Design and Implementation of An Improved K-Best Detection Algorithm

Jie Zhu1,*, Bing Xi2 and Zhizhong Zhang3
1Chongqing University of Posts and Telecommunications, ChongQing, China
2Chongqing University of Posts and Telecommunications, ChongQing, China
3Chongqing University of Posts and Telecommunications, ChongQing, China
*Corresponding author e-mail: 279976891@qq.com

Abstract. In MIMO system, with high modulation order and large number of antennas, the complexity of the traditional K-Best detection algorithm is still high. For solving this problem, the theory of the inter-layer reservation in the algorithm is further extended, and propose a new child node expansion strategy which is based on Euclidean distance sorting. By reducing the access nodes, get a low-complexity improved algorithm. Considering the performance of the algorithm, a fast sorting QR decomposition based on SINR is introduced for the pre-detection preprocessing. The simulation proves that the improved algorithm performs well while reducing the complexity. Then design a parallel pipeline of the detector which is implemented by Xilinx XC7Z045 FPGA. The results show that this scheme performs well and has less hardware resource consumption. Now it has been applied to the LTE-A air interface monitoring instrument project.

1. Introduction
In MIMO-OFDM system, good receiver can guarantee the performance of the entire system. The signal detection is almost the most important part, so the optimization of algorithm and design of the detection module are very significant.

Among the MIMO signal detection algorithm, ML (Maximum Likelihood) detection is the best performance algorithm. As a sub-optimal detection algorithm, the spherical detection algorithm not only retains good detection performance but also greatly reduces complexity [1]. This algorithm can adopt the depth-first and breadth-first strategies for tree search. The performance of the depth-first strategy can approach the ML detection. However, it has difficulty in application due to the existence of the feedback path. The breadth-first strategy, which is the k-best algorithm, has some loss in performance. But it has constant complexity, which makes it easy to use the parallel pipeline structure for VLSI design. However, as the number of antennas and the modulation order increases, the number of searching child nodes increases, which leads to an increase in hardware complexity [2]. Therefore, in case of reducing the complexity, a new sub-node expansion strategy is adopted on the basis of the original algorithm. At the same time, for the pre-detection preprocessing, an improved fast-sequencing QR decomposition based on SINR is introduced. Under the premise of maintaining good detection performance, the purpose of reducing computational complexity is achieved. The proposed pipeline structure scheme in this paper can also be well applied to the MIMO detection in the spatial multiplexing scenario of the current LTE-A system.
2. Analysis of Algorithms

2.1. Traditional k-best algorithm

In a MIMO communication system with N transmit antennas and N receive antennas, if the channel is quasi-static flat fading, the mathematical model of the system can be expressed as:

\[ y = Hx + n \]  

(1)

H is the equivalent channel transmission matrix, x is the data column vector with N dimension, and y is the received data column vector, n represents the additive white Gaussian noise [3]. In order to avoid complex operation in actual processing, the system model is transformed into an equivalent real-value model by sacrificing dimensions, which is:

\[ y = Hx + n \]  

(2)

Solving the above formula by the maximum likelihood detection criterion is:

\[ \hat{x} = \arg\min_{x \in \Omega} ||y - Hx||^2 \]  

(3)

Ω represents the set of modulation constellation points. In the traditional k-best algorithm, H array is taken QR decomposition, and Equation 3 is:

\[ \hat{x} = \arg\min_{x \in \Omega} ||y' - R^T x||^2 \]  

(4)

Where Q is a 2N-dimensional orthogonal matrix, and R is a 2N-dimensional upper triangular matrix, \( y' \) is a dimensional vector.

Define the parameter k as the number of nodes reserved for each layer. The weighting function (PED) of a node on the i-th layer is:

\[ |y_i' - \sum_{j=1}^{2N} R_{ij} x_j |^2 \]  

(5)

The weighting function sum (PEDS) from the i-th layer to the 2N-layer 1 node is:

\[ \sum_{i=1}^{2N} |y_i' - \sum_{j=1}^{2N} R_{ji} x_j |^2 \]  

(6)

The detection starts from the 2Nth layer, first calculate the Euclidean distance increment of each layer, and add the accumulated PEDS of upper layer to get the PEDS of the current layer, then judge by ordering to retain the smaller k PEDS and corresponding k nodes, delete all information of the abandoned nodes [4]. When the last layer is detected, selecting the path with the smallest PEDS from the candidate vector as the detection result in the retained result set.

2.2. Improved k-best algorithm

The improvement is mainly for the pre-processing and the search detection. For the pre-processing before detection, in order to ensure the detection performance of the algorithm, it is considered to introduce an improved fast-sequencing QR decomposition based on SINR. The basic idea is to achieve an optimal ranking of the upper triangular matrix R in the process of calculating the QR decomposition. Namely:

The first step: Calculate the QR decomposition of the H array to obtain the R array.

The second step: obtaining the corresponding optimal detection order and the corresponding standard matrix P according to the upper triangular matrix R. Assume the transmit power of each layer of data stream is 1, and the SINR of each layer is:

\[ \text{SINR} = \frac{1}{\sum_{i=1}^{2N} |y_i' - \sum_{j=1}^{2N} R_{ji} x_j |^2} \]  

(7)

Use the QR substitutes for H, Equation 7 is:
\[
\text{SINR} = \frac{1}{\sigma^2[(QR)^*(QR)]} \\
= \frac{1}{\sigma^2[(R^*(R^*)^*)]} = \frac{1}{\sigma^2\|R^*\|_2},
\]

(8)

It can be seen from Equation 8 that the optimal detection order of SINR depends on the modulus of the diagonal elements. The inverse matrix of R is also an upper triangular matrix, so we can obtain the standard matrix P.

The third step: multiply R by the mark matrix P, the result is RP, then do the QR decomposition. Since it contains a large number of zero elements, the complexity is very low, the result of optimal detection QR decomposition is obtained, and the result of the final QR decomposition preprocessing is:
\[
H' = HP = QR = Q_2QR = QR
\]

(9)

For the traditional k-best detection algorithm, the analysis shows that the complexity is mainly concentrated on the Euclidean distance calculation and inter-layer ordering[5]. Based on the original algorithm, this paper proposes a new sub-node expansion strategy of the interlayer reservation points, which reduces the complexity of the algorithm.

For the solution of PED and sorting in the algorithm, the improved SE strategy is used. It means, when the child node expansion is determined, it is not necessary to calculate the Euclidean distance of all the extended child nodes, but only need to calculate the floating points of this layer which makes PED value equals to zero, then find the constellation point which is closest to the floating point to obtain the extended child nodes of the parent node[6]. The floating point is defined as:
\[
y = \frac{1}{\sum_{j=1}^{2N} R_{ij}x_j} \sum_{j=1}^{2N} R_{ij}x_j
\]

(10)

Considering that reserved nodes of the upper layer have been sorted, the above strategy can be further extended. The inter-layer ordering is arranged by the magnitude of the influence on the signal in descending order. For the part of upper layer’s nodes with small PED, it is possible that the PED value of their extending nodes is still small. It has been verified in the existing literature that if one layer’s parent node has sorted, the trend of the number of reserved expansion points should be generally downward, so we can adopt the sub-node expansion method with descending order, which is an arithmetic progression. And the loss of the performance can be basically ignored. The algorithm achieves the effect of reducing complexity by deleting some access sub-nodes, and then proposes an implementation structure with variable number of nodes per layer. However, for this algorithm the complexity of each layer is not fixed. In order to facilitate the hardware implementation and reduce the complexity, this paper introduces a new sub-node expansion strategy based on Euclidean distance sorting. Which is, in the existing sorted parent node of a certain layer, find the previous M nodes according to PED value, and only extend these M nodes. If M takes the limit value 1, which is the k-best algorithm only extend the FBC (best sub-node)[7], the impact of the value on performance is evaluated by simulation.

The improved k-best algorithm flow is as follows:
1. Realize the system model and preprocessing: optimal QR fast decomposition;
2. Initialize the root node, start detection from the 2N layer. In this layer, each node expands certain extended child nodes (depending on the modulation mode), sorts the extended child nodes in ascending order, and reserves the previous k nodes; only the previous M nodes are extended, during the extension still using improved SE strategy, each layer gets M node groups in ascending order, and each group has k child nodes;
3. Devide the above nodes into groups, selecting the previous k extended child nodes by PED value a subgroup, then pass it to the next layer;
4. Detecting each layer, repeating the above steps, and when it comes to the first layer, i=1, make the judgment and output the subgroup correspond to the smallest PEDs.

The whole process of the improved algorithm is shown in Figure1.
3. Simulation and performance analysis

According to the analysis of algorithm, as the number of access nodes decreases, the number of additions and multiplications executed in the algorithm will also decrease significantly, which will reduce the complexity of the algorithm. Since the k-best algorithm is based on the breadth-first search strategy, only k candidate points with the smallest weight are retained in each layer. Different detection performance can be obtained by adjusting the magnitude of the k value.

Based on the above improved algorithm, this paper carries out MATLAB performance simulation. The simulation environment is a 4×4 MIMO system, 16QAM modulation mode and the channel is quasi-static Rayleigh flat fading channel. When K is 4 and 8, the parameter M is taken different values. The performance comparison between the improved algorithm and the traditional k-best algorithm is shown in figure 2 and figure 3.

In the simulation process, the value of the parameter M is decremented from k-1 to 1. From Fig. 2 and Fig. 3, it can be seen that the detection performance is better when the value of k is larger, which is consistent with the traditional k-best algorithm. For the improved k-best algorithm, the performance difference is small when the SNR is low. However when the SNR is high, as the value of the parameter M decreases, the performance decreases to a certain extent but not evidently. Only when its value approaches the limit value of 1, the performance declines more obviously. The specific value for the implementation may depends on the communication system. For some FFT systems and wavelet systems, the improved k-best algorithm with M=1 can greatly reduce the resource consume of the detector.

The superiority of the improved algorithm is verified by the simulation, and the complexity is successfully reduced under the premise that the difference of detection performance is negligible.
As for the complexity, in original k-best algorithm, except for 2N layer, each layer needs to sort $Mc \times k$ child nodes and retains previous $K$ child nodes. $Mc$ depends on the modulation method. After using the improved SE strategy, in addition to the 2N layer, each layer needs to sort $k \times k$ child nodes and retains $k$ child nodes too.

The improved algorithm only needs to sort $k \times M$ ($M < k$) sub-node sorting for each layer except the 2N layer, and retains the former $M$ sub-nodes for expansion. The calculation mainly contains PEDS value calculation and sorting. At each layer, the count of addition and multiplication decreases. So the computational complexity can be reduced.

4. FPGA design and implementation of the improved algorithm

4.1. Whole structure of the design
According to the analysis of the algorithm, the design mainly includes three parts: preprocessing and pre-computation and search detection. Considering the data throughput rate, a pipeline parallel structure can be implemented.

The complexity of the traditional k-best algorithm is concentrated on the PED value calculation and the inter-layer ordering. The improved algorithm reduces the complexity for the above problems, and thus greatly reduces the hardware resources consumed when designing the detector.
The QR decomposition of the preprocessing module is implemented by Givens rotation combined with the Cordic algorithm. The Cordic algorithm has hardware-friendly features. The basic idea is to use a series of constant yaw to approximate the required rotation angle. Only use the addition, subtraction and shift operations during the implementation, which can balance the accuracy efficient and other aspects at the same time. The algorithm provides two working ways based on two different convergence criteria, which can achieve the diagonalization of real vectors. The pre-calculation part generates a R×X set before the PED value is calculated, R is the upper triangular matrix. This part can avoid repeated multiple calculations per layer. The core search detection section is common to each layer, requiring only one processing unit per layer, and each processing unit includes an Euclidean distance calculation unit (DCU) and a sorting unit (SSU). A feasible pipeline parallel structure is shown in the following figure:

![Figure 4: implementation block diagram of the structure](image)

### 4.2. Experimental Results

To verify the implementation complexity and performance of the improved algorithm, this paper carries out the hardware simulation. The integrated implementation and functional test are carried out by Vivado, and the XC7Z045-1FFG676C FPGA chip is used as the hardware platform for testing and verification. For the 4×4 16QAM MIMO system, while k=8,M=k-1, the resource consumption of the detection algorithm is shown in the following table:

| Resource name  | Resource Consumption |
|----------------|---------------------|
| Slice Registers | 31642 (554800)      |
| Slice LUTs      | 35646 (277400)      |
| Slice           | 12212 (69350)       |
| DSPs            | 120 (2020)          |

Compared with the traditional algorithm, the hardware resources consumed by the improved algorithm have a certain degree of decline. At the same time, compared with some existing detector designs, the algorithm maintains a high throughput rate with less resource consumption.

### 5. Conclusion

In this paper, analyze the k-best detection algorithm in MIMO system. The complexity of original algorithm is still high with high modulation order. For solving this problem, adopt a new child node expansion strategy based on Euclidean distance sorting, which can reduce the number of accessed node and reduce the computational complexity.

The improved algorithm is applicable to the signal detection module in MIMO-OFDM system. For the transmit diversity mode, the traditional MMSE detector can be applied. However for the space division multiplexing mode, the improved k-best algorithm performs better because of the higher...
requirements for detection performance. As for the design of the detector, a pipeline parallel structure is proposed for the improved algorithm. The simulation results show that the detector performs well and cause less resource consumption. The effectiveness and feasibility of the scheme are verified. Now the improved algorithm is used in the development of the LTE-A air interface monitoring analyzer. For further work, the preprocessing module can be optimized. Using two QR decompositions still has the problem of high complexity.

Acknowledgments
This work is supported by National Science and Technology Major Project(2015ZX3001013),Chongqing University Innovation Team Project (KJTD201312), Chongqing Science and Technology Commission-Chongqing Municipal Key Industry Common Key Technological Innovation Major Theme Special Project, Development and Application of 5G Terminal Simulator (cstc 2017 zdcy-zdzx0030).

References
[1] Dan Wu, Chao Shao. Improved sphere decoding detecting algorithm for MIMO, J.Transducer and Microsystem Technologies, 2016, 35(8): 123-126.
[2] X.J.Ma, L.Liu , F.Ye, MIMO Signal Detector Based on Configurable K-Best, J. Computer Engineering, 2009, 35(24): 236-238.
[3] J.J.Zhou , Research on channel capacity of MIMO system under Rayleigh correlation fading, J.Science & Technology Innovation Review, 2010(11): 4-4.
[4] L.Z.Jia , Z.X.Li, Improved QRD-M Detection Algorithm in Multiple Input Multiple Output System, J. Computer Engineering, 2011, 37(3): 77-79.
[5] Y.Zhu,Y.S Gui,Z.B.Wang, Research on Improvement of Downlink MIMO Detection Algorithm in LTE-A System, J.Journal of Nanjing University of Posts and Telecommunications, 2015, 35(3): 76-80.
[6] El-Banna A A A, El -Sabrouty M. CIR-based adaptive K-Best Sphere Decoder for DVB-T2,C.International Symposium on Modeling & Optimization in Mobile, Ad Hoc & Wireless Networks. IEEE, 2013:615-618.
[7] Ramasamy P, Ahmedkhan M B,Ramasamy M. Design and FPGA-Implentation of Minimum PED Based K-Best Algorithm in MIMO Decoder, J.Circuits and Systems,7,612-621.nalytical Modelling and Computer Simulation (AMS), IEEE, 2010,pp:40-45.