Big Data Optimization for Modern Communication Networks

Lanchao Liu
Advisor: Dr. Zhu Han
Co-Advisor: Dr. Wei-Chuan Shih

Wireless Networking, Signal Processing and Security Lab
Electrical and Computer Engineering Department
University of Houston

November 30, 2014

Table of Contents

Outline

The era of 'big data'

The "big data" presents us opportunities:

- Personal advertisement.
- Intelligent social network analysis.
- Smart city development.
- Medical data management.
- Smart grid evolution.
- Geophysics.

Figure: Big data '3V'.

- Big data requires big models and novel methods.
  - Thousands of parameters on TBs of data.
- Big data needs systems built for it.
  - Hadoop, Apache Spark, Storm, Yahoo! S4, Parameter sever.

Motivations and Contributions

- New computational mathematical models and methodologies must be explored.
  - Respect the inherent structure of the data. (Sparse, low rank, prior...).
  - Enjoys robustness and scalability.
- Review the parallel and distributed optimization algorithms based on ADMM.
- Investigated the 'big data' optimization methods for modern communication networks.
  - Smart grid security
  - Mobile data traffic management

Outline

Dual Ascent Methods

Consider an optimization problem of the form

$$\min_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{c},$$

- The Lagrangian: $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^\top(\mathbf{A}\mathbf{x} - \mathbf{c})$,
- Dual function: $g(\boldsymbol{\lambda}) = \inf_x \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda})$,
- Dual Problem: $\max_{\boldsymbol{\lambda}} g(\boldsymbol{\lambda})$,
- Optimal solution: $\mathbf{x}^* = \arg\min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*)$,
- Dual Ascent: $\begin{cases} \mathbf{x}^{k+1} = \arg\min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^k), \\ \boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \rho^k(\mathbf{A}\mathbf{x}^{k+1} - \mathbf{c}). \end{cases}$

Require an appropriate step size ρ and assumptions of strong convexity of the objective function f.

Method of Multipliers

Introduce an augmentation $\|\mathbf{A}\mathbf{x} - \mathbf{c}\|_2^2$ to the Lagrangian:

- The Augmented Lagrangian:
  $\mathcal{L}_\rho(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^\top(\mathbf{A}\mathbf{x} - \mathbf{c}) + \frac{\rho}{2}\|\mathbf{A}\mathbf{x} - \mathbf{c}\|_2^2,$
- Method of Multipliers: $\begin{cases} \mathbf{x}^{k+1} = \arg\min_{\mathbf{x}} \mathcal{L}_\rho(\mathbf{x}, \boldsymbol{\lambda}^k), \\ \boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \rho(\mathbf{A}\mathbf{x}^{k+1} - \mathbf{c}), \end{cases}$

Pros:

- Stable, robust and fast compare with the dual ascent method.
- No need to tune the parameter ρ during each iteration.

Cons:

- Difficult to decouple and parallelize due to the augmentation $\|\mathbf{A}\mathbf{x} - \mathbf{c}\|_2^2$.

Alternating Direction Method of Multipliers(ADMM)

The general form of ADMM is expressed as

$$\min_{\mathbf{x}_1 \in \mathcal{X}_1, \mathbf{x}_2 \in \mathcal{X}_2} f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2) \quad \text{s.t.} \quad \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 = \mathbf{c}. \tag{1}$$

The augmented Lagrangian for (1) is

$$\begin{aligned} \mathcal{L}_\rho(\mathbf{x}_1, \mathbf{x}_2, \boldsymbol{\lambda}) = f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2) + \boldsymbol{\lambda}^\top(\mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 - \mathbf{c}) \\ + \frac{\rho}{2}\|\mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 - \mathbf{c}\|_2^2, \end{aligned}$$

A Gauss-Seidel iterations of $\mathbf{x}_1$ and $\mathbf{x}_2$ as follows

$$\begin{cases} \mathbf{x}_1^{k+1} = \arg\min_{\mathbf{x}_1} \mathcal{L}_\rho(\mathbf{x}_1, \mathbf{x}_2^k, \boldsymbol{\lambda}^k), \\ \mathbf{x}_2^{k+1} = \arg\min_{\mathbf{x}_2} \mathcal{L}_\rho(\mathbf{x}_1^{k+1}, \mathbf{x}_2, \boldsymbol{\lambda}^k), \\ \boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \rho(\mathbf{A}_1\mathbf{x}_1^{k+1} + \mathbf{A}_2\mathbf{x}_2^{k+1} - \mathbf{c}). \end{cases}$$

Global convergence for convex optimization with a convergence rate $O(1/k)$

Outline

Direct Extension

Consider the following convex optimization problem

$$\begin{aligned} \min_{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N} \quad & f(\mathbf{x}) = f_i(\mathbf{x}_i) + \ldots + f_N(\mathbf{x}_N), \\ \text{s.t.} \quad & \mathbf{A}_i\mathbf{x}_i + \ldots + \mathbf{A}_N\mathbf{x}_N = \mathbf{c}, \\ & \mathbf{x}_i \in \mathcal{X}_i, \quad i = 1, \ldots, N. \end{aligned} \tag{2}$$

- The augmented Lagrangian:

$$\mathcal{L}_\rho(\{\mathbf{x}_i\}_{i=1}^N, \boldsymbol{\lambda}) = \sum_{i=1}^N f_i(\mathbf{x}_i) + \boldsymbol{\lambda}^\top(\sum_{i=1}^N \mathbf{A}_i\mathbf{x}_i - \mathbf{c}) + \frac{\rho}{2}\|\sum_{i=1}^N \mathbf{A}_i\mathbf{x}_i - \mathbf{c}\|_2^2$$

- Gauss-Seidel Multi-block ADMM:

$$\begin{cases} \mathbf{x}_i = \arg\min_{\mathbf{x}_i} \mathcal{L}_\rho(\{\mathbf{x}_j^{k+1}\}_{j<i}, \mathbf{x}_i, \{\mathbf{x}_j^k\}_{j>i}, \boldsymbol{\lambda}^k), & i = 1, \ldots, N. \\ \boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \rho(\sum_{i=1}^N \mathbf{A}_i\mathbf{x}_i^{k+1} - \mathbf{c}). \end{cases}$$

Gauss-Seidel mulit-block ADMM is not necessarily convergent

Variable Splitting Multi-block ADMM

Reformulate (2) by introducing auxiliary variable $\mathbf{z}$

$$\min_{\mathbf{x},\mathbf{z}} \quad \sum_{i=1}^{N} f_i(\mathbf{x}_i) + I_{\mathcal{Z}}(\mathbf{z}) \quad \text{s.t.} \quad \mathbf{A}_i\mathbf{x}_i + \mathbf{z}_i = \frac{\mathbf{c}}{N}, \quad i = 1, \ldots, N,$$

where $\mathcal{Z} = \{\mathbf{z} | \sum_{i=1}^{N} \mathbf{z}_i = 0\}$, $I_{\mathcal{Z}}(\mathbf{z})$ is the indicator function.

- The augmented Lagrangian

$$\mathcal{L}_\rho = \sum_{i=1}^{N} f_i(\mathbf{x}_i) + I_{\mathcal{Z}}(\mathbf{z}) + \sum_{i=1}^{N} \boldsymbol{\lambda}_i^\top (\mathbf{A}_i\mathbf{x}_i + \mathbf{z}_i - \frac{\mathbf{c}}{N}) + \frac{\rho}{2} \sum_{i=1}^{N} \|\mathbf{A}_i\mathbf{x}_i + \mathbf{z}_i - \frac{\mathbf{c}}{N}\|_2^2.$$

- Variable splitting multi-block ADMM

$$\begin{cases} \mathbf{x}_i^{k+1} = \arg\min_{\mathbf{x}_i} \mathcal{L}_\rho(\mathbf{x}_i, \mathbf{z}_i^k, \boldsymbol{\lambda}_i^k), \\ \mathbf{z}_i^{k+1} = \arg\min_{\mathbf{z}_i} \mathcal{L}_\rho(\mathbf{x}_i^{k+1}, \mathbf{z}_i, \boldsymbol{\lambda}_i^k), \quad \forall i = 1, \ldots, N, \\ \boldsymbol{\lambda}_i^{k+1} = \boldsymbol{\lambda}_i^k + \rho(\mathbf{A}_i\mathbf{x}_i + \mathbf{z}_i - \frac{\mathbf{c}}{N}). \end{cases}$$

Converge as two-block setting, but the number of variables and constraints will increase substantially when N is large.

Proximal Jacobian ADMM

Recall the augmented Lagrangian:

$$\mathcal{L}_\rho(\{\mathbf{x}_i\}_{i=1}^N, \boldsymbol{\lambda}) = \sum_{i=1}^N f_i(\mathbf{x}_i) + \boldsymbol{\lambda}^\top(\sum_{i=1}^N \mathbf{A}_i\mathbf{x}_i - \mathbf{c}) + \frac{\rho}{2}\|\sum_{i=1}^N \mathbf{A}_i\mathbf{x}_i - \mathbf{c}\|_2^2$$

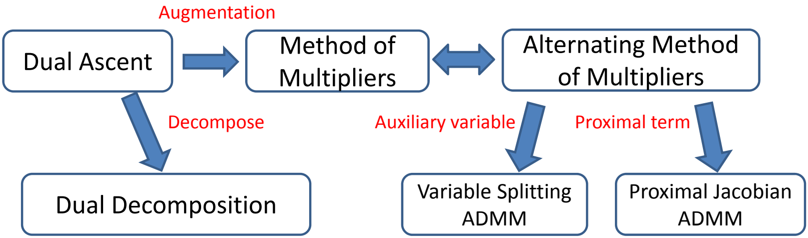
A proximal term is added to the augmented Lagrangian, and the update of $\mathbf{x}_i$ is performed concurrently:

$$\begin{cases} \mathbf{x}_i^{k+1} = \arg\min_{\mathbf{x}_i} \mathcal{L}_\rho(\mathbf{x}_i, \{\mathbf{x}_j^k\}_{j\neq i}, \boldsymbol{\lambda}^k) + \frac{1}{2}\|\mathbf{x}_i - \mathbf{x}_i^k\|_{\mathbf{P}_i}^2, \\ \boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \gamma\rho(\sum_{i=1}^N \mathbf{A}_i\mathbf{x}_i^{k+1} - \mathbf{c}), \end{cases} \quad \forall i = 1, \ldots, N.$$

where $\|\mathbf{x}_i\|_{\mathbf{P}_i}^2 = \mathbf{x}_i^\top \mathbf{P}_i\mathbf{x}_i$ for some symmetric and positive semi-definite matrix $\mathbf{P}_i \succeq 0$.
The involvement of the proximal term

- Make subproblem of $\mathbf{x}_i$ strictly or strongly convex
- Ensure the convergence. Easier to solve.

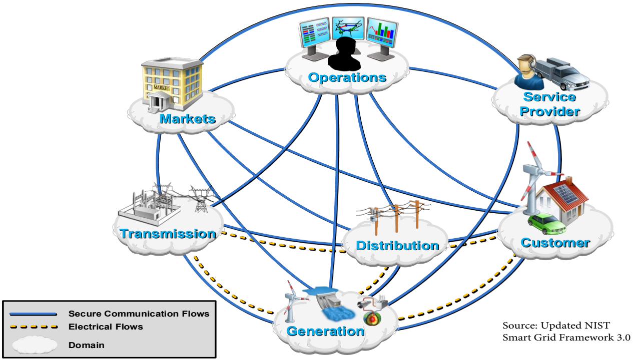
Recap: From Dual Ascent to Multi-block ADMM



1. Stability and robustness are the utmost concern for an optimization algorithm.
2. Should better be distributed and parallel.
3. If possible, we want it converge fast.

Outline

Smart Grid 'Big Data'



The anticipated smart grid data deluge:

1. The deployment of phasor measurement units for future North American power grid will generate 4.15 TB data per day.
2. 61.8 million smart meters are deployed in the U.S. by the end of 2013. Every one million users will produce 27.3TB per year.

Blackout

- Increasing integration between cyber operations and physical infrastructures for generation, transmission, and distribution control.
- The security and reliability are not guaranteed

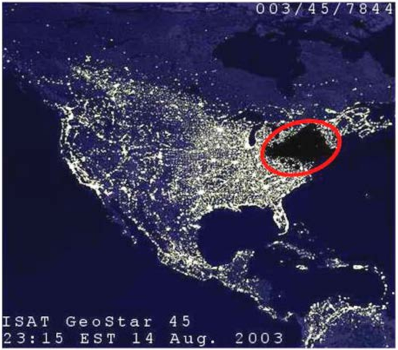
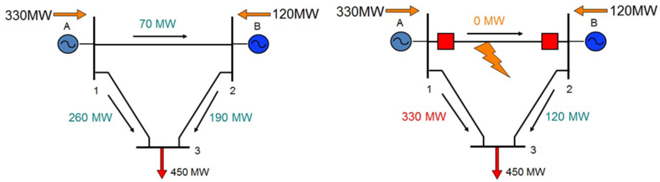


Figure: 2003 blackout

Contingency Analysis

Simple example



1. Assume line 1-2 is disconnected.
2. Generators A and B cannot change productions quickly.
3. The flows over other lines would increase.
4. Trigger cascading failure.

Security Constrained Optimal Power Flow (SCOPF)

SCOPF: Minimizing the cost of system operation while satisfying a set of postulated contingency constraints.

$$
\begin{aligned}
\min_{\{\mathbf{x}^c\}_0^C;\{\mathbf{u}^c\}_0^C} \quad & f^0(\mathbf{u}^0) \quad \text{scheduling objective} \\
\text{s.t.} \quad & \mathbf{g}^0(\mathbf{x}^0, \mathbf{u}^0) = 0, \text{power flow equations} \\
& \mathbf{h}^0(\mathbf{x}^0, \mathbf{u}^0) \leq 0, \text{operating limits for base case} \\
& \mathbf{g}^c(\mathbf{x}^c, \mathbf{u}^c) = 0, \text{power flow equations} \\
& \mathbf{h}^c(\mathbf{x}^c, \mathbf{u}^c) \leq 0, \text{operating limits for contingency k} \\
& \|\mathbf{u}^0 - \mathbf{u}^c\|^2 \leq \boldsymbol{\Delta}_c, c = 1, \ldots, C, \text{security constrains}
\end{aligned}
$$

Challenges:

1. Number of constraints is prohibitive.
2. How to find the best operating point with a scalable algorithm?

DC Approximation

Power flow equations:

$$P_i = \sum_{k=1}^{N} |V_i||V_k|(G_{ik}\cos\theta_{ik} + B_{ik}\sin\theta_{ik}),$$

$$Q_i = \sum_{k=1}^{N} |V_i||V_k|(G_{ik}\sin\theta_{ik} - B_{ik}\cos\theta_{ik}),$$

where G_{ik} and B_{ik} are the real and imaginary part of $(i, k)^{th}$ element of the bus admittance matrix.
DC power flow approximation:

1. Neglect the reactive power.
2. Neglect resistance of the branches.
3. Assume all voltage magnitudes = 1.0 p.u.
4. Assume all angles are small.

DC SCOPF

'N-1' contingency, corrective setting:

$$
\begin{aligned}
\min_{\{\theta^c\}_{c=0}^{C};\{\mathbf{P}^{g,c}\}_{c=0}^{C}} \quad & \sum_{i\in\mathcal{G}} f_i^g(\mathbf{P}_i^{g,0}) \\
\text{subject to} \quad & \mathbf{B}_{bus}^0\boldsymbol{\theta}^0 + \mathbf{P}^{d,0} - \mathbf{A}^{g,0}\mathbf{P}^{g,0} = 0, \\
& \mathbf{B}_{bus}^c\boldsymbol{\theta}^c + \mathbf{P}^{d,c} - \mathbf{A}^{g,c}\mathbf{P}^{g,c} = 0, \\
& |\mathbf{B}_f^0\boldsymbol{\theta}^0| - \mathbf{F}_{max} \leq 0, \\
& |\mathbf{B}_f^c\boldsymbol{\theta}^c| - \mathbf{F}_{max} \leq 0, \\
& \underline{\mathbf{P}^{g,0}} \leq \mathbf{P}^{g,0} \leq \overline{\mathbf{P}^{g,0}}, \\
& \underline{\mathbf{P}^{g,c}} \leq \mathbf{P}^{g,c} \leq \overline{\mathbf{P}^{g,c}}, \\
& |\mathbf{P}^{g,0} - \mathbf{P}^{g,c}| \leq \boldsymbol{\Delta}_c, \\
& i \in \mathcal{G}, \quad c = 1, \ldots, C,
\end{aligned}
$$

where $\mathbf{B}_{bus}$ and $\mathbf{B}_f$ can be modified from the bus admittance matrix $\mathbf{Y}_{bus}$. $\mathbf{A}^{g,c}$ is the generator connection matrix.

A Distributed Approach by ADMM

Introduce a slack variable $\mathbf{p}^c$ to rewrite $|\mathbf{P}^{g,0} - \mathbf{P}^{g,c}| \leq \mathbf{\Delta}_c$ as:

$$\begin{aligned} &\mathbf{P}^{g,0} - \mathbf{P}^{g,c} + \mathbf{p}^c = \mathbf{\Delta}_c \\ &0 \leq \mathbf{p}^c \leq 2\mathbf{\Delta}_c, \quad c = 1, \ldots, C. \end{aligned} \tag{3}$$

The partial scaled augmented Lagrangian associated with (3) can be calculated with follows

$$\mathcal{L}_\rho(\{\mathbf{P}^{g,c}\}_{c=0}^C; \{\mathbf{p}^c\}_{c=1}^C; \{\boldsymbol{\mu}^c\}_{c=1}^C)$$
$$= \sum_{i \in \mathcal{G}} f_i^g(\mathbf{P}_i^{g,0}) + \sum_{c=1}^C \frac{\rho^c}{2} \|\mathbf{P}^{g,0} - \mathbf{P}^{g,c} + \mathbf{p}^c - \mathbf{\Delta}_c + \boldsymbol{\mu}^c\|_2^2.$$

Iterate till convergence

1. Update $\{\mathbf{P}^{g,0}\}$.
2. Update $\{\mathbf{P}^{g,c}, \mathbf{p}^c\}$.
3. Update dual variable $\boldsymbol{\mu}^c$.

A Distributed Approach by ADMM (cont.)

The update for base case (Modified OPF problem):

$$\mathbf{P}^{g,0}[k+1] = \arg\min_{\mathbf{P}^{g,0}} \sum_{i \in \mathcal{G}} f_i^g(\mathbf{P}_i^{g,0})$$

$$+ \sum_{c=1}^{C} \frac{\rho^c}{2} \|\mathbf{P}^{g,0} - \mathbf{P}^{g,c}[k] + \mathbf{p}^c[k] - \boldsymbol{\Delta}_c + \boldsymbol{\mu}^c[k]\|_2^2,$$

The update for contingency case c

$$\mathbf{P}^{g,c}[k+1] = \arg\min_{\mathbf{P}^{g,c},\mathbf{p}^c} \frac{\rho^c}{2} \|\mathbf{P}^{g,0}[k+1] - \mathbf{P}^{g,c} + \mathbf{p}^c - \boldsymbol{\Delta}_c + \boldsymbol{\mu}^c[k]\|_2^2,$$

The scaled dual variable is updated by:

$$\boldsymbol{\mu}^c[k+1] = \boldsymbol{\mu}^c[k] + \mathbf{P}^{g,0}[k+1] - \mathbf{P}^{g,c}[k+1] + \mathbf{p}^c[k+1] - \boldsymbol{\Delta}_c.$$

Distributed Implementation

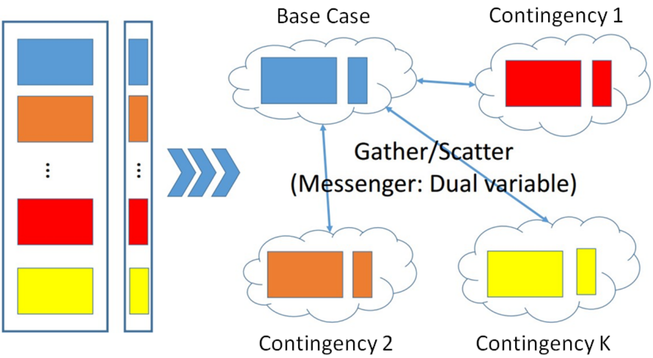


Figure: Distributed Implementation.

1. On multi-core machine.
2. High performance computer cluster using MPI (message passing interface).
3. On cloud using Hadoop or Apache Spark.

Numerical Results

Evaluation setup: Modified data of IEEE 57 bus, IEEE 118 bus and IEEE 300 bus generated by MATPOWER

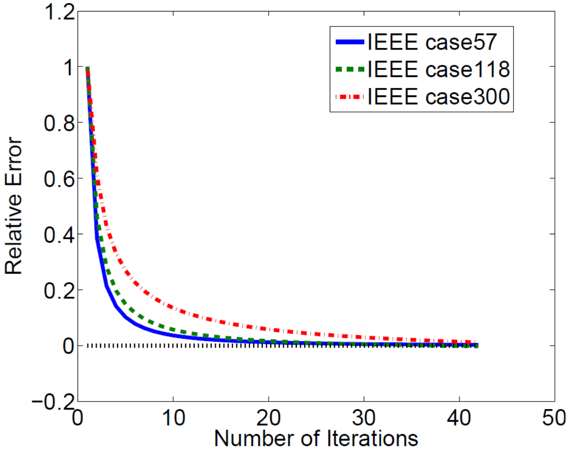


Figure: Convergence performance.

Numerical Results (cont.)

Using the Matlab distributed and parallel toolbox, Subproblems are solved by CVX.

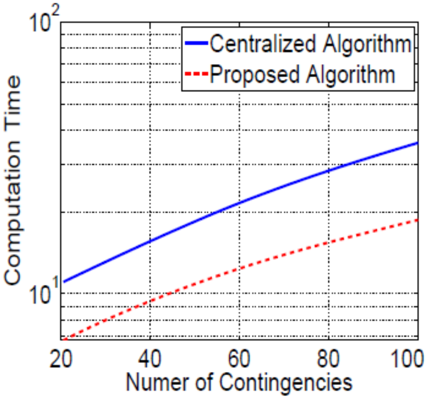


Figure: Computing time for IEEE 57 bus.

Outline

Global Mobile Data Traffic

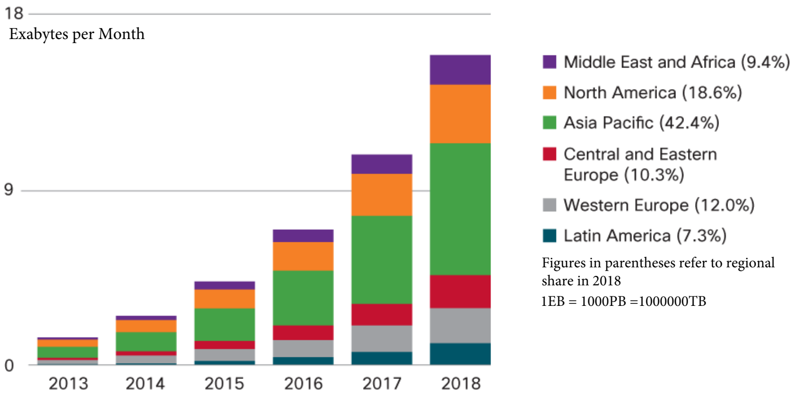


Figure: The global mobile data traffic forecast by region.

- Expected to reach 15.9 EB per month by 2018.
- A 11-fold increase over 2013.

Mind the Gap

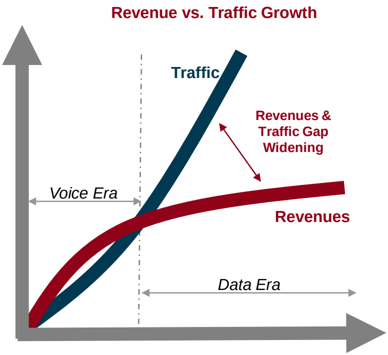


Figure: Revenue & traffic gap. Source:radisys.com

The sheer volume of the mobile ‘big data’ traffic far exceeds

- The growth in service revenues.
- The budgets required to address the new demands.

Mobile Data Offloading

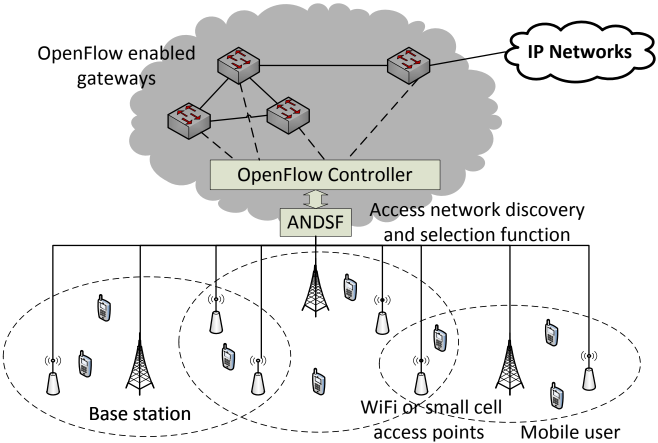


Figure: An illustration of the network model

- Mobile data offloading: Offload traffic from cellular networks to alternate wireless technologies.
- Software defined network (SDN) at the edge: Dynamically route the traffic in a mobile network.

Mobile Data Offloading in SDN

Consider a mobile network

- A BS $b \in \{1, \ldots, B\}$ serves a group of mobile users.
- An AP $a \in \{1, \ldots, A\}$ provides bandwidth for data offloading.
- $\mathbf{x}_b = [x_{b1}, \ldots, x_{bA}]^\top$ represent the offloaded traffic of BS b.
  - x_{ba} denotes the data of BS b offloaded through AP a.
- BS b's utility: $U_b(\mathbf{x}_b)$. Non-decreasing, non-negative and concave.
- $\mathbf{y}_a = [y_{a1}, \ldots, y_{aB}]^\top$ represents the admitted traffic of AP a
  - y_{ab} represents the admitted data traffic from BS b.
- AP a's cost: $L_a(\mathbf{y}_a)$. Non-decreasing, non-negative and convex.
- Feasible mobile data offloading decision: $x_{ba} = y_{ab}, \forall a$ and $\forall b$.

Mobile Data Offloading in SDN (Cont.)

- Utility of base stations: $\sum_{b=1}^{B} U_b(\mathbf{x}_b)$.
- Cost of access points: $\sum_{a=1}^{A} L_a(\mathbf{y}_a)$.
- Total revenue: $\sum_{b=1}^{B} U_b(\mathbf{x}_b) - \sum_{a=1}^{A} L_a(\mathbf{y}_a)$.
- Equivalent revenue maximization problem:

$$\begin{aligned} \min_{\{\mathbf{x}_1,\ldots,\mathbf{x}_B\},\{\mathbf{y}_1,\ldots,\mathbf{y}_A\}} \quad & \sum_{a=1}^{A} L_a(\mathbf{y}_a) - \sum_{b=1}^{B} U_b(\mathbf{x}_b), \quad \text{Service revenue} \\ \text{s.t} \quad & \sum_{b=1}^{B} y_{ab} \le C_a, \quad \forall a, \quad \text{Capacity constraint} \\ & x_{ba} = y_{ab}. \quad \forall a, b \quad \text{Consensus} \end{aligned}$$

Challenges

- Privacy preserving: Utility functions at BSs and cost functions at APs should be kept private.
  - How to address the information asymmetry?
- Concurrent update: The updating process at the BSs and APs should be performed concurrently.
  - How to get concurrency?
- Scalability: The operations at the SDN controller should be simple to alleviate the computation burden.
  - How to design a scalable scheme?

Proximal Jacobian ADMM

The Lagrangian function

$$\mathcal{L}_{\rho}(\mathbf{x}, \mathbf{y}, \boldsymbol{\lambda}) = \sum_{a=1}^{A} L_a(\mathbf{y}_a) - \sum_{b=1}^{B} U_b(\mathbf{x}_b) + \sum_{a=1}^{A}\sum_{b=1}^{B} \lambda_{ab}(x_{ba} - y_{ab}) + \frac{\rho}{2}\sum_{a=1}^{A}\sum_{b=1}^{B} \|x_{ba} - y_{ab}\|_2^2.$$

Base Station Update:

$$\mathbf{x}_b^{k+1} = \arg\min_{\mathbf{x}_b} -U_b(\mathbf{x}_b) + \frac{\rho}{2}\sum_{a=1}^{A} \|x_{ba} - y_{ab}^k + \frac{\lambda_{ab}^k}{\rho}\|_2^2 + \frac{1}{2}\|\mathbf{x}_b - \mathbf{x}_b^k\|_{\mathbf{P}_i}^2.$$

Access Point Update:

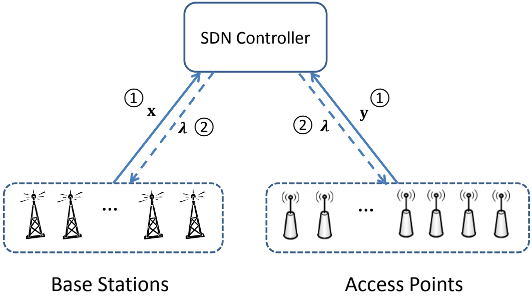
$$\mathbf{y}_a^{k+1} = \arg\min_{\mathbf{y}_b} (L_a(\mathbf{y}_a) + \frac{\rho}{2}\sum_{b=1}^{B} \|y_{ab} - x_{ba}^k - \frac{\lambda_{ab}^k}{\rho}\|_2^2 + \frac{1}{2}\|\mathbf{y}_a - \mathbf{y}_a^k\|_{\mathbf{P}_i}^2),$$

$$\text{s.t} \quad \sum_{b=1}^{B} y_{ab} \leq C_a.$$

SDN Controller Update:

$$\lambda_{ab}^{k+1} = \lambda_{ab}^k + \gamma\rho\sum_{b=1}^{B}\sum_{a=1}^{A}(x_{ba}^{k+1} - y_{ab}^{k+1}).$$

Distributed Update Scheme



① Gather: BSs and APs concurrently update $\mathbf{x}$ and $\mathbf{y}$, which are gathered by controller.

② Scatter: Controller simply updates $\boldsymbol{\lambda}$, which are scattered to BSs and APs

Figure: Distritbuted update scheme

- Iterative gather-scatter scheme (Map-reduce).
- Signaling: $p_{ab}^{k} = (y_{ab}^{k} - \frac{\lambda_{ab}^{k}}{\rho}),\ q_{ba}^{k} = (x_{ba}^{k} + \frac{\lambda_{ab}^{k}}{\rho})$

Numerical Results

Evaluation setup: $B = 5$ base stations and $A = \{5, 10\}$ access points. $C_a = 10Mbps$

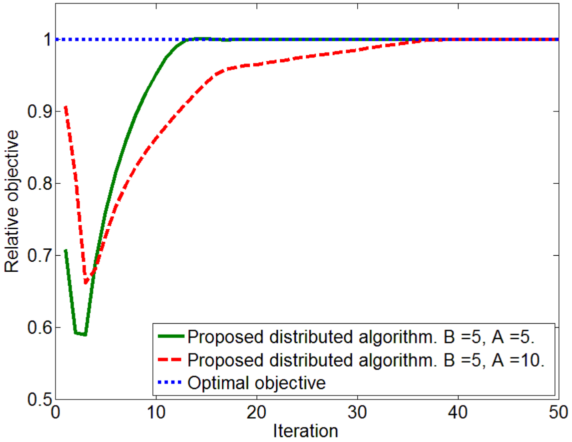


Figure: Convergence performance

Numerical Results(Cont.)

Total market gap: $\sum_{a=1}^{A} \sum_{b=1}^{B} (x_{ba} - y_{ab})$

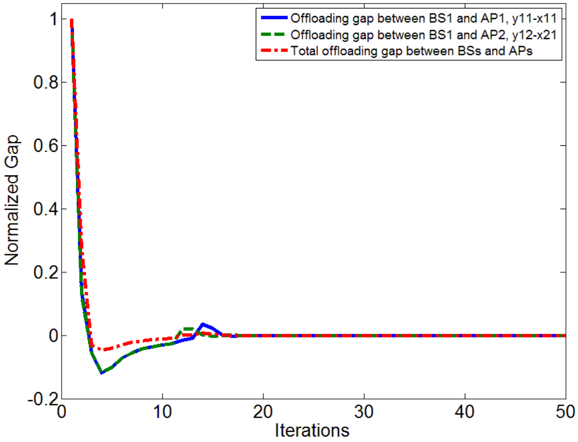


Figure: Offloading gap

Outline

Scalable service management in mobile cloud computing

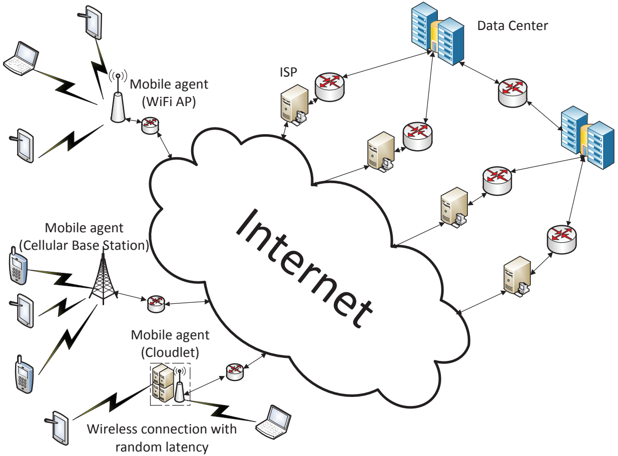


Figure: An illustration of mobile cloud computing infrastructure.

Sparse optimization for false data injection attacks detection

$\mathbf{Z_0}$ is the normal temporal state measurement, $\mathbf{A}$ is the sparse attack matrix.

$$\min_{\mathbf{Z_0},\mathbf{A}} \|\mathbf{Z_0}\|_* + \lambda\|\mathbf{A}\|_1, \quad s.t. \quad \mathbf{Z_a} = \mathbf{Z_0} + \mathbf{A}, \tag{4}$$

where

- $\|\cdot\|_*$: nuclear norm. Sum of singular values of a matrix.
- $\|\cdot\|_1$: l_1 *norm*. Sum of absolute values of matrix entries.

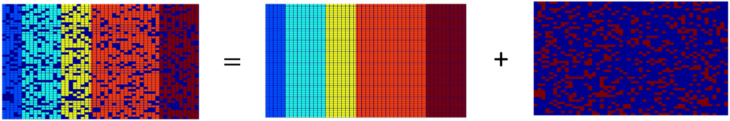


Figure: Temporal measurements illustration.

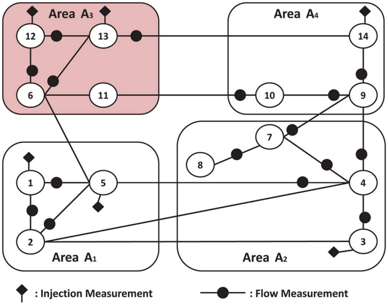
Outline

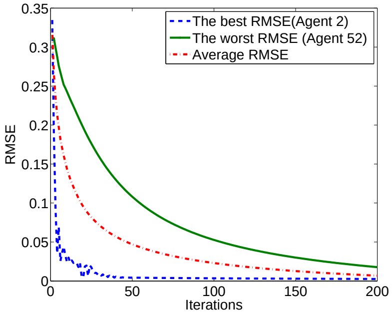
Distributed State Estimation

A fully distributed Gauss-Newton method for state estimation in power system.

$$\min_{\mathbf{x}_i \dots \mathbf{x}_N} \quad f(\mathbf{x}) = \sum_{i=1}^{N} (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i))^\top \mathbf{R}_i^{-1} (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i)),$$

$$\text{s.t.} \quad \mathbf{x}_i = \ldots = \mathbf{x}_N,$$





Smart City

Air quality monitoring using big data techniques.

Figure: Taxi route.

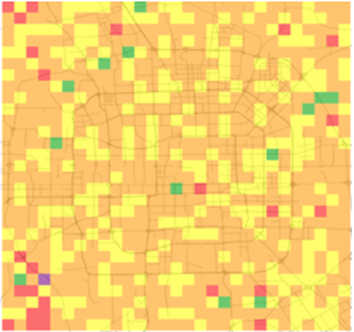


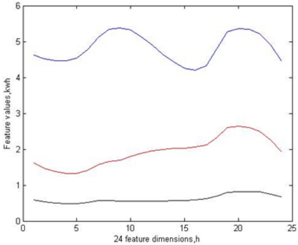
Figure: Air monitoring map

- Ultra fine particles and PM2.5 monitoring.
- Taxi-carried air monitoring sensor.
- How to infer the air quality index spatially and temporally?

Smart Meter Data Clustering

From smart meter data, try to tell users usage behaviors

1. Housewives?
2. Commute workers?
3. Ph.D students?



Outline

Conclusions

- A brief review of the alternating direction method of multipliers.
- Zoomed in two applications of big data optimization in modern communication networks.
  - Security constrained optimal power flow in smart grid.
  - Data offloading in software defined networks.
- Effective management and processing of 'big data' has the potential to significantly improve network security and management efficiency.

Awards and Research Summary

Adwards:

- Candidate for excellent overseas Chinese Ph.D. student.

Book Chapter:

- L. Liu, Z. Han, S. Cui, and V. H. Poor, "Big Data Processing for Smart Grid Security," Big Data over Networks, *Cambridge University Press*, edited by A. Hero, J. Moura, T. Luo, and S. Cui.

Journal:

- L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han, "Detecting False Data Injection Attacks on Power Gird by Sparse Optimization," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 612-621, Mar. 2014.
- M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting Stealthy False Data Injection Using Machine Learning in Smart Grid," to appear *IEEE Systems Journal*.
- L. Liu, D. Niyato, P. Wang, and Z. Han, "Scalable Mobile Cloud Services Management with Stochastic Wireless Link Latency," submitted to *IEEE Transaction on Mobile Computing*.
- L. Liu, A. Khodaei, W. Yin, and Z. Han, "Distribute Approach for Security Constrained Optimal Power Flow with Renewable Generation," submitted to *IEEE Transaction on Power System*.

Conference:

- L. Liu, A. Khodaei, W. Yin, and Z. Han, "Distribute Parallel Approach for Big Data Scale Optimal Power Flow with Security Constraints," *IEEE SmartGridComm*, Vancouver, Canada, October 2013.
- L. Liu, M. Esmalifalak, and Z. Han, "Detection of False Data Injection in Power Grid Exploiting Low Rank and Sparsity," *IEEE International Conference on Communication*, Budapest, Hungary, June 2013.
- L. Liu, H. Li, and Z. Han, "Sampling Spectrum Occupancy Data over Random Fields: A Matrix Completion Approach," *IEEE International Conference on Communications,* Ottawa, Canada, June 2012.
- L. Liu, Z. Han, Z. Wu, and L. Qian, "Collaborative Compressive Sensing based Dynamic Spectrum Sensing and Mobile Primary User Localization in Cognitive Radio Networks," *IEEE Globe Communication Conference*, Houston, December 2011.

Thanks!