Knowing the air you breath - the Signal Processing way

Vijay A.

IIT Bombay

Acknowledgments : Prof. Animesh Kumar, Prof. Vivek Borkar

Knowing the Air ...

Economic Times, 2017

Delhi pollution: Air quality dips sharply ahead of Diwali

(Business Today 06 Nov)

Not just Delhi, 70 Indian cities reel under air pollution

(Down To Earth 05 Nov)

Clean air is a human right: WHO

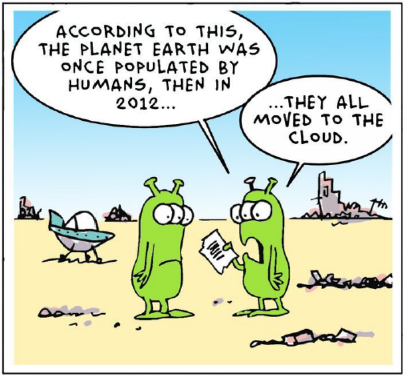
(Down To Earth 02 Nov)

Investments in Engineering

- Centre of Excellence for Research on Clean Air (CERCA) set up in IIT Delhi
- Number of ground based air quality (AQ) stations increased by Pollution Control Boards
- Sprinklers and artificial rains sought as options to help settle particulate matter (PM)
- **Deficits** in present AQ measurement/sensing
  - Ground station equipments require regular repair
  - Stations mostly concentrated in industrial belts
  - Average error in PM monitoring between 10-26 %

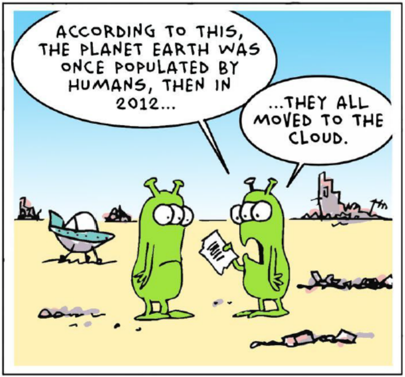
Signal Analytics for Rescue

- With emergence of IoT devices, shift has been from 'ground to the cloud'



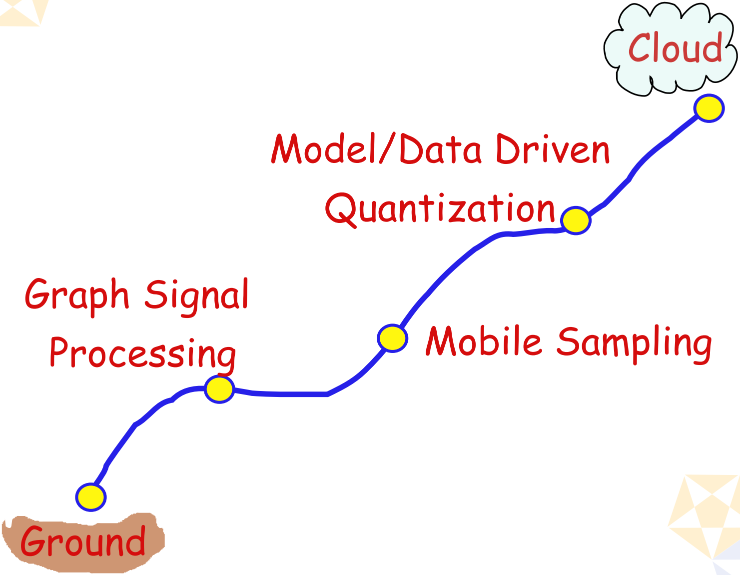
Signal Analytics for Rescue

- With emergence of IoT devices, shift has been from 'ground to the cloud'

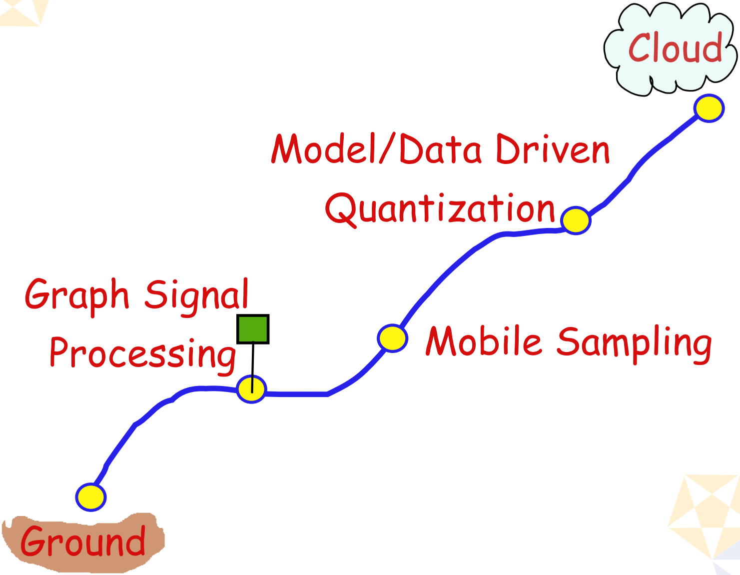


Data has become all pervasive due to smart devices

Roadmap to the Cloud

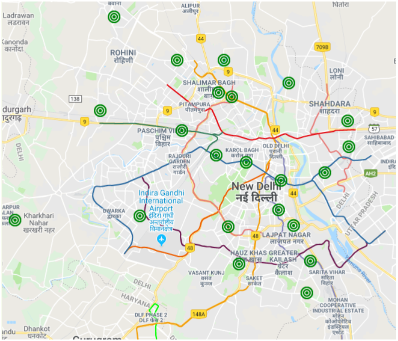


Roadmap to the Cloud



Graph SP : Basics

- Each station is a node on graph

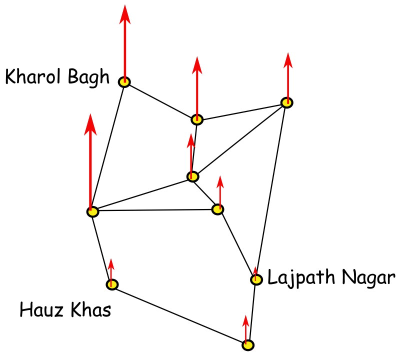


- $\mathcal{G}(V, E, W)$; edges represents correlation (distance)
- At any time t, graph signal is a vector $\vec{f}$
- $\vec{f}$ represents and snapshot of (say) PM 2.5 measurement
- An example of the edge weights is;

$$W_{i,j} = \begin{cases} \exp\left(\frac{\text{dist}(i,j)^2}{2\theta^2}\right) & \text{if dist}(i,j) \leq \kappa \\ 0 & \text{otherwise.} \end{cases}$$

Graph SP : Basics

- Each station is a node on graph
- $\mathcal{G}(V, E, W)$; edges represents correlation (distance)
- At any time t, graph signal is a vector $\vec{f}$



- $\vec{f}$ represents and snapshot of (say) PM 2.5 measurement
- An example of the edge weights is;

$$W_{i,j} = \begin{cases} \exp\left(\frac{\text{dist}(i,j)^2}{2\theta^2}\right) & \text{if dist}(i,j) \leq \kappa \\ 0 & \text{otherwise.} \end{cases}$$

Spectral is better?

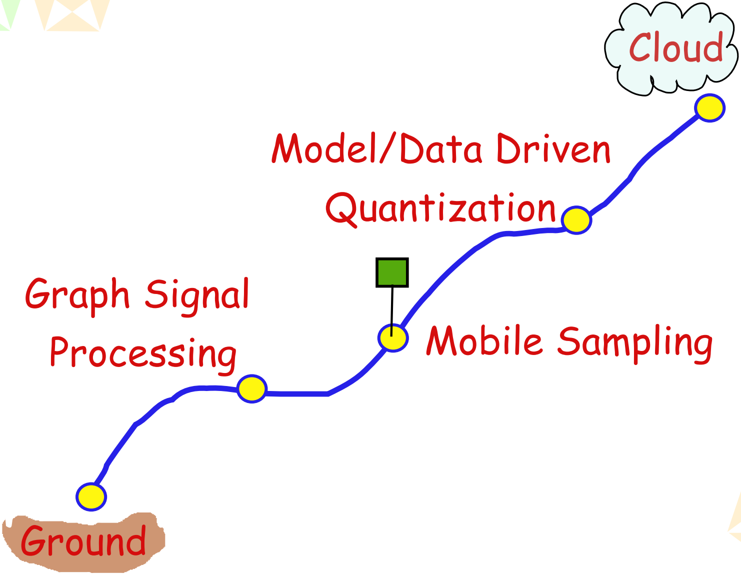
- Like in classical signal processing, we can define a Graph Fourier Transform [Shuman-Narang-Frossard-Ortega-Vandergheynst'13]
- (The Graph Laplacian) $\mathcal{L} := D - W$
  - Laplacian is positive semi-definite
  - Eignenvalues are bounded $0 := \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} := \lambda_{\mathsf{max}}$
  - Corresponding eigenvectors are frequency vectors (increasing zero crossings); $\vec{u}_0, \vec{u}_1, \cdots, \vec{u}_{N-1}$
- GFT is defined as;

$$\widehat{f}(\lambda_l) := \left\langle \vec{f}, \vec{u}_l \right\rangle$$

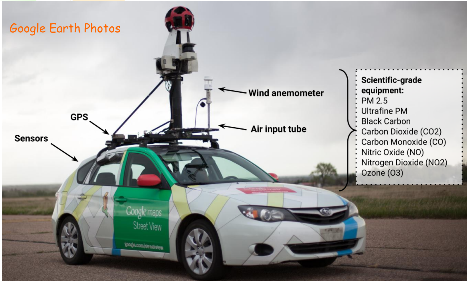
Recent trends...

- Graph Filtering: helps in denoising, ensuring sparsity, localization (zooming)
- Graph Wavelet Transforms : multiresolution analysis of graphs, sparse representations [Narang-Ortega'12]
- Spatio-Temporal Processing : deals with time variation of the graph signal. Joint space-time transforms are studied
- Graph Learning : learning graph weights from (training) data [Dong-Thanou-Frossard-Vandergheynst'16]

Roadmap to the cloud



Sampling Spatial Fields



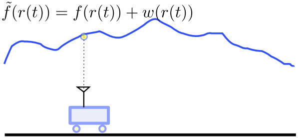
http://www.pigeonairpatrol.com

- Sensors are mounted on vehicles or birds to collect large amounts of data
- Mobile Sampling comes with measurement noise/ delay, but a large number of samples would come to the rescue !

A familiar Bandlimited model

- To start, model assumes time-invariance and limited bandwidth [Unnikrishnan-Vetterli'13]

- Let sampling vehicle have uniform velocity; $r(t) = u + vt$



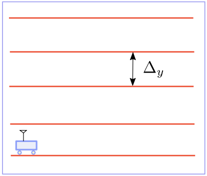
- 1-D Field: BW $\in [-\rho, \rho]$; time domain signal will see a bandwidth of $[-v\rho, v\rho]$

- (Nyquist-Shannon Sampling):

$$\widehat{f}(x) = \sum_{j \in \mathbb{Z}} s(jT) \frac{vT\rho}{\pi} \text{sinc}\left(\frac{\rho(x - jvT)}{\pi}\right)$$

More into practice

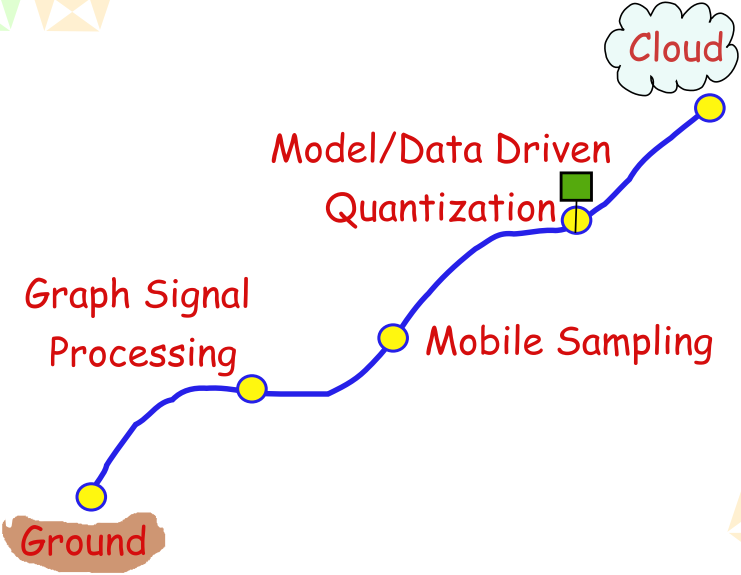
- 2-D Field: BW: $[-\rho, \rho] \times [-\rho, \rho]$.
- 2-D mobile sampling results in lower mean-square error than static sampling approach



Other Extensions:

- Sampling high dimensions using lower dimension manifolds [Unnikrishnan-Vetterli'13]
- Location unaware sampling [A. Kumar'17]
- Sampling of spatio-temporal fields [ongoing]

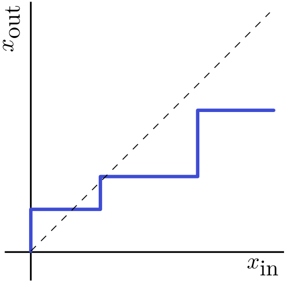
Roadmap to the cloud

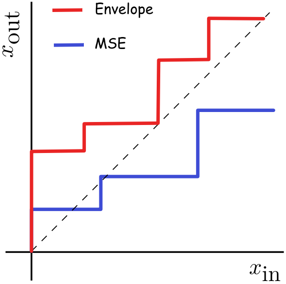


Envelope to the cloud

- Traditional mean square error quantizer (a.k.a Lloyd-Max quantizer) will not provide reliable approximation

- A comparative picture:





- Envelope quantization applicable in protection region database (TV whitespace), weather monitoring

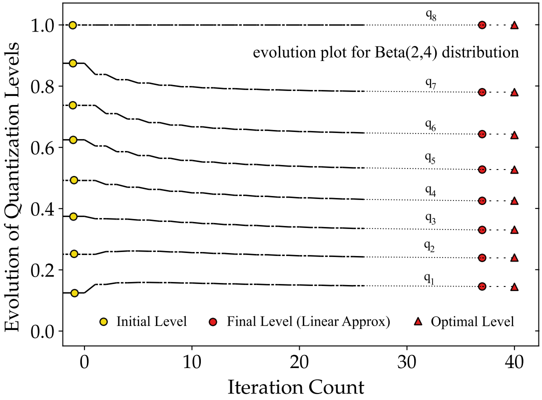
Drive using Model

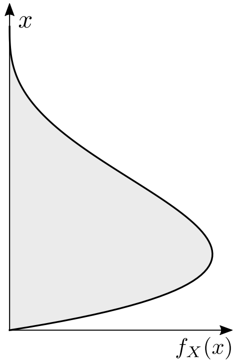
- Model driven approaches assumes the probability density is known
- Cost function (mean square error) optimization

$$\min_{Q(.)} \quad \text{MSE}(Q(X), X) \quad \text{subject to } Q(X) \geq X$$

- 1-parameter/ scalar quantization [Anavangot-Kumar'18]
  - Using localized MSE optimization a recursive algorithm is developed
  - In low computational settings, using piecewise-linear approx can be applied
  - Algorithm is shown to have exponential (error) decay rate and global optimality

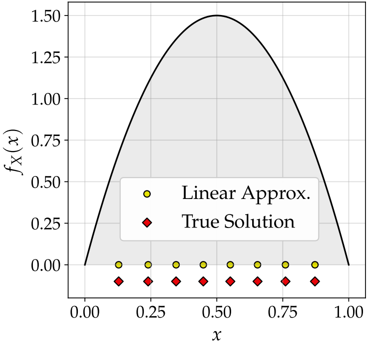
Alternating Recurrence



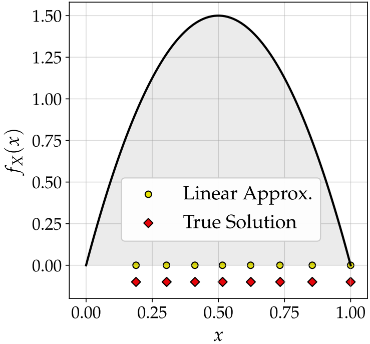


Algorithm details on the board

In simulation



ALM for symmetric Beta(2,2)



AEQ for symmetric Beta(2,2)

- Approximate Lloyd-Max (ALM) uses MSE cost (unconstrained)
- Approximate Envelope Quantizer (AEQ) uses envelope constrained MSE

Driven by data

- For traditional MSE; Learning Vector Quantization (LVQ) and K-means are popular
- Data driven envelope quantization approaches are not straight extensions
- High-dimensional quantization can be challenging

Summary... On cloud 9

- We discussed three state-of-the-art signal processing methods for analyzing sensor data
- Statistical learning tools are expected to dominate in each case
- Modern sensor technology and signal acquisition need to be reliable and power efficient

Summary... On cloud 9

- We discussed three state-of-the-art signal processing methods for analyzing sensor data
- Statistical learning tools are expected to dominate in each case
- Modern sensor technology and signal acquisition need to be reliable and power efficient

Thank you!