

Compressive Sensing Based Multiuser Detection via Iterative Reweighed Approach in M2M Communications

Xiaoxu Zhang, *Student Member, IEEE*, Fabrice Labeau, *Senior Member, IEEE*,
Ying-Chang Liang, *Fellow, IEEE* and Jun Fang, *Member, IEEE*

Abstract—Massive machine-to-machine (M2M) is an important application for IoT in 5G. In this letter, we focus on solving the multiuser detection (MUD) problem supported by Low-Activity Code Division Multiple Access (LA-CDMA) for M2M communications. To address the user activity factor unknown issue in the optimal maximum a posterior probability (MAP) and improve the signal reconstruction ability, we propose iterative reweighed (IR) and minimum mean-square-error iterative reweighed (MMSEIR) algorithms based on compressive sensing (CS) theory. The simulation results demonstrate that the proposed algorithms achieve substantial performance gain over traditional detectors.

Index Terms—5G, M2M Communications, Multiuser Detection, Compressive Sensing, Iterative Reweighed Algorithm.

I. INTRODUCTION

The next generation (5G) is a true sense of the integration network, it mainly contains two networks, i.e., the mobile Internet and the Internet of Thing (IoT) [1]. Among them, the new paradigm IoT has been receiving a lot of attention, which contains human-to-machine and machine-to-machine (M2M). With its low-cost, low power consumption and narrow bandwidth advantages, M2M is an inevitable trend to transform various vertical sectors such as smart home, car networking, disaster scenarios, etc.

Meanwhile, we notice that there are some technical issues to be solved, for instance, communications among machines require suitable multiple access method. Code Division Multiple Access (CDMA) system is flexible to support variable

This work is supported by Hydro-Quebec, the Natural Sciences and Engineering Research Council of Canada, and McGill University in the framework of the NSERC/Hydro-Quebec Industrial Research Chair in Interactive Information Infrastructure for the Power Grid (IRCPJ406021-14), the National Natural Science Foundation of China (Grant No. 61571100, Grant No. 61631005), the National Science Foundation of China (Grant 61522104) and China Scholarship Council (No. 201706070089).

X. Zhang is with the National Key Lab of Science and Technology on Communications, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, R. P. China, and also with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: xiaoxu.zhang@mail.mcgill.ca).

F. Labeau is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: fabrice.labeau@mcgill.ca).

Y.-C. Liang is with University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China, and also with the School of Electrical and Information Engineering, University of Sydney, NSW 2006, Australia, and (e-mail: liangyc@ieee.org).

J. Fang is with the National Key Lab of Science and Technology on Communications, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, R. P. China (e-mail: JunFang@uestc.edu.cn).

rate services and scalable number of users, which makes it very attractive for M2M communications. Note that the total number of users is typically large and just a small part of them (i.e., 10%) will transmit signal at any given time interval, in this letter, we exploit the Low-Activity Code Division Multiple Access (LA-CDMA) for M2M communications.

Accurate multiuser detection (MUD) in M2M communications help us to address the challenges such as reducing the cost of machine terminals, resource allocation and low-cost M2M terminal coverage. Indeed, there are some studies to solve the MUD problems recently. In [2], Zhu and Giannakis proposed Ridge detector (RD) and Lasso detector (LD) to approach the optimal maximum a posterior probability (MAP) for data detection. However, the two detectors require to know the value of user activity factor p_a (the probability of active user), which is unknown in the practical MUD scenarios, because any user enters/leaves the communication systems randomly. In our previous work [3], we proposed Bayesian inference algorithms for data detection but the computational complexity is relatively large.

The objective of this letter focuses on the MUD problem in the uplink of LA-CDMA based on compressive sensing (CS). Encouraged by the observation of the total number of users is very large and the rate of active users is typically low, we consider the transmitted signal as *sparse* vector (many elements in the signal vector are zero). The recovery of sparse signal in CS has been received a great deal of attention recently from both academic and industry [4], [5]. Note that the transmitted signal in M2M communications has sparse feature, we are able to address the MUD problem via the theory of CS. For the special non-sparse signal situation, we first convert the system into sparse format, then do estimation.

The main contributions of this letter are summarized as follows: 1) We develop two computationally efficient algorithms for recovering the transmitted signals, respectively, iterative reweighed (IR) and minimum mean-square-error iterative reweighed (MMSEIR) algorithms; 2) For sparse transmitted signal, we first reformulate the traditional input-output system model under CS framework, and replace 0-norm with log-sum function. Then we propose the IR algorithm based on majorization-minimization (MM) strategy and transform the non-convex MUD optimization problem into an iterative problem [6], [7]; 3) For non-sparse situation, we first convert the system into sparse framework, then embed the minimum mean-square-error (MMSE) into IR, and propose MMSEIR

algorithm [8]; 4) The two schemes overcome the user activity factor p_a unknown difficulty in data detection. The efficacy is largely enhanced and the computational complexity is substantially reduced compared with the traditional MUD algorithms.

II. SYSTEM MODEL

We consider the LA-CDMA uplink system where there are K users, the spreading factor is N . In this system, the relationship between the transmitted signal \mathbf{b} and the received signal \mathbf{y} can be formulated as follows

$$\mathbf{y} = \mathbf{H}\mathbf{b} + \mathbf{w}, \quad (1)$$

which are expressed by

$$\begin{aligned} \mathbf{y} &= [y_1, y_2, \dots, y_N]^T, \\ \mathbf{b} &= [b_1, b_2, \dots, b_K]^T, \\ \mathbf{w} &= [w_1, w_2, \dots, w_N]^T, \\ \mathbf{H} &= [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K], \end{aligned} \quad (2)$$

where $\mathbf{y} \in \mathbb{C}^{N \times 1}$ is the received signal vector, $\mathbf{H} \in \mathbb{C}^{N \times K}$ is a complex matrix containing the spreading process and the channel information, $\mathbf{w} \in \mathbb{C}^{N \times 1}$ is the complex Gaussian noise ($\mathbf{w} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$). At the transmitter, $\mathbf{b} \in R^{K \times 1} \in (\mathcal{B} \cup \{0\})$ is the transmitted signal vector for the K devices whose elements are selected from a finite alphabet set. Let b_i be the i th signal to be transmitted by user i ($i = 1, 2, \dots, K$). When the device is active, $b_i \in \mathcal{B}$, where \mathcal{B} stands for the modulation alphabet such as Binary Phase Shift Keying (BPSK), i.e., $\mathcal{B} = \{-1, +1\}$. When the device is inactive, $b_i \in \{0\}$. Note that in such transmission, if most devices are inactive, namely, most of the elements in \mathbf{b} are zero, then the transmitted signal vector to be sparse; otherwise, the signal is non-sparse. Let p_a be the user activity factor, which defines the probability of active user. The main objective of this letter is to detect the transmitted signal \mathbf{b} using the given channel matrix \mathbf{H} and the received signal \mathbf{y} .

Indeed, there are some traditional MUD methods to solve the MUD problems recently. Linear receivers with low complexity for zero-forcing (ZF) and MMSE detectors and nonlinear detectors based on MAP for RD and LD are listed as follows [2], [3]:

$$\tilde{\mathbf{b}} = \begin{cases} (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{y}, & \text{for ZF,} \\ (\mathbf{H}^H \mathbf{H} + \sigma^2 \mathbf{I})^{-1} \mathbf{H}^H \mathbf{y}, & \text{for MMSE,} \\ (\mathbf{H}^H \mathbf{H} + 2\lambda \mathbf{I})^{-1} \mathbf{H}^H \mathbf{y}, & \text{for RD,} \\ \arg \min_{\mathbf{b}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{b}\|_2^2 + \lambda \|\mathbf{b}\|_1 \right\}, & \text{for LD,} \end{cases} \quad (3)$$

in which $\lambda \triangleq \log\left(\frac{1-p_a}{p_a}\right)$.

In general, the nonlinear detectors have better recovery ability than the linear receivers, however, they need to know the value of user activity factor p_a , which is hard to be obtained in a practical system, because any user enters/leaves the communication systems randomly. In addition, the value of user activity factor p_a for RD and LD should be very small [2]. In the following, we propose the IR and MMSEIR algorithms based on CS theory to address the challenges.

III. IR FOR MUD

In this section, we exploit the CS based scheme to solve the MUD problem for LA-CDMA uplink in massive M2M communications. For the traditional system model (1), \mathbf{H} is assumed to be an $N \times K$ dictionary matrix, \mathbf{b} is the recover vector, and \mathbf{y} is the measurement vector. In massive M2M communications, the number of users is large, and the rate of active users is typically low, which makes the transmitted signal vector has sparse feature. The sparsity can be defined as the total number of active devices, i.e., $S = \lceil p_a \times K \rceil$, where $\lceil a \rceil$ is a smallest integer greater than or equal to a , then signal vector \mathbf{b} is treated as S-sparse. For a better CS estimate of the transmitted signal \mathbf{b} , the form of our MUD system model based on CS can be written as

$$\begin{aligned} \min_{\mathbf{b}} \|\mathbf{b}\|_0 \\ \text{s.t. } \|\mathbf{y} - \mathbf{H}\mathbf{b}\|_2 \leq \varsigma, \end{aligned} \quad (4)$$

where the objective function $\|\mathbf{b}\|_0$ denotes the number of active devices in \mathbf{b} , and ς is a specified error tolerance parameter and related to the noise. In this way, the MUD problem is reduced to the recovery of sparse signal in CS. Problem in (4) is a non-convex and NP-hard problem. Conventionally, we can relax the non-convex 0-norm to the convex 1-norm MUD problem, and obtain a sub-optimal solution. Thus we can associate the relaxed 1-norm MUD problem with traditional CS recovery method such as Basis Pursuit (BP) technology [9]. However, it has been proved that the log-sum function can be used to promote sparsity, and has better sparse recovery ability than 1-norm [10]. In order to have computationally efficient solution, we replace 0-norm by log-sum function to detect the transmitted signal \mathbf{b} in M2M communications

$$\begin{aligned} \min_{\mathbf{b}} \left(\sum_{i=1}^K \log(|b_i|^2 + \varepsilon) \right) \\ \text{s.t. } \|\mathbf{y} - \mathbf{H}\mathbf{b}\|_2 \leq \varsigma, \end{aligned} \quad (5)$$

in which $\varepsilon > 0$ is small regularization constant, and is used to make sure that the log-function does not go to negative infinity. Based on the theory of convex optimization, we combine the objective function and constraint condition, the optimization problem (5) can be formulated as a Lagrange function

$$\min_{\mathbf{b}} L(\mathbf{b}) \triangleq \sum_{i=1}^K \log(|b_i|^2 + \varepsilon) + \eta \|\mathbf{y} - \mathbf{H}\mathbf{b}\|_2^2, \quad (6)$$

in which η is a lagrangian multiplier of $\sum_{i=1}^K \log(|b_i|^2 + \varepsilon)$ and $\|\mathbf{y} - \mathbf{H}\mathbf{b}\|_2^2$. Since then, the subject has been reduced to minimize the Lagrange function with respect to the transmitted signal \mathbf{b} in M2M communications.

In the following, we exploit IR algorithm based on (5) and (6) to address the MUD problem. The introduced IR algorithm is based on MM strategy to recover sparse transmitted signal \mathbf{b} . The MM algorithm helps to minimize a relative simple surrogate function majorizing the given aim function iteratively.

We first assume that a surrogate function $Q(\mathbf{b}|\tilde{\mathbf{b}}^{(t)})$ is the upper bound for the objective log-sum function, i.e., $\sum_{i=1}^K \log(|b_i|^2 + \varepsilon) \leq Q(\mathbf{b}|\tilde{\mathbf{b}}^{(t)})$. It is easy to verify that an

appropriate choice of such surrogate function for log-sum function $\sum_{i=1}^K \log(|b_i|^2 + \varepsilon)$ is given by

$$Q(\mathbf{b}|\tilde{\mathbf{b}}^{(t)}) \triangleq \sum_{i=1}^K \left(\log \left(|\tilde{b}_i^{(t)}|^2 + \varepsilon \right) + \frac{|b_i|^2 + \varepsilon}{|\tilde{b}_i^{(t)}|^2 + \varepsilon} - 1 \right), \quad (7)$$

in which $\tilde{b}_i^{(t)}$ is the estimated signal b_i at t th iteration.

Proof: To prove the upper bound inequation, which amounts to

$$\sum_{i=1}^K \left(\log \frac{|\tilde{b}_i^{(t)}|^2 + \varepsilon}{|b_i|^2 + \varepsilon} + \frac{|b_i|^2 + \varepsilon}{|\tilde{b}_i^{(t)}|^2 + \varepsilon} - 1 \right) \geq 0, \quad (8)$$

we set $x = \frac{|b_i|^2 + \varepsilon}{|\tilde{b}_i^{(t)}|^2 + \varepsilon}$, where $x > 0$. Then the left-hand side of (8) $f(x) = \log x + 1/x - 1$ and $f'(x) = (1 - 1/x)/x$. When $x > 1$, $f'(x) > 0$, then $f(x) > f(1)$. When $0 < x < 1$, $f'(x) < 0$, then $f(1) < f(x)$. So when $x = 1$, $f(x)$ has the minimum value. As $f(1) = 0$, we can conclude $f(x) \geq 0$. Thus the upper bound inequation is proven.

Note that the above surrogate function $Q(\mathbf{b}|\tilde{\mathbf{b}}^{(t)})$ is convex and differentiable. Consequently, the optimization problem can be formulated as follows

$$\begin{aligned} & \min_{\mathbf{b}} (Q(\mathbf{b}|\tilde{\mathbf{b}}^{(t)}) + \eta \|\mathbf{y} - \mathbf{H}\mathbf{b}\|_2^2) \\ & = \min_{\mathbf{b}} \left(\sum_{i=1}^K (\log(|\tilde{b}_i^{(t)}|^2 + \varepsilon) + \frac{|b_i|^2 + \varepsilon}{|\tilde{b}_i^{(t)}|^2 + \varepsilon} - 1) \right. \\ & \quad \left. + \eta \|\mathbf{y} - \mathbf{H}\mathbf{b}\|_2^2 \right). \end{aligned} \quad (9)$$

Take the first derivation of (9) over the transmitted signal \mathbf{b} , we can easily obtain the IR solution

$$\mathbf{b}^{IR} = (\mathbf{H}^H \mathbf{H} + \eta^{-1} \mathbf{D}^{(t)})^{-1} \mathbf{H}^H \mathbf{y}, \quad (10)$$

where $\mathbf{D}^{(t)}$ is a diagonal matrix defined as $\mathbf{D}^{(t)} \triangleq \text{diag} \left(\left(\left(|\tilde{b}_1^{(t)}|^2 + \varepsilon \right)^{-1}, \dots, \left(|\tilde{b}_K^{(t)}|^2 + \varepsilon \right)^{-1} \right) \right)$.

IV. MMSEIR FOR MUD

As mentioned above, the advantage of IR is that we can detect the optimal \mathbf{b} without the activity factor information. However, the IR method only makes use of the sparsity property of the transmitted signal \mathbf{b} , thus making the recovery result accurate only if the user activity is very low. In this section, we try to recover the special non-sparse signal, which is applicable for the large user activity factor cases in M2M communications. The proposed detector is called MMSEIR and derived in two steps: 1) transfer the non-sparse system model into sparse error format with the help of MMSE detector; 2) apply the IR algorithm to estimate the error vector. Here are the details.

Referring to (1), the transmitted signal \mathbf{b} is non-sparse. In order to transform the non-sparse system (1) into a sparse error format, here we remove the observed signal $\mathbf{H}\hat{\mathbf{b}}$ from

the received signal \mathbf{y} , and obtain a new sparse error system, that is,

$$\mathbf{y}^{new} = \mathbf{y} - \mathbf{H}\hat{\mathbf{b}} = \mathbf{H}(\mathbf{b} - \hat{\mathbf{b}}) + \mathbf{w} = \mathbf{H}\mathbf{e} + \mathbf{w}, \quad (11)$$

in which \mathbf{y}^{new} is the newly built received signal, the error vector \mathbf{e} is defined as $\mathbf{e} = \mathbf{b} - \hat{\mathbf{b}}$. For the signal $\hat{\mathbf{b}}$, it obtained with the help of the estimated signal $\tilde{\mathbf{b}}$ in Section II. To ease the notations, here we take the conventional linear receiver MMSE as an example

$$\hat{\mathbf{b}} = S(\tilde{\mathbf{b}}^{MMSE}) = S((\mathbf{H}^H \mathbf{H} + \sigma^2)^{-1} \mathbf{H}^H \mathbf{y}), \quad (12)$$

in which $S(\cdot)$ is a slicing function which maps the input $\tilde{\mathbf{b}}^{MMSE}$ to a finite set $\{-1, 0, +1\}$. Obviously, if the conventional method MMSE leads to perfect hard decisions, all elements in error vector \mathbf{e} will be zero. In practice, however, the elements might not be all zero. It is natural to consider the error vector \mathbf{e} as a sparse one. Additionally, it is worthy to mention that the linear MMSE tries to maximize the Signal to Interference-plus-Noise Ratio, it is a low complexity detector and easy to implement, but the performance is not as efficient as nonlinear detectors.

As we discussed above, the newly built (11) is a sparse system model. Now we consider to recover the error vector \mathbf{e} in (11), and the form of the sparse error system model based on CS can be reformulated as follows

$$\begin{aligned} & \min_{\mathbf{e}} \|\mathbf{e}\|_0 \\ & \text{s.t. } \mathbf{y}^{new} = \mathbf{H}\mathbf{e} + \mathbf{w}. \end{aligned} \quad (13)$$

The following problem reduces to estimate the error vector \mathbf{e} from the new received signal \mathbf{y}^{new} by the IR approach proposed in Section III, we get the IR solution of the error vector

$$\mathbf{e}^{IR} = (\mathbf{H}^H \mathbf{H} + \eta^{-1} \mathbf{E}^{(t)})^{-1} \mathbf{H}^H \mathbf{y}^{new}, \quad (14)$$

where $\mathbf{E}^{(t)}$ is a diagonal matrix defined as $\mathbf{E}^{(t)} \triangleq \text{diag} \left(\left(\left(|\tilde{e}_1^{(t)}|^2 + \varepsilon \right)^{-1}, \dots, \left(|\tilde{e}_K^{(t)}|^2 + \varepsilon \right)^{-1} \right) \right)$.

Accordingly, we obtain a more precise transmitted signal vector based on (12) and (14)

$$\mathbf{b}^{MMSEIR} = \mathbf{b}^{MMSE} + \mathbf{e}^{IR}. \quad (15)$$

V. SIMULATION RESULTS

We consider the uplink of CDMA system with $K = 20$ users. We exploit random sequences with length $N = 32$ for spreading. We consider additive white Gaussian noise (AWGN) channels and each user has the same received power when the user is active. The regularization parameter ε is 1, the Lagrangian multiplier η is 1 and the number of iterations rpt is 1000. The metric to evaluate the reconstruction ability is the symbol error rate (SER), which is defined as

$$\text{SER} \triangleq \frac{\text{ERR}}{\text{rpt} \times K}, \quad (16)$$

in which rpt stands for the number of iterations of each Monte Carlo simulation. K is the number of devices, which is used at each Monte Carlo round. ERR denotes the number of devices

