Towards Elastic Sensor Networks
A Position Statement for Information Management in Security Applications and Beyond

Prof. Yike Guo, Dr. Moustafa Ghanem,
Department of Computing, Imperial College London

Context within Grand Challenges

Our submission focuses primarily on the new **security** challenge area. It also relates to the **ubiquitous computing** challenge area. Our focus is on dealing with challenges related to information overload and associated resource management issues in large-scale sensor networks where large amounts of information is captured and processed. Our goal here is to highlight the need for investigating and developing novel theoretical frameworks and associated computational models for information management in Large-scale Sensor Networks (LSSN).

Applications in Security and Beyond

Our own motivations for research in this area are drawn from ongoing research experience and collaborations in two application domains for large scale sensor networks: security monitoring and urban pollution monitoring.

Our work in the security area is in collaboration with the Institute of Security Science and Technology (ISST) at Imperial College London. The aims of the Institute are to apply leading edge science to develop new technologies for increased safety and security in society across a range of scales, from protecting the individual to ensuring the security of whole populations. The activities of ISST encompass the physical, engineering and biomedical disciplines as well as computer science. Our collaboration with the institute covers various topics relating to applying analytics to the analysis of large amounts of data to focus on the UK's national priorities. These include applications for **combating cyber crime**, applications for **national infrastructure monitoring scenario,** applications for **crowded places monitoring** and also **terrorist activity tracking.**

The analysis of networked data, and in particular data captured by sensor network technologies, plays an important role in all these applications. Our previous work in the area of urban pollution monitoring using mobile sensors provides us with a starting point for infrastructure and technology. These have been developed conducted in the context of the EPSRC funded Discovery Net and MESSAGE projects. In Discovery Net it was based on collecting data from static sensors and designing a centralized warehousing and analysis framework. The work was extended within the MESSAGE project by developing the infrastructure and algorithms that support in-network mining techniques. A key contribution of this work has been demonstrating how to support distributed data management and analysis using a mix of mobile and static sensors. Another key contribution was the design of real-time distributed data analysis algorithms whose accuracy results are competitive with that of centralized algorithms. These developments can be direct applicable to large scale application with significant social importance. For example, a mobile MESSAGE-like system measuring multi-modal sensor data deployed across national infrastructure areas, e.g. the London Olympics area, can be easily developed and deployed for monitoring the pollutions as well as other environmental measurements.

Challenges for Information Management in Large-Scale Sensor Networks

The key challenge for using large-scale sensor networks in security and other applications is not just an engineering challenge relating to deployment of the network or collecting the data. Rather it how to analyze, and reach on the complex and distributed information in real-time. Sensor units are deployed over wide geographic areas. Each sensor can be equipped with multiple sensing devices that measure different variables, Each of the sensor units typically has on-board processing units and can perform local data processing and analysis operations as well as exchange data with other sensor units and base stations. The two key challenges in this context are:

- **Sensor Network Information Management:** As the amount of information monitored by sensors increases a key challenge becomes avoiding information overload in the sensor network, i.e. organizing information collection and processing to focus on analyzing only information relevant to the user needs. Typical decisions include what information each of the units should be collecting, at what rates, how and where it is processed, summarized and stored, what information should be exchanged between the sensor units and how all such information should flow within the network itself. Tradeoffs in addressing these decisions arise based on whether only local information collected from one unit or global information collected from many units is needed to inform the required decisions.
- **Sensor Network Resource Allocation:** With a finite number of sensor units, and each having finite resources (processing capacity, memory and storage size, communication bandwidth and battery power), a second challenge is maximizing the value of collected information under resource and real-time constraints. Tradeoffs occur in addressing this challenge since allocating a group of sensors to exploring a particular geographic region would mean fewer resources available for exploring other areas. Similarly, within one region, assigning more computational resources to exploring the features of a particular event imply less capacity for others.

Developing Novel and Sound Theoretical Frameworks

Our own research is focusing on investigating a number of new mechanisms for

1. Developing methods to identify, and focus on, relevant information within a complex sensor network environment to **avoid information overload**; and
2. Developing an **elastic computation model** that realizes an **on-demand and real-time resource provision** model for data acquisition and information processing in large-scale sensor networks.

We have been recently investigating a number of formalisms for addressing both problems. These include investigating the use of other biologically-inspired models including those based on swarm intelligence including honey bee foraging, ant foraging bacterial quorum sensing. They also include Bayesian formalisms that model human attention mechanisms. The mathematical foundations of the Bayesian approach enable us and other researchers to reason about correctness of the attention mechanism and reason about its extensions. All these techniques offer different advantages. Evaluating their utility success can be achieved only through close collaboration with end users and based on real-case studies and real data sets.