SOME EXPERIMENTS WITH ARRAY DATA COLLECTED IN ACTUAL URBAN AND SUBURBAN ENVIRONMENTS

A. Lee Swindlehurst1  Malcolm J. Goris2  Björn Ottersten3

1Dept. of Signals, Sensors, & Systems, Royal Institute of Technology, Stockholm, Sweden (On leave from Dept. of Electrical Engineering, Brigham Young University, Provo, UT, USA).
2Netherlands Foundation for Research in Astronomy, Dwingeloo, the Netherlands.
3ArrayComm, Inc., San Jose, CA, USA (On leave from Dept. of Signals, Sensors, & Systems, Royal Institute of Technology, Stockholm, Sweden.)

ABSTRACT

We compare the performance of several algorithms for signal separation based on actual mobile cellular radio data. The data were collected by basestations in two different environments: using an eight element linear array on a hillside overlooking a suburban area, and using a four element square array at the top of a ten story building in a dense urban area. Calibration data were available for one of the arrays used for the suburban data, allowing for performance comparisons between DOA-based and blind signal separation algorithms. Experiments were conducted with three different types of signal waveforms: sinewaves, simple FM waveforms, and 9/4-DQPSK signals, each of which necessitated the use of a different performance metric. We focus here on the results obtained for cases where the data were approximately low rank, with little delay spread (although not all data sets from the urban environment were of this type).

1. INTRODUCTION

Multiple-antenna arrays have been considered as a means of accommodating increasing numbers of users in cellular radio systems. An array can be used to service multiple, spatially-separated users in a common cell and on a common frequency by adaptively amplifying the signal from each user while rejecting interference from the other users. An array can reject interference from adjacent cells and increase the signal-to-noise ratio of the received and transmitted signals. These properties can lead to more users per cell, re-use of frequencies in adjacent cells, smaller re-use distances, and lower mobile-handset power requirements. In order to realize these benefits, algorithms that estimate the spatial signature of each of the received signals must be used.

Many different techniques have been developed for separating co-channel communications signals. These techniques can be grouped into two categories: those that use spatial calibration information, and those that do not. Algorithms in the latter category are often referred to as “blind” beam-formers, and usually rely on some type of temporal information about the transmitted signals. The term “blind” is also used to refer to techniques that do not use sequences for training either spatial or temporal filters. In either case, little work has been published on the performance of these algorithms with real mobile cellular data, in real communication environments. This is the principal goal of the paper.

Data were collected by ArrayComm, Inc. and Allgon Systems AB, with help from the authors, in two realistic settings. In the first set of experiments, an eight element linear array was mounted on a tower on a hillside overlooking a suburban area, and was used to collect a variety of different signals at 852.27 MHz from vehicles in normal traffic (0-50 km/h). In the second, a four element square array was mounted on the top of a building in an urban environment, and received DQPSK data centered at approximately 1.9 GHz from a hand-held mobile phone below. A more detailed description of the experimental setup is given in the next section. The data were stored and processed off-line using a number of different techniques, which are briefly outlined in Section 3. To conclude the paper, some selected results from this processing are presented in Section 4.

Before continuing, we note that it would be difficult, not to mention unwise, to draw too many general conclusions about the algorithms studied from the results presented below. Though the scenarios considered are representative of typical mobile cellular networks, the relative performance of the algorithms may be quite different under other circumstances. Indeed, the results presented may tell more about the validity of the modeling assumptions behind the algorithms than the performance of the algorithms themselves.

2. EXPERIMENT DESCRIPTIONS

Results will be reported for data collected in three separate experiments, two conducted at the suburban site, and one in the urban environment. The test parameters for each case are described below.

2.1. Suburban Collect 1 (SUB1)

For this set of experiments, a 12-element dual-polarized linear array constructed by Allgon System AB of Sweden was mounted at the top of a 15 meter tower on a hillside overlooking a relatively flat residential area. The combined height of the antenna (hill + tower) was several hundred feet. The array elements were uniformly spaced 17.7 cm apart, and each consisted of cross-polarized antennas oriented at ±45 degrees. Only a single polarization output for each of the inner eight elements of the array (elements 3-10) were used in the experiments reported here. Gain and phase patterns for both the azimuth and elevation dimensions were measured at Allgon's test facility in Sweden prior to the array's shipment to the US for use in the experiments. In addition to this antenna calibration data, a receiver calibration was performed by measuring the on-site response of the array/receiver combination to a strong far-field broadside source with (nominally) no other sources present. Data from the array's RF front end were sampled at 71.4 kHz, and stored in blocks of 512 samples.

The results presented below for this collect were obtained using two mobile transmitters in vehicles moving through residential traffic at speeds ranging from 0-50 km/h. Both sources transmitted data centered at $f_c = 852.27$ MHz, and both were located at a range of roughly 2-3 km from the array. At this frequency, the array elements were just under one half wavelength apart. The first source transmitted a simple sinewave at a frequency of $f_s = 850$ Hz. Its azimuth position varied between approximately $-8^\circ$ and $-12^\circ$ relative to broadside throughout the collect, and its mean (single sensor) SNR was estimated to be 9.0 dB with a standard deviation of 8.2 dB due to fading. The second source broadcast a standard IS-54 9/4-DQPSK signal (35% square-root raised cosine pulse) with a pseudo-random symbol stream.
and a band rate of 24.3 kHz; its azimuth angle varied between roughly \(-25^\circ\) and \(-38^\circ\), and its mean SNR estimate was 20.3 dB with a standard deviation of 8.3 dB. While the array overlooked the area where the mobiles were located, there was seldom a line-of-sight (LOS) path between them due to trees and small buildings. Any multipath present in the data appears to be confined to a small angular sector surrounding each mobile, and is coherent (essentially zero delay spread).

Because the sampling rate of the array was not an integer multiple of the baud rate of the DQPSK signal, it was difficult to accurately measure the quality of the digital source after separation. For this reason, the performance metric used with this collect was based on how closely the signal copied from the sinewave transmitter approximated the estimated sinewave, the standard deviation of the signal’s (unwrapped) phase difference \(\delta \phi(t + 1) - \delta \phi(t)\) was computed in each trial.

2.2. Suburban Collect 2 (SUB2)

The location and parameters of this collect were essentially identical to the first, with the following exceptions: a 12-element array of dipoles spaced 18 cm apart was mounted on the tower (only the inner eight were used, as before), and both mobiles transmitted analog FM modulated 1 kHz sinewaves. In addition, no element gain nor phase calibration was available for this array, although an on-site receiver calibration was performed, as above. The average sensor SNR for both sources was estimated to be in excess of 27 dB, with a standard deviation of 6 dB, and the azimuth angles of the two sources varied in the intervals \([-5^\circ, 5^\circ]\) and \([-35^\circ, -25^\circ]\).

The center frequencies of the signals were at \(f_c \pm 7.5\) kHz, and since their bandwidths were only a few kilohertz, they were essentially spectrally disjoint. Since the SNR is quite high, one can easily measure the amount of interference rejection by simply taking the highest peaks in the frequency source after separation. For this reason, the performance metric used with this collect was based on how closely the signal copied from the sinewave transmitter approximated the estimated sinewave, the standard deviation of the signal’s (unwrapped) phase difference \(\delta \phi(t + 1) - \delta \phi(t)\) was computed in each trial.

2.3. Urban Collect (URB)

The urban data sets were collected by a four-element square array mounted on the top edge of a ten story building in a metropolitan area. The mobiles were hand-held phones transmitting PHS (for Personal Handphone System) data with a carrier frequency near 1.9 Ghz. At this frequency, the elements of the array were separated by over five wavelengths. The PHS standard is similar to IS-54, except the noise term \(N\) is defined similarly to \(X\) and \(S\). By the term “low-rank” we mean that each source makes only a rank one contribution to \(X\), and that the total number of sources \(d\) is less than the number of antennas. We also define the covariance of the data as

\[ R_{xx} = \lim_{N \to \infty} \frac{1}{N} \sum_{N-iw} X^* X = A R_{ss} A^* + R_{nn}, \]

where \(R_{ss}\) and \(R_{nn}\) represent the covariance of the signals and noise, respectively.

The algorithms studied below attempt to use certain assumptions about the array and received signals to factor the low-rank portion of the data into terms involving \(A\) and \(S\) (or \(A\) and \(R_{ss}\) when \(R_{xx}\) is used instead of \(X\)) that are consistent with these assumptions. The algorithms may be grouped into the categories listed below based on their defining features.

**Direction of arrival (DOA) estimation** – These techniques assume the availability of calibration data for the array and use it to estimate the DOAs \(\theta = [\theta_1, \ldots, \theta_d]\) of all signals. Once the estimate \(\hat{\theta}\) is found, the columns of \(A\) are set equal to the corresponding vectors from the so-called array manifold: \(\hat{A} = A(\hat{\theta})\). In the case of multipath, it may be necessary to fit a linear combination of several calibration vectors to each source. The calibration data may be obtained empirically, or by assuming identical antennas in known locations (sometimes referred to as an “analytical” calibration). The DOA-based methods studied in this paper were MUSIC [1], ESPRIT [2], Weighted Subspace Fitting (WSF) [3], and the classical delay-and-sum beamformer (DSB).

**On-line auto-calibration (AC)** – There are a number of such algorithms, each based on a different set of assumptions. Two algorithms from this class were considered, and will be referred to as AC1 [4] and AC2 [5]. Both assume that the signals are uncorrelated (\(R_{ss}\) is diagonal), and that their spatial signatures have angle independent gain responses. These methods operate by trying to iteratively decompose the signal subspace part of \(R\) into a product \( \Gamma \Gamma^* \), where \(\Gamma\) is real and diagonal (the angle independent antenna gains), the elements of \(\Gamma\) have unit amplitude and arbitrary phase (the antenna element phases), and \(\Gamma\) is real and diagonal (the signal powers).
Second- and fourth-order cumulants to factor out were used to process the data, the "zero-forcing" solution from which a set of beamformer or "signal copy" weights were derived. The algorithms tested in this category were the normalized LMS adaptive CM algorithm (NCMA) [6], the iterative LS-CMA method [7], and the CM factorization (CMF) approach [8].

Finite alphabet (FA) signals - Algorithms from this class assume that the sources transmit digital communications signals with known modulation. Of the available FA techniques, only the decision-directed (DD) beamformer presented in [9] was implemented.

Higher order statistics (HOS) - A number of algorithms for blind source separation based on HOS have also been developed, but only the JADE algorithm of [10] was considered here. JADE assumes only that the signals are uncorrelated and non-Gaussian, and uses the structure of the resulting second- and fourth-order cumulants to factor out the unknown factorization matrix. An advantage of the DOA methods is that in high SNR situations, very little data and relatively little computation are required for them to "converge" to a reasonable solution. Figure 2 shows that although they worked well with only rudimentary calibration data, the CM algorithms offer nearly 10 dB more interference rejection. When accurate calibration data were available, the performance difference was much smaller, as evidenced by the results using the Allgon array. Note that no attempt was made to use the DOA-based algorithms with the urban data sets.

4. SOME RESULTS

Figures 1-6 display some results obtained by processing data from each of the three collectives. Most of the figures show performance as a function of \( N \), the number of data samples used by each algorithm to determine \( W \). Once computed, the weights were applied to the entire burst, which typically contains many more than \( N \) samples. Note that Figures 3-4 show the performance of each method sorted from worst to best to give an indication of algorithm robustness. Some comments about the results for each algorithm class are given below.

DOA methods - Due to the fact that the sources had high SNR, low correlation, and were sufficiently well separated, there was little difference in performance between the various DOA-based methods (except the DSB). For this reason, only the results of one or two algorithms are shown in the figures. An advantage of the DOA methods is that in high SNR situations, very little data and relatively little computation are required for them to "converge" to a reasonable solution.

Figure 1. Results from Collect SUB1.

Figure 2. Results from Collect SUB2.

Figure 3. Results from Collect SUB2 sorted from worst to best, \( N = 8 \).

AC methods - The performance of the AC1 and AC2 algorithms was very similar, which is not surprising since both attempt to factor the covariance matrix in the same way. In some of the figures then, only the performance of AC2 is shown since it was slightly better. It was also noted that the AC2 iterations tended to converge somewhat more quickly than AC1. The angle-independent gain assumption used by the AC methods is least likely to hold in a multipath environment with large angular spread, which is probably the explanation behind the algorithms' poor performance with...
Figure 4. Results from Collect SUB2 sorted from worst to best, $N = 500$.

Figure 5. Results from Collect URB, 2 sources present, results for source at 500m shown.

the urban data.

CM methods - The LS-CMA algorithm was initialized by WSF in Collects SUB1 and SUB2, and by AC1 in Collect URB. Both LS-CMA and CMF performed very well, yielding accurate source separation with very little data. The only exception was perhaps the SUB1 data set, where the more severe IS-54 pulse shaping filter reduces the CM property of the DQPSK signal. The results shown for the adaptive CMA algorithm in Figures 2-4 were obtained by initializing it with the nominal array response vector corresponding to a point roughly between the two sources.

DD method - As observed with other DD techniques, a certain data length threshold must be passed before an improvement in performance is achieved. In the urban data sets processed, this threshold was on the order of 40-50 symbols, beyond which the SER was reduced by 1-2 orders of magnitude over the CM methods.

JADE - In most cases, JADE was the slowest algorithm to converge, due to its reliance on higher order statistics of the data. However, it is interesting to note the algorithm's excellent performance and fast convergence with the SUB1 data. A possible explanation is the fact that the signals in this data set were the least correlated of those processed.

Acknowledgment
The authors wish to thank Richard Roy and Marc Goldberg of ArrayComm, Inc. and Claes Beckman of Allgon System AB for generously providing the data used in the experiments.

REFERENCES