to be done in this aspect and our processing methodology aims to advance in this direction.

One aspect we have to address next related to event processing is the temporal order of the observations. In our scenario, we request data from a Web service which collects measurements from different sensors. If the sensor presents problems, the observa-

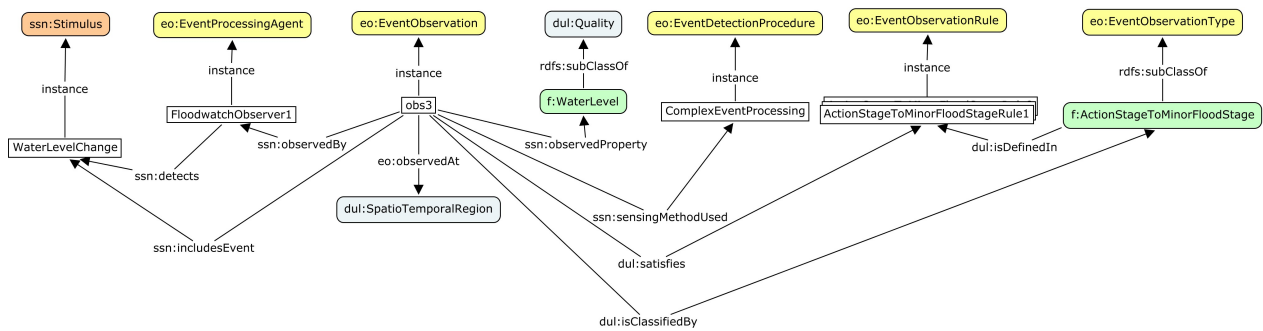


Figure 3: Representation of observed change “Action stage to Minor flooding stage” at Prattsville monitoring station

tions might be published in the Web service with some delay. An event processing agent responsible for sorting the individual observations before sending them to the pattern detection agent has to be developed. As part of the future work, the idea is to develop a semantic notification infrastructure to allow users subscribing to alerts and receive notifications in near real-time. Another line of research points to making available our event repository online so that it will be populated as the observations are processed. This way, we can connect our dataset to the Linked Data cloud. Event records can be used later to get a better understanding of complex environmental phenomena, e.g. by correlating flood occurrences with heavy rainfalls or by analyzing the relations between flooding situations from upstream to downstream locations.

Acknowledgments

The presented work is funded by the European project ENVISION¹⁸ (FP7-249170).

References

[1] M. Botts, G. Percivall, C. Reed, and J. Davidson. Ogc sensor web enablement: Overview and high level architecture. *Lecture Notes In Computer Science*, 4540:175–190, 2008.

[2] Arne Broering, Johannes Echterhoff, Simon Jirka, Ingo Simonis, Thomas Everding, Christoph Stasch, Steve Liang, and Rob Lemmens. New Generation Sensor Web Enablement. *Sensors*, 11(3):2652–2699, March 2011.

[3] Kevin A Delin. Sensor web: a new instrument concept. *Proceedings of SPIE*, 4284(January):1–9, 2001.

[4] Opher Etzion and Peter Niblett. *Event Processing in Action*. Manning Publications, 2010.

[5] Cory A. Henson, Josh K. Pschorr, Amit P. Sheth, and Krishnaprasad Thirunarayan. SemSOS: Semantic sensor observation service. *Collaborative Technologies and Systems, International Symposium on*, 0:44–53, 2009.

[6] David Luckham. *The Power Of Events*, pages 146–147. Addison-Wesley, Boston, USA, 2002.

[7] David Luckham and Roy Schulte. Event processing glossary - version 2.0, 2011.

[8] Florian Probst. Ontological Analysis of Observations and Measurements. In M. Raubal, H. Miller, A. Frank, and M. Goodchild, editors, *Geographic Information Science, 4th International Conference, GIScience 2006*, pages 304 – 320. Springer, Number 4197 in LNCS, 2006.

[9] Amit Sheth, Cory Henson, and Satya S Sahoo. Semantic sensor web. *IEEE Internet Computing*, 12(4):78–83, 2008.

[10] Michael Worboys and Kathleen Hornsby. From objects to events : GEM , the geospatial event model. In M Egenhofer, C Freksa, and H Miller, editors, *Third International Conference on GIScience 2004*, pages 327–344, Berlin, 2004. SpringerVerlag.

[11] Yuejian Zhu and Zoltan Toth. Extreme weather events and their probabilistic prediction by the ncep ensemble forecast system. In *Preprints of the Symposium on Precipitation Extremes: Prediction, Impact, and Responses*, Albuquerque, New Mexico, 2001.

¹⁸<http://www.envision-project.eu>