Sensed Outlier Detection for Water Monitoring Data and a Comparative Analysis of Quantization Error Using Kohonen Self-Organizing Maps

E. M. Dogo, N. I. Nwulu, B. Twala, C.O. Aigbavboa

Measurement values obtained from sensors deployed in the field are sometimes prone to deviation from known patterns of the sensed data which is referred to as outlier or anomalous readings. The reasons for this outlier may include noise, faulty sensor errors, environmental events and cyber-attack on the sensor network, resulting in faulty and missing data that greatly affects quality of the raw data and its subsequent analysis. This paper employs the Self-Organizing Maps (SOM) algorithm to visualise and interpret clusters of sensed data obtained from fresh water monitoring sites, with patterns of similar expressions in a graphical form. With the aim of detecting potential anomalous sensed data, so that they could be investigated and possibly removed to guarantee the quality of the overall dataset. Furthermore, a comparative study of the effects of four different well known neighborhood functions (gaussian, bubble, triangle and mexican hat) with varying neighborhood radius (σ) and learning rate (η) values on Quantization Error (QE) metric was conducted. From the experiment conducted a 3.45% potentially anomalous sensed data were discovered from the entire dataset, in addition, our initial finding suggests a very insignificant variation of the QE based on our dataset and the experiments conducted.

https://ieeexplore.ieee.org/abstract/document/8769276