Recent Progress in Non-orthogonal Multiple Access

Zhiguo Ding
School of Computing and Communications Lancaster University
A collaboration with Princeton, SWJTU, FAU, USTC, BUPT, AUTH, CMCC, UT Dallas, XJTU, QMUL, KCL, Southampton, Tsinghua, Peking and Bell Labs.

24 September 2017
Outline

Overview and Motivation

A General Framework
   Single-Carrier NOMA
   Multi-carrier NOMA

NOMA Assisted Wireless Caching
   Motivation and Introduction
   NOMA Caching Strategies

Promising Research Challenges
Non-orthogonal Multiple Access (NOMA)

- **What is multiple access (MA)?**
  - Techniques to serve multiple users with limited bandwidth.
  - An example for downlink multiple access

![Diagram of NOMA](image)

- BS: multiplexing the signals
- Signals sent by the BS
- Receiver for the $k$-th user
- User 1, 2, ..., $k$-th, $K$-th user

Non-Orthogonal Multiple Access Overview & Motivation
Non-orthogonal Multiple Access (NOMA)

- What kind of multiple access techniques have been used?
  - We have been using orthogonal multiple access (OMA).
  - TDMA: Orthogonal (non-overlapping) time slots are allocated to users.
  - FDMA: Orthogonal (non-overlapping) frequency channels are allocated to users.
  - ...

[Diagrams showing TDMA and FDMA]
Non-orthogonal Multiple Access (NOMA) - (1/2)

Disadvantages of OMA

- Dilemma to realize a better trade-off between throughput and user fairness, illustrated in the following example:
  - A user with a poor connection to the base station (BS) is served by using OMA.
  - Spectral efficiency is low since this user cannot utilize the allocated bandwidth efficiently.
  - Since OMA is used, the bandwidth resources occupied by this user cannot be shared by the others.

- Difficult to support massive connectivity
  - Recall that the three key requirements for 5G are to support high throughput, low latency and massive connectivity
Non-orthogonal Multiple Access (NOMA) - (2/2)

- A promising solution is to break orthogonality → NOMA
  - The key idea of NOMA is to encourage spectrum sharing
  - Details for the advantages of NOMA are to be given in the remaining of this tutorial.

- NOMA is gaining ground on the competition of multiple access techniques for the next generation wireless networks
  - Adopted by many 5G MA concepts, including power-domain (PD) NOMA, sparse code multiple access (SCMA), multi-user sharing access (MUSA), pattern division multiple access (PDMA), lattice partition multiple access (LPMA), etc.
  - Used by 4G LTE-A, termed multi-user superposition transmission (MUST).
  - Included in the forthcoming digital TV standard (ATSC 3.0).
Outline

Overview and Motivation

A General Framework
  Single-Carrier NOMA
  Multi-carrier NOMA

NOMA Assisted Wireless Caching
  Motivation and Introduction
  NOMA Caching Strategies

Promising Research Challenges
Single-Carrier NOMA - Power Domain NOMA

- All the users are served at the same time, frequency and code, but with different power levels.
- Users with better channel conditions get less power.
- Successive interference cancellation (SIC) is used.

Conventional power-domain NOMA allocates more power to the user with poor channel conditions to ensure user fairness. But power domain NOMA cannot strictly guarantee the users’ QoS targets.

CR-NOAM can strictly guarantee the users’ QoS requirements by using the fact that NOMA can be viewed as a special case of CR networks.

For example, consider the following two-user scenario:
User 1 can be viewed as a primary user in a CR network:
- If OMA is used, the orthogonal bandwidth allocated to user 1 cannot be accessed by other users.
- Spectral efficiency is low since user 1 has a poor channel.

The use of NOMA is equivalent to the application of the cognitive radio concept:
- Specifically user 2, a user with better channel conditions, is admitted to the channel occupied by user 1.
- Although user 1 causes extra interference at user 2 and hence reduces user 2’s rate, the overall system throughput will be increased significantly since user 1 has a stronger connection to the BS.

A new form of NOMA, CR-NOMA, can be developed by using the synergy between NOMA and CR.

Outline

Overview and Motivation

A General Framework
  Single-Carrier NOMA
  Multi-carrier NOMA

NOMA Assisted Wireless Caching
  Motivation and Introduction
  NOMA Caching Strategies

Promising Research Challenges
Multi-carrier NOMA - Fundamental Limits

- Why to adopt hybrid NOMA
  - Asking all users to participate in NOMA can cause significant complexity
  - This motivates the hybrid NOMA scheme
    - Users in a cell are divided into small groups
    - OMA is used to avoid inter-group interference
    - NOMA is implemented among the users within a single group

- Who is to be grouped with whom?
  - User pairing has an important impact on the performance of NOMA.

Multi-carrier NOMA - Resource Allocation

- Resource allocation for hybrid NOMA is very difficult

\[
\text{max} \quad \text{(Weighted) sum rate} \\
\text{s.t.} \quad \text{Power and bandwidth constraints}
\]  

- How to solve such a mixed integer non-convex optimization problem
  - Monotonic optimisation is applied to solve such a non-convex optimisation problem to obtain an optimal solution
  - A low-complexity suboptimal solution based on SCO can be obtained with a performance gain close to the optimal.

Multi-carrier NOMA - 5G Variants

- Various practical forms of multi-carrier NOMA have been proposed for the 5G standard.
  - Multi-carrier NOMA achieves a favourable tradeoff between system performance and complexity.

- Both low density spreading (LDS) and sparse code multiple access (SCMA) are based on the idea that one user’s information is spread over multiple subcarriers.

- However, the number of subcarriers assigned to each user is smaller than the total number of subcarriers.
  - This is referred to as the low spreading (sparse) feature of these two versions of NOMA.
  - This feature ensures that the number of users utilizing the same subcarrier is not too large, such that the system complexity remains manageable.
SCMA: An Example of MC-NOMA (1/3)

- The key step of SCMA is how to map users to subcarriers.

\[
\begin{bmatrix}
1 & 0 & 0 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 1
\end{bmatrix}
= F
\]

Consider an SCMA system with 6 users and 4 subcarriers.

The key step to implement SCMA is to design the factor graph matrix, which specifies which user's encoded messages are allocated to which subcarriers.

A typical factor graph matrix for SCMA with 6 users and 4 subcarriers is the following:

\[
F = \begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 1
\end{bmatrix}
\]

where \([F]_{i,j} = 1\) means that the \(j\)-th user can use the \(i\)-th subcarrier, and \([F]_{i,j} = 0\) means that this user cannot use the subcarrier.
SCMA: An Example of MC-NOMA (3/3)

- The sparse feature of SCMA is reflected by the fact that there are only two non-zero entries in each column of $F$, i.e., each user employs only two subcarriers.

- Since one user can use multiple subcarriers, SCMA employs multi-dimensional coding in order to ensure that the user’s information is effectively spread over the subcarriers.

- Because one user’s messages at different subcarriers are jointly encoded, SCMA requires joint decoding at the receiver, where the message passing algorithm (MPA) is used to ensure low complexity.
  - Joint decoding is an important feature of SCMA, which distinguishes it from power-domain NOMA, where SIC is employed.
Pattern Division Multiple Access (1/2)

- PDMA is another type of multi-carrier NOMA, but the low density spreading (sparse) feature is no longer strictly present
  - The number of subcarriers occupied by one user is not necessarily much smaller than the total number of subcarriers.
- Similar to the factor graph matrix for SCMA, the performance of PDMA is largely determined by the design of the subcarrier allocation matrix, referred to as the PDMA pattern matrix.
- Consider a case with five users and three subcarriers, and PDMA pattern matrix

\[
Q = \begin{bmatrix}
1 & 1 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 & 0
\end{bmatrix},
\]

(2)

where the entries of this matrix indicate how the subcarriers are allocated to the users.

Pattern Division Multiple Access (2/2)

Consider a case with five users and three subcarriers, and PDMA pattern matrix

\[
Q = \begin{bmatrix}
1 & 1 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 & 0
\end{bmatrix}
\]

- User 1 is able to transmit or receive on all subcarriers.
- User 5 uses the 6-th subcarrier only.
- Therefore, different from LDS and SCMA, some users might be able to use all the subcarriers.
  - SCMA requires that each user occupies the same number of subcarriers.
  - This constraint is not required by PDMA and hence makes PDMA more flexible.
Outline

Overview and Motivation

A General Framework
Single-Carrier NOMA
Multi-carrier NOMA

NOMA Assisted Wireless Caching
Motivation and Introduction
NOMA Caching Strategies

Promising Research Challenges
Motivation and Introduction

- NOMA is shown to be compatible to many other important communication techniques, such as mmWave, relaying, MIMO, etc.
- Little is known about the coexistence between NOMA and wireless caching
  - How to realize content pushing in a timely and robust manner?
    - Content can be pushed to caching infrastructure during *off-peak hours* through *wired connections*.
  - How to cope with the non-ideal situation for wireless caching, when some users’ requests have to be fetched from the BS directly?
    - When this situation happens, the spectral efficiency of wireless caching is reduced.
    - An inevitable but frequently happening situation.

The NOMA principle is applied to the content pushing and content delivery phases, respectively.

Both the cache hit probability and content delivery probability are improved.
Push-then-delivery Strategy - Content Pushing

- An illustration of the impact of NOMA on content pushing, where $1 < m < t$.
  - CSs denotes content servers.
  - $CS_m$ is closer to the BS than $CS_t$.
  - In OMA, a single file is pushed to $CS_t$.
  - In NOMA, the BS pushes three additional files by using NOMA.
  - A content server closer to the BS is likely able to decode more pushed file.
Push-then-delivery Strategy - Content Delivery

- An illustration of the impact of NOMA on content delivery.
  - In OMA, each content server serves a single user.
  - By using NOMA, an additional user can be served.
Push-and-delivery Strategy

- When this situation happens, the spectral efficiency of wireless caching is reduced, since the users’ requests cannot be accommodated locally.
- The proposed push-and-delivery strategy treats this situation as an opportunity for the application of NOMA to improve the spectrum efficiency of wireless caching.
The use of NOMA can ensure new content files can be continuously pushed to content servers, in a more timely manner, compared to OMA.

This strategy can also been applied to D2D caching.
Promising Research Challenges

- Different variants of NOMA
- New coding and modulation for NOMA
- Hybrid multiple access
- NOMA assisted caching
- Coexistence between mmWave and NOMA
- MIMO and cooperative NOMA
- Interplay between NOMA and cognitive radio
- Imperfect CSI and limited channel feedback
- Security provisioning in NOMA
- Efficient resource allocation for NOMA
- Implementation issues of NOMA
- Emerging applications of NOMA
Thank you for your attention!


Email: z.ding@lancaster.ac.uk