Cloud Computing with 4G LTE-A

國立雲林科技大學 資工系

Ben-Jye Chang, Ph.D., Chair
張本杰 教授兼系主任

Outline

- **Introduction**
  - Cloud Computing
  - 4G LTE-A
- **Related Works**
- **Motivations & Goals**
- **Network Model**
- **Proposed Approach**
  - Phase 1: the Loads of UMTS and LTE Determination Phase (LDP)
  - Phase 2: the connection Admission control and Classification Phase (ACP)
  - Phase 3: the Dynamic resource Allocations of UMTS and LTE Phase (DAP)
- **Numerical results**
- **Conclusions**

Cloud Computing Application

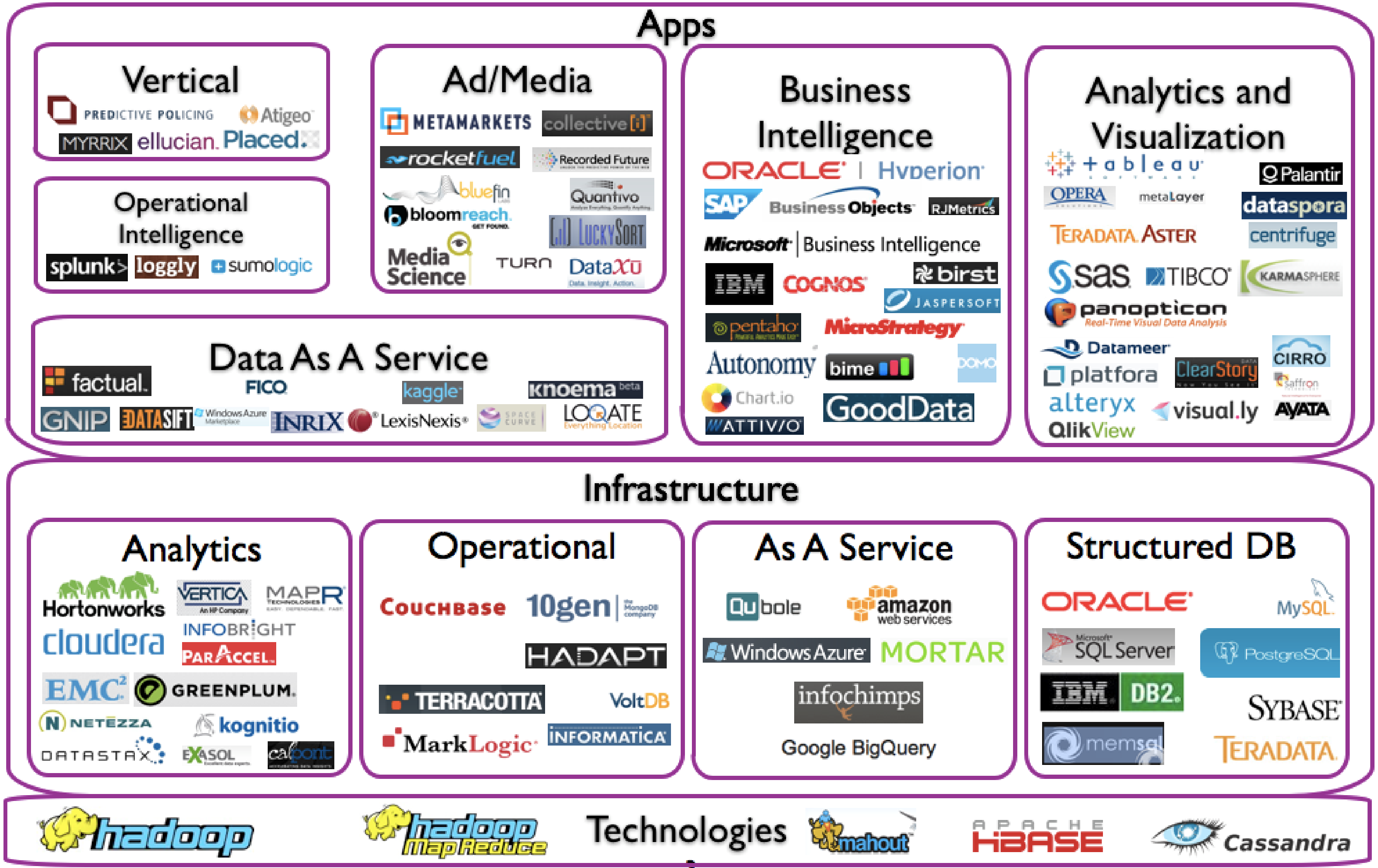
Gmail

Google Drive

iCloud

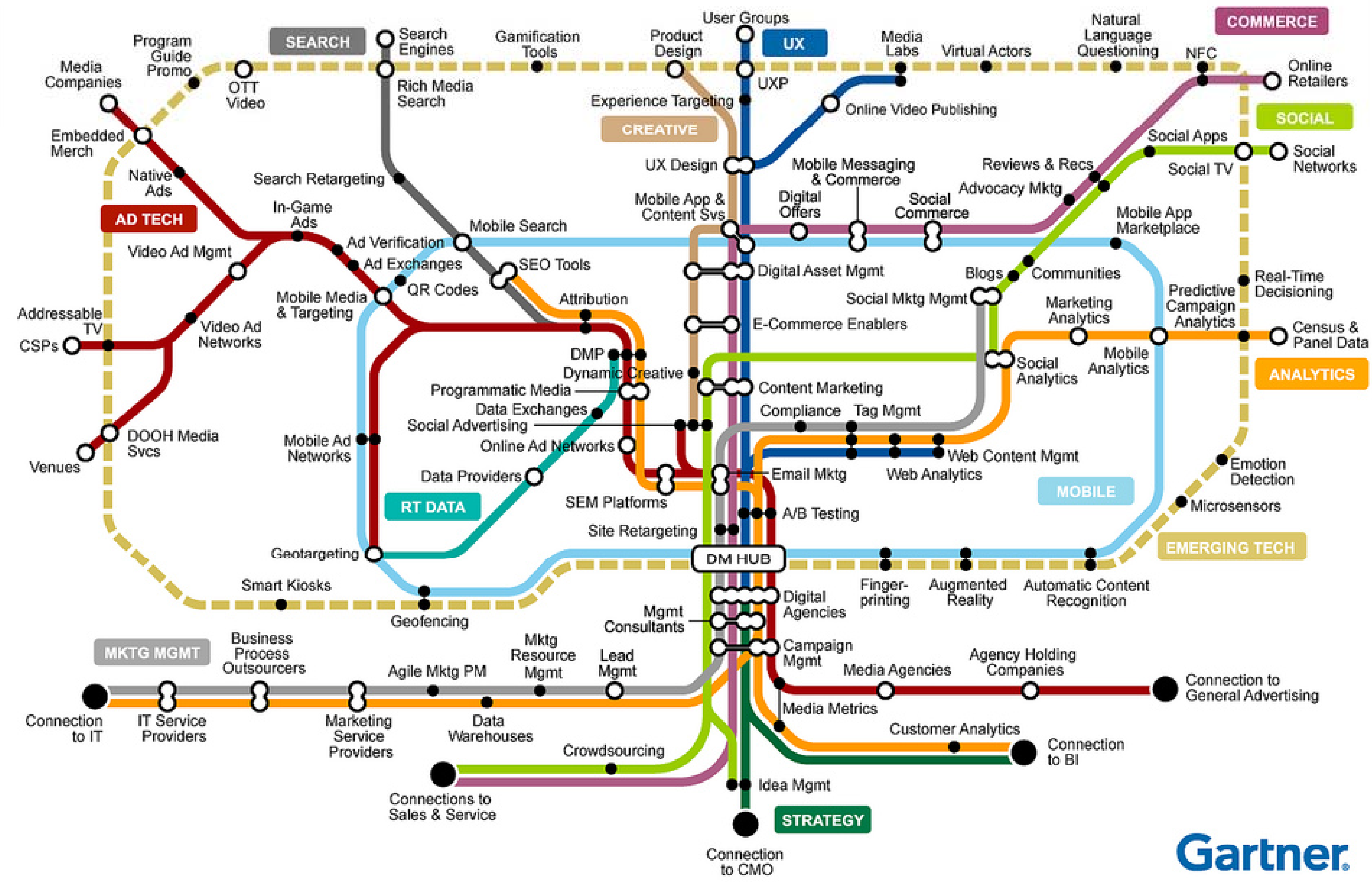
Dropbox

The Big Data Landscape



Gartner Digital Marketing Transit Map

The digital sphere is always evolving. Gartner for Marketing Leaders keeps the CMO and her team connected to the research, so they can keep the competition guessing. For more information: gartner.com/dmtransitmap



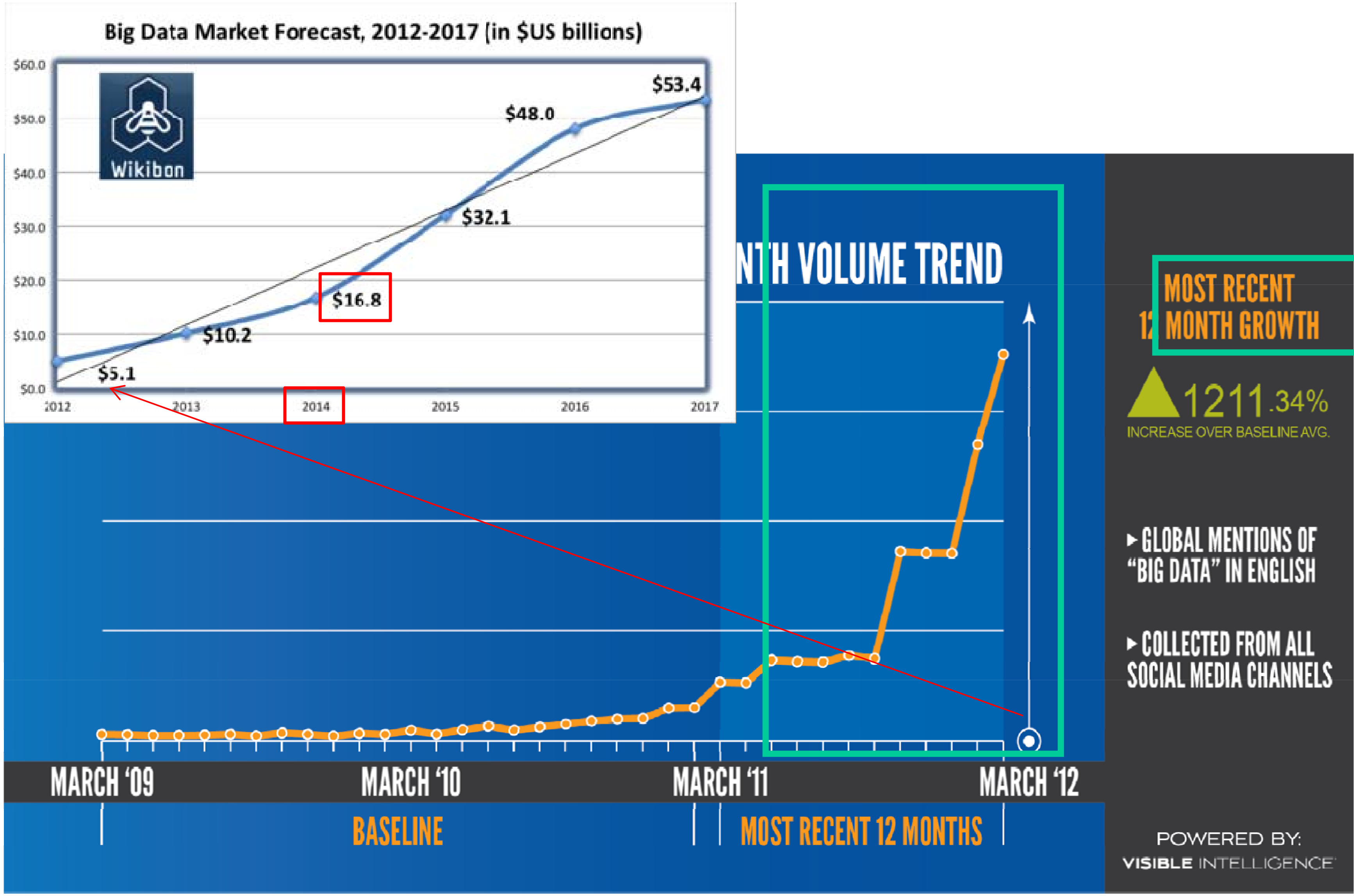
Off-line Connection
Vendor Station
Product Station

Why needs Cloud Computing

- Big Data!
  - A collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications.
    - EX: In 2012 , every day 2.5 quintillion (2.5×10^{19}) bytes of data were created.

  - Definition: Big data are high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization."

Why needs Cloud Computing(cont.)



Why needs Cloud Computing(cont.)

- Big Data: some examples
  - Big science: The Large Hadron Collider experiments represent about 150 million sensors delivering data 40 million times per second.
  - Science and research: collecting astronomical data, amassed more than 140 terabytes of information.
  - Decoding the human genome
  - Government: Big data analysis played a large role in Barack Obama's successful 2012 re-election campaign!

Why needs Cloud Computing(cont.)

Need gigabit rate wireless networking

LTE-A, LTE, 802.11ac, …

Why needs Cloud Computing(cont.)

- **10 reasons to consider the cloud computing**
  1. No need to worry about upfront costs.
  2. No need to worry about upgrades, patches or staffing resources.
  3. No need to worry about security (backups).
  4. No need to hire or train IT personnel.
  5. No need to worry about long deployment times.
  6. No need to worry about scalability.
  7. No need to worry about unknown service fees.
  8. No need to worry about data accessibility.
  9. No need to worry about time to value.
  10. No need to worry about managing your technology.

What is Cloud Computing

- Cloud Computing - is a colloquial expression used to describe a variety of different computing concepts that involve a large number of computers that are connected through a real-time communication network (**Internet**).

- **Cloud Computing aims to...**
  - Provide easy, scalable access to computing resources and IT services on-line, instead of building your own IT infrastructure.
  - Rent the storage, computing power and applications from the provider.

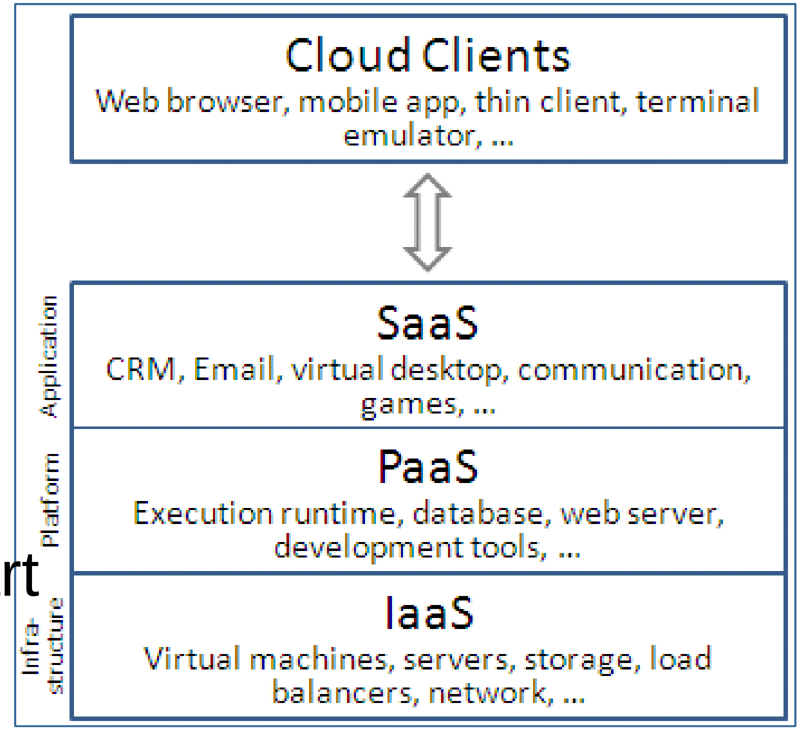
What is Cloud Computing (cont.)

Benefits of Cloud …

- 1. Flexibility
- 2. Disaster Recovery
- 3. Automatic Software Update
- 4. Capital Expenditure Free
- 5. Increased Collaboration
- 6. Work from Anywhere
- 7. Document Control
- 8. Security
- 9. Competitiveness
- 10. Environmentally Friendly

Service models

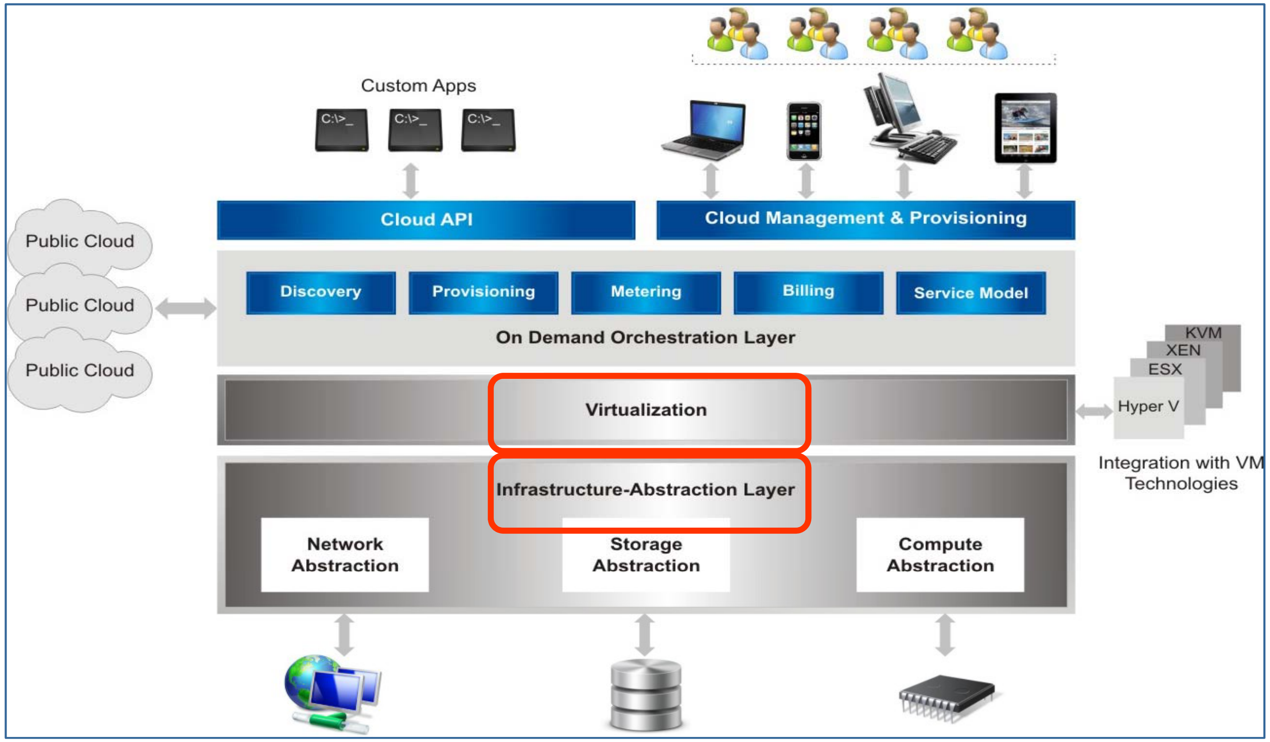
- Cloud computing providers offer their services according to several fundamental models:
  - Infrastructure as a Service (IaaS)
  - Platform as a Service (PaaS)
  - Software as a Service (SaaS)
- In 2012, Network as a Service (NaaS) and Communication as a Service (CaaS) were officially included by ITU (International Telecommunication Union) as part of the basic cloud computing models.



Infrastructure as a Service (IaaS)

- In the most basic cloud-service model, providers of IaaS offer computers - physical or (more often) virtual machines - and other resources.
- To deploy their applications, cloud users install operating-system images and their application software on the cloud infrastructure. In this model, the cloud user patches and maintains the operating systems and the application software.

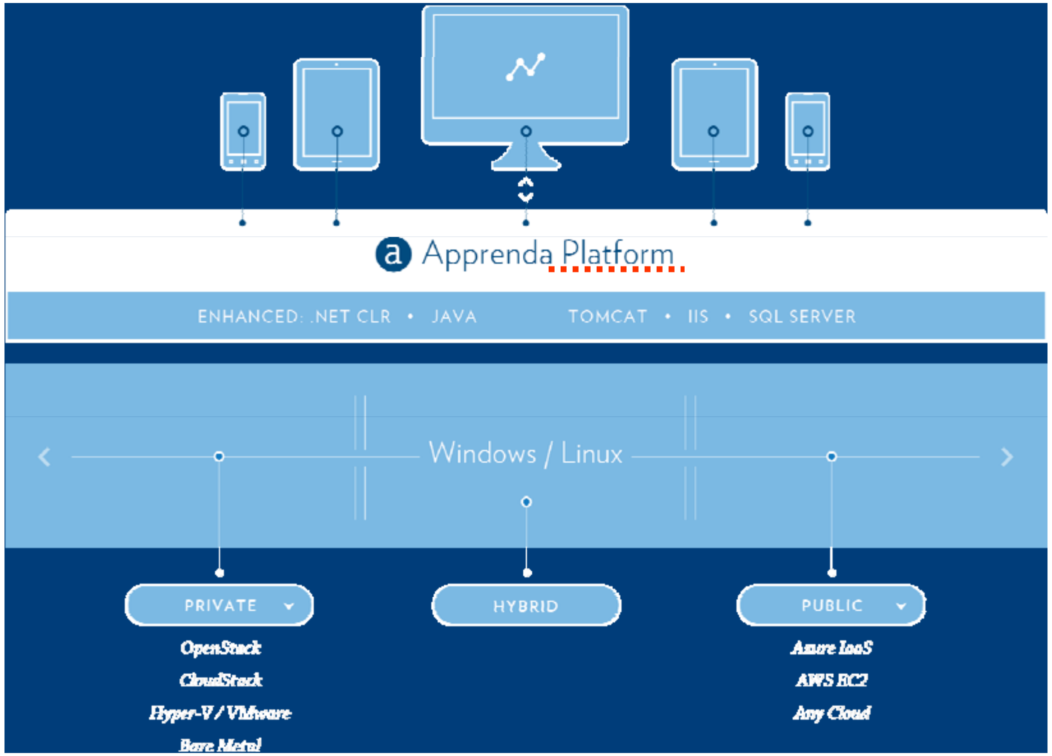
IaaS providers:
Amazon EC2
Google Compute Engine
Azure Services Platform
HP Cloud . . .



Platform as a Service (PaaS)

- In the PaaS model, cloud providers deliver a computing platform typically including operating system, programming language execution environment, database, and web server.
- Application developers can develop and run their software solutions on a cloud platform without the cost and complexity of buying and managing the underlying hardware and software layers.

PaaS providers:
AWS Elastic Beanstalk
Google APP Engine
Cloud Foundry
Windows Azure Cloud Services
. . .

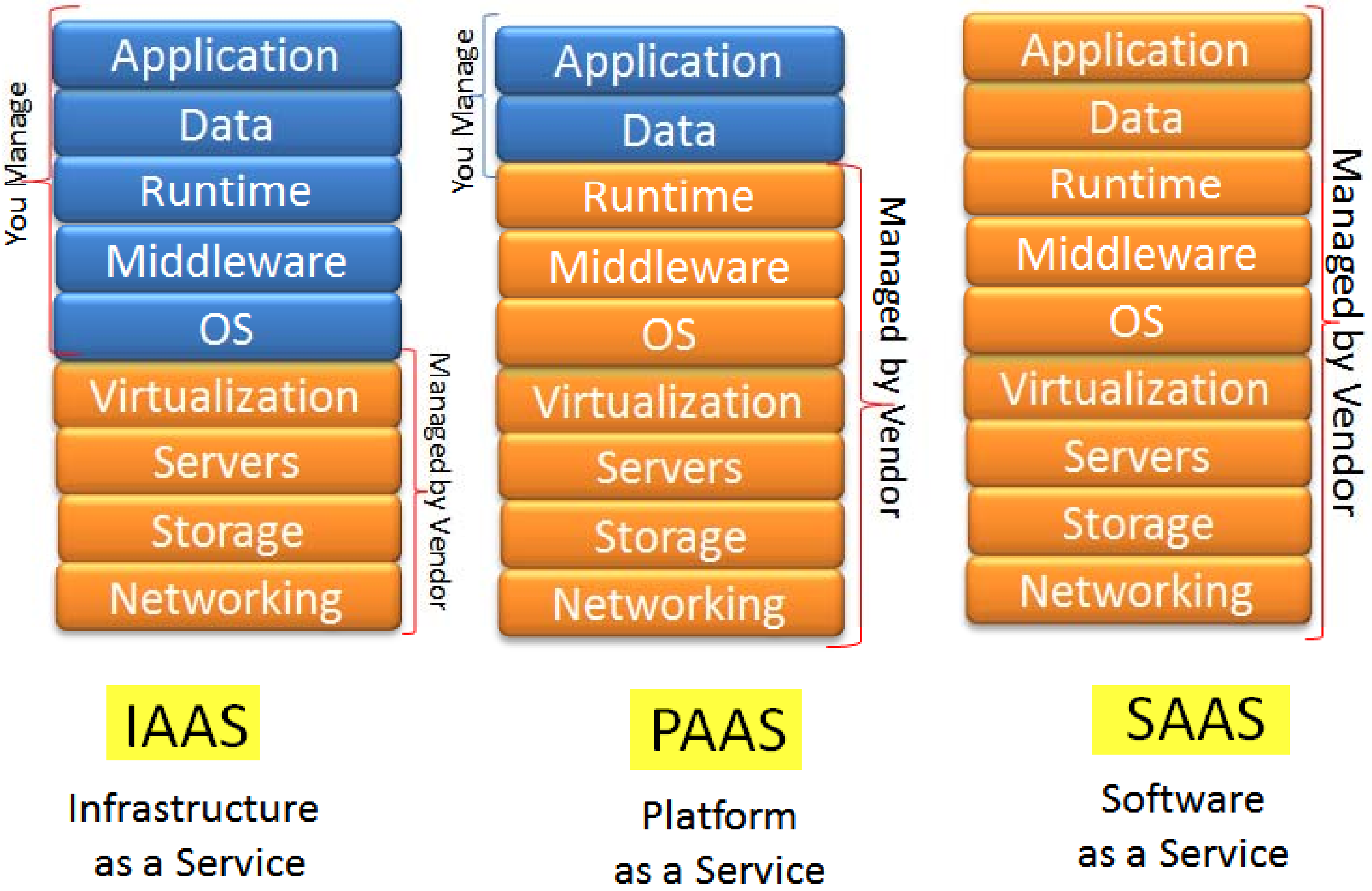


Software as a Service (SaaS)

- In the business model using software as a service (SaaS), users are provided access to application software and databases. Cloud providers manage the infrastructure and platforms that run the applications.
- Cloud users do not manage the cloud infrastructure and platform where the application runs. This eliminates the need to install and run the application on the cloud user's own computers, which simplifies maintenance and support.
- Cloud applications are different from other applications in their scalability - which can be achieved by cloning tasks onto multiple virtual machines at run-time to meet changing work demand.

SaaS providers:
Google APP
Microsoft Office 365
Petrosoft
GT Nexus . . .

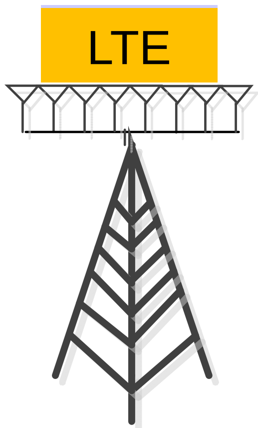
IaaS vs. PaaS vs. SaaS



Four Trends of Big Data in CC

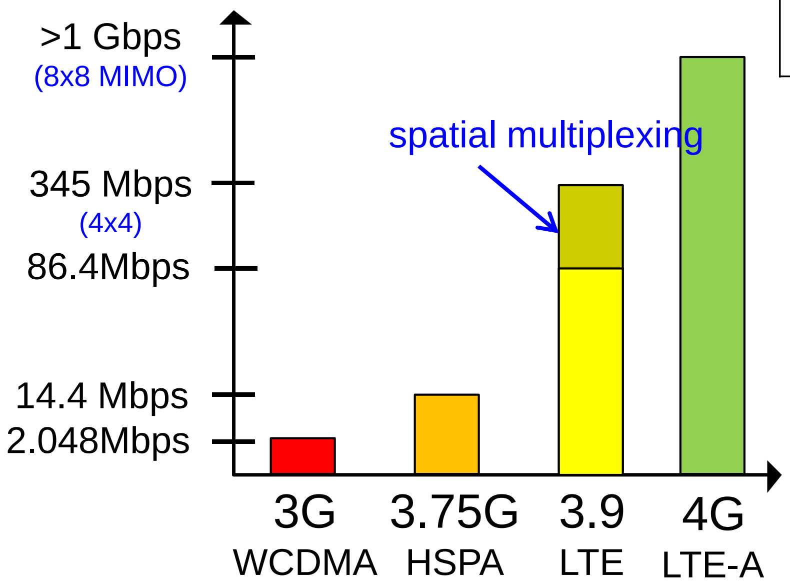
- T1. Big Data Will Transition From Hype to Actionable Insights (to realize applications)
- T2. Traditional Companies Will Derive Revenue ($) From Data
- T3. Visualization Tools (VM/Data Center) Will Become Essential Enterprise IT Investments
- T4. More Companies Will Implement Machine Learning (Sensors/Vehicles/Smart Machine) and Predictive Analytics

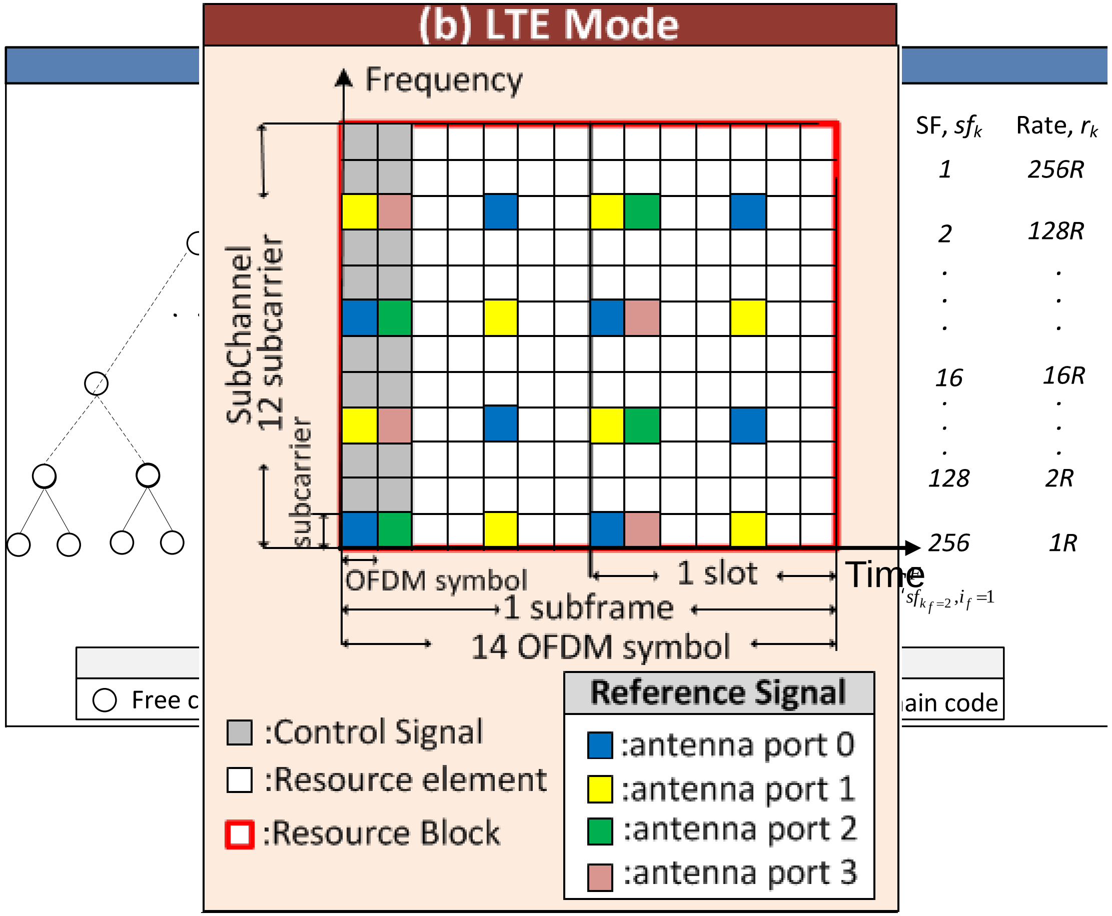
4G LTE-A



Base Station

Max.Network Speed

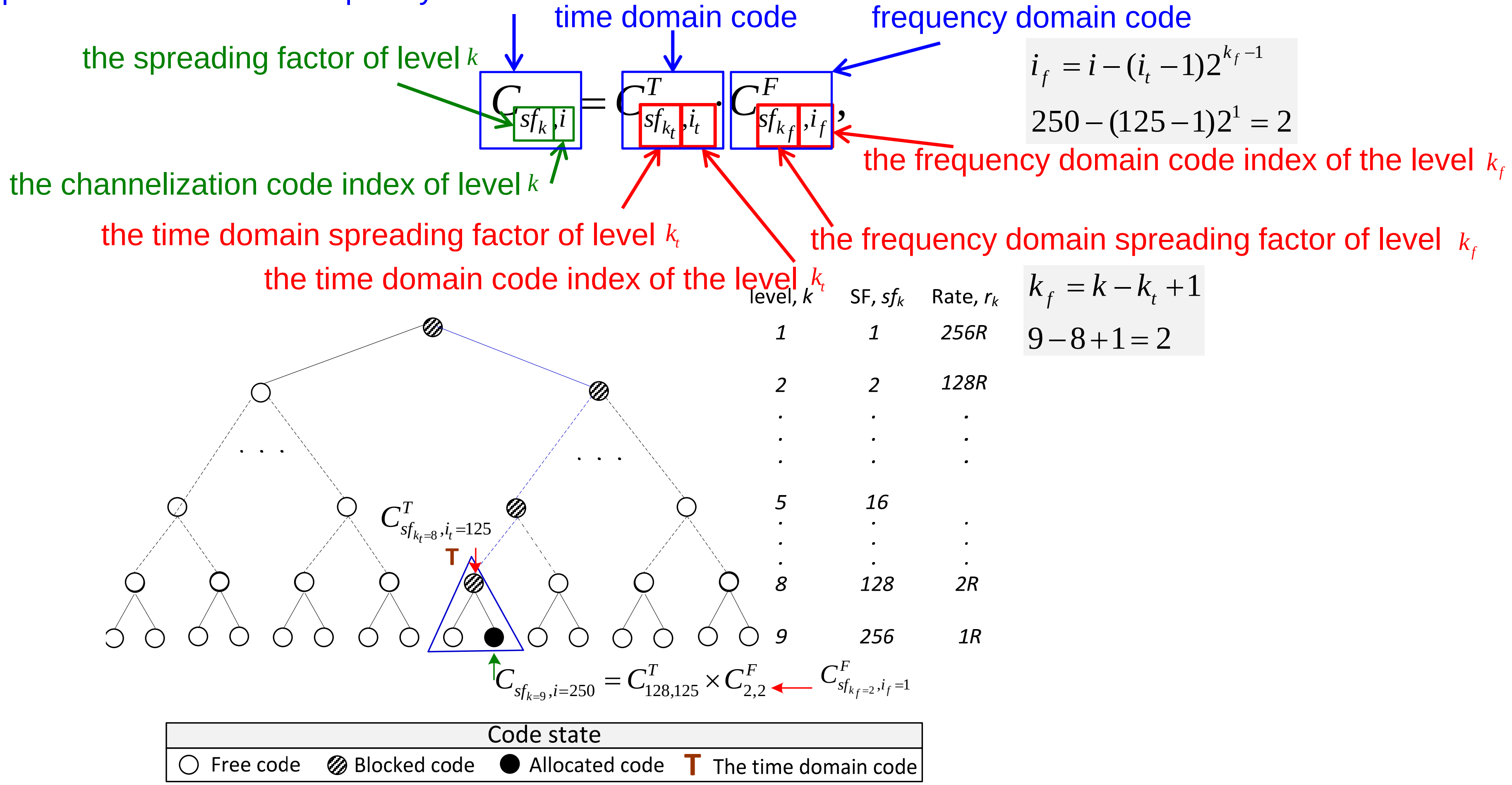




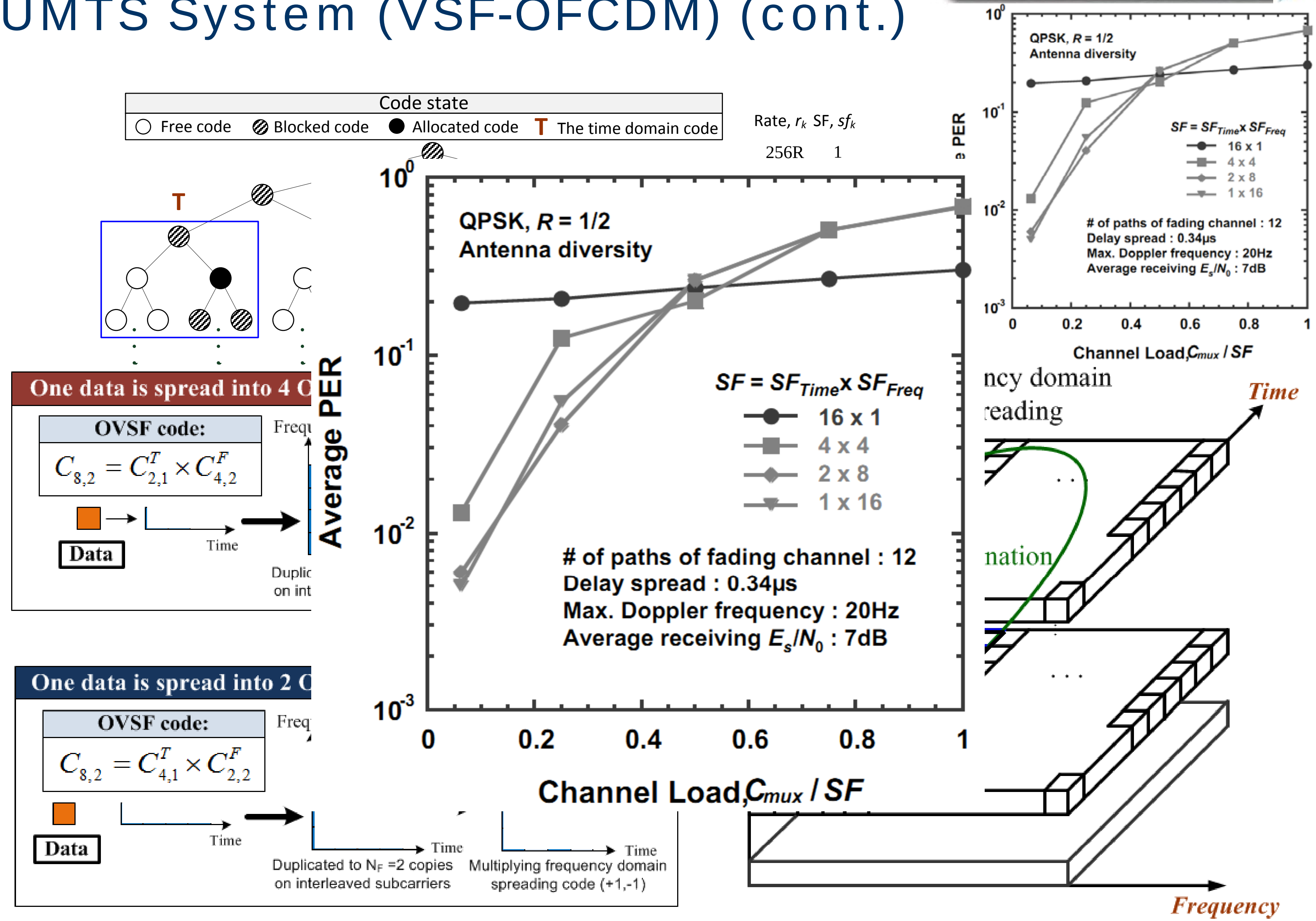
UMTS System (VSF-OFCDM)

- the OVSF code tree has height K. A channelization code of level is denoted as $C_{sf_k,i}$ $C_{256,250} = C^T_{128,125} \cdot C^F_{2,2}$

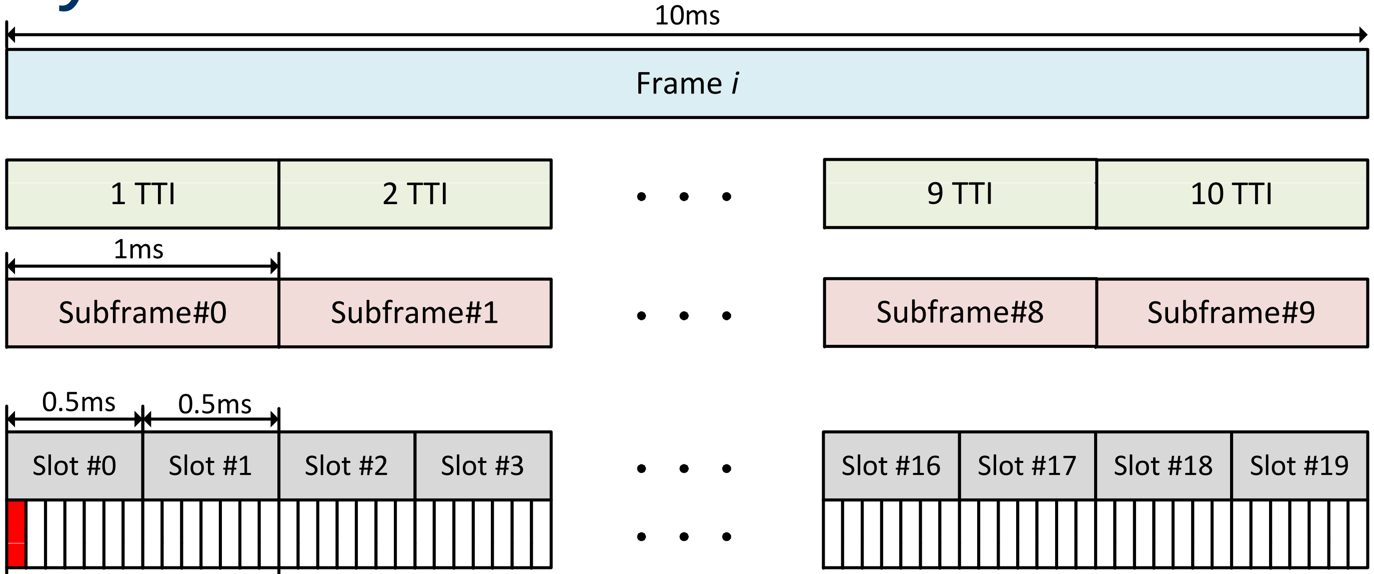
spread in the time and frequency domains

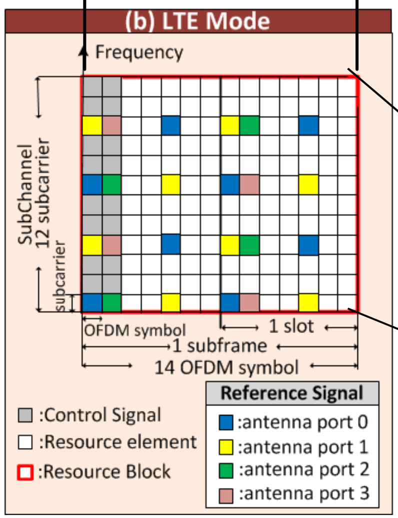


UMTS System (VSF-OFCDM) (cont.)

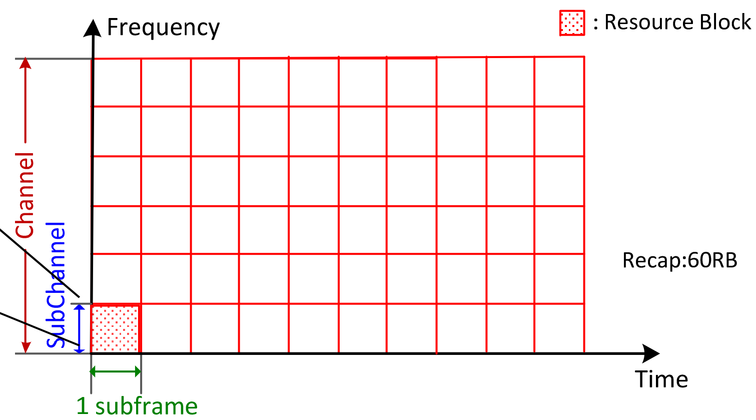


LTE System





Channel Bandwidth (MHz)	1.4	3	5	10	15	20
Number of RBs within a subframe time	6	15	25	50	75	100

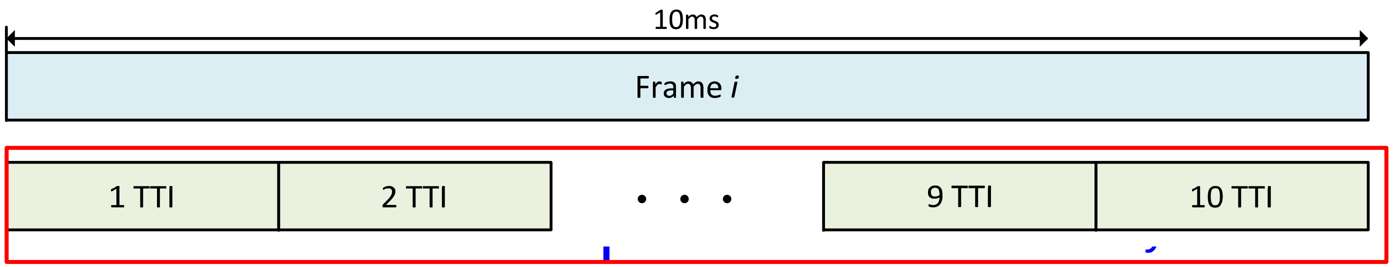


LTE System (cont.)

Modulation		Antenna	Bit/symbol	1 RB (Kbps)
QPSK	½	Single antenna	1	144
16QAM	½	Single antenna	2	288
16QAM	¾	Single antenna	3	432
64QAM	1	Single antenna	6	864
64QAM	1	2 x 2 MIMO	12	1728
64QAM	1	4 x 4 MIMO	24	3456

ls

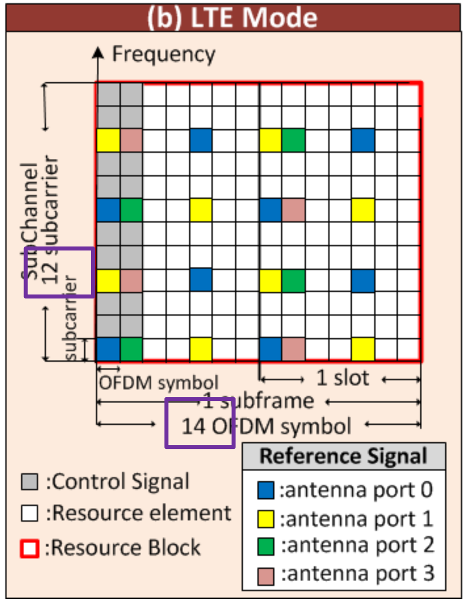
Total of dat



number of subcarriers

example

Channel Bandwidth (MHz)	1.4	3	5	10	15	20
Number of RBs within a subframe time	6	15	25	50	75	100



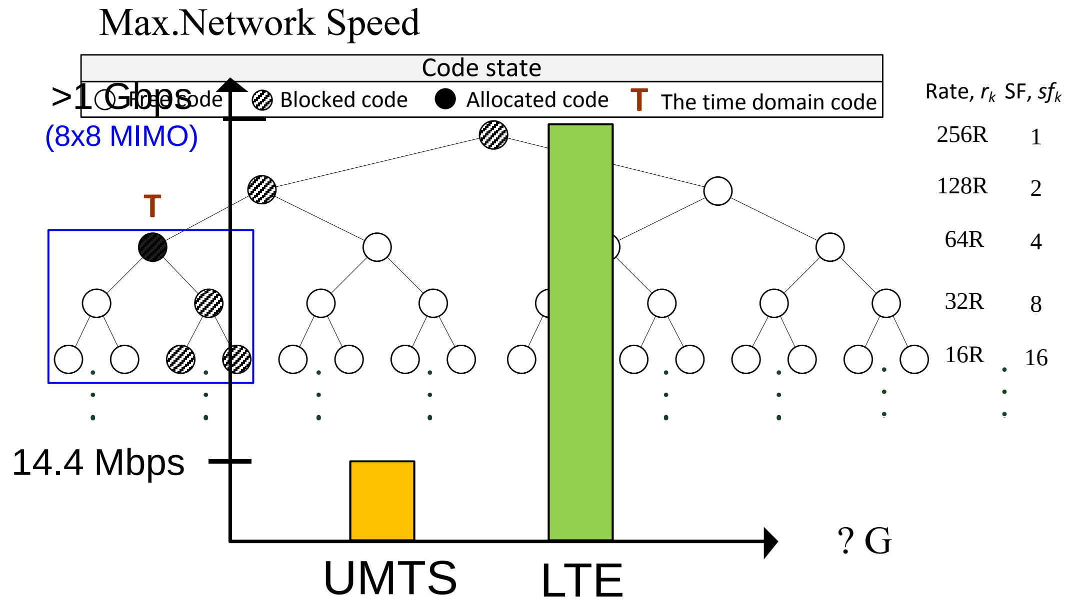
R_{DL}

$$1RB = 1(RBs)\cdot 12(subcarriers)\cdot 12(symbols)$$
$$\cdot 24(bits/symbol)\cdot 10(TTIs)\cdot 100$$
$$= 3456Kbps,$$

Issue

- **UMTS system**
  - high packet loss rate.
  - high bandwidth waste rate,
  - not consider user area
- **LTE system**
  - Suffers from unbalanced loads and unfair bandwidth utilization.

Single Code
Request：33R
Waste rate：31R



Related Works

- The code allocation algorithms of the VSF-OFCDM scheme can be classified into three types:
  - the allocations of 2D-based single code [13], [14]
    - 2D-single
  - single code [15]-[21]
    - CF
    - LM

Disadvantage

they suffer from wasting the bandwidth when the request data rate does not meet the data rate of a channelization code of the VSF-OFCDM code .

  - multicode [22]-[29], [38]
    - 3GPP_MS
    - AVCS

Disadvantage

the most important process in the multicode allocation is to partition the required data rate of a new incoming connection according to the residual radio resources

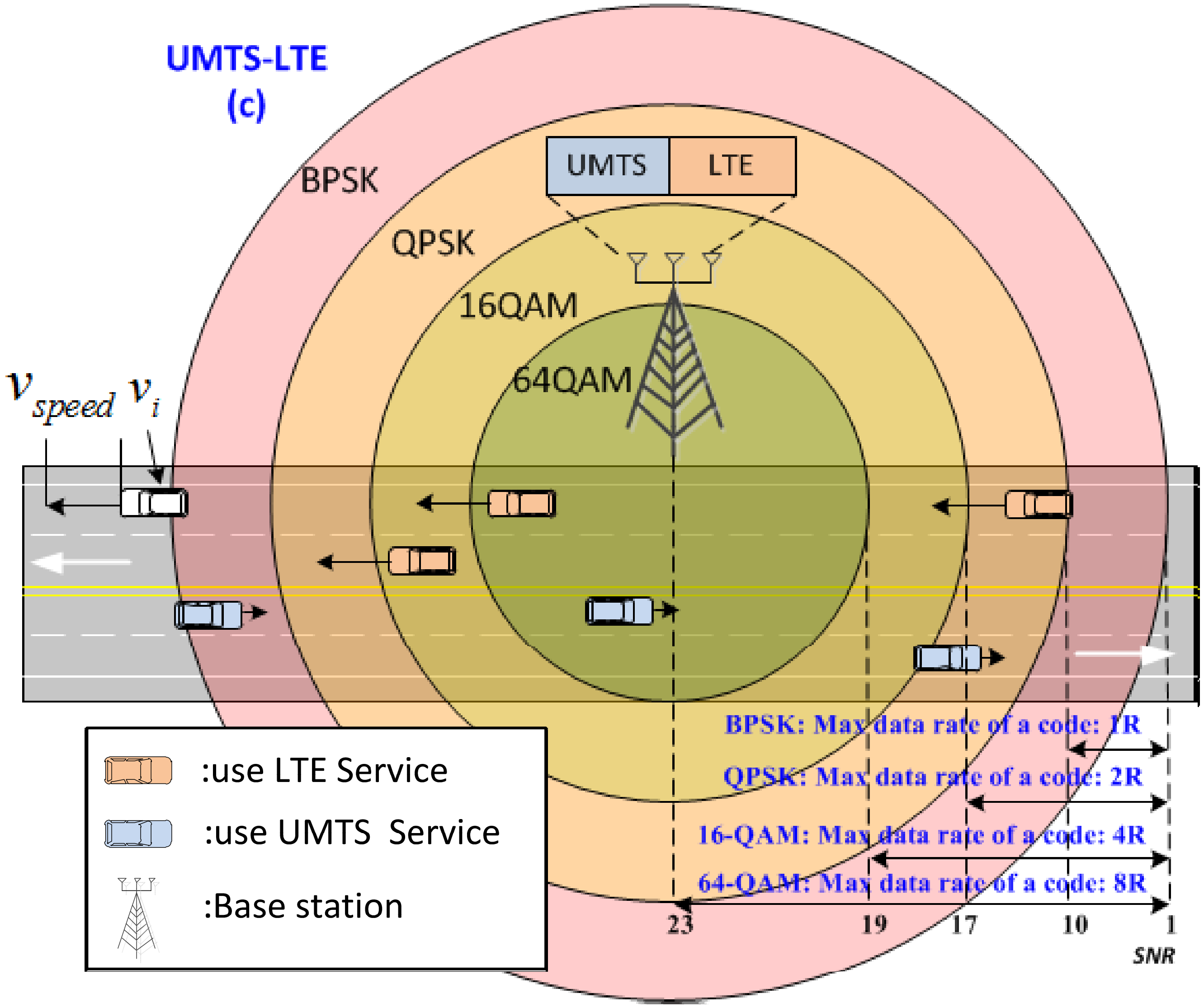
  - **Disadvantage**
    - not consider user locale.

Motivations & Goals

- **Motivations**
  - This paper proposes an Adaptive Radio-resource Allocation (ARA)
    - the LTE and UMTS interfaces according to the conditions of a mobile node

- **Goals**
  - An efficient radio resource management required
    - maximize reward
    - balance loads
    - minimize bandwidth waste rate

Network model



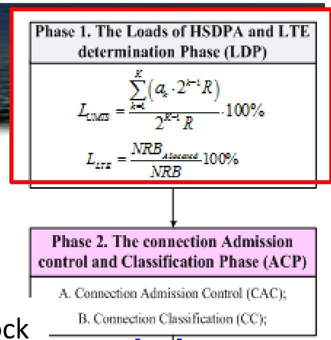
ARA Approach

- This paper proposes an Adaptive Radio-resource Allocation (**ARA**) approach for managing radio resource in the next generation high-speed cellular communication. The ARA approach consists of three main phases:
  - Phase 1: the Loads of UMTS and LTE Determination Phase (LDP)
  - Phase 2: the connection Admission control and Classification Phase (ACP)

    Goal: maximize the interface reward and to minimize the interface overhead
  - Phase 3: the Dynamic resource Allocations of UMTS and LTE Phase (DAP)

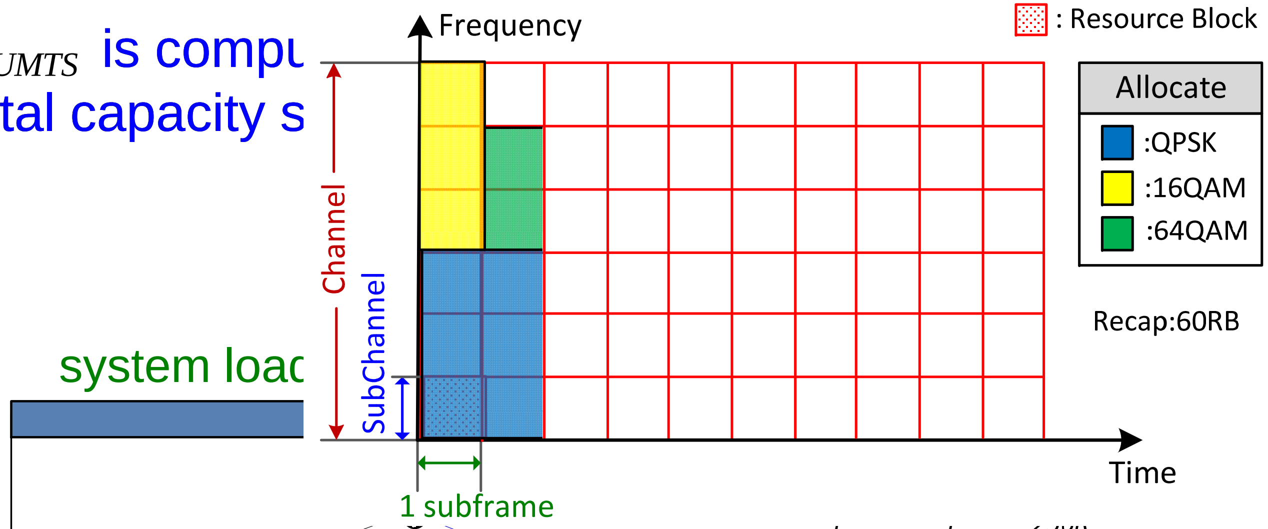
    Goal: guaranteeing the QoS and achieving BS's loads balancing

Phase 1: the Loads of UMTS and LTE Determination Phase (LDP)



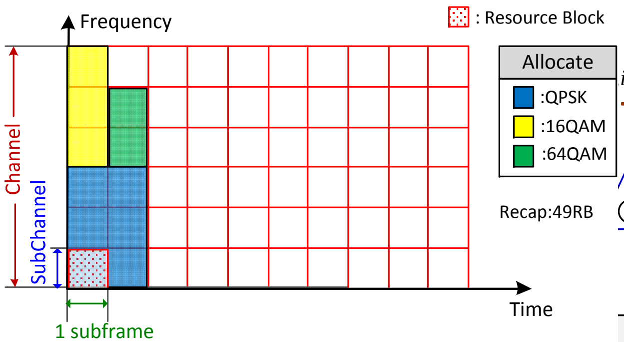
- L_{UMTS} is compu… and the total capacity s… tal of …nelization codes

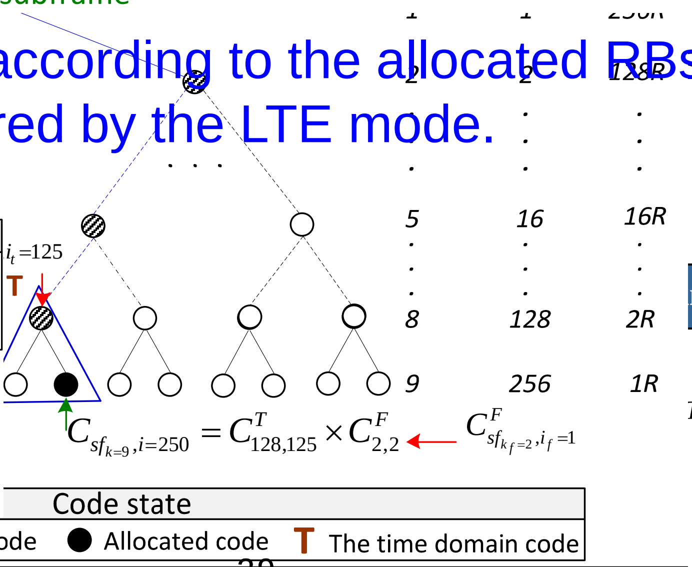
system loac



$\cdot 100\% = 0.390625\%,$

- L_{LTE} is formulated according to the allocated RBs and the total number of RBs offered by the LTE mode.

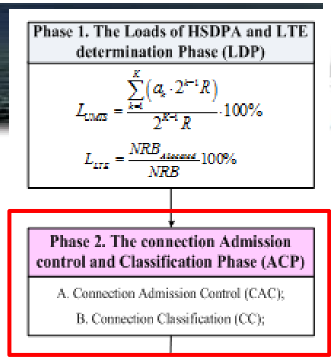




mple

$L_{LTE} = \frac{11}{60} \cdot 100\% = 18.33\%,$

Phase 2: The connection Admission control and Classification Phase (ACP)



- Connection Admission Control (CAC)
  - capacity-based

- Connection Classification (CC)
  - 1) the interface loads of the eNodeB (L_{UMTS} and L_{LTE}).
  - 2) the reward (r_m) of the required traffic class (m),
  - 3) the velocity (v_i) of mobile node i .

Phase 2: The connection Admission control and Classification Phase (ACP) (cont.)

- **First**, if both loads, L_{UMTS} and L_{LTE} , of the eNodeB are all below or all above their individual thresholds,

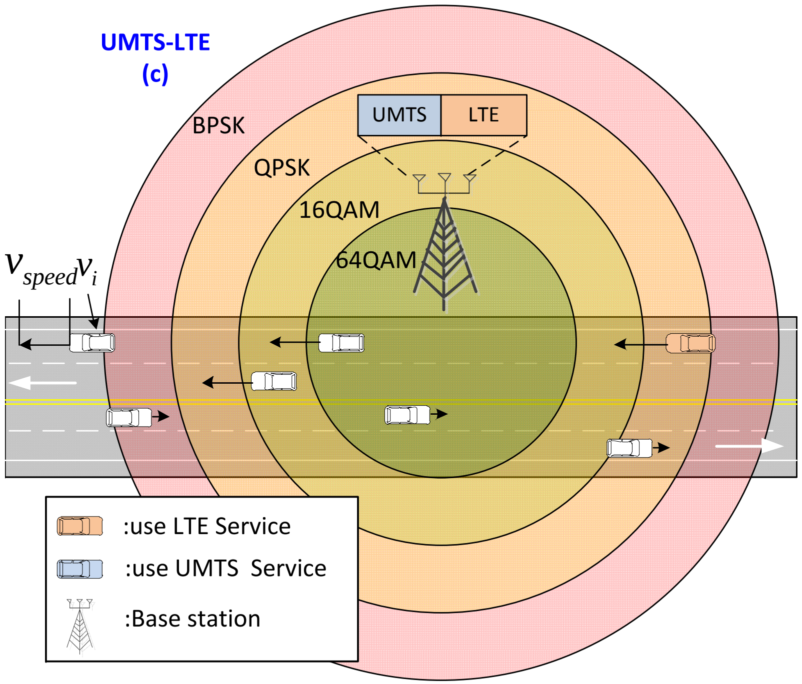
$$L_{LTE} < L_{LTE}^{S} \qquad L_{UMTS} < L_{UMTS}^{S}$$

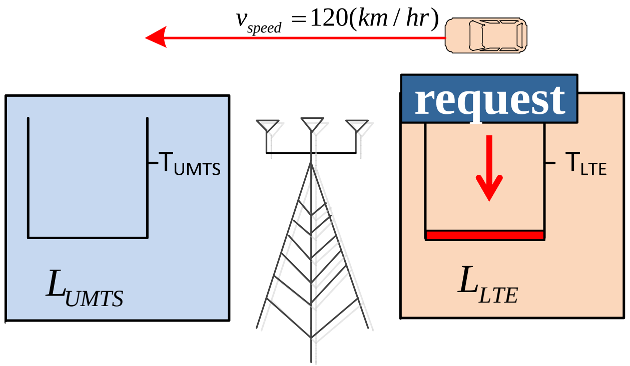
$$L_{LTE} \geq L_{LTE}^{S} \qquad L_{UMTS} \geq L_{UMTS}^{S}$$

- the eNodeB determines the communication interface by using the factors of mobile node velocity (v_i) and the carrying reward (r_m).

Phase 2: The connection Admission control and Classification Phase (ACP) (cont.)

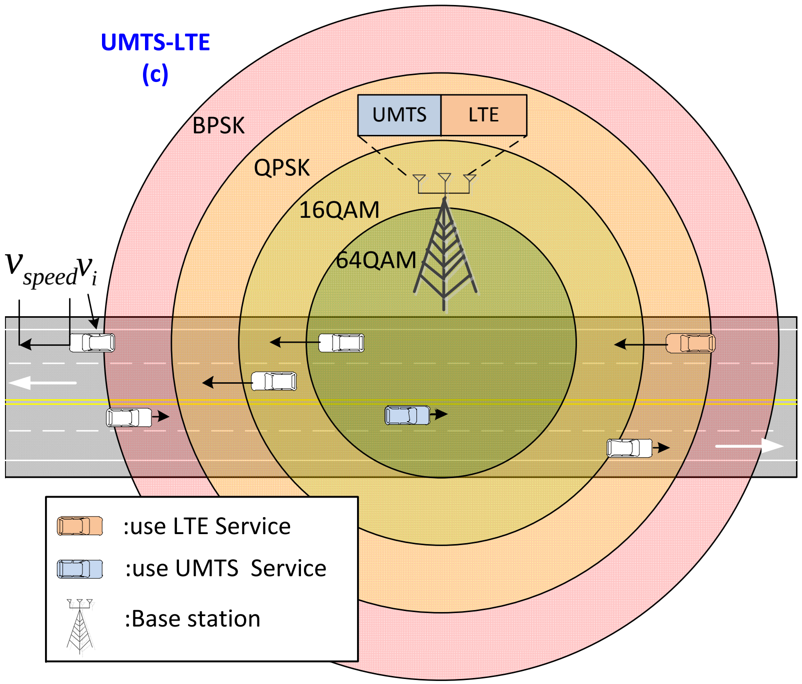
- ACP allocates LTE to a new connection if the carrying reward of using LTE is larger than that of using UMTS (i.e., $r_m^{LTE} > r_m^{UMTS}$), or the mobile node is with a high speed (e.g., $v_i \geq v^S$). Otherwise, ACP allocates UMTS to the new connection.

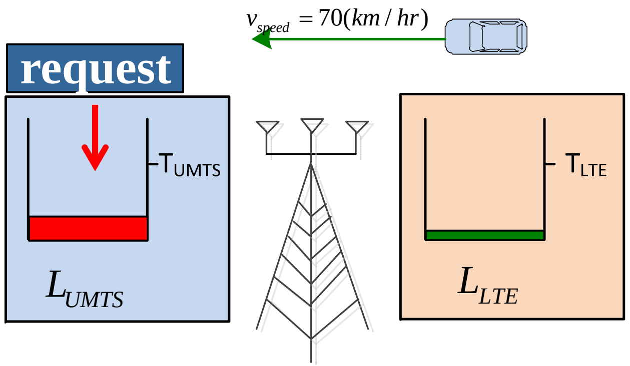




Phase 2: The connection Admission control and Classification Phase (ACP) (cont.)

- ACP allocates LTE to a new connection if the carrying reward of using LTE is larger than that of using UMTS (i.e., $r_m^{LTE} > r_m^{UMTS}$), or the mobile node is with a high speed (e.g., $v_i \geq v^S$). Otherwise, ACP allocates UMTS to the new connection.



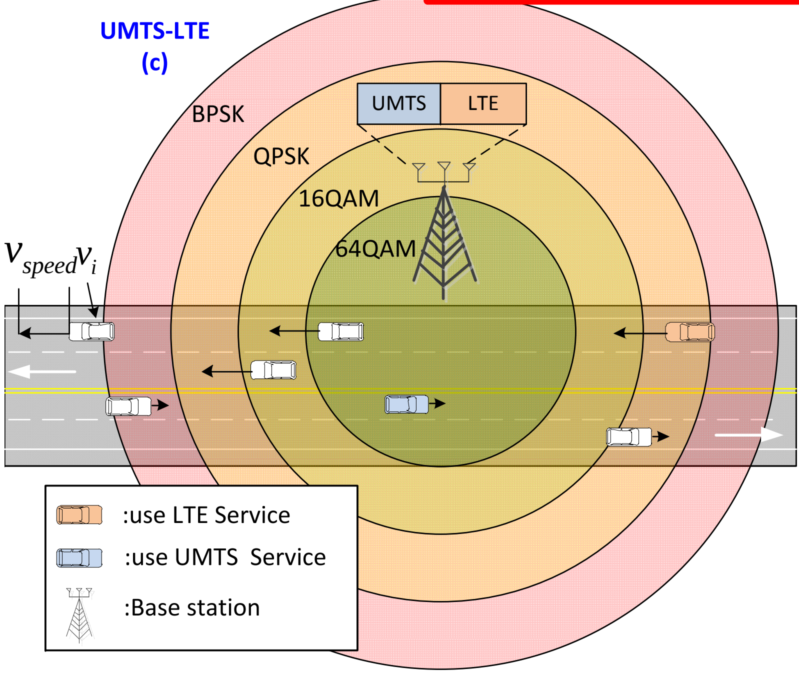


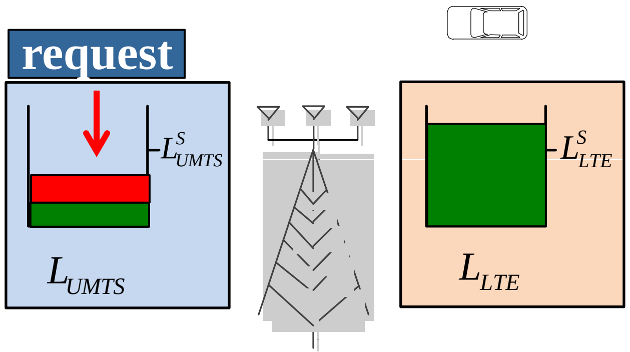
Phase 2: The connection Admission control and Classification Phase (ACP) (cont.)

- **Second**, if only one interface exceeds its threshold, ACP allocates another interface to the new incoming connection.

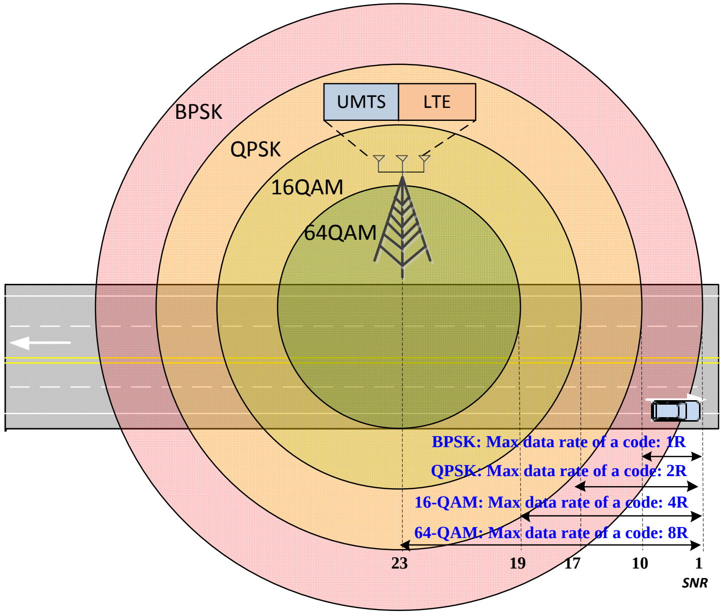
$$L_{LTE} < L_{LTE}^{S} \qquad L_{UMTS} \geq L_{UMTS}^{S}$$

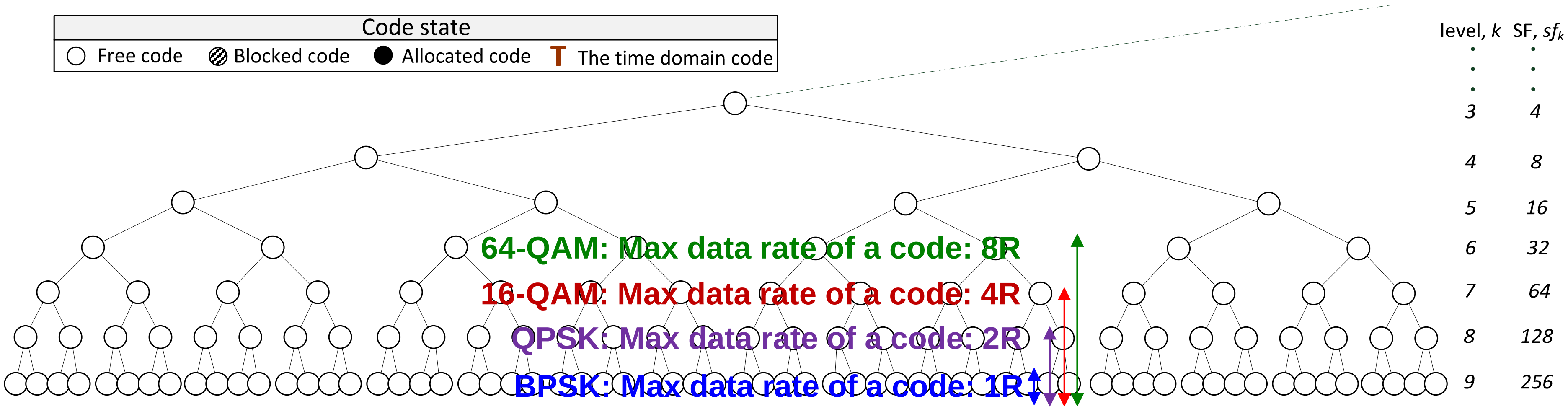
$$L_{LTE} \geq L_{LTE}^{S} \qquad L_{UMTS} < L_{UMTS}^{S}$$





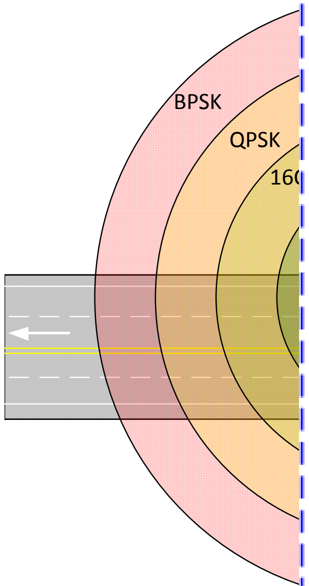
Phase 3: the Dynamic resource Allocations of UMTS and LTE Phase (DAP)

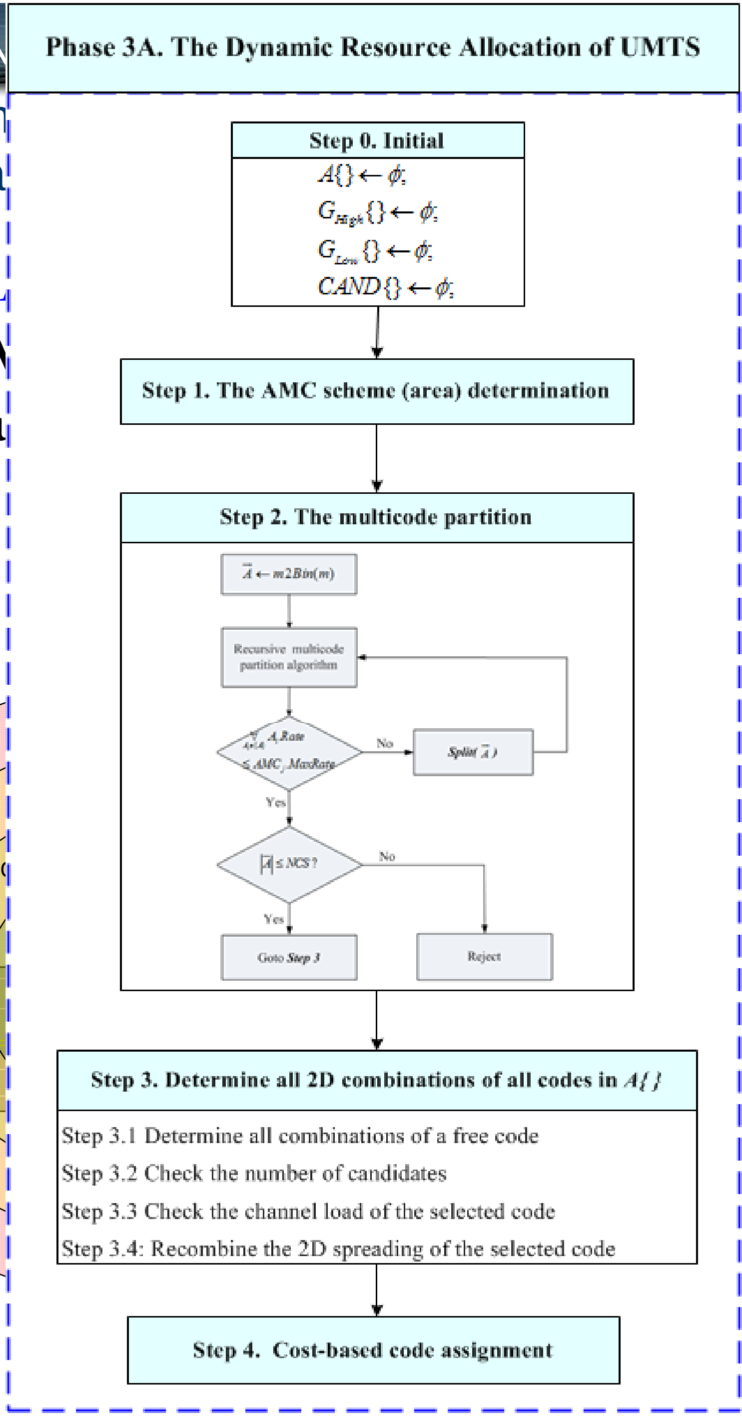


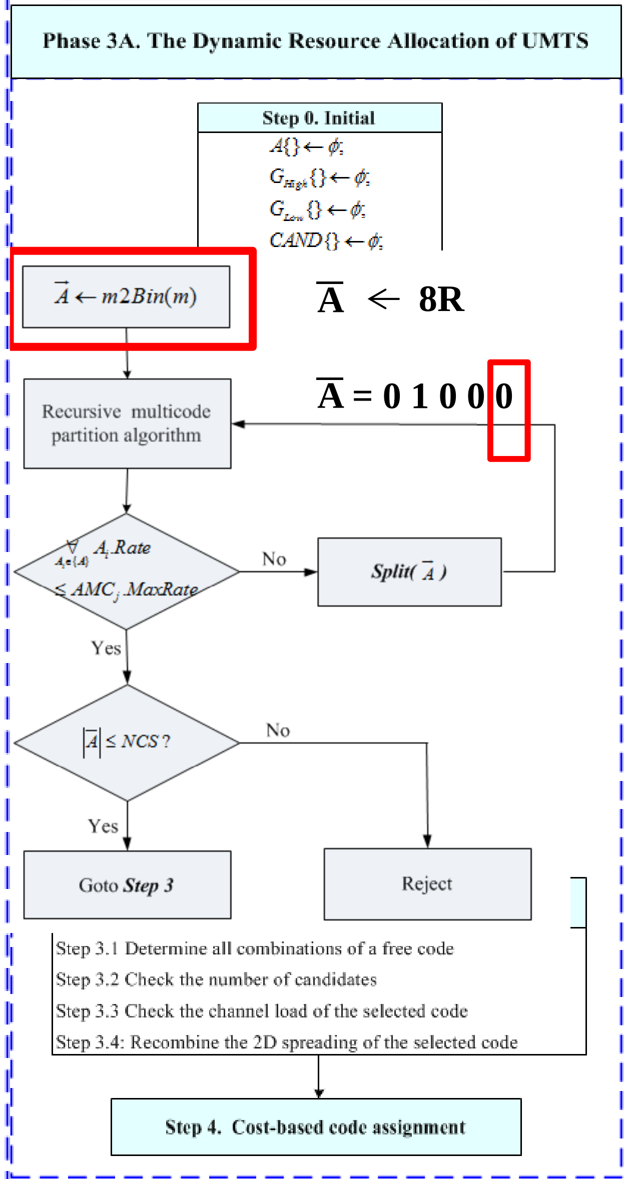


Phase 3: the Dynan
UMTS and LTE Pha

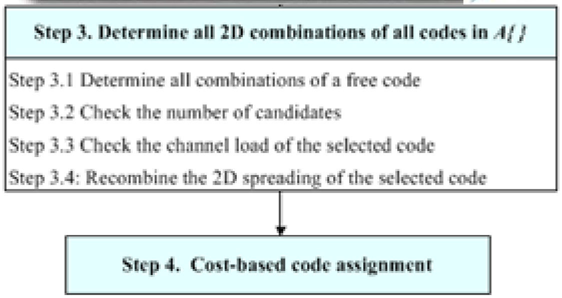
- UMTS Resource
  - **Step 1. The AN**
  - **Step 2. The mu**





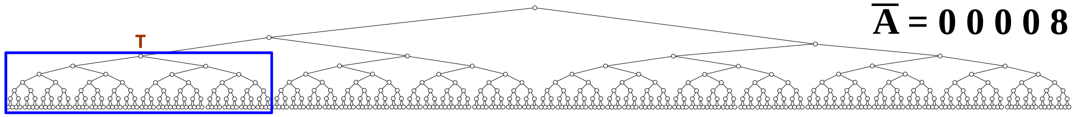


Phase 3: the Dynamic resource Allocations of UMTS and LTE Phase (DAP) (cont.)



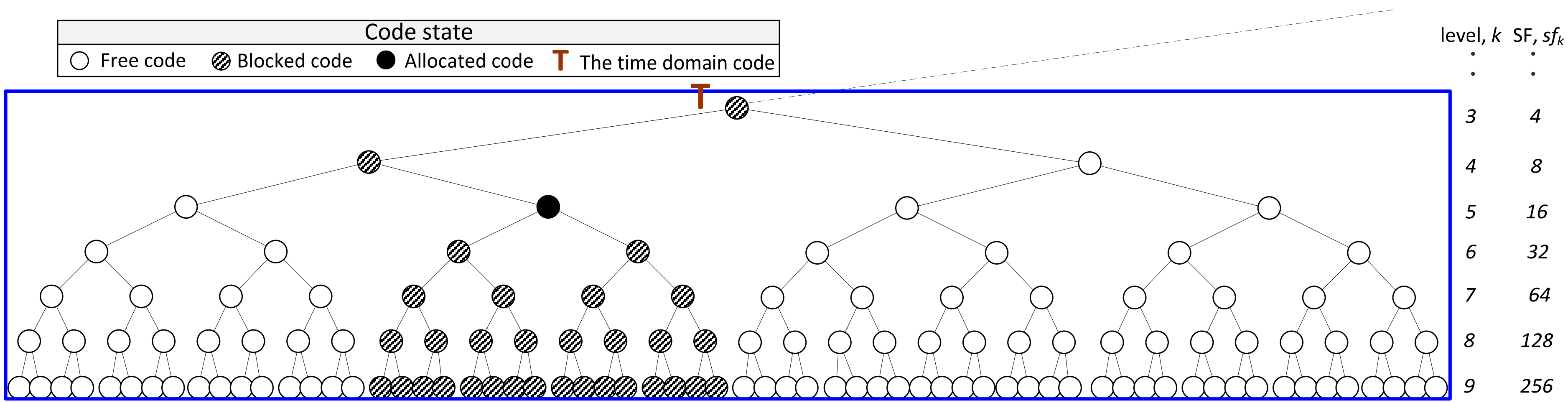
- Consists of four processes:
  1. to determine all combinations of a free channelization code
  2. to check the number of candidate
  3. to check the channel load of the selected code
  4. to recombine the 2D spreading of the selected code.

Step3.1 Determine all combinations of a free channelization code



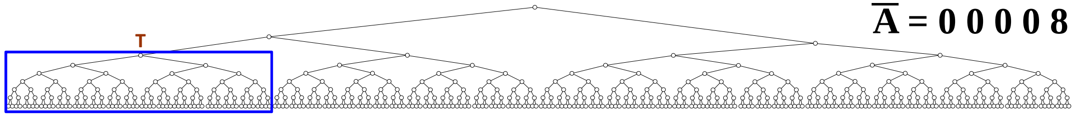
- all 2D spreading combinations of each free-state channelization code that is in $A\{\ \}$

- a lower channel load are then selected as the candidates based on two factors
  - the combination with the largest frequency spreading factor
  - the time domain code should not have a high channel load



Step3.1 Determine all combinations of a free channelization code

$\overline{\mathbf{A}} = 00008$



- After determining all 2D spreading combinations of the free-state channelization codes, the combinations are prioritized into two groups according to the low and high thresholds of a channel load.

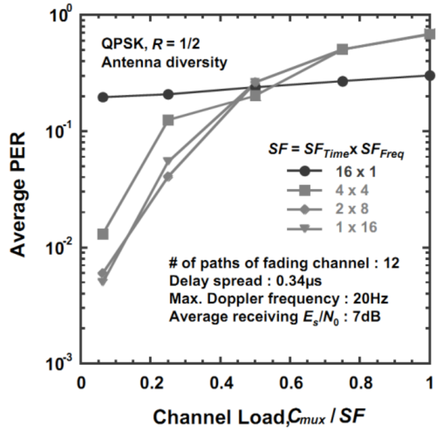
$$CL_{sf_{k_t},i_t} < CL_{sf_{k_t},Low} \qquad G_{High}\{\ \} \leftarrow C_{sf_k,i} = C^T_{sf_{k_t},i_t} \cdot C^F_{sf_{k_f},i_f}$$

0.375

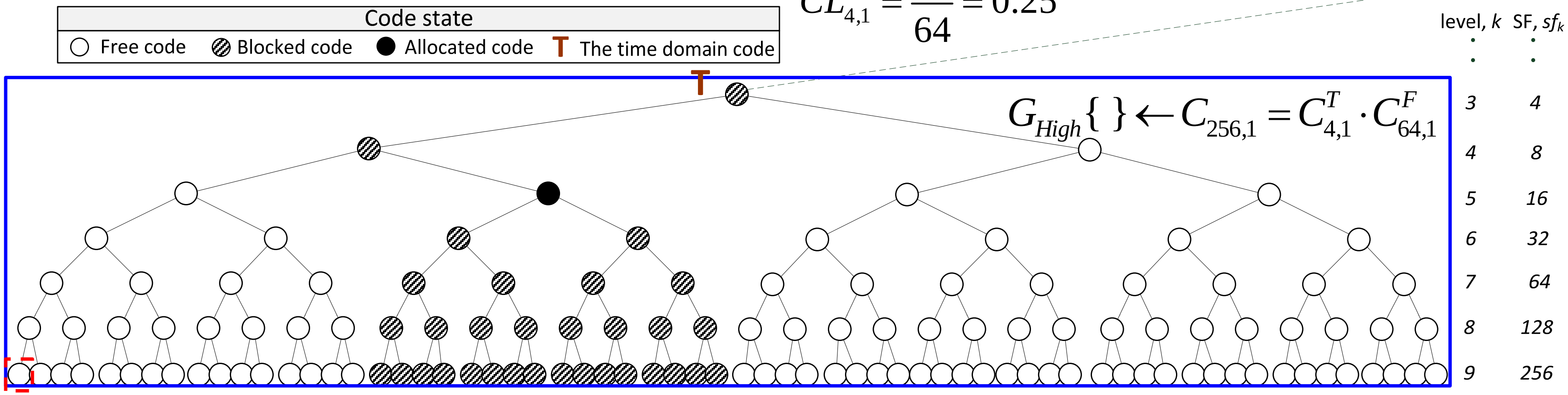
$$CL_{sf_{k_t},Low} \le CL_{sf_{k_t},i_t} < CL_{sf_{k_t},High} \qquad G_{Low}\{\ \} \leftarrow C_{sf_k,i} = C^T_{sf_{k_t},i_t} \cdot C^F_{sf_{k_f},i_f}$$

$$CL_{sf_{k_t},i_t} > CL_{sf_{k_t},High}$$ re-combination

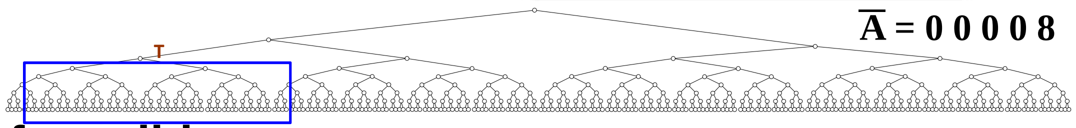
0.75



$$CL_{4,1} = \frac{16}{64} = 0.25$$



$\overline{A} = 0\,0\,0\,0\,8$



Step3.2 Check the number of candidates

- After performing ***Step 2*** and ***Step 3.1***, the number of candidates is at least one. If the size of the candidate set is equal to one, only one optimal 2D spreading exists

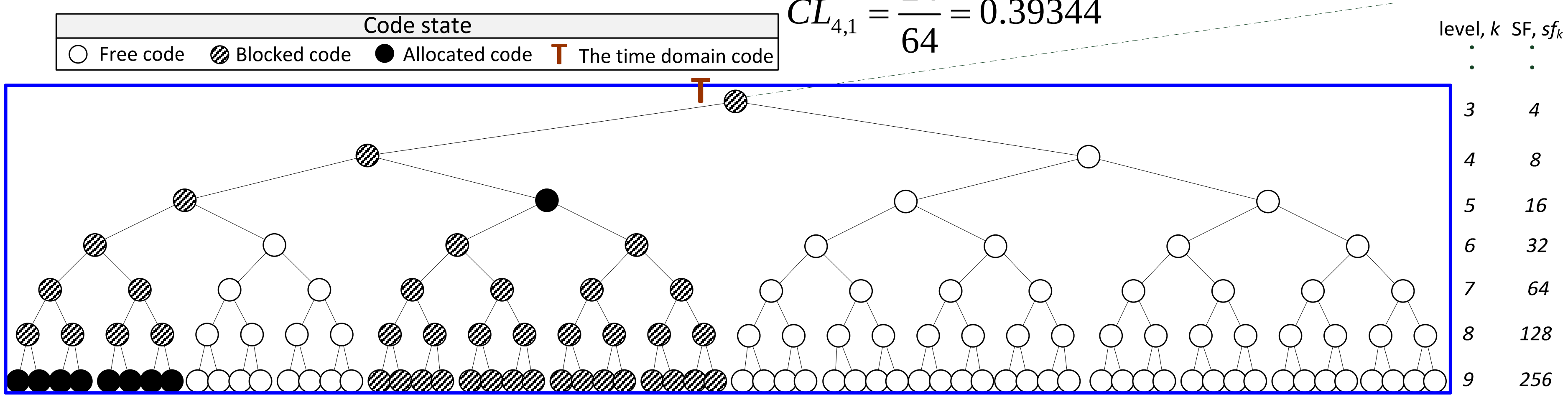
Step3.3 Check the channel load of the selected code

- The system checks whether the channel load of this combination exceeds the high threshold or not

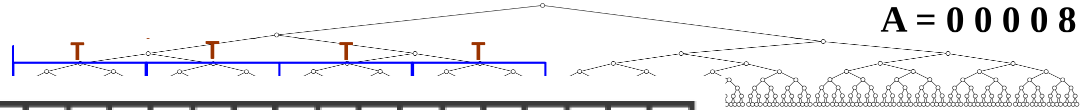
$$CL_{sf_{k_t},i_t} > CL_{sf_{k_t},High}$$

0.75

$$CL_{4,1} = \frac{24}{64} = 0.39344$$

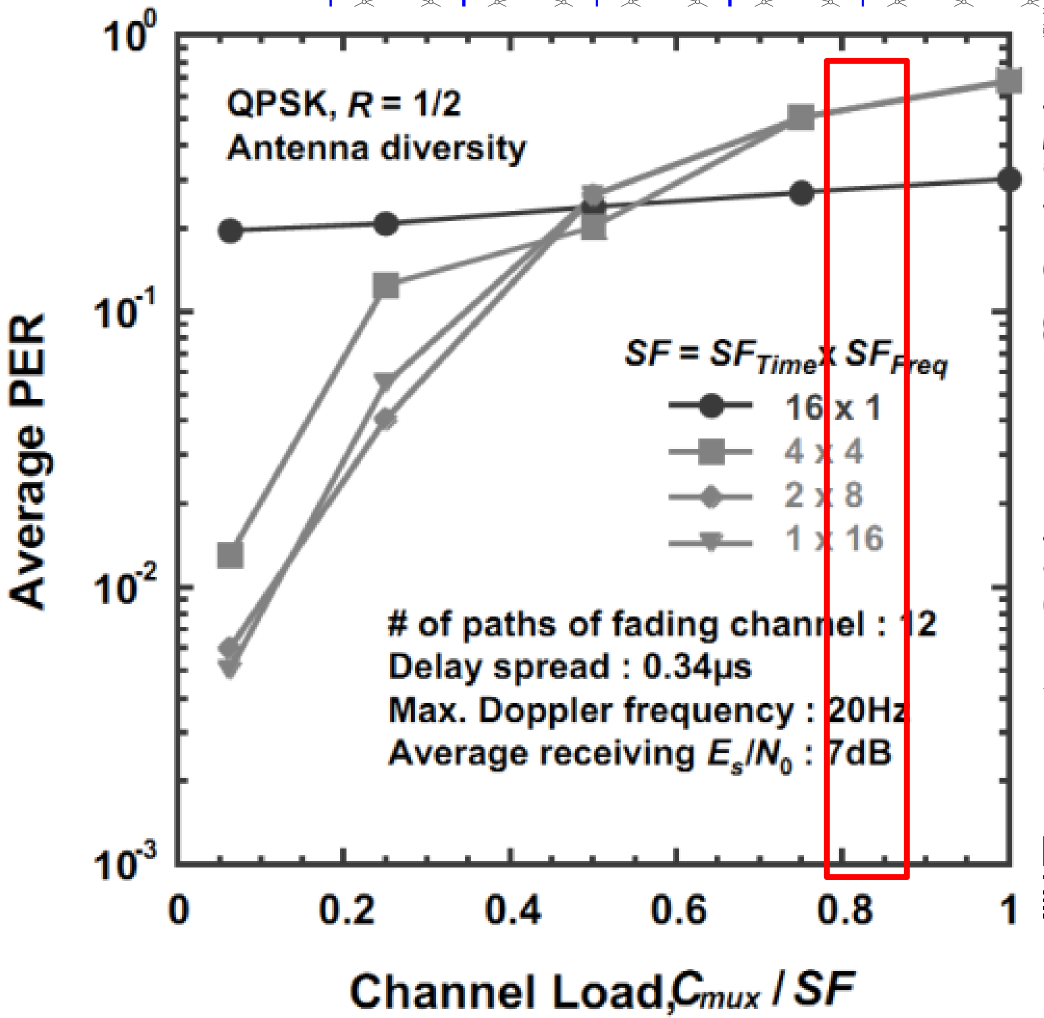


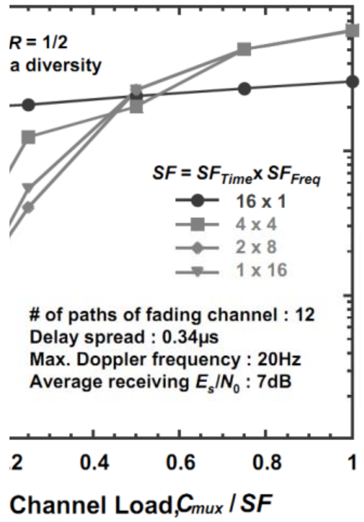
A = 0 0 0 0 8



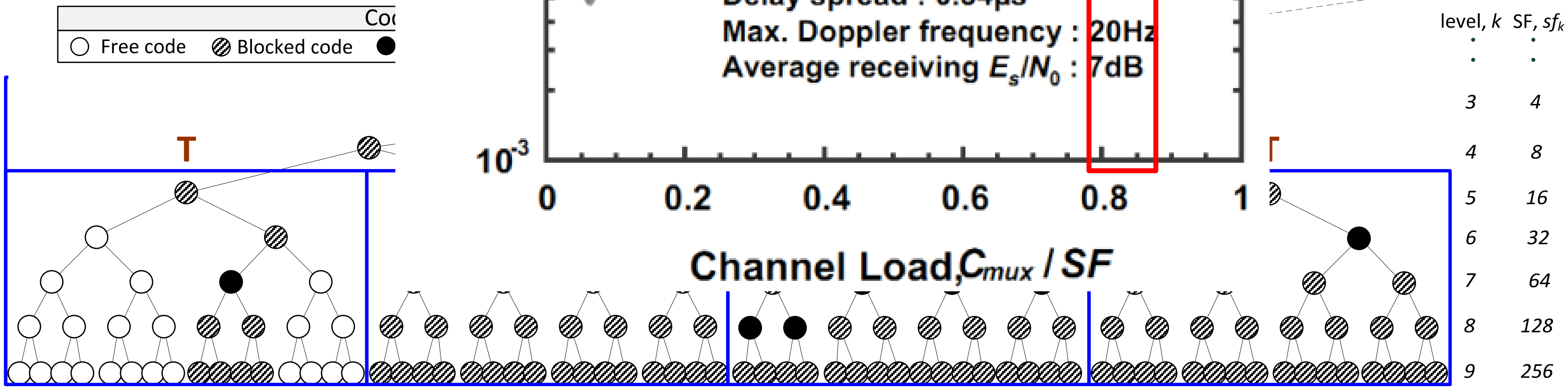
- Step3.4 Rec … code

  CL_s





- Step 4 Cost



Phase 3: the Dynamic resource Allocations of UMTS and LTE Phase (DAP) (cont.)

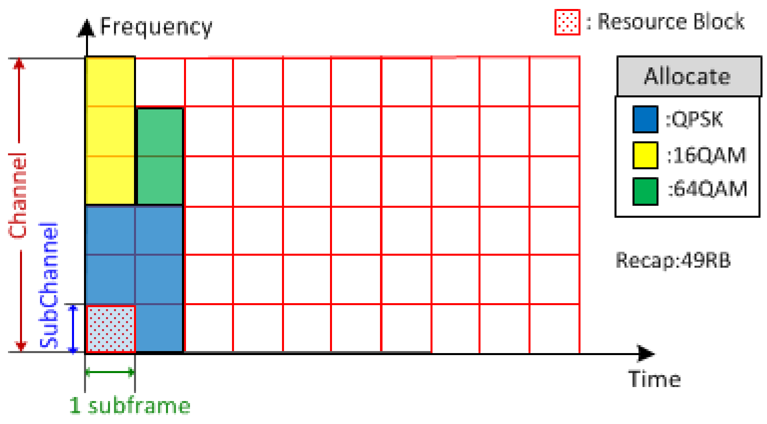
Phase 3B. The Dynamic Resource Allocation of LTE
Step.1 Determine user request and local.
Step.2 Assign user request.

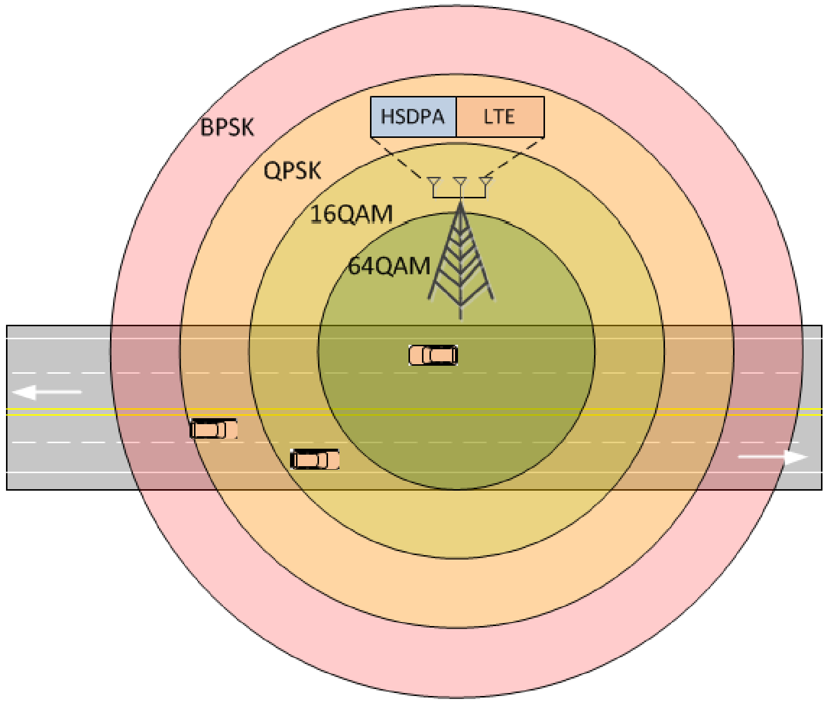
LTE Resource Allocation

- A RB can adopt any type of AMC coding. Thus, the minimum and maximum data rates of a RB are 144 Kbps (QPSK) and 3.456 Mbps (64-QAM).

$$R_{RB} = 12(subcarriers)\cdot 12(symbols)\cdot 1(bits/symbol)\cdot 10(TTs)\cdot 100 = 144Kbps.$$

- **example**
  - If a UE requires a 12R bandwidth. Then, the eNodeB can allocate 6 RBs with the QPSK coding, 3 RBs with the 16QAM coding, or 2 RBs with 64QAM coding.





Phase 3: the Dynamic resource Allocations of UMTS and LTE Phase (DAP) (cont.)

- the number of RBs needed for a 16R required rate under different AMC modulations can be determined

$$N_{R_{RB}} = \frac{16 \cdot 144kbps}{864kbps} = \frac{2304kbps}{864kbps} = 2.666 \approx 3RBs$$

- Although the eNodeB can allocate sufficient number of RBs for a 16R connection located in QPSK, it brings too much overhead and degrades the system reward and interface utilization

Modulation		Antenna	Bit/symbol	1 RB (Kbps)
QPSK	½	Single antenna	1	144
16QAM	½	Single antenna	2	288
16QAM	¾	Single antenna	3	432
64QAM	1	Single antenna	6	864
64QAM	1	2 x 2 MIMO	12	1728
64QAM	1	4 x 4 MIMO	24	3456

- For instance, we assume that MRB is set to 4. If a vehicle located at the QPSK location requires 16R data rate, the incoming connection will be rejected. The reason is the number of required RBs is 16 that is more than MRB

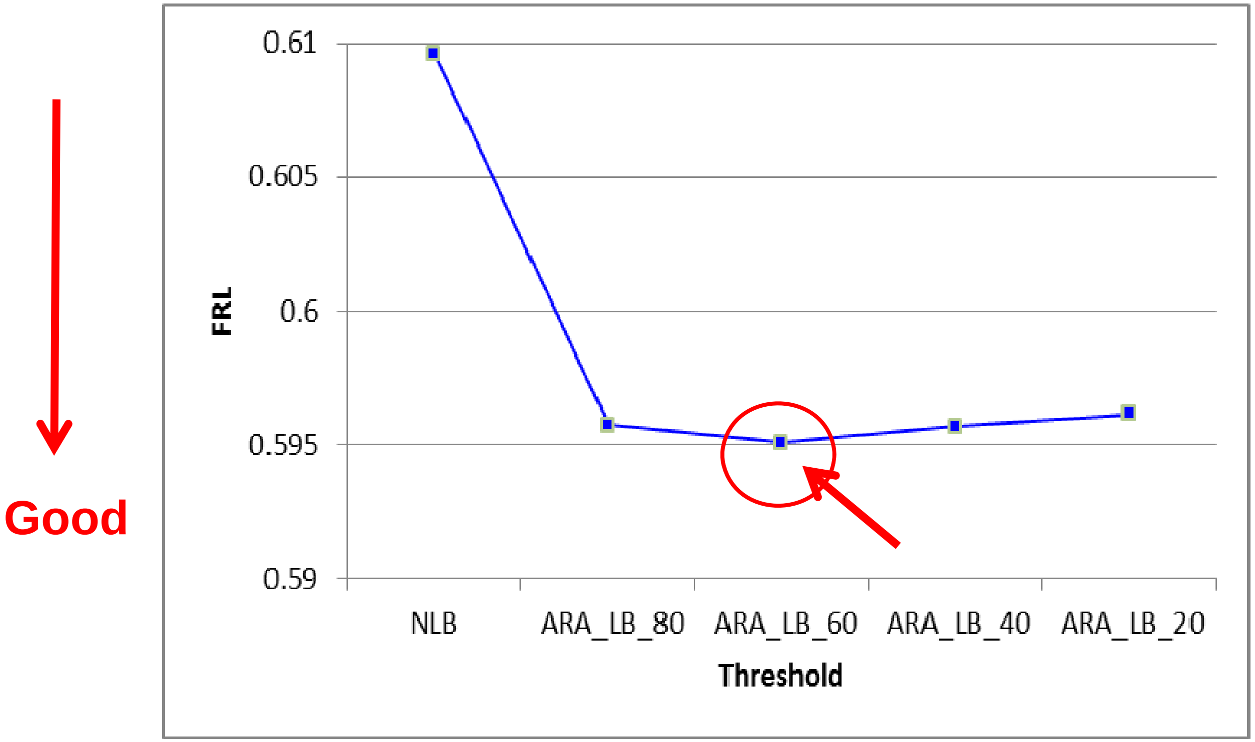
Parameters for simulations

Network Model

Parameters	Values
LTE Capacity [39]	345.6 Mb/s
UMTS Capacity	3.84 Mb/s=(256R)
Number of rake-combiners (NCS)	1(single code), 2~8
Total number of RBs	1000 RBs
AMC coding schemes	QPSK/16QAM(1/2)//64QAM(1)
Channel bandwidth	20 MHz
Mode load threshold (S)	0.6
Spreading factor, sf_k	1 ~ 256
Time domain spreading factor, sf_{k_t}	4, 16
Frequency domain spreading factor, sf_{k_f}	1~64
Low threshold of the channel load in time domain, $CL^T_{sf_{k_t},Low}$	0.375, if $sf_k > sf_{k_t}$, 0.5, if $sf_k = sf_{k_t}$
High threshold of the channel load in time domain, $CL^T_{sf_{k_t},High}$	0.75, if $sf_k > sf_{k_t}$, ∞, if $sf_k = sf_{k_t}$
The weight parameter of the exponential smoothing model, α	0.85
Traffic Model	
Number of traffic classes (m)	16 (i.e., 1~16)
Average arrival rate (λ)	2 ~ 22
Arrival rate (λ_m)	λ / m
Bandwidth request (b_m)	$m \cdot 1R$
Average holding time (μ)	1

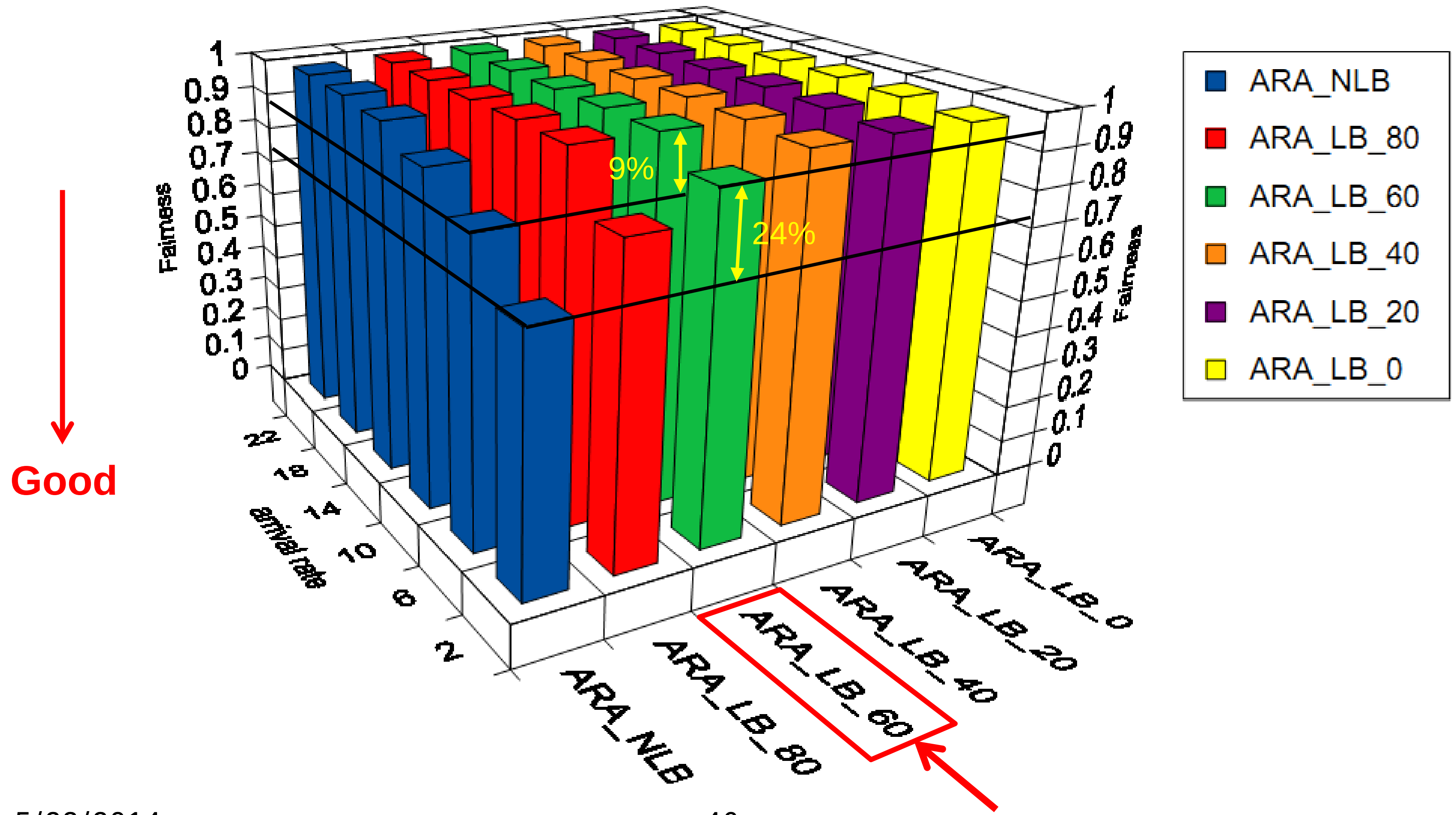
Numerical results

- **A. The analysis of optimal threshold of load balancing**
  - FRL under different thresholds of load balancing



Numerical results (cont.)

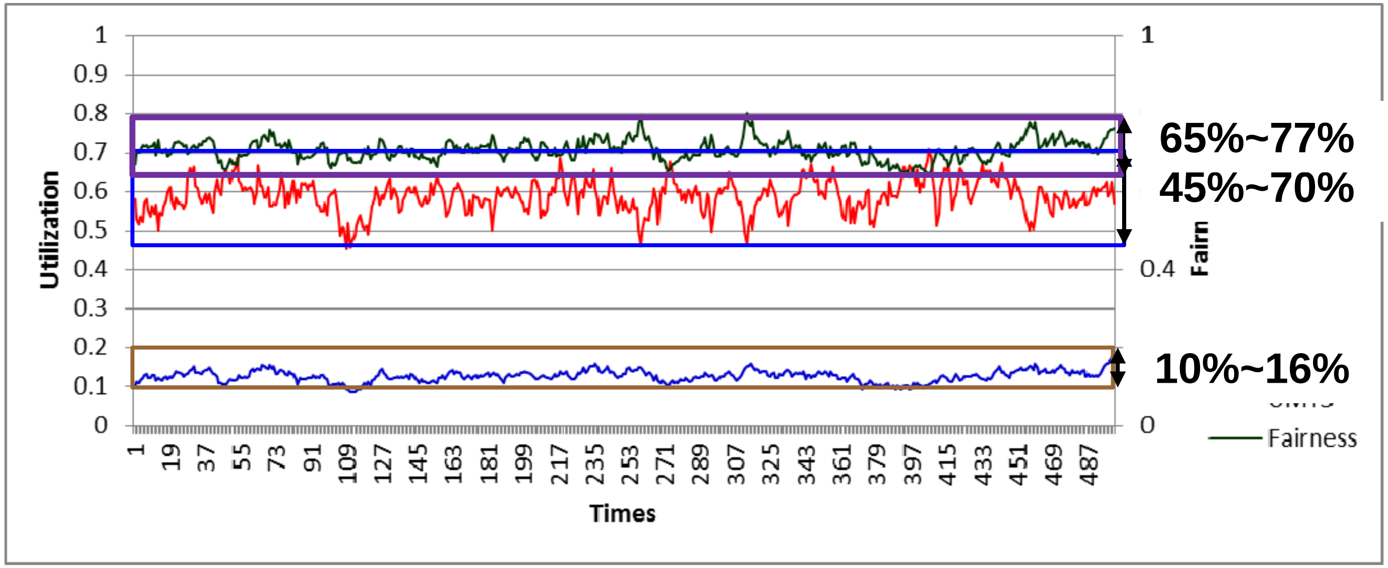
- **Fairness of different thresholds of the load balancing under various arrival rates**



Numerical results (cont.)

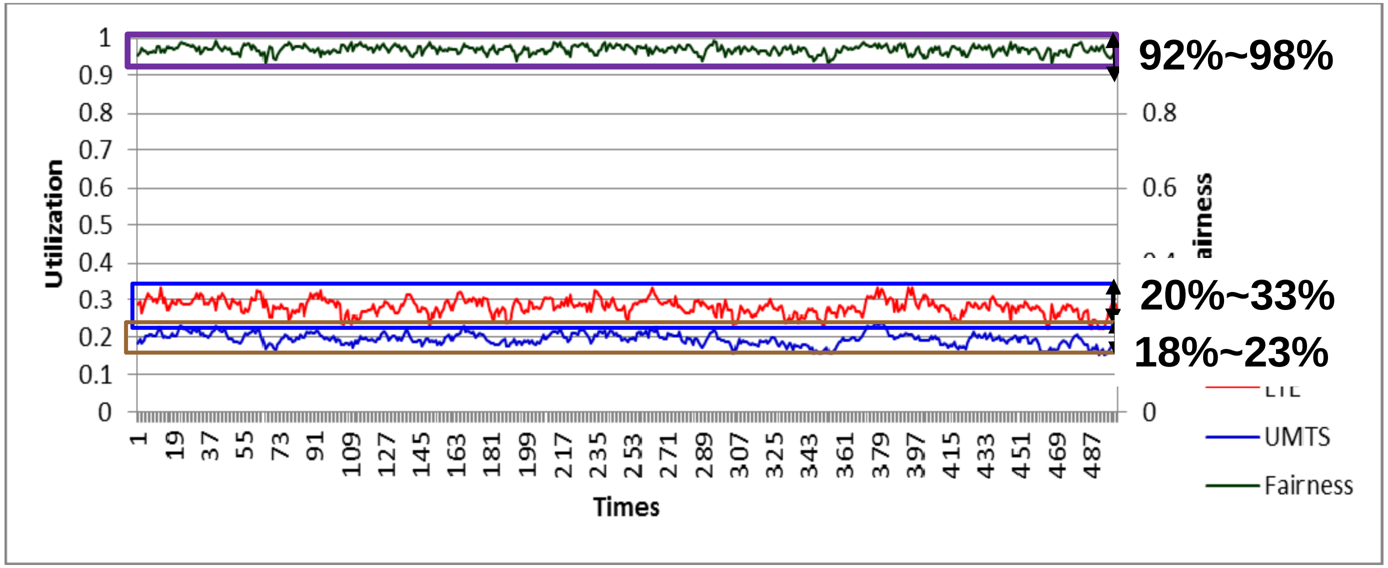
ARA_NLB

Good



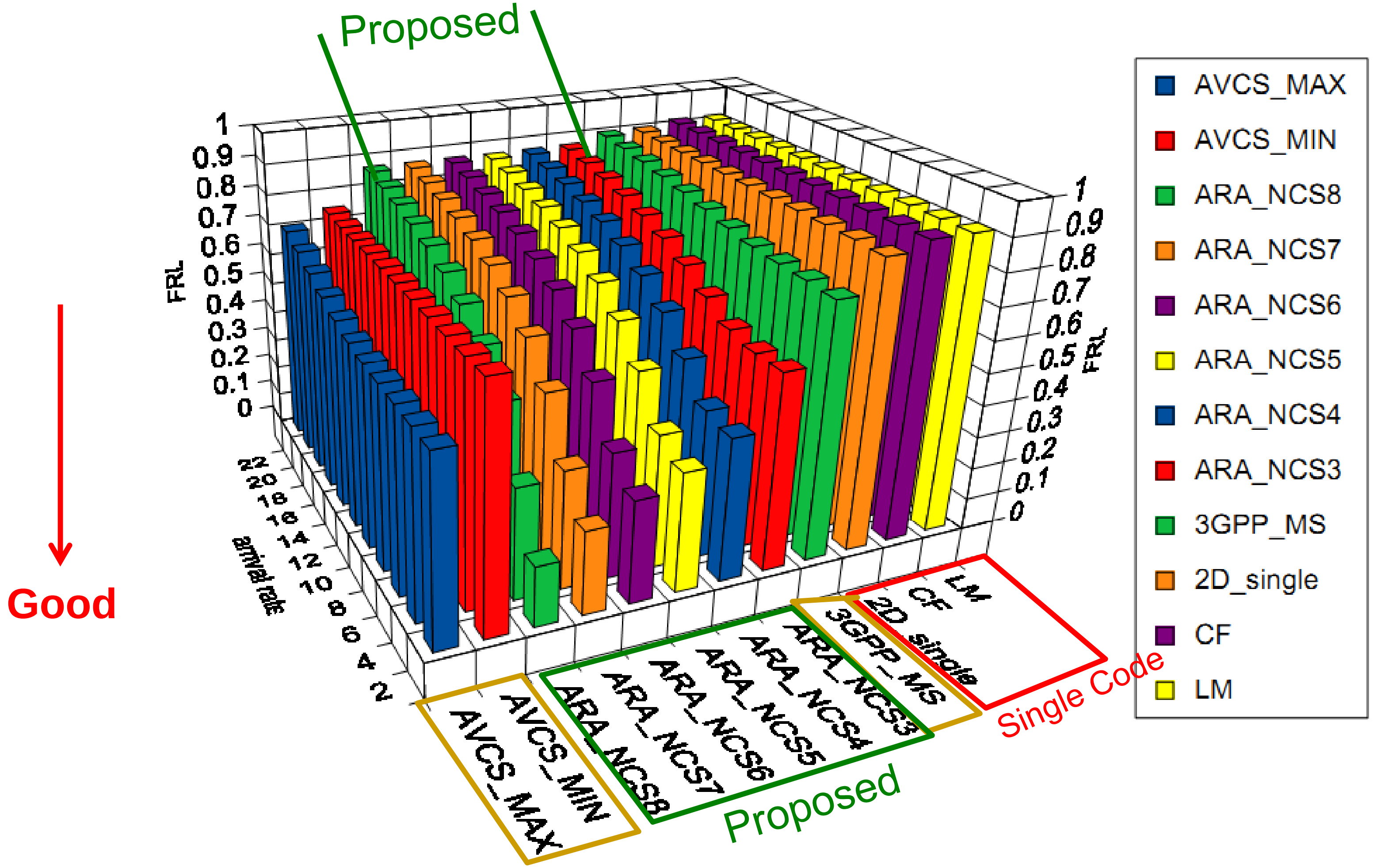
ARA_LB_0

Good



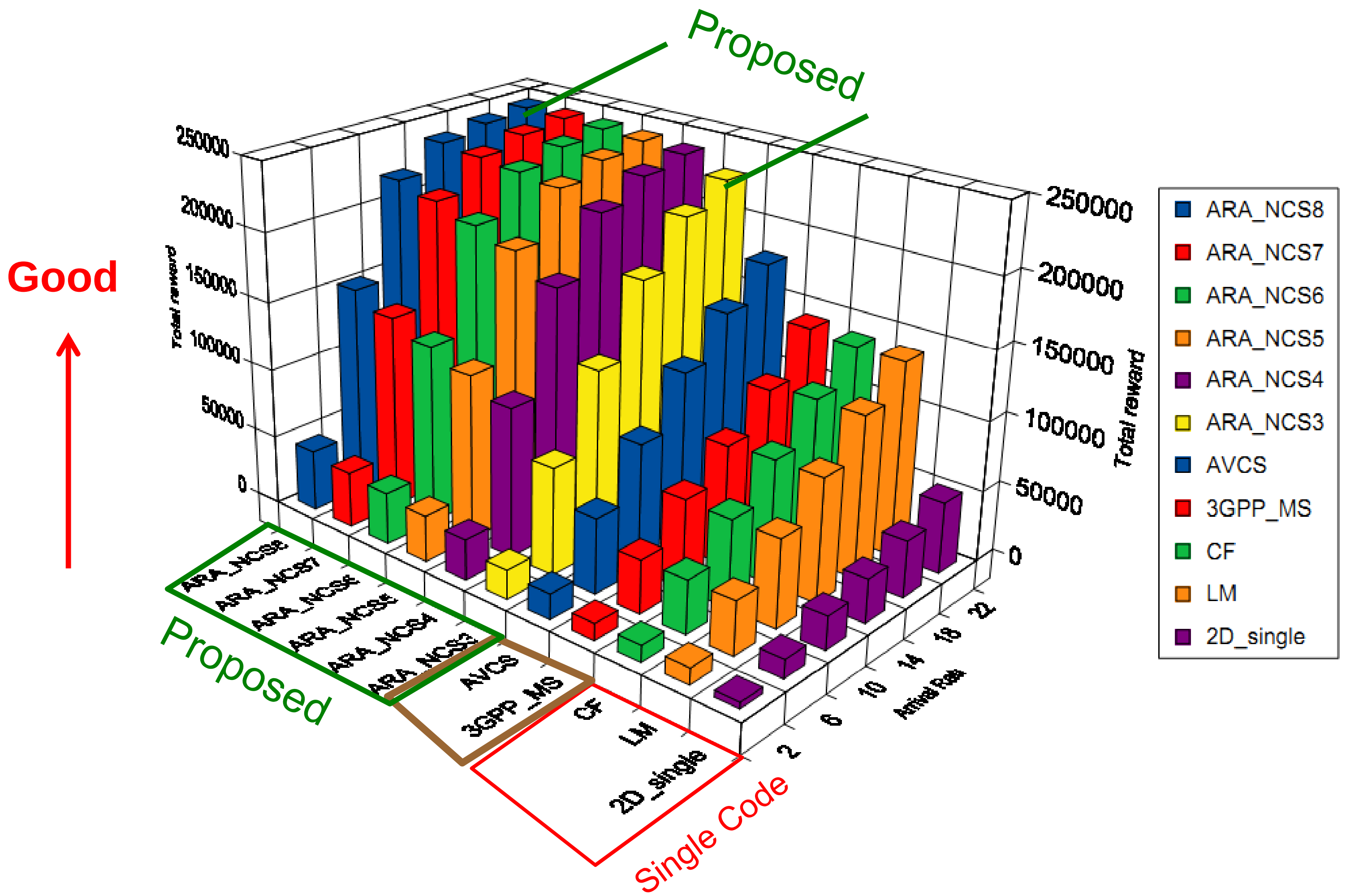
Numerical results (cont.)

- B. Performance evaluations among different compared approaches



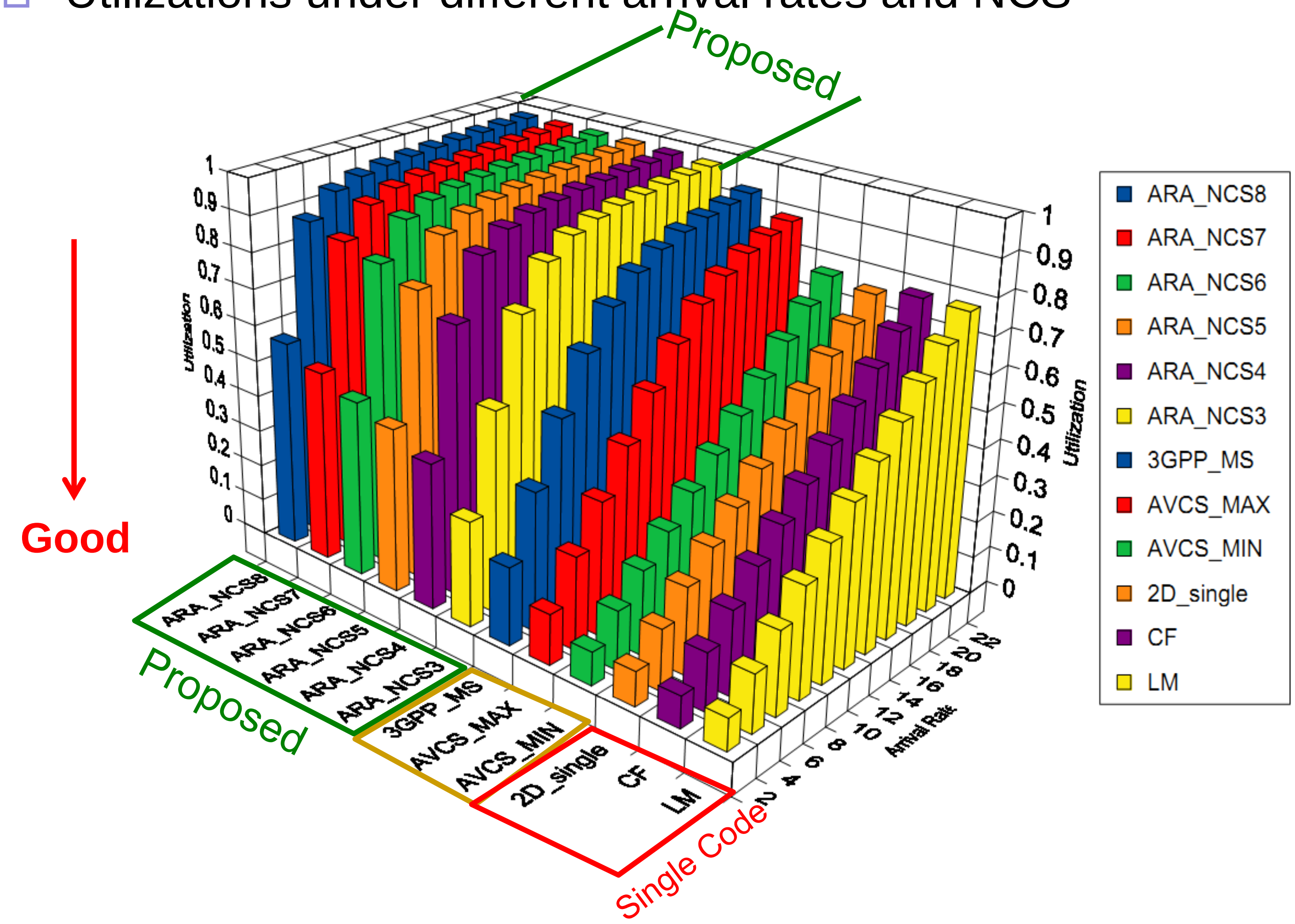
Numerical results (cont.)

- Total reward under different arrival rates



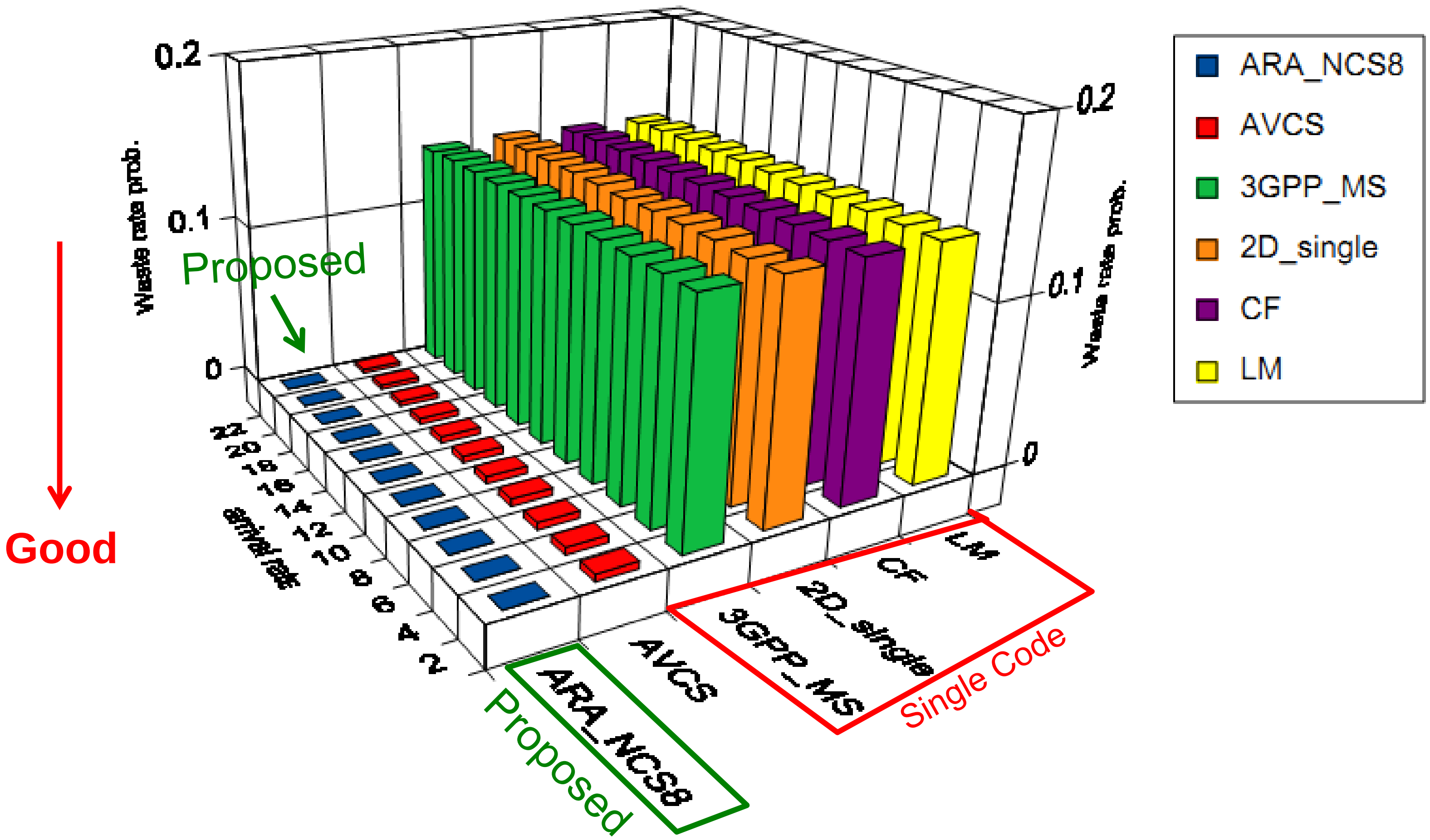
Numerical results (cont.)

- Utilizations under different arrival rates and NCS



Numerical results (cont.)

- Bandwidth waste rate of all compared approaches under different arrival rates



Conclusions

- the ARA approach can maximize reward, balance loads, and minimize bandwidth waste rate in 4G UMTS and LTE Communications.

- this paper thus proposed the **A**daptive **R**adio-resource **A**llocation (**ARA**) approach that consists of three phases:
  - 1) the **L**oads of UMTS and LTE **D**etermination **P**hase (**LDP**),
  - 2) the connection **A**dmission control and **C**lassification **P**hase (**ACP**)
  - 3) the **D**ynamic resource **A**llocations of UMTS and LTE **P**hase (**DAP**).

- The first two phases: LDP and ACP efficiently achieve the load balancing feature.

Q & A

Thanks