Using compression for the analysis of distributed sensor networks

January 7, 2014

1 Motivation

We consider situations where we have a sensor network that is distributed over a large area. Our goal is to perform continuous analysis of the data generated by the sensors. The goal of the analysis is to detect abnormal behaviour and to create predictive models of the behaviour of the physical system (PS) that is monitored by the sensors.

We assume the following generic architecture, consisting of sensor nodes and compute nodes. Sensor nodes consist of some sensors a general purpose computer and local storatge. Compute nodes consist only of computers, potentially stronger than the computers in the sensor nodes. The compute nodes communicate with each other and with the sensor nodes. The compute nodes aggregate information from many sensors in order to estimate the global state of the PS, and to create predictive models for the dynamics of that PS. The main constrained resource is the communication bandwidth between the nodes.

As an example consider an ad-hoc sensor network consisting of smart-phones. The smart-phones communicate with a network of computers through cell phone connections and the internet. The amount of data generated by a sensor such as a video camera easily exhausts the bandwidth cellular communication network. Even when the bandwidth is not exhousted it is usually the most expensive part of operating the system, both in terms of data-communication costs and in terms of battery life.

It is therefor very desirable to design a system which operates in such a way as to minimize the amount of communication between the sensors and the compute nodes. A common approach is to use *lossless compression*. This is a good solution when possible, but it rarely decreases the communication volume by a factor bigger than 4. Our goal is to design method that will decrease the communication volumes by a factor of ten or more.

To achieve such rates we need to look towards *lossy compression* methods. When data is compressed and decompressed using a lossy compression method, the result is a *distorted* version of the original data. We say that a compression method is good if a small data *rate* (i.e. the bandwidth required to carry the compressed data) is enough to achieve low expected *distortion*. The foundational theory of lossy compression is Shannon's Rate-Distortion theory, which characterizes the achievable rate-distortion pairs.

The way in which distortion is measured is sometimes of critical importance. For example the use of *perceptual coding* in MP3 reduces the number of bit devoted to encoding frequencies to which the human ear is less sensitive.

The basic idea behind this work is that the data coming from sensors about a PS can usually be seen as a sum of two parts: signal and noise. Usually written as:

$$f(t) = s(t) + \sigma(t)w(t)$$

where $s(t)$ is the signal as a function of time $w(t)$ is white noise and $\sigma(t)$ is the amplitude of the noise. White noise is not compressable, in other words, no method of coding can decrease the bandwidth required to send it. Happily, white noise also carries no useful information about the PS, only $s(t)$ and $\sigma(t)$ carry useful information. In addition, $s(t)$ and $\sigma(t)$ are usually highly compressible.

Our plan is therefor to partition the signal from a sensor (or a set of sensors) into the useless and uncompressible noise part and the useful and compressible signal and noise-amplitude part.