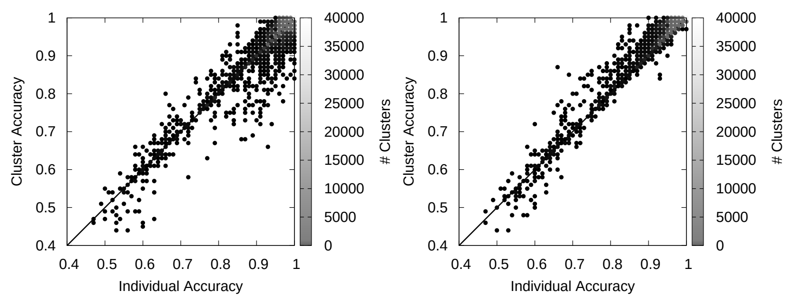
Confident Sensor Collaboration with Machine Learning

Gang Zhou
Computer Science Department, College of William and Mary

This white paper visions two rising research challenges that are fundamental to pervasive computing at scale: (1)how to collaborate sensors of the same modality as well as different modalities for achieving user specified sensing confidence? This is very challenging since real sensors' sensing ranges as well as capabilities are very irregular [1] or diversified [2] and hence can not be well modeled [3] but can be carefully learned. (2)how to exploit different machine learning methods for sensor collaboration in different network topologies for providing such confidence?

1 Initial Evidence: Impact of Sensing Diversity on Collaboration

To understand how sensing diversity impacts sensor collaboration, we compare Nearest Centroid [2], a variant of K-Means clustering, with another machine learning technique, Fisher's Linear Discriminant. While Nearest Centroid assumes that each reading in a tuple of sensor cluster data is independent, Fisher's Linear Discriminant attempts to capture the dependencies among different sensors in a cluster. We use the trace data from the Wisconsin vehicle tracking deployment [4] in which 23 sensor nodes are deployed along a road, each containing an acoustic, seismic, and infrared sensor. We select 103 vehicle detection locations throughout the deployment area. For each target location we form random clusters of size 2 through 25, with up to 30 clusters of each size, using sensors within 100m of each target location. For each generated cluster we compare the detection accuracy of the best individual member sensor as a singleton cluster to the generated cluster accuracy and plot the results in Figure 1.



(a) Nearest Centroid (b) Fisher's Linear Discriminant

Figure 1: Cluster accuracy vs. best individual accuracy

We discover that sensor collaboration should be conducted on-demand. First, when an individual sensor meets the user detection requirements for a target location, collaboration is unnecessary. We show that in over 36,000 cases, for both Fisher's Linear Discriminant and Nearest Centroid, individual sensors have perfect accuracy. Individual sensors have over 95% accuracy in 2,000 more cases. When the user requirements fall in this accuracy range, collaboration is not needed. When an individual sensor cannot meet the user requirements for a given target location, collaboration is needed, but there are individual sensors that can be excluded from the collaboration process to reduce the search space size. In Figure 1, with this specific deployment, it is clear that no individual sensor can boost cluster accuracy as a cluster member by more than 20% points. These sensors can be excluded from detection and collaboration as any cluster consisting entirely of sensors below this threshold will not meet the user detection requirements. Lastly, we show that when a sensor has a detection accuracy above the sensitivity threshold, but below the user requirements, it is a candidate for collaboration in a sensor cluster. Therefore, to provide confident event detection and classification, sensors need to to be carefully collaborated using machine learning to meet user defined accuracy requirement while at the same time minimize minimize the involving sensors to save energy and extend system lifetime.

2 Exploiting Machine Learning for Sensor Collaboration

Different network topologies may be used for different pervasive computing applications. Taking a smart healthcare and assisted living application as an example, a cluster topology is usually used for individual Body Sensor Networks (BSN) to monitor human behavior, and an ad-hoc topology is usually used for Emplaced Sensor Networks to assist monitoring the context environmental. When we view the whole system vertically, it consists of a Hierarchical Sensor Network. One key challenge is to use the right machine learning methods to collaborate the right sensors in the right networks for providing user requested detection and classification accuracy while at the same time minimizing energy consumption. Due to space limit, we can not elaborate all the detailed plans but just illustrates one detailed issue we plan to investigate for BSN: in different BSN applications, how to trade off accuracy and cost between different machine learning methods? Such methods include K-Means, Nearest Centroid, K-Nearest Neighbor, Hidden Markov Model (HMM), Fisher's Linear Discriminant, and even Support Vector Machine (SVM). K-Means and its variations are simple and energy efficient, while HMM and its variations capture time dependencies without ground truth, but are more costly. Fisher's Linear Discriminant addresses data dependencies due to spatial and time correlation, but is also more expensive. SVM has solid theoretical foundation and empirically shows very good performance in most scenarios, but is much more expensive and hence usually used for offline training.

3 The Background and Experience of the Participant

One notable system Dr. Zhou co-built is VigilNet which integrates network-level tripwire, section-level sentry service, and node-level duty cycle scheduling to provide long-term military surveillance. Dr. Zhou also has strong system design and field deployment experiences related to assisted living and smart healthcare which is reflected by the AlarmNet assisted-living and residential monitoring network he co-built. In addition, Dr. Zhou also has expertise in other related areas like body sensor networks, confident sensing and event detection, and quality of service of low power communication and networking. Dr. Zhou's strong system background is also reflected by his highly cited system papers in MobiSys, SenSys and RTSS. Dr. Zhou also has a solid background in generalizing the system practices into principles and theories, as signified by his 9 INFOCOM papers from 2005 to 2010 and his Best Paper Award in ICNP 2010.

References

[1] J. Hwang, T. He, and Y. Kim, "Exploring In-Situ Sensing Irregularity in Wireless Sensor Networks," in *ACM SenSys*, 2007.

[2] M. Keally, G. Zhou, and G. Xing, "Exploiting Sensing Diversity for Confident Coverage in Wireless Sensor Networks," in *IEEE INFOCOM*, 2011.

[3] M. Keally, G. Zhou, and G. Xing, "Watchdog: Confident Event Detection in Heterogeneous Sensor Networks," in *IEEE RTAS*, 2010.

[4] M. Duarte and Y. Hu, "Vehicle Classification in Distributed Sensor Networks," *JPDC*, 2004.