Crowd Sourced Environmental Monitoring

Briana Ausby, Paul Stevenson, Dr. Sule Ozev

Abstract— **As respiratory disease patients will have our wearable devices to monitor their health it will also monitor their current environment for current temperature, ozone, and particulates. Inexpensive sensors will be used to collect the data, larger numbers of the sensors will ensure reliability. With constant collecting of sizable amounts of environmental data we propose to investigate a proper algorithm to gather and correlate the spatial and temporal data of the outdoor pollution.**

I. Description of Your Project

Crowd Sourced Environmental Monitoring (CSEM) is the gathering current environmental data from an engaged crowd using their smart devices. A group of users that will have these wearable Air Care System will not only monitor their health but the outdoor pollution surrounding them. Specifically, we will keep track of the temperature, ground-level ozone, and particulates, monitoring when these pollutants are high. These devices will be using inexpensive sensors to make these devices more affordable. However, to ensure accuracy the sensors will be used in large numbers. Constantly receiving sizable amounts of data requires a strong algorithm to correlate the information we are receiving and remove the outliers to return the most accurate information for the users.[3]

With crowd-sourced data gathering we will create a fine-grained air-quality map for public use. We are providing sensors for multi-sensing system and tracking data with free range people. We upload the data to the cloud and interpret data to create a 3D map. It's necessary to have accurate information. We will face issues such as lack of control over nodes (lcon), data fusion (df), sensor drift (sd), autonomous node operation (ano), and correlating spatiotemporal awareness.[3] Therefore we need to investigate a proper algorithm for sever and node level that is low computation power and efficient.

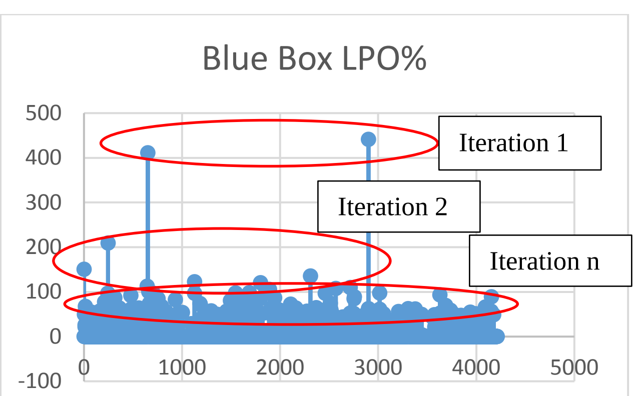

Figure 1: Initial particle test of one device. There are obvious outlier. However the device will iterate over and over to remove less obvious outlier.

Outliers will have large effect on our data collection. For example, sd, where a sensor will return inaccurate information in the data set. We need to identify it and provide feedback of sensor calibration or replacement. Sometimes the sensors will input information that will affect the collection. A person wearing the device could walk by a fire or someone who is smoking and variate the information. These are examples of outliers.

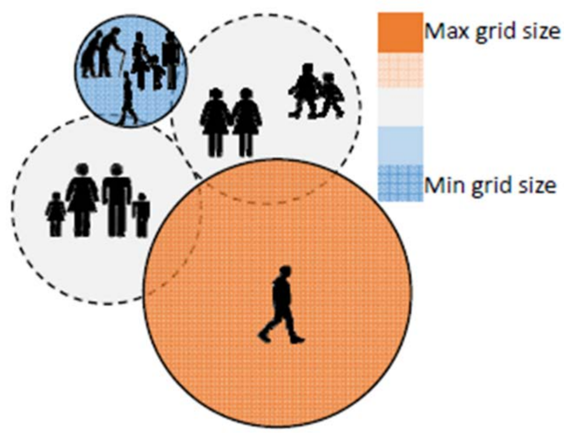

Figure 2: Dynamic grid size determination based on minimum and maximum grid size and the number of users inside the grid [3]

With our prototype algorithm we will take in the sample data and attempt the window approach. We will have a confidence interval within a sigma range that will continuously adjust as we iterate over the data. The code will eliminate all samples outside the confidence interval. The sigma will continue to change and it will eliminate more outliers until the data is strong and confidently accurate. (Figure 1)

References

[1] Alvear, O., Zamora, W., Calafate, C., Cano, J. C., & Manzoni, P. (2016). An Architecture Offering Mobile Pollution Sensing with High Spatial Resolution. Journal of Sensors, 2016.

[2] Cheng, Y., Li, X., Li, Z., Jiang, S., Li, Y., Jia, J., & Jiang, X. (2014, November). AirCloud: a cloud-based air-quality monitoring system for everyone. In *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems* (pp. 251-265). ACM.

[3] Christen, J. B., PhD, Ozev, S., PhD, & Ross, H., PhD. (2017, March 06). CYBER-PHYSICAL SYSTEMS (CPS) [Scholarly project]. Retrieved June 05, 2017.

[4] Cuckov,Filip. "Dependability in the Context of Human-in-the-Loop CyberPhysical Systems".Presentation.