International Journal of Hybrid Information Technology
Vol.8, No.11 (2015), pp.27-34
http://dx.doi.org/10.14257/ijhit.2015.8.11.03

Sphere Decoding in Parallel Mode and its Performance

Weiliang Fan, Zhijun Wang and Xinyu Mao

Peking University
xymao@pku.edu.cn

Abstract

Sphere decoding is a very powerful algorithm in searching the optimal solution of multiple input and multiple output systems. However, it cannot perform in parallel directly. Sphere decoding can be depicted as searching in a tree. In this paper, we propose a parallel mode of the sphere decoding algorithm. We proposed that the searching tree can be partitioned into several sub-trees. The searching is divided into two stages. In the first stage, the partial Euclidean distances of sub-tree root nodes are calculated. In the second stage, several sub-trees perform their searching simultaneously. The Euclidean distance of the early finished sub-tree helps to reduce the calculation in the later finished sub-trees search. Simulation results show that the search time can be reduced and the calculation fluctuation can be stabilized.

Keywords: *multiple input multiple output systems, parallel calculation, sphere decoding*

1. Introduction

Along with the boost of integrated circuit industry and computer technology, parallel calculation shows great prospect in many applications. Cloud computing, even known by persons who only know computer, is attracting attentions from industry to academic [1-3]. Many calculations are written to accommodate the parallel calculation. The calculation of multi-input multi-output (MIMO) is a NP hard problem. Many algorithms are developed to approach its optimal or suboptimal solutions. Among them, Sphere decoding is a very powerful algorithm to find the optimal solution in many scenarios [4-12]. Many works have been done to promote this algorithm [13-29]. To utilize the benefit of parallel calculation in signal process is a good method to promote the performance [30-31]. In this paper, we propose to utilize the SD in parallel calculation in contrast to most popular SDs, and then explore the performance of parallel SD.

2. Sphere Decoding

Let us consider a multi-input multi-output system with *Nt* inputs and *Nr* outputs like Figure 1. The data streams of *Nt* inputs are independent, and indicate as a vector $\mathbf{x} = (x_1 \quad x_2 \quad \dots \quad x_{Nt})^T$, where x_i is the draw from a complex QAM constellation, where the modulation order is *m*. The *Nt*×*Nr* channel gains of the system are independent identically distributed (i.i.d.), and denoted as *hij*. They form a matrix

ISSN: 1738-9968 IJHIT

$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1Nt} \\ h_{21} & h_{22} & \cdots & h_{2Nt} \\ \vdots & \vdots & \ddots & \vdots \\ h_{Nr1} & h_{Nr2} & \cdots & h_{NrNt} \end{pmatrix}$. The *Nr* outputs form a vector $\mathbf{y} = (y_1 \quad y_2 \quad \ldots \quad y_{Nr})^T$. The relationship between them can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} \tag{1}$$

where $\mathbf{v} = (v_1 \quad v_2 \quad \ldots \quad v_{Nr})^T$ is the additional Gaussian noise. They also satisfy i.i.d. When *Nt*<=*Nr*, it is easy to express the optimal solution of equation (1) as

$$\hat{\mathbf{x}}_{ML} = \arg\min_{\mathbf{x}\in\aleph} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \tag{2}$$

where $\aleph$ denotes the set of all possible inputs. Without losing the generality, we set *Nt*=*Nr*=*N/2* in the rest of this paper.

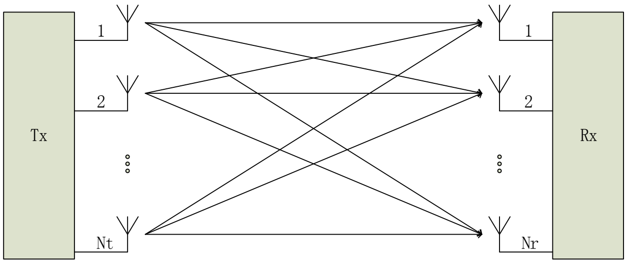


Figure 1. Structure of Multi-input Multi-output System

The complex system is often transferred into a real system for simplification purpose.

$$\begin{cases} \tilde{\mathbf{y}} = \left(real(\mathbf{y})^{\mathrm{T}} \quad image(\mathbf{y})^{\mathrm{T}} \right)^{\mathrm{T}} \\ \mathbf{s} = \left(real(\mathbf{x})^{\mathrm{T}} \quad image(\mathbf{x})^{\mathrm{T}} \right)^{\mathrm{T}} \\ \tilde{\mathbf{v}} = \left(real(\mathbf{v})^{\mathrm{T}} \quad image(\mathbf{v})^{\mathrm{T}} \right)^{\mathrm{T}} \\ \tilde{\mathbf{H}} = \begin{pmatrix} real(\mathbf{H}) & -image(\mathbf{H}) \\ image(\mathbf{H}) & real(\mathbf{H}) \end{pmatrix} \end{cases} \tag{3}$$

Then we have

$$\tilde{\mathbf{y}} = \tilde{\mathbf{H}}\mathbf{s} + \tilde{\mathbf{v}} \tag{4}$$

where all vectors and matrix are real.

Equation (4) is often disposed in its triangle mode. The channel matrix can be discomposed into two matrix $\tilde{\mathbf{H}} = \mathbf{Q}\mathbf{R}$, where $\mathbf{Q}$ is a unitary matrix, and $\mathbf{R}$ is a upper triangle matrix. Then equation (4) can be transferred into

$$\boldsymbol{\rho} = \mathbf{R}\mathbf{s} + \boldsymbol{\eta} \tag{5}$$

where $\boldsymbol{\rho} = \mathbf{Q}^H \tilde{\mathbf{y}}$, $\boldsymbol{\eta} = \mathbf{Q}^H \tilde{\mathbf{v}}$. Because $\mathbf{Q}$ is a unitary matrix, the statistical characters of $\boldsymbol{\rho} = \mathbf{Q}^H \tilde{\mathbf{y}}$ and $\boldsymbol{\eta} = \mathbf{Q}^H \tilde{\mathbf{v}}$ remain the same.

Equation (5) can be rewritten as

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_N \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ 0 & r_{22} & \cdots & r_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{NN} \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_N \end{pmatrix} + \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_N \end{pmatrix} \tag{6}$$

We can find in equation (6) the facts. In the last row, ρ_N is only determined by η_N. And in the next last row, ρ_{N-1} in only determined by η_{N-1} and η_N. ρ_i is determined by η_k with k>=i. In another point of view, we can calculate ρ_i iteratively. Based on the iterative calculation, the calculation of the optimal solution can be depicted as seek the best solution in a tree, as in Figure 2. Equation (1) means searching in a tree from the root node with Euclidean distance (ED) 0 to one leaf node with the smallest ED.

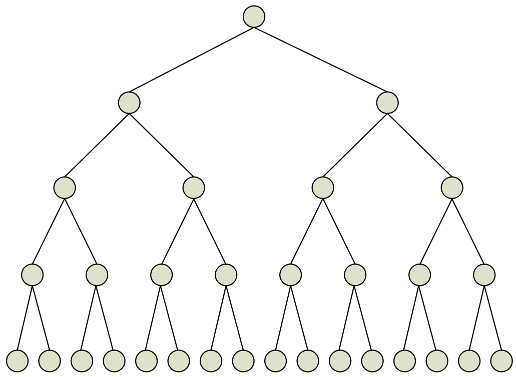


Figure 2. Tree Search Structure

We define the Euclidean distance of a possible node to the received signal as

$$\begin{aligned} D &= \sum_{j=1}^{N} \left(\rho_j - \sum_{k=j}^{N} r_{jk} \hat{s}_k \right)^2 \\ &= \left| \rho_N - r_{NN} \hat{s}_N \right|^2 + \left| \rho_{N-1} - \sum_{k=N-1}^{N} r_{N-1k} \hat{s}_k \right|^2 + \ldots + \left| \rho_1 - \sum_{k=1}^{N} r_{1k} \hat{s}_k \right|^2 \end{aligned} \tag{7}$$

And we also define the partial Euclidean distance (PED) as

$$D_i = \left| \rho_N - r_{NN} \hat{s}_N \right|^2 + \ldots + \left| \rho_i - \sum_{k=i}^{N} r_{ik} \hat{s}_k \right|^2 \tag{8}$$

The problem (1) is equal to the problem to find a node with the smallest ED.
The SD can be described as followers.

1. Calculate $\boldsymbol{\rho}$, $\boldsymbol{\eta}$ and **R** respectively. Set the root node as the focus node, set PED of the current focus node is 0, set cut radius C is infinity, set le=N.
2. If all the children nodes of the focus node have been visited, le=le+1, go to step 6.
3. Expand one unvisited children node, calculate its PED, le=le-1.
4. If the expanded children node PED is large than C, le=le+1, go to step 6. Else, go to step 5.
5. Substitute the focus node and its PED with the expanded children node and its PED. If the focus node is a leaf node, update the C with the PED of the focus node and save the node as the candidate solution, go to step 2.
6. If le<=N, go to step 2. Else end the search.

3. The Sphere Decoding in Parallel Mode

The SD is a very powerful algorithm in searching tree for the best solution. However, it has the drawback that its calculation complexity is not fixed. In other word, it fluctuates largely. This limited its application. From the description above, the SD can perform only in serial mode. It is also a serious drawback too. We propose that the searching tree can be divided into sub-trees, the whole searching can be partitioned into two stages, like Figure 3. We define nodes in the N-1^{th} or N-2^{th}level as the root nodes of sub-trees. All nodes expanded from one sub-tree node compose a sub-tree.

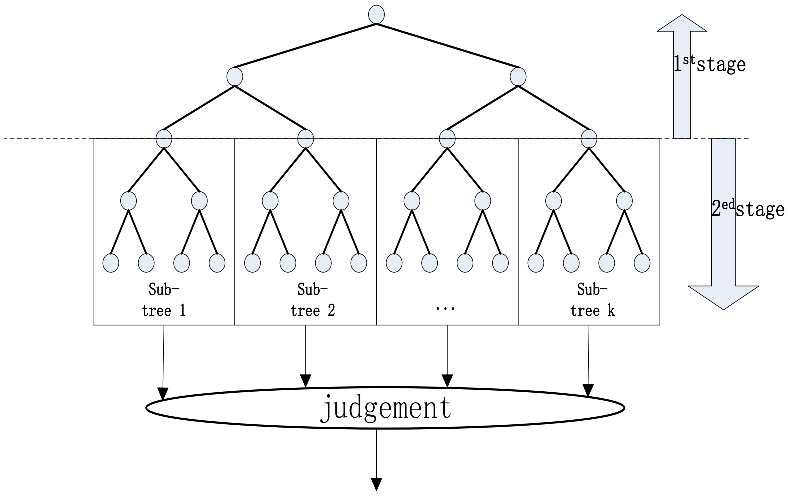


Figure 3. Sub-tree Search Structure

Sub-trees begin their searching from the sub-tree root nodes as the original whole tree search. The tree searches N-1^{th} or N-2^{th} level at the first stage. In rest of this paper, nodes of the N-2^{th} level are defined as the sub-tree root nodes for each sub-tree. The PEDs of the nodes of the level are served as the PEDs of sub-tree root nodes. Because of the complex-real transformation, the number of children nodes of the tree is equal to $2^{m/2}$, where m is the scale of modulation constellations. The number of the nodes of the N-2^{th} level, or say, the number of sub-trees is $2^{m/2}$.

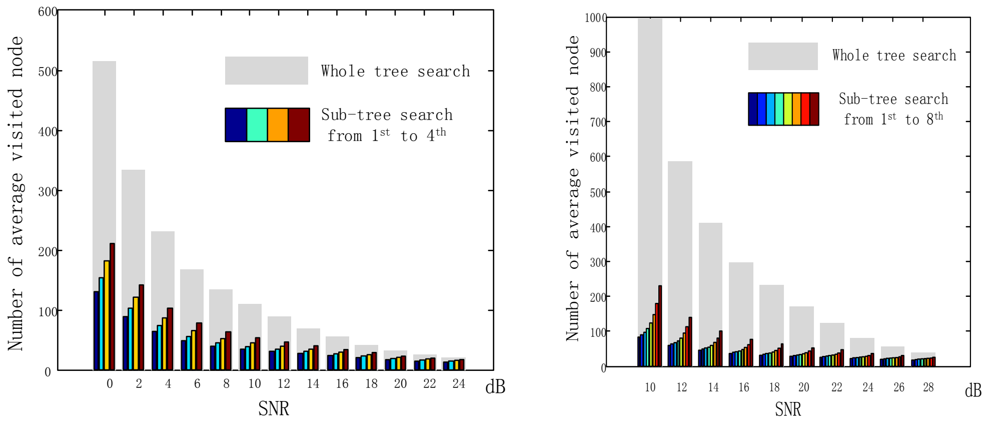
At the second stage, the search is divided into $2^{m/2}$ calculations, and the search is performed in sub-trees parallel. Once one sub-tree searching has been finished, the solution of it is the candidate solution of the whole tree, and the PED of it will update the initial cut radius of all other unfinished sub-trees to simplify the total calculations. As a result, the number of searching nodes will be reduced.

The size of the sub-tree is smaller than the whole tree, so the searching time is reduced consequently. Furthermore, in the second stage several sub-trees perform their searching simultaneously. The calculation of first finished sub-tree must be less than the average calculation of the tree with the same size. Although the later finished sub-tree may be more complex than the average tree, the cut radius offered by the early ending sub-tree can reduce the complexity largely. So that the other sub-trees search will also end earlier than the average tree with the same size as one sub-tree.

4. Simulation Results

In this section, we provide some simulation results according to Monte Carlo method. The channel matrix is uncorrelated Rayleigh flat-fading channel. The number of input and output is 4. The modulation order of input is 4 or 6. The channel matrix, all the input and output vectors are complex. They are transferred into their real counterparts at first before the dispose.

Figure 4 gives the number of average visited nodes of the traditional whole tree search and the proposed scheme. There are two sub-figures in Figure 4. In both figures, the gray bars means the number of the visited nodes of whole tree search, and the color bars means the different number of visited nodes in sub-tree search. The number of sub-tree search is ordered as ascending order. Both figures show that the number of the visited nodes of the sub-tree search is smaller than that of the whole tree search in corresponding scenario, which means that the disposing time of the sub-tree search is smaller than that of the whole tree search. They also show that the saved disposing time in lower SNR or lower modulation order is less than that in higher SNR or higher modulation order. For example, in Figure 4.a, 57.7% calculation spending time is saved in SNR=0, 9% calculation spending time is saved in SNR=24. While in the Figure 4.a, those numbers are 76.6% in SNR=10 and 33.4% in SNR=28.



a. N=4, m=4 b. N=4, m=6

Figure 4. Number of Average Visited Node

Figure 5 shows the distribution of visited nodes in SNR=0 and N=4, m=4. The ordinate is the normalized distribution density, the horizontal axis is the number of total visited nodes. The density curves concentrate together and cannot be distinguished. We enlarge the part of the peak of curves to observe the detail, which is shown in the center of Figure 5. From Figure 5, it can be found that the distribution density of the whole tree search is lower than that of the sub-tree search. The peak of the distribution of the whole tree is about 0.077, while those of the sub-trees are about 0.30, 0.27, 0.26, and 0.22.

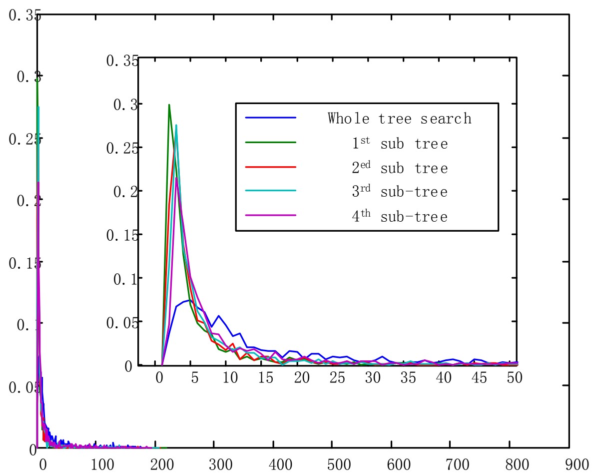


Figure 5. Distribution of Visited Node in SNR=0 and N=4, m=4

Figure 6 shows the accumulate distribution of visited node of Figure 5. It is clearer to show the difference of distribution in this Figure: the accumulate distribution of visited node of the sub-tree search accesses 100% sooner than that of the whole tree search. If we take 90% as a flag, the 1st sub-tree search reaches it in about 40, the 4th sub-tree search reaches it in about 65, while the whole tree search reaches it in about 144. Both Figure 5 and .6 show that the calculation of the proposed algorithm concentrates more than the traditional SD, or we can say that the fluctuation of the SD can be stabilized.

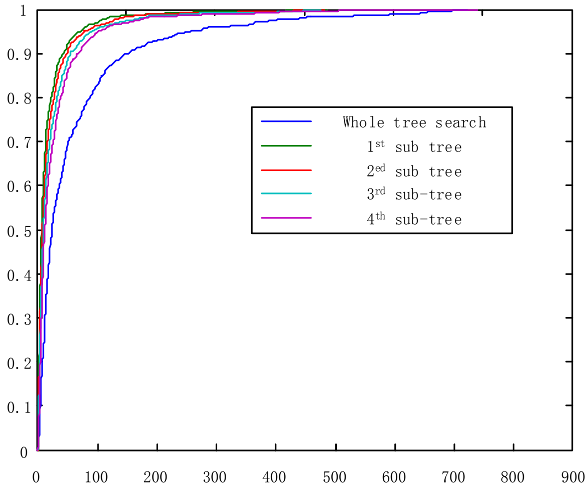


Figure 6. Accumulate Distribution of Visited Node of Figure 5

5. Conclusion

An algorithm to realize the parallel mode of SD is proposed and its performance is discussed and simulated. In the proposed algorithm, the traditional SD structure is divided into two stages, and parallel calculation can be operated in different sub-trees in the second stage. The radius obtained from the earlier finished sub-tree is used to simplify the later finished sub-trees' search. As a result, the total calculation is reduced. Simulation results show that the proposed algorithm can reduce the calculation time 57.7% or 9% in SNR=0 or 24 when modulation order is 4, or 76.6%, 33.4% in SNR=10 or 28 when modulation order is 6. It also show that the proposed algorithm stabilized the calculation fluctuation of the SD.

Acknowledgment

This work was supported by the China High-Tech 863 Program under Grants No. 2014AA01A704 and National Natural Science Foundation of China (NSFC) Program under Grants No. 61471405.

References

[1] B. T Ograph and Y. R. Morgens, "Cloud Computing", Communications of the ACM, vol. 51, no. 7, **(2008)**.
[2] P. Mell and T. Grance, "The NIST Definition of Cloud Computing", National Institute of Standards and Technology, vol. 53, no. 6, **(2009)**.
[3] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, *et al.*, "A View of Cloud Computing", Communications of the ACM, vol. 53, no. 4, **(2010)**.
[4] E. Agrell, T. Eriksson, A. Vardy and K. Zeger, "Closest Point Search in Lattices", IEEE Transactions on Information Theory, vol. 48, no. 8, **(2002)**.
[5] X. Chu and J. McAllister, "Software-Defined Sphere Decoding for FPGA-Based MIMO Detection", IEEE Transactions on Signal Processing, vol. 60, no. 11, **(2012)**.
[6] J. Jalden and P. Elia, "Sphere Decoding Complexity Exponent for Decoding Full-Rate Codes Over the Quasi-Static MIMO Channel", IEEE Transactions on Information Theory, vol. 58, no. 9, **(2012)**.

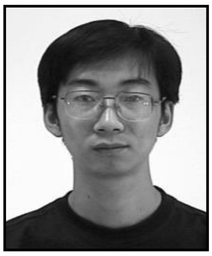
[7] L. Liu, J. Lofgren and P. Nilsson, "Area-Efficient Configurable High-Throughput Signal Detector Supporting Multiple MIMO Modes", IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 59, no. 9, **(2012)**.
[8] L. Liu, J. Lofgren, P. Nilsson and V. Owall, "VLSI Implementation of a Soft-Output Signal Detector for Multimode Adaptive Multiple-Input Multiple-Output Systems", IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 21, no. 12, **(2013)**.
[9] X. Chen, G. He and J. Ma, "VLSI Implementation of a High-Throughput Iterative Fixed-Complexity Sphere Decoder", IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 60, no. 5, **(2013)**.
[10] A. Younis, S. Sinanovic, M. Di Renzo, R. Mesleh and H. Haas, "Generalised Sphere Decoding for Spatial Modulation", IEEE Transactions on Communications, vol. 61, no. 7, **(2013)**.
[11] B. Y. Kong and I.-C. Park, "Efficient Tree-Traversal Strategy for Soft-Output MIMO Detection Based on Candidate-Set Reorganization", IEEE Communications Letters, vol. 17, no. 9, **(2013)**.
[12] S. Wu, L. Kuang, Z. Ni, J. Lu, D. D. Huang and Q. Guo, "Low-Complexity Iterative Detection for Large-Scale Multiuser MIMO-OFDM Systems Using Approximate Message Passing", IEEE Journal of Selected Topics in Signal Processing, vol. 8, no. 5, **(2014)**.
[13] Y. Zhang, "Effective reduction for sphere decoder in linear multi-input multi-output channel systems", IET Communications, vol. 6, no. 11, **(2012)**.
[14] K. Nikitopoulos and G. Ascheid, "Approximate MIMO Iterative Processing With Adjustable Complexity Requirements", IEEE Transactions on Vehicular Technology, vol. 61, no. 2, **(2012)**.
[15] R. Y. Chang, S.-J. Lin and W.-H. Chung, "Efficient Implementation of the MIMO Sphere Detector: Architecture and Complexity Analysis", IEEE Transactions on Vehicular Technology, vol. 61, no. 7, **(2012)**.
[16] R. H. Chen and W.-H. Chung, "Reduced Complexity MIMO Detection Scheme Using Statistical Search Space Reduction", IEEE Communications Letters, vol. 16, no. 3, **(2012)**.
[17] J. Ahn, H. N. Lee and K. Kim, "A Near-ML Decoding with Improved Complexity over Wider Ranges of SNR and System Dimension in MIMO Systems", IEEE Transactions on Wireless Communications, vol. 11, no. 1, **(2012)**.
[18] J. W. Choi, B. Shim and A. C. Singer, "Efficient Soft-Input Soft-Output Tree Detection via an Improved Path Metric", IEEE Transactions on Information Theory, vol. 58, no. 3, **(2012)**.
[19] R. H. Chen and W. H. Chung, "Reduced Complexity MIMO Detection Scheme Using Statistical Search Space Reduction", IEEE Communications Letters, vol. 16, no. 3, **(2012)**.
[20] L. P. Natarajan, K. P. Srinath and B. S. Rajan, "On the Sphere Decoding Complexity of High-Rate Multigroup Decodable STBCs in Asymmetric MIMO Systems", IEEE Transactions on Information Theory, vol. 59, no. 9, **(2013)**.
[21] A. Younis, S. Sinanovic, M. Di Renzo, R. Mesleh and H. Haas, "Generalised Sphere Decoding for Spatial Modulation", IEEE Transactions on Communications, vol. 61, no. 7, **(2013)**.
[22] Z. Yue, G. He, J. Ma and Z. Mao, "A memory reduced stack algorithm for MIMO detection", In IEEE International Conference on Signal Processing, Communication and Computing, **(2013)**.
[23] T. Cui, S. Han and C. Tellambura, "Probability-Distribution-Based Node Pruning for Sphere Decoding", IEEE Transactions on Vehicular Technology, vol. 62, no. 4, **(2013)**.
[24] S.-L. Shieh, R.-D. Chiu, S.-L. Feng and P.-N. Chen, "Low-Complexity Soft-Output Sphere Decoding with Modified Repeated Tree Search Strategy", IEEE Communications Letters, vol. 17, no. 1, **(2013)**.
[25] S. Songlin and S. Wang, "Two novel reordering methods for MIMO sphere detection based on MMSE detection", In IEEE Wireless Communications and Networking Conference, **(2013)**.
[26] Y. Wang and H. Leib, "Sphere Decoding for MIMO Systems with Newton Iterative Matrix Inversion", IEEE Communications Letters, vol. 17, no. 2, **(2013)**.
[27] S. Barik and H. Vikalo, "Sparsity-Aware Sphere Decoding: Algorithms and Complexity Analysis", IEEE Transactions on Signal Processing, vol. 62, no. 9, **(2014)**.
[28] M. Abbasi, A. A. Tadaion and S. Gazor, "Reduced complexity sphere decoding using a geometrical approach", In IEEE International Conference on Acoustics, Speech and Signal Processing, **(2014)**.
[29] T. H. Kim, "Low-Complexity Sorted QR Decomposition for MIMO Systems Based on Pairwise Column Symmetrization", IEEE Transactions on Wireless Communications, vol. 13, no. 3, **(2014)**.
[30] H.-W. Liang, W.-H. Chung, H. Zhang and S.-Y. Kuo, "A parallel processing algorithm for Schnorr-Euchner sphere decoder", IEEE Wireless Communications and Networking Conference, **(2012)**.
[31] S. Roger, C. Ramiro, A. Gonzalez, V. Almenar and A. M. Vidal, "Fully Parallel GPU Implementation of a Fixed-Complexity Soft-Output MIMO Detector", IEEE Transactions on Vehicular Technology, vol. 61, no. 8, **(2012)**.

Authors

Weiliang Fan, He is currently working toward the M.S. degree at Peking University, Beijing, China. His research interests include communication theory and signal processing techniques for multi-user MIMO wireless networks.

Zhijun Wang, He is professor and director of experimental teaching center in the School of Electronics Engineering and Computer Science at Peking University.

His research interests include circuit and systems, emboarded systems, and communications.



Xinyu Mao, He received the Doctor of Engineering in Electronics from Peking University. Now he is teacher in the School of Electronics Engineering and Computer Science at Peking University.

His research interests include wireless communications and communications signal processing, especially MIMO and OFDM.

Mr. Mao is a member in IEEE.