MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE
IEEE COMMUNICATIONS SOCIETY
http://www.comsoc.org/~mmc

E-LETTER

Vol. 9, No. 6, November 2014

IEEE COMMUNICATIONS SOCIETY

CONTENTS

Message from MMTC Chair

Dear MMTC colleagues:

November is always a special month! It is the month that reminds us that we are approaching the end of the year, fall season with its colorful trees and cold breeze, and that a GLOBECOM conference is coming soon! If you plan to attend this year conference in Austin, then I hope to meet you there!

It is a great honor to continue to be part of the IEEE ComSoc MMTC team. As the elected North America Vice-Chair for 2014-2016, I am excited to get the chance to work with MMTC leadership team to continue our mission of providing great value to our colleagues, from both academia and industry, focusing on multimedia communication technologies.

In the September issue of the E-Letter, my colleague Liang Zhou, the Asia Vice-Chair of MMTC has emphasized the importance of the Interest Groups (IGs) as the catalyst for the success of our technical committee to effectively reaching out to the community of multimedia communication researchers and engineers.

One of my main goals in the coming months is to focus on increasing the level of participation of colleagues from industry in all activities of MMTC, with emphasis on North America. This includes submitting papers to E-Letter and other ComSoc conferences/publications, sponsorship of events, and sharing industry experience on developing innovative multimedia communication products. This effort will be coordinated with the IGs chairs and the MMTC leadership team.

We always welcome your feedback on how we can improve the operation of MMTC. Your active participation will strengthen the reach and impact of the technical committee.

I look forward to another fruitful year in 2015!

Khaled El-Maleh
North America Vice-Chair, IEEE ComSoc Multimedia Communications Technical Committee

EMERGING TOPICS: SPECIAL ISSUE ON MASSIVE MIMO FOR 5G CELLULAR SYSTEMS

Guest Editors: Lingjia Liu[1] and Guosen Yue[2]

[1]University of Kansas (KU), Lawrence, KS, USA, [2]Broadcom Corporation, Matawan, NJ, USA

lingjialiu@ittc.ku.edu , guosenyue@gmail.com

Cellular communications have experienced an unprecedented growth-rate of mobile data traffic due to the rapid introduction of various connected mobile devices and excessively data-hungry applications run on those devices. According to the International Telecommunication Union (ITU), the total number of mobile cellular subscriptions is fast approaching the total number of the people living on earth, and is expected to reach 7 billion by the end of 2014, corresponding to a penetration rate of 96%.

To cater for this demand, many advanced physical layer techniques have been developed, e.g., multiple-input-multiple-output (MIMO) with orthogonal frequency division multiplexing (OFDM). However, with linear throughput improvement but the exponential growth on the data traffic, the gap between the demand and supply has been increasingly widened. To solve the problem, the next technology we could resort to is massive MIMO (a.k.a. large-scale MIMO, full-dimension MIMO, or hyper MIMO), which significantly increases the system capacity by employing a large number of antennas at the base station. As an emerging and promising technology, massive MIMO also enjoys many advantages such as low-power, robust transmissions, simplified transceiver design, and simplified multiple access layer.

This special issue of E-Letter focuses on the recent progresses of massive MIMO systems. It is the great honor of us to have seven invited contributions from both academia and industry to discuss various aspects/issues of massive MIMO/full-dimension MIMO, report their ideas/solutions for attacking those issues, and share their latest results.

The first letter with the title of "Full-dimension MIMO Cellular Systems Realizing Potential of Massive-MIMO" is written by Yang, Md Saifur, and Young-Han from Samsung Research America. Currently, Samsung is leading the work of full-dimension MIMO (FD-MIMO) within 3GPP community for Rel-12 LTE-Advanced systems. This letter gives a systematic overview of Full-Dimension MIMO (FD-MIMO) including design challenges, channel knowledge acquisition, and system level evaluations.

"3-Dimensional Channel Characteristics on Active 2-Dimensional Antenna Array System" is contributed by Hyoungju, Choelkyu, Hoondong, Young-Han, and Younsun from DMC R&D Center from Samsung Electronics. 3D massive MIMO/FD-MIMO is a paradigm shift from the conventional 2D MIMO systems to a 3D system. Accordingly, channel modeling becomes a critical issue. This letter presents a comprehensive modeling of the underlying 3D stochastic channel.

The third letter is contributed by Rubayet (University of Kansas), Lingjia (University of Kansas), and Charlie (Samsung Research America) with the title of "On the Channel Estimation for 3D Massive MIMO Systems". In this article, various channel estimation methods including channel transfer function estimation and channel estimation based on parametric channel models (direction-of-arrival and direction-of-departure estimations) are discussed for 3D massive MIMO/FD-MIMO systems. 3D beam-forming and impacts of channel estimation are also illustrated.

The fourth letter with the title of "On Channel Acquisition for Massive MIMO System" is written by Runhua, Qiubin, Hui, Rakesh, and Shaohui from China Academy of Telecommunication Technology (CATT). This letter studies channel knowledge acquisition for massive MIMO system in LTE-Advanced. For deployment scenarios where UL/DL reciprocity holds (e.g. TDD), the DL channel can be inferred from UL measurement, allowing greater beam-forming flexibility, reduced system overhead, and lower UE complexity. For scenarios without UL/DL reciprocity (e.g. FDD), a two-dimension feedback mechanism is introduced to take advantage of both the elevation and azimuth degrees of freedom.

"Exploiting Adaptive Downtilt and Vertical Sectorization in LTE Advanced Networks using Active Antenna Systems" is contributed by Meilong, Moon-il, Ananth, Mohsen, and Janet from InterDigital. The letter showed that adaptive downtilt and vertical sectorization are promising techniques for 3D massive MIMO/FD-MIMO systems. When elevation degrees of freedom is utilized, the evaluation presented in the letter showed that adaptive downtilt achieves up to 11% cell edge and 5% cell average spectral efficiency gain compared to the baseline system with a fixed downtilt. This result provides strong evidence on utilizing

elevation degrees of freedom for 3D massive MIMO/FD-MIMO systems.

Starting from the sixth letter, we begin to look at networking related issues for massive MIMO/FD-MIMO systems. To be specific, the sixth letter is contributed by Yi from Auburn University, Guosen from Broadcom, and Shiwen from Auburn University with the title of "User Grouping and Load Balancing for FDD Massive MIMO Systems". This paper studies user grouping and scheduling problems based on a two-stage precoding framework for FDD massive MIMO systems. The weighted likelihood similarity measure and hierarchical clustering for user grouping are proposed. A dynamic user scheduling scheme and a user grouping algorithm to achieve load balancing and user fairness for FDD massive MIMO systems are introduced. The efficacy of the introduced schemes has been validated with analysis and simulations.

The last paper with the title of "Massive MIMO Operation in Fronthaul-capacity Limited Cloud Radio Access Networks" is contributed by Sangkyu, Chan-Byoung, and Saewoong from Seoul National University and Yonsei University. In this letter, the authors investigate the effect of a limited number of antennas or users due to a fronthaul link constraint on the wireless sum-rate, which also depends on the beamforming strategy and the radio signal transport method in a fronthaul link. It is shown in the letter that for a given fronthaul link capacity and user environments, the beam-forming and the fronthaul transport methods can be jointly optimized to maximize the sum-rate.

We would like to thank all authors for their contributions and great efforts. We hope you enjoy reading this special issue and also hope these articles can stimulate further research works in this area.

Lingjia Liu (S'03–M'08) received the B.S. degree in Electronic engineering from Shanghai Jiao Tong University, China, and the Ph.D. degree in Electrical Engineering from Texas A&M University, USA. He is currently working as an Assistant Professor in the EECS Department at the University of Kansas (KU), USA. Prior to that he spent more than three years in Samsung Research America leading Samsung's work on multi-input-multiple-output (MIMO), coordinated multi-point (CoMP) transmission, and heterogeneous networks for 3GPP LTE and LTE-Advanced standards. His general research interests lie in the areas of wireless communication systems, statistical signal processing, queuing theory, information theory, with emphasis on delay-sensitive communication over wireless networks, and smart energy networks.

Dr. Liu served as Technical Program Committee (TPC) member/chairs of various international conferences including IEEE ICC, IEEE GLOBECOM, IEEE ICNC, WCSP, etc. He currently serves as an Editor for IEEE Trans. on Wireless Comm., and Associate Editors for EURASIP Journal on Wireless Comm. and Networking and International Journal on Comm. Systems.

GUOSEN YUE (S'99-M'04-SM'09) received the B.S. degree in physics and the M.S. degree in electrical engineering from Nanjing University, China in 1994 and 1997, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, in 2004. He was a senior research staff with the Mobile Communications and Networking Research Department, NEC Laboratories America, Princeton, New Jersey. In August 2013, he joined Broadcom Corporation as a system design scientist. Dr. Yue serves as an Associate Editor for the IEEE Transactions on Wireless Communications. He has served as the associate Editor for Research Letters in Communications, the Guest Editors for EURASIP Journal of Wireless Communication and Networking special issue on interference management, ELSEVIER PHYCOM special issue on signal processing and coding. He served as the Symposium Co-chair for IEEE ICC 2010, the Track Co-chair for IEEE ICCCN 2008, the steering committee member for IEEE RWS 2009. He is a senior member of the IEEE.

Full-dimension MIMO Cellular Systems Realizing Potential of Massive-MIMO

Yang Li, Md Saifur Rahman and Young-Han Nam
Samsung Research America, United States
{yang.l, md.rahman, younghan.nam}@samsung.com

1. Introduction

The rapid growth of data traffic driven by mobile devices poses challenges on capacity of cellular networks. One promising technology for meeting the demands is massive MIMO, by employing a large number of antennas to exploit MU-MIMO (multi-user MIMO) [1]. However, most of previous works have considered massive number of antennas being arranged along the horizontal axis, which is not applicable in practice, owing to the restriction on the base station (BS) form factor. Recently, a concept of full-dimension MIMO (FD-MIMO) adopting 2-dimensional (2D) array antenna has been proposed with intent to cope with these practical challenges [2]. A typical FD-MIMO deployment scenario is illustrated in Figure 1, for a macro base station (BS) with 2D active antenna array (AAA) panels.

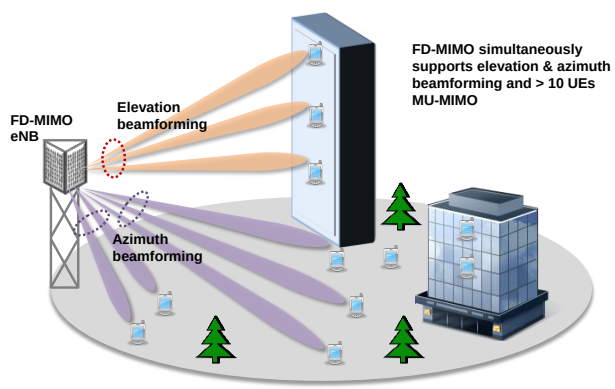


Figure 1 FD-MIMO deployment scenario example

This paper presents recent advances in FD-MIMO in cellular systems, and shows that FD-MIMO attains 2-5 times cell average throughput gain and cell-edge throughput gain using system-level simulations. Key challenges are discussed for implementing FD-MIMO in LTE/LTE-advanced, especially channel state information (CSI) estimation and channel quality indicator (CQI) prediction.

2. System Model

In this section, we first introduce 2D AAA, the key component of FD-MIMO and then present a system model.

FD-MIMO Architecture.

One of the most important features of the FD-MIMO BS is that the BS can perform 3D beam steering in both azimuth and elevation angles. An FD-MIMO BS equipped with 2D antenna array, which is fed by multiple TRX chains as illustrated in Figure 2, enables the multiple 3D beam steering toward multiple users. In one extreme case, each element in the 2D array is fed by a separate TRX chain, in which case the number of TRX chains is the same as the number of antenna elements. In typical deployment scenarios, however, a TRX feeds multiple antenna elements simultaneously, which reduces hardware cost and implementation complexity. For beam-steering, the number of TRX chains in each dimension determines the steering precision; the larger the better.

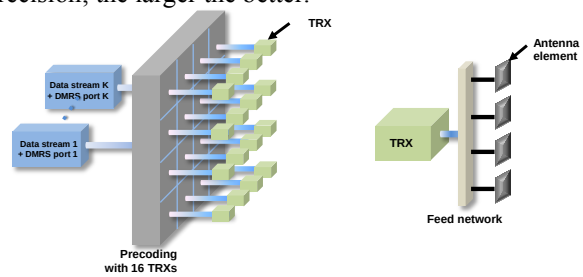


Figure 2 FD-MIMO system architecture

Signal Model.

Let Nt, Nr, K , and Nd be the number of the transmit antennas (or number of TRX) at BS, the number of the receive antennas at a user, the number of UEs, and the number of data streams at a user. The signal transmitted by the BS at subcarrier j is:

$$\boldsymbol{x}j=k=1K\boldsymbol{W}kj\boldsymbol{s}kj, \qquad (1)$$

where $\boldsymbol{W}_k^j$ is an $Nt\times Nd$ precoding matrix used for user i, $\boldsymbol{s}_i^j$ is a $N_d\times 1$ signal vector transmitted for user i. The transmit power is constrained by $\mathbf{E}j\mathrm{Tr}(\boldsymbol{xj}(\boldsymbol{x}j)H=1$, where $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix. For notation simplicity, otherwise specified the subcarrier index j is dropped in the rest of the paper.

The received signal at user k is:

$$\boldsymbol{y}k=\boldsymbol{H}k\boldsymbol{W}k\boldsymbol{s}k+\boldsymbol{H}k\Sigma i\neq k\boldsymbol{W}i\boldsymbol{s}i+\boldsymbol{n}k,\ k=1,...,K\ , \quad (2)$$

where $\boldsymbol{H}_k$ is an $Nr\times Nt$ channel matrix between the BS and user k, and n_k is the noise at user k.

3. FD-MIMO Design Challenges

An FD-MIMO BS perform 3-dimension (3D) precoding to efficiently deliver signals to serving users while minimizing mutual interference among users,

and therefore significantly increase spectral efficiency by high-order MU-MIMO transmission. To harvest gain of FD-MIMO, the BS must acquire CSI of users, which is typically obtained by uplink sounding reference signals (SRS) in time division duplex (TDD) systems or via CSI feedback in FDD systems. Large number of antennas used in FD-MIMO gives new challenges on circuit and CSI feedback design.

CSI Acquisition by Reciprocity in TDD.
In TDD systems, BS can obtain downlink CSI using uplink channel sounding if channel reciprocity holds. To exploit reciprocity in practice, however, phase and amplitude mismatch in different antenna branches must be calibrated and compensated. Otherwise, the channel directions estimated based on SRS is not aligned with actual downlink channels, as illustrated in Figure 3.

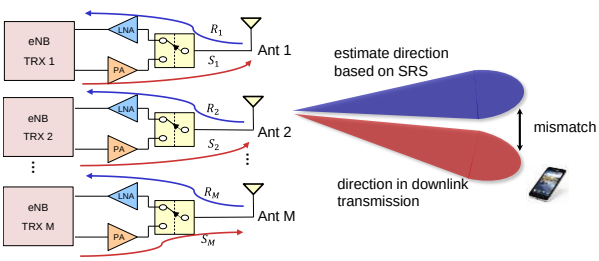


Figure 3 Illustration of antenna calibration issues

In FD-MIMO, the calibration circuit needs to interconnect a large number of antennas, which makes it difficult to manufacture these interconnections with uniform high precision. Calibration circuits with insufficient phase accuracy will result in large residual errors and degraded beam steering performance. One example is shown in Figure 4, where cascaded conventional splitters are used to connect antennas to a calibration transceiver. Error is accumulated during each stage, which may result in intolerable inaccuracy (e.g., 14° in Figure 4). Therefore, in TDD FD-MIMO new reliable and scalable calibration architecture and algorithm are needed to maintain the reciprocity.

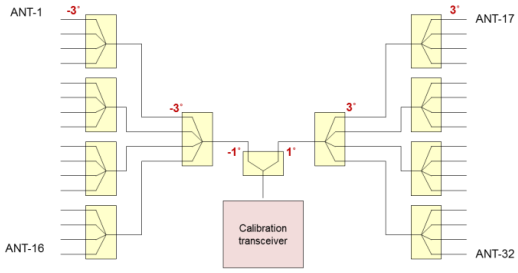


Figure 4 Illustration of error propagation in calibrating a large array

CSI Acquisition by Feedback in FDD.
In FDD systems, precoding based on channel-reciprocity is challenging, and thus CSI feedback is typically required. In current 3GPP LTE/LTE-advanced systems, CSI includes PMI (precoding matrix indicator), RI (rank indicator), and CQI (channel quality indicator). For FD-MIMO, overhead of CSI-RS (CSI reference signal) required for downlink channel estimation will be large, which is usually proportional to the number of TRX in the FD-MIMO systems. In addition, CSI feedback overhead, CSI estimation and PMI selection complexity increase. For example, if 16 bit codebook is used for 16-Tx antenna systems, the number of codewords in the codebook is 216 = 65536, which imposes huge complexity in implementation, and may not be easily fit in in the currently-designed CSI feedback channels in 3GPP LTE/LTE-A. Hence, it is necessary to maintain a small codebook size while achieving FD-MIMO throughput gain.

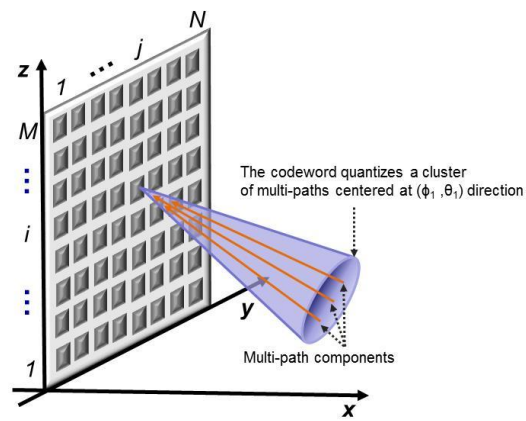


Figure 5 Codeword (PMI) selection for 2D AAA in FD-MIMO

One potential approach is to exploit channel correlations among antennas in 2D array panel. As shown in Figure 5, a preferred PMI shall effectively capture power of the multi-path components centered around a certain elevation and azimuth angle. The selected precoder (PMI) can be represented by Kronecker product of a horizontal and vertical channel component precoders. For example, for a FD-MIMO systems with NM antennas, N antennas in a horizontal row and M antennas in a vertical column, BS transmits CSI-RS in the first row and the first column only, and a user estimates and feeds back the corresponding Nx1 horizontal channel h_H and Mx1 vertical channel h_V, respectively. The BS reconstructs the channel as:

$$\boldsymbol{h}=\boldsymbol{h}H\otimes\boldsymbol{h}V \qquad (3)$$

In this case, the CSI-RS resource used and the feedback overhead is proportional to $N + M$, instead of NM, and a significant resource and complexity reduction is attained when N, M is large.

Other Possibility in CSI Acquisition in FDD.
Although for FDD system, channel reciprocity does not hold, downlink and uplink channels are not independent, because in either direction the

electromagnetic waves propagate through similar environment, as shown in Figure 6. In [3], measurement campaign on FDD UMTS band shows that power-delay profile (PDP) and root-mean-square (RMS) of delay spread is very similar in both downlink and uplink. Therefore, in many cases uplink and downlink have similar values in the number of multipath components, angle of arrival (AoA) and departure (AoD), power and delay of each path.

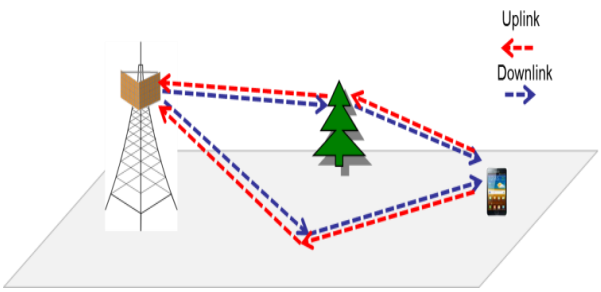


Figure 6 Dependency between downlink and uplink for FDD systems

We now study the correlation between the uplink and downlink covariance matrices. Figure 7 shows that dominant eigenvalue between uplink and downlink covariance is strongly correlated, regardless of duplex distance. For WCS band with small duplex distance 50MHz (downlink 2300MHz, uplink 2250MHz), the correlation is 0.995, and for AWS band with large duplex distance 400MHz (downlink 2100MHz, uplink 1700MHz) the correlation is still as high as 0.827. In contrast, we observe that the correlation of directions of the dominant eigenvectors highly depends on duplex distance, as shown in Figure 8, where the correlation is measured as angle between two vectors (0-degree means fully correlated, while 90-degree means uncorrelated). In WCS band the probability of the angle being less than 30 degree is 99.6%, while in AWS band such probability is less then 1%.

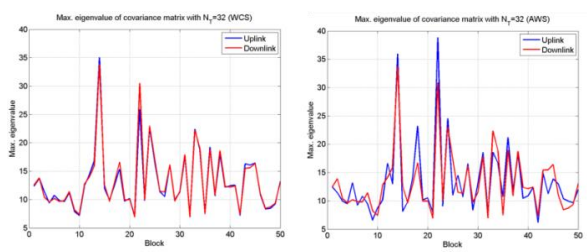


Figure 7 Maximum eigenvalue of uplink and downlink covariance matrices

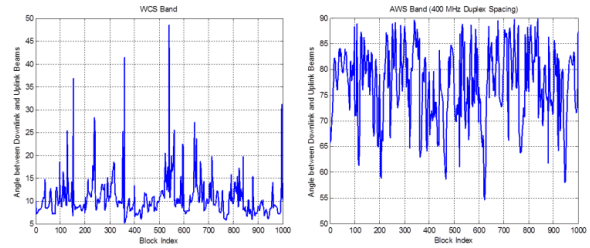


Figure 8 Angle between dominant eigenvector of uplink and downlink covariance matrices

CQI Compensation and Prediction.

In cellular networks, CQI is required in channel-dependent scheduling and is reported by the users after processing channel estimates on cell-specific reference signals (CRS) or CSI-RS. For instance, the user estimates the best precoder (or, PMI) that matches the measured channel information using the CSI-RS received and estimates the CQI presuming the PMI is used. In FD-MIMO, CQI mismatch is expected due to e.g. the number of antennas used in precoding is different in downlink estimation in TDD systems and channel reconstruction inaccuracy in FDD.

In this paper, a CQI mismatch compensation scheme is proposed for CSI-RS-based CQI in TDD systems, in which user k estimates downlink channel direction based on uplink SRS: $\boldsymbol{H}_k, k = 1, \ldots, K$. The received CSI-RS symbols with antenna virtualization $\boldsymbol{W}_0$ at user k is

$$\boldsymbol{y}k=\boldsymbol{H}k\boldsymbol{W}0\boldsymbol{s}0+\boldsymbol{n}k,\ k=1,\ldots,K\ . \quad (4)$$

For simplicity, it is assumed that the single-user (SU) CQI for user k is equal to SINR (signal-to- interference and noise ratio) ρ_{0k}:

$$\rho 0k=||\boldsymbol{H}k\boldsymbol{W}0||2\sigma k2, \quad (5)$$

where σ_k^2 is the receiver noise plus inter-cell interference, which is unknown by the BS. The SINR ρ_k for the data symbols when precoding Wk is applied is:

$$\rho k=||\boldsymbol{H}k\boldsymbol{W}k||2\sigma k2. \quad (6)$$

Given $\boldsymbol{H}_k$ is known by the BS, it can obtain SINR for data channels as:

$$\rho k=||\boldsymbol{H}k\boldsymbol{W}k||2||\boldsymbol{H}k\boldsymbol{W}0||2\ \rho 0k. \quad (7)$$

In the above equation, $\boldsymbol{W}_0$ and $\boldsymbol{W}_k$ are known at the BS and $\boldsymbol{H}k$ can be estimated based on SRS in the uplink transmission for TDD systems, and accordingly the BS can predict the CQI for link adaptation. Figure 9 shows the above prediction works well if there is no MU-MIMO transmission, achieving 10% normalized prediction error. In case of MU-MIMO, the CQI prediction is more difficult and designing reliable estimation algorithms is still open issue.

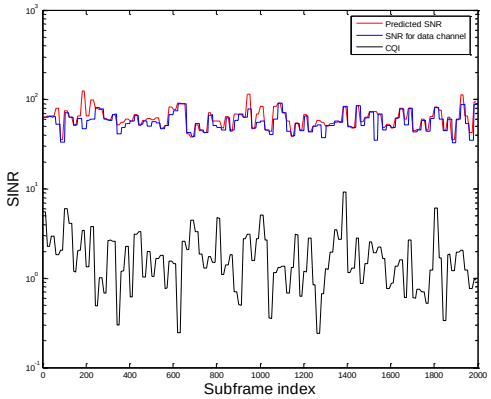


Figure 9 CQI prediction for SU-MIMO

4. System-level Simualtion

This section presents system-level simulation results for the proposed FD-MIMO system based on 3D ITU channel model [4]. A 4 antenna system is considered as a baseline. Figure 11 and Figure 10show FD-MIMO capacity gain with two precoding methods, conjugate beamforming and signal-to-leakage-and-noise ratio (SLNR) [5]. For conjugate beamforming, around 2.5 times gain of cell average throughput and 5 times gain of cell edge throughput are achieved. For SLNR, both cell average throughput and cell edge throughput achieve 5 times gain.

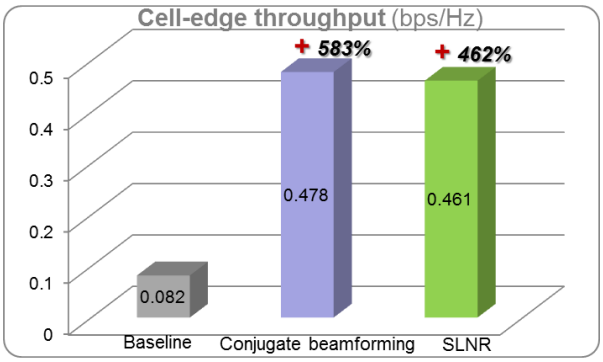


Figure 10 FD-MIMO cell edge capacity gain

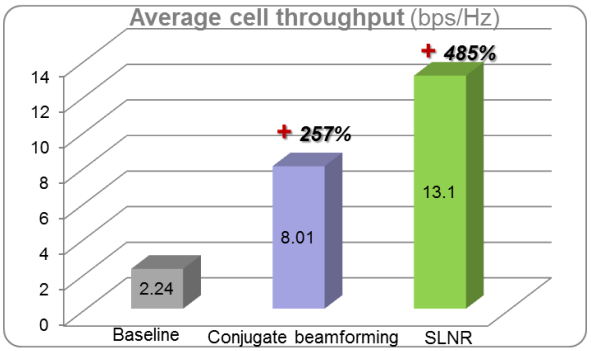


Figure 11 FD-MIMO cell average capacity gain

5. Conclusion

In this paper, challenges in FD-MIMO are presented with preliminary solutions for CSI acquisition in both TDD and FDD, and CQI prediction. System simulation results show a 2-5 times cell average throughput gain as well as cell-edge throughput gain. Based on our results, FD-MIMO has great potential to improve spectral efficiency in LTE and LTE-advanced systems.

References

[1] T. L. Marzetta, "Non-cooperative cellular wireless with unlimited numbers of base station antennas", IEEE Trans. Wireless Communications, vol. 9, Issue 11, pp. 3590 – 3600, November 2010.

[2] Y. H. Nam, B. L. Ng, K. Sayana, Y. Li, J. Zhang, Y. Kim and J. Lee, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology", IEEE Commun. Mag., 2013.

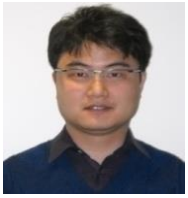
[3] Sana Salous and Hulya Gokalp, "Medium- and Large-Scale Characterization of UMTS-Allocated Frequency Division Duplex Channels", IEEE Trans. On Vehicular Technology.

[4] 3GPP TR36.873 v2.0.0 Study on 3D channel model for LTE (2014)

[5] M. Sadek, A.Tarighat, and A. H. Sayed, "Leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, May 2007.

Yang Li a Staff Research Engineer in Samsung Research America, Dallas. He received a B.S. and a M.S. degree, in 2005 and 2008, in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, and a Ph.D. degree in 2012, in electrical engineering from The University of Texas at Dallas, Richardson, TX, USA. His research interests include MIMO, interference management, cooperative communication and cognitive radio.

Md Saifur Rahman received his M.S. and Ph.D. degrees in Electrical and Computer Engineering from Cornell University in 2011 and 2012, respectively, and received his B.Tech. Degree in Instrumentation Engineering from Indian Institute of Technology Kharagpur in 2006. He worked as a Post-Doctoral Associate in the School of Electrical and Computer Engineering at Cornell University during January - October 2012. Since November 2012, he has been with Samsung Research America - Dallas, where he is currently a Senior Standards Engineer working on next generation communication standards. His areas of interest include wireless communication, optimization, and compression. He is a recipient of Irwin D. Jacobs fellowship in 2008. He was nominated for the ECE director's PhD thesis award for his doctoral work.



Young-Han Nam is currently leading a team working on a MIMO and algorithm in Samsung Research America at Dallas (SRA-D), Richardson Texas, in which he has been working ever since 2008. He has been engaged in standardization, design and analysis of the 3GPP LTE and LTE-Advanced from Release 8 through 12. He has received a Ph.D. in electrical engineering from the Ohio State University, Columbus, in 2008, and B.S. and M.S from Seoul National University, Korea, in 1998 and 2002, respectively. His research interests include channel models, mmWave, full-dimension MIMO, massive MIMO, MIMO transceiver algorithm.

3-Dimensional Channel Characteristics on Active 2-Dimensional Antenna Array System

Hyoungju Ji[†], Choelkyu Shin[†], Hoondong Noh[†], Young-Han Nam[‡] and Younsun Kim[†]
[†]Samsung Electronics DMC R&D Center, Korea; [‡]Samsung Research America, USA
{hyoungju.ji, ck13.shin, hoondong.noh, younghan.n and younsun}@samsung.com

1. Introduction

Massive MIMO or large-scale antenna system is one of the key technologies currently studied and the 3rd Generation Partnership Project (3GPP) Long Term Evolution (LTE) standard prepares this for the next generation systems based on LTE-advanced system [1] [2]. As a first step, a study item [3] has been initiated to study new channel model and a new user deployment scenarios under which evaluation of the antenna technologies will be performed. Follow-up 3GPP study and work items on FD-MIMO are expected in 3rd quarter of 2014. The focus of the study and work items will be to identify key areas in the LTE-Advanced standards that need to be enhanced in supporting more than 8 antenna ports placed in a 2-D array structure. Most of massive MIMO literatures assume that the antenna ports are horizontally placed in the transmitter-side. However, 64TX system in horizontal has almost 4m wide, which is not proper to practical system deployment. By incorporating 2-D active array into LTE systems, it is expected that system throughput will be drastically improved beyond what is possible in conventional LTE systems with passive antennas at a cost of reasonable array size.

With 2-D active array antenna, FD-MIMO system can have two advantages in MIMO transmission perspective. In multiple-user MIMO point of view, FD-MIMO utilizes multiple antennas placed in a 2-D antenna array panel to realize high order multi-user MIMO (MU-MIMO) transmissions. High order MU-MIMO refers to the use of a large number of antennas at the base station to transmit or receive spatially multiplexed signals to or from a large number of terminals by 3-D user-specific beamforming in both horizontal and vertical direction. The other point of view, such as single-user MIMO, 2-D antenna array can radiate the signal into 3-D space and the channel independency and de-correlation between vertical ports as well as horizontal ports can have rank for UE depending on UE's location [10][11].

To evaluate 3-D beamforming and MIMO technics, it is necessary to model large and small scale parameters in vertical domain as well as in horizontal domain. In this paper, we discuss spatial channel characteristics of vertical domain as well as horizontal domain under different 2-D active array configuration. Impact on separate and full channel measurement and feedback for horizontal and vertical antennas are also discussed.

2. 3-Dimensional Stochastic Channel Model

In this section, we first present the formulation of the 3-Dimensional stochastic channels with considering both vertical and horizontal domain.

Channel Coefficient Generation

3-Dimensional channel can be modeled with N clusters, modeling scatter between transmitter and receiver and M rays per a cluster. Conceptual 3-Dimensional channel is shown in Figure 1 with an exemplary view of one of the clusters. Channel coefficient for each cluster n and each receiver and transmitter element pair u, s is given by [4-8]:

$$H_{u,s,n}(t) = \sqrt{P_n / M} \sum_{m=1}^{M} \left(\begin{array}{l} \begin{bmatrix} F_{rx,u\theta}(\theta_{n,m,ZOA}, \varphi_{n,m,AOA}) \\ F_{rx,u\varphi}(\theta_{n,m,ZOA}, \varphi_{n,m,AOA}) \end{bmatrix}^T \times \\ \begin{bmatrix} exp(j\Phi_{n,m}^{\theta\theta}) & \sqrt{\kappa_{n,m}^{-1}} exp(j\Phi_{n,m}^{\theta\varphi}) \\ \sqrt{\kappa_{n,m}^{-1}} exp(j\Phi_{n,m}^{\varphi\theta}) & exp(j\Phi_{n,m}^{\varphi\varphi}) \end{bmatrix} \times \\ \begin{bmatrix} F_{tx,s\theta}(\theta_{n,m,ZOD}, \varphi_{n,m,AOD}) \\ F_{tx,s\varphi}(\theta_{n,m,ZOD}, \varphi_{n,m,AOD}) \end{bmatrix} \times \\ exp(j2\pi\lambda_0^{-1}(\hat{r}_{rx,n,m}^T \cdot \bar{d}_{rx,u})) \times \\ exp(j2\pi\lambda_0^{-1}(\hat{r}_{tx,n,m}^T \cdot \bar{d}_{tx,s})) exp(j2\pi)_{n,m} t) \end{array} \right)$$

where,

P_n is the power of nth cluster

$F_{rx,u,\theta}$, $F_{rx,u,\theta}$ are the receive antenna element n field patterns in the direction of azimuth angle and zenith angle, respectively

$F_{tx,s,\theta}$, $F_{tx,s,\theta}$ are the transmit antenna element s field patterns in the direction of azimuth angle and zenith angle, respectively

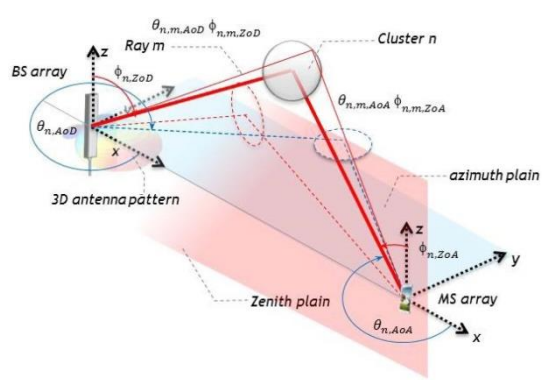


Figure 1. 3-Dimensional Channel.

$\Phi_{n,m}^{\theta\theta}, \Phi_{n,m}^{\theta\phi}, \Phi_{n,m}^{\phi\theta}, \Phi_{n,m}^{\phi\phi}$ are the initial phase for each ray m of each cluster n for different polarization combinations

$\theta_{n,m,ZOA}, \phi_{n,m,AOA}$ are zenith and azimuth arrival angle for ray m of cluster n, respectively

$\theta_{n,m,ZOD}, \phi_{n,m,AOD}$ are zenith and azimuth departure angle for ray m of cluster n, respectively

$\hat{r}_{rx,n,m}, \hat{r}_{tx,n,m}$ are the spherical unit vector of receiver and transmitter, respectively

$\bar{d}_{rx,u}, \bar{d}_{tx,s}$ are the location vector of receiver antenna elements u and transmit antenna element s, respectively

$\kappa_{n,m}$ is the cross polarization power ratio

λ_0 is the wave length of the carrier frequency

$v_{n,m}$ is the Doppler frequency component using speed and travel angle of receiver

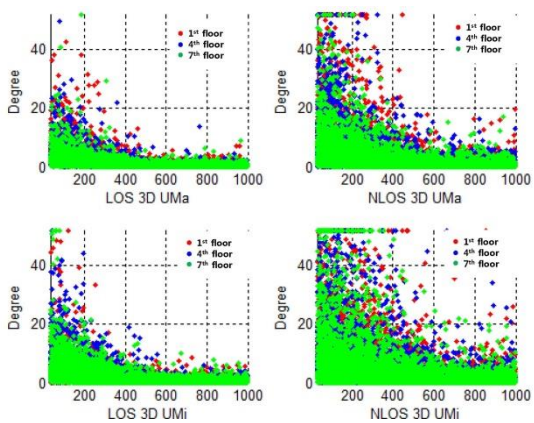


Figure 2. Distance Dependent RMS of ZSD

In equation above, we can observe that phase shifted channel between antenna elements which depends on relative distance of antenna elements $\bar{d}_{rx,u}, \bar{d}_{tx,s}$, can vary with different angular spread of clusters in vertical and horizontal direction. The different variance of angular spread between vertical and horizontal domain cause a significant impact on the channel spatial multiplexing capacity using different 2D active array structure. The generation of arrival and departure angles w_n of each cluster n follows random angles in proportional of the power of cluster P_n as

$$w_n = (X_n w_n' + Y_n) + \overline{w}$$

where w_n' is $\eta\sqrt{-\ln(P_n/\max(P_n))/C}$ for horizontal domain and $-\eta - \ln(P_n/\max(P_n))/C$ for vertical domain, respectively and C is a scaling factor related to the total number of cluster in the given channel condition and X_n is integer random number from set {-1, 1}. Other parameters are listed in Table 1 and [7, 8]. The angular spread for each cluster is derived by RMS angular spread η and angles of rays within a cluster randomly spread centering on a cluster angle. Even with the same RMS angular spread for azimuth and zenith angles, the average azimuth cluster angle is more than a square of the zenith cluster angle.

For example, UE has one antenna element and eNB has two antenna elements in vertical direction with coordination of (x, y, z)=(0, 0, 0) and (0, 0, 0.5λ). Channel of second antenna elements is phase shifted from the reference (first) antenna in proportional to zenith angle of departure. RMS distribution of zenith angular spread of departure (ZSD) is shown in the Figure 2. The figure shows that we can have similar angular spread in elevation domain and horizontal domain (15º-30º) in case of UEs placed center of cells. This reveals that high rank transmissions can happen with vertically spaced array than legacy array structure which placed horizontal only. It is noted that there is marginal different of angle distribution with the different floors, however, smaller angles are observed when UE located in higher floor.

3. Performance Assessment

In this section, we investigate the characteristics of 3-Dimensional channel based on different UE feedback schemes for 2D active array system. For TDD system, the eNB can have knowledge of channel of massive antennas relying on channel reciprocity. However, reciprocity does not hold for FDD system and need to have a proper scheme to feedback channel of antennas.

Channel Measurement for 2-Dimensional Array.

Scheme #1: Full channel measurement and single feedback - The receiver configures $N_T = N_H \times N_V$ antennas from the transmitter with designed precoder matrix for measurement. Search complexity for precoder p_n can significantly increase in proportional to the total number of antenna in the system.

$$p_n = \underset{n\in\mathbf{N}, p\in\mathbf{P}_n,}{\arg\max} \left\| \mathbf{H_n P_n} \right\|$$

where $\mathbf{H_n}$ is $N_{RX} \times N_V$ channel matrix, $\mathbf{N}$ is set of rank, $\mathbf{P_n}$ is the set of rank n precoder matrix and p_n is precoder with 2-D DFT $N_T \times n$ matrix.

Scheme #2: Full channel measurement and separate feedback - The receiver configures N_T antennas from the transmitter as in Scheme #1. To reduce complexity of selecting best precoder and preferred rank, horizontal precoder p_n^H and vertical precoder p_n^V can be separately searched and feedback.

$$p_n^H = \underset{n\in\mathbf{N}^H, p\in\mathbf{P}_n^H}{\arg\max} \sum^{N_V} \left\| \mathbf{H}_n^H \mathbf{P}_n^H \right\|, \quad p_n^V = \underset{n\in\mathbf{N}^V, p\in\mathbf{P}_n^V}{\arg\max} \sum^{N_H} \left\| \mathbf{H}_n^V \mathbf{P}_n^V \right\|$$

Scheme #3: Separate channel measurement and separate feedback - To reduce measurement and best precoder selection complexity, the receiver only configures $N_H + N_V$ measurement resources out of N_T antennas from the transmitter. The receiver feedbacks two precoders; one from N_H antennas and another from N_T antennas. With this scheme, measurement overhead can be decreased in order of half, however, UE has partial knowledge of overall channel and this leads mismatch of selected precoder from the Scheme 1 and 2.

$$p_n^H = \underset{n \in \mathbf{N}^H, p \in \mathbf{P}_n^H}{\arg\max} \left\| \mathbf{H}_n^H \mathbf{P}_n^H \right\|, \quad p_n^V = \underset{n \in \mathbf{N}^V, p \in \mathbf{P}_n^V}{\arg\max} \left\| \mathbf{H}_n^V \mathbf{P}_n^V \right\|$$

After receiving CSI feedback from the UE, the transmitter selects weight for each active array using p_n for Scheme #1 and the overall weight $\hat{p}_n$ can be calculated using kronecker product of two precoders for Scheme #2 and #3.

$$\hat{p}_n = p_n^H \otimes p_n^V$$

Two Dimensional Rank

This section presents the link-level performance for a 2-D array antenna system based on the channel measurement and feedback methods introduced in previous section. The simulation is conducted with 3-D channel models where the eNB is equipped with eight transceiver unit with 32 antenna elements and each transceiver unit has one antenna element in horizontal and 4 in vertical direction. Equivalently, the eNB has 2-D transceiver units with 4 transceiver unit in horizontal and 2 transceiver units in vertical with 32 antenna active elements. One UE with two receive antennas is dropped newly over 40msec time interval for different distances with d=50, 500, and 1000 meters and different heights with 1, 4 and 7th floor.

In Figure 3, we can see that the performance loss from the separate feedback (gap between scheme 1 and scheme 2) is up to 3dB. Figure 4 suggests a reason for the performance gap. Shown in Figure 4, schemes 2 chooses full rank in both horizontal and vertical domain (H, V)=(2, 2) with high ratio in all SNR ranges even though (H, V)=(2, 2) cannot be used in precoder selection of separate feedback. In addition, the number of codebook for scheme 2 (64+48+64=176) is smaller than that for scheme 1 ($2x2^8$=512). As a result, inaccurate precoder selection results in performance degradation of separate feedback.

In Figure 3, we can observe that scheme 2 achieves some performance gain over scheme 3 by using full channel measurement. In other words, the performance of scheme 3 can decrease due to feedback mismatch using measuring 56% of channels. However, we can also observe that the performance of scheme 3 is very close to that of scheme 2 with separated feedback mechanism. It is observed that if UE select feedback in each direction separately, knowledge of channel information has less impact on overall performance.

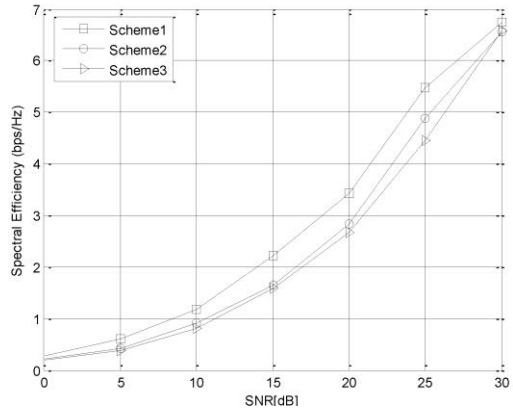


Figure 3. Spectral Efficiency Results.

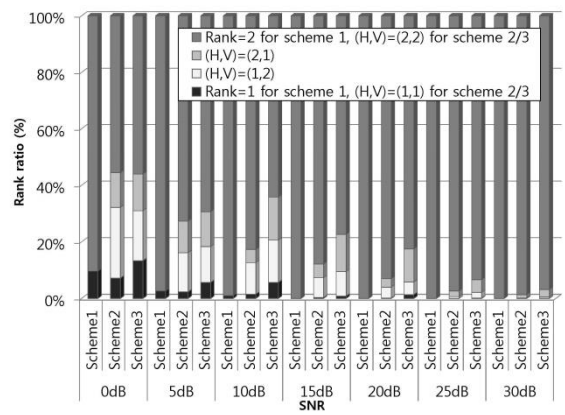


Figure 4. Rank Ratio Results

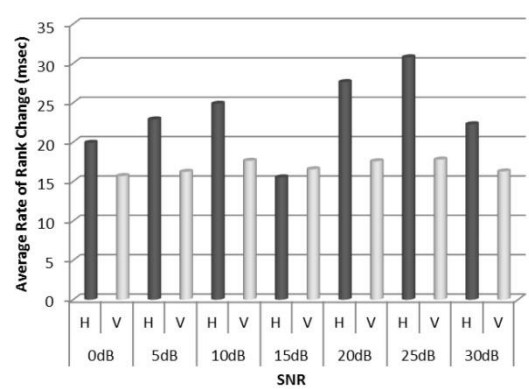


Figure 5. Change Rate of Vertical (V) and Horizontal (H) Ranks (d=500m, 4th floor).

Lastly, we evaluate rate of rank change estimated with CSI-RS whose duty cycle is 5 msec. The average rate of change is shown in Figure 5 for both horizontal and

vertical domain, when scheme 2 is used. As shown in the figure, the vertical rank changes quite often, once in as small as 15msec in all SNR range. Therefore, in 2-D active array antenna systems, it is important to exploit both vertical and horizontal rank with proper feedback interval.

4. Conclusion

We have investigated 3-D channel characteristics with various 2-D active array antenna systems. We presented the link-level performance of 2-D active array antenna system using different channel measurement and feedback methods. Simulation result shows that 2-D antenna structure based on full channel measurement and feedback outperforms than full or partial channel measurement and separate feedback mechanism but small degradation can be obtained with separate measurement with reduced feedback.

References

[1] Younsun Kim, Hyoungju Ji, Hyojin Lee, Juho Lee, Boon Loong Ng and Jianjhong(Charlie) Zhang "Evolution Beyond LTE-Advanced with Full Dimension MIMO," ICC 2013, Jun 2013.

[2] M.T. Feeney and J.D. Parsons, "Cross-correlation between 900MHz signals received on vertically separated antennas in small-cell mobile radio systems," IEE Proceeding, Vol.128, No.2, Apr 1991.

[3] 3GPP TSG RAN Plenary #58, R1-122034, "Study on 3D-channel model for Elevation Beamforming and FD-MIMO studies for LTE," Barcelona, Spain, 4th – 11th Dec 2012.
C. Schneider et al., "Large Scale Parameter for the WINNER II Channel Model at 2.53 GHz in Urban Macro Cell", IEEE VTC (Spring), 2010.

[4] K. Baltzis et al., "A Simple 3D Geometric Channel Model for Macrocell Mobile Communications", Wireless Personal Communications, 2009.

[5] X. Lu et al., "Comparison of Antenna Arrays in a 3-D Multiuser Multicell Network", ICC 2011.

[6] 3GPP TR 36.873 V1.0.0 Release 10.

[7] 3GPP TR 36.819 V10.0.0 Release 10.

[8] Boon Loong Ng, Younsun Kim, Juho Lee, Yang Li, Young-Han Nam, Jianjhong(Charlie) Zhang, and Krishna Sayana, "Fulfilling the Promise of Massive MIMO with 2D Active Antenna Array," Globecom, Dec 2012.

[9] Fredirk Athley, Martin Alm, Ola Kaspersson, Karl Werner, John Furuskog and Bo Hagerman, "Dual-polarized base station antenna configurations for LTE," IEEE, 2010.

[10] Weiliang Xie, Tao Yang, Xuetian Zhu, Fengyi Yang and Qi Bi, "Measurement-Based Evaluation of Vertical Separation MIMO Antennas for Base Station," IEEE Antenna And Wireless Propagation Letters, Vol.11, 2012.

Hyoungju Ji received his B.S and M.S degrees from Sogang University, Korea in 2005 and 2007. He is currently with Samsung Electronics DMC R&D center, Suwon, Korea. His research interests include large-scale antennas, Relay, Small cell and Heterogeneous network. He has been working on standardization of LTE/LTE-Advanced and is rapporteur of elevation beamforming and FD-MIMO study in 3GPP.

Choelkyu shin received the B.S. degree from Inha University, Korea, in 2007, and the M.S. and Ph.D degrees from the Korea Advanced Institute of Science and Technology (KAIST), Korea, in 2009 and 2013, respectively. He is currently a senior engineer at Samsung Electronics Co., Suwon, Korea. His research interests include MIMO BIC-OFDM systems, link-adaptation techniques, PHY-MAC cross-layer design, and the general area of communication theory.

Hoondong Noh received the B.S. and Ph.D. degrees in electronic engineering from Yonsei University, Korea, in 2008 and 2014, respectively. He joined Samsung Electronics in 2014, and has been working on the standardization of wireless communication systems. His current research interests include wireless network resource allocation and signal processing for massive MIMO system.

Young-Han Nam received his B.S and M.S degrees from Seoul National University, Korea and a Ph.D. degree in electrical engineering from the Ohio State University, Columbus. Since February 2008, he has been with Samsung Telecommunications America, Dallas, Texas, engaged in the design of physical-layer for the LTE/LTE-A standards. His research interests include Massive and multiuser MIMO schemes, cross-layer design and information theory.

Younsun Kim received his B.S and M.S degrees in electronic engineering from Yonsei University, Korea, and his Ph.D degree in electrical engineering from the University of Washington in 1996, 1999, and 2009, respectively. He joined Samsung Electronics in 1999, and has since been working on the standardization of wireless communication systems such as cdma2000, HRPD and LTE/LTE-A. His research interests include multiple access schemes, network coordination and Massive MIMO for next generation systems.

On the Channel Estimation for 3D Massive MIMO Systems

Rubayet Shafin[1], Lingjia Liu[1], and Jianzhong (Charlie) Zhang[2]
[1]Department of Electrical Engineering and Computer Science, University of Kansas, USA
[2]Samsung Research America – Dallas, USA
{rshafin, lingjialiu}@ittc.ku.edu, jianzhong.z@samsung.com

1. Introduction

In recent years, wireless communications has experienced an unprecedented growth-rate of wireless data traffic due to the rapid introduction of various connected mobile devices and excessively data-hungry applications run on those devices. If the trend is observed, it is not very difficult to predict the demand for more data volume in future. According to the International Telecommunication Union (ITU) [1], the total number of mobile cellular subscription is fast approaching the total number of the people living on earth, and is expected to reach 7 billion by the end of 2014, which corresponds to a penetration rate of 96%. Cisco Systems predicts that Global Mobile data traffic will increase nearly 11 fold between 2013 and 2018 reaching 15.9 exabytes per month in 2018 [2]. In order to fulfill this huge capacity demand, an incrementally improved multi-user multiple-input-multiple-output (MU-MIMO) could be thought of as a primary solution to increase the spectral-efficiency of the underlying communication system. However, the presently available MIMO transmissions, allowed by LTE-Advanced systems [3], do not support more than 8 antenna ports at the base station.

Recently, Massive MIMO, also known as Large Scale MIMO or Full Dimension MIMO (FD-MIMO) ([4], [5]), has created much interests, both in academia and industry, with the promise of meeting the future capacity demand by providing increased spectral-efficiency achieved through aggressive spatial multiplexing. Considering the form factor limitation at the base station (BS), instead of placing a large number of antennas horizontally, 3D Massive MIMO system employs those antennas in a 2D antenna array panel enabling the exploration of the degrees of freedom in elevation domain along with those in the azimuth domain.

Apart from the huge potential of providing excellent spatial resolution and array gains, Massive MIMO can also offer a significant reduction of latency, a simplified multiple access layer, and robustness to interference [6]. With the help of a large number of antennas, this system can concentrate more energy in a particular direction leading to a dramatic increase in energy-efficiency. Furthermore, Massive MIMO is the key enabling technology for gigabit per second data transmission in the millimeter wave (mmW) wireless communications with carrier frequency between 30 and 300 GHz. In mmW communications, it becomes feasible to pack a greater number of antennas at the base station. However, the benefits of Massive MIMO are limited by the accuracy of the channel state information (CSI) obtained at the transmitter. The CSI is critical for functionalities such as downlink beam-forming, transmit precoding, user scheduling, etc.

In this letter, we discuss different channel estimation methods used for 3D Massive MIMO systems. We also describe the potential downlink beam-forming strategies for such systems.

2. Channel Estimation Methods

A typical 3D massive MIMO system is shown in Fig.1.

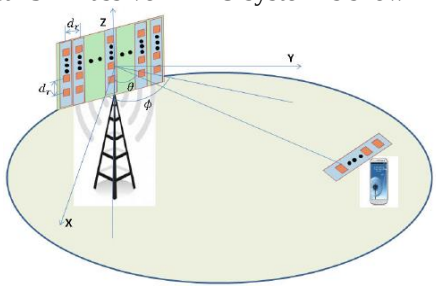


Fig.1: Model of 3D massive MIMO system

The antenna array at the base station is placed in the X-Z plane with M antennas in the horizontal direction and N antenna elements in the vertical direction. The spacing between adjacent antenna elements is assumed to be d_r.

In general, there are two methods to estimate the MIMO channel. First is the traditional way where the channel estimation is done by estimating the channel transfer function. Alternatively, channel estimation can be conducted based on the parametric channel models to estimate the direction of arrival (DoA) and direction of departure (DoD) of different paths.

2.1 Estimation of Channel Transfer Function

Channel state information can be obtained by sending some predefined pilot signals/reference signals, and estimating the channel matrix from the received signal. If the channel statistics are known, the instantaneous channel matrix can be obtained by applying the Bayesian minimum mean square (MMSE) estimator on the received signals [7]. On the other hand, minimum

variance unbiased (MVU) estimator can be used if the channel statistics is not available [8]. In these channel estimators, a linear system of equations is solved, which actually multiplies the received pilot signal with an inverse of the covariance matrix [9]. This gives rise to cubic computational complexity, and is extremely computationally expensive for Massive MIMO systems.

In order to address this complexity issue, a polynomial expansion technique, coined as PEACH (Polynomial Expansion Channel) estimators, has been introduced in [9]. A set of low complexity channel estimator is obtained based on approximating the inversion of covariance matrices in the MMSE estimator by an L-degree matrix polynomial. It is shown that as the value of L becomes large, the PEACH estimator converges to the MMSE estimator, and L does not scale with the system dimension. Therefore, these estimators are supposed to be beneficial for practically large systems.

In estimating the channel, most the previous works are based on time division duplex (TDD) operation which assumes channel reciprocity and reciprocity calibration. If frequency division duplex (FDD) is employed, the channel estimation process becomes very complex because the MIMO channel sounding requires a great deal of overhead which increase with the number of antennas. Here, a dedicated feedback mechanism for the receiver to report the channel state information is required. In [10], compressive sensing (CS) is used to reduce the CSI feedback load in large-scale MIMO systems by exploiting sparsity in the spatial frequency domain. In order in increase the efficiency, an adaptive CS based feedback scheme is also proposed where the feedback can be dynamically rearranged depending on the channel condition.

A joint spatial division and multiplexing (JSDM) approach is adopted in [11] in order to achieve the throughput gain and system operation simplification similar to Massive MIMO. It is shown that by JSDM, users can be segmented into groups with similar transmit correlation. The tall unitary structure among the transmit correlation matrices can be satisfied by the system geometry.

2.2 Parametric Channel Estimation

Channel estimation can also be conducted by estimating channel parameters, where spatial correlation of wireless channel is exploited for the better accuracy of the estimation. For a calibrated system, it has been observed that parametric approach-based channel estimation outperforms the simple unstructured interpolation schemes. In [12], a parametric channel estimation method is presented for sparse multipath fading channel using pilot subcarrier. The channel frequency response is modeled as the Fourier Transform of multipath finite impulse response. The estimator estimates the channel parameters such as delays, gains, and phases of the paths.

The uplink sounding reference signals go through a series of reflection, refraction and diffraction before arriving at the base station, where the received signal is the superposition of many resolvable signals coming from different signal paths. This parametric channel model is also the basis for the virtual channel representation for MIMO systems introduced in [13]. For a 3D massive MIMO system where there are P resolvable paths, the received signal at the (m, n)th antenna element, $1 \leq m \leq M$ and $1 \leq n \leq N$, of an $M \times N$ antenna array, can be expressed as [14]:

$$x_{m,n}=\sum_{i=1}^{P} s\alpha_i e^{-jM-12u_i+N-12v_i} e^{jm-1u_i+n-1v_i}+w_{m,n} \quad (1)$$

Where s is the transmitted signal, α_i denotes the complex channel gain, θ_i denotes the elevation DoA, and ϕ_i denotes the azimuth DoA of the ith path, $1 \leq i \leq P$. Here, $u_i=2\pi d r\cos\theta_i/\lambda$, $v_i=2\pi d r\sin\theta_i\cos\phi_i/\lambda$, $w_{m,n}$ represents the additive white Gaussian noise (AWGN), and λ is the wavelength.

If we denote the received data matrix of the antenna array by $X \in C^{M \times N}$, then X can be expressed as:

$$X=\sum_{i=1}^{P} s\alpha_i e^{-jM-12u_i+N-12v_i} \mathbf{a}(u_i)\mathbf{a}^T(v_i)+N \quad (2)$$

where $\mathbf{a}(u_i)=[1, e^{ju_i}, \ldots, e^{jM-1u_i}]^T$, $\mathbf{a}(v_i)=[1, e^{jv_i}, \ldots, e^{jN-1v_i}]^T$, and N is the AWGN noise matrix. The vectors $\mathbf{a}(u_i)$ and $\mathbf{a}(v_i)$ can be viewed as the steering vectors of elevation angle and azimuth angle respectively. The Cramer Rao Lower Bound (CRLB) for various antenna configurations is also evaluated in [14]. For the SNR ranging from -6 dB to 25 dB it is observed that Mean Square Error (MSE) reduces as the SNR increases (Fig. 2).

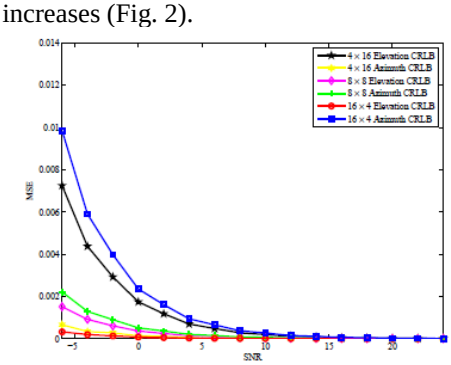


Fig. 2. CRLB for various Antenna configuration.

The root mean square error (RMSE) of the angle estimation of various base station antenna array configuration can be shown as in Fig. 3. Here it is

assumed that the base station is at a height of 35 meters while the mobile station is at 3.5 meters height. For the underlying 3D massive MIMO systems, it is generally assumed that the spacing between two adjacent antenna elements is equal to half wavelength.

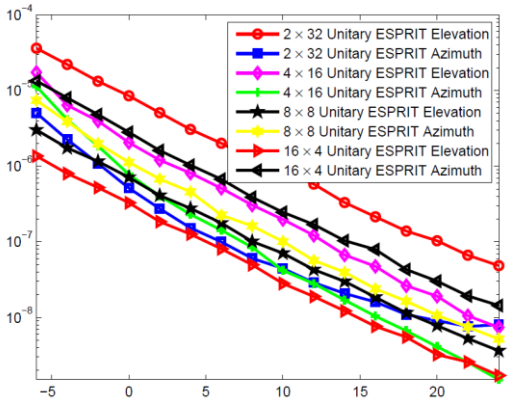


Fig. 3. RMSE of 3D DoA Estimation for Various Antenna Configuration

We can see from Fig. 3 that the performances of elevation angle estimation of different antenna array configuration are almost parallel to each other. The MSE decreases as the SNR increases. However, it is interesting to note that the performance of azimuth angle estimation does not scale proportionally to the number of antennas horizontally. We observe that the MSE of azimuth angle estimation of a 2 × 32 array is even larger than that of a 4 × 16 array, which seems a little bit counter-intuitive. The reason for this phenomenon to happen is because azimuth angle estimation is actually coupled with elevation angle estimation. In the case of 2 × 32 antenna configuration, the performance of elevation angle estimation is so poor that it hits that of azimuth angle estimation.

In scenarios like that of massive MIMO systems, where the base stations are equipped with hundreds of antennas, because of the narrow angular spread, it becomes very difficult for the base stations to perform the beam-forming by acquiring channel state information. It is necessary to conduct the beam-forming in both azimuth and elevation domain. Under parametric channel modeling, the estimation of channel essentially becomes estimation of direction of arrival (DoA) or direction of departure (DoD) and delay of resolvable paths.

There exist many subspace based techniques such as Multiple Signal Classification (MUSIC), Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) or Matrix Pencil technique for estimating the DoA/DoD for a 2D mobile wireless system. However, its counterpart in 3D has not yet been well explored in the literature of wireless communication. The multipath propagation channel is characterized not only by the DoA's but also by the time delays of different propagation paths.

In [15], an iterative algorithm is presented for joint estimation of the time delays and direction of arrivals of overlapping reflections of a known signal. This algorithm approximates the maximum likelihood estimator (MLE), and it is shown that the accuracy of resulting DoA surpasses the Cramer Rao Bound (CRB) of the DoA-only estimation. However, due to the iterative process, this algorithm has very high computational complexity. A technique of relatively low complexity for estimating known signal parameters with the assumption of asynchronous reception of signals in the array of sensors is proposed in [16]. A disadvantage of this model is that the pairing of the 3D angles and delay cannot be determined automatically. Therefore, signals in close parameters become indistinguishable.

In [17], a low complexity and high accuracy MUSIC based method—namely, TST-MUSIC (Time-Space-Time MUSIC) is proposed which offers great performance in estimating DoA and delay of a wireless channel. It is shown that the tree-structured TST-MUSIC algorithm, in addition to rendering automatic pairing of the estimated delay and DoA, is able to resolve incoming rays with very close DoA or very close delays. However, this TST-Music method still has a higher complexity and it can only be applied when the parameters are very close.

Analytical performance evaluation of the standard ESPRIT first appeared in [18]. Even though, it introduces a dramatic reduction of computation and storage cost by requiring that the sensor arrays possesses a displacement invariance, the result obtained is essentially based on the distribution of eigenvectors of the sample covariance matrix. In [19], a different approach is introduced where a first order expression of the DoA statistics in terms of fundamental parameters such as array manifold, signal covariance matrix, and number of snapshot and sensors rather than singular values and vectors is presented. However, in this work, the authors only considered the 1D standard ESPRIT method with the assumption of additive white Gaussian noise. [20] gives a more general framework for the MSE analysis of multi-dimension cases, where it is shown that MSE expression only depends on the antenna array configuration as well as the second order moments of the noise.

3. MIMO 3D Beam-forming

To take advantage of the additional degrees of freedom

provided by 3D Massive MIMO systems, 3D beam-forming is introduced as one of the most important design considerations.

In a conventional 2D MIMO system, beam-forming can only adapt its transmission in the horizontal domain, thus restricting the capabilities of interference avoidance and throughput optimization. On the other hand, in 3D MIMO systems, 3D beam-forming technology combines the horizontal beam pattern adaptation together with the vertical beam pattern adaptation, making it possible to exploit an additional degree of freedom for interference avoidance and beam coordination leading to a significant increase in throughput and coverage.

In the recent years, there have been few works on 3D beam-forming. An investigation into joint-optimization of BS tilting angle and precoding design for multi-user active antenna system is carried out in [21]. With a view to maximizing the cell average rate by employing the maximum ratio transmission (MRT) as the precoding scheme, a novel vertical beam-forming technique is proposed. For the multi-user systems, the authors gave a new 3D beam-forming technique which uses the parameter separation method for active beam-forming systems, thus resulting in a huge reduction in computational complexity. It is shown that the proposed beam-forming scheme outperforms the conventional 2D beam-forming methods. A DFT based 3D beam-forming codebook-design algorithm is proposed in [22], where the vertical codebook size is decided according to the distribution of elevation angle.

A thorough performance analysis of 3D beam-forming is carried out in [23]. For both with and without coordinated 3D beam-forming, investigations are carried out for the cases of single cell scenario representing the noise-limited system with negligible inter-cell interference, and multi-cell scenarios reflecting the interference-limited system. For the single cell scenario, it is observed that the maximal spectral efficiency does not depend on the radius of the cell border, and despite of the increased path loss, spectral efficiency can be improved with 3D beam-forming. For the multi-user case, it is shown that cell edge user throughput as well as the spectral efficiency can be improved at the same time when the appropriate combinations of near- and far downtilts are chosen. Even though the direct steering method is very effective for increasing the spectral efficiency and maximizing the desired user signal at the UE, it is to be noted that the minimal possible downtilt must be restricted for avoiding extreme interference in the neighboring cells for the situation when a very flat beam is steered to a UE located close to the cell border. The DoA estimation for the 3D Massive MIMO systems is related to 3D beam-forming in [24]. To be specific, the performance of DoA estimation and its impact on 3D beam-forming is analyzed. It is shown that the optimal beam-forming will utilize both the azimuth and elevation DoA information. Furthermore, the DoA estimation errors have a significant impact on the performance of the underlying 3D Massive MIMO systems.

4. Conclusions

The massive MIMO or FD-MIMO has the potential of meeting the future data traffic evolution. Along with the coordinated multipoint and small cell, the massive MIMO will play a key role in the enhancement of the spectral efficiency with reasonable complexity, and is considered as one of the enabling technologies for the fifth generation (5G) mobile communication. Channel estimation is the main limiting factor for completely exploiting the benefits of massive MIMO. In this E-letter, we investigated different types of channel estimations falling under the headings of estimating channel matrix and parametric-based channel estimation. We have also discussed about the MIMO 3D beam-forming technology which possesses a great potential of exploiting the benefits of massive MIMO system in both elevation and azimuth domains.

References

[1] ITU-D, The World in 2014: ICT Facts and Figures. International Telecommunication Union (ITU), April, 2014.

[2] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013-2018. Cisco, February, 2014.

[3] H. Holma and A. Toskala, “LTE Advanced: 3GPP Solution for IMT-Advanced,” 1st ed. Wiley, 2012

[4] Samsung, and Nokia Networks, “RP-141644 New SID Proposal: Study on Elevation Beamforming/Full-Dimension (FD) MIMO for LTE”, September 2014.

[5] 3GPP TR 36.873 V12.0.0, “Study on 3D channel model for LTE”, September 2014.

[6] E. G. Larsson, O. Edfors, and T. L. Marzetta, “Massive MIMO for Next Generation Wireless System,” *IEEE Commun. Mag*., Vol. 52, no. 2, pp 186-195, 2014.

[7] N. Shariati and J. Wang, “Robust Training Sequence Design for correlated MIMO Channel Estimation,” *IEEE Trans. Signal Process.*, vol. 62, no.1, pp. 107-120, 2014.

[8] S. M. Kay, “Fundamentals of Statistical Signal Processing: Estimation Theory,” Prentice Hall, 1993.

[9] N. Shariati, E. Bjornson, M. Bengtsson, and M. Debbah, “Low Complexity Polynomial Channel Estimation in Large Scale MIMO with Arbitrary Statistics,” *IEEE Journal of Selected Topics on Signal Processing*. vol. 8, no. 5, pp. 815-830, 2014.

[10] P. H. Kuo, H. T. Kung, and P. Ting, “Compressive Sensing based Channel Feedback Protocols for Spatially Correlated Massive Antenna Arrays,” in Proc. *IEEE* WCNC, Paris, France, pp. 492-497, 2012.

[11] J. Nam, J. Ahn, A. Adhikary, and Giuseppe Caire, “Joint

Spatial Division and Multiplexing: Realizing Massive MIMO Gains with Limited Channel State Information," in Proc. *IEEE* CISS, Pacific Grove, CA, USA, pp. 1-6, 2013.

[12] B. Yang, K. Letaief, R. Cheng, and Z. Cao, "Channel Estimation for OFDM Transmission in Multipath Fading Channels Based on Parametric Channel Modeling," *IEEE Trans. on Communi.*, vol. 49, no. 3, pp. 467-479, 2001.

[13] Raghavan, V.; Sayeed, A.M., "Sublinear Capacity Scaling Laws for Sparse MIMO Channels," *IEEE Trans. on Inf. Theory*, vol.57, no.1, pp.345,364, Jan. 2011

[14] A. Wang, L. Liu, and J. Zhang, "Low Complexity Direction of Arrival (DoA) Estimation for 2D Massive MIMO Systems," *IEEE* GLOBECOM International Workshop on Emerging Technologies for the LTE-Advanced and Beyond-4G, 2012.

[15] M. Wax and A. Leshem, "Joint Estimation of Time delays and Directions of Arrival of Multiple Reflections of a Known Signal," *IEEE Trans. on Signal Process.*, vol. 45, no. 10, pp. 2477-2484, 1997.

[16] A. L. Swindlehurst, "Time-Delay and Spatial Signature Estimation Using Known Asynchronous Signals," *IEEE Trans. on Signal Process.*, vol. 46, no. 2, pp. 449-461, 1998.

[17] Y. –Y. Wang, J. –T. Chen, and W. –H. Fang, "TST-MUSIC for Joint DOA –Delay Estimation," *IEEE Trans. on Signal Process.*, vol. 49, no. 4, pp. 721-729, 2001.

[18] R. Roy and T. Kailath, "ESPRIT- Estimation of Signal Parameters via Rotational Invariance Techniques," *IEEE Trans. on Acoustics, Speech and Signal Process.*, vol. 37, no. 7, pp. 984-995, 1989.

[19] F. Li, H. Liu, and R. J. Vaccaro, "Performance Analysis for DOA Estimation Algorithms: Unification, Simplification, and Observation," *IEEE Trans. Aerospace and Electronic Systems*, Vol. 29. No. 4, pp. 1170-1184, 1993.

[20] F. Roamer, M. Haardt, and G. D. Galdo, "Analytical Performance Assessment of Multi-Dimensional Matrix- and Tensor-Based ESPRIT Type Algorithms," *IEEE Trans. on Signal Process.*, Vol. 62, no. 10, pp. 2611-2625, 2014.

[21] W. Lee, S. –R. Lee, H. –B. Kong, S. Lee, and I. Lee, "Downlink Vertical Beamforming Designs for Active Antenna Systems," *IEEE Trans. on Communi.*, Vol. 62, no. 6, pp. 1897-1907, 2014.

[22] Y. Yuan, Y. Wang, W. Zhang, and F. Peng, "Separate Horizontal & Vertical Codebook Based 3D MIMO Beamforming Scheme in LTE-A Networks," *IEEE* 78th Vehicular Technology Conference, pp. 1-5, 2014

[23] H. Halbauer, S. Saur, J. Koppenborg, and C. Hoek, "3D Beamforming: Performance Improvement for Cellular Networks," *Bell Labs Technical Journal*, Vol. 18, no. 2, pp. 37-56, 2013.

[24] L. Liu, Y. Li, and J. Zhang, "DoA Estimation and Achievable Rate Analysis for 3D Millimeter Wave Massive MIMO Systems," the *15th IEEE International Workshop on Signal Processing Advances in Wireless Communications*, June 2014.

Rubayet Shafin was born in Bangladesh in 1990. He was educated at Bangladesh University of Engineering and Technology (BUET), where he received his Bachelor of Science in Electrical and Electronic Engineering degree in 2013. After his graduation, he worked as a Research Engineer in the Institute of Information and Communication Technology (IICT), BUET.

Currently, Rubayet is pursuing his Ph.D. degree in Electrical Engineering at the University of Kansas, and working as a Graduate Research Assistant in the Information and Telecommunication Technology Center (ITTC). His research interests include MIMO wireless communication and array signal processing for communication systems.

Lingjia Liu (S'03–M'08) received the B.S. degree in Electronic engineering from Shanghai Jiao Tong University, China, and the Ph.D. degree in Electrical Engineering from Texas A&M University, USA. He is currently working as an Assistant Professor in the EECS Department at the University of Kansas (KU), USA. Prior to that he spent more than three years in Samsung Research America leading Samsung's work on multi-input-multiple-output (MIMO), coordinated multi-point (CoMP) transmission, and heterogeneous networks for 3GPP LTE and LTE-Advanced standards. His general research interests lie in the areas of wireless communication systems, statistical signal processing, queuing theory, information theory, with emphasis on delay-sensitive communication over wireless networks, and smart energy networks.

Dr. Liu served as Technical Program Committee (TPC) member/chairs of various international conferences including IEEE ICC, IEEE GLOBECOM, IEEE ICNC, WCSP, etc. He currently serves as an Editor for IEEE Trans. on Wireless Comm., and Associate Editors for EURASIP Journal on Wireless Comm. and Networking and International Journal on Comm. Systems.

Jianzhong (Charlie) Zhang (S'96, M'02, SM'09) is currently senior director and head of Wireless Communications Lab with Samsung Research America at Dallas, where he leads technology development, prototyping and standardization for Beyond 4G and 5G wireless systems. From Aug 2009 to Aug 2013, he served as the Vice Chairman of the 3GPP RAN1 working group and led development of LTE and LTE-Advanced technologies such as 3D channel modeling, UL-MIMO and CoMP, Carrier Aggregation for TD-LTE, etc. Before joining Samsung, he was with Motorola from 2006 to 2007 working on 3GPP HSPA standards, and with Nokia Research Center from 2001 to 2006 working on IEEE 802.16e (WiMAX) standard and EDGE/CDMA receiver algorithms. He received his Ph.D. degree from University of Wisconsin, Madison.

On Channel Acquisition for Massive MIMO System

Runhua Chen, Qiubin Gao, Hui Li, Rakesh Tamrakar, Shaohui Sun
China Academy of Telecommunication Technology (CATT)
{chenrunhua, gaoqiubin, lihui, rakesh, sunshaohui}@catt.cn

1. Introduction

MIMO (multi-input multi-output) technology is a fundamental component in any throughput-driven wireless system [1]. In recent year, much interest has been observed in very large antenna arrays at the base stations [2], i.e., massive MIMO. Theoretically, as the number of antennas gets sufficiently large, interference between co-scheduled users and the impact of fast fading vanish, while the simplest precoder and receiver (e.g., MRT and MRC) become optimal [2].

The advantages of massive MIMO reply on accurate acquisition of CSI at the base stations. In FDD systems, user equipment (UE) provides quantized CSI through feedback channel. The overhead, including DL reference symbol overhead and UL CSI feedback overhead, scales with the number of antennas and the quantization methods greatly impact the CSI accuracy. This renders the use of very large antenna arrays quite challenging. In TDD systems where UL/DL channel reciprocity can be exploited, DL channel can be acquired by simple processing of estimated uplink channel. Furthermore, UE quantization is avoided and better CSI accuracy is possible.

In this paper, we discuss CSI acquisition methods for very large antenna arrays targeting LTE standard. The implementation issue of channel reciprocity based feedback in TDD is discussed and a two-dimension feedback mechanism for FDD is proposed. The potential of channel reciprocity feedback in TDD is demonstrated in the evaluations results. The proposed feedback mechanism for FDD is also shown to be feasible by system-level evaluation compliant with 3GPP simulation methodologies.

2. System Model

We present the system model of a massive MIMO system where a 2D planar antenna array is equipped at each eNB. An OFDM based air-interface is assumed where the system bandwidth is divided into L subcarriers. A multi-cell topology is considered where the network comprises K cells. Without loss of generality, we consider cell 0 as the serving cell, and the received signal of a UE located in cell 0 at subcarrier k is given by

$$\mathbf{y}_0^k = \mathbf{H}_0^k \mathbf{w}_0^k \mathbf{x}_0^k + \sum_{l=1}^{K} \mathbf{H}_l^k \mathbf{w}_l^k \mathbf{x}_l^k + \mathbf{n}_0^k, \qquad (1)$$

wherein

- $\mathbf{y}_0^k$ is an $N_r \times 1$ received signal vector, N_r is the number of receive antennas;
- $\mathbf{H}_l^k$ is an $N_r \times N_1 N_2$ complex channel response matrix from eNB l to the UE in cell 0, where N_1, N_2 denote the number of transmit antennas in the elevation (i.e. vertical) and azimuth (i.e. horizontal) domains, respectively. Herein $\mathbf{H}_l^k$ reflects the composite channel response comprising both large scale pathloss and shadowing fading as well as small-scale fading;
- n is an $N_r \times 1$ complex additive white Gaussian noise vector with element-wise variance σ^2;
- $\mathbf{w}_l^k$ is a $N_1 N_2 \times R_l$ precoding matrix at cell l, where $R_l \le \min(N_1 N_2, N_r)$ is the transmission rank in cell l, and $\mathbf{x}_l^k$ is the $R_l \times 1$ data vector for UE in cell l, subject to total power constraint $\mathrm{Trace}\left(E\left(\mathbf{x}_l^k \mathbf{x}_l^{k\,H} \right) \right) = P$.

For each cell l, the objective is to obtain the optimal beamforming matrix $\mathbf{w}_l^k$ that maximizes the system performance, based on information of the DL channel and noise/interference covariance matrix. In practical wireless system without any inter-cell coordination, beamforming for each cell is independently optimized and usually achieved by feedback of CSI measured on the downlink (DL) reference symbols. Specifically, denote the noise and interference covariance matrix seen by UE in cell 0 as

$$\mathbf{R}_{nn} == \sum_{l=1}^{K} \mathbf{H}_l^k \mathbf{w}_l^k \left(\mathbf{w}_l^k\right)^H \left(\mathbf{H}_l^k\right)^H + \sigma^2 \mathbf{I}_{N_r}, \qquad (2)$$

the optimal precoding matrix $\mathbf{w}_0^k$ for UE in cell 0, after DL channel estimation, is calculated as

$$\mathbf{w}_0^k = \sqrt{\mathbf{P}_0}\, eig\left(\mathbf{H}_0^{k\,H} \left(\mathbf{R}_{nn}\right)^{-1} \mathbf{H}_0^k \right), \qquad (3)$$

where $\mathbf{P}_0 = diag\,(P_{0,1}, P_{0,2} ... P_{0,R_0})$ is the power allocation matrix, $P_{0,r} = \max(u - 1/\lambda_{0,r}\ 0)$ is the power assigned to the r-th layer, u is the water filling level, and $\lambda_{0,r}$ is the r-th eigenvalue of the equivalent channel $\mathbf{H}_0^{k\,H} \left(\mathbf{R}_{nn}\right)^{-1} \mathbf{H}_0^k$, $r = 1,..., R_0$ Precoder $\mathbf{w}_0^k$ is then quantized and reported in the uplink feedback channel (e.g. PUSCH/PUCCH in LTE [9]).

Prevalent CSI quantization framework often

replies on codebook-based feedback, where $\mathbf{w}_0^k$ is quantized in the form of a recommend precoding matrix indicator (PMI) in a pre-defined codebook. The aforementioned process then mandates two robust system requirements, e.g. DL reference symbol for channel estimation, and feedback codebook for channel quantization. With a very large number of antennas, the DL pilot overhead will significantly impact the system throughput, whereas codebook design becomes increasingly complicated. How to efficiently acquire CSI at the eNB is therefore a major challenge. Compared to FDD system, a major advantage is found in TDD system where UL/DL channel reciprocity can be exploited to substantially reduce the feedback overhead and UE complexity. The rest of this paper is dedicated to the issue of CSI acquisition for massive MIMO, in both TDD and FDD networks.

3. CSI Acquisition Method

In this section we discuss practical CSI acquisition methods in the context of massive MIMO deployment.

UL/DL recipriocity with realisitic interference feedback - TDD

If UL/DL reciprocity holds, the DL channel $\mathbf{H}_0^k$ measured in the uplink, the eNB can readily calculate the optimal precoding matrix $\mathbf{w}_0^k$ (3).

It is worth noting that the majority of exiting literature assumes ideal UL/DL reciprocity including the channel $\mathbf{H}_0^k$ and interference covariance $\mathbf{R}_{nn}$. Although channel reciprocity holds in TDD, interference covariance $\mathbf{R}_{nn}$ in (2) cannot be estimated by eNB in the uplink and needs to be reported. In one example, the average received noise/interference power P_{int} can be reported as

$$P_{\text{int}} = avg(diag(\mathbf{R}_{nn})) , \tag{4}$$

This incurs feedback of a single quantized real-value scalar P_{int}, whose overhead is equivalent to conventional CQI.

Another alternative is to report the quantized diagonal component of the interference matrix as

$$\overline{\mathbf{R}}_{nn} = diag(\mathbf{R}_{nn}) , \tag{5}$$

resulting in an overhead of N_r real-value scalars. Given that practical UE has a rather small number of receive antennas (e.g. $N_r = 2$ receive chains in most commercial UE), the feedback overhead and UE complexity is still well acceptable.

Two-dimension CSI quantization - FDD

For FDD system with a very large number of eNB antennas (e.g. $N_1 N_2$ =64 antennas), the most straightforward approach is to design a high-dimension codebook of size- $N_1 N_2$, and report a single PMI corresponding to the total array size $N_1 N_2$. However, codebook design is generally a very complicated task and lacks future extendibility. Secondly, to ensure satisfactory accuracy, massive MIMO requires a gigantic quantization codebook, which increases the PMI search complexity, UE power consumption and UL overhead tremendously.

To circumvent these problems, in this paper we study a practical two-dimension feedback approach, where the UE quantizes the elevation domain CSI and azimuth domain CSI jointly to maximize the system throughput. Specifically, the UE reports two PMIs, where PMI_1 indicating the elevation domain, and PMI_2 indicating the horizontal domain.

First it is assumed that the system has two pre-defined codebooks C_1 and C_2 for elevation and azimuth quantization respectively. Codebook C_1 comprises a set of $|C_1|$ complex precoding matrices $C_1 = \{\mathbf{W}_{N_1 \times r_1}, r_1 = 1,...N_1\}$, where each precoder has N_1 rows, and r_1 is the number of columns (e.g. rank). Likewise, codebook C_2 is denoted as $C_2 = \{\mathbf{W}_{N_2 \times r_2}, r_2 = 1,...N_2\}$.

For simplicity of notation, the subscript *l* denoting the cell index is omitted in the ensuing discussion. It is assumed that DL reference symbol allows channel estimation of the full DL channel estimation $\overline{\mathbf{H}}^k$. The UE then jointly selects the optimal precoder pair $\{\mathbf{w}_1 \in C_1, \mathbf{w}_2 \in C_2\}$, such that the total throughput for serving cell *0* is maximized as

$$(\mathbf{w}_{1,opt}, \mathbf{w}_{2,opt}) = \arg\max_{\left\{\substack{\mathbf{w}_1 \in C_1, \mathbf{w}_2 \in C_2, \\ r_1 = 1,... N_1, r_2 = 1,... N_2}\right\}} \sum_{r=1}^{r_1 r_2} \log_2(1 + CQI_r). \tag{6}$$

Herein CQI_r is the post-beamforming CQI for the *r*-th data stream, $r = 1,...r_1 r_2$. For instance for a UE with MMSE-IRC receiver, CQI_r can be calculated as

$$CQI_r = \frac{1}{\left(\left(\mathbf{w}^{k\,H} \mathbf{H}^{k\,H} (\mathbf{R}_{nn})^{-1} \mathbf{H}^k \mathbf{w}^k + \mathbf{I}_{R\,,}\right)^{-1}\right)_{r,r}} - 1, \tag{7}$$

and $\mathbf{w}^k$ denotes the composite 3D-MIMO precoding matrix of dimension $N_1 N_2 \times R$ corresponding to each $\{\mathbf{w}_1 \in C_1, \mathbf{w}_2 \in C_2\}$ combination denoted as

$$\mathbf{w}^k = \mathbf{w}_1 \otimes \mathbf{w}_2 \tag{8}$$

Herein $\otimes$ represents the Kronecker product, $(.)_{r,r}$ denotes the *r*-th diagonal element.

Intuitively, it can be seen that this two-dimension feedback involves searching two individual precoders $\mathbf{w}_1$ of dimension N_1 and $\mathbf{w}_2$ of dimension N_2 in two codebooks. The Kronecker product of $\mathbf{w}_1/\mathbf{w}_2$ is specifically based on the assumption of 2D planar antenna arrangement. Such a two step precoder method provides several practical benefits for system implementation. Firstly, a much smaller codebook can be used for each spatial dimension separately, instead of reporting the full channel with a gigantic codebook. It then becomes possible to use the existing LTE codebooks for each dimension. Secondly, PMI1 and PMI2 can be reported separately, making it possible to reuse the existing CSI feedback channel (e.g. PUCCH/PUSCH in LTE) with very minor system impact.

Antenna virtualization

Massive MIMO can be deployed as an eNB implementation technique, transparently from the UE's perspective. Specifically, the N_1 physical antennas in each column can be virtualized into a single virtual antenna port and associated with one reference pilot. From the UE's perspective, each column of physical antennas appear as a single virtual antenna because they all transmit the same pilot symbols, hence the eNB antenna panel appears as a conventional N_2 -antenna array to the UE.

4. Performance evaluation

TABLE I. EVALUATION ASSUMPTION

eNB antenna configuration	Horizontal: 8，X-pol (+/-45), 0.5λ Vertical: 10，0.5λ
UE antenna configuration	2, X-pol (0/+90)
Scenarios	3D-UMi 3D-UMa [8]
Bandwidth and carrier	10MHz
Carrier frequency	2GHz
CSI feedback	Subband size (6PRB) 5ms feedback periodicity
MU-MIMO maximum number of pairable UEs	2/4/6/8
Traffic model	Full buffer

System-level simulation results are presented to demonstrate the performance with various CSI acquisition methods discussed in the previous sections. DL/UL reciprocity is assumed for TDD system and the diagonal components of interference matrix are fed back by UE. Two-dimension CSI quantization and antenna virtualization schemes are evaluated for FDD system. For antenna virtualization, each column of antenna array is virtualized to one vertical antenna port using the following beamforming vector

$$\mathbf{w}_m = \frac{1}{\sqrt{N_1}}\exp(-j\frac{2\pi}{\lambda}(m-1)d_v\cos\theta_{etilt}), m = 1,\cdots,N_1 \quad (9)$$

where N_1 is the number of physical antennas in the elevation domain (Fig. 1), λ and d_v denote the wave length and antenna elements spacing. The electric antenna down-tilting θ_{etilt} equals 102 degree in the simulation. $\pm 45^\circ$ antenna polarization is assumed with 8 horizontal antenna elements in each row and 10 vertical antenna elements in each column.

The simulated scenario is compliant with the 3GPP system-level simulation guideline with 19 cell sites, 57 sectors and 10 UEs dropped per sector. Urban macro (UMa) and urban micro (UMi) spatial channel models recently endorsed in 3GPP for massive MIMO evaluation [8] are used to realistically capture the spatial domain propagation property. Dynamic SU/MU-MIMO switching is assumed as the transmission scheme. For MU-MIMO, the number of paired UEs in any subframe is dynamically computed by the eNB, upper bounded by the maximum number of pairable UEs in Table I. For two-dimension CSI feedback, LTE Rel-10 codebook is used in the horizontal domain and DFT codebook is used in the elevation domain.

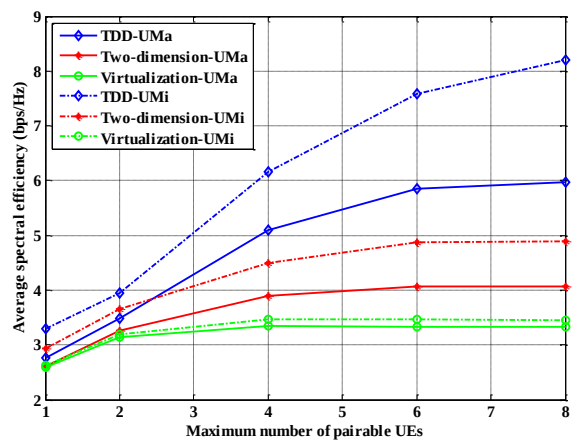


Figure 1. Cell average SE

The cell average spectral efficiency (SE) and 5% cell-edge SE are evaluated for each CSI acquisition methods and channel models. Fig.2 illustrates the cell average SE results, where 'TDD', 'Two-dimension' and 'Virtualization' denote TDD system with reciprocity, FDD system with two-dimension CSI quantization and antenna virtualization, respectively. According to Fig.2, the cell average SE is a monotonically increasing function of the maximum

number of pairable UEs; however, for FDD, the average SE quickly saturates when the maximum number of pairable UEs becomes larger than 6. Comparing different CSI acquisition methods, TDD system based on reciprocity achieves the best average SE performance. When the maximum number of pairable UE is 6, TDD system demonstrates 44% gain compared to two-dimension CSI quantization in UMa and 56% gain in UMi channel. For FDD system, two-dimension CSI quantization scheme outperforms antenna virtualization scheme, achieving throughput gain of 22% gain in UMa and 41% gain in UMi.

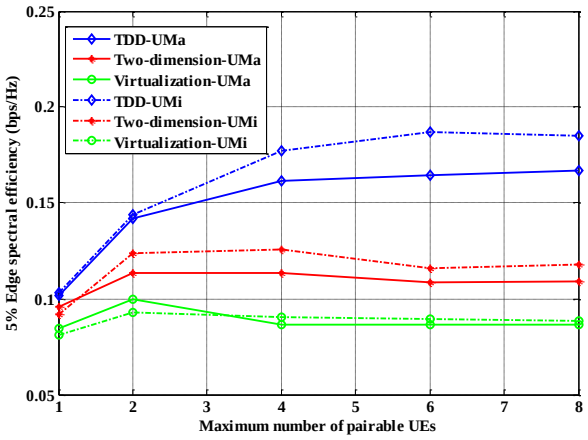


Figure 2. 5% cell-edge SE

5. Conclusions

This paper studied CSI acquisition for massive MIMO system in LTE. For deployment scenarios where UL/DL reciprocity holds (e.g. TDD), the DL channel can be inferred from UL measurement, allowing greater beamforming flexibility, reduced system overhead, and lower UE complexity. For scenarios without UL/DL reciprocity (e.g. FDD), a two-dimension feedback mechanism can achieve good trade-off between feedback complexity and performance.

References

[1] S. Sun, Q. Gao, Y. Peng, Y. Wang, L. Song, "Interference management through CoMP in 3GPP LTE-Advanced networks", *IEEE Wireless Communications Magazine*, vol. 20, no. 1, pp. 59-66, 2013.

[2] T. L. Marzetta, "Non-cooperative cellular wireless with unlimited numbers of base station antennas", *IEEE Trans. Wireless Communications*, vol. 9, no. 11, pp. 3590 – 3600, Nov. 2010.

[3] H. Q. Ngo, E. G. Larsson, T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. on Communications*, vol. 61, no. 4, pp. 1436-1449, 2013.

[4] J. Hoydis, S. Brinkz and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?", *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160-171, 2013.

[5] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. on Wireless Communications*, no. 99, pp. 2640–2651, Aug. 2011.

[6] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. Asilomar Conf. on Signals, Systems and Computers*, Urbana-Champaign, Illinois, US, Sep. 2006, pp. 359–363.

[7] 3GPP RP-141644, "New SID Proposal: Study on Elevation Beamforming/Full-Dimension (FD) MIMO for LTE", Samsung.

[8] 3GPP TR 36.873, "Study on 3D channel model for LTE."

[9] 3GPP TS 36.213, "3GPP E-UTRA: Physical layer procedures."

Runhua Chen received his Ph.D. from The University of Texas at Austin, M.Phil. from The Hong Kong University of Science and Technology, and B.E. from Tsinghua University, all in Electrical Engineering. He is currently a Senior Expert Consultant at Datang Telecom, working on 3GPP LTE standardization. Previously he has been with the Wireless Basestation Infrastructure Business Unit and DSPS R&D Center at Texas Instruments, and AT&T Labs.

Qiubin Gao received his B.S. and Ph.D. in control science and engineering from Tsinghua University. He is currently a senior research engineer at Datang Wireless Mobile Innovation Center of China Academy of Telecommunications Technology (CATT). His current research interests include physical layer design for mobile communication, multiple-antenna technology, CoMP, D2D, and system performance evaluation.

Hui Li received his B.E. degree from Beijing Jiaotong University, and Ph.D. degree from Tsinghua University, China, all in electrical engineering, in 2002 and 2007 respectively. He joined Fujitsu R&D Center Co. Ltd., Beijing, China in 2007. Since then he has been working on signal processing in wireless communication systems. In 2014, he joined China Academy of Telecommunication Technology (CATT). His current research interests include massive MIMO and communication signal processing.

Rakesh Tamrakar is currently with Datang Wireless Mobile Innovation Center of China Academy of Telecommunications Technology (CATT).

Shaohui Sun received the M.S. degree in computer engineering and Ph.D. degree in communication and information systems from Xidian University, Xi'an, China, in 1999 and 2003, respectively. Form March 2003 to June 2006, he was a postdoctoral fellow with the China Academy of Telecommunication Technology, Beijing, China. From June 2006 to December 2010, he was with Datang Mobile Communications Equipment company Ltd., Beijing, where he has been deeply involved in the development and standardization of the Third-Generation Partnership Project Long-Term Evolution (3GPP LTE). His research areas include multiple antenna technology, heterogeneous wireless networks, and relay.

Exploiting Adaptive Downtilt and Vertical Sectorization in LTE Advanced Networks using Active Antenna Systems

Meilong Jiang, Moon-il Lee, Ananth Kini, Mohsen Hosseinian, Janet Stern-Berkowitz
InterDigital, Inc., USA
{meilong.jiang, moonil.lee, ananth.kini, mohsen.hosseinian, janet.stern}@interdigital.com

1. Introduction

Significant interest in enhancing wireless cellular network performance through the use of active antenna systems (AAS) has recently developed [1, 2], with AAS seen as an enabler for achieving both cell edge and cell average gains through advanced control of the elevation (i.e., vertical) dimension. Unlike traditional passive antennas, in AAS, the complete radio frequency (RF) transceiver chain is integrated into each antenna element, allowing the amplitude and phase of each element to be adaptively adjusted. Using a 2-dimensional (2D) array, AAS provides control for the elevation dimension as well as the traditional azimuth dimension. This control enables a variety of new advanced multiple-input multiple-output (MIMO) techniques such as adaptive downtilt, vertical sectorization, and full dimension MIMO (FD-MIMO) [3, 4].

In existing cellular networks, without AAS capability at the base station (BS), a fixed vertical downtilt (achieved by mechanical or electrical control) has typically been used. The fixed downtilt can provide a certain level of inter-cell interference reduction and elevation beamforming gain [3], but is not optimal, especially in dense urban deployment scenarios in which users can be located in a wide range of elevations and the users' distribution can change dramatically in 3-dimensional (3D) space over time. As a result, enhancement of system performance by leveraging the dynamic control of vertical beams that AAS enables has been actively studied in the Third Generation Partnership Project (3GPP) in the context of Long-Term Evolution Advanced (LTE-A).

In line with the AAS study, a new 3D channel model has recently been developed in 3GPP which introduces the channel characteristics for the elevation domain for both low and high rise scenarios. The new model, which considers departure and arrival of multipath components for elevation angles as well as for the azimuth angles supported by the conventional 2D model, enables the elevation domain to be fully exploited when optimizing and evaluating advanced AAS-based MIMO techniques.

In this letter, we evaluate the benefits of AAS using adaptive downtilt and vertical sectorization techniques in an LTE-A network. The 3D channel considered in our performance analysis follows the latest 3GPP assumptions and agreements [5].

2. Adaptive Downtilt and Vertical Sectorization

Adaptive downtilt and vertical sectorization are promising techniques for capitalizing on 2D antenna arrays and the independent control of antenna elements enabled by AAS. Adaptive downtilt dynamically optimizes the downtilt angle used in a cell according to the cell's user equipment (UE) distribution to maximize elevation beamforming gain while reducing inter-cell interference. Vertical sectorization increases spatial reuse of time and frequency resources by forming multiple vertical sectors, each with a different vertical beam. Deployment examples for adaptive downtilt and vertical sectorization are shown in Figure 1.

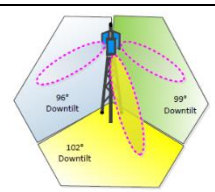
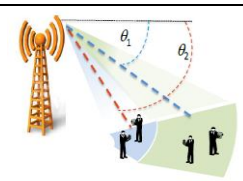
(a) Adaptive Downtilt	(b) Vertical Sectorization
	
3 cells, each with adaptive vertical downtilt	Single cell with multiple vertical sectors

Figure 1. A cellular system supporting adaptive downtilt or vertical sectorization.

For our evaluation, we consider a multi-cell downlink LTE-A system equipped with AAS. Each BS (Evolved Node B (eNB) in LTE-A) controlling one or more cells is equipped with a 2D uniform rectangular array (URA) with n_T antenna elements, where $n_T = n_H \times n_V$ and n_V and n_H denote the number of antenna elements in the vertical and horizontal domains, respectively. Each cell has uniformly distributed UEs in both horizontal and elevation domains [5], with each UE having n_R antennas. The system has adaptive downtilt capability, where each cell can apply a cell specific downtilt in the elevation domain, and vertical sectorization capability, where each cell can simultaneously form two or more vertical sectors by applying different vertical beamforming vectors to the antenna elements in the vertical domain.

2.1 Adaptive Downtilt

We consider a system with N cells using adaptive downtilt and optimize the downtilt angle for each cell to maximize wideband signal-to-interference-plus-noise-ratio (SINR) as a representative measure of system performance. Each cell is equipped with a URA which is capable of forming a vertical beam with a downtilt angle θ_n in each cell n, $n \in \{1, \dots, N\}$. The beamforming weights to form each vertical beam, which are applied as a vector to the antenna elements of the URA in the vertical domain, can be found in [6]. The SINR of the UE u served by cell n, $\gamma_{u,n}$, is given by

$$\gamma_{u,n} = \frac{g_{u,n}(\theta_n)}{\sigma_N^2 + \sum_{\substack{n'=1 \\ n' \neq n}}^{N} g_{u,n'}(\theta_{n'})}, \qquad (1)$$

where $g_{u,n}(\theta_n)$ denotes the received signal power measured at UE u from its serving cell n. In (1), the denominator consists of the noise power σ_N^2 and the co-channel interference from the interfering cells $n' \in \{1, \dots, N\}, n' \neq n$. We observe from (1) that $\gamma_{u,n}$ is a function of downtilt θ_n as well as the downtilt angles from all interfering cells n'. Using SINR for optimization therefore requires joint optimization across the whole system.

The system-wide optimization problem can be very complex due to the fact that all cells are coupled. We therefore use an approximation which is a relatively simple and efficient closed-loop scheme. We first assume a candidate downtilt set consisting of J downtilt angles, $\hat{\theta}_{set} = [\hat{\theta}_1, \dots, \hat{\theta}_J]$, from which downtilt angles may be selected in each cell. We then restrict the joint design to be over a neighboring cell set (a subset of the system set), $S_{set} = [1, \dots, S]$ where $S < N$ since the interference from the neighboring cells will be the dominant interference. For other interfering cells not belonging to the neighboring cell set, the downtilt angles can be randomly chosen as one entry of $\hat{\theta}_{set}$ to approximate the interference. Within the neighboring cell set, the optimal downtilt angle for each cell can now be obtained by maximizing the sum capacity based on wideband SINR, $\gamma_{u,s}$, in which inter-cell interference is limited to the neighboring cell set, i.e.,

$$\left[\theta_1^{opt}, \dots, \theta_s^{opt}, \dots, \theta_S^{opt}\right] = \arg\max_{\theta_s, s=1,\dots,S} \sum_{s=1}^{S} \sum_{u=1}^{U_s} \log_2\left(1 + \gamma_{u,s}\right), \qquad (2)$$

subject to $\theta_s \in \hat{\theta}_{set}$ and $s \in S_{set}$, where S is the cardinality of the neighboring cell set and U_s is the number of UEs within cell s.

We can obtain the solution for (2) by exhaustively searching among the neighboring cell set S_{set} and the downtilt candidate set $\hat{\theta}_{set}$. For this scheme, coordination among cells can be limited to coordination of downtilt angles among the cells of a neighboring cell set, the size of which can be small to reduce the complexity. In our simulation, we assume $S = 3$ and evaluate performance for $J = 2, 3, 4$. In the end, by repeating the process for each non-overlapping neighboring cell set, S_{set}, the downtilt angles for the whole system can be obtained.

To achieve an adaptive downtilt implementation in an LTE-A system, we use the procedure shown Figure 2, involving UE feedback of the existing LTE-A measurement of reference signal received quality (RSRQ) which can be used to approximate wideband SINR in a practical system. As shown in the figure, the eNB sends reference signals applying each downtilt angle from a pre-defined downtilt candidate set. UEs measure RSRQ for each candidate downtilt and feed this information back to the eNB. Finally the eNB selects the best downtilt angle to maximize the sum capacity according to (2).

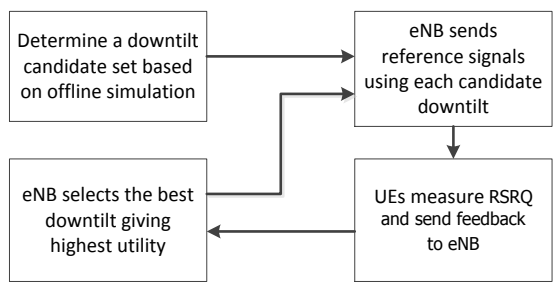


Figure 2: An adaptive downtilt procedure in an LTE-A network.

Table I shows the performance benefit of the adaptive downtilt over the fixed downtilt with respect to cell average and cell edge spectral efficiencies. The system level simulation assumptions are summarized in the attached Annex.

Table I: Cell average and cell edge spectral efficiency of fixed and adaptive downtilt in bits/s/Hz and % gain.

Fixed downtilt				
Fixed downtilt	98°	100°	102°	104°
Cell average	1.96 (-10%)	2.10 (-3.7%)	2.18 (0%)	1.82 (-16%)
Cell edge	0.075 (-0.1%)	0.059 (-20%)	0.074 (0%)	0.053 (-28%)
Adaptive downtilt				
Downtilt candidate set	[100°,102°]	[98°,100°,102°]		[96°,98°, 100°,102°]
Cell average	2.25 (3.2%)	2.28 (+4.6%)		2.29 (+5.1%)
Cell edge	0.076 (+2.7%)	0.080 (+8.1%)		0.082 (+11%)

The performance for three adaptive downtilt candidate sets and four fixed downtilts are compared where gain

(in percentage) is with respect to the baseline of a fixed downtilt of 102°. From the results, it can be seen that the adaptive downtilt scheme achieves up to 5% cell average and 11% cell edge spectral efficiency gain over the fixed downtilt baseline.

2.2 Vertical Sectorization

We now consider a system using vertical sectorization in which there are N cells and M vertical sectors for each cell. The vertical sector m in cell n can be formed with a vertical beamforming vector with elevation angle θ_n^m; thus, M vertical beamforming vectors with different elevation angles are required to form M vertical sectors. Since multiple vertical beams are transmitted through the same set of transmit antennas, the total transmission power should be split into M vertical sectors with a power split ratio which can be flexible.

A UE served by a cell n with a vertical sector m experiences interference from the other cells $n' \in \{1, \dots, N\}, n' \neq n$ and the other vertical sectors $m' \in \{1, \dots, M\}, m' \neq m$ in the same cell n. Assuming equal power split for M vertical sectors, the wideband SINR of UE u in cell n served by the vertical sector m, $\gamma_{u,n,m}$, is given by

$$\gamma_{u,n,m} = \frac{\frac{1}{M} g_{u,n,m}(\theta_n^m)}{\sigma_N^2 + \frac{1}{M}\sum_{\substack{m'=1 \\ m' \neq m}}^{M} g_{u,n,m'}(\theta_n^{m'}) + \sum_{\substack{n'=1 \\ n' \neq n}}^{N} \tilde{g}_{u,n'}}, \qquad (3)$$

where $\frac{1}{M} g_{u,n,m}(\theta_n^m)$ denotes the received signal power measured at UE u from its serving vertical sector m in cell n. In (3), the denominator consists of the noise power σ_N^2, the inter-sector interference in the same cell $\frac{1}{M}\sum_{\substack{m'=1 \\ m' \neq m}}^{M} g_{u,n,m'}(\theta_n^{m'})$, and the sum of the inter-cell interference from each cell $n', \tilde{g}_{u,n'}$. Inter-cell interference $\tilde{g}_{u,n'}$, can be expressed as the summation of interferences from all vertical sectors in cell n', i.e., $\tilde{g}_{u,n'} = \frac{1}{M}\sum_{m=1}^{M} g_{u,n',m}(\theta_{n'}^m)$.

For simplicity, we consider a system with two vertical sectors ($M = 2$) respectively formed with elevation angles $\theta_n^1 = \theta_1$ (outer sector) and $\theta_n^2 = \theta_2$ (inner sector) as shown in Figure 1(b). In order to achieve the desired spatial reuse gain, the elevation angles for the sectors should be properly chosen.

Table II shows the system performance according to various sets of elevation angles when an MMSE receiver is used at a UE. As seen in the table, the vertical sectorization with the elevation angle set ($\theta_1 = 99^\circ, \theta_2 = 102^\circ$) provides the best cell average performance among the sets in the system we considered.

Table II: System performance according to the set of elevation angles for two vertical sectors in bits/s/Hz.

Vertical Sectorization (M=2)				
Elevation angle set [θ_1, θ_2]	[93°,96°]	[96°,99°]	[99°,102°]	[102°,105°]
Cell average	2.58	2.59	2.80	2.77
Cell edge	0.079	0.068	0.076	0.088

Although vertical sectorization increases signal strength per UE and spatial reuse of time/frequency resources, both of which improve system performance, it suffers from strong intra-cell interference even with single-user MIMO (SU-MIMO) operation due to the increased number of vertical sectors. To overcome the dominant interference from other vertical sectors, LTE-A interference mitigation schemes can be used such as Network Assisted Interference Cancellation and Suppression (NAICS) in which an advanced UE receiver can cancel out dominant interference using symbol level interference cancellation (SLIC). Network assistance for the cancellation process involving providing interferer information to the UE is currently being discussed in 3GPP [7]. In the case of NAICS, the receiver peformance improves when the dominant interference power becomes stronger since the estimation error for the dominant interference signal gets reduced. A system with vertical sectorization will, therefore, provide higher system throughput as the number of UEs equipped with NAICS receivers increases in the cell.

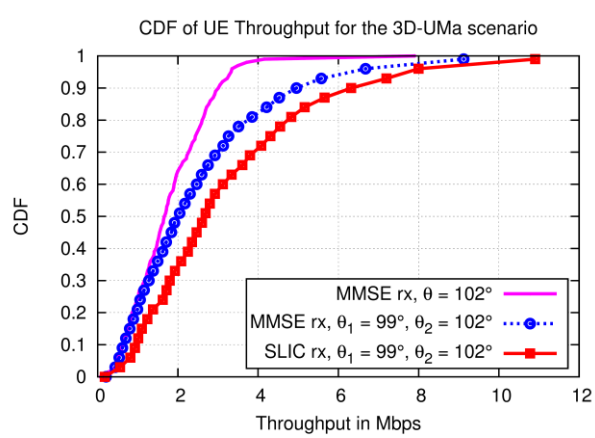


Figure 3: UE throughput performance according to the number of vertical sectors and UE receiver type.

Figure 3 shows the performance improvement that can be achieved when vertical sectorization and an advanced receiver (SLIC) are utilized. As shown in the figure, the vertical sectorization provides significant

UE throughput gain as compared with that for the single vertical sector case. In addition, the UE throughput distribution is even further improved when the NAICS receiver is used at the UE.

3. Conclusion

In this letter, we showed that adaptive downtilt and vertical sectorization are promising techniques for future wireless systems to exploit the benefit of additional control in the vertical domain enabled by AAS. When optimizing downtilt based on wideband SINR, adaptive downtilt achieves up to 11% cell edge and 5% cell average spectral efficiency gain compared to the baseline system with fixed downtilt. In addition, vertical sectorization provides even higher performance gain by increasing spatial reuse of time and frequency resources using multiple vertical beams. Use of advanced receivers and NAICS by the UE provides significant additional gain as the dominant interference from the vertical sector can be removed.

Annex

Table A: System Level Simulation Assumptions

Parameter	Assumption
Network layout	7-site 21-cell wraparound
Channel model	3D Urban Macro (UMa) [5]
eNB antenna configuration	$n_H = 4$, $n_V = 10$, $\lambda/2$ spacing in H/V, cross-polarization
UE antenna configuration	$n_R = 2$, cross-polarization
UE attachment	Based on reference signal received power formula in [6]
Adaptive downtilt angle candidates	[96°, 98°, 100°,102°]
Vertical sectorization	[93°, 96°], [96°,99°], [99°,102°], [102°,105°]
Number of UEs per cell	10
UE distribution	uniformly dropped
Traffic model	full buffer
Scheduler	proportional fair (PF)
Transmission scheme	SU-MIMO with rank adaptation
Feedback	Subband channel quality index, wideband precoding matrix index
Link adaptation	Practical with open loop link adaptation
Receiver	MMSE, NAICS

References

[25] J. Koppenborg, et al., "3D beamforming trials with an active antenna array," in 2012 International ITG Workshop on Smart Antennas (WSA), pp.110-114, Mar. 2012.

[26] 3GPP TR 37.840, "Study of RF and EMC requirements for Active Antenna Array System (AAS) base station," V12.1.0, Release 12, Jan. 2014.

[27] F. Athley, et al, "Impact of electrical and mechanical antenna tilt on LTE downlink system performance," Vehicular Technology Conference, IEEE VTC 2010-Spring, pp. 1-5, May 2010.

[28] Young-Han Nam et al. "Full-Dimension MIMO (FD-MIMO) for Next Generation Cellular Technology," pp. 172-179, June 2013.

[29] 3GPP TR 36.873, "Study on 3D channel model for LTE," V2.0.0, Release 12, Mar. 2014.

[30] M. Jiang, et al, "Efficient Adaptive Vertical Downtilt Schemes in LTE-Advanced Networks," to appear in IEEE PIMRC 2014.

[31] 3GPP TR 36.866, "Study on Network Assisted Interference Cancellation and Suppression (NAICS) for LTE Release 12," v12.0.1, 2014.

Meilong Jiang received his Ph.D from the University of Hong Kong in electrical engineering in 2006. He was a research staff member with NEC Labs America from 2008 to 2012. He then joined InterDigital where he is currently a senior staff engineer. His current research interests include 3D channel model, FD-MIMO, massive MIMO and other 5G techniques. He is a senior member of IEEE.

Moon-il Lee received his M.S. degree in electrical engineering from Ajou University, Korea, in 2005. He joined InterDigital to work in mobile communication technology research in 2011. Previously, he worked for LG Electronics, Korea. He has been actively participating in 3GPP standardization for LTE radio access technologies, primarily in the area of advanced antenna techniques since 2005. He is active in concept development and 3GPP standardization of LTE-A and future wireless technologies.

Ananth Kini received his Ph.D. in Computer Engineering from Drexel University in January 2009. His research interests have spanned several areas of Computer and Wireless Networking, including 3GPP standards evolution, wireless network multicasting and broadcasting, and ad hoc network transmission coordination. He is currently employed at InterDigital where his work focuses on researching future enhancements to 3GPP LTE standards.

Mohsen Hosseinian received his Ph.D. in electrical engineering from The University of Calgary, Canada, in 1999. He joined InterDigital in 2005, where he is currently leading simulation activities for the PHY and MAC layers of LTE and LTE-A. He previously worked at TRLabs, WiLAN, and Harris communications, where he designed and developed WCDMA, OFDM, and high-speed radio systems. His current research interests include 3D channel modeling, massive MIMO, and small cell enhancements.

Janet Stern-Berkowitz received her B.E. and M.E. in electrical engineering from The Cooper Union. She designed beamforming and signal processing systems for sonar applications at EDO Corporation until she joined InterDigital in 1998 where her work has been focused on design, development, and standardization of wireless communications systems. She is currently working in InterDigital's Innovation Labs leading research and development for the MAC and PHY layers for ongoing and future LTE-A enhancements.

User Grouping and Load Balancing for FDD Massive MIMO Systems

Yi Xu[1], Guosen Yue[2] and Shiwen Mao[1]
[1]Auburn University, Auburn, AL, USA
[2]Broadcom Corporation, Matawan, NJ, USA
yzx0010@auburn.edu, guosenyue@gmail.com, smao@ieee.org

1. Introduction

Last decades have witnessed ever-increasing demand for higher data rates in wireless networks. To cater for this demand, many advanced physical layer techniques have been developed, e.g., multiple input multiple output (MIMO) with orthogonal frequency division multiplexing (OFDM). However, with linear throughput improvement but the exponential growth on the data traffic, the gap between the demand and supply has been increasingly widened. To solve the problem, the next technology we could resort to is massive MIMO (a.k.a. large-scale MIMO, full-dimension MIMO, or hyper MIMO), which significantly increases the system capacity by employing a large number of antennas at the base station. As an emerging and promising technology, large-scale MIMO also enjoys many advantages such as low-power, robust transmissions, simplified transceiver design, and simplified multiple access layer [1], [2], in addition to enhanced capacity.

In general, the more transmit antennas, the more degrees of freedom a massive MIMO system can provide, resulting in higher reliability or larger throughput. However, due to the difficulties of acquiring channel state information at the transmitter side (CSIT), it is challenging to simultaneously support a large number of users [2]. Most of the existing works on massive MIMO systems consider the time division duplexing (TDD) mode [3]-[5], within which by exploiting channel reciprocity, the downlink channel can be estimated through uplink training. Unfortunately, there is no such privilege in frequency-division-duplexing (FDD) systems.

There are much more FDD (≥ 300) than TDD (≤ 40) LTE licenses worldwide. It is therefore of great importance to investigate the massive MIMO design for FDD systems. To reduce pilot resources and the channel state information (CSI) feedback in FDD systems, a two-stage precoding scheme has been proposed in [6] recently. Firstly, the users in service are divided into groups, while each group of users have similar second-order channel statistics (i.e., transmit correlation). The same pre-beamforming, or the first-stage precoding, is then used for each group of users semi-statically. Next, with reduced dimensions on the effective channel, simplified channel feedback can be realized and the second-stage dynamic precoding can be applied. The performance of such system design is largely dependent on user grouping. In [7], a K-means clustering scheme, based on chordal distance as the clustering metric, is introduced for user grouping. In this letter, instead of chordal distance, we propose weighted likelihood similarity measure and hierarchical clustering. By theoretical analysis and simulations, we validate the proposed approaches.

Once user groups are formed, another important issue is user scheduling, i.e., selecting users for transmission based on instantaneous channel conditions. In this letter, we propose a dynamic user scheduling method. If there are only a few active users, some groups may barely have users while some other groups are overloaded. Therefore, we also consider the load balancing problem and develop an effective solution algorithm. Note that some results can also be found in [8][9].

2. Problem Statement and Main Results

We consider a downlink system with M antennas at the base station (BS) and a single antenna at each user terminal (UT). Denote y_k as the received signal at user$k$, $k = 1,2,\ldots,K$. The signals received by all UTs y can be written as

$$y=\boldsymbol{HHV}d+z=\boldsymbol{HHBP}d+z, \quad (1)$$

where $()^H$ denotes the Hermitian of a matrix; **H**, of dimension $M \times K$, is the actual channel between the BS and the users; $\boldsymbol{V}$ is the precoding matrix of dimension $M \times S$; d is the data vector of dimension $S\times 1$; and z is the zero mean circulant symmetric complex Gaussian noise vector. The key idea is to decompose $\boldsymbol{V}$ into $\boldsymbol{B}$ and $\boldsymbol{P}$, where $\boldsymbol{B}$ is pre-beamforming matrix of dimension $M \times b$; $\boldsymbol{P}$ of dimension $b \times S$, is designed to suppress the interferences within each group.

1) User grouping schemes.

To design the pre-beamforming matrix $\boldsymbol{B}$, we need to group users. Different from the chordal distance based K-means user grouping scheme [7], we first propose a weighted likelihood function as the similarity measure between a user and a group, which is defined as:

$$\mathcal{L}(Rk,Vg)=\|(Uk\Lambda k1/2)HVg\|2 \quad (2)$$

In addition to the new similarity measure, we also propose new user grouping schemes, which employs the agglomerative hierarchical clustering method. Different from the K-means approach, which essentially looks at all possible combinations of users and groups, the agglomerative hierarchical clustering method starts with each individual user forming a user group. It then proceeds by a series of successive mergers based on certain criteria. Eventually, all users can form one single group. We can terminate the scheme when the desired number of groups is reached.

One important issue in hierarchical clustering is how to define the similarity measure or distance between existing groups and newly defined groups (called linkage methods). We propose to use weighted average linkage as follows.

$$dvi,vj,vq=(dvi,vq+dvj,vq)/2, \quad (3)$$

where v_i, v_j and vq are groups.

Given chordal distance, weighted likelihood similarity, K-means clustering and hierarchical clustering, we could combine either two of them to form a complete user grouping scheme.

Proposition 1: The complexity of K-means clustering is $O(GKKG(2M3+M2))$ for chordal distance and $O(GKKG[(r*)3+(Mr*2))$ for weighted likelihood similarity measure, where r^* is the effective rank of R_k, i.e., the number of columns in Uk.

Proposition 2: The complexity of hierarchical clustering is $O(K(K-1)(2M^3+M^2)/2)$ for chordal distance, and $O(K(K-1)[(r*)3+(Mr*2)/2)$ for weighted likelihood similarity measure.

2) User scheduling scheme.

With user groups being formed, we can obtain the pre-beamforming matrix $\boldsymbol{B}_g$ for each group g. At a particular time slot, based on the instantaneous channel conditions of the users, we dynamically schedule a subset of users in each group for the transmissions in this time slot.

In [7], a MAX and an ALL user scheduling algorithm are presented. Different from these approaches, we propose a dynamic user scheduling algorithm that schedules users in a greedy manner. In particular, at each step, the proposed algorithm only schedules the user that can achieve the largest gain in the system throughput. The proposed algorithm is presented in the Algorithm 1.

3) Load balancing scheme.

In practical applications, many users may gather at one geographic location (e.g., in a skyscraper). If we design the precoder exactly as discussed, these users will form a big group. It would be desirable to offload some of the users to other groups, to achieve fairness among the users. This is because with more members in a group, each member's chance of getting scheduled for transmission will be smaller.

User groups $\{\mathcal{S}_g\}$ are given ;
Initially set $\mathcal{U}=\{1,2,\cdots,K\}$, $\mathcal{C}=0$, and $\mathcal{K}_g=\emptyset, \forall g$;
while *Termination conditions* ($\sum_g |\mathcal{K}_g| = \sum_g b_g$, $\kappa(k^*, g_{k^*})=0$, *or* $\mathcal{U}=\emptyset$) *are not satisfied* **do**
  for $k \in \mathcal{U}$ **do**
    if $|\mathcal{K}_{g_k}| < S_g$ **then**
      Set $\mathcal{K}'_g = \mathcal{K}_g \cup \{k\}$ if $k \in \mathcal{S}_g$, and $\mathcal{K}'_{g'} = \mathcal{K}_{g'}, \forall g' \neq g$;
      Perform ZFBF or RZFBF based on $\{\mathcal{K}'_g\}$ and $\{\mathbf{B}_g\}$;
      Compute the gain $\kappa(k,g)=\max\{0, \mathcal{C}(\{\mathcal{K}'_g\},\{\mathbf{B}_g\})-\mathcal{C}(\{\mathcal{K}_g\},\{\mathbf{B}_g\})\}$;
    end
  end
  Obtain $(k^*, g_{k^*}) = \arg\max_{k\in\mathcal{U}} \kappa(k,g)$;
  if $(k^*, g_{k^*}) \neq \emptyset$ **then**
    $\mathcal{U} \leftarrow \mathcal{U}\backslash k^*$;
    $\mathcal{K}_{g_{k^*}} \leftarrow \mathcal{K}_{g_{k^*}} \cup \{k^*\}$;
  end
end

Algorithm 1. Greedy Algorithm for Dynamic User Selection and Beamforming.

We develop a user grouping method considering group load balancing and user proportional fairness. The problem can be formulated as:

$$\max_{\{x_{kg}\}} \quad \mathcal{J} = \sum_{k=1}^{K}\sum_{g=1}^{G} x_{kg} \log\left(\frac{\overline{\eta}_{g_k}}{\sum_i x_{ig}}\right)$$
$$s.t. \quad \sum_g x_{kg} = 1, \quad \forall k \in \{1,2,\cdots,K\}$$
$$x_{kg} \in \{0,1\}, \quad \forall k, g,$$

where xkg is *1* if user k is connected to group g, and *0* otherwise; $\bar{\eta}_{k_g}$ is the rate for user k in group g. By adopting a two-tier dual decomposition approach, we could obtain the optimal solution to the above problem.

3. Performance Evaluation

Simulations are performed to evaluate the proposed schemes. We fix $M = 100$ and group number G=6.

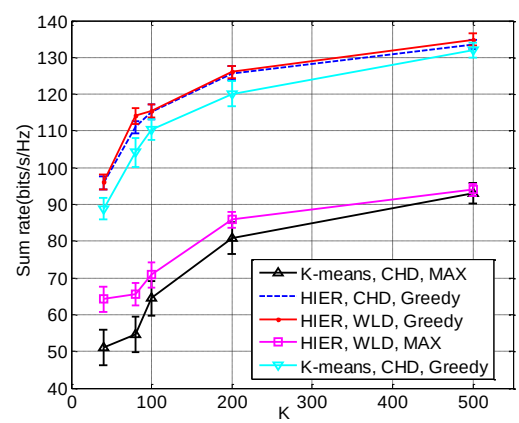


Fig. 1 System sum rate versus the number of users.

We can see from Fig. 1 that all our proposed schemes outperform the scheme in [7]. In particular, hierarchical clustering greedy user selection with weighted likelihood has the highest system throughput.

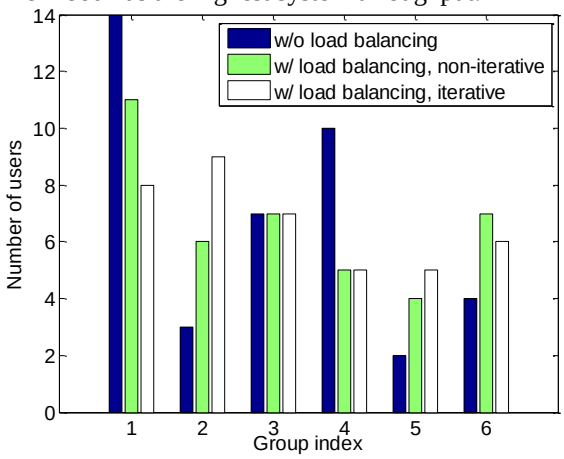


Fig. 2 Group sizes for user grouping with joint group load balancing and precoding design when $K = 40$.

We can see from Fig. 2 that the maximum difference of the scheme without considering load balancing is 12, while this number is 4 in our proposed iterative scheme. Thus the loading of users is much more balanced in our proposed scheme.

4. Conclusions

In this letter, we have studied the user grouping and scheduling problems based on a two-stage precoding framework for FDD massive MIMO systems. We have proposed weighted likelihood similarity measure and hierarchical clustering for user grouping. We have also proposed a dynamic user scheduling scheme and a user grouping algorithm to achieve load balancing and user fairness for FDD massive MIMO systems. The efficacy of the proposed schemes has been validated with analysis and simulations.

ACKNOWLEDGMENT

This work was supported in part by the U.S. National Science Foundation under Grants CNS-1247955 and CNS-1320664.

References

[1] E. Larsson *et al*, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 185--195, Feb. 2014.

[2] F. Rusek *et al*., "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40--60, Jan. 2013.

[3] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640--2651, Aug. 2011.

[4] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 192--201, Feb. 2013.

[5] J. Hoydis, K. Hosseini, S. ten Brink, and M. Debbah, "Making smart use of excess antennas: Massive MIMO, small cells, and TDD," *Bell Labs Tech. J.*, vol. 18, no. 2, pp. 5--21, Sep. 2013.

[6] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441--6463, Oct. 2013.

[7] A. Adhikary and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming and user grouping," *IEEE J. Sel. Topics Signal Process.*, vol. 8, iss. 5, pp. 876--890, Mar. 2014.

[8] Y. Xu, G. Yue, N. Prasad, S. Rangarajan, and S. Mao, "User grouping and scheduling for large scale MIMO systems with two-stage precoding," in *Proc. IEEE ICC*, Sydney, Australia, Jun. 2014, pp. 5208--5213.

[9] Y. X., G. Yue, and S. Mao, "User grouping for Massive MIMO in FDD systems: New design methods and analysis," *IEEE Access Journal, Special Section on 5G Wireless Technologies: Perspectives of the Next Generation Mobile Communications and Networking*, vol.2, pp.947-959, Aug. 2014.

YI XU (S'11) received the M.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2010, and the B.S. degree in electronic information engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2007. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA. His research interests include optimization, game theory, MIMO, OFDM, IDMA, and cognitive radio networks.

GUOSEN YUE (S'99-M'04-SM'09) received the B.S. degree in physics and the M.S. degree in electrical engineering from Nanjing University, Nanjing, China in 1994 and 1997, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, in 2004. He was a senior research staff with the Mobile Communications and Networking Research Department, NEC

Laboratories America, Princeton, New Jersey, conducting research for broadband wireless systems and mobile networks. His research interests are in the general areas of wireless communications and signal processing. In August 2013, he joined Broadcom Corporation as a system design scientist. Dr. Yue serves as an Associate Editor for the IEEE Transactions on Wireless Communications. He has served as the associate Editor for Research Letters in Communications, the Guest Editors for EURASIP Journal of Wireless Communication and Networking special issue on interference management, ELSEVIER PHYCOM special issue on signal processing and coding. He served as the Symposium Co-chair for IEEE ICC 2010, the Track Co-chair for IEEE ICCCN 2008, the steering committee member for IEEE RWS 2009. He is a senior member of the IEEE.

SHIWEN MAO (S'99-M'04-SM'09) received the Ph.D. degree in electrical and computer engineering from Polytechnic University, Brooklyn, NY, USA. Currently, he is the McWane Associate Professor with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA. His current research interests include cross-layer optimization of wireless networks and multimedia communications, with current focus on cognitive radios, femtocells, 60 GHz mmWave networks, free space optical networks, and smart grid. He is on the Editorial Board of the IEEE Transactions on Wireless Communications, IEEE Internet of Things Journal, IEEE Communications Surveys and Tutorials, and several other journals. He received the 2013 IEEE ComSoc MMTC Outstanding Leadership Award and the NSF CAREER Award in 2010. He is a co-recipient of the IEEE ICC 2013 Best Paper Award and the 2004 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems.

Massive MIMO Operation in Fronthaul-capacity Limited Cloud Radio Access Networks

Sangkyu Park, Chan-Byoung Chae †, and Saewoong Bahk**
*Department of EECS, INMC, Seoul National University
†School of Integrated Technology, Yonsei University
skaprk@netlab.snu.ac.kr, cbchae@yonsei.ac.kr , and sbahk@netlab.snu.ac.kr

1. Introduction

Massive multiple-input multiple-output (MIMO) is a key enabler to improve the spectral efficiency in future wireless systems. The evolution to cellular systems with a larger number of antennas at the base station is essential to enhance the cell capacity without extra spectral resources. In the literature, it was shown that, in a single-cell system, it is always advantageous to have an unlimited number of antennas at the transmitter [2] and also at the receiver [4]. In [2], the author proposed massive MIMO systems using a simple linear algorithm such as maximal ratio transmission (MRT) in downlink and maximal ratio combining (MRC) in uplink. It was shown that the system capacity with an unlimited number of antennas is only degraded by pilot contamination in a multi-cell network. In [5], the downlink performance of MRT and zero-forcing (ZF) beamforming for massive MIMO systems were investigated. With MRT precoding, inter-user interference is eliminated when the transmitter has unlimited number of antennas.

The assumption of an infinite number of antennas, however, is not feasible in practice. In [6] and [7], the authors investigated the capacity of the numerous but finite antenna systems for perfect channel state information (CSI) and imperfect CSI scenarios, respectively. The authors in [3] implemented a many antenna infrastructure system with 64 transmit antennas. From the experimental results of [3] and theoretical results of [6], it was shown that there is a large performance gap between MRT and ZF with respect to the number of served users when the number of transmit antennas is large but limited.

Another approach to enhance the performance of cellular networks performance has been considered from architecture, that is, cloud radio access networks (C-RAN). In these networks, the base band unit (BBU) processes Layer-2/Layer-1 functions and generates (or decodes in uplink) in-phase and quadrature phase (IQ) data for transmission at the remote radio head (RRH) (or reception in uplink).

The BBUs and RRHs are connected via a high speed fronthaul link such as optical fibers [8]. With such centralized BBU and distributed RRH structures, network-wide performance can be more enhanced by exploiting real-time joint scheduling or network MIMO as in [6], especially, in managing inter-cell interference. In addition, due to the centralized computational resources at the BBU, mobile operators can reduce the installation and operation cost. There are few researches on fronthaul data transport issue in massive MIMO employed C-RANs..

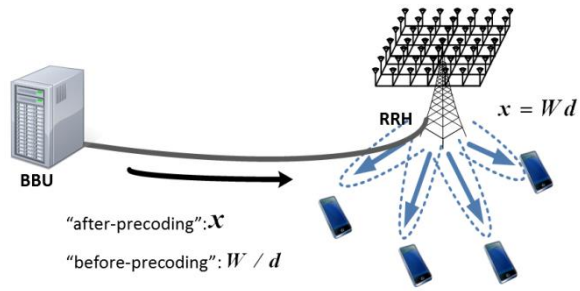


Figure 1. Two fronthaul data transport solutions.

2. Fronthaul Transport Methods in Cloud MIMO

In C-RAN, the transmit signals for each RRH antenna are processed in the BBU pool and transported to the RRH through a fronthaul link. The transmit radio signals are the mix of each stream's signal for each antenna, which is, determined by the beamforming technique chosen by the BBU, and the precoded data symbol with a weight vector.

IQ Data Transport.

For MIMO transmission, the simplest way for a BBU to inform an RRH of transmit signals is to transport the IQ-data samples of radio signal. That is, the BBU, "after precoding" the data symbols, transports the radio signal vector $\boldsymbol{x}$ to the RRH in the form of IQ-data.
Then, the required fronthaul bit-rate when radio signals are transported "after-precoding" is

$$R_{\text{after}} = N_{\text{sub}} f_{\text{sym}} b_{\text{IQ}} M. \qquad (1)$$

N_{sub} is the number OFDM subcarriers during a symbol duration and f_{sym} is the symbol frequency. The number of required bits to represent one pair of IQ sample and the number of active antennas at the RRH are denoted by b_{IQ} and M, respectively.

Separate Transport of Precoder and Data Symbol.

Another fronthaul transport solution for MIMO transmission at RRHs is separate delivery of

unmodulated data symbols and the precoder "before precoding" them. Fig.1 shows the difference between "after-precoding" and "before-precoding" methods in a fronthaul link to transport radio signal information for MIMO transmission at an RRH. The data symbol vector $\boldsymbol{d}$ and the precoding matrix $\boldsymbol{W}$ can be carried separately before the data symbols are ever precoded. In this case, the length of the data symbol vector equals the number of users, K.

By separately transporting precoding vectors and data symbols, the precoding vectors does not need to be updated with the identical time scale of data symbol. Considering that $K \times M$ precoding matrices are delivered with frequency of f_{pre}, the bit-rates required for precoder delivery in the fronthaul link is

$$R_{\text{before}} = N_{\text{sub}} f_{\text{pre}} b_{\text{IQ}} MK + N_{\text{sub}} f_{\text{sym}} b_{\text{DS}} K. \qquad (2)$$

The number of bits to represent data symbol, b_{DS}, depends on the possible number of modulation symbols. The data symbols are updated every symbol time as the transmit signals in "after-precoding", but the data symbols for each user are commonly used in generating the transmit signal of all the transmitting antennas. This is effective in the amount of fronthaul data especially when the number of antennas is much larger than that of users (i.e., $K \ll M$).

In heterogeneous channel scenario where users experience different channel coherence times, the precoding vector of each user terminal can be updated according to its channel coherence time.

3. Wireless capacity under given fronthaul capacity. Wireless sum-rates of MRT/ZF beamforming.

The received signal at the j-th user is expressed as

$$y_j = \sqrt{\rho_f} h_j^T w_j d_j + \sum_{\ell=1,\ell\neq j}^{k} \sqrt{\rho_f} h_j^T w_\ell d_\ell + n_j, \qquad (3)$$

where $\boldsymbol{h}_j^T$ is the column vector of the j-th user's channel.

We use the following precoding matrix for MRT and ZF beamforming methods.

$$\boldsymbol{W}_{\text{MRT}} = c_1 \hat{\boldsymbol{H}}^* = \left[\boldsymbol{w}_1\ \boldsymbol{w}_2 \cdots \boldsymbol{w}_K\right]. \qquad (4)$$

$$\boldsymbol{W}_{\text{ZF}} = c_2 \hat{\boldsymbol{H}}^* \left(\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^*\right)^{-1} = \left[\boldsymbol{w}_1\ \boldsymbol{w}_2 \cdots \boldsymbol{w}_K\right], \qquad (5)$$

where $\hat{\boldsymbol{H}}$ estimated channel coefficient matrix from pilot signals, and c_1 and c_2 are constants for power normalization. For simplicity, we assume identical average SNR and precoding matrix update interval of considered users. Then, the ergodic capacity lower bounds of MRT and ZF beamforming in a single cell are given by [7]

$$C_{\text{MRT}} = K \log_2\left(1 + \frac{M}{K} \frac{\rho_f \tau_r \rho_r}{(\rho_f + 1)(\tau_r \rho_r + 1)}\right) \qquad (6)$$

$$C_{\text{ZF}} = K \log_2\left(1 + \frac{M-K}{K} \frac{\rho_f \tau_r \rho_r}{\rho_f + 1 + \tau_r \rho_r}\right) \qquad (7)$$

The optimal number of active antennas with a limited fronthaul capacity.

From (6) and (7), it can be noticed, that the sum-rates of both beamforming schemes monotonically increase with the number of active antennas M. This means that if other variables are constant, using a larger M is always better. In this sense, for the given fronthaul link capacity, C_{front}, the number of active antennas that maximizes the wireless capacity with "after-precoding" fronthaul transport is

$$M^*_{\text{after}} = \min\left\{\lfloor \theta_1 \rfloor, M_{\text{tot}}\right\}, \qquad (8)$$

where M_{tot} is the total number of deployed antennas at the RRH and $\theta_1 = C_{\text{front}} / (N_{\text{sub}} f_{\text{sym}} b_{\text{IQ}})$.

On the other hand, the required bit-rates of "before-precoding" transport depends on the number of users and the precoding update frequency as

$$M^*_{\text{before}} = \min\left\{\left\lfloor \frac{\theta_2}{K} - \theta_3 \right\rfloor, M_{\text{tot}}\right\}. \qquad (9)$$

Here, $\theta_2 = C_{\text{front}} (N_{\text{sub}} f_{\text{pre}} b_{\text{IQ}})$ and $\theta_3 = f_{\text{sym}} b_{\text{DS}} / (f_{\text{pre}} b_{\text{IQ}})$.

Thus, as more users are simultaneously served, less number of antennas can be activated. Under the following condition, total number of deployed antennas, M_{tot} , can be utilized.

$$C_{\text{front}} \geq N_{\text{sub}} \left(f_{\text{pre}} b_{\text{IQ}} M_{\text{tot}} + f_{\text{sym}} b_{\text{DS}}\right) K. \qquad (10)$$

By substituting (8) and (9) into (6) and (7), we can obtain the capacity lower bounds of MRT and ZF of a single RRH in cloud MIMO system. The sum-rate curves with different average SNRs and precoding update interval are shown in Figs. 2,3,4, and 5. The used parameters are summarized in Table 1.

N_{sub}	2048
T_{sym}	66.7 μsec
b_{IQ}	40 bits
b_{DS}	7 bits
ϕ	39.28 (4x9.82) Gbps
M_{tot}	64
K_{tot}	64

Table 1. Evaluation Parameters.

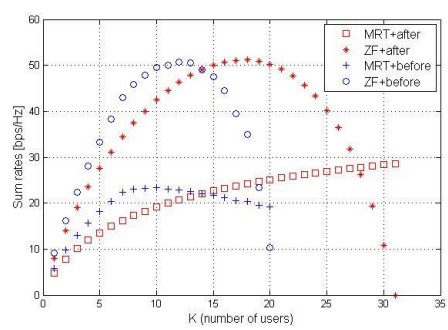


Figure 2. Sum rates (M_{tot} =64, T_{pre} =1 msec).

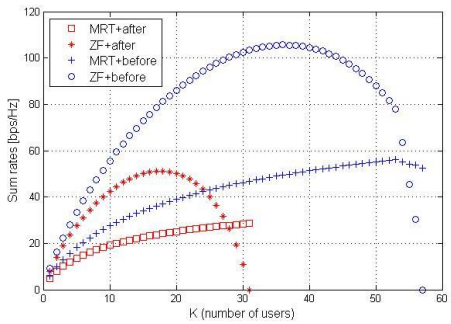


Figure 3. Sum rates (M_{tot} =64, T_{pre} =10 msec).

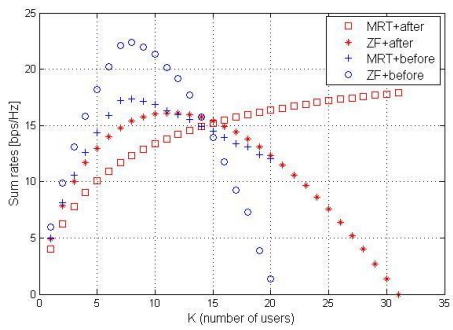


Figure 4. Sum rates (SNR = 0 dB, T_{pre} =1 msec,).

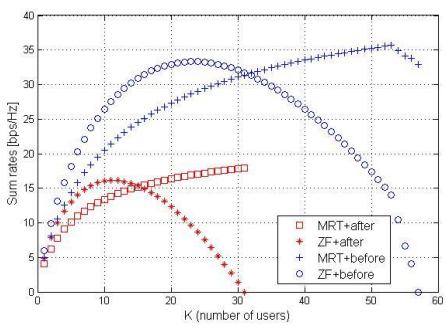


Figure 5. Sum rates (SNR = 0 dB, T_{pre} =10 msec,).

4. Discussions

The sum-rate curve is plotted with feasible K. We can see that with the "after-precoding" method, the sum-rate curve is independent of the precoding matrix update interval and has the same tendency as was found in [6]. This is because the required IQ-data transport rate is independent of the precoding matrix update interval. On the other hand, the sum rates of the "before-precoding" method in Figs. 3 and 5 are larger than those in Figs. 2 and 4. This is because as f_{pre} decreases, M increases, and the sum-rate monotonically increases with the number of antennas used for transmission. We see that with the "after-precoding" method, the maximum sum-rate of MRT is higher than that of ZF in the low SNR scenario, but, in the high SNR scenario, ZF has higher maximum sum-rates than MRT–a trend already shown in [6] and [7].

	K_{max}	$K^\star$	$M^\star$
MRT, after	31	31	31
ZF, after	31	11	31
MRT, before	20	8	57
ZF, before	20	8	57

Table 2. Selected variables.

However, with the "before-precoding" method, ZF shows, in both high and low SNR scenarios, a higher maximum sum-rate than MRT. This is due to the flexibility of "before-precoding" in choosing M and K, whereas M is fixed with the "after-precoding" method. Table 2 represents K_{max} and the optimal values of M and K, in Fig. 4. It is remarkable that with the "before-precoding" method, lower K and larger M are chosen to maximize the sum-rate.

This can be interpreted as that more antennas can be used at the cost of reduced multiplexing order to achieve higher sum-rates for low SNR users. Compare Fig. 2 and Fig. 4 and check that less number of users and therefore more antennas maximizes the sum-rates. Thus, we can derive desirable beamforming strategys according to fronthaul transport method and average user SNR.

Note that we assumed different precoding weights are applied for each subcarrier. In practice, the same precoding weights can be applied over several subcarriers that have similar channel coefficients, leading to much lower required fronthaul rates of "before-precoding" for the given number antennas and users.

	"after-precoding"	"before-precoding"
Cell-center	ZF	ZF
Cell-boundary	MRT	MRT or ZF

Table 3. Desired beamforming technique in massive MIMO C-RAN.

5. Conclusion

In this article, based on our prior work, we investigated the effect of a limited number of antennas or users due to a fronthaul link constraint on the wireless sum-rate, which also depends on the beamforming strategy and the radio signal transport method in a fronthaul link.

For the given fronthaul link capacity and user environments, the beamforming and the fronthaul transport methods can be jointly optimized to maximize the sum-rate.

References

[1] S. Park, C.-B. Chae, and S. Bahk, "Before/after precoding massive MIMO systems for cloud radio access networks," *IEEE/KICS Jour. Comm. Networks*,

vol. 15, no. 4, pp. 398–406, Aug. 2013.
[2] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," IEEE Trans. Wireless Comm., vol. 9, no. 11, pp. 3590-3600, Nov. 2010.
[3] C. Shepard, H. Yu, N. Anand, L. E. Li, T. Marzetta, R. Yang, L. Zhong, "Argos: Practical many-antenna base stations," in Proc. *ACM Mobicom*, 2012.
[4] C.-B. Chae and R. W. Heath, Jr., "On the optimality of linear multiuser MIMO beamforming for a two-user two-input multiple-output broadcast system," *IEEE Sig. Proc. Lett.*, vol. 16, no. 2, pp. 117-120, Feb. 2009.
[5] B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, F. Rusek, D. Persson, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with large arrays," *IEEE Sig. Proc. Mag.*, vol. 30, no. 1, pp. 40-60, Jan. 2013.
[6] C. Lee, C.-B. Chae, T. Kim, S. Choi, and J. Lee, "Network massive MIMO for cell-boundary users: From a precoding normalization perspective," in Proc. *IEEE Glob. Telecom. Conf.*, 2012, pp. 1-5.
[7] H. Yang and T. L. Marzetta Lee, "Performance of conjugate and zeroforcing beamforming in large-scale antenna systems," *IEEE Jour. Select. Areas in Comm.*, vol. 31, no. 2, pp. 172-179, Feb. 2013.
[8] Chaina Mobile Research Institute, "C-RAN: The road towards green RAN," White Paper, http://labs.chinamobile.com/article download. php?id=104035.
[9] "CPRI specification v5.0 (2011-09-21)," http://www.cpri.info/downloads /CPRI v 5 0 2011-09-21.pdf.

Sangkyu Park received B.S degree in Electrical Engineering from Yonsei University, Seoul, Korea in 2009. He is currently working towards the Unified Master's and Doctor's Course in Seoul National University. His current research interests include network optimization and radio resource management for cloud radio access network, large-scale antenna system, and heterogeneous network.

Chan-Byoung Chae is an Assistant Professor in the School of Integrated Technology, Yonsei University, Korea. He was a MTS at Bell Labs., Murray Hill, NJ, USA. Before joining Bell Labs, he was with the SEAS at Harvard University, MA, USA. He received the Ph.D. degree in Electrical and Computer Engineering from the Univ. of Texas (UT), Austin, TX, USA in 2008. Prior to joining UT, he was a Research Engineer at the Telecom. R&D Center, Samsung Electronics, Suwon, Korea, from 2001 to 2005. His current research interests include capacity analysis and interference management in energy-efficient wireless mobile networks and nano (molecular) comm. He has served/serves as an Editor for the IEEE TRANS. ON WIRELESS COMM., IEEE TRANS. ON SMART GRID, and IEEE/KICS JOUR. OF COMM. NETWORKS. He is also an Area Editor for the IEEE JOUR. SELEC. AREAS IN COMM. He is an IEEE Senior Member.

Saewoong Bahk received B.S. and M.S. degrees in Electrical Engineering from Seoul National University in 1984 and 1986, respectively, and the Ph.D. degree from the University of Pennsylvania in 1991. From 1991 through 1994, he was with AT&T Bell Laboratories as a Member of Technical Staff where he worked for AT&T network management. In 1994, he joined the school of Electrical Engineering at Seoul National University and currently serves as a Professor. He has been serving as TPC Members for various conferences including ICC, GLOBECOM, INFOCOM, PIMRC, WCNC, etc. He is on the Editorial Boards of IEEE Transaction on Wireless Communications (TWireless), Computer Networks Journal (COMNET), and Journal of Communications and Networks (JCN). His areas of interests include performance analysis of communication networks and network security. He is an IEEE Senior Member and a Member of Who's Who Professional in Science and Engineering.

INDUSTRIAL COLUMN: BIG DATA IN 5G NETWORKS: FROM END DEVICE TO PERSONAL CLOUD, EDGE CLOUD AND INTERNET CLOUD

Guest Editor: Hassnaa Moustafa, Intel Corporation, USA
hassnaa.moustafa@intel.com

With the advent and proliferations of end user personal devices, sensors technologies and cloud infrastructure, big data is becoming an essential pillar for many applications making end-users status and their environment trigger the applications behavior in a context-aware manner. These applications are mostly real-time and they include, but not limited to, surveillance and monitoring, medical and wellness, multimedia and entertainment.

Real-time applications require high throughput and low latency on one hand. On the other hand, the increasingly big data streams created by end-user devices, IoT, and M2M, require huge processing and storage capacity. This necessitates selecting the right database and storage level (personal cloud, edge cloud or Internet Cloud) combining data base solutions with big data technologies.

5G from end device to personal cloud, edge cloud and internet cloud Newsletter associates big data with storage and network infrastructure technologies showing new approaches in storage (e.g. semantic storage based on big data) and new network infrastructure (e.g. private Internet, personal and edge cloud). Context-aware content adaptation techniques are also presented making use of big data.

Similar to the continued evolution of the Internet, there is consensus amongst analysts that "Big Data" associated with the network, from the Cloud to the edge device, will drive a tremendous amount of innovation and disrupt existing business practices across industries.

This special issue of E-Letter focuses on the recent progresses of big data usage, manipulation and storage in 5G applications. The E-Letter includes five papers from leading research groups, from both academia and industry laboratories, to report their solutions for meeting these challenges and share their latest results.
David Cohen and Eve Schooler from Intel authored the first article, entitled "*Data Inversion and SDN Peering: Harbingers of Edge Cloud Migration*". The article presents a data inversion model for IoT showing how the assumption that data will continue to be managed largely in the back-end data-center Cloud will not be the perfect match for Internet of Things (IoT). Yet, the Internet of Things is not exclusively about end-users who consume "Big Data" from the Cloud but increasingly is about interconnected devices deployed at the edge of the network who are the producers of "Little Data" flowing upstream toward the Cloud. The vision involves moving more of the intelligence, for example smart data analysis algorithms, closer to the edge and in the process enabling smart services, smart storage and smart data retrieval. The paper presents the expected changes in network infrastructure resulting from the massive deployment of IoT applications. These include CDNs, peering, overlays, and computation models, physical and virtual IoT clouds, multi-tenancy and Software Defines Network (SDN).

The second article, entitled "*Enabling Massive MTC Deployments with Stringent Performance Requirements*", authored by Tamper University and Ericsson Research and presents Machine Type Communication (MTC) that a very distinct emerging capability in 5G networks. The current and future long-range LTE-powered IoT deployments are expected to see a huge influx of traffic collected by diverse MTC (machine to machine) devices. This allows information exchange between a device and another entity in the Internet or the core network, or between the devices themselves, which does not necessarily require human interaction. Even though the data from one particular device may be rather small, MTC devices are expected to generate massive amount of information "big data". This paper reviews recent research efforts along these lines, with the ultimate goal of delivering efficient and affordable wireless connectivity at low power and with moderate large-scale deployment effort. In this context, 3GPP LTE technology proved to support large-scale MTC installations in wide coverage, relatively low deployment costs, and simplicity of management, achieving "anytime, anywhere" wireless connectivity across a plethora of prospective MTC applications.

The third article, entitled "*Device to Device for Wearable Communication in 5G Networks*", presents the core technology of Green Communication. The article shows a new telecommunication model for wearable devices communication in 5G networks alleviating the disadvantages of the classical telecommunication model from 2G to 4G in terms of capacity and energy consumption. The paper introduces the concept of Embedded Internet where devices hold a full TCP/IP environment, an open operating system and a sharing storage. This allows Device to Device (D2D) network to be established in

few seconds and operates a local Internet. This solution guarantees an increase in the capacity of the networks, low power communication, local cloud storage matching the massive wearable devices needs and offloading the internet and core network from the big data transfer to and from Internet Cloud.

The fourth article, entitled "*Big data for Cloud based Video Adaptation*", is contributed by Orange Labs and it presents a new approach for using big data to optimize video service delivery in terms of saving network resources and enhancing end-user experience through creating new service. In this approach, Big data allows network operators and service providers to assess Quality of Experience (QoE) for video services and adapt the video delivery accordingly in terms of content adaptation (e.g. personalized content, content recommendation) and network/resources adaptation (e.g. adaptive bit rate, anticipated caching, multicast delivery ..). Big data is also applied in Cloud-as-a-Service offers by Over the Top (OTT) actors and Telcos, allowing taking adequate decisions on infrastructure virtualization considering QoE coupled with context data as big data on the user, devices, network and the content itself. In this approach, big data coupled with video content delivery is promising for network operators and service providers in terms of maintaining the Average Revenue per User (ARPU) through QoE guarantees and reducing the allocated CAPEX/OPEX. In addition, OTTs find benefits in terms of service differentiation.

The last article of this special issue is from AvidBeam and the title is *"Overview of Big Data in Video Analytics"*. This article provides an excellent overview on the video analytics applications making use of big data. The author first how big data can be useful in video analytics to spot unexpected problems (e.g. in surveillance applications). Then the article presents video analytics applications, principles, system components and data processing and describes the technical approaches and challenges in video analytics, which are mainly management of compressed video, accuracy of image detection and recognition. The paper shows the details of big data analytics for several real-life use-cases.

While this special issue is far from delivering a complete coverage on this exciting research area, we hope that the five invited letters give the audiences a taste of the main activities in this area, and provide them an opportunity to explore and collaborate in the related fields. Finally, we would like to thank all the authors for their great contribution and the E-Letter Board for making this special issue possible.

Hassnaa Moustafa is a Senior Research Scientist in Intel PC Client Division, leading innovations in user's experience for multimedia and video applications. Her research interests include wireless networks (including WPANs and IoT), optimized video delivery and context-awareness. She previously worked in France Telecom (Orange), where she led projects on low-cost wireless networks and services convergences in next generation networks. Hassnaa is a senior IEEE member and is qualified as a research director by the University of Paris XI in 2010. She is contributing to the IETF standardization and has over 70 publications in international conferences and journals and she is a co-author of two books published by the CRC press. Hassnaa obtained a PhD in Computer and Networks from Telecom ParisTech in 2004, a Master degree in distributed systems from the University of Paris XI in 2001.

Data Inversion and SDN Peering: Harbingers of Edge Cloud Migration

David E. Cohen and Eve M. Schooler
Intel Corporation, USA
{*david.e.cohen, eve.m.schooler*}*@intel.com*

Introduction. There is consensus amongst analysts that "Big Data" associated with the network, from the Cloud to the edge device, will drive a tremendous amount of innovation and disrupt existing business practices across industries. Sound like the continued evolution of the Internet? Not quite. The assumption is that data will continue to be managed largely in the back-end data-center Cloud. Yet, the Internet of Things is not exclusively about end-users who consume "Big Data" from the Cloud but increasingly is about interconnected devices deployed at the edge of the network who are the producers of "Little Data" flowing upstream toward the Cloud. The vision involves moving more of the intelligence, for example smart data analysis algorithms, closer to the edge and in the process enabling smart services, smart storage and smart data retrieval.

The Internet of Things. Gartner defines the Internet of Things (IoT) as "the network of physical objects that contain embedded technology to communicate and sense or interact with their internal states or the external environment." They assert that the number of connected Things is expected to climb from billions of devices to 10s of billions by 2020. By that time, IoT product and service suppliers are expected to generate $100s of billions in incremental revenue, mostly in Services. Needless to say such potential for growth has not gone unnoticed amongst competitors in the technology sector. For example, Intel's third quarter 2014 financial report explicitly calls out the Internet of Things Group revenue of $530 million, representing 14% year-to-year growth.

Internet Evolution. The emergence of this fast-growing IoT sector harkens back to the emergence of the Internet sector in the late 1990's. Is the Internet a harbinger of the IoT to come? In the Internet global-scale services like Google's search and Facebook's News Feed are consumed by end-users via their smart phones. None of this existed 20 years ago. For example, Akamai Technologies is a leader in the $4 billion Content Distribution Network (CDN) market. Prior to the late 1990's this market didn't exist. Further, content producers such as Google, Amazon, Facebook, among others, now run their own internal CDNs, which underscores the scope and importance of CDNs to the operation of the evolving Internet and emergent IoT. Another capability crucial to the operation of the modern Internet is the notion of peering and the use of Internet Exchange Points (IXP). The general perception is that tier-1 transit providers such as AT&T, Verizon, British Telecom, and China Telecom make up the backbone of the network and as such carry the bulk of the Internet's traffic. In this model tier-1 providers peer amongst themselves and do not charge each other for transit costs. Instead, the tier-1s make their money in peering relationships with regional/tier-2 and local/tier-3 network providers. Ultimately, these network providers provide connectivity to Internet Service Providers (ISP), who connect our businesses and residences at the very edges of the Internet to the rest of the infrastructure.

While this model was prevalent in the early days of the Internet, today the Internet works quite differently. The vast majority of Internet content never leaves a region. Google, Netflix, Facebook, and other content providers like them all use peering relationships through regional IXPs to connect to the ISPs in that region. This so-called "flattening of the Internet" trend began to emerge in the mid-2000's. Today, this flattening is a global phenomenon with hundreds of IXPs operating around the world, which changes the notion that tier-1 carriers provide the backbone of the Internet.

These ingredients (end-devices, CDNs, peering, overlays, and computation) - have defined the leadership of the Internet. The leadership of the IoT will refactor these ingredients with a priority toward localization.

IoT Service Providers. Discussion of the Internet evolution gives rise to an interesting question: Who will dominate the content distribution and network connectivity capabilities in the IoT? For many it is a foregone conclusion that leaders in the Telecommunications sector will lead in this market. Certainly, they have a clear advantage with their ownership of metro-area-based wireless and cellular networks along real-estate and right-of-way in these locations. However, these are the same key assets that should have given them the advantage as the Internet emerged. Yet, today no one argues that it is companies like Google, Facebook, Microsoft, and Amazon who have provided significant innovation and leadership to rival the Telcos in the Internet sector and stand just as good, if not a better chance of winning the IoT segment.

Though one must wonder if new disruptive IoT services may up end all of the current Internet players. Such uncertainty presents a challenge for mobile/cellular network operators, as an example. Over the next few years, they face having to make substantial investments to upgrade their current infrastructure to next-generation 5G technologies in order to support IoT, particularly for mobile multimedia applications.

So what strategy to adopt to win in the IoT? The wireless and cellular networks are key components to winning the IoT. Currently they provide the main method by which edge devices attach to the IoT. Notably the role of phones as a driver of the architecture has never been greater, as mobile phones have surpassed laptops as the predominant means by which users access the web. However, as the warehouse-scale Cloud Service Providers (CSPs) have shown, whomever finds ways to gain control over end device services has an added advantage, as they showed by exploiting peering to go "over-the-top" of the Telcos' back-end network services. This Over-the-Top approach has the potential danger to, yet again, relegate the Telcos to being merely "Pipe Providers." This brings into question the wisdom of investing in improvements to enable 5G when there is uncertainty on where the return on the investment will come from.

Data Inversion. IoT will increasingly encourage localization of content, a shift in where data is produced and consumed. In fact, the IoT promises to bring with it "Data inversion", where a growing number of end devices at the edge of the network are producers of content while consumers of the content will reside upstream, not only in back-end data center clouds, but also closer to the edge, at aggregation points or devices co-resident to where the data is being created. Effectively this represents an inversion of the content distribution model. In this inverted model, the IoT employs IP-peering techniques to prioritize localization in contrast to the Internet where peering is used to regionalize content.

Industry-specific Partners and Multi-Tenant Infrastructure. Within a region or local area, network infrastructure must evolve to support the decentralized IoT and physical resources will need to be managed and shared. While such multi-tenant sharing is common in the broader Internet, within the IoT this capability must be scaled further, to be distributed across the local and regional environments. Assuming the Telcos take the lead on building out this physical connectivity, it is critical that they make this connectivity accessible to the broadest constituency of consumers and partners, who have tremendous expertise in specific verticals and as such can offer vertical-specific IoT services: for example, General Electric who is driving what they refer to as the "Industrial Internet"; Shell Oil who is reshaping its business by employing IoT technologies to change its business processes; John Deere who is aggressively employing IoT technologies to drive innovation in the agriculture/farming market. The reason why partnerships with leading enterprises will distinguish leadership in the IoT is that openness and collaboration will encourage needed innovation. As seen already, trying to "go it alone" or to limit access to only those willing to pay high tariffs and sign up for exclusivity are questionable strategies.

Multi-Tenancy, Localization, and Computation-Near-the-Data. The implication of Data Inversion is that increasingly much of the IoT content will remain near to where it is produced. The challenge becomes one of building "mini-Clouds" where Industry partners (aka "tenants") can deploy their services, effectively migrating computation to the data. Unlike the MapReduce of the Internet era, moving computation close to the data means moving it out of warehouse-scale data centers. The question is to where?

The answer is close to the IoT devices producing the data. However, as discussed there are 10s of billions of these devices distributed at the farthest most outer boundaries of the network. Today, no such computational facilities exist in this "IoT wilderness".

Physical and Virtual IoT Clouds. One possibility is that the Telcos, Utilities, Municipalities and other property owners will transform their metro-area real-estate assets into multi-tenant computational facilities where they can migrate compute closer to data generation. Then they can offer these facilities to their partner/customers like GE, Shell, and John Deere to deploy/operate IoT-centric services. Granted, these distributed resources may be considerably scaled down local versions of data centers. There are also opportunities for the even further scaled back version to be found in the Edge Cloud, which might reside physically in the smart neighborhood or smart home, or in the Personal Cloud, which might be embodied by all of the federated devices in the smart car or encompass the devices that are being carried by an individual on a daily basis. Cloud migration out of the data center will likely be disruptive in that cloud locations and resources may be owned and managed by more local proprietors, small businesses, households or individuals.

Software Defined Networking for IoT. Bringing IoT capabilities to bear on a particular problem also will require a level of virtualization, such as decoupling mini-Clouds from their physical locations. A

company/tenant is then free to reassemble resources in support of the services they operate over their virtualized cloud.

In the Internet, Content Providers have already done this. They deliver content to consumers via the use of "Overlay Networks." The warehouse-scale CSPs leverage smart phones (tablets, laptops, etc.) as a way to "own" content distribution in a manner analogous to Rockefeller's Standard Oil with its service model of refineries and gas stations. The smart phone is the Internet's car, the CDN's cache is the gas station, and the CSPs' warehouse-scale data centers the refineries. Connecting all of this are IP-based overlay networks.

These overlays and the means by which they are operated are increasingly managed via logically centralized, software-based services, collectively referred to as Software-Defined Networking (SDN). An SDN includes a controller that can be thought of as the overlay's master mind. In a world where cloud control migrates closer to the edge, something like the SDN controller will play a strong role not only in managing the network for the cloud, but also in managing a diverse collection of additional services associated with the cloud, and also in creating peering relationships with other clouds through their SDN controllers. Furthermore, IoT overlay networks will be key to enabling edge devices finding IoT services operating in relatively close proximity.

Vision for the Future. Although more questions have been raised than answered, these musings identify several important trends to heed as the IoT is under design and construction. Earlier the question was asked: who will own the CDN and connectivity services for the IoT, given who owns these services in the current Interent? Yet the more fundamental question is who will be the service providers of the IoT? Meaning, what broader range of services will and should the IoT offer? And what will they require of an increasingly distributed services architecture?

If the pervasiveness of CDNs is any indicator, the IoT will be a place where content continues to reign. Thus, the need for a content-centric IoT architecture, possibly one where the functionality of CDNs is built in natively to the lower layers of the network, so that data is liberated to be shared across IoT applications and market verticals. IoT data is increasingly more likely to remain local to where it was generated and certainly within much smaller geographical regions than in the present Internet, where it might be used in a more localized fashion by IoT applications such as driverless cars and smart energy, and where data privacy has the opportunity to be managed more transparently. Thus, computation must be able to migrate to where the data and services reside. IoT services such as edge analytics, smart data pipes, trusted data as a service, data exchanges, are all data-driven services to enable. The vision is to repurpose SDNs and overlays as software-defined clouds, not only to elevate and liberate data for re-use within the content-centric IoT, but also to support migration and localization of services to the farthest reaches of the network edge.

References

[1] Rivera, "Gartner Says the Internet of Things Installed Base Will Grow to 26 Billion Units By 2020," 2013, http://www.gartner.com/newsroom/id/2636073

[2] Feamster et al, "SDX: A Software Defined Internet Exchange," 2013, http://www.cs.umd.edu/~dml/papers/sdx-ons13.pdf

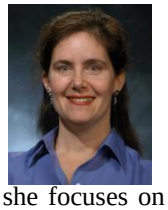
[3] McCauley et al, "Extending SDN to Large-Scale Networks," 2013, http://www.cs.columbia.edu/~lierranli/coms6998-10SDNFall2014/papers/Xbar-ONS2013.pdf

[4] Schmid et al, "Exploiting Locality in Distributed SDN Control," 2013, http://users.ics.aalto.fi/suomela/doc/local-sdn.pdf

[5] Shavitt et al, "Topological Trends of Internet Content Providers," 2012, http://arxiv.org/pdf/1201.0834.pdf

[6] Gill et al, "The Flattening Internet Topology: Natural Evolution, Unsightly Barnacles or Contrived Collapse?", 2008, http://www.hpl.hp.com/techreports/2008/HPL-2008-47.pdf

David E. Cohen received a Bachelor's Degree in Accounting from the University of Texas at Austin in 1989. He has been a distributed systems practitioner for over 30 years. His research interests include exploiting the network to manage data at scale. He is currently with Intel where he is a Senior Principal Engineer and Architect in the Communication and Storage Infrastructure Group.



Eve M. Schooler received her BS from Yale, MS from UCLA, and PhD from Caltech, all in Computer Science. She joined Intel Labs in 2005 and is now a Principal Engineer and Architect in the Software and Services Group, where she focuses on data privacy, trusted data as a service, edge and personal clouds. Prior to Intel, she held positions at Apollo Computer, Information Sciences Institute, AT&T Labs-Research and Pollere LLC.

Enabling Massive MTC Deployments with Stringent Performance Requirements

Sergey Andreev[1], Tuomas Tirronen[2], Johan Torsner[2], and Yevgeni Koucheryavy[1]
[1]Tampere University of Technology, Finland; [2]Ericsson Research, Finland
sergey.andreev@tut.fi, {tuomas.tirronen, johan.torsner}@ericsson.com, yk@cs.tut.fi

1. Introduction

The *Internet of Things* (IoT) has recently developed as a novel concept where physical things (such as our surrounding objects) are extended with actuators, sensors, and identifiers, made discoverable through the network, and are tightly integrated into the infrastructure of the future Internet [1]. As traditional voice service revenues continue to shrink, mobile network operators are increasingly interested in the IoT applications to bridge in the growing revenue gap.

A very distinct emerging capability that enables the implementation of the IoT is machine-to-machine or *Machine-Type Communications* (MTC). Industry reports indicate the considerable potential of the MTC market, with millions of devices connected over the following years, resulting in trillion-dollar revenues by the year 2020. According to [2], the concept of MTC may be defined as information exchange between a *device* and another entity in the Internet or the core network, or between the devices themselves, which does not necessarily require human interaction.

The prospective MTC applications are numerous and generally cover the areas such as (i) secured access and surveillance, (ii) tracking, tracing, and recovery, (iii) public safety, (iv) electronic payment, (v) healthcare, (vi) remote maintenance and control, (vii) smart metering, (viii) consumer devices, and (ix) retail. Due to this impressive variety, existing wireless systems are targeting air interface enhancements to efficiently support MTC connectivity [3]. Primarily, 3GPP *Long Term Evolution* (LTE) cellular technology has seen numerous improvements over the last years across several study and work items defined on MTC.

The corresponding 3GPP efforts have revealed that ubiquity of MTC deployments is challenged by typical characteristics of MTC. Among these are extremely large numbers of devices, their small-size data transmissions, and infrequent traffic patterns. Hence, the primary standardization focus is to satisfy a range of unique requirements that are very different from those in conventional human-to-human communication, but common to many MTC use cases. Correspondingly, existing MTC-related research generally targets to support efficient transmission of small and infrequent MTC data bursts with minimal overheads across massive unattended deployments with a very large number of potentially active devices.

This paper comprehensively reviews our recent research efforts along these lines, with the ultimate goal of delivering efficient and affordable wireless connectivity at low power and with moderate large-scale deployment effort. As the result, we confirm that 3GPP LTE technology is becoming increasingly attractive for supporting large-scale MTC installations due to its wide coverage, relatively low deployment costs, and simplicity of management. Hence, it has capability to achieve "anytime, anywhere" wireless connectivity across a plethora of prospective MTC applications.

With further evolution of 3GPP LTE (Rel-12 and beyond), which is currently regarded as the mainstream activity towards fifth-generation (5G) wireless networks, the information sharing capabilities between diverse interacting devices would improve dramatically. Given its potential to enable ubiquitous connectivity between various communicating objects, as well as collection and sharing of the massive amounts of data, MTC is likely to become the centerpiece of the emerging 5G ecosystem with its associated challenges, such as extreme heterogeneity, large-scale unattended wireless connectivity, and unprecedentedly large volumes of information.

2. Challenges of massive large-scale MTC scenarios

Our analysis in [4] indicates that *smart grid* may become one of the key MTC use cases that involves meters autonomously reporting usage and alarm information to grid infrastructure to help reduce operational cost, as well as to regulate a customer's utility use based on load-dependent pricing signals received from the grid. The motivating smart metering use case therefore serves as a valuable reference MTC scenario (see Figure 1) covering many characterizing MTC features.

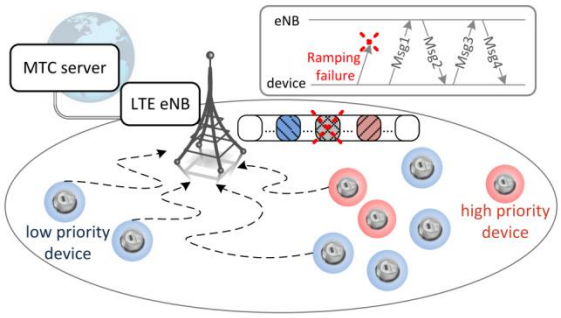


Figure 1. Envisioned long-range MTC deployment [5].

We expect that cellular technologies, such as 3GPP LTE, will play a pivotal role in enabling future smart metering applications. Together with effective measures for overload control in smart grid, the LTE system shall also provide mechanisms to lower power consumption of small-scale battery-powered wireless meters. As transmitted data bursts may be extremely small in size, the network should additionally support efficient transmission of such packets with very low overhead. In what follows, we summarize our related research, which is a combined pursuit of analysis and detailed protocol-level simulations.

Enabling device cooperation for MTC.
Our initial work in [6] studied a typical smart metering MTC application scenario, which features a very large number of devices connecting to the network. We focused on enhancing the performance of cell-edge MTC devices with poor communication link and proposed a simple and feasible *client relay* scheme to improve link performance. With client relay technique, MTC devices with better channel conditions may relay data on behalf of other proximate devices and thus improve the resulting communication performance.

The outcomes of our analysis indicate that latency and energy expenditure of cell-edge MTC devices may be dramatically lowered with our proposed communication scheme, even when there is a surge in near simultaneous network entry attempts by a large number of meters. Such surge in network access attempts may occur, for example, in a power outage scenario where a large number of smart meters attempts to connect to the network to report the outage event and again when they reconnect to the network upon restoration of power.

MTC-specific overload control mechanisms.
With this research, we continue to consider a typical smart metering MTC application scenario in 3GPP LTE wireless cellular system [7] featuring a large number of devices connecting to the network near-simultaneously and then sending their data through the network. We target thorough analysis of the random access channel (RACH) within the LTE technology with respect to the congested MTC scenario.

More specifically, the RACH procedure is decomposed into two stages. The first one is the uplink timing correction stage (known as Msg1/Msg2), where the power ramping technique may be used to adjust the transmit power of a random-access preamble to particular channel conditions. Further, a meaningful uplink message (subject to appropriate contention resolution) is transmitted to the base station (termed eNB, see Figure 1) for the purposes of initial network access or bandwidth requesting (known as Msg3/Msg4).

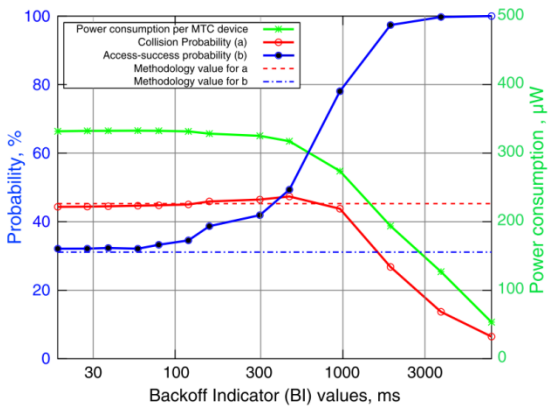


Figure 2. Performance of overload control schemes [8].

Our approach in [8] (and prior work in [9]) allows investigating the performance of MTC devices, the impact of RACH settings, and the overload control mechanisms in terms of conventional metrics, such as medium access delay and access success probability (see Figure 2). In particular, the limitation of existing access protocol in case of correlated network entry attempts has been indicated and the benefits of several potential enhancements were highlighted. These modifications do not require major protocol change and feature the pre-backoff technique complemented by the usage of larger MTC-specific backoff indicator values.

Reducing MTC device power consumption.
Due to the fact that the MTC devices are typically small-scale and battery-powered, considering their power consumption is of paramount importance. By including power consumption into our framework [8], [9] together with the conventional performance metrics mentioned above, we aim at providing a complete and unified insight into MTC device operation (see Figure 2).

Further, the research in [10], [11] studies the power consumption aspects of LTE MTC devices. We discuss characteristics of possible MTC traffic and devices to propose an appropriate power consumption model. We further investigate how various model parameters affect the battery lifetime and the power consumption of the devices. Our results indicate that making the current maximum discontinuous reception (DRX) and paging cycle length longer would lead to significant gains in the energy consumption of MTC devices.

Efficient MTC-centric data access mechanisms.
Whereas our earlier research concentrated on the analysis of the overloaded random access channel in the LTE network, we continue our investigations with an emphasis on small data access when the network is not experiencing an MTC overload. We propose and detail an efficient small data transmission mechanism which may be used as an alternative to the

conventional signaling thus significantly improving the MTC performance. In particular, the contributions of our work in [12] are (i) a novel integrated simulation-analytical framework for evaluating MTC data access mechanisms over various uplink LTE channels (see Figure 3); and (ii) an efficient MTC-specific data access scheme, which we name **co**ntention-**ba**sed **L**TE **t**ransmission (COBALT).

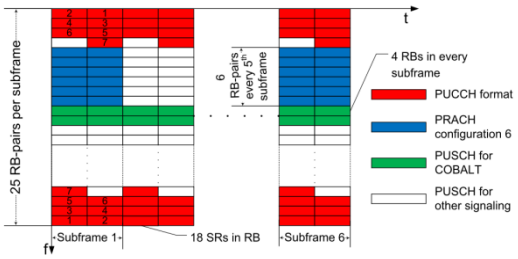


Figure 3. Uplink channels in 3GPP LTE [12].

Summarizing, in [12] we have reviewed the existing data access mechanisms, which may be used by MTC devices to transmit their data over LTE. We emphasized that neither the default PUCCH-based scheme, nor the alternative PRACH-based scheme is optimal for supporting massive MTC deployments, where the traffic arrivals are infrequent and small. By contrast, our proposed COBALT scheme takes advantage of the simple implementation and thus fewer number of LTE signaling messages. Consequently, it demonstrates significantly better usage of network resources, lower power consumption for the MTC devices, and often reduced latency performance.

Accommodating connected-mode MTC devices. In this research track, our attention shifts to investigating the situations when MTC devices have already established their LTE connection and send their meaningful data rather than the network entry requests. In [5], we investigate whether a surge in simultaneous transmission attempts by numerous *connected-mode* MTC devices actually threatens the radio network. Our main contribution is detailed characterization of an overloaded scenario where a large number of MTC devices transmit their information into the LTE network.

In particular, we introduce the scenario incorporating a mixture of *diverse device classes* transmitting to a common eNB (see Figure 1), and then assess it with both analysis and protocol-level simulations. The MTC classes correspond to the priority of the transmitted information, e.g., high priority (alarm messages), low priority (measurement data), etc. We conclude that appropriate overload control mechanisms may also be necessary for connected-mode devices and especially for high priority MTC devices to mitigate the deleterious effects of the random access procedure.

Random access for a large number of MTC devices. Recent publications have thoroughly characterized the effects of RACH overloads to understand the consequences of near-simultaneous initial network entry attempts by many MTC devices. However, past disjoint research efforts did not provide a uniform analytical view of contention-based behavior, especially when a large number of MTC devices transmit their dynamic (unsaturated) traffic. In [13] and [14], we bridge the indicated gap by proposing a novel mathematical model that essentially captures contention in typical MTC environment and further tailor it to several important MTC-over-LTE use cases. More specifically, the main contribution of our work is rigorous analytical characterization of dynamic contention in the multi-channel environment across a range of channel access algorithms and for extremely large numbers of MTC devices.

3. Conclusion and current work

Our world is rapidly developing into a *networked society* [15], where people, knowledge, devices, and information are tightly integrated across several key markets, including utilities, vehicular telematics, healthcare, and consumer electronics. This vision suggests that on the order of 50 billion devices be connected by 2020 [16], thus making the concept of the IoT reality and fueling the prospective 5G-ready 3GPP LTE installations with unprecedentedly massive amounts of information that will be gathered by various connected objects.

In recent years, LTE technology has seen numerous MTC-specific modifications, from overload protection schemes and lightweight signaling procedures, to efficient small data transmission and coverage extension mechanisms, as well as advanced approaches to MTC-aware radio resource allocation [17]. Current work in the standards (e.g., in the latest Rel-12 of LTE and beyond) targets further decisive modifications in a wide range of aspects, from handling very large number of devices with ultra-low energy consumption to delivering extended cellular coverage to low-cost and low-complexity MTC devices [15].

An emerging research angle, particularly important for industry-grade 5G systems, is focusing on feasible improvements for MTC in current and forthcoming releases of 3GPP LTE technology with respect to longer communication range, higher transmission reliability, and lower network access latency. Another pressing demand is to explore novel opportunities offered by recent progress of radio access technology for MTC connectivity in light of envisioned Industrial Internet challenges and applications, including the industrial automation use case. The above constitutes the directions of our current work.

Acknowledgment

This work is supported by the IoT SRA program of Digile, funded by Tekes. The authors also acknowledge contribution of their many colleagues from Ericsson Research and Tampere University of Technology. The work of the first author is supported with a Postdoctoral Researcher grant by the Academy of Finland as well as with Jorma Ollila grant by Nokia Foundation.

References

[1] Finnish Strategic Centre for Science, Technology, and Innovation, Internet of Things Strategic Research Agenda, 2011.

[2] J. Alonso-Zarate, M. Dohler, and T. Watteyne, "Machine-to-Machine: An emerging communication paradigm," Wireless World Research Forum (WWRF), November 2010.

[3] G. Wu, S. Talwar, K. Johnsson, N. Himayat, and K. Johnson, "M2M: From mobile to embedded Internet," IEEE Communications Magazine, vol. 49, pp. 36–43, 2011.

[4] N. Himayat, S. Talwar, K. Johnsson, S. Andreev, O. Galinina, and A. Turlikov, Proposed IEEE 802.16p Performance Requirements for Network Entry by Large Number of Devices, IEEE 802.16p-10/0006, November 2010.

[5] O. Dementev, O. Galinina, M. Gerasimenko, T. Tirronen, J. Torsner, S. Andreev, and Y. Koucheryavy, "Analyzing the overload of 3GPP LTE system by diverse classes of connected-mode MTC devices," in Proc. of the IEEE World Forum on Internet of Things, 2014.

[6] S. Andreev, O. Galinina, and Y. Koucheryavy, "Energy-Efficient Client Relay Scheme for Machine-to-Machine Communication," in Proc. of the IEEE GLOBECOM, 2011.

[7] Study on RAN Improvements for Machine-Type Communications. 3GPP Technical Report (TR) 37.868, 2011.

[8] M. Gerasimenko, V. Petrov, O. Galinina, S. Andreev, and Y. Koucheryavy, "Impact of MTC on energy and delay performance of random-access channel in LTE-Advanced," Transactions on Emerging Telecommunications Technologies, vol. 24, pp. 366–377, 2013.

[9] M. Gerasimenko, V. Petrov, O. Galinina, S. Andreev, and Y. Koucheryavy, "Energy and Delay Analysis of LTE-Advanced RACH Performance under MTC Overload," in Proc. of the IEEE GLOBECOM Workshops, pp. 1632–1637, 2012.

[10] T. Tirronen, A. Larmo, J. Sachs, B. Lindoff, and N. Wiberg, "Reducing energy consumption of LTE devices for machine-to-machine communication," in Proc. of the IEEE GLOBECOM Workshops, pp. 1650–1656, 2012.

[11] T. Tirronen, A. Larmo, J. Sachs, B. Lindoff, and N. Wiberg, "Machine-to-machine communication with LTE with reduced device energy consumption," Transactions on Emerging Telecommunications Technologies, vol. 24, pp. 413–426, 2013.

[12] S. Andreev, A. Larmo, M. Gerasimenko, V. Petrov, O. Galinina, T. Tirronen, J. Torsner, and Y. Koucheryavy, "Efficient small data access for machine-type communications in LTE," in Proc. of the IEEE ICC, pp. 3569–3574, 2013.

[13] O. Galinina, A. Turlikov, S. Andreev, and Y. Koucheryavy, "Stabilizing multi-channel slotted Aloha for machine-type communications," in Proc. of the IEEE ISIT, pp. 2119–2123, 2013.

[14] A. Ahmadian, O. Galinina, S. Andreev, and Y. Koucheryavy, "Modeling Contention-Based M2M Transmissions over 3GPP LTE Cellular Networks," in Proc. of the IEEE ICC Workshops, pp. 441–447, 2014.

[15] D. Astely, E. Dahlman, G. Fodor, S. Parkvall, and J. Sachs, "LTE Release 12 and beyond," IEEE Communications Magazine, vol. 51, pp. 154–160, 2013.

[16] Ericsson, More than 50 billion connected devices. White Paper, February 2011.

[17] K. Zheng, F. Hu, W. Wang, W. Xiang, and M. Dohler, "Radio resource allocation in LTE-Advanced cellular networks with M2M communications," IEEE Communications Magazine, vol. 50, pp. 184–192, 2012.

Sergey Andreev is a Senior Research Scientist in the Department of Electronics and Communications Engineering at Tampere University of Technology, Finland. He received the Ph.D. degree (2012) from Tampere University of Technology. Sergey (co-)authored more than 80 published research works on wireless communications, energy efficiency, heterogeneous networking, cooperative communications, and machine-to-machine applications.

Tuomas Tirronen has been working as wireless access networks researcher at Ericsson Research since 2012. He received his D.Sc. in Communications Engineering in 2010 from Aalto University. His current research interests include 4G and 5G technologies, Internet of Things, performance evaluation, radio protocols and resources. He is also active in 3GPP standardization work and innovation and patenting.

Johan Torsner is a research manager in Ericsson Research and is currently leading Ericsson's research activities in Finland. He joined Ericsson in 1998 and has held several positions within research and R&D. He has been deeply involved in the development and standardization of 3G and 4G systems and has filed over 100 patent applications. His current research interests include 4G evolution, 5G and machine-type communication.

Yevgeni Koucheryavy is a Full Professor and Lab Director at the Department of Electronics and Communications Engineering of Tampere University of Technology (TUT), Finland. He received his Ph.D. degree (2004) from TUT. He is the author of numerous publications in the field of advanced wired and wireless networking and communications. His current research interests include various aspects in heterogeneous wireless communication networks and systems, the Internet of Things and its standardization, as well as nano-communications.

Device To Device For Wearable Communication in 5G Networks

Khaldoun Al Agha
Green Communications
khaldoun.alagha@green-communications.fr

1. Introduction

Telecommunication technologies in all their forms (2G to 4G, Wi-Fi, Wi-Max…) [1] use a unique model to communicate between two entities. It consists in forwarding the packets to an access point which sends the information back to the destination (Figure 12-a). This happens when source and destination are in the same cell. If source and destination are located in two different cells, then the information uses a hierarchical path based on a tree topology going up and then down to the final destination as shown on Figure 12-b.

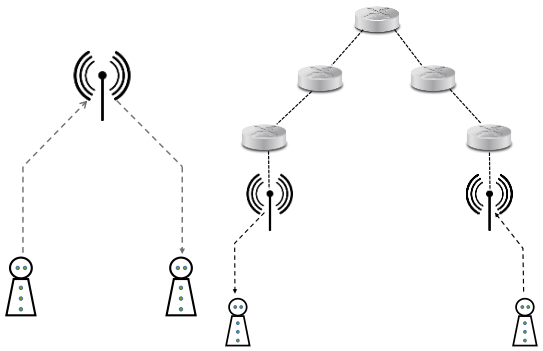


Figure 12. (a) Communication in one cell; (b) in different cells

This model implies many disadvantages in terms of capacity and energy consumption:

1) When many devices want to access to the network in the same cell (Figure 13), they have to send the information to the access point serving the cell. However, the access point can only process one request at the same time. If many requests arrive simultaneously, they generate collisions and interferences. A scheduling mechanism needs to be elaborated at the access point in order to organize the access to the network, request after request. Parallel request cannot be served which limits the capacity of the cell.

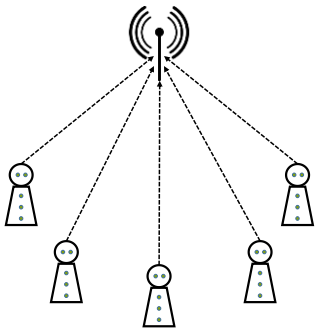


Figure 13. Scheduling at the access point

2) In terms of energy consumption, transferring the information to an intermediate access point instead of direct transfer doubles the number of transmissions/receptions. Also, when using a tree-topology path to climb up then down on a vertical way will increase the energy per bit comparing to a direct and horizontal path. Figure 14 illustrates the difference between horizontal and vertical communications.

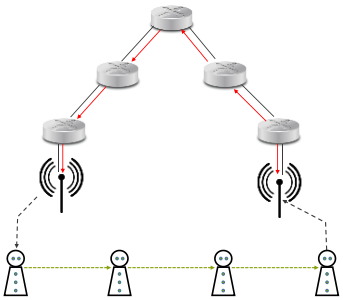


Figure 14. Horizontal versus Vertical communications.

The objectives of the next generation 5G is to increase the network capacity by introducing a Device to Device technology. Using direct communications would increase the capacity of the networks since parallel communications could happen comparing to an access point centralized scheduling. Direct communications do not need to communicate with high power since the destination is very close to the source. The low power reduces the energy consumption and increases the capacity by reducing the interferences.

2. Observation on the traffic

When scanning the traffic on telecommunication networks, one could observe that an important part of the traffic is local. People during events or in smart cities and smart homes, share pictures and videos, exchange information on their environment. The information in the majority of cases do not need to travel through the infrastructure of the network and the cloud and to consume a huge amount of resources and energy to come back to the neighborhood. Although, the video traffic [2] on mobile devices is doubling every year, damaging seriously the capacity at access points by creating bottlenecks.

3. Embedded Internet and 5G

The idea to go directly from source to destination is not new. Its implementation was never realized in the past because operators of networks had to control the traffic for accounting and monitoring. Nowadays, the majority of subscriptions include an unlimited usage of the network. Hence, the more the traffic is out of the operator network, the less the infrastructure of the operator is saturated while its income stays the same. Introducing horizontal and direct communications is a win-win policy for operators, customers and also the entire eco-system. Why, should local traffic travel over the entire network and data centers to just reach a close destination?

5G networks include in their requirements the faculty that two devices should communicate directly with no need of any network infrastructure. This implies that the device should contain a minimum of network intelligence to have the capacity to detect destinations and to also transfer the traffic of its peers. In addition, Device2Device communication should continue to work even when the global network is not operating or when damages push the global network offline.

Efforts to make intelligence in the applications of Smart Phones is an option to introduce the Device2Device. Some products are already working on different platforms (IOS, Android) such as AirDrop [4] or Firechat [5]. In those cases, the Device2Device is managed at the application layer and with a proprietary protocol. It means that Smart Phones should first install the application to run a Device2Device. Furthermore, each new service requires the development and download of a new application. This model cannot scale.

4. Wearable and Embedded Internet

Green Communications [3] introduces the concept of Embedded Internet where devices hold a full TCP/IP environment, an open operating system (Buildroot-based Linux) and a sharing storage. Each device contains two wireless network interfaces. One to establish the network between the devices and the second to offer an access to SmartPhones, tablets and computers. All type of applications will be active immediately because the network is running a TCP/IP locally. It is also possible to provide many applications over http using a local web server and latest web technologies (htmls). This gives the opportunity to anyone to communicate even if they didn't download a native application.

The device is light enough to be held or wearable. It could be distributed on-site or could come directly within the visitors. When turning the devices ON, the network is established in few seconds and then a local Internet is operated. A subset of the devices could be configured as gateways to the global Internet to create a hybrid model where local traffic stay local and global traffic is transferred across the gateways.

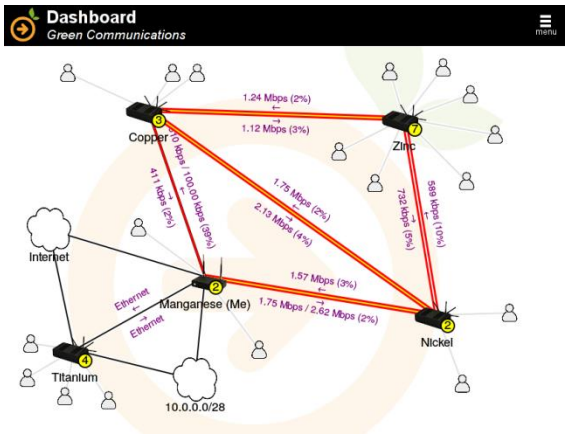


Figure 15. Monitoring interface with open data.

5. Local Cloud

The Green Communications device also includes storage with security options to offer the possibility to share information and also to install new specific applications related to the area where the network is running. The network provides access to its raw open data to help people in their connections and enables application developers to build more appropriate applications related to on-site usage. Figure 15 shows the different kind data that are open: network topology, number of users per access point, quality of links between access points, location of the Internet Gateways, access point profiles through the vCard Standard

6. Distributed TCP/IP

In order for an embedded Internet to work correctly, it is mandatory to distribute all the services over the devices. Services shouldn't depend on one specific machine to avoid any network blackout. When a device enters or leaves the network, services must continue to

work correctly. The solution should also provide distributed TCP/IP over the devices by enhancing them with Zeroconf (mDNS, DNS-SD), anycast addresses, handoff management, distributed DHCP implementation, distributed database.

7. Conclusions

Communicating locally for local traffic is simply the common sense in the telecommunication systems. Until now, all networks never proposed any Device2Device communications. Within the 5G, this type of communication will be standardized allowing an enhancement of the network capacity and its energy consumption.

References

[1] http://en.wikipedia.org/wiki/Cellular_network
[2] Cisco's Visual Network Index Cisco's Visual Network Index 2013-2018
[3] http://www.green-communications.fr
[4] http://en.wikipedia.org/wiki/AirDrop
[5] http://en.wikipedia.org/wiki/FireChat

Khaldoun Al Agha is CEO/Co-Founder of Green Communications and full professor at Paris-Sud University.
Khaldoun Al Agha received his habilitation degree (2002) from Paris XI University, his PhD (1998) and his Master degree (1995) from Versailles University and his engineering degree (1993) from the Ecole Supérieure d'Electricité (1993).
From 2010-2013, he was leading at EIT ICTLABS, the European action line "Digital Cities of the Future". In that action line, he was developing a new model, citizen-centric based to improve the urban environment. Khaldoun Al Agha is leading many projects on telecommunication networks and published more than 150 papers in journals and conference.

Big data for Cloud based Video Adaptation

Mamadou Tourad Diallo[1], Frédéric Fieau[1], Hossam Afifi[2]
[1]Orange Labs, France, [2]Institut Mines-Télécom France
{mamadoutourad.diallo, frederic.fieau}@orange.com, hossam.afifi@mines-telecom.fr

1. Introduction

Nowadays, video service delivery faces two challenges: how to use big data to optimize decisions? How to create new services? Many service providers and operators are struggling to maintain the Average Revenue per User (ARPU) and reduce hardware CAPEX/OPEX despite the high competition in the video services market. Assessing Quality of Experience can allow them to understand how the end-users perceive the quality of video and audio. They are more and more using big data techniques to store data and assess Quality of Experience (QoE) aiming at providing new dynamic adaptation methods.

Two types of adaptation can be considered to optimize QoE: content based adaptation and network/resources based adaptation. On the one hand, content based adaptation allows adapting the video choosing the right video codec, bitrate and layers that will be sent to the player. On the other hand, network and resources based adaptation relates to which mode will be used to deliver content to the end-users (e.g. unicast/multicast), through which network, and what is the required hardware.

Furthermore, underlying architectures and hardware resources can affect QoE. Since the early stages of the Internet, CDN (Content Delivery Network) service providers have been deploying geographically distributed servers to extend their footprint, locating them closer to end-users in order to reduce the delay and improve QoE. Nowadays, OTT (Over The Top) actors (like Amazon, Microsoft) and Telcos provide Cloud "X-as–a-Service" (XaaS) offers where X may for instance, stand for Infrastructure, Platform or Software. At a technical level, these offers are mainly supported by centralized infrastructures using automated processes and virtualization as an enabler. Infrastructure virtualization facilitates server operations allowing fast instantiation of servers (scale in/out) and upgrade/downgrade (scale up/down).

As said in [1], to take an adequate adaptation decision, we should consider QoE coupled with context data on the user, devices, network, and the content itself to take the adequate adaptation decision. In this work, we first propose to collect information in big data approach, next, we use these data for better assessing Quality of Experience regarding collected information on user, devices, network, hardware and contents. Finally a Cloud based architecture for Video Delivery Adaptation is suggested to optimize networks resources and hardware and improve QoE.

2. Big data information

For an accurate estimate of QoE, we consider context data. The notion of context allows taking into account information coming from a user and his environment. Data need to be dynamically gathered in real-time during service delivery. Collected data may include: i): device context (capacity, screen size …), ii) network context (jitter, packet loss, available bandwidth …), iii) user context (who is the user, his preferences, his consumption style, gender, age…), iv) service context (type of contents, available video bitrates) and v) and hardware context (available memory, CPU, storage capacity …).

Figure 1 shows the considered context data in our study.

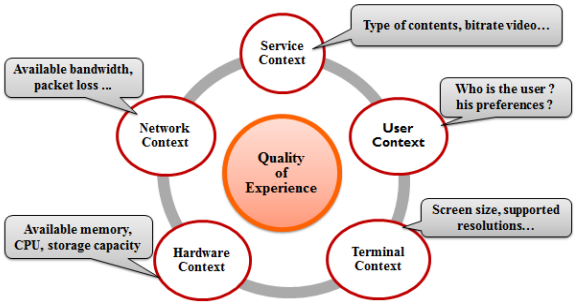


Figure 16. Considered context data

3. Quality of Experience assessment and video optimization

In this section, we first introduce classical objective and subjective methodologies in the literature. Secondly, we describe our proposal for QoE assessment and thirdly, we describe the video service optimization problem.

3.1 Related work

The following works either use objective or subjective approaches to measure Quality of Experience (QoE).

First of all, objective approaches don't require end-user's feedback to predict the quality of the perceived content. An example of comparative method is the PSNR (Peak Signal-to-Noise Ratio) which is based on the power balance between the native information and

the background noise generated by distortion after compression or recompression. It is defined via the Mean Squared Error (MSE) between an original frame o and the distorted frame d as presented in [2]. On the contrary, some direct approaches do not require reference information on the original content. For instance, the method showed in [3] is based on a metric that quantifies blurring phenomena along edges and contours in video images due to video degradation. Please note that the required parameters and inputs are all available on client side.

Subjective methodologies are the most fundamental approaches for evaluating QoE. They are based on surveys, interviews of customers and statistical sampling of their answers to analyze their perception and needs with respect to the service and network quality. Several subjective assessment methods suitable for video application have been recommended, the P.910 [4] and ITU-R Recommendation BT 500-11[5].

Furthermore, these techniques have some drawbacks. The objectives ones do not consider the user's opinion about streamed content and also context data are not considered. For the subjective ones, it might not be realistic to ask all viewers their scores as it can be very annoying for them.

In this work, we thus propose to use context data in order to assess Quality of Experience.

3.2 Proposed Quality of Experience assessment model and Optimization problem

The proposed QoE assessment model takes into account the data showed in the precedent section. We divided the model in three parts: the network effects, the video quality effects and the hardware impacts.

First, we analyze the impact of network degradation which is function of network with respect to different content types and devices types as our work in [6].

By using subjective data coming from past experiments, we build a parametric model. It takes into account context parameters such as the device type, the video content type, video bitrate and the quality of the network link. The analytical is showed by the equation below:

$$Nsi,j = \alpha+\beta*e-\delta*Dv(j)Dr(i) \quad (1)$$

- $N_{s(i,j)}$ corresponds to the satisfaction obtained by the end-users in network i for flow j.
- α, β and δ, are the model parameters calculated by using subjective test data from different experiments.
- D_r, is the available throughput and D_v, the bitrate video.

Secondly, we analyze the effect of video quality which is function of encoding variation with respect to different content types and device categories As in the network impact study, by using past experiments data, the developed model is:

$$VQsj=\mu1* \log Dvj+\mu2 \quad (2)$$

- $VQ_{s(j)}$ is the satisfaction obtained in video quality.
- Dvj) is the video bitrate for flow j.
- μ_1 and $\mu2$ are the model parameters

And finally, the score from hardware can be computed using the formula below:

$$Hsk = k=1n\alpha k*Hk \quad (3)$$

- $H_{s(k)}$ corresponds to the hardware score for server k
- Hk is the hardware parameter (CPU, storage capacity, memory, inbound and outbound bandwidth) for server k
- α_k is the weight of each hardware parameter.

The proposed general QoE assessment model is:

$$\text{QoE}(i,j,k)=\delta1*Nsi,j+\delta2* VQ_{s(j)}+ \delta3* Hsk \quad (4)$$

where**:** $\delta1$, $\delta2$ and $\delta3$, are the weights of entities in the global Quality of Experience and $\delta1+\delta2+\delta3=1$.

After computing scores on network, hardware and video quality, we describe the optimization problem with the constraints:

Objective: maximize (QoE(i,j,k))

Subject to: $D_r < D_{rmax}$

$VQ \epsilon [VQ1 \ VQ2.....VQmax]$

$CPU < CPU_{max}$

$mem < mem_{max}$

$storage < storagemax$

where:

- D_r, is the bandwidth that should be allocated and D_{rmax} is the maximum bandwidth the network can allocate.
- VQ is the video quality. For each content, the chosen quality belongs to the finite set [$VQ1$ $VQ2.....VQmax$]
- $CPUmax$,$memmax$,$storagemax$are respectively the maximum of CPU, memory and storage that can be allocated.

4. Video Delivery Adaptation

We can then formalize the possible choices leading to the optimal decision, with the following triplet (N_i, VQ_j, H_k). The decision is based on the previous parameters. It includes the following possibilities:

1) Deliver content from another network (Ni). e.g., network1 offers better conditions for delivering the considered content.
2) Change video quality: VQj, the encoding quality among available content encoding qualities.
3) Deliver content from another server (H_k), including:

- Deliver content using another server, or,
- Upgrade (e.g., scaling-up) the current virtual server characteristics. For instance this would mean increasing the amount of allocated memory or number CPUs, or bandwidth, if support hardware is available.

H_k is the server number k, for delivery, where its characteristics vary following an n-uplet (CPU, storage capacity, memory, inbound and outbound bandwidth).

The optimal decision may be computed for each session request. The terminal/player would be either redirected to another server, network or would be requesting another video encoding rate. The decision enables to have optimal QoE in given conditions.

The following figure describes the proposed solution architecture, where $\mathbf{N_{opt}}$, $\mathbf{VQ_{opt}}$ and **Hopt** are optimal parameters.

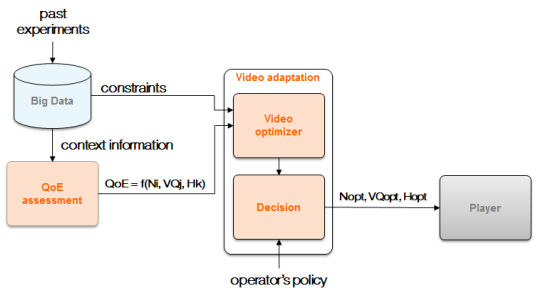


Figure 2. Proposed solution architecture

5. Conclusion

With the explosion of multimedia and audiovisual services and also the competition between service providers, Quality of Experience has become critical for them to continue gaining users' satisfaction.

In this work, we proposed to use context data to assess Quality of Experience. The proposed model is function of device, network, service contexts and hardware parameters. Secondly, we proposed an optimization decision, which enables having optimal Quality of Experience in given conditions. It takes into account the network status, hardware parameters and video encoding. The optimization decision can be ensured following three options: redirecting the terminal to another server/network, requesting another video encoding rate, or upgrading the characteristics of the server (storage capacity, memory …).

The optimization of the virtual server's characteristics may also reduce the needed CAPEX/OPEX allocated to a given server while respecting the desired QoE, which is essential in the operational context.

References:

[1] M.T. Diallo, H. Moustafa, H. Afifi, N. Maréchal: "Adaptation of audiovisual contents and their delivery means" Commun. ACM 56(11): 86-93, 2013.

[2] A. N. Netravali and B. G. Haskell, "Digital Pictures: Representation, Compression, and Standards (2nd Ed)," 1995.

[3] P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahimi, "A no-reference perceptual blur metric," Image Processing. International Conference on Vol. 3, 2002.

[4] ITU-T, P.910, "Subjective Video Quality Assessment Methods for Multimedia Applications", 1999

[5] ITU-R,BT 500-11, "Methodology for the Subjective assessment of the quality of Television Pictures," 2002

[6] M.T. Diallo, N. Maréchal, and H. Afifi, "A Hybrid Contextual User Perception Model for Streamed Video Quality Assessment", IEEE International Symposium on Multimedia (ISM), pp.518,519, 9-11 Dec. 2013.

Mamadou Tourad Diallo received his M.Sc. from Rennes University in 2011. He is currently working on his Ph.D. within Orange Labs. His thesis work concerns Quality of Experience measurement, new video content and delivery adaptation, based on context information.

Frédéric Fieau is a R&D engineer in Orange Labs, he received his engineer degree from Telecom Bretagne, France. He is now working as a network architect on Cloud and Content Delivery Networks architectures, video delivery adaptation and infrastructures virtualization.

Hossam Afifi is a professor at Institut Mines Telecom, Saclay, France. He works on multimedia services and applications.

Overview of Big Data Video Analytics

Mohamed Rehan, AvidBeam Technologies
Mohamed.rehan@avidbeam.com

1. Introduction

Big data analytics is becoming more and more important. These analytics provide substantial information that can be used for finding out problems, identify important patterns or behavior, or provide statistical information regarding certain product or service. There are currently several framework dedicated for processing big data and providing the required analytics at high efficiency such as Apache Hadoop framework[1,2].

With the growing popularity of using video contents in many of these applications and the availability of high speed networks. Video data is becoming the major portion of internet traffic. In addition, the emergence of new technologies that enable capturing, encoding, and transmitting or storing video at higher quality and higher resolution, many video related applications, such as video surveillance, video broadcasting, require the processing of huge video data on a continuous bases. As a result, the need for intelligent video analytics platform for big data is become more and more urgent.

Video analytics can be very important and effective for many clients who depends on video contents as their primary data. It can be used to spot unexpected problems or serious or alarming situations. By identifying problems, solutions can be proposed, problems can be avoided, and revenues can increase.

The following sections presents typical use cases for video analytics solutions, the main building blocks of any video analytics solutions, and the main challenges presents in today's video analytics systems.

2. Video Analytics Applications

There are many applications that can be improved with the use of video analytics. Examples of those applications include

- Surveillance (traffic, crowded areas, private areas): many useful applications such as intruder detection and tracking, detection of unattended luggage in airports, suspicious activities, theft, loiters, Finding out lost items, missing persons. In traffic, identify traffic patterns and analysis of traffic flow which could help divert car and trucks from heavy traffic roads to roads with lighter traffic. It can help identify traffic violators such as speeding car, car driving in the wrong or unallowable direction, hit and run accidents. Location of stolen cars, etc. In parking lots as shown in Figure 1, video analytics can be used to analyze client pattern, car counting
- Broadcasting (sports, streaming): in sports video analytics can be used to analyze player activities and performance, detecting improper behavior. In streaming, Analytics can provide details about popular streamed contents or visited web site per age or gender group.
- Marketing: Analysis of customer's behavior in large malls or stores can help identify customer preference (store, product) and speed up purchase and service process. Identify frequently needed products
- Others: Medical, education, etc.: similar use cases can be derived for different applications.

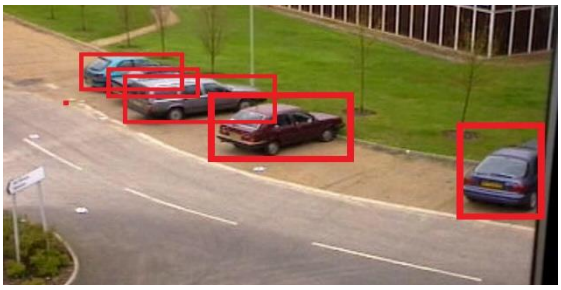


Figure 1.Video Analytics Example: Counting Cars in a Parking Lot (PETS DataSet[3,4])

3. Video Analytics System Components

A typical video analytic system has both hardware and software requirements. The hardware system consists of a group of camera connected to a central server(s). The cameras are used to capture video from certain position and the captured video can then be encoded and streamed to the analytics server. In some cases, camera can perform other tasks such as motion detection, line or area crossing, etc. The analytics server is used to operate or control the cameras, process or store incoming video feeds and perform other required tasks such as intruder detection, recognition, and alarming if necessary. Big data processing is accomplished by using numbers of servers in addition to the central server in order to improve the processing performance. The central

server will be responsible for data and processing distribution. Cameras are connected to the server(s) with high speed network as shown in Figure 2. The choice of the camera capabilities, their placement and orientation as well as the server(s) capabilities is usually related to target applications.

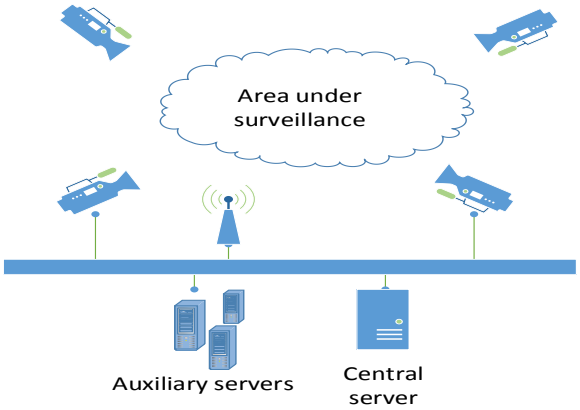


Figure 2 Video Analytics Setup.

Figure 3 shows the main software components of a typical video analytics software. In more complex systems some of those components can further be split.

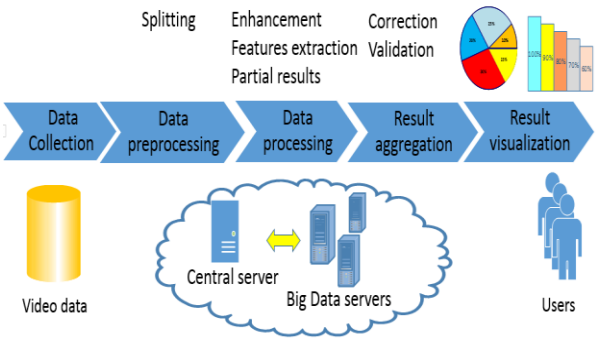


Figure 3 Components of Video Analytics Platform.

The use of each component can be described as follows

A. **Video Data Collection**: Data can be online or offline from stored files or cameras. Video data can be stored in local or remote files. Number of cameras or video feeds vary from small or limited, as in case of small office surveillance, to extremely large or huge, as the case of city traffic management system.

B. **Video data preparation**: Video data needed to be prepared so it can be processed in parallel or in multiple device. The scalability of big data and algorithms at this stage is very critical. The use of big data platform should facilitate this process and provide the necessary scalability to achieve target performance.

C. **Video data processing:** This is where the actual processing of split data chunks occurs. The choice of a process to be executed on data depends on the required feature(s), The main process in the central server should balance and manage process data among existing server(s). Each process handles data as a standalone set and produces partial results. In many scenarios, the processing of data is used to extract the important or required features from this video data.

D. **Result Aggregation:** When all chunks are processed, collecting results together to produce the combined or final results is applied. The aggregation represents the validation, summarization, and other processing on the extracted features from data processing step.

E. **Result Visualization**: This is the final step where the produced results is represented in a simplified form such as presentation, chart, or summary report.

Hadoop platform, for example, uses MAPReduce[5] approach to process big data in multiple servers. In this case, data is split into chunks of defined size. In the MAP process, chunks are distributed to multiple servers, and processed as intended. In the reduce process, the results from separate MAP processes are collected, validated, and combined to produce the intended outcome.

Some of the simple examples of MAP/Reduce approach are face or object counting or histogram calculation for Gigabyte images. In this case, the image is split into blocks, each block is processed in a separate MAP process which computes number of faces, objects, or colors in each block separately. Once the MAP processes are completed, the partial results are aggregated in the Reduce process to provide the total results which can later be presented by a graph.

4. Video Analytics Challenges

There are several common problems associated with the generic big data, such as log files, and Video big data. For example, all big data needed to be stored, transferred, and processed at reasonable time. However, video big data includes additional challenges that needed to be tackled independently. This section provides highlights on some of these challenges.

A. **Management Video Data Size:** Video data size can grow rapidly to Gigabytes, Terabytes, or Petabytes range based on many parameters such as number of video feeds, resolution and durations of those videos. In other big data types, few Gigabytes can represent a collection of data over a month duration while for video big data, the same amount of Gigabytes could just represents few hours of videos even if video encoding technique is being used[6].

This challenge puts huge pressure on storage and transmission resources and dictates whether the video analytics processes should work remotely or locally.

In addition to video compression, there are several approaches that can be used to reduce the need to process or transfer large amount of videos such as preprocessing video on capturing devices, or identifying important video frames only. For example, motion sensors can detect motion after which key video frames should be captured.

B. **Management of Compressed Video:** Compressing video content is a must when dealing with big video data. However, compression algorithms rely heavily on removing video redundancies which results in having the encoded frames dependents on other frames, usually prior frames in time.

Splitting video big data into chunks should remove the inter-chunk dependencies so each chunk can be processed as a single video stream. This process could require large amount of computation and parsing through the video streams for proper split position.

Another inherited problem in compressed video is the degradation of video quality due to the use of lossy compression algorithms such as H.264 or HEVC[6]. One of the possible solution to this problem is the select the proper recording quality based on the activity under observation. When such activity occur, the video should be recorded at highest possible quality.

C. **Accuracy of Computer Vision Algorithms:** Most video analytics solutions require the use of computer vision algorithms for detection, recognition, and tracking of different types of objects such as faces, cars, pedestrian, or any other objects of interest [7,8]. The efficiency of computer vision algorithms can be dramatically affected by many factors such as

- Crowded areas or locations of objects moving in complex motion patterns. It is very likely that the object under investigation can be blocked with other objects in the scene.
- Face orientation change or simple changes in facial features such as taking off or putting on glasses, hats, or head cover. Those simple items can fails most of current recognition algorithm
- Environmental conditions such as rainy, foggy, cloudy weather which can affect video capture quality. Other daily lighting changes or presence of shades can also result in false detection.
- Equipment conditions such as camera zooming or angle of placement.
- Performance requirements. Both Computer vision and video coding algorithms are very computationally intensive and require high performance computing.

The above described problems often results in approximated results which can be useful in many applications. However, in applications which require higher accuracy, other solutions are needed. For example, use of semi-automatic analytics where human involvement is used to improve the accuracy, use of additional sensors such as RFID, or increase number of cameras or their capabilities and provide strong coordination between them. On top of that, advanced research that yields computer vision algorithms with higher accuracy is still needed.

5. Conclusions

This paper provided an overview of big data video analytics. The paper discussed the basic building components of a typical video analytics platform and highlighted the most common challenges and limitations associated with existing solutions.

References

[1] Apache Hadoop. Available at http://hadoop.apache.org.

[2] M. Zaharia, M. Chowdhury, M. J. Franklin, S. Shenker, and I. Stoica, "Spark: Cluster Computing with Working Sets", HotCloud 2010. June 2010.

[3] MapReduce: Simplified Data Processing on Large Clusters. Available at http://labs.google.com/papers/mapreduce-osdi04.pdf.

[4] D. Thirde, L. Li, and F. Ferryman. Overview of the PETS2006 challenge. In Proceedings of IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, pages 47-50, 2006.

[5] R. Fisher, The PETS04 surveillance ground-truth data sets. In Proceedings of IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, pages 1-5, 2004.

[6] G. J. Sullivan, J. Ohm, W. Han, and T, Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Transaction on Circuits and Systems for Video Technology, vol. 22, No. 12, Dec 2012.

[7] Vizzari, G.; and Bandini, S. "Studying Pedestrian and Crowd Dynamics through Integrated Analysis and Synthesis", IEEE Intelligent Systems, page(s): 56 - 60 Volume: 28, Issue: 5, Sept.-Oct. 2013.

[8] X. Wang, X. Ma and W. E. L. Grimson,"Unsupervised activity perception in crowded and complicated scenes using hierarchical Bayesian models", IEEE Trans.Pattern Analysis and Machine Intelligence, vol. 31, no. 3, pp.539 -555 2009.

Mohamed Rehan Dr. Rehan is the chief technical officer of AvidBeam Technologies which specialized in intelligent video surveillance and analytics solutions, He obtained his Ph.D. from the University of Victoria, Victoria, BC, Canada in video coding in 2006. He received his B.S. in communications and his M.S. in computer graphics in 1991 and 1994 respectively from the department of electronics and communications, Cairo University, Cairo, Egypt. Dr. Rehan has been involved in multimedia research and development for more than 20 years. He has an extensive set of publications in video coding and multimedia.

Call for Papers

IEEE Transactions on Multimedia

Special Issue on "Multimedia: The Biggest Big Data"

Summary

Multimedia is increasingly becoming the "biggest big data" as the most important and valuable source for insights and information. It covers from everyone's experiences to everything happening in the world. There will be lots of multimedia big data --- surveillance video, entertainment and social media, medical images, consumer images, voice and video, to name a few, only if their volumes grow to the extent that the traditional multimedia processing and analysis systems cannot handle effectively. As such, multimedia big data will emerge as the next "must have" competency in our society, and is spurring on tremendous amounts of research and development of related technologies and applications. As an active and inter-disciplinary research field, multimedia big data also presents a great opportunity for multimedia computing in the big data era. The challenges and opportunities highlighted in this field will foster some interesting future developments in the multimedia research and applications.

Scope

The goal of this special issue is to provide a premier forum for researchers working on the aforementioned multimedia big data aspects to present their recent research results. It also provides an important opportunity for multidisciplinary works connecting big data to multimedia computing. Topics of interest include, but are not limited to:

- New theory and models for multimedia big data computing
- Ultra-high efficiency compression, coding and transmission for multimedia big data
- Content analysis and mining for multimedia big data
- Semantic retrieval of multimedia big data
- Deep learning and cloud computing for multimedia big data
- Green computing for multimedia big data (e.g., high efficiency storage)
- Security and privacy in multimedia big data
- Interaction, access, visualization of multimedia big data
- Multimedia big data systems
- Novel and incentive applications of multimedia big data in various fields (e.g., search, healthcare, transportation, and retail)

Important dates

Submission deadline: **February 28, 2015**
First notification: April 28, 2015
Revision due: May 31, 2015
Final notification of acceptance: July 5, 2015
Camera-ready manuscript due: July 21, 2015
Tentative publication date: August 2015

Submission procedure

Papers should be formatted according to the IEEE Transactions on Multimedia guidelines for authors (see: http://www.signalprocessingsociety.org/tmm/tmm-author-info/). By submitting/resubmitting your manuscript to this transactions, you are acknowledging that you accept the rules established for publication of manuscripts, including agreement to pay all over-length page charges, color charges, and any other charges and fees associated with publication of the manuscript. Manuscripts (both 1-column and 2-column versions are required) should be submitted electronically through the online IEEE manuscript submission system at http://mc.manuscriptcentral.com/tmm-ieee. When selecting a manuscript type, the authors must click on BigMM Special Issue. All the submitted papers will go through the same review process as that for the regular TMM paper submissions. Referees will consider originality, significance, technical soundness, clarity of exposition, and relevance to the special issue topics above.

Guest Editors

Shu-Ching Chen, Florida International University, USA (chens@cs.fiu.edu)
Ramesh Jain, University of California, Irvine, USA (jain@ics.uci.edu)
Yonghong Tian, Peking University, China (yhtian@pku.edu.cn)
Haohong Wang, TCL Research America, USA (haohongwang@gmail.com)

Call for Papers

IEEE Transactions on Cloud Computing

Special Issue on "Mobile Clouds"

Mobile cloud computing represents one of the latest developments in cloud computing advancement. In particular, mobile cloud computing extends cloud computing services to the mobile domain by enabling mobile applications to access extern al computing and storage resources available in the cloud. Not only mobile applications are no longer limited by the computing and data storage limitations within mobile devices, nevertheless adequate offloading of computation intensive processes also has the potential to prolong the battery life.

Besides, there is also an incentive for mobile devices to host foreign processes. This represents a new type of mobile cloud computing services. Ad-hoc mobile cloud is one instance that mobile users sharing common interest in a particular task such as image processing of a local happening can seek collaborative effort to share processing and outcomes. Vehicular cloud computing is another instance of mobile cloud computing that exploits local sensing data and processing of vehicles to enhance Intelligent Transportation Systems.

This Special Issue will collect papers on new technologies to achieve realization of mobile cloud computing as well as new ideas in mobile cloud computing applications and services. The contributions to this Special Issue may present novel ideas, models, methodologies, system design, experiments and benchmarks for performance evaluation. This special issue also welcomes relevant research surveys. Topics of interest include, but are not limited to:

- Trends in Mobile cloud applications and services
- Architectures for mobile cloud applications and services
- Mobile cloud computing for rich media applications
- Service discovery and interest matching in mobile cloud
- Collaboration in mobile clouds
- Process offloading for mobile cloud computing
- Mobile device virtualization
- Mobile networks for cloud computing Mobile cloud monitoring and management
- Security and privacy in mobile clouds
- Performance evaluation of mobile cloud computing and networks
- Scalability of mobile cloud networks
- Software defined systems for mobile clouds
- Self-organising mobile clouds
- Mobile vehicular clouds
- Disaster recovery in mobile clouds
- Economic, social and environmental impact of mobile clouds
- Mobile cloud software architecture

Important Dates

Paper submission: **February 1, 2015**
First Round Decisions: May 15, 2015
Major Revisions Due: June 15, 2015
Final Decisions: August 15, 2015
Publication: 2016

Submission & Major Guidelines

This special issue invites original research papers that present novel ideas and encourages submission of "extended versions" of 2-3 Best Papers from the IEEE Mobile Cloud 2015 conference. Every submitted paper will receive at least three reviews and will be selected based on the originality, technical contribution, and scope. Submitted articles must not have been previously published or currently submitted for publication elsewhere. Papers should be submitted directly to the IEEE TCC at https://mc.manuscriptcentral.com/tcc, and must follow TCC formatting guidelines. For additional information, please contact Chuan Heng Foh (c.foh@surrey.ac.uk).

Editor-in-Chief

Rajkumar Buyya, the University of Melbourne, Australia

Guest Editors

- Chuan Heng Foh, University of Surrey, UK
- Satish Narayana Srirama, University of Tartu, Estonia
- Elhadj Benkhelifa, Staffordshire University, UK
- Burak Kantarci, Clarkson University, USA
- Periklis Chatzimisios, Alexander TEI of Thessaloniki, Greece
- Jinsong Wu, Alcatel Lucent, China

Call for Papers

IEEE Workshop on LTE in Unlicensed Bands: Potentials and Challenges

http://www.lte-u.com

Organized in conjunction with

IEEE International Conference on Communications (IEEE ICC 2015)

8-12 June 2015, London, UK

The exponential growth of mobile data traffic and the scarcity and costliness of licensed spectrum are driving mobile network operators (MNOs) to consider offloading at least part of their traffic onto the unlicensed spectrum. Most recently, the 3GPP is considering extending the use of LTE into the unlicensed spectrum as a seamless approach to enable traffic offload. This new approach is dubbed LTE Unlicensed (LTE-U). Compared to Wi-Fi, LTE-U offers MNOs a way to offload traffic onto the unlicensed spectrum with a technology that seamlessly integrates into their existing LTE evolved packet core (EPC) architecture. Furthermore, LTE-U promises higher throughout and spectral efficiency than Wi-Fi, with estimates ranging from 2x to 5x improvement over Wi-Fi. Currently two operating modes are under discussion: 1) unlicensed spectrum is aggregated with existing licensed channels and 2) unlicensed spectrum acts as the only carrier for LTE-U where both data and control channels reside. The liberal non-exclusive use of unlicensed spectrum has spurred innovation on the one hand, but has also created the need for coexistence measures when various uncoordinated wireless networks operate on the same frequency. In this case, LTE-U introduces new coexistence challenges for other technologies operating in the same unlicensed bands particularly for legacy Wi-Fi. Wi-Fi is designed to coexist with other technologies through channel sensing and random backoff, while LTE is designed with the assumption that one operator has exclusive control of a given spectrum. Furthermore, LTE traffic channels are designed to continuously transmit with minimum time gap even in the absence of data traffic. Consequently, Wi-Fi users will have little chance to sense a clear channel and deem it suitable for transmission. The goal of this full-day workshop is to bring together academics, researchers, and practitioners to discuss the opportunities, challenges and potential solutions for operation of LTE in the unlicensed bands. The topics of interest include, but are not limited to:

- Coexistence of schedule-based and contention-based networks in unlicensed bands
- Fairness considerations for coexistence of LTE and Wi-Fi
- Performance impact of LTE on networks operating in the unlicensed band
- Radio resource management, dynamic channel selection and band steering for LTE/WiFi coexistence
- Traffic demand-aware coexistence
- Distributed and centralized techniques for coexistence of heterogeneous networks
- Technical challenges and solutions of operating LTE solely on the unlicensed bands
- QoS model for standalone LTE-U access model

Paper submission link: https://edas.info/newPaper.php?c=18693&track=66309

Information for Authors: Prospective authors are invited to submit original technical papers by the deadline **January 20, 2015**. All submissions should be written in English with a maximum paper length of Six (6) printed pages (10-point font) including figures without incurring additional page charges (maximum 1 additional page with over length page charge if accepted). Please also see the Section in the main ICC 2015 website for Authors Guidelines.

Registration: Please see the Section in the main ICC 2015 website on Registration.

Workshop Organizers

Dr. Andrea Goldsmith, Stanford University, USA (andrea@wsl.stanford.edu)
Dr. Alireza Babaei, CableLabs, USA (a.babaei@cablelabs.com)
Dr. Jennifer Andreoli-Fang, CableLabs, USA (j.fang@cablelabs.com)
Dr. Klaus Doppler, Nokia Research Center, USA (klaus.doppler@nokia.com)

TPC Chairs

Dr. Shiwen Mao, Auburn University, USA (smao@ieee.org)
Dr. Karthik Sundaresan, NEC Laboratories America, (karthiks@nec-labs.com)
Dr. Sayantan Choudhury, Nokia Research Center, USA (sayantan.choudhury@nokia.com)

Call for Papers

IEEE Workshop on Quality of Experience-based Management for Future Internet Applications & Services

Sponsored by the IEEE ComSoc Technical Committee on Multimedia Communications

http://qoe-fi.diee.unica.it

Organized in conjunction with
IEEE International Conference on Communications (IEEE ICC 2015)
8-12 June 2015, London, UK

Recent technological advances have enabled a constant proliferation of novel immersive and interactive services that pose ever-increasing demands to our communication networks and add to their load. Examples are: social TV, immersive environments, mobile gaming, HDTV over mobile, 3D virtual world, book/newspaper consumption, social networking, IPTV applications, just to cite a few. Some of these services have already reached a major market success especially because a *user-centered* approach has been followed to design the whole process of content production, service activation, content consumption, and service (and network) management.

In addition, we witness the trend of migrating end-to-end multimedia communication systems/platforms to the cloud. Media processing and consumption in the cloud requires attention from two main perspectives: maintenance of processing-related cloud operations over the execution time considering the end-user and application-related QoS/QoE requirements via dynamic resource provisioning; and the parallelization and abstraction of media processing tasks for the optimization of limited and shared cloud resources. Furthermore, the domain of Smart Cities offers new opportunities and use cases, but at the same time poses new challenges for keeping users engaged and interested in those services. This also includes other aspects such as quality of life as well as critical considerations such as user safety, particularly when it comes to urban transport and emergency scenarios.

In this dynamically evolving context, network operators and service providers are struggling to keep their increasingly sophisticated customers content while remaining profitable at the same time. Consequently, optimization and management of QoE has become crucial issue in the deployment of successful services and products. However, even if the concept itself may seem straightforward to understand, it requires rather complex implementation processes for efficient performances in real end-to-end systems/networks. The complexity of QoE is mainly due to the difficulties in its modeling, evaluation, and mapping to objective Quality of Service (QoS) parameters, which, for more than a decade, has been used as a partial substitution to QoE, and due to its multi-dimensional end-to-end nature covering a wide range of networks, applications, systems, devices, contexts and expertise.

On this background, the QoE-FI Workshop is aimed at bringing together researchers from academia and industry to identify and discuss technical challenges, exchange novel ideas, explore enabling technologies, and report latest research efforts that cover a variety of topics including, but not limited to:

- QoE evaluation methodologies and metrics
- Frameworks and testbeds for QoE evaluation (crowd-sourcing, field testing, etc.)
- QoE studies & trials in the context of Smart Cities
- QoE models, their applications and use cases
- QoE for immersive audio-video and interactive multimedia communication environments
- QoE-aware cross-layer design
- QoE-driven media processing and transmission over the cloud and over the top (OTT)
- QoE control, monitoring and management strategies
- QoE in community-focused interactive systems
- KPI and KQI definition for QoE optimization in different environments
- Integration of QoE in infrastructure and service quality monitoring solutions
- Media analytics from QoE Big Data
- Standards for media coding (HEVC, HEVC for 3D, etc.) and transport (DASH, MMT, XMPP, etc.)
- Future Media Internet architectures

General Chairs:
Ahmet Kondoz, Loughborough Univ. London, UK
Luigi Atzori, University of Cagliari, Italy
Raimund Schatz, FTW, Austria

Panel Chairs:
Weisi Lin, Nanyang Technological University, Singapore
Periklis Chatzimisios, Alexander TEI of Thessaloniki, Greece

Technical Program Chairs:
Markus Fiedler, BTH, Sweden
Pedro Assuncao, Institute of Telecommunications/IPL, Portugal
Tasos Dagiuklas, Hellenic Open University, Greece

Publicity Chairs:
Adlen Ksentini, University of Rennes, France
Nabil J. Sarhan, Wayne State University, USA
Evangelos Pallis, TEI of Crete, Greece

Important Dates
Submission deadline: **January 20, 2015** Author notification: **February 14, 2015** Camera ready due: **February 28, 2015**

MMTC OFFICERS (Term 2014—2016)

CHAIR

Yonggang Wen
Nanyang Technological University
Singapore

STEERING COMMITTEE CHAIR

Luigi Atzori
University of Cagliari
Italy

VICE CHAIRS

Khaled El-Maleh
Qualcomm
USA

Liang Zhou
Nanjing University of Posts and Telecommunications
China

Maria G. Martini
Kingston University,
UK

Shiwen Mao
Auburn University
USA

SECRETARY

Fen Hou
University of Macau, Macao
China

E-LETTER BOARD MEMBERS (Term 2014—2016)

Periklis Chatzimisios	Director	Alexander TEI of Thessaloniki	Greece
Guosen Yue	Co-Director	Broadcom	USA
Honggang Wang	Co-Director	UMass Dartmouth	USA
Tuncer Baykas	Editor	Medipol University	Turkey
Tasos Dagiuklas	Editor	Hellenic Open University	Greece
Chuan Heng Foh	Editor	University of Surrey	UK
Melike Erol-Kantarci	Editor	Clarkson University	USA
Adlen Ksentini	Editor	University of Rennes 1	France
Kejie Lu	Editor	University of Puerto Rico at Mayagüez	Puerto Rico
Muriel Medard	Editor	Massachusetts Institute of Technology	USA
Nathalie Mitton	Editor	Inria Lille-Nord Europe	France
Zhengang Pan	Editor	China Mobile	China
David Soldani	Editor	Huawei	Germany
Shaoen Wu	Editor	Ball State University	USA
Kan Zheng	Editor	Beijing University of Posts & Telecommunications	China