Malm, Nicolas; Zhou, Liang; Menta, Estifanos; Ruttik, Kalle; Jäntti, Riku; Tirkkonen, Olav; Costa, Mario; Leppänen, Kari

User Localization Enabled Ultra-dense Network Testbed

Published in:
IEEE 5G World Forum, 5GWF 2018 - Conference Proceedings

DOI:
10.1109/5GWF.2018.8517071

Published: 31/10/2018

Please cite the original version:
Abstract—We present an over-the-air testbed designed for assessing the performance of ultra-dense networks and related mobility management schemes based on tracking the location of user equipment. Location information of UEs at the physical layer of the radio access network allows for proactive and seamless handovers. This is particularly important for 5G ultra-dense networks where small cells may otherwise be overwhelmed by handover related signalling. We exploit sounding reference signals transmitted by UEs to track their location. Novel mobility management schemes can then be designed such that handover triggering is done by the network and without UE involvement. In current cellular networks, handover is triggered by downlink measurements carried out at the UE, and subsequent reporting. This work also validates our testbed’s proof-of-concept software architecture based on stateless unordered worker threads. In particular, a Cloud-RAN of up to four radio transceiver points can be combined into a network overseen by one gNB. Also, dataplane support allows for experimentation with real-time data handovers. The obtained results demonstrate the feasibility of location-based handover schemes as a tool for future cellular network design.

Index Terms—SDR, UDN, low-latency, location based handover

I. INTRODUCTION

Ultra-dense networks (UDN) have emerged as a major trend in the development of fifth generation (5G) cellular systems. The predicted growth [1] in both the number of users and data volume places great pressure on future networks. Reaching future mean rate requirements for vehicular users in cities, e.g. for applications such as virtual reality, will require significant densification of outdoor city networks [2]. The result is a network structure comprised of smaller, densely packed cells with high rates of handovers for moving users. In a Cloud-RAN architecture, transmission-reception points (TRP) form cells, and multiple cells are controlled by a gNB. Frequent transfer of connections between TRPs using legacy mobility management schemes will result in excessive burden on the system due to handover signalling [3].

Addressing the aforementioned challenges is essential for any practical implementation to be feasible. Previous research [4] has studied the possibility of utilizing location data to assist in handover decisions. These schemes were however deemed too complex to implement and provided accuracy of approximately 25-30 m at best, which is close to UDN TRP cell size. Positioning mechanisms in 2G and 3G networks also suffer from higher overhead due to a lack of support at the physical layer of the radio access network (RAN). One proposed solution for 5G is uplink (UL) sounding reference signals (SRS) based localization [2]. In this scheme, UEs periodically transmit SRSs, which are received by TRPs and exploited to determine the angle-of-arrival of the line-of-sight (LoS) path. Angles from multiple TRPs can then be fused in a central controller to obtain an accurate (sub 1 m) estimate of the user’s position [5].

This paper presents a novel testbed for assessing the performance and feasibility of location based mobility management schemes in UDNs. In doing so, we identify the challenges that arise. Furthermore, the architectural solution proposed in this work is presented.

The UDN architecture selected for study was proposed in [2]. We present its main features in Section II while Section III describes the version implemented in the testbed, which strives to be as close to the selected architecture as feasible within technical constraints. In Section IV we report initial measurement results acquired from the testbed. Section V concludes this paper.

II. UDN ARCHITECTURE

A. The Need for a New Type of Radio Interface for UDN

As indicated in [2], massive MIMO from macro cells has great challenges to serve vehicular users with 5G data rates in dense urban areas, whereas a UDN was shown to have significantly better mobility performance. In order to save costs, the UDN nodes should be connectable with a low-capacity wireless link to a central unit. With such an arrangement, and an inter-site distance of 30-50 m, the current mobility schemes of LTE need to be improved [6]. Mobility management based on uplink SRS from the UE, has been proposed e.g. in [3]. This solution is promising for UDN as there is no UE measurement reporting signalling overhead over-the-air. In this paper a testbed to verify these ideas in practice is presented.
B. Overview of Selected Architecture

In the selected architecture [2], each TRP is expected to have line-of-sight (LOS) conditions with most UEs in the cell, for most of the time. TRPs serve the users in their local area that have been assigned to them by a controller (labeled gNB). This node collects reports on UEs’ SRSs from TRPs and computes each user’s location. This information is then used by the gNB to decide when to re-assign a UE from one TRP to another.

The transfer of a UE from one TRP to the next within an area should be seamless from the UE’s perspective. Since the UE continuously transmits an SRS allowing the network to track its location, it does not need to perform handover measurement and reporting. When the need for a change of serving TRP arises, the gNB sends an appropriate set of instructions to the current and target TRPs. The latter may need to synchronize the state information it holds on the UE. When the transfer from source to target TRP takes effect, the UE will simply receive its transmissions from another TRP and computes its new location. This benefit provided is twofold. Firstly, air interface handover signalling is eliminated. Secondly, mobility becomes seamless from the UE perspective as it does not need to perform any procedure with the target TRP.

A further benefit of using high-accuracy location information for handovers is the ability to distinguish cases of low SNR due to pathloss from those due to fast fading. Figure 1 illustrates a stationary UE in the middle of two TRP cell borders (dashed lines). In a system with conventional handovers, a drop in the SNR might have initiated a handover procedure, or at the very least neighbor measurements. In a UDN, knowledge of the UE’s geographic position allows the gNB to determine that the SNR drop is not due to increasing distance from the serving TRP. Knowing that the UE is already being served by the closest base station enables the gNB to avoid an unnecessary handover in case of a temporary weakening of the received signal strength and no better TRP being available.

In [2], TRPs use a circular array of 25 antennas. The bandwidth of the system in both DL and UL is 200 MHz with a sub 1 ms latency. To achieve the latter goal, the TTI length is set to 200 µs.
time extensions are required. The platform does not require
on an umodified Linux kernel. In particular, no hard–real–
C. Software Implementation
received signal strength (RSS) of the UE uplink signal.
acknowledgment status of past data transmissions and the
frontend clock rates. In the ARF platform, the TRPs report
the system to tolerate some level of divergence in OS and
ically increasing logical time. Each TRP maintains its own
is challenging, the gNB continuously broadcasts a monoton–
time synchronization across many devices with varying loads
between gNB and TRPs. These links have limited capacity
from all TRPs. The limitiations come from the links be–

A. Hardware Configuration
A TRP is comprised of a PC and two USRP X-series [9] ra–
dio frontends with UBX-160 daughterboards connected using
10G Ethernet, each supporting two antennas. Communication
between gNB and TRP occurs over a 1G Ethernet link. The
limiting factors for testing different bandwidths are the
fronthaul link capacity and processing speed of the computer.
The latter was equipped with an Intel Xeon D-1541 [10] CPU,
running Ubuntu 16.04 as the operating system.

B. Functional Split
Three node types are implemented: gNB, TRP and a UE. In
an effort to increase scalability, each node type is designed
to be as independent of the others as possible. This helps to
reduce the need for communication between nodes. Figure 1
depicts the division of positioning related tasks amongst nodes.
Since one gNB can potentially be responsible for a large
number of TRPs, it cannot realistically process all samples
from all TRPs. The limitations come from the links be–
tween gNB and TRPs. These links have limited capacity
and introduce latency. Once received at the TRP, the data
must undergo network protocol processing, adding further
latency. A subset of the information available at the TRP must
therefore be selected for transmission to the gNB. The latter
handles centralized control and sends instructions to the TRP,
which is then responsible for their execution. New information
generated at the TRP is reported immediately while the gNB
issues instructions once per frame (equal to 5 TTIs). Since
time synchronization across many devices with varying loads
is challenging, the gNB continuously broadcasts a monoton–
ically increasing logical time. Each TRP maintains its own
local mapping from this system time to its own local time.
This enables independent conversion factors, thus allowing
the system to tolerate some level of divergence in OS and
frontend clock rates. In the ARF platform, the TRPs report
a list of currently detected UEs, azimuth angle to the UE,
acknowledgment status of past data transmissions and the
received signal strength (RSS) of the UE uplink signal.

C. Software Implementation
The ARF is implemented in C++14 and designed to run
on an unmodified Linux kernel. In particular, no hard–real–
time extensions are required. The platform does not require
any specialized hardware other than a radio frontend. Use of
accelerators, such as offloading network cards, is nonetheless
supported. Performance can also be improved by upgrading the
processor as the ARF can scale to exploit an increased CPU
core count. Figure 2 presents the main software components
of a TRP.

Node management logic handles communication with the
gNB over Stream Control Transmission Protocol (SCTP). It
transmits updates on the status of the TRP and receives
instructions. This sets an execution speed constraint equal to
the gNB frame length (1 ms). Another task handled by the
management logic is the user interface, for example to turn
sample recording on and off. These tasks run independently
of the RAN code.

The air interface implementation must meet an execution
deadline equal to the TTI length of 184 µs. The radio protocol
logic and state machine execute in one thread. Decisions
made in this component are passed to the entities performing
the actual execution of the tasks via timestamped buffers.
The messages passed through these buffers carry a copy of
all the necessary information for the processing of the TTI.
Consequently, all coupling between protocol logic and L1 can
be eliminated.

L1 functionality is organized into separate TX and RX
instances. The transmit side instances perform the tasks required
to create the IQ samples for one TTI, while the RX instances
decode information from received IQ samples. Decisions made
in the protocol logic thread are implemented based on the
messages received from the radio protocol logic. Each L1
instance runs in its own thread. L1 entities have no thread–
local memory making them independent and interchangeable
to avoid introducing any implementation related dependencies.
Any TTI can therefore be handled on any one of the processing
threads. This helps make the platform more scalable vis–`a–vis
CPU core count. Additionally, since no couplings between L1
instances and other parts of the system exist, implementation
may be swapped without affecting the rest of the platform.
For example, it would be possible to change from a CPU-based
implementation to a graphics processing unit (GPU)-based one
with no changes to other parts of the codebase.

The virtual hardware enhancement layer (VHEL) [11]
prents a generic interface to L1. The primary function of

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TX-1</td>
<td>Manages frontends</td>
</tr>
<tr>
<td>TX-2</td>
<td>Manages protocol state machine</td>
</tr>
<tr>
<td>RX-1</td>
<td>Sends reports to gNB</td>
</tr>
<tr>
<td>RX-2</td>
<td>Receives instructions from gNB</td>
</tr>
<tr>
<td>L1 TTI buffers</td>
<td>Sends and receives messages from L1</td>
</tr>
<tr>
<td>VHEL</td>
<td>Manages protocol state machine</td>
</tr>
<tr>
<td>gNB frames (1 ms)</td>
<td>Sends and receives messages from L1</td>
</tr>
<tr>
<td>Air interface TTI</td>
<td>Sends reports to gNB</td>
</tr>
<tr>
<td>(184 µs)</td>
<td>Receives instructions from gNB</td>
</tr>
</tbody>
</table>

**Fig. 2. Main components of the ARF software architecture.**
the VHEL is to mask timing constraints and platform non-idealities from L1 entities. Lost or late samples are replaced with zeros. The VHEL also handles the grouping of IQ samples into TTIs or conversely splits TTI buffers into the correct number of samples to send to radio frontends. The presence of the VHEL allows protocol code to operate unmodified on any supported hardware model or using a simulator.

D. Issues and Solutions

Several issues were encountered during the development of the platform. First, it was observed that the frontends used will start with a random phase offset after each restart of the software. This significantly impacts the achievable positioning accuracy since a known phase relationship among the antennas is crucial for angle-of-arrival based positioning. Note that in the TRP of the testbed, there are two USRPs, each controlling two RX antennas. The relative phase between the antennas controlled by one USRP are fixed for a given UE position, whereas the phase difference between antennas of different USRPs is a priori random, and fluctuates with clock drift. Accordingly, such a random offset leads to (random) changes in the measured angles even though the physical position of transmitters and receivers does not change. This is due to the clock drift at the frontends. To remedy local oscillator stability shortcomings, we introduced an additional transmitter that serves as a reference self-calibration device. This device synchronizes over-the-air to the TRP. Once synchronized, it transmits a known reference signal from a known position. The TRP receives this transmission through all of its antennas and computes the differences in phase. This result can then be compared to that predicted based on the position of the calibration reference antenna. Online calibration also helps combat the effects of dissimilar phase drift in each frontend.

Another challenge was presented by the synchronization of the UE to the TRPs frame structure. Even small offsets greatly degraded localization performance. The testbed uses a TDD frame structure, where each TTI has both a downlink and uplink portion. Initial synchronization proved easier in this regard as the UE itself does not transmit. During tracking synchronization, TX leakage into the RX chain causes strong transceiver self noise that can compromise the results. This issue becomes especially pronounced at the cell edge, where the SNR is low. In order to take the reliability of measurements into account, the TRP reports a confidence value, along with the AoA and the RSSI, to the gNB when making a UE position report. Doing so also helps offset the loss of information incurred by only transmitting the final computed angle instead of the raw IQ samples. A balance is therefore reached between the quantity of data to transfer and the accuracy of positioning.

IV. MEASUREMENTS AND RESULTS

We collected angle-of-arrival data from the testbed in order to validate the ability of the UL SRS based UDN concept to provide sufficiently accurate localization information for handover decisions. The obtained AoA estimates are compared to a physical measurement of the true angle to determine the achieved accuracy of the testbed. For each true angle, estimates from 500 TTIs were collected. The platform was operated using a TTI duration of 184 µs with a bandwidth of 15.36 MHz (12.48 MHz effective). TTI duration was chosen to match the system proposed in [2] as closely as possible. This decision was taken before the 3GPP selected New Radio (NR) air interface parameters for the TS 38.211 draft standard [12]. The main differences pertain to TTI duration, supported bandwidth configurations and uplink-downlink split. Since our measurements were conducted using a single static UE transmitting an UL signal in only one OFDM symbol, these differences do not have a significant impact on the applicability of the results obtained. Positioning accuracy was tested at five different angles of arrival: -45°, -22.5°, 0°, 22.5° and 45°. The broadside of the array is at 0 degrees. The measurements were carried out outdoors in a parking lot using a linear array of four dipole antennas. A single UE equipped a dipole antenna at a height of 1.5 m was used. Transmitter gain values were set at the maximum supported by the hardware resulting in approximately 20 dBm [13] of power sent to the antenna before cable and connector losses.

Figure 3 shows the estimated and measured actual AoA of the UE with regards to the TRP. The 50 % and 90 % confidence intervals are shown. The interquartile range indicates that UL pilot based localization can estimate the AoA to within 2° to 3° precision. These measurements were made using only one OFDM subcarrier. Significant gains are expected to be achieved by improving the algorithm to utilize the full system bandwidth. The presented single-OFDM subcarrier results are already sufficient to validate the ability of the ARF UDN platform to support the desired functionality and carry out the intended localization measurements.

<table>
<thead>
<tr>
<th>Angle</th>
<th>Median</th>
<th>Bias</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>-45°</td>
<td>-61°</td>
<td>-16°</td>
<td>3.5°</td>
</tr>
<tr>
<td>-22.5°</td>
<td>-38°</td>
<td>-15.5°</td>
<td>2°</td>
</tr>
<tr>
<td>0°</td>
<td>-5.5°</td>
<td>-6.5°</td>
<td>1.5°</td>
</tr>
<tr>
<td>22.5°</td>
<td>16.5°</td>
<td>-6°</td>
<td>2.5°</td>
</tr>
<tr>
<td>45°</td>
<td>52°</td>
<td>7°</td>
<td>3°</td>
</tr>
</tbody>
</table>

Fig. 3. Estimated and actual angle of arrival for five UE positions.

TABLE II

Comparison of True and Estimated Angles-of-Arrival. The estimates include the median, the bias and interquartile range (IQR).
The results exhibit a systematic, angle-dependent bias. This bias is likely due to a discrepancy between the actual position of the reference antenna and its nominal one used in the calibration algorithm. Table II contains the median, bias and interquartile range for each tested angle. The data shows AoA estimation accuracy being superior closer to the broadside of the array, as expected.

V. CONCLUSION

Hardware-in-the-loop testing of current and future proposed 5G algorithms is possible using the developed over-the-air testbed. It helps to validate algorithm performance under operational conditions and gives feedback in the form of physical layer measurements. The testbed provides a tool to develop improved algorithms. Yet, the results obtained show promise as to the ability of the ARF UDN platform to provide sufficiently accurate positioning for handover decision making even when using a relatively simple AoA estimation method. The accuracy obtained exceeds that of legacy cellular systems even using only a single subcarrier. Performance can be further increased by employing all subcarriers as well as averaging over time. Future work include utilizing the data gathered during the measurements to optimize the algorithms employed, including elimination of the observed systematic bias. Further work aims to implement the aforementioned improvements along with repeating the measurements using a network comprised of multiple TRPs and a gNB controlling them.

REFERENCES


