Dagstuhl Seminar: Semantic Challenges in Sensor Networks

Karl Aberer* Avigdor Gal† Manfred Hauswirth‡
Kai-Uwe Sattler§ Amit P. Sheth¶

January 24–29, 2010

1 Executive Summary

There has been significant progress in the number and capabilities of mobile devices, wireless sensors, and sensor networks. These developments, combined with the improved ability to bridge between the physical and cyber world in a more seamless way, have fostered the broad availability of sensor data capturing the state of the physical world. Promising and already successful examples are applications in environmental monitoring, agriculture, surveillance and intrusion detection, public security, and supply chain management. Furthermore, ideas towards a Web of sensors have been proposed, which is to be understood as a (large scale) network of spatially distributed sensors. In particular, terms like "Internet of Things", "Collaborating Objects" and "Ambient Intelligence" emphasize the trend towards a tighter connection between the cyber space and the physical world.

The existence of a huge number of sensor sources – 40 billion by one estimate – may of which continuously produce data results in tremendous data volumes. The observations, reported by many sensors, are often valid or useful only for a certain period of time and are never inspected by humans. In order to make sensor data useful, even without human supervision, semantic annotation and analysis becomes a key component in setting up sensor data-based applications: Only if sensors and sensor data are annotated and enriched by information describing their meaning, as well as provenance, including source, and validity scope, they can be automatically discovered, processed and combined with other data in an open world. The types of useful semantics range from technical metadata describing the sensors and the measurements (time, location, sensor type, validity, measurement error, etc., as partially captured by standardization proposals like SensorML) to emergent semantics derived by aggregating, combining, analyzing, and enriching the raw data, e.g., in the form of analytical models, annotations, correlations, etc. In comparison, the data collected by human-in-the-loop sensing is small but of significant variety and complexity (e.g., language nuances, and capturing sentiments and emotions), which add more challenges (together with more opportunities) to the problems of annotation, integration and analysis of such data.

Modeling, representing, discovering and deriving as well as using semantics for sensor data raises several challenges which are related to different aspects of de-

*EPFL, Lausanne, CH
†Technion,Haifa, IL
‡DERI, Galway, IE
§TU Ilmenau, Ilmenau, DE
¶Wright State University, Dayton, US

veloping, deploying, and using sensor-based applications. Thus, the goal of this seminar was to bring together researchers from relevant areas, such as:

- sensor node providers and sensor networking,
- data fusion and data stream processing,
- sensor middleware,
- geo-spatial and uncertain data management,
- semantic integration and Semantic Web, and
- social computing and collective intelligence.

Semantics plays an important role in all of these areas, either by producing and enriching data with explicit semantics or by exploiting semantics for data processing. Therefore, sharing and exchanging knowledge and experiences among disciplines could result in significant synergy effects.

The seminar focused on the following major issues:

- methodologies and languages for modeling and representation, issues of sensing-perception-semantics,
- standards, ontologies, and middleware for semantic sensor networks, semantic annotation of sensor data as well as social/human-in-the-loop sensing data,
- emergent semantics in sensor networks and sensor data processing,
- exploiting sensor data semantics for geo-spatial and uncertainty data management,
- specific use cases and applications of semantic sensor networks, and
- review of related community efforts that directly relevant to the seminar topic, especially W3C's XG on Semantic Sensor Networking,

The objectives set out by the organizers were to analyze the state of the art in the different areas with respect to semantics, discuss problems, specific methodologies and applications of semantic-aware sensor networks and emergent semantics as well as to identify future trends and research directions.

The presentations and follow-up discussions and working sessions helped the participants understand the different viewpoints on problems and provided interesting insights on how different communities represented in the seminar try to address some of the problems and what the effects on other research areas would be. The highlighted areas were:

Uncertainty: How can uncertainty be described and managed and what are the effects on logical statements based on aggregations of uncertain data and statements?

Infrastructures: To decrease deployment costs and cope with the involved complexities, a number of approaches and abstractions for infrastructures at various levels – data representation, query processing, metadata (semantic annotations), middleware, virtual observatories, etc. – where presented and discussed. It was interesting to observe similarities of approaches across disciplines while the abstractions differed significantly. We recognized it will take additional collaborative effort to understand how to reconcile these differences.

Applications: Specific application domains of sensors and sensor data have very heterogeneous, often contradicting requirements. A number of application examples, ranging from environmental monitoring to disaster management to the personal sphere, where presented. A number of efforts represented significant scale in terms of number of sensor and amount of data generated, with additional opportunity for better integration and analysis of sensor data.

Semantics: Value of using semantics to integrate, analyze and understand sensor data as well as to devise more efficient infrastructures based on this understanding was largely agreed on. The area of semantic descriptions of sensors and sensor data is currently an area of very active research which tries to define expressive ontologies which can represent basic concepts like observations and modalities like space and time while still supporting efficient processing. One specific relevant activity discussed in a panel was that of W3C Incubator on Semantic Sensor Networking, which is developing a Sensor Ontology and guidelines for semantic annotations of sensor data.

Social Sensing: There are a number of case studies on using humans and their associated devices as sensors. The seminar showcased a number of interesting applications of social sensing and proposals for infrastructures to support this.

The seminar was well attended: 27 researchers from Europe, Asia, and North America actively contributed to the seminar. During the week, two special discussion sections were organized where the seminar was split in smaller groups. The topics of these discussion groups were "Data Representation & Semantics", "Query Models", "Architectures for Semantic Sensor Networks", and "Application Requirements".

At the end of the seminar a joint session with the parallel Dagstuhl seminar "Digital Social Networks" was held to explore research topics at the intersection of both research domains, for example, the use of social network infrastructures to discover and publish sensor data, the problem of privacy at the intersection of sensor networks, mobile phones and social networks, and the use of social networking in social sensing. There was enthusiasm to organize a follow-on seminar bringing together the two area and a proposal for a Dagstuhl Seminar is planned.