UNIVERSITY of HOUSTON
YOU ARE THE PRIDE
Department of Electrical and Computer Engineering

Big Data Analysis and Cross-Layer Optimization for Communication, Caching and Computing (C^3) Networks

Zhu Han

ECE and CS Departments, University of Houston

Globecom 2017 Singapore

Thanks for NSF,

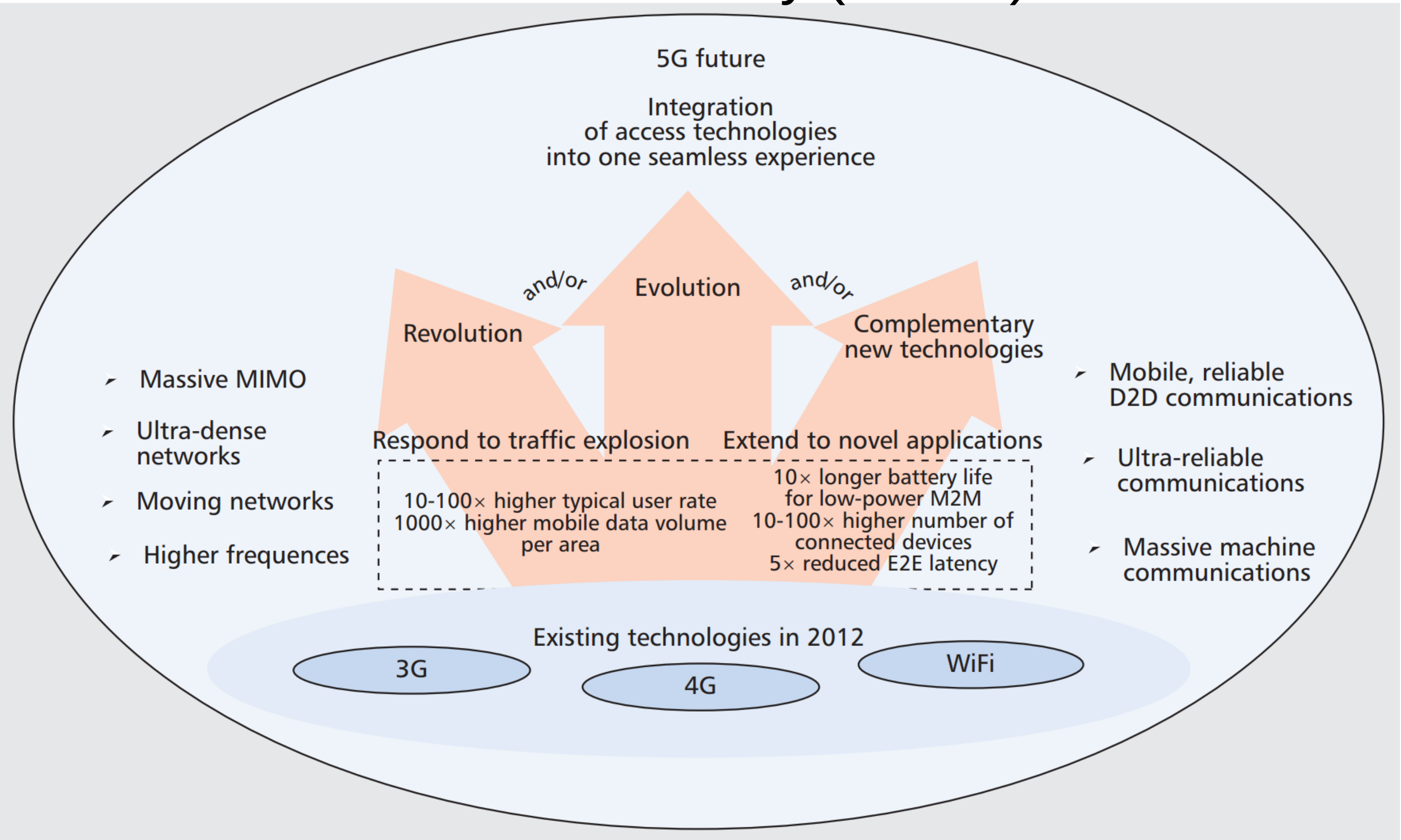
work by Ye Yu, Prof. Li Wang, Xunsheng Du and Kevin Tsai

Outline

- Introduction and Motivation for C^3 Networks
- Big Data Analysis and Cross Layer Optimization
  - Wireless Network Function Virtualization
  - Mobile Social Networks over D2D
  - Deep Learning Analysis
- Conclusions

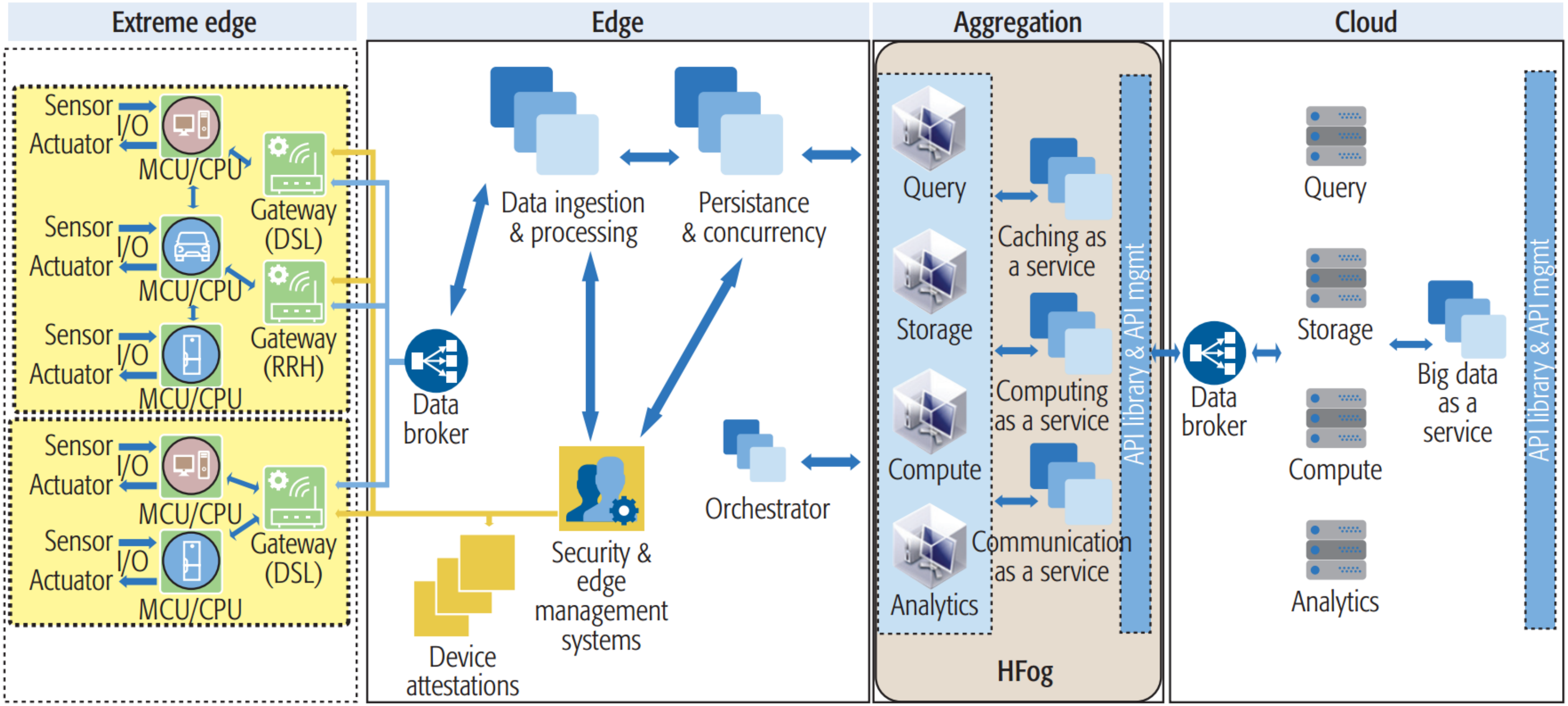
Future 5G Networks

1. Hyper-connected society
2. High data rates at the network edge (1–10 Gb/s)
3. Ultra low end-to-end latency (~1 ms).



* A. Osseiran et al., "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," in IEEE Communications Magazine, vol. 52, no. 5, pp. 26-35, May 2014.

Communication, Caching and Computing



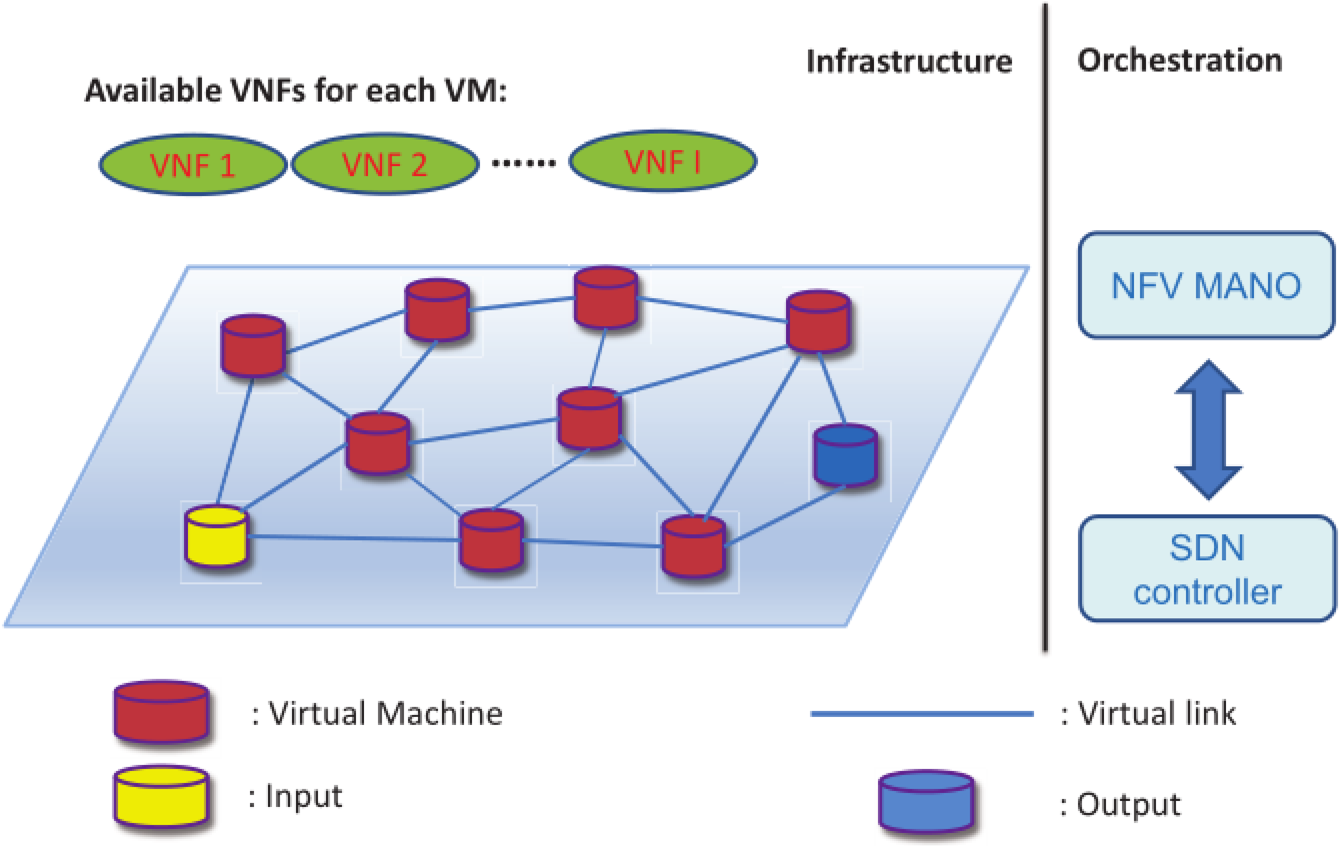
* A. Osseiran et al., "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," in IEEE Communications Magazine, vol. 52, no. 5, pp. 26-35, May 2014.

Outline

- ❑ Introduction and Motivation for C^3 Networks
- ❑ Big Data Analysis and Cross Layer Optimization
  - ❑ Wireless Network Function Virtualization
  - ❑ Mobile Social Networks over D2D
  - ❑ Deep Learning Analysis
- ❑ Conclusions

Network (Function) Virtualization

Virtualize entire classes of network node functions into building blocks that may connect, or chain together, to create communication services



Advantages

Reduce Expenditure; Accelerate Time-to-Market; Deliver Agility and Flexibility

Problem Formulation Example

$$\delta_{n,k}^{m} = \begin{cases} 1, if\ f_k^m\ is\ implemented\ on\ VM\ n \\ 0, otherwise. \end{cases}$$

number of the instances: $$e_k^m = \sum_{n=1}^{N} \delta_{n,k}^m.$$

VNF placement cost: $$Q_p = \sum_{k=1}^{K} e_k^m \xi_k^m,$$

Traffic cost: $$Q_t = \sum_{k=1}^{K} \gamma_k \sum_{n=1}^{N} v_{n,k}^m,$$

Total cost function: $$Q_{total} = Q_p + Q_t$$

Variables: $\delta_{n,k}^m$ (Integer)

$v_{n,k}^m$ (continuous)

$$\min_{\boldsymbol{\delta}^{m,(l)} \mathbf{v}^{m,(l)}} \quad Q_{total}$$

s.t $$1 \leq \sum_{n=1}^{N} \delta_{n,k}^m \leq N \quad \forall k, m,$$

VNF number

$$\sum_{k=1}^{K} \delta_{n,k}^m v_{n,k}^m \phi_k \leq C_n \quad \forall n, m,$$

VM's capacity

$$\sum_{n=1}^{N} v_{n,k}^m \geq R_k \quad \forall k,$$

Data rate requirement

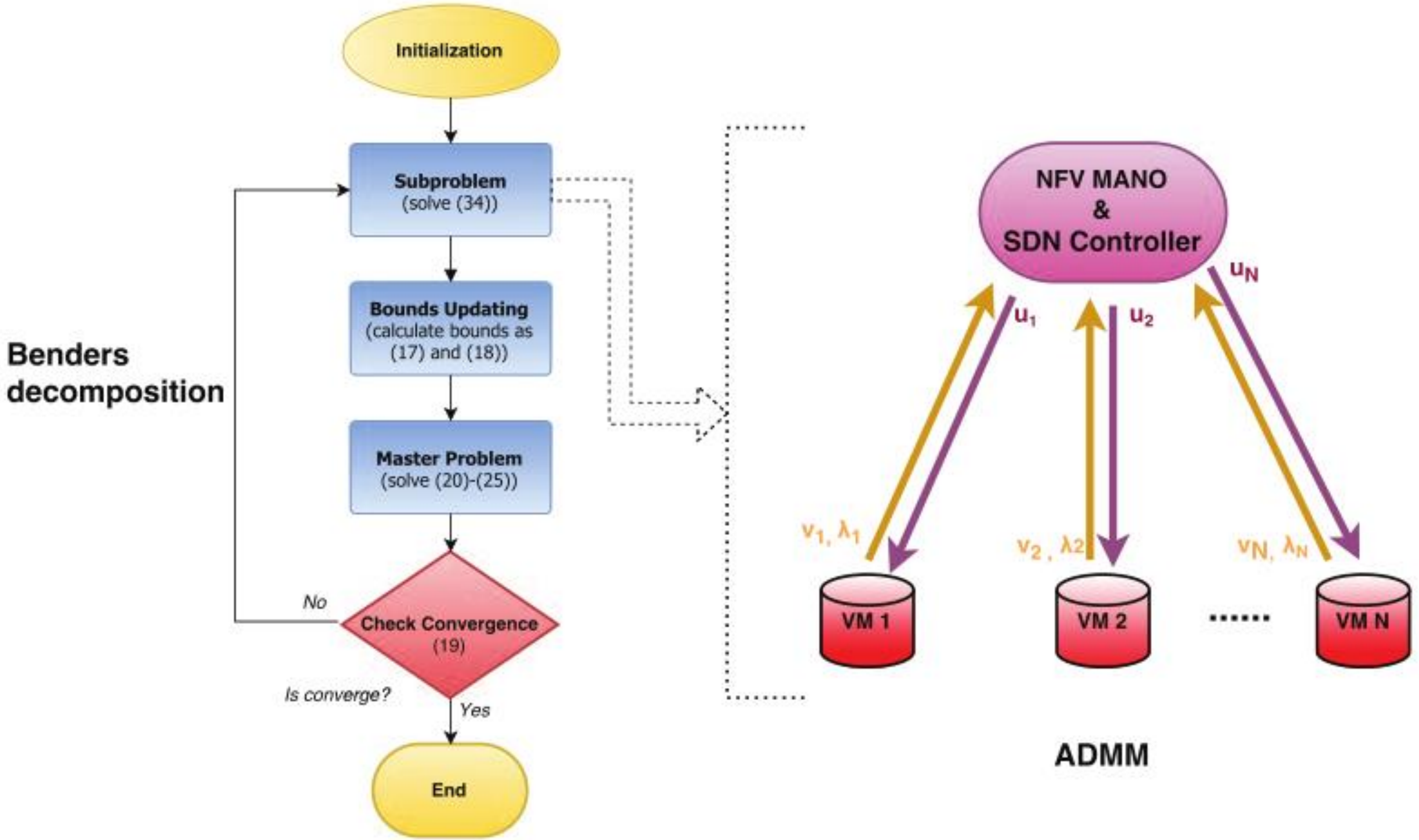
$$\mu_{f_k} \leq \frac{\sum_{n=1}^{N} v_{n,k+1}^m}{\sum_{n=1}^{N} v_{n,k}^m} \quad \forall k,$$

Streaming constraint

$$\delta_{n,k}^m \in \{0, 1\} \quad \forall m, n, k.$$

UNIVERSITY of HOUSTON
YOU ARE THE PRIDE
Department of Electrical and Computer Engineering

Proposed Algorithm



Benders Decomposition

- General MILP

$$\underset{x_1,\ldots,x_n;\, y_1,\ldots,y_m}{\text{minimize}} \quad \sum_{i=1}^{n} c_i\, x_i + \sum_{j=1}^{m} d_j\, y_j$$

subject to

$$\sum_{i=1}^{n} a_{\ell i}\, x_i + \sum_{j=1}^{m} e_{\ell j}\, y_j = b_\ell; \quad \ell = 1,\ldots,q$$

$$x_i^{\text{down}} \le x_i \le x_i^{\text{up}}, \qquad x_i \in \mathbb{N}; \qquad i = 1,\ldots,n$$

$$y_j^{\text{down}} \le y_j \le y_j^{\text{up}}, \quad y_j \in \mathbb{R}; \qquad j = 1,\ldots,m\,.$$

Benders Decomposition

Step 0: Initialization. Initialize the iteration counter, $\nu = 1$, and let

$$x_i^{(\nu)} = \begin{cases} x_i^{\text{down}} & \text{if } c_i \geq 0 \\ x_i^{\text{up}} & \text{if } c_i < 0 \end{cases}$$

$$\alpha^{(\nu)} = \alpha^{\text{down}}$$

Step 1: Subproblem solution. Solve the LP subproblem

$$\underset{y_1,\ldots,y_m}{\text{minimize}} \quad \sum_{j=1}^{m} d_j \, y_j$$

subject to

$$\sum_{j=1}^{m} e_{\ell j} \, y_j = b_\ell - \sum_{i=1}^{n} a_{\ell i} \, x_i; \quad \ell = 1, \ldots, q$$

$$y_j^{\text{down}} \leq y_j \leq y_j^{\text{up}}, \qquad y_j \in \mathbb{R}; \quad j = 1, \ldots, m$$

$$x_i = x_i^{(\nu)} : \quad \lambda_i; \quad i = 1, \ldots, n \, .$$

The solution of this problem is $y_1^{(\nu)}, \ldots, y_m^{(\nu)}$ with dual variable values $\lambda_1^{(\nu)}, \ldots, \lambda_n^{(\nu)}$.

Benders Decomposition

Step 2: Convergence checking. Compute upper and lower bounds of the optimal value of the objective function of the original problem

$$z_{\text{up}}^{(\nu)} = \sum_{i=1}^{n} c_i \, x_i^{(\nu)} + \sum_{j=1}^{m} d_j \, y_j^{(\nu)}$$

$$z_{\text{down}}^{(\nu)} = \sum_{i=1}^{n} c_i \, x_i^{(\nu)} + \alpha^{(\nu)} \, .$$

If $z_{\text{up}}^{(\nu)} - z_{\text{down}}^{(\nu)}$ is smaller than a pre-specified tolerance, stop, the optimal solution is $x_1^{(\nu)}, \ldots, x_n^{(\nu)}$ and $y_1^{(\nu)}, \ldots, y_m^{(\nu)}$. Otherwise, the algorithm continues with the next step.

Benders Decomposition

Step 3: Master problem solution. Update the iteration counter, $\nu \leftarrow \nu+1$. Solve the MILP master problem

$$\underset{x_1,\ldots,x_n,\alpha}{\text{minimize}} \quad \sum_{i=1}^{n} c_i x_i + \alpha$$

subject to

Benders cut

$$\alpha \geq \sum_{j=1}^{m} d_j y_j^{(k)} + \sum_{i=1}^{n} \lambda_i^{(k)}(x_i - x_i^{(k)}); \quad k = 1,\ldots,\nu-1$$

$$x_i^{\text{down}} \leq x_i \leq x_i^{\text{up}}, \qquad x_i \in \mathbb{N}; \quad i = 1,\ldots,n$$

$$\alpha \geq \alpha^{\text{down}} \ .$$

The solution of this problem is $x_1^{(\nu)},\ldots,x_n^{(\nu)}$ and $\alpha^{(\nu)}$. The algorithm continues with Step 1. $\square$

UNIVERSITY of HOUSTON
YOU ARE THE PRIDE
Department of Electrical and Computer Engineering

Alternating Direction Method of Multipliers (ADMM) to solve subproblems

$$\boxed{\begin{array}{ll} \min & f(x) + g(z) \\ \text{s.t.} & Ax + Bz = c. \end{array}}$$

Augmented Lagrangian function

$$\mathcal{L}(x, z, \lambda) = f(x) + g(z) + \lambda^T(Ax + Bz - c) + \frac{\rho}{2}||Ax + Bz - c||^2$$

Iterative procedure to solve an optimization problem using ADMM

$$x[t+1] := \arg\min_x \mathcal{L}(x, z[t], \lambda[t]),$$

California government

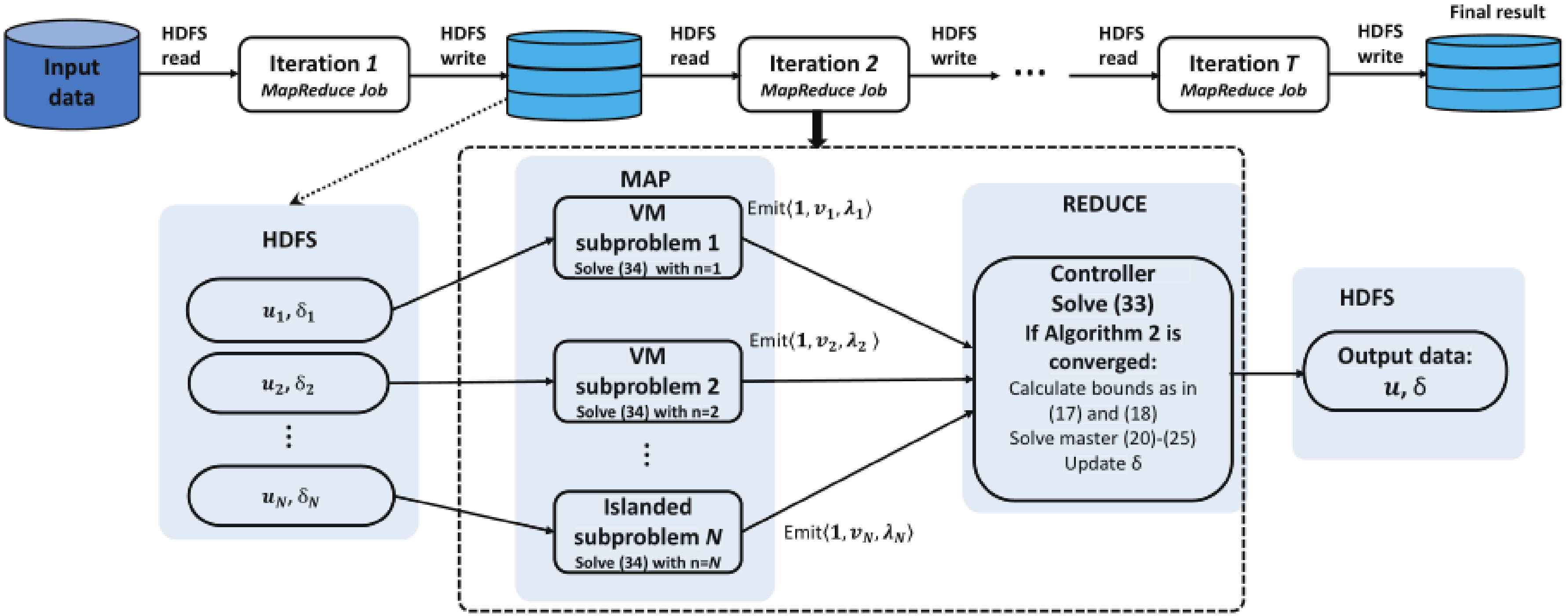
$$z[t+1] := \arg\min_z \mathcal{L}(x[t+1], z, \lambda[t]),$$

Texas government

$$\lambda[t+1] := \lambda[t] + \rho\left(Ax[t+1] + Bz[t+1] - c\right),$$

US Congrress

Implementation using Hadoop MapReduce



Simulation Results

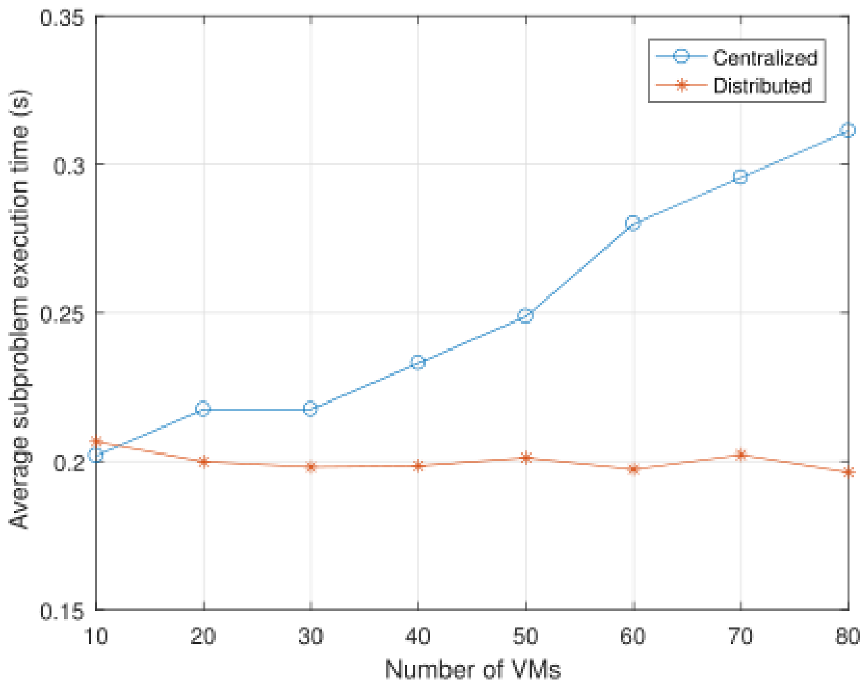


Fig. 9: The average execution time of subproblem versus number of VMs

UNIVERSITY of HOUSTON
YOU ARE THE PRIDE
Department of Electrical and Computer Engineering

Outline

- ❑ **Introduction and Motivation for C^3 Networks**
- ❑ **Big Data Analysis and Cross Layer Optimization**
  - ❑ **Wireless Network Function Virtualization**
  - ❑ **Mobile Social Networks over D2D**
  - ❑ **Deep Learning Analysis**
- ❑ **Conclusions**

D2D Communication

- Device-to-Device (D2D) Communications/Sidelink
  - Technology that enables devices to connect directly without relying on infrastructure of access points or base stations.

① Increase network capacity
② Extend (edge) coverage
③ Offload data
④ Improve energy efficiency
⑤ Create new applications

licensed (Inband) spectrum
  - Underlay vs. Overlay
  - UpLink vs. DownLink
  - Mode Selection (D2D mode vs. Cellular Mode)

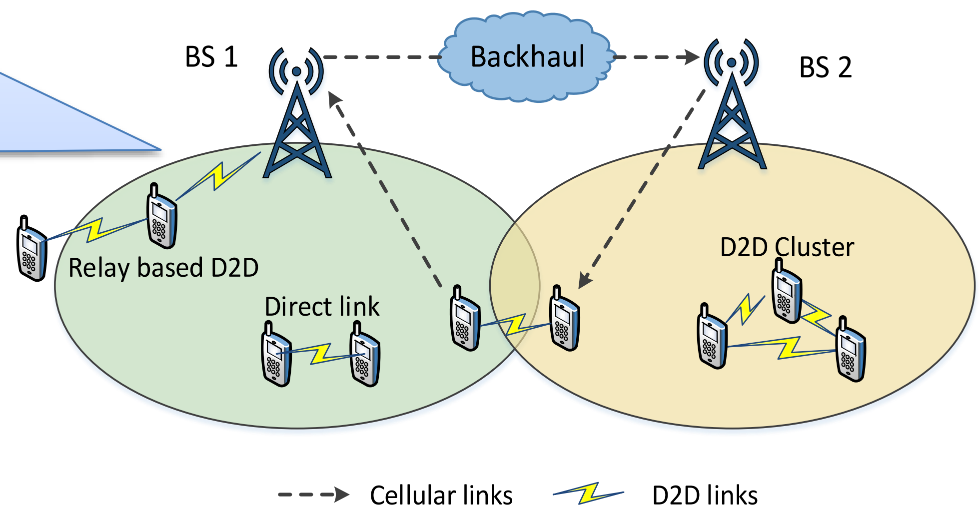


Figure: Scenario of D2D Communications

TS36.300 v12 (2015.02)

section 23.10 for D2D Communications

section 23.11 for D2D Discovery (proximity detection for commercial services)

TS36.211 v12 (2015.10)

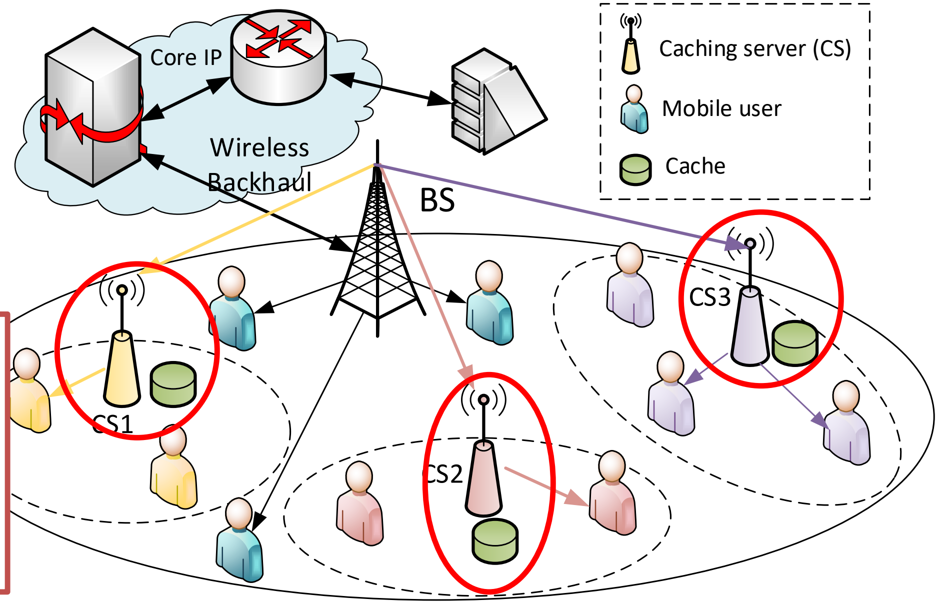
section 9: a sidelink is used for ProSe direct communication and ProSe direct discovery between UEs. (ProSe – Proximity Services)

Exploiting D2D in Mobile Social Network

- Imminent wireless capacity crunch
  - Smartphones with larger storage and higher computing capability
  - Mobile social platforms, e.g.,
  - Virtual and Social Community
    - Social tie (family members, club members…)
    - Social relationship with common interest
    - Social interactions
  - Continuous (uninterrupted) wireless connection
  - File sharing & Online gaming & Video dissemination
- D2D---exploit benefits from social networking
  - in terms of pairs and clusters
  - to offload the increasing traffic from base stations (BSs)
  - to meet the higher speed demands for mobile users

Wireless Distributed Storage Systems

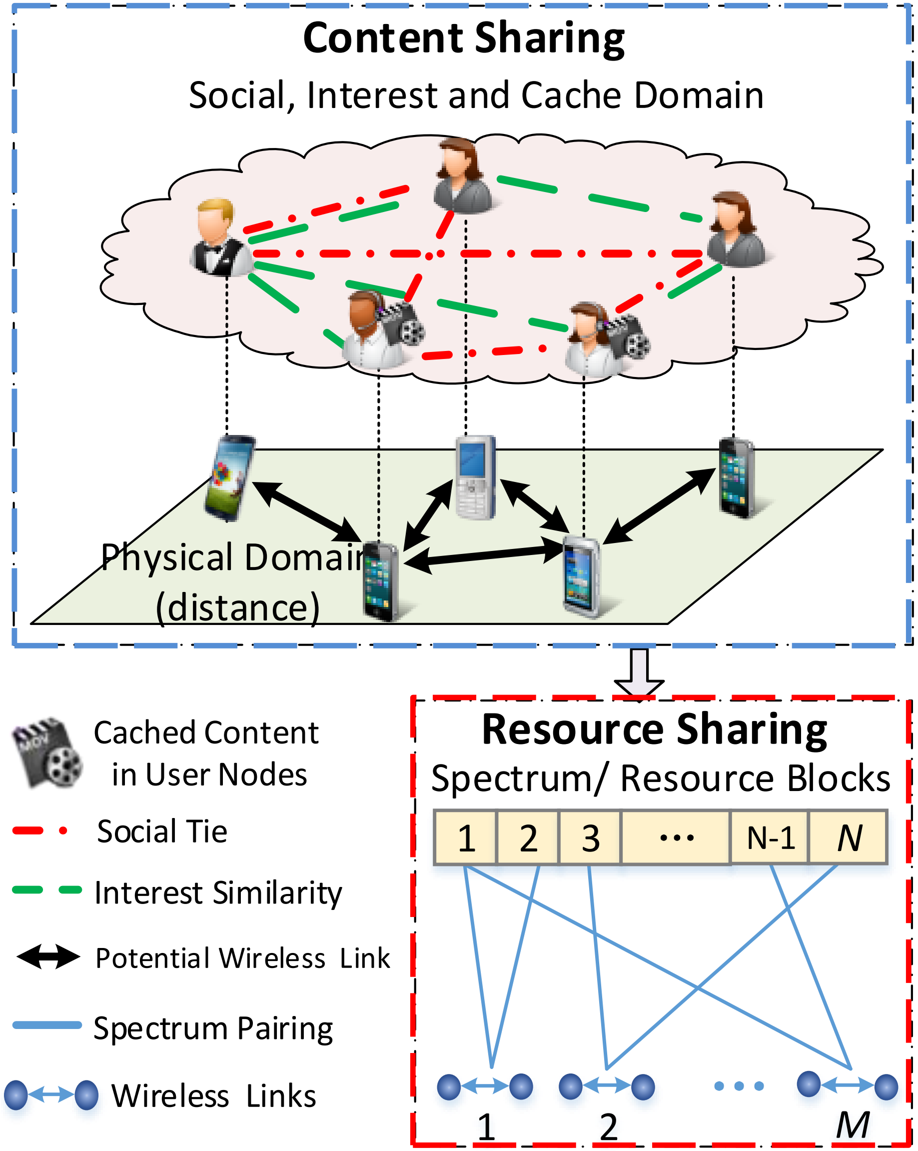
- **Wireless distributed storage:**
  - Store the popular content files at the BS/mobile devices during off-hours to improve the end-to-end performance and reduce backhaul loading at peak-hours
  - Small BS or powerful user nodes
  - Popular files can be received from the BS and stored in their own caching
  - Files in its cache can be accessed by other users within its coverage at a later time
- **Benefits of wireless storage:**
  - Capacity enhancement
  - Backhaul traffic offloading
  - Interference mitigation



- **Wireless caching schemes:**
  - **Coded caching** to create coded multicast opportunities
  - **Proactive caching** to exploit both the spatial and social structure of the wireless networks
  - **D2D caching networks** to reduce number of hops and balance load
  - **Asymptotic scaling laws** in large wireless networks

Focus:
In D2D-based wireless distributed storage systems, resource allocation and content sharing are investigated in different scenarios with different objective functions

D2D-Based Distributed Storage Systems



Upper layer: Content sharing for D2D partners to form Links

Main considerations:

- Social relationship
- Interest similarity
- Physical proximity
- Caching capability
- Computing capability
- Content coding

Lower layer: Resource sharing for CUE and D2D Links in D2D Underlay

Main considerations:

- Channel state information
- Power control
- Co-channel interference

Matching Game Definitions

Basic elements (Stable Marriage (SM)):

1. Agents: A set of men, and a set of women;

2. Preference list: A sorted list of men/women based on her/his preferences;

3. Blocking pair (BP) (m,w):

m prefers w to his current partner;

w prefers m to her current partner;

4. Stable matching: A matching admit no BPs.

5. Gale-Shapley (GS) algorithm: find a stable matching in SM

UNIVERSITY of HOUSTON
YOU ARE THE PRIDE
Department of Electrical and Computer Engineering

GS Algorithm


Adam

Geeta, Heiki, Irina, Fran


Fran


Bob

Irina, Fran, Heiki, Geeta


Geeta

Carl > Adam

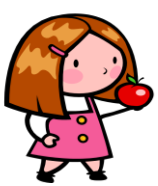

Carl

Geeta, Fran, Heiki, Irina


Heiki


David

Irina, Heiki, Geeta, Fran


Irina

David > Bob

We reach a stable marriage!

UNIVERSITY of HOUSTON
YOU ARE THE PRIDE
Department of Electrical and Computer Engineering

Outline

- ❑ Introduction and Motivation for C^3 Networks
- ❑ Big Data Analysis and Cross Layer Optimization
    - ❑ Wireless Network Function Virtualization
    - ❑ Mobile Social Networks over D2D
    - ❑ Deep Learning Analysis
- ❑ Conclusions

Reinforcement Learning

Effective and Efficient Method

Dummy

Policy
(Machine)

Expert

An Example

Today

1

1

Have Fun

-100

Fencing Event

0

0

1

Study & Discussion

0

IEEE Globecom

Submit Paper

100

Rewards

a: action

s: state

$$
R(s,a) = \begin{array}{c|c}
 & \begin{array}{ccccc} \text{Today} & \text{Study} & \text{Submit} & \text{Fun} & \text{Fencing} \end{array} \\
\begin{array}{c} \text{Today} \\ \text{Study} \\ \text{Submit} \\ \text{Fun} \\ \text{Fencing} \end{array} &
\begin{bmatrix}
-1 & 0 & -1 & 1 & -1 \\
-1 & -1 & 0 & 1 & -1 \\
-1 & -1 & -1 & -1 & 100 \\
-1 & 1 & -1 & 1 & -100 \\
-1 & -1 & -1 & -1 & -1
\end{bmatrix}
\end{array}
$$

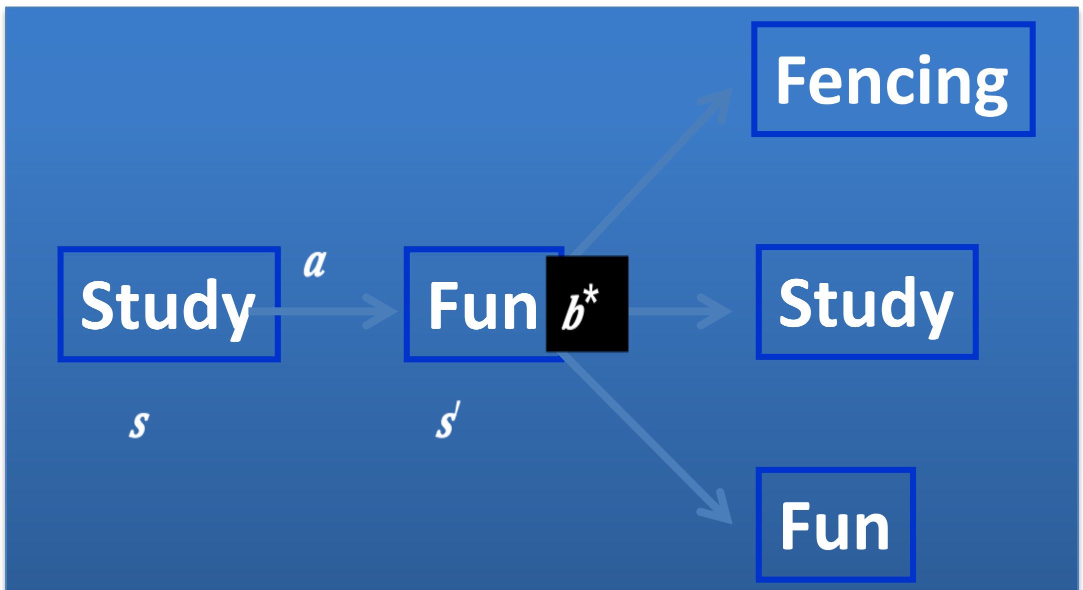
Initially set to zero

Q-Learning

Learning Rate

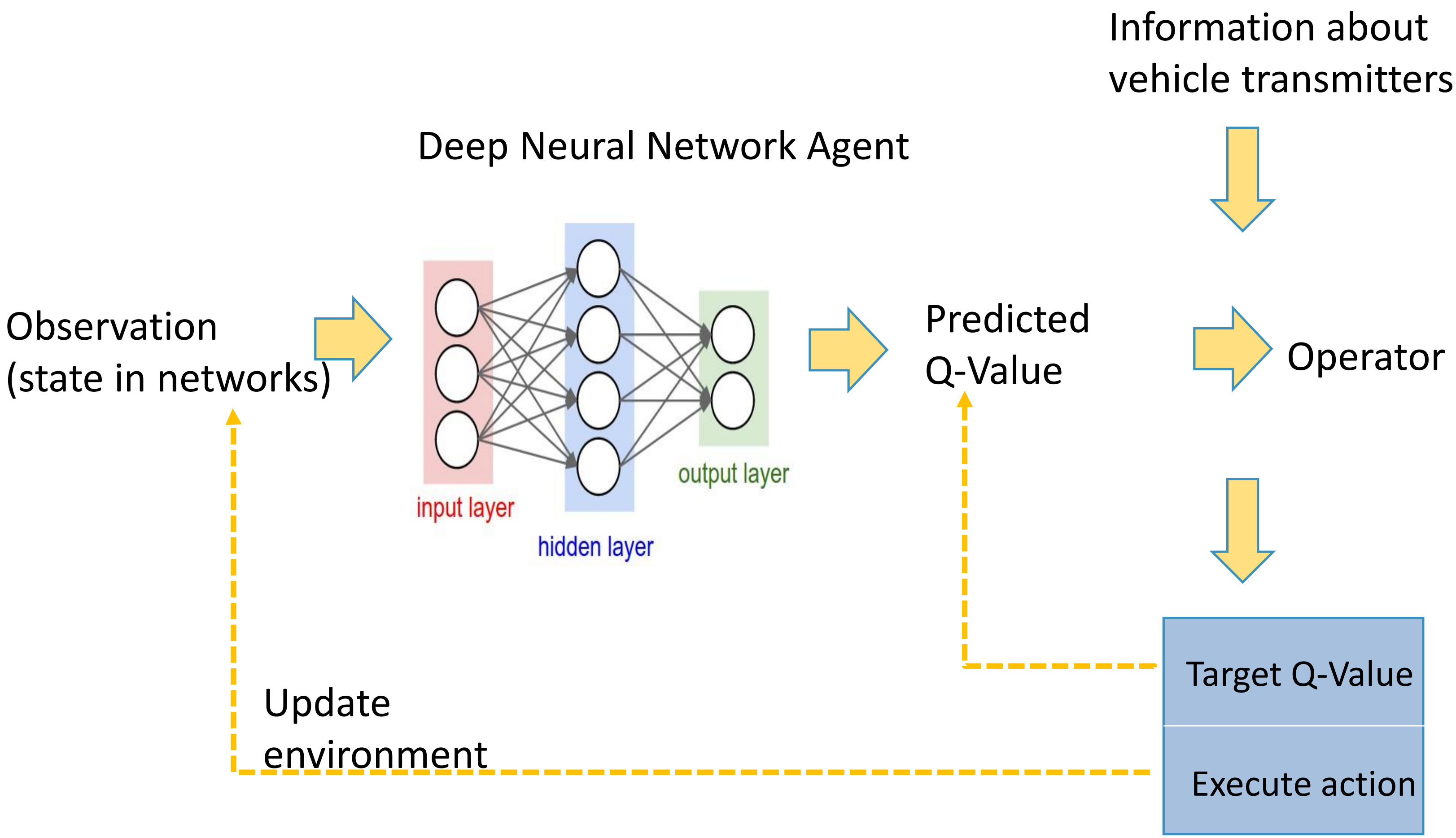
Discount Factor

$$Q_{t+1}(s,a) = (1-\alpha_t)Q_t(s,a) + \alpha_t\left[R_t(s,a) + \beta \max_b Q_t(s',b)\right]$$




HOUSTON
SWORD S

Deep Reinforcement Learning Framework



AlphaGo Zero beats AlphaGo beats Human

UNIVERSITY of HOUSTON
YOU ARE THE PRIDE
Department of Electrical and Computer Engineering

Conclusions

- Communication, Caching and Computing are potential solutions to achieve 5G
- Challenge is how to write utility to link them
- Scenarios vs. solutions
  - Network virtualization based on Benders decomposition and ADMM
  - Mobile social networks over D2D based on matching
  - Deep reinforcement learning
- Just a small peek on a new paradigm

UNIVERSITY of HOUSTON
YOU ARE THE PRIDE
Department of Electrical and Computer Engineering

Thank You!

Proposed Algorithm

$$
\begin{aligned}
\min_{\boldsymbol{\delta}^m \alpha} \quad & Q_p + \alpha \\
\text{s.t} \quad & 1 \leq \sum_{n=1}^{N} \delta_{n,k}^m \leq N \quad \forall k, m, \\
& Q_t(\mathbf{v}^{m,(j)}) + \boldsymbol{\theta}^{m,(j)}(\boldsymbol{\delta}^m - \boldsymbol{\delta}^{m,(j)}) \leq \alpha \quad j = 1, \ldots, l-1, \\
& \sum_{k=1}^{K} \delta_{n,k}^m v_{n,k}^m \phi_k \leq C_n \quad \forall n, m, \\
& \alpha \geq \alpha^{down} \\
& \delta_{n,k}^m \in \{0, 1\} \quad \forall m, n, k,
\end{aligned}
$$

Algorithm 1 NFV Resource al

Initialize: loop index l, $U_{\delta^m_n}$

while $Q_{up}^l - Q_{down}^l \geq \varepsilon$ **do**

Subproblem

acquire $\mathbf{v}^{m(l)}$ and $\boldsymbol{\theta}^m$

Bounds calculation

calculate upper and l

Master problem

step 1: update loop index $l = l + 1$

step 2: add new Benders cut to the problem (20)

step 3: solve the problem (20) to obtain the optimal value of $\boldsymbol{\delta}^m$ and α

end while

Proposed Algorithm

- Transformation of subproblem

$$
\begin{aligned}
\min_{\mathbf{v}^m} \quad & Q_t \\
\text{s.t} \quad & \sum_{k=1}^{K} \delta_{n,k}^m v_{n,k}^m \phi_k \le C_n \quad \forall n, m, \\
& \sum_{n=1}^{N} v_{n,k}^m \ge R_k \quad \forall k, \\
& \mu_{f_k} \le \frac{\sum_{n=1}^{N} v_{n,k+1}^m}{\sum_{n=1}^{N} v_{n,k}^m} \quad \forall k, \\
& \delta_{n,k}^m = \delta_{n,k}^{m,(l)} : \theta_{n,k}^{m,(l)} \quad \forall m, n, k.
\end{aligned}
$$

$\Rightarrow$

$$
\begin{aligned}
\min_{\mathbf{v}^m} \quad & Q_t \\
\text{s.t} \quad & \sum_{k=1}^{K} \delta_{n,k}^m v_{n,k}^m \phi_k \le C_n \quad \forall n, m, \\
& \sum_{n=1}^{N} u_{n,k}^m \ge R_k \quad \forall k, \\
& \mu_{f_k} \sum_{n=1}^{N} u_{n,k}^m - \sum_{n=1}^{N} u_{n,k+1}^m \le 0 \quad \forall k \\
& \delta_{n,k}^m = \delta_{n,k}^{m,(l)} : \theta_{n,k}^{m,(l)} \quad \forall m, n, k \\
& v_{n,k}^m = u_{n,k}^m \quad \forall n, k.
\end{aligned}
$$

Global copy variable

Proposed Algorithm

Algorithm 2 Distributed algorithm for the subproblem based on ADMM

Initialize: t, $\boldsymbol{\lambda}$, $\mathbf{v}$
while the stopping criterion is not satisfied **do**
  Controller update
  repeat
    wait
  until receive updated $\boldsymbol{\lambda}$, $\mathbf{v}$ from all N distributed VMs
   step 1: Solve the problem (44) and obtain the optimal solution $\tilde{\mathbf{u}}$
   step 2: Then send $\tilde{\mathbf{u}}$ to the VMs
   step 3: Update
———
  Each VM u
  repeat
    wait
  until receive
   step 1: So
   step 2: Up
     $\boldsymbol{\lambda}[$
   step 3: Se
end while

$$\min_{\mathbf{v}^{m,(l)}} \quad \sum_{k=1}^{K} \gamma_k v_{n,k}^m + \sum_{k=1}^{K} \lambda_{n,k} + \frac{\rho}{2} \sum_{k=1}^{K} \| v_{n,k}^m - \tilde{u}_{n,k}^m \|_2^2 \tag{45}$$

$$\text{s.t} \quad \sum_{k=1}^{K} \delta_{n,k}^m v_{n,k}^m \phi_k \le C_n \quad \forall m.$$

$$\min_{\mathbf{u}^{m,}}$$

$$\text{s.t} \quad \sum_{n=1} u_{n,k}^m \ge R_k \quad \forall k,$$

$$\mu_{f_k} \sum_{n=1}^{N} u_{n,k}^m - \sum_{n=1}^{N} u_{n,k+1}^m \le 0 \quad \forall k,$$

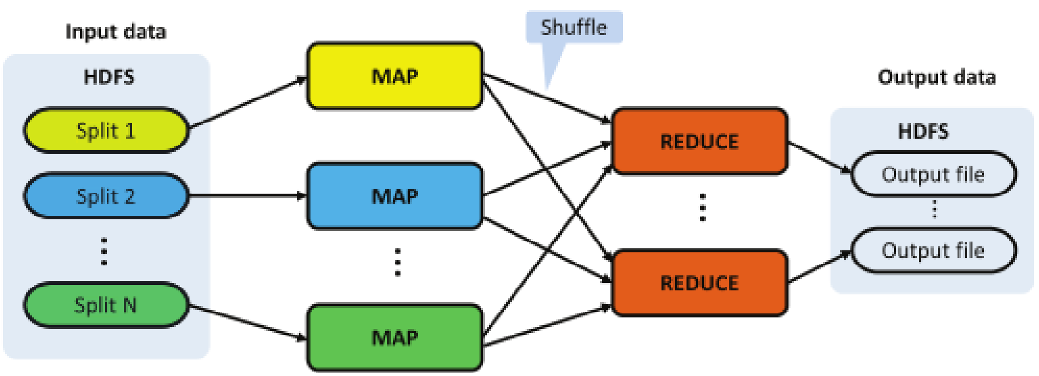
Implementation using Hadoop MapReduce

Algorithm 3 ADMM-based Bender decomposition using Hadoop MapReduce

function MAP(vm_ID, inputData)
  Load data of previous iteration from HDFS corresponding to vm_ID
  Solve subproblem corresponding to each VM in (45)
  Update $\boldsymbol{\lambda}_n$ using (46)
  EMIT $\langle 1, \{\boldsymbol{v}_n, \boldsymbol{\lambda}_n\}\rangle$
end function

function REDUCE(key, Data from Mappers)
  Concatenate $\{\boldsymbol{v}_n, \boldsymbol{\lambda}_n\}$ from all VMs
  Solve controller subproblem (44) for $\boldsymbol{u}$
  if Algorithm 2 is converged **then**
    Calculate Upper bound and Lower bound as in (17) and (18)
    Sove master problem in (20)
    Update $\boldsymbol{\delta}$
  end if
  EMIT $\langle\{\boldsymbol{u}, \boldsymbol{\delta}\}\rangle$
end function

function MAIN(inputPath, outputPath)
  Initialization
  while $Q^l_{up} - Q^l_{down} \geq \varepsilon$ **do**
    run MapReduceJob (inputPath, outputPath)
    $t \leftarrow t + 1$
  end while
end function



Simulation Results

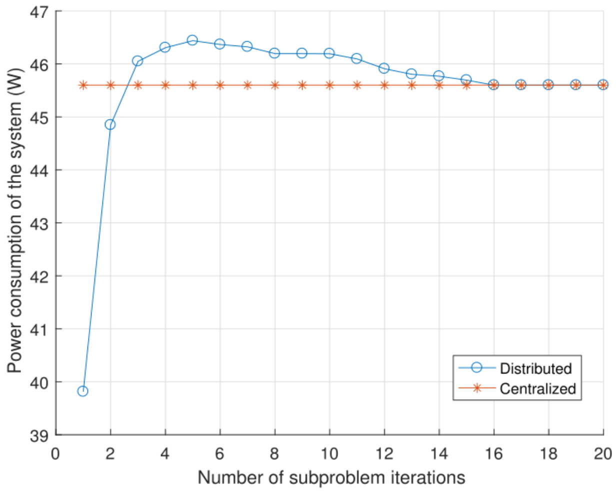


Fig. 6: The convergence performance of ADMM

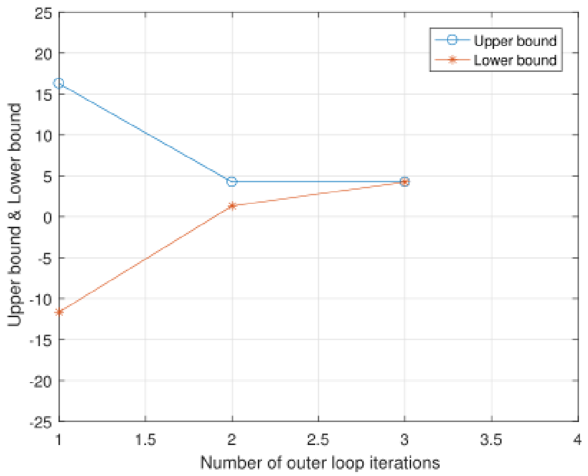


Fig. 7: The convergence performance of Benders decomposition