This is a repository copy of *Network Diversity Multiple Access in Rayleigh Fading Correlated Channels with Imperfect Channel and Collision Multiplicity Estimation*.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/108889/

Version: Accepted Version

**Proceedings Paper:**

https://doi.org/10.1109/TELFOR.2016.7818814

© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

**Reuse**
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Network Diversity Multiple Access in Rayleigh Fading Correlated Channels with Imperfect Channel and Collision Multiplicity Estimation

Ramiro Robles, Eduardo Tovar, Mauricio Lara, Aldo Orozco, Desmond C. McLernon, and Mounir Ghogho

Abstract—Network diversity multiple access or NDMA is the family of algorithms with the highest potential throughput in the literature of signal-processing-assisted random access. NDMA uses the concept of protocol-induced retransmissions to create an adaptive source of physical (PHY) layer diversity. This adaptive diversity is used to resolve packet collisions (via signal separation) without the explicit need (or as a complement) of a multiple antenna receiver. This paper proposes a further improvement on the modelling of NDMA by considering the effects of imperfect channel and collision multiplicity estimation. In addition, this work considers channel correlation between consecutive retransmissions (i.e., temporal correlation). Conventionally, the analysis of NDMA assumes that any error in the collision multiplicity estimation translates into the loss of all contending packets. This is an optimistic assumption because even when the multiplicity has been correctly estimated, errors can still occur. On the other hand, it is also pessimistic because correct reception can also occur when the multiplicity has been incorrectly estimated. This paper presents a more detailed study of the performance of NDMA considering these more specific detection/reception cases.

I. INTRODUCTION

Cross-layer design is an important tool in future random access networks. Correct reception now depends on physical (PHY)-layer performance, as well as traffic load conditions [1]. A breakthrough in this topic was the work in [2], where collisions were resolved using a new type of diversity based on retransmissions. The algorithm was called network diversity multiple access (NDMA). In NDMA, retransmissions are used to create a virtual MIMO (multiple-input multiple-output) system from which colliding signals can be recovered via source separation. Signals with collisions that cannot be resolved immediately are not discarded as in conventional protocols. They are initially used to estimate the collision multiplicity. Based on this information, the base station (BS) requests further retransmissions from the contending terminals in an attempt to create a full-rank MIMO system. The BS uses the stored signals to resolve the collision via source separation. A cooperative NDMA protocol was later proposed in [3]. NDMA with multi-packet reception was proposed in [4] with a finite user model. Stability analysis of NDMA with perfect collision multiplicity estimation and packet reception can be found in [5]. A Markov model for NDMA stability analysis was presented in [6].

In NDMA, the collision multiplicity estimation is used to determine the number of retransmissions that are necessary to resolve the collision. Too many retransmissions translates into a waste of resources and throughput degradation. Too few retransmissions means that full-rank conditions will be probably lost rendering the incorrect decoding of signals. Conventional modelling of NDMA is based upon the assumption that any error in the collision multiplicity estimation means the loss of all the contending packets [2]. However, this assumption is both optimistic and pessimistic at the same time. It is optimistic because even in the case of correct estimation of the collision multiplicity, packet decoding errors can still occur. It is also pessimistic because some packets can still be correctly decoded in case of incorrect estimation of the collision multiplicity. In addition, the protocol has only been analysed under the assumption of uncorrelated retransmissions and perfect channel estimation. This paper addresses these issues by reformulating all protocol expressions based on a more accurate model with all the potential cases of correct or incorrect packet reception, and under the assumption of incorrect estimation of the collision multiplicity. The model also includes the effects of channel estimation errors, as well as the effects of correlation between consecutive retransmissions (temporal correlation).

The remainder of this paper is organised as follows. Section II describes the system scenario. Section III deals with the signal model. Section IV presents the performance analysis of the protocol. Section V presents results of the performance of the protocol, and finally Section VI presents the conclusions.

II. SYSTEM MODEL

A. System scenario and channel model

Consider the slotted random access network with retransmission diversity depicted in Fig. 1 with a set of $J$ buffered one-antenna terminals and one central node or base station (BS) with one receiver antenna. The channel between terminal $j$ and the BS in time-slot $n$ is denoted by $h_j(n)$. All channel envelopes are assumed to be non-dispersive with Rayleigh...
statistics: $h_j(n) \sim C\mathcal{N}(0, \gamma)$. Signals experience identical correlation across (re)transmissions (i.e., temporal correlation). This means that $E[h_j^*(n)h_k(n)] = \rho_{j,k} \gamma$, where $\rho_{j,k}$ is the temporal correlation coefficient, $(\cdot)^*$ is the complex conjugate operator, and $E[\cdot]$ is the statistical average operator. For simplicity in analysis, time-slot notation $n$ in all variables will be dropped in subsequent derivations.

In NDMA, having more diversity sources than contending terminals is needed to potentially resolve the collision. Note that in this first epoch the set of detected terminals is identical to the set of contending terminals ($\mathcal{T}_d = \mathcal{T}$), which means no presence detection errors occurred. In this case, the number of collected signals is equal to two, which is enough to attempt the recovery of the two contending signals. Also note that the binary feedback is only set to $\xi = 1$ at the end of the first time-slot. Once the first retransmission has been received, the value is set to $\xi = 0$, which means that the current epoch has finalized so the contending terminals stop retransmitting while a new epoch-slot starts.

The second epoch ($e = 2$) experiences three contending terminals given by $\mathcal{T} = \{3, 8, 9\}$, which ideally requires two retransmissions plus the original transmission to be resolved. However, only 2 terminals given by $\mathcal{T} = \{3, 8\}$ were detected as active (terminal $j = 9$ has not been correctly detected). Therefore, the system only requests one more retransmission instead of two. This leads to a rank-deficient MIMO system, which in turn can lead to excessive decoding errors. Only a subset of the contending signals have been actually correctly decoded: $\mathcal{T}_d = \{3\}$. The next epoch ($e = 3$) experiences $K = 2$ contending terminals given by $\mathcal{T} = \{1, 8\}$, but the BS has detected $\tilde{K} = 3$ terminals given by $\mathcal{T} = \{1, 3, 8\}$, thus falsely considering terminal $j = 3$ as active. Note that detecting one more terminal has caused the BS to request one more retransmission than actually needed. This is a waste of resources. However, it can be observed that in this case all the signals were correctly decoded by the BS, even with an error of collision multiplicity estimation. The last epoch ($e = 4$) shows the case where all terminals were correctly detected, but the detection process was still incorrect for one of the contending terminals. These examples of epoch realizations aim to illustrate the variety of cases of correct/incorrect detection and decoding that might arise in NDMA.

### III. Signal Models

#### A. Signal model for terminal detection

Each terminal is pre-assigned with a unique orthogonal code that is used for purposes of terminal activity detection and channel estimation [2]. The BS uses a matched-filter operation to extract the detection indicator of terminal $j$, denoted here by

![Fig. 1. Random access network assisted by retransmission diversity.](image-url)
z_j. This indicator is then compared to a detection threshold \( \beta \) to decide whether terminal \( j \) is active or not. If \( z_j < \beta \) then the terminal is detected as inactive: \( j \notin \mathcal{T} \). Otherwise, if \( z_j \geq \beta \) then the terminal is detected as active (\( j \in \mathcal{T} \)). This means that \( \mathcal{T} = \{ j : z_j \geq \beta \} \) is the set of all terminals whose detection indicator exceeds the threshold \( \beta \). Since this detection process is prone to errors due to fading and noise, two cases of presence detection (\( j \in \mathcal{T} \)) can be identified: 1) terminal \( j \) can be correctly detected as active with probability \( P_D \) (probability of correct detection) provided the terminal has transmitted a packet (\( j \in \mathcal{T} \cup \mathcal{F} \)), and 2) terminal \( j \) is incorrectly detected as active with probability \( P_F \) (probability of false alarm) provided the terminal did not transmit a packet (\( j \in \mathcal{T} \). \( j \notin \mathcal{T} \)). Analytical expressions for \( P_D \) and \( P_F \) have been obtained for Rayleigh fading channels in [2].

**B. Signal model for multi-packet reception**

Each terminal \( j \) transmits packets with \( Q \) QAM symbols denoted by \( x_j = [x_j(0), x_j(1), \ldots, x_j(Q-1)]^T \), where \( (\cdot)^T \) is the transpose operator. Considering unitary power transmission \( E[x_j^H x_j] = 1 \), where \( (\cdot)^H \) is the Hermitian transpose operator, the signal vector received at the beginning of an epoch is given by \( \mathbf{y} = \sum_{j \in \mathcal{T}} h_j x_j + v \), where \( v = [v(0), v(1), \ldots, v(Q-1)]^T \) is the zero-mean and white complex Gaussian noise vector with variance \( \sigma_v^2 \): \( v \sim CN(0_Q, \sigma_v^2 I_Q) \) where \( 0_Q \) and \( I_Q \) denote, respectively, the vector of \( Q \) zeroes and the identity matrix of order \( Q \). The BS proceeds to estimate the collision multiplicity by means of terminal activity detection (explained in the previous subsection) and requests the number of necessary retransmissions (given by \( K - 1 \)) to resolve the collision. All the collected (re)transmissions are stored in memory to create a virtual MIMO system that can be expressed as follows [2] [5]: \( \mathbf{Y}_{K \times Q} = \mathbf{H}_{K \times K} \mathbf{S}_{K \times Q} \mathbf{V}_{K \times Q} \), where \( \mathbf{Y} \) is the array formed by the collection of all received signals from all the \( K \) time-slots of the epoch, \( \mathbf{H} \) is the mixing matrix or MIMO (multiple-input multiple-output) channel, \( \mathbf{S} \) is the array of stacked packets from all the contending terminals, each one with \( Q \) symbols, and finally \( \mathbf{V} \) represents the collected Gaussian noise components. The mixing matrix \( \mathbf{H} \) can be estimated by using the outcome of the matched filter operation from each retransmission [2]. The estimate \( \hat{\mathbf{H}} \) can be used to recover the contending packets. The contending signals can be estimated at the BS means by a linear decoding matrix \( \mathbf{G} \): \( \hat{\mathbf{S}} = \mathbf{G} \mathbf{Y} = \mathbf{GHS} + \mathbf{GV} \). This expression can be rewritten as follows: \( \hat{\mathbf{S}} = \mathbf{W}_1 \mathbf{S}_1 + \mathbf{W}_2 \mathbf{S}_2 + \mathbf{GV} \), where \( \mathbf{W}_1 = \mathbf{GH}_1 \), \( \mathbf{W}_2 = \mathbf{GH}_2 \), \( \mathbf{H}_1 \) is the mixing matrix of the contending terminals that have been detected as active (\( j \in \mathcal{T}, j \notin \mathcal{T} \)), and \( \mathbf{H}_2 \) is the mixing matrix of the contending terminals that have not been detected as active (\( j \notin \mathcal{T}, j \notin \mathcal{T} \)). The decoding matrix can be calculated using zero-forcing (ZF) or minimum mean square error (MMSE) equalization. The decoding signal for terminal \( j \) will experience a signal-to-interference-plus-noise ratio (SINR) given by:

\[
\Gamma_j = \frac{|W_1(j, k)|^2}{\sum_{k \neq j} |W_1(j, k)|^2 + \sum_k |W_2(j, k)|^2 + |g_j|^2 \sigma_v^2},
\]

where \( W_1(j, k) \) and \( W_2(j, k) \) denote the entries of matrix \( \mathbf{W}_1 \) and \( \mathbf{W}_2 \), respectively, that correspond to row and column of terminal \( j \) and terminal \( k \), respectively, and \( g_j \) is the row of matrix \( \mathbf{G} \) corresponding to terminal \( j \). It is assumed that a packet is correctly received when the SINR in (1) exceeds a decoding threshold denoted here by \( \beta_d \). The probability of a terminal transmission to be correctly decoded is thus denoted by \( \Pr(\Gamma_j > \beta_d) \).

**IV. Reception model and performance metrics**

The correct packet reception probability of \( q \) out of \( K \) contending signals provided \( K_d \) contending signals were correctly detected as active and \( K_f \) inactive terminals were incorrectly detected as active (false alarm) can be expressed as follows:

\[
\Pr(\mathcal{C}(q, K)_{K_d, K_f}) = \binom{K}{K_d} \left( \frac{J - K}{K_f} \right) \times \Pr(\cup_{j \in \mathcal{T}} \mathcal{G}_j > \beta_d | K_d, K_f), \quad q = |T_d|, \quad T_d \subset \mathcal{T} \cap \mathcal{F},
\]

where \( \binom{n_1}{n_2} \) is the combinatorial number of \( n_1 \) elements in \( n_2 \) positions, and \( \Pr(\cup_{j \in \mathcal{T}} \mathcal{G}_j > \beta_d | K_d, K_f) \) is the probability that \( q \) contending signals are correctly decoded as their experienced SINR exceeds the decoding threshold \( \beta_d \) conditional on the number of contending terminals \( K \), the number of contending terminals correctly detected as active \( K_d \) and the number of inactive terminals incorrectly detected as active \( K_f \) (false alarm).

**A. Throughput**

Packet throughput can be defined here as the ratio of the average number of packets correctly received (denoted by \( S \)) to the average length of an epoch-slot (\( E[l] \)):

\[
T = \frac{S}{E[l]},
\]

With the help of the reception model in (2), the numerator of (3) can be mathematically written as:

\[
S = \sum_{K=1}^{J} \sum_{q=1}^{K} \sum_{K_d=0}^{J-K} \sum_{K_f=0}^{K-K_d} \left( \frac{J}{K} \right) q(pP_d)^K_d (\bar{p}P_f)^K_f \mathcal{C}(q, K)_{K_d, K_f},
\]

where \( (\dagger) = 1 - (\cdot) \), which means that \( \bar{p} = 1 - p \). The expression in (4) represents the average of correct packet reception over all possible cases of transmission and terminal activity detection (correct and incorrect). The average length of an epoch in the denominator of (3) can be obtained by averaging over all possible cases of terminal activity detection, i.e., when an active terminal is correctly detected as active, or when an inactive terminal is incorrectly detected as active (false alarm). We recall the reader that the number of time-slots of each epoch is determined by the number of retransmissions necessary to make the MIMO system full-rank, which in our setting is given by \( \bar{K} \) [6]. The probability mass function (PMF) of length of an epoch \( l \) is thus given by:

\[
\Pr(l = m) = \begin{cases} \Pr(\bar{K} = m), & m > 1 \\ \Pr(\bar{K} = 0) + \Pr(\bar{K} = 1), & m = 1 \end{cases}
\]
It can be also proved that $\hat{K}$ has a binomial distribution with parameter $P_A = pP_D + \bar{p}P_F$, which can be written as
\[
\Pr\{\hat{K} = k\} = \binom{J}{k} \bar{p}^k P_A^{1-k}, \quad k = 0, \ldots, J.
\]
Therefore, $E[\hat{K}]$ can be obtained by averaging over the PMF of $l$ in 5, which yields $E[\hat{K}] = J\bar{p} + \bar{p}P_A$, where the second term $\bar{p}P_A$ stands for the contribution of one time slot in case any terminal is detected as active: $\Pr\{\hat{K} = 0\} = \bar{p}P_A$. The parameter $P_A$ is thus regarded as the total probability of terminal activity detection, and is given by the probability of correct detection in case of transmission plus the probability of false alarm in case of no transmission: $P_A = \Pr\{j \in T\} = \Pr\{j \in T|j \in T\} \Pr\{j \in T\} + \Pr\{j \in T|j \notin T\} \Pr\{j \notin T\} = pP_D + \bar{p}P_F$.

V. RESULTS

Let us now present some results that show the concepts explored in the previous sections. Consider a scenario with $J = 16$ terminals with an average signal-to-noise ratio (SNR) of $\frac{E_s}{N_0} = 10$ dB. All simulation results assume a packet decoding threshold for the signal-to-interference-plus-noise (SINR) ratio of $\beta = 2.5$, above which a packet is considered to be correctly received by the BS when there is no detection errors. It does not include the potential errors due to multi-user decoding. This curve is useful as reference for all other results. The results that include the errors due to multi-user decoding are labelled in Fig. 1 with the subscript $0$.

For example, the results using ZF decoder with correlation coefficient $\rho_r = 0.2$ are simply labelled $ZF_{\rho_r}$. The results show that further decoding errors of the multi-user decoding stage reduce the throughput with respect to the predicted value given by the detection throughput. Also note that the effect of correlation tends to reduce throughput performance at high values of traffic load, even affecting the stability bound (the maximum value of traffic load before the throughput curve rapidly decreases). However, at low traffic load, the results show that the highly correlated case can even slightly surpass the case with low correlation. The results without subscript have been obtained by using the concepts developed in this paper. Even in the case of incorrect detection of the collision multiplicity, the system attempts the decoding of the contending signals. It can be observed that for both types of decoder ZF and particularly for MMSE, the throughput of the protocol surpasses that of the conventional assumption (with subscript 0). In one of the cases the obtained throughput can even surpass the detector throughput, which means that more potential gains in NDMA can be obtained by using signal processing post-collision multiplicity estimation. The effects of correlated retransmissions are similar to the previous case analysed here. These results suggest that there can be some cases where NDMA can obtain benefit from correlated retransmissions.

VI. CONCLUSIONS

This paper has presented a more detailed analysis of the operation of a class of random access protocols assisted by retransmission diversity and signal processing tools for multi-user detection. The conventional analysis of the protocol ignores several details of correct detection and packet decoding that have been addressed in this work. It was found that temporal correlation and imperfect channel and collision multiplicity estimation can affect the stability and throughput performance at high traffic loads. By comparison, low temporal correlation and MMSE decoding can even surpass the predicted detection throughput of the protocol, which opens further possibilities for improvement of the performance of this type of algorithm.

REFERENCES