INVITED

FROM INFORMATION THEORY TO INFORMATION PRACTICE

DR SINAN SINANOVIC
GLASGOW CALEDONIAN UNIVERSITY

Modern society and science in particular are inundated with the data: due to wireless communications, megabits of data are riding electromagnetic waves around us each second; in genetics, gigabytes of DNA data are obtained at ever increasing rates; in finance, constant activity of stock market and foreign exchange provides a wealth of data; in medicine, huge amount of the data is collected from X-rays, MRI (magnetic resonance imaging), EEG (electroencephalography) and ECG (electrocardiogram). One of the key questions common to all of the above areas is: Where is the meaning in all the data? More specifically: Can we somehow automate the process of extraction of the interesting information from the wealth of noisy data when a change, an anomaly or a new trend appears? The traditional approach is to employ experts to analyse the data and extract useful information but this is expensive, time-consuming or sometimes simply impossible. That is why it is crucial to further apply the computing power which is available today in order to help us find the features, trends, anomalies and other interesting information which may be buried in the data. Since the data in general tends to be noisy, multidimensional and otherwise complex in its structure, the question becomes: How do we harness the computing power to process huge amounts of available data in order to extract the meaning and augment our understanding of the underlying hidden phenomena? One promising answer lies in using the information-theoretic tools. One possibility is to use the Kullback-Leibler divergence for comparison of the empirical pdfs (probability density functions). The data may have pilot ('clean' or 'normal') segment which is used to build the empirical pdf via histogram or, where smoothness is the issue, kernel density estimation method. Then, as the new data arrives, its empirical pdf is also obtained. The two pdfs can be compared in real time via divergence. Two similar sets of data will have small divergence value while dissimilar data will have larger value. The advantage here is that one does not have to decode the data or have a priori knowledge of the underlying data in order to capture new trends, anomalies or features which started to evolve or appeared momentarily in the data. Another practical advantage is that the data can be severely downsampled (by a factor of 100 or more in some instances) without compromising the quality of the detection. This is especially useful for the hardware implementation since the sampling rates can be on the order of tens of MHz. Our method, which has also been implemented in hardware, verifies robustness of the information-theoretic approach when used on the actual (not simulated) wireless data recordings. In cases where the comparison of two segments of data is infeasible, perhaps due to the inherent non-stationarity of the data, the information-theoretic approach can be extended to detection of the cluster of rare events. The method has already been successfully applied on the actual wireless data recording. Finally, we briefly look at fields outside of wireless communication, such as particle physics, image recognition and plagiarism detection where the information-theoretic tools we have developed have also been successfully applied.

Sinan Sinanovic has joined Glasgow Caledonian University in 2013. He has obtained his Ph.D. in electrical and computer engineering from Rice University, Houston, Texas, in 2006. In the same year, he joined Jacobs University Bremen in Germany as a post doctoral fellow. In 2007, he joined the University of Edinburgh in the UK where he has worked as a Research Fellow in the Institute for Digital Communications. While working with Halliburton Energy Services, he has developed acoustic telemetry receiver which was patented. He has also worked for Texas Instruments on development of ASIC testing. He has published over 50 papers in the areas of information theory, MIMO, interference management and visible light communication. He is a member of the Tau Beta Pi engineering honour society and a member of Eta Kappa Nu electrical engineering honour society. He won an honourable mention at the International Mathematics Olympiad in 1994.