**OptiMoS: Optimal Sensing for Mobile Sensors**

Zhixian Yan\(^1\), Julien Eberle\(^2\), Karl Aberer\(^1\)

\(^1\) LSIR-IC-EPFL \quad \(^2\) Nokia Research Center - Lausanne

---

**OpenSense: An open environmental sensing platform with massive, heterogeneous, mobile and increasingly miniaturized sensors.**

- Wireless fixed nodes
- Mobile nodes
- Wireless communication
- Internet
- Base stations
- Servers, etc.

**Sensors:** CO, CO\(_2\), NO, NO\(_2\), Ozone, fine particles, ... \(\text{and increasingly miniaturized sensors.}

---

**Model-Driven Segmentation of Mobile Sensing**

**Problem Statement of Segmentation**

- **Initial Sensor Readings**
  - Sensor reading records \(R = \{R_1, \cdots, R_N\}\)
  - Each record with time, location, measurements

- **Data-driven Modeling**
  - Various models: linear, SVM regression, etc.
  - Model errors RSS – Residual Sum of Square

- **Optimal Segmentation**
  \[
  \arg\min_{K, \delta_1, \delta_2, \cdots, \delta_{K-1}} \sum_{l=1}^{K} \text{RSS}(R_{\delta_{l-1}, \cdots, \delta_l})
  \]

---

**Optimal Segmentation**

- Dynamic programming, expensive \(O(K^2)\), over-fitting
- **Top-down Binary Segmentation**
  - Binary: \(O(K\log N)\)
  - Binary*: better strategy in finding division segment

**Error-based Heuristic Segmentation**

- Heuristic: division by absolute errors
- Heuristic+: division by relative errors

**Near-Optimal Segmentation**

- B'**: Binary with Heuristic*, \(O(KN\log N)\)

---

**Near-Optimal Sampling for Individual Segments**

**Sampling Strategies**

- **Optimal Sampling**
  \[
  \arg\min_{R_{\delta_{l-1}, \cdots, \delta_l}} L(R, R_{\delta_{l-1}, \cdots, \delta_l}) \text{ s.t. } \|R_{\delta_{l-1}, \cdots, \delta_l}\|/|R| \leq \delta
  \]

- **NP-hard problem**
  \[
  \arg\min_{R_{\delta_{l-1}, \cdots, \delta_l}} |R_{\delta_{l-1}, \cdots, \delta_l}|/|R| \text{ s.t. } L(R, R_{\delta_{l-1}, \cdots, \delta_l}) \geq \epsilon
  \]

- **Distribution-based sampling**
  - Uniform: regular duty cycle readings
  - Random: irregular duty cycle readings

- **Entropy (Error) based sampling**
  - Selecting points with top entropy

- **Mutual Information based sampling**
  - Remove information redundancy
  - Recalculate entropy after each selection/sampling

---

**Integration with GSN for real-time data prediction**

**Visualization:** create pollution map by model perdition from limited measurements.