IEEE P802.11
Wireless LANs

High-Accuracy Indoor Geolocation using Collaborative Time of Arrival (CToA) - Whitepaper

Date: 2017-09-07

Author(s):

Name	Affiliation	Address	Phone	email
Leor Banin	Intel Corporation	94 Em Hamoshavot Rd. Petah Tikva, Israel, 49527		leor.banin@intel.com
Ofer Bar-Shalom	Intel Corporation	94 Em Hamoshavot Rd. Petah Tikva, Israel, 49527		ofer.bar-shalom@intel.com
Nir Dvorecki	Intel Corporation	94 Em Hamoshavot Rd. Petah Tikva, Israel, 49527		nir.dvorecki@intel.com
Yuval Amizur	Intel Corporation	94 Em Hamoshavot Rd. Petah Tikva, Israel, 49527		yuval.amizur@intel.com

Abstract

Collaborative time of arrival (CToA) is the next generation, indoor geolocation method, which is designed for enabling scalability of the existing IEEE802.11/Wi-Fi-based, geolocation systems. The technique leverages on the IEEE802.11 fine timing measurements (FTM) capabilities, enabled in state-of-the-art Wi-Fi chipsets, and supports two *concurrent* operation modes; the CToA "client-mode" enables "GPS-like" operation indoors, and allows an *unlimited* number of clients to privately estimate their position and navigate indoors, without exposing their presence to the network. The CToA "network-mode" is designed for large-scale asset-tracking applications, and enables a centric positioning server to pinpoint objects equipped with wireless, Wi-Fi-based, low-power electronic tags (e-Tags).

The CToA method is broadcast-based and operates over an un-managed network, built out of cheap, unsynchronized units called "CToA broadcasting stations" (bSTA). The bSTAs, which are stationed at known locations, periodically broadcast a unique beacon transmission and publish its time of departure (ToD). Neighbor bSTA units and clients that receive the beacon broadcast, measure and log its time of arrival (ToA). Every bSTA publishes its most recent timing measurement log as part of its next beacon broadcast. CToA clients combine their own ToA measurements with those published by the bSTAs, in order to estimate and track their location.
CToA e-Tag clients act similar to bSTAs, and simply wake-up sporadically to broadcast a CToA beacon. The ToA of that broadcast is measured and logged by the receiving bSTAs similarly to beacons broadcast by other bSTAs. The timing measurement report is then delivered to a centric positioning server that can estimate and track the location of numerous CToA-based e-Tags, simultaneously.

The paper outlines the principles of the CToA method and the mathematical background of the position estimation algorithms. In addition, performance examples as well as theoretical analysis of the expected positioning accuracy are provided.

High-Accuracy Indoor Geolocation using Collaborative Time of Arrival

Leor Banin, Ofer Bar-Shalom, Nir Dvorecki, and Yuval Amizur

***Abstract*—Collaborative time of arrival (CToA) is the next generation, indoor geolocation protocol, which is designed for enabling scalability of the existing IEEE802.11/Wi-Fi-based, geolocation systems. The protocol leverages on the IEEE802.11 fine timing measurements (FTM) capabilities, enabled in state-of-the-art Wi-Fi chipsets, and supports two *concurrent* operation modes; the CToA "client-mode" enables "GPS-like" operation indoors, and allows an *unlimited* number of clients to privately estimate their position and navigate indoors, without exposing their presence to the network. The CToA "network-mode" is designed for large-scale asset-tracking applications, and enables a centric positioning server to pinpoint objects equipped with wireless, Wi-Fi-based, low-power electronic tags (e-Tags).**

The CToA protocol is a broadcast-based protocol that operates over an un-managed network, built out of cheap, unsynchronized units called "CToA broadcasting stations" (bSTA). The bSTAs, which are stationed at known locations, periodically broadcast a unique beacon transmission and publish its time of departure (ToD). Neighbor bSTA units and clients that receive the beacon broadcast, measure and log its time of arrival (ToA). Every bSTA publishes its most recent timing measurement log as part of its next beacon broadcast. CToA clients combine their own ToA measurements with those published by the bSTAs, in order to estimate and track their location. CToA e-Tag clients act similar to bSTAs, and simply wake-up sporadically to broadcast a CToA beacon. The ToA of that broadcast is measured and logged by the receiving bSTAs similarly to beacons broadcast by other bSTAs. The timing measurement report is then delivered to a centric positioning server that can estimate and track the location of numerous CToA-based e-Tags, simultaneously.

The paper outlines the principles of the CToA protocol and the mathematical background of the position estimation algorithms. In addition, performance examples as well as theoretical analysis of the expected positioning accuracy are provided.

***Index Terms*—geolocation, Indoor navigation, fine timing measurement, FTM, time delay estimation, Maximum likelihood estimation, WLAN, Wi-Fi, IEEE 802.11**

I. Introduction

THE challenge of accurate indoor location and navigation has been attracting an increasing amount of attention since the mid 1990's. Cultivated by the cellular revolution and the U.S. federal communication committee (FCC) enhanced 911 services (E911) [1], indoor location has ignited a rapid development of mobile location technologies. The ubiquity of IEEE802.11™ wireless local area network (WLAN) technology in mobile devices, which to date, has already reached an attach-rate of 100% in the smart-device segment [2], facilitated the development of WLAN-based indoor location systems. Due to the lack of standard infrastructure for high-resolution timing measurement capabilities in its early releases, existing WLAN-based location technology relies, to a great extent, on the WLAN received signal strength indicator (RSSI) infrastructure. The RSSI is a measure of the RF energy received by the station. WLAN stations estimate the RSSI of the beacons broadcast by access points (AP), and use this metric to sort between the APs based on their signal quality and proximity. The RSSI metric is measured in units of [dBm], and in general is inversely proportional to the logarithm of the squared distance between the transmitter and the receiver [16]. RSSI-based mobile device positioning exists in two main flavors: path-loss models, and "fingerprinting". Path-loss models relate the received signal power to the propagation distance. A set of RSSI measurements obtained from different WLAN APs in the vicinity of the client station, enable it to estimate its position via trilateration methods [18], [19]. While this method is relatively simple to implement, it is prone to yield fairly inaccurate positioning results due to the large variations in the RSSI measurements [6]. The alternative approach is to correlate the RSSI measurements against a pre-calibrated database of RSSI "fingerprints", measured over a pre-defined grid and stored in a server. The fingerprint approach provides better accuracy compared to the path-loss based RSSI. However, as the method's accuracy is sensitive to even minor changes in the propagation channel (e.g., a placement of a new sales-stand in a shopping mall), frequent re-calibrations and updates of the fingerprints database are required. The high-maintenance incurred by this type of positioning systems obviously limits their scalability.

Facing the limited positioning accuracy enabled by RSSI/path-loss based location technologies and the limited scalability of fingerprint-based systems, industry vendors began seeking alternative WLAN-based positioning technologies, which will enable to achieve higher positioning accuracy. Taking advantage of the high bandwidth supported by the WLAN systems (ranging between 20-160 MHz), the approach pursued was geolocation based on time-delay estimation [11]. Though the early releases of the IEEE802.11™ standard included means for time delay estimation, the timing resolution enabled by these mechanisms was in the microseconds level - too coarse for any practical indoor positioning applications. High-accuracy positioning in a dense multipath environment imposed several hardware design changes in the existing WLAN chipsets, in order to increase the timing resolution from the microseconds level to the nanosecond level (or even sub-nanosecond level). The solution that was endorsed by the IEEE802.11™ group, was a novel time-delay based ranging protocol called "fine-timing measurement" (FTM). The FTM protocol enables a WLAN station to measure its

L. Banin, O. Bar-Shalom, N. Dvorecki and Y. Amizur are with Intel's Location Core Division, 94 Em Hamoshavot Rd., Petah Tikva 49527, Israel. Corresponding author's e-mail: oferbarshalom@gmail.com.

distance w.r.t. another station[1]. The range measurement is based on high-resolution, time delay estimation, which also accounts for the latency imposed by the chipset hardware. The hardware-imposed latency (e.g. the receive/transmit filters' group-delay and other hardware latencies), is measured and pre-calibrated by the chipset in order to reach the required timing resolution. Obtaining an accurate time delay estimate in a dense-multipath environment is typically implemented using some super-resolution method , which are applied to the estimated channel response, [9], [10]. FTM is a point-to-point (P2P), single-user protocol, which includes an exchange of multiple message frames between an initiating WLAN station (STA) and a responding STA. The initiating STA attempts to measure its range w.r.t. the responding station (e.g., WLAN AP or a dedicated FTM responder). The FTM message sequence chart is illustrated in Fig. 1. The time of flight (ToF) between

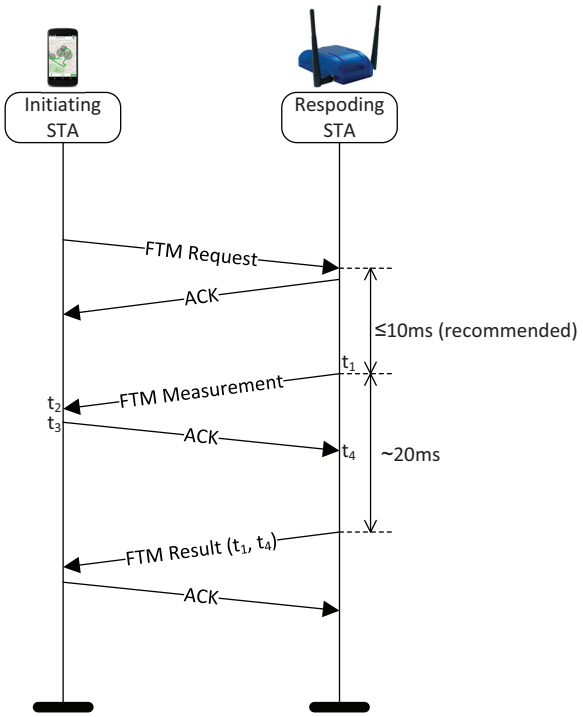

Fig. 1. FTM Protocol Message Flow Example

the two stations is calculated using (1),

$$\text{ToF} = \frac{(t_4 - t_1) - (t_3 - t_2)}{2} \tag{1}$$

where t_1 denotes the time of departure (ToD) measured by responding station, and t_4 denotes the time of arrival (ToA), which is estimated by the responding station. The values of t_1 and t_4 are reported back to the initiating station[2] after the completion of the FTM measurement phase. The initiating station combines these parameters along with its own estimated ToA, t_2, and measured ToD, t_3 values, to obtain a range estimate w.r.t. the responding station[3]. The FTM protocol has first appeared in the 2016 release of the IEEE802.11™ standard [23] (formerly known as IEEE802.11REVmc). Followed the standard release, the Wi-Fi alliance - the organization that promotes Wi-Fi technology - has announced in February 2017 on a Wi-Fi location certification program to certify WLAN devices complying with the FTM protocol.

Being a P2P, single-user protocol, the FTM protocol is limited in scenarios where an extremely large number of users are requesting positioning services simultaneously. Provided no FTM message transactions are lost on the way due to temporal channel interruptions, the initiating station should be able to obtain a range estimate w.r.t. the responding station within about 30ms. Hence, obtaining its position, which involves ranges estimation towards 3 additional stations, should ideally take about 100-120ms. This implies that each FTM responder may be able to serve about 30 client stations per second. Clearly, with more and more navigating stations attempting to execute FTM sessions, the collision likelihood increases, which effectively decreases the number of stations that can be serviced. Consider for example, large stadiums hosting rock concerts or major sports events. In such occasions it is easy to imagine tens of thousands of users navigating throughout the stadium area using location-based services. Servicing all these users might require to deploy a network of thousands of FTM responders around the stadium. The protocol described in the sequel, dubbed "collaborative time of arrival" (CToA), is aimed to provide a more cost-effective solution for such use cases.

Paper Organization: The remainder of the paper is organized as follows. Section II gives an overview of the CToA protocol and its challenges. The mathematical model for the positioning problem considered, is formulated in section III. This section is divided into three parts; the first part, addressed in section III-A, outlines the measurement models and maximum-likelihood position estimators for client-mode CToA in the absence of clock-skewness. These measurement models are then used in section III-B for obtaining measures for the expected positioning accuracy. The third part, which is outlined in section III-C, introduces the effect of the clock-drifts on the measurement models. This section also details the Kalman filter algorithm that is executed by the client device and used for estimating and tracking all the time-varying parameters in the system. System-level simulation performance are described in section IV. In the appendix we derive the Cramér-Rao lower bounds for the positioning problem and the associated concentration ellipses, both of which are used in the main text to illustrate the theoretical system performance discussed in section III-B.

Notation: We use lower-case letters to denote scalars, lower-case, boldface letters to denote vectors, and upper-case, boldface letters to denote matrices. We further

[1]Notice that FTM only enables to measure the range between two stations. Obtaining a position estimate based on multiple range measurements is out of the standard scope. However, the standard does define mechanisms for the responding stations to provide their location information (such as, absolute or relative position coordinates, floor level etc.), in an information element (IE), called location configuration information (LCI). The LCI of the responding stations may be used by the initiating station to estimate its absolute or relative position.

[2]The values of t_1 and t_4 are reported at picosecond granularity.

[3]The exchange of the FTM measurement message and its acknowledgement (ACK) frame, which has to be sent out after exactly a short inter-frame spacing (SIFS) of 16μs, is assumed to finish within a short period, during which the clocks of the two stations do not drift appreciably.

use the following nomenclature throughout the paper.

$\{.\}^T$	transpose
$\mathrm{diag}\{\mathbf{x}\}$	$N \times N$ diagonal matrix, whose diagonal is the vector $\mathbf{x}$
$\mathbf{I}_N$	$N \times N$ identity matrix
$\mathbf{1}_N$	$N \times 1$ vector of 1's
$\mathbf{0}_N$	$N \times 1$ vector of 0's
$\otimes$	Kronecker product
$\|\mathbf{x}\|$	norm of the vector $\mathbf{x}$, i.e. $\sqrt{\mathbf{x}^T\mathbf{x}} = \sqrt{\sum_i x_i^2}$
$\underset{\mathbf{x}}{\mathrm{argmin}}\|\mathbf{y}(\mathbf{x})\|$	search for the value of $\mathbf{x}$ that minimizes the norm of $\mathbf{y}(\mathbf{x})$
$n \sim \mathcal{N}(\mu, \sigma)$	Gaussian-distributed noise with mean, μ, and standard-deviation, σ.
$E\{.\}$	Expectation operator
c	The speed of light, $c = 2.99792458 \cdot 10^8 \mathrm{m/s}$.

II. Scaling-up the FTM Protocol

A. CToA Overview

CToA is a geolocation protocol designed for scaling up the number of clients that could be serviced simultaneously. This can be achieved through the use of broadcast approach rather than a P2P or a point to multi-point ranging approach. The protocol operates over an un-managed network of *unsynchronized* and independent units called "CToA broadcasting stations" (bSTA), which together form a high-precision, geolocation network. The bSTAs, which are deployed at known locations, are implemented using either standard WLAN APs that have the ability to measure accurate ToA or network-detached, FTM-responders.

According to the CToA protocol, the bSTA units serve several purposes. Every bSTA:

- Periodically broadcasts a CToA "beacon" and measures the ToD of that beacon.
- Listens for CToA beacons broadcast by its neighbor bSTAs, and measures their ToA.
- Maintains a log of its current ToD and ToA measurements, and publishes its most recent measurements log as part of its next CToA beacon broadcast.
- Periodically announces its location as part of its CToA beacon broadcasts.

The CToA protocol supports two modes: a client-mode and a network-mode. While both modes rely on the same protocol principles, they are targeted towards different usage models:

- **Client-Mode CToA** - may be visioned as the indoor counterpart of the global navigation satellite systems (GNSS). It is designed for enabling an *unlimited* number of clients to estimate their location and navigate, simultaneously, while maintaining their privacy. The CToA client stations (cSTAs) only listen to the bSTA broadcasts. Once a cSTA receives a broadcast, it measures its ToA and combines it with the ToD/ToA measurements log published by the bSTA in the CToA beacons, in order to determine its position. Since the cSTAs do not transmit, their presence is not exposed and their privacy is maintained.
- **Network-Mode CToA** - designed to enable a network administrator to simultaneously track the position of a large number of clients. This mode is useful for large scale asset tracking (e.g., using eTags[4]), fleet management, law-enforcement, etc. CToA clients in operating in the network-mode do not listen for CToA beacons, but only transmit CToA beacons (at rather low rate), in order to enable the network administrator to track their position. The sporadic, short transmissions executed by such devices enables them to operate for long periods using small, coin-cell batteries.

As bSTAs activity in the two modes is identical, the CToA network can support both modes simultaneously.

Since the protocol is based on broadcasting of CToA beacons, it is uni-directional; the CToA beacons are unacknowledged if not received. Yet, the fact that beacons can be received by multiple neighbor bSTAs in the vicinity of the broadcasting bSTA, gives a level of redundancy and immunity against frame losses. Each CToA beacon is associated with a unique packet identification index (PID), which is assigned to it by the broadcasting unit. The PID is typically implemented as a running counter, and is independently maintained by every bSTA. Each CToA beacon has a ToD time-stamp (measured by the broadcasting unit), and multiple ToA measurements - all of which are associated with the same PID. The PID enables the CToA clients (operating in "client-mode"), or the positioning server (in "network-mode") to associate between the ToD and its corresponding ToA measurements, collected either by the client itself or different bSTAs.

In addition to the ToD, some of the CToA beacons broadcast by every bSTA include also a data log of the timing measurements collected by the bSTA during the past n-seconds. This data log is called "CToA location measurement report" (CLMR). The timing measurements included in the CLMR reports are used by the cSTA, which combines them with its own ToA measurements in order to estimate its position. Although the CLMR logs are maintained by each of the bSTAs independently, the protocol also enables the CLMR logs broadcast by one bSTA to be aggregated by its neighbors, thereby providing an immunity mechanism against "hidden-nodes" in the wireless network.

Figures 2-3 illustrate an example of CToA beacon broadcast and its reception. The example assumes a CToA network, which consists of 3 bSTA units and a single cSTA. These units are assigned with (simplified) medium-access control (MAC) addresses: 10:01, 10:02 and 10:03, while the cSTA has a MAC address of 55:55. As depicted in Fig. 2, at some point indicated by ToD time-stamp of "199678" (measured in picoseconds and referenced to the time-base of bSTA#1), bSTA#1 broadcasts a CToA beacon associated with PID "1551". The ToD and the PID are logged in the CLMR log maintained by bSTA#1. This log also includes the MAC address of the broadcasting bSTA. The beacon propagates, and as illustrated in Fig. 3, it is received by the neighbor bSTAs (#2 & #3), and by the cSTA

[4] An "e-Tag" refers to an electronic tag - a small, wireless-enabled device that is attached to a larger object, and enables to remotely monitor various measures related to the object including its location. In the context of this paper the e-Tag is assumed to be Wi-Fi enabled.

- all of which update their CLMR logs accordingly: bSTA#2 measured the ToA of the beacon to be “329673” (referenced to its own time-base) and updates that value in its CLMR log, along with it MAC address as the receiving unit. Similarly do bSTA#3 and the cSTA, which estimate the ToAs of that same beacon to be “341006”, and “133564”, respectively.

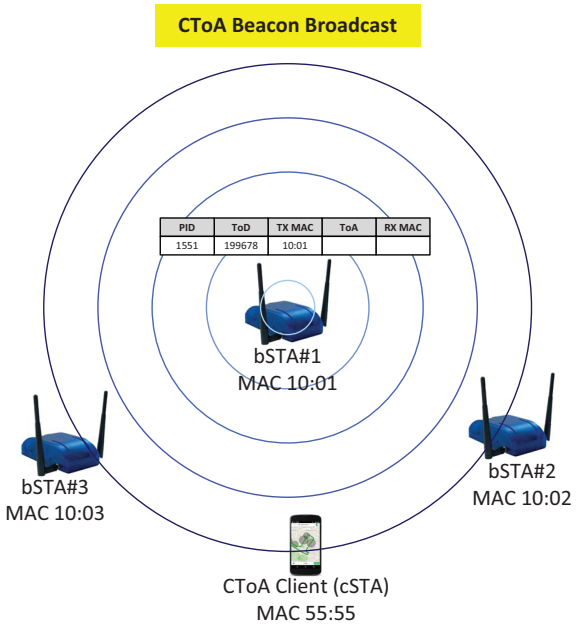

Fig. 2. CToA Beacon Broadcast Example

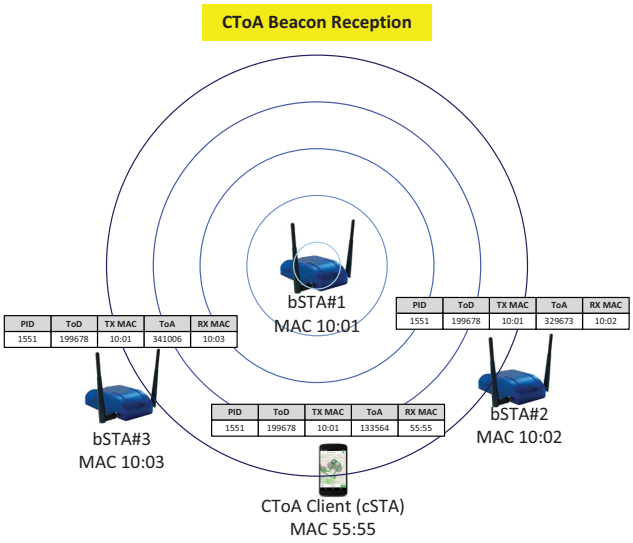

Fig. 3. CToA Beacon Reception Example

The CLMR logged by each of the bSTAs will be broadcast as part of their next CToA beacon broadcast event, and the CLMR logs will be updated accordingly. As will be explained in sectionIII-C, the entries of the CLMR log table will undergo a time-stamp matching step, in which the ToA measurements will be matched with their corresponding ToD value. This information, combined with the position information of the bSTAs, (which is also provided as part of their CToA beacon broadcasts), will be fed into a Kalman filter that will produce an updated estimate of the cSTA’s current position.

B. The CToA Protocol

The CToA protocol follows the principles of the channel sounding mechanism introduced for IEEE802.11ac standard (a.k.a very high-throughput/VHT) [23]. The channel sounding protocol, which was originally proposed for determining the optimal beamforming weights at the transmitter side [17], relies on a transmission of a null-data packet (NDP), which consists of only a known sequence of OFDM symbols, but with no data payload. The transmission of the NDP is preceded by a packet called NDP announcement (NDPA), which informs the receivers of an NDP frame that is about to be transmitted after a standard, short inter-frame spacing/SIFS of 16μs from the end of the NDPA packet. The NDPA contains information for the receiver for estimating the channel response.

CToA relies on a similar protocol structure; the CToA beacons broadcast by the bSTAs consist of an NDPA frame announcing the upcoming transmission of the NDP frame, which is used by the receivers for measuring its ToA. As mentioned above, the NDPA also includes data that enables the receiving clients to estimate their location. The protocol messaging sequence for the client-mode is illustrated in Fig. 4. As shown, cSTAs in this mode only receive, but do not transmit any data, maintains their privacy. The protocol messaging

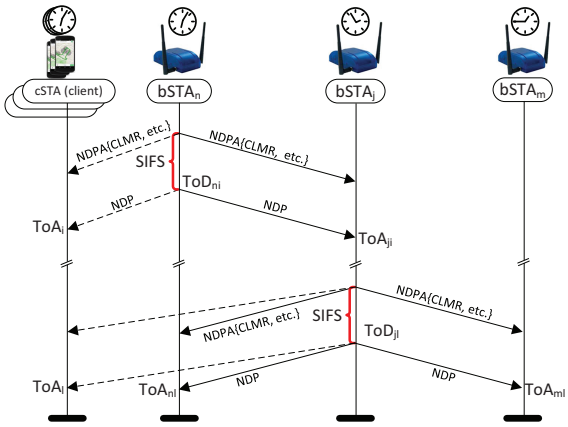

Fig. 4. The CToA Protocol - Client-Centric Operation

sequence for the network-mode CToA is illustrated in Fig. 5.

C. The Un-managed Operation of the CToA Network

Besides using unsynchronized broadcasting units, the CToA network is also un-managed in the sense no that coordination or scheduling protocol between the bSTAs (or the receiving clients) is required for its operation. Ideally, if the bSTAs are implemented as dedicated units, the CToA network could be allocated with a unique operation channel (with bandwidth of 20MHz, 40MHz or 80MHz, depending on the spectrum band in which the system operates). Any bSTA can first scan the spectrum to detect prior CToA broadcast activity, and once detected - the bSTA can contend on accessing that channel

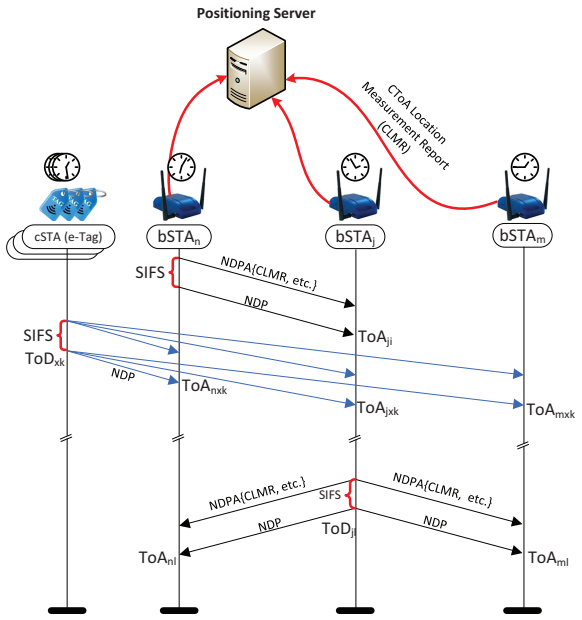

Fig. 5. The CToA Protocol - Network-Centric Operation

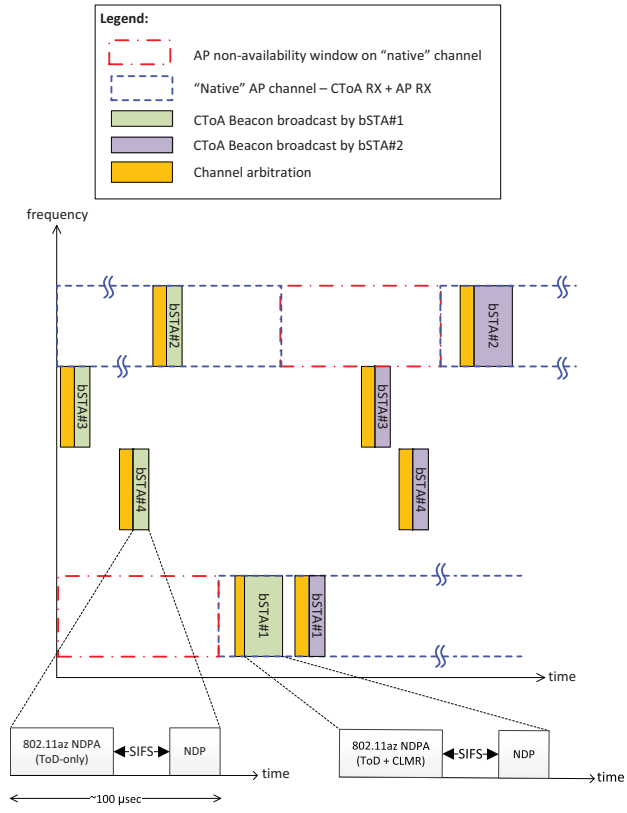

Fig. 6. bSTA Frequency Management with AP-based bSTAs

for broadcasting its beacons, and listen to broadcasts of its neighbor during the remaining time.

A more challenging case is when the bSTA functionality is implemented as part of a standard Wi-Fi AP. In such case, the AP is obligated to provide data transaction services to its associated STAs. To ensure high-data throughput and capacity, the Wi-Fi network typically consists of a grid of APs, where each of them is set to operate at a different channel (aka, its "native channel"). To enable an un-managed CToA network to coexist with a live Wi-Fi data network, the scheme illustrated in Fig. 6 may be used. According to this scheme, the AP/bSTA periodically (e.g., once every few minutes) scans the spectrum to detect CToA broadcast activity. Following the scan results, the AP/bSTA announces to its associated STAs, on a short "un-availability window" (such window should typically last 1 millisecond or less). During that window, the bSTA/AP hops between the native channels of its neighbor bSTA/APs. On each of these channels it broadcasts a short CToA beacon, which only includes its ToD but lacks the CLMR. When the AP returns to its native channel, it broadcasts a longer CToA beacon that includes both its ToD and its recent CLMR.

This process is depicted in Fig. 6. This example illustrates the frequency hopping process of two AP-based bSTAs (#1 and #2), in a CToA network that consists of a total of 4 AP-based bSTAs. As shown, while bSTA#1 goes off of its native channel, it hops between the channels occupied by AP/bSTAs #2,#3 and #4, and on each channel it broadcasts a short beacon. The broadcasting process, including the standard WLAN channel arbitration, lasts about 200μs while the beacon itself, (which consists of NDPA-SIFS-NDP) consumes about 100μs. When bSTA#1 returns to its native channel, it broadcasts a longer beacon that includes also the CLMR log. A similar process is executed by bSTA#2.

By scanning the medium, the CToA clients can detect the CToA activity and infer the broadcast periodicity of the bSTAs. Once the cSTA figures out this information, it simply needs to hop between the native channels used by the bSTA/APs, and collect the CLMRs broadcast on each of these channels. Under the assumption of 5Hz beacon broadcasting rate, the unavailability of the AP for its associated STAs is about 0.5% or less. Accordingly, within 200 milliseconds the client should be able to gather sufficient information for estimating its position.

D. The CToA Time-Tracking Challenge

The general approach described above, of synchronizing a network using time-stamped broadcast transmissions, is well-known [20]. However, whilst most network-based applications (e.g., audio or video distribution) would be satisfied with microsecond or even millisecond-level synchronization, accurate geolocation applications require sub nanosecond-level accuracy. Attaining such a high timing accuracy in a network built out of unsynchronized, and independent broadcasting units, which rely on low accuracy oscillators as their clock source, is extremely challenging.

To understand the challenge, let us compare the problem of client geolocation in a CToA network to a client geolocation in a GNSS network. GNSS networks implement a similar time-stamped broadcasts approach for enabling an unlimited number of client receivers to navigate simultaneously, worldwide. Yet, there are two fundamental differences between the time-tracking of receivers in a GNSS network compared to CToA network.

1) The GNSS network is fully synchronized, whereas the CToA network is not.

2) The broadcasts in a GNSS network are received simultaneously, while in a CToA network the broadcasts are staggered in time.

Let's delve into these two differences and explore them in a bit more details. In GNSS networks the satellite vehicles (SV) are synchronized using on-board atomic clocks, which have a frequency stability of approx. 10^{-14}. This frequency stability translates into a clock-drift of roughly 1 nanosecond per day [21], (which is equivalent to a ranging error of about 35cm - an error, which is further corrected by the GNSS system). Since GNSS networks are fully-synchronized, in terms of timing parameters, the GNSS client receiver needs to estimate only the offset and the drift between its internal clock, and the GNSS network clock. The GNSS client receiver's clock is typically generated using a crystal oscillator/XO with a frequency stability in the order of 10^{-6}, which is commonly expressed in units of parts-per-million/ppm). Tracking these parameters (along with additional system states such as position and velocity), is done using a Kalman filter algorithm [22].

In the CToA network, since the bSTAs are unsynchronized, each bSTA contributes a clock-offset and drift that need to be estimated and tracked. Furthermore, the different medium-access control (MAC) methods used by GNSS and CToA impose an additional challenge. In GNSS networks, the multiplexing at the code space (CDMA) or the frequency space (FDMA), ensures that broadcast transmissions from all SVs are received simultaneously at the client. On the other hand, CToA relies on the "listen-before-talk" MAC of the IEEE802.11™, which effectively results in timing measurements being staggered in time. Given that a typical Wi-Fi XO has an accuracy of $\pm$25ppm, consecutive timing measurements taken of the same broadcasting source may accumulate significant time-drift [14]. This effectively means that while one bSTA clock offset is being measured, the other bSTAs clock offsets keep on drifting apart. As an example, consider two bSTA broadcast timing measurements taken by a static receiver, while the beacons are being broadcast at rate of 5Hz. The drift of the second ToD time-stamp accumulates to up to 5μs (w.r.t. its nominal value). This effectively translates into a ranging error of: $\approx 5\mu\text{s} \cdot 3 \cdot 10^8 \text{m/s} = 1500\text{m}$! Clearly, the clock skewness poses a major challenge for the receiver, which requires the application of filtering techniques for tracking these changes over time. As will be described in the sequel, the CToA client uses a Kalman filter for tracking the various timing-related parameters as well as it own location. Assuming the rate at which the clock offsets vary in time is slow enough, a beacon broadcasting rate of 3-5Hz by each bSTA is sufficient for the cSTAs to accurately track the clock behavior of the bSTAs.

III. CToA Problem Formulation & Positioning Algorithms

In the following section, the mathematical background of the CToA client position estimation is established. To facilitate the explanation we split the derivation into two parts; first, in section III-A, we address the position estimation problem under the idealistic assumption that the bSTA clocks do not drift over time, such that their offsets w.r.t. to the client's clock are time-invariant. Under this assumption we derive the position estimators for two cases:

- "1st Fix" - corresponds to the scenario, at which the client first attempts to sync and estimate its position.
- "Tracking" - corresponds to the scenario where the client already has an estimate of its position and the bSTA timing-related parameters, and continues tracking them using a Kalman filter.

These measurement models are used for developing approximate performance bounds that can predict the expected positioning accuracy. Then, in section III-C, we define the Kalman filter that enables the client to simultaneously estimate and track its own location coordinates, as well as the clocks parameters the bSTAs, (both of which it receives directly, as well as of bSTAs received indirectly via other bSTA).

A. CToA MLE Solution in the Absence of Clock Drift

We shall now derive the maximum-likelihood estimates (MLE) of the client position under the assumption that the clocks of the client and the bSTA do not drift over time, so that the clock offsets are fixed. For simplicity, the derivation assumes a horizontal position only, (which is typically of most interest in indoor-positioning scenarios). The extension to 3-dimensional positioning is straightforward.

We define a "measurement" as the time-of-flight (ToF) of a broadcast transmission between two endpoints, A and B. The transmitting endpoint, A, measures the broadcast's time of departure (ToD), while the receiving endpoint, B, measures its time of arrival (ToA). Both timing measurements are referenced to a 3rd party clock, and thus have offsets marked by ν_A and ν_B, respectively.

$$z \triangleq \text{ToF}_{AB} = (\text{ToA}_B + \nu_B) - (\text{ToD}_A + \nu_A) \tag{2}$$

By denoting the coordinates vectors of the endpoints as $\mathbf{q}_A$ and $\mathbf{q}_B$, respectively, and ignoring any non-line of sight (NLoS) timing biases, the ToF between the two endpoints may be expressed as,

$$\text{ToF}_{AB} \cong \frac{1}{c}\|\mathbf{q}_B - \mathbf{q}_A\| \equiv \text{ToA}_B - \text{ToD}_A \tag{3}$$

Combining (2) with (3) we obtain the definition of a noiseless ToF measurement as,

$$\tilde{z} \triangleq \frac{1}{c}\|\mathbf{q}_B - \mathbf{q}_A\| + \nu_B - \nu_A \tag{4}$$

If the 3rd party also acts as the receiving endpoint, then $\nu_B = 0$ and the noiseless ToF measurement is defined as,

$$\bar{z} \triangleq \frac{1}{c}\|\mathbf{q}_B - \mathbf{q}_A\| - \nu_A \tag{5}$$

1) MLE Solution for CToA Client's First Fix: Assume that a single CToA client station (cSTA), located at: $\mathbf{p} = [x, y]^T$, attempts to estimate its position using time-delay measurements it gathers from M CToA bSTAs, whose locations are known to the cSTA, where the mth bSTA is located at $\mathbf{q}_m = [x_m, y_m]^T$.

The cSTA collects two types of time delay measurements:

- bSTA$_i$ $\rightarrow$ bSTA$_j$ measurements, where $i,j \in 1\ldots M, \forall i \neq j$. These time-delays are measured by the bSTAs and published in their CToA beacon broadcast. The cSTA collects L measurements of this type, where the lth measurement is denoted by $\tilde{z}_l$. Each measurement is subjected to additive measurement error denoted by $\tilde{n}_l \sim \mathcal{N}(0,\tilde{\sigma})$.
- bSTA$_i$ $\rightarrow$ cSTA, where $i \in 1\ldots M$. These time-delays are measured by the cSTA itself, and the cSTA collects $\mathcal{L}$ measurements of this type, where the ℓth measurement is denoted by $\bar{z}_\ell$. Each measurement is subjected to additive measurement error denoted by $\bar{n}_\ell \sim \mathcal{N}(0,\bar{\sigma})$. Typically, $\bar{\sigma} > \tilde{\sigma}$.

Let ν_i denote the (unknown) offset between the cSTA & bSTA$_i$ clocks. A single bSTA$_i$ $\rightarrow$ cSTA, ToA measurement of the ℓth broadcast made by the ith bSTA, may be modeled as,

$$\bar{z}_\ell = \frac{1}{c}\|\mathbf{p}-\mathbf{q}_i\| - \nu_i + \bar{n}_\ell,\ \ell = 1,\ldots,\mathcal{L} \tag{6}$$

Similarly, the measurement of the lth broadcast that is received by the jth bSTA, may be modeled as,

$$\begin{aligned}\tilde{z}_l &= \mathrm{ToF}_{ij} - \nu_i + \nu_j + \tilde{n}_l \\ &= \mathrm{ToA}_j - \mathrm{ToD}_i - \nu_i + \nu_j + \tilde{n}_l \\ &= \frac{1}{c}\|\mathbf{q}_j-\mathbf{q}_i\| + \beta_{ij} - \nu_i + \nu_j + \tilde{n}_l,\ l = 1,\ldots,L\end{aligned} \tag{7}$$

Notice that the ToF (scaled by the propagation speed, c), may represent a biased version of the range between the two bSTAs. This may happen due to some obstruction in the propagation medium, resulting in a non-line of sight (NLoS) link between the two endpoints. The scalar $\beta_{ij} > 0$ represents the NLoS ranging bias between the ith and jth bSTAs. However, since the position of the bSTAs is assumed to be known *a* priori, β_{ij} can be easily estimated and eliminated. For the sake of derivation clarity, it shall be further assumed that $\beta_{ij} = 0, \forall i,j$. An example for timing measurements collected under the assumptions that the bSTA clocks are stable and do not drift over time is illustrated in Fig. 7.

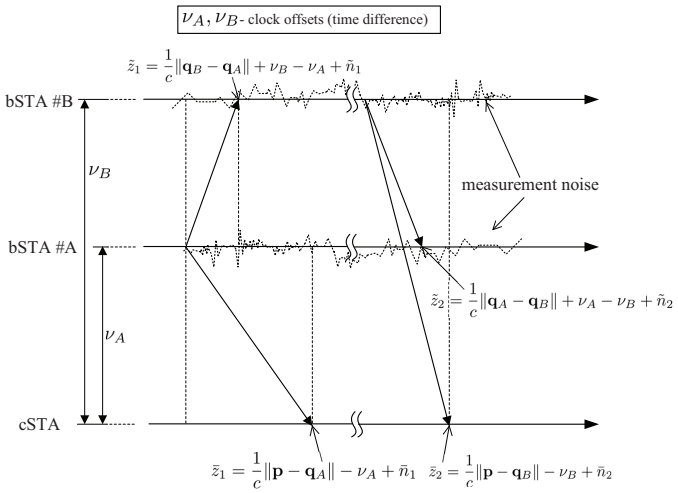

Fig. 7. Timing Measurements with Stable Clock

Let $\mathbf{e}_i$ denote an $M \times 1$ vector of zeros, whose ith entry is 1. Using this notation,

$$\nu_i = \mathbf{e}_i^T\boldsymbol{\nu} \tag{8}$$

$$\nu_j - \nu_i = (\mathbf{e}_j^T - \mathbf{e}_i^T)\boldsymbol{\nu} \tag{9}$$

where, $\boldsymbol{\nu} \triangleq [\nu_1,\ldots,\nu_M]^T$. Now, the timing measurements may be recast as,

$$\bar{z}_\ell = \frac{1}{c}\|\mathbf{p}-\mathbf{q}_i\| - \mathbf{e}_i^T\boldsymbol{\nu} + \bar{n}_\ell \tag{10}$$

$$\begin{aligned}\tilde{z}_l &= \mathrm{ToF}_{ij} + (\mathbf{e}_j^T - \mathbf{e}_i^T)\boldsymbol{\nu} + \tilde{n}_l, \\ &\approx \frac{1}{c}\|\mathbf{q}_j-\mathbf{q}_i\| + (\mathbf{e}_j^T - \mathbf{e}_i^T)\boldsymbol{\nu} + \tilde{n}_l\end{aligned} \tag{11}$$

We may further define the following vectors and matrices,

$$\begin{aligned}\bar{\mathbf{z}} &\triangleq [\bar{z}_1,\ldots,\bar{z}_\mathcal{L}]^T, \\ \tilde{\mathbf{z}} &\triangleq [\tilde{z}_1,\ldots,\tilde{z}_L]^T, \\ \mathbf{z} &\triangleq \begin{bmatrix}\bar{\mathbf{z}}\\ \tilde{\mathbf{z}}\end{bmatrix} \\ \bar{d}_\ell(\mathbf{p}) &\triangleq \|\mathbf{p}-\mathbf{q}_i\| \\ \tilde{d}_l &\triangleq \|\mathbf{q}_i-\mathbf{q}_j\| \\ \bar{\mathbf{d}}(\mathbf{p}) &\triangleq [\bar{d}_1(\mathbf{p}),\ldots,\bar{d}_\mathcal{L}(\mathbf{p})]^T \\ \tilde{\mathbf{d}} &\triangleq [\tilde{d}_1,\ldots,\tilde{d}_L]^T \\ \mathbf{d}(\mathbf{p}) &\triangleq \begin{bmatrix}\bar{\mathbf{d}}\\ \tilde{\mathbf{d}}\end{bmatrix} \\ \bar{\mathbf{E}} &\triangleq [-\mathbf{e}_{i,1},\ldots,-\mathbf{e}_{i,\mathcal{L}}]^T \\ \tilde{\mathbf{E}} &\triangleq [(\mathbf{e}_{j,1}-\mathbf{e}_{i,1}),\ldots,(\mathbf{e}_{j,L}-\mathbf{e}_{i,L})]^T \\ \mathbf{E} &\triangleq \begin{bmatrix}\bar{\mathbf{E}}\\ \tilde{\mathbf{E}}\end{bmatrix} \\ \bar{\mathbf{n}} &\triangleq [\bar{n}_1,\ldots,\bar{n}_\mathcal{L}]^T \\ \tilde{\mathbf{n}} &\triangleq [\tilde{n}_1,\ldots,\tilde{n}_L]^T \\ \mathbf{n} &\triangleq \begin{bmatrix}\bar{\mathbf{n}}\\ \tilde{\mathbf{n}}\end{bmatrix}\end{aligned} \tag{12}$$

Using the definitions of (12), we may recast (10)-(11) as,

$$\mathbf{z} = c^{-1}\mathbf{d}(\mathbf{p}) + \mathbf{E}\boldsymbol{\nu} + \mathbf{n} \tag{13}$$

Assuming that the measurement noise is Gaussian-distributed with the mean and covariance as follows,

$$\begin{aligned}E\{\mathbf{n}\} &= \mathbf{0} \\ E\{\mathbf{n}\mathbf{n}^T\} &= \begin{bmatrix}\bar{\sigma}^2\mathbf{I}_\mathcal{L} & \mathbf{0}\\ \mathbf{0} & \tilde{\sigma}^2\mathbf{I}_L\end{bmatrix} \triangleq \mathbf{W}\end{aligned} \tag{14}$$

Then the maximum likelihood estimate (MLE) of the cSTA position vector, $\hat{\mathbf{p}}$, may be obtained as,

$$\hat{\mathbf{p}} = \underset{\mathbf{p},\boldsymbol{\nu}}{\mathrm{argmin}}(\mathbf{z} - c^{-1}\mathbf{d} - \mathbf{E}\boldsymbol{\nu})^T\mathbf{W}^{-1}(\mathbf{z} - c^{-1}\mathbf{d} - \mathbf{E}\boldsymbol{\nu}) \tag{15}$$

The estimate of the clock offsets vector may be found using weighted least-squares (WLS) criteria,

$$\hat{\boldsymbol{\nu}} = (\mathbf{E}^T\mathbf{W}^{-1}\mathbf{E})^{-1}\mathbf{E}^T\mathbf{W}^{-1}(\mathbf{z} - c^{-1}\mathbf{d}) \tag{16}$$

Define,

$$\mathbf{B} \triangleq [\mathbf{W}^{-1} - \mathbf{W}^{-1}\mathbf{E}(\mathbf{E}^T\mathbf{W}^{-1}\mathbf{E})^{-1}\mathbf{E}^T\mathbf{W}^{-1}] \tag{17}$$

Then, by substituting (16) back in (15) we get,

$$\hat{\mathbf{p}} = \underset{\mathbf{p}}{\operatorname{argmin}} (\mathbf{z} - c^{-1}\mathbf{d})^T \mathbf{B} (\mathbf{z} - c^{-1}\mathbf{d}) \tag{18}$$

The nonlinear minimization problem in (18) can be solved via 2-dimensional grid search (or 3-dimensional search, in case of 3-positioning), over all the possible locations.

2) MLE Solution for a CToA Client in "Tracking" Mode: Once the CToA client receiver has converged to the true values of the bSTA clock offsets and continuously tracks them, these clock offsets may be considered as "known" (up to some estimation error). In such case it would be reasonable to assume that $\mathbf{z} \simeq \bar{\mathbf{z}}$, $\mathbf{d} \simeq \bar{\mathbf{d}}$. Define,

$$\boldsymbol{\zeta} \triangleq \bar{\mathbf{z}} - \bar{\mathbf{E}}\boldsymbol{\nu} \tag{19}$$

The resulting measurement model in this case may be recast as,

$$\boldsymbol{\zeta} = c^{-1}\bar{\mathbf{d}}(\mathbf{p}) + \check{\mathbf{n}} \tag{20}$$

The additive noise vector, $\check{\mathbf{n}}$, is assumed to be Gaussian distributed with the following properties:

$$\begin{aligned} E\{\check{\mathbf{n}}\} &= \mathbf{0} \\ E\{\check{\mathbf{n}}\check{\mathbf{n}}^T\} &\triangleq \sqrt{\bar{\sigma}^2 + \sigma_r^2} \cdot \mathbf{I}_{\mathcal{L}} \end{aligned} \tag{21}$$

where σ_r corresponds to the standard deviation residual estimation error of the clock offsets. The client position MLE in this case is obtained by minimizing the following cost function

$$\hat{\mathbf{p}} = \underset{\mathbf{p}}{\operatorname{argmin}} \left\| \boldsymbol{\zeta} - c^{-1}\bar{\mathbf{d}}(\mathbf{p}) \right\|^2 \tag{22}$$

where again, the nonlinear minimization problem in (22) can be solved via grid search over the position coordinates space.

B. Approximate CToA Performance Analysis

In order to obtain theoretical performance bounds, the skew-less measurement models derived in section III-A were used for calculating the respective Cramér-Rao lower bounds (CRLB). The derived bounds are affected mainly on the geometrical properties of the network deployment, as well as the additive noise levels. Yet, these bounds ignore propagation models which may take into account the type of materials through which the signals propagate.

To illustrate the expected positioning accuracy, the bounds are depicted as "heat-maps", where the colors are mapped to the location according to the size of the minimal theoretical positioning error predicted by the CRLB. Under the assumption that the additive measurement noise is Gaussian-distributed, for each location on a given grid of locations one can calculate the concentration ellipse that defines the smallest area at which the CToA receiver is contained with a given probability. As explained in Appendix C, the concentration ellipse is tightly related to the position estimation error covariance matrix predicted by the CRLB, which is derived in Appendix A. Fig. 8- 9 describe the geometric-dependent accuracy of CToA in a typical office network deployment, where the magenta rings denote the position of the bSTAs. The color mapping corresponds to the size of the major axis of the 95% percentile concentration ellipse at that position (namely, 2σ), as the measure of accuracy. The bound is calculated using,

$$\text{Bound @ } 95\% = \sqrt{\kappa \cdot \lambda_{max}} \tag{23}$$

where κ is calculated using (77) and λ_{max} is evaluated for the measurement models (13).

The bound was evaluated over a pre-defined grid covering the office floor with a resolution of 0.5m×0.5m. The theoretical error was computed under the assumption is that at each grid point the CToA client receives timing measurements from only the 4 nearest bSTAs. Fig. 8 illustrates the expected accuracy for a "1st-Fix" scenario, while Fig. 9 corresponds to the "tracking" scenario. The measurement noise standard deviations assumed in the "1st-Fix" scenario were $\bar{\sigma}$ =1.5m, and $\tilde{\sigma}$ =0.6m. For the "tracking" scenario the assumed residual error was σ_r =0.3m.

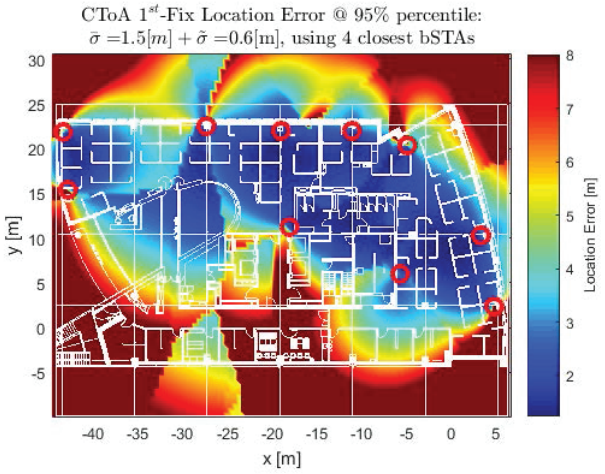

Fig. 8. Accuracy of CToA at "1st-Fix" Mode in a Typical Office Environment

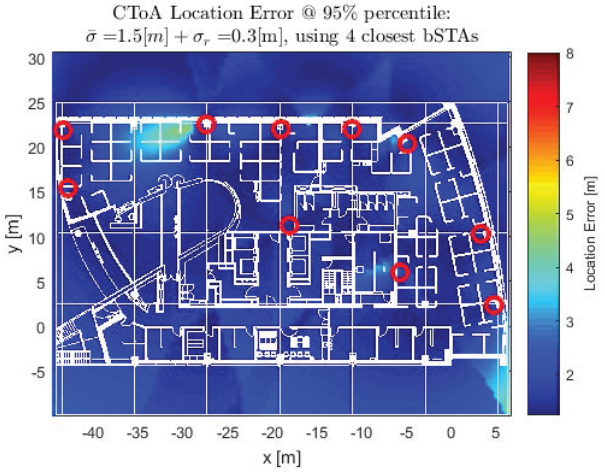

Fig. 9. Accuracy of CToA at Tracking Mode in a Typical Office Environment

As can be seen, once the CToA client has its EKF converged, the positioning error over the entire area drops significantly compared to the "1st-Fix" scenario. It can be further be

shown that positioning accuracy of the "1st-Fix" resembles to the accuracy of a hyperbolic, time-difference of arrival (TDoA) based, positioning system. When the client converges its time tracking, the performance is similar to the performance that can be achieved using ToA-based positioning system (namely, an FTM-based system in which the client estimates ranges (or round-trip time/RTT) to individual bSTAs). The following proposition proves that the asymptotic accuracy of the client-based CToA system is equivalent to "tracking" mode and hence to the achievable RTT accuracy.

Proposition 1 (Asymptotic CToA Performance): The asymptotic positioning accuracy for a CToA client in "1st-Fix" mode, approaches the accuracy attained in "tracking" mode, given $N \to \infty$ replicas of the bSTA→bSTA timing measurements.

Proof: See Appendix B. ∎

C. Coping with Clock Drifts Using Kalman Filtering

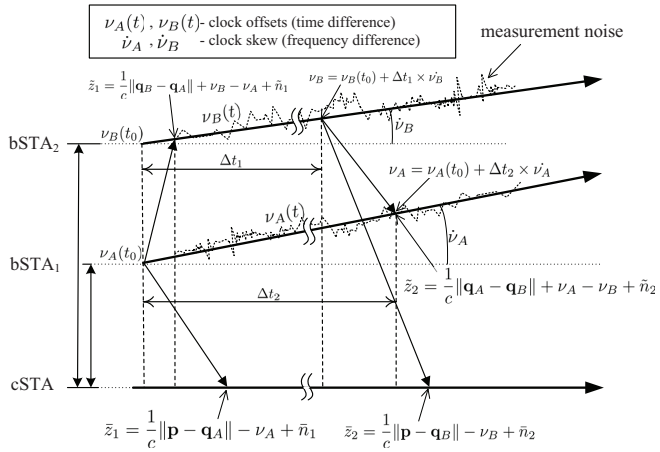

Fig. 10. Timing Measurements with Skewed Clocks

The analysis outlined in the former section ignored any clock skews, which result from the XO's frequency deviations. Such deviations may be caused due to multiple effects such as: ambient temperature changes, phase noise, thermal noise, aging, and so on. To illustrate the timing measurements collected under clock skews, consider the example depicted in Fig. 10. In this example there is a single client station (cSTA) and two bSTAs marked #A and #B. Each of the bSTAs has an initial timing offset w.r.t. the cSTA clock, which is denoted by $\nu_A(t_0)$ and $\nu_B(t_0)$, respectively. In addition, the bSTAs XOs' output frequencies are skewed, such that the clock offsets drift over time at rates, which are denoted by $\dot{\nu}_A$ and $\dot{\nu}_B$, respectively. To facilitate the explanation, the assumption in this illustration is that the skew rate is time-invariant. Yet, in practice it may change over time (e.g., due to ambient effects). Each bSTA measures the broadcast events timing (ToD or ToA) relative to its native time base. Once the measurements are conveyed to the cSTA for enabling it to compute its location, then the cSTA needs to resolve the instantaneous clock offset of the bSTA, which is associated with that measurements and is a function of the offset drift rate. Under the assumed 1st-order clock skew model, the instantaneous value of the clock-offset of the nth bSTA is calculated using,

$$\nu_n(t_i) = \nu_n(t_{i-1}) + \dot{\nu}_n \times \Delta t \tag{24}$$

where $\nu_n(t_{i-1})$ corresponds to the previous estimated value of the clock offset, (whose $\nu_n(t_0)$, is its initial value), $\dot{\nu}_n$ denotes the clock-skew (or the change rate of the clock offset), and $\Delta t = t_i - t_{i-1}$.

In the following section we outline the algorithm, which enables the client station to estimate and track its location. The Kalman Filter is the optimal estimate for linear system models with additive independent white noise in both the transition and the measurement system models. Yet, in many systems, including navigation systems, the measurement model is not linearly dependent in the parameters of interest. In such cases, the extended Kalman filter (EKF), which is the nonlinear version of the Kalman filter, is widely used [22], [24]. In the EKF, the state transition and observation models are not required to be linear functions of the states, but instead, may be only differentiable functions. In the client-mode CToA, the EKF is executed by the client and is used by the client to estimate and track its own position coordinates, as well as the timing parameters of the stray bSTA units, from which it receives the measurement broadcasts. In the network-mode CToA, the EKF is executed at a centric positioning server, connected to the CToA network, and is used for tracking the position of multiple clients simultaneously, as well as tracking the timing parameters of all the network bSTAs. Consequently, the system states tracked by the EKF in each of the modes are different; the network-mode EKF needs to track position *per client*, as well as timing parameters of all network bSTA. In the following section we focus on the client-mode CToA EKF.

1) CToA EKF System Model: The EKF is described by two equations: a system model equation and an observation (measurement) model with additive noise. The system model is defined by the following recursive equation,

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{w}_k, \quad k \geq 0 \tag{25}$$

where the index k denotes the discrete time-step. The vector $\mathbf{x}_k$ denotes an $N \times 1$ states vector, which describes the parameters being estimated and tracked by the filter. The states vector for the client-mode CToA consists of the client's position coordinates and per-bSTA clock parameters (clock offset and clock offset change rate (or skew)). This vector is defined as follows. The size of the EKF state vector is thus: $N = 3 + 2M$, where M denotes the number of bSTAs being

received by the cSTA (both directly and indirectly)[5].

$$\begin{aligned}\mathbf{p}_k &\triangleq [x_k, y_k, z_k]^T \\ \boldsymbol{\nu}_k &\triangleq [\nu_{1,k}, \ldots, \nu_{M,k}]^T \\ \dot{\boldsymbol{\nu}}_k &\triangleq [\dot{\nu}_{1,k}, \ldots, \dot{\nu}_{M,k}]^T \\ \mathbf{x}_k &\triangleq \left[\mathbf{p}_k^T, \boldsymbol{\nu}_k^T, \dot{\boldsymbol{\nu}}_k^T\right]^T \end{aligned} \tag{26}$$

The state-vector $\mathbf{x}_k$ is associated with a covariance matrix,

$$\mathbf{P}_k = E\{(\mathbf{x}_k - \bar{\mathbf{x}}_k)(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T\} \tag{27}$$

where $\bar{\mathbf{x}}_k \triangleq E\{\mathbf{x}_k\}$. When the filter is initialized the state-covariance matrix is assumed to be,

$$\mathbf{P}_0 = \begin{bmatrix} \tilde{\mathbf{P}}_{\mathbf{p},0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_{\nu,0}^2 \mathbf{I}_M & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_{\dot{\nu},0}^2 \mathbf{I}_M \end{bmatrix} \tag{28}$$

where $\sigma_{\nu,0}$, $\sigma_{\dot{\nu},0}^2$ denote the initial values for the standard deviations of the clock offsets and drifts, and $\tilde{\mathbf{P}}_{\mathbf{p},0}$ denotes the initial value of states covariance matrix, given by

$$\tilde{\mathbf{P}}_{\mathbf{p},0} \triangleq \begin{bmatrix} \sigma_{x,0}^2 & 0 & 0 \\ 0 & \sigma_{y,0}^2 & 0 \\ 0 & 0 & \sigma_{z,0}^2 \end{bmatrix} \tag{29}$$

The initial values of the standard deviations constructing the initial states covariance matrix are commonly determined empirically.

The dynamic system-model linear transfer function is denoted by $\mathbf{F}_k$, an $N \times N$ block-diagonal matrix defined as follows,

$$\mathbf{F}_k \triangleq \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_M & \Delta t \mathbf{I}_M \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_M \end{bmatrix} \tag{30}$$

where Δt corresponds to the elapsed time between two consecutive discrete time steps.

The vector $\mathbf{w}_k$ denotes a random $N \times 1$ model noise vector, which described the uncertainties in the system model and has the following statistical properties:

$$\begin{aligned} E\{\mathbf{w}_k\} &= \mathbf{0} \\ E\{\mathbf{w}_k \mathbf{w}_k^T\} &= \mathbf{Q}_k \\ E\{\mathbf{w}_k \mathbf{w}_j^T\} &= \mathbf{0},\ \forall k \neq j \\ E\{\mathbf{w}_k \mathbf{x}_k^T\} &= \mathbf{0},\ \forall k \end{aligned} \tag{31}$$

In the CToA EKF system model, the process noise, $\mathbf{w}_k$ is assumed to be distributed as, $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$, where system model noise covariance matrix, $\mathbf{Q}_k$, is a block-diagonal matrix given by,

$$\mathbf{Q}_k = \Delta t \cdot \begin{bmatrix} \mathbf{Q}_{\mathbf{p},k} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tilde{\sigma}_\nu^2 \mathbf{I}_M & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \tilde{\sigma}_{\dot{\nu}}^2 \mathbf{I}_M \end{bmatrix} \tag{32}$$

[5]In a GPS system, since the entire GPS network is synchronized and uses highly-stable atomic clocks, the receiver has to track only its clock offset and drift w.r.t. the GPS system clock. Furthermore, the SVs orbital motion generates a substantial Doppler offset on the GPS carrier frequency, which enables to estimate the GPS receiver 3-dimensional speed. Thus, the EKF state vector in the GPS receiver typically includes a total of 8 states: 3 states for the receiver position, 3 states for the 3-dimensional receiver speed and 2 states for the clock model.

where,

$$\mathbf{Q}_{\mathbf{p},k} \triangleq \begin{bmatrix} \tilde{\sigma}_x^2 & 0 & 0 \\ 0 & \tilde{\sigma}_y^2 & 0 \\ 0 & 0 & \tilde{\sigma}_z^2 \end{bmatrix} \tag{33}$$

In general, the determination of the noise variance values of $\mathbf{Q}_k$, is challenging, and is often resorted to some heuristic methods. Commonly it is assumed that most of the clock deviation is dictated by the clock skew, rather than clock measurement noise. The values of $\{\tilde{\sigma}_x^2,\ \tilde{\sigma}_y^2,\ \tilde{\sigma}_z^2\}$ are determined according to the motion assumptions of the cSTA device (e.g., pedestrian, vehicle/drone etc.).

2) CToA EKF Measurement Model: The measurement model is defined as,

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \tag{34}$$

where $\mathbf{z}_k$ is a $J \times 1$ vector of measurements, in which each entry corresponds to a ToF measurement that follows the definition in (2). The vector $\mathbf{h}(\mathbf{x}) \triangleq [h_1(\mathbf{x}), h_2(\mathbf{x}), \ldots, h_J(\mathbf{x})]^T$, denoting the nonlinear measurement model vector transfer function, and $\mathbf{v}_k$ denotes the additive measurement noise that has the following statistical properties:

$$\begin{aligned} E\{\mathbf{v}_k\} &= \mathbf{0} \\ E\{\mathbf{v}_k \mathbf{v}_k^T\} &= \mathbf{R}_k = \sigma_m^2 \mathbf{I} \\ E\{\mathbf{v}_k \mathbf{v}_j^T\} &= \mathbf{0},\ \forall k \neq j \\ E\{\mathbf{v}_k \mathbf{x}_k^T\} &= \mathbf{0},\ \forall k \\ E\{\mathbf{w}_k \mathbf{v}_j^T\} &= \mathbf{0},\ \forall k, j \end{aligned} \tag{35}$$

As discussed in sec. III-A, there are two types of transfer functions, which depend on the type of the measurement (bSTA$_i$ $\rightarrow$ cSTA or bSTA$_i$ $\rightarrow$ bSTA$_j$). From (10)-(11) it is easy to see that the corresponding measurement transfer functions are given by,

$$h_\ell(\mathbf{x}_k) = \frac{1}{c}\|\mathbf{p}_k - \mathbf{q}_i\| - \mathbf{e}_i^T \boldsymbol{\nu}_k \tag{36}$$

$$h_l(\mathbf{x}_k) = \frac{1}{c}\|\mathbf{q}_j - \mathbf{q}_i\| + (\mathbf{e}_j - \mathbf{e}_i)^T \boldsymbol{\nu}_k \tag{37}$$

Since the measurement transfer function, $\mathbf{h}(\cdot)$ is nonlinear, it cannot be applied to estimate the measurements covariance matrix directly. Instead we linearize $\mathbf{h}(\cdot)$ by replacing it with its first order Taylor series expansion, calculated around $\hat{\mathbf{x}}_{k|k-1}$:

$$\mathbf{h}(\mathbf{x}_k) \cong \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{H}_k \cdot (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \tag{38}$$

where the notation, $\hat{\mathbf{x}}_{n|m}$ represents the estimate of $\mathbf{x}$ at time n given observations up to and including time $m \leq n$. The matrix $\mathbf{H}_k$ denotes the Jacobian of the measurement model function vector $\mathbf{h}(\cdot)$, which is a $J \times N$ matrix defined as,

$$\begin{aligned} \mathbf{H}_k &\triangleq \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_N} \\ \vdots & & \ddots & \vdots \\ \frac{\partial h_J}{\partial x_1} & \frac{\partial h_J}{\partial x_2} & \cdots & \frac{\partial h_J}{\partial x_N} \end{bmatrix} \\ [\mathbf{H}_k]_{ij} &\equiv \left.\frac{\partial h_i}{\partial x_j}\right|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k-1}} \end{aligned} \tag{39}$$

The Jacobian is obtained by calculating the partial derivatives of (36)-(37). Equations (40)-(41) define the corresponding lines of the matrix $\mathbf{H}_k$.

$$[\mathbf{H}_k]_\ell = \left[\frac{(\mathbf{p}_k - \mathbf{q}_n)^T}{c\|\mathbf{p}_k - \mathbf{q}_n\|}, -\mathbf{e}_i^T, \mathbf{0}_M^T\right] \quad (40)$$

$$[\mathbf{H}_k]_l = \left[\mathbf{0}_3^T, (\mathbf{e}_j - \mathbf{e}_i)^T, \mathbf{0}_M^T\right] \quad (41)$$

The EKF is a recursive estimator, in which only the estimated state from the previous time step and the current measurement are required for the computation of the estimate for the current state. The state of the filter is represented by two variables: the vector $\hat{\mathbf{x}}_{k|k}$, which denotes the *a* posteriori state estimate at time k given observations up to and including at time k, and $\mathbf{P}_{k|k}$, the *a* posteriori error covariance matrix. The CToA EKF algorithm is summarized in Algorithm 1.

Initialize
1. Use (16) and (18) for obtaining an estimate for $\hat{\mathbf{p}}_0$ and $\hat{\boldsymbol{\nu}}_0$.
2. Set $\hat{\mathbf{x}}_0 = [\hat{\mathbf{p}}_0^T, \hat{\boldsymbol{\nu}}_0^T, \mathbf{0}_M^T]^T$

Predict
EKF time as well as EKF states, are predicted according to the ToA of the received NDP packet.
Predicted state estimate:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{x}}_{k-1|k-1} \quad (42)$$

Predicted covariance estimate:

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1} \quad (43)$$

Update
The measurements included in the LMR conveyed by the packet are updated according to the new EKF predicted time.
Innovation (measurement residual):

$$\tilde{\mathbf{y}}_k = \mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}) \quad (44)$$

Innovation covariance:

$$\mathbf{S}_k = \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k \quad (45)$$

Near-Optimal Kalman gain:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T\mathbf{S}_k^{-1} \quad (46)$$

Updated state estimate:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\tilde{\mathbf{y}}_k \quad (47)$$

Updated estimate covariance:

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\,\mathbf{P}_{k|k-1} \quad (48)$$

Algorithm 1: CToA Client-Mode EKF Algorithm

IV. CToA Performance in an Indoor Network

In the following section we outline some simulation performance of an indoor CToA network. To analyze the accuracy of the client position estimates, the EKF position estimates were compared against a "ground truth" trajectory. The reference trajectory was generated using a light detection and ranging (LIDAR)-based ground truth tool [13]. The LIDAR system used integrates a 270° laser scanner, which uses a dedicated map and laser measurements to estimate its position. The LIDAR output is a series of position reports generated at a rate of 20Hz with an accuracy of 10-30 cm. The map is obtained in advance by performing a survey of the venue using the LIDAR, during which a structure map is created using a simultaneous localization and mapping (SLAM) algorithm. This is a one-time procedure, and the generated map is then used in subsequent sessions for localization of the device.

Fig. 11 depicts the reference client trajectory (denoted by $\mathbf{p}$ and marked by red dots), and the position estimates (denoted by $\hat{\mathbf{p}}$ and marked by line-connected blue dots). The location of the bSTAs is marked by the red rings. The simulation generated independent clock sources for each of the bSTAs. These clocks were generated with an accuracy of $\pm$10ppm and with Gaussian-distributed clock noise with zero mean and standard deviation of $\pm 10^{-9}$ (1 ppb). The bSTAs were set to broadcast CToA beacons at 2Hz. Each bSTA exchange measurements its 4 nearest neighbor bSTAs. On every point along the trajectory, the CToA client used measurement from the 4 nearest bSTAs. The cumulative distribution function (CDF) of the positioning errors is depicted in Fig. 12. As can be seen, the estimation accuracy is equal or better to 1.5m for 95% of the position estimates along the trajectory.

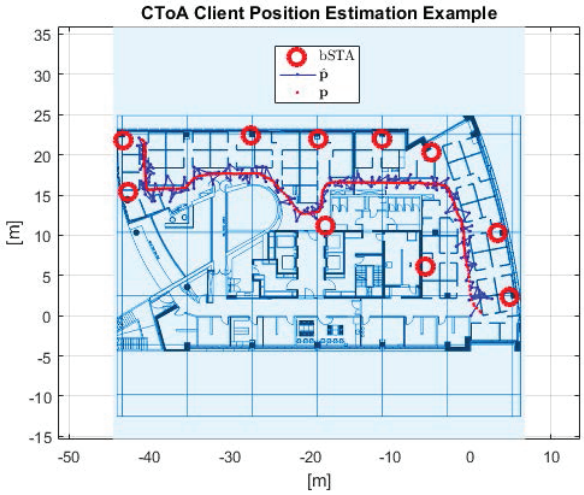

Fig. 11. CToA bSTA Deployment a Typical Office Environment

V. Conclusions

Collaborative time of arrival (CToA), which is the next generation, indoor geolocation protocol was presented. The protocol is designed for enabling scalability of the existing IEEE802.11/Wi-Fi-based, geolocation systems. The protocol uses beacon broadcast-based fine-time delay measurements, collected by both clients and un-managed bSTA units, in order to concurrently enable an unlimited number of clients to privately navigate indoors, and enable a positioning server to track a plethora of clients of a different type (e-Tags).

Fig. 12. CToA Client Positioning Error CDF in a Typical Office Environment

Due to the infrequent nature of the beacon broadcasts, and the skewness of the clocks used by the network units, the estimation of the position, as well as the clock-related parameters is implemented using a Kalman filter, which is executed by every CToA client independently (or by the positioning server, for tracking the position of the CToA e-Tags). System simulations indicate that the network is capable of reaching a positioning accuracy of about 1.5m in a typical office environment in 95% of the cases. These results also match the theoretical performance predicated by the CRLB.

APPENDIX

A. CToA Cramér Rao Lower Bound

We shall now derive the Cramér-Rao lower bound (CRLB) for the CToA method in the absence of clock drifts. The CRLB provides a lower bound on the covariance matrix of any unbiased estimator.

1) CRLB CToA Client in "1st Fix" Mode: Since the observations vector, $\mathbf{m}$, is distributed $\mathbf{m} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{W})$, the ijth entry of the Fisher information matrix (FIM) may be obtained as [29],

$$\mathbf{J}_{ij} = \mathrm{tr}\left\{ \mathbf{W}^{-1}\frac{\partial \mathbf{W}}{\partial \psi_i}\mathbf{W}^{-1}\frac{\partial \mathbf{W}}{\partial \psi_j} \right\} + 2\left\{ \frac{\partial \boldsymbol{\mu}^T}{\partial \psi_i}\mathbf{W}^{-1}\frac{\partial \boldsymbol{\mu}}{\partial \psi_j} \right\} \quad (49)$$

where ψ_i is the ith element of the unknown parameters vector, $\boldsymbol{\psi} \triangleq [\mathbf{p}^T, \boldsymbol{\nu}^T]^T$. Noticing that $\mathbf{W}$ is free of any unknown parameters, then

$$\mathbf{J}_{ij} = 2\left\{ \frac{\partial \boldsymbol{\mu}^T}{\partial \psi_i}\mathbf{W}^{-1}\frac{\partial \boldsymbol{\mu}}{\partial \psi_j} \right\} \quad (50)$$

For, $\boldsymbol{\mu} \triangleq c^{-1}\mathbf{d} + \mathbf{E}\boldsymbol{\nu}$, the partial derivatives w.r.t. the client's position coordinates are given by,

$$\begin{aligned} \frac{\partial \boldsymbol{\mu}}{\partial x} &= c^{-1}\dot{\mathbf{d}}_x \equiv c^{-1}[\dot{\bar{\mathbf{d}}}_x^T \mathbf{0}_L^T]^T \\ \frac{\partial \boldsymbol{\mu}}{\partial y} &= c^{-1}\dot{\mathbf{d}}_y \equiv c^{-1}[\dot{\bar{\mathbf{d}}}_y^T \mathbf{0}_L^T]^T \\ \frac{\partial \boldsymbol{\mu}}{\partial \nu_i} &= \mathbf{E}\mathbf{e}_i \end{aligned} \quad (51)$$

where $\dot{\bar{\mathbf{d}}}_x, \dot{\bar{\mathbf{d}}}_y$ denote the vectors containing the partial derivatives w.r.t. the client's position coordinates, which are given by,

$$\frac{\partial \bar{d}_i}{\partial x} = -\frac{(x_i - x)}{\sqrt{(x_i - x)^2 + (y_i - y)^2}} \quad (52)$$

$$\frac{\partial \bar{d}_i}{\partial y} = -\frac{(y_i - y)}{\sqrt{(x_i - x)^2 + (y_i - y)^2}} \quad (53)$$

Using (51), the FIM elements can be found as,

$$\begin{aligned} J_{xx} &= 2c^{-2}\dot{\mathbf{d}}_x^T\mathbf{W}^{-1}\dot{\mathbf{d}}_x = \frac{2}{c^2\bar{\sigma}^2}\dot{\bar{\mathbf{d}}}_x^T\dot{\bar{\mathbf{d}}}_x \\ J_{xy} &= J_{yx} = 2c^{-2}\dot{\mathbf{d}}_x^T\mathbf{W}^{-1}\dot{\mathbf{d}}_y = \frac{2}{c^2\bar{\sigma}^2}\dot{\bar{\mathbf{d}}}_x^T\dot{\bar{\mathbf{d}}}_y \\ J_{yy} &= 2c^{-2}\dot{\mathbf{d}}_y^T\mathbf{W}^{-1}\dot{\mathbf{d}}_y = \frac{2}{c^2\bar{\sigma}^2}\dot{\bar{\mathbf{d}}}_y^T\dot{\bar{\mathbf{d}}}_y \\ J_{x\nu_i} &= 2c^{-1}\dot{\mathbf{d}}_x^T\mathbf{W}^{-1}\mathbf{E}\mathbf{e}_i \\ J_{y\nu_i} &= 2c^{-1}\dot{\mathbf{d}}_y^T\mathbf{W}^{-1}\mathbf{E}\mathbf{e}_i \\ J_{\nu_i\nu_j} &= 2\mathbf{e}_i^T\mathbf{E}^T\mathbf{W}^{-1}\mathbf{E}\mathbf{e}_j \end{aligned} \quad (54)$$

Define,

$$\mathbf{J}_{\mathbf{pp}} \triangleq \begin{bmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{bmatrix} \quad (55)$$

$$\mathbf{J}_{\mathbf{p}\boldsymbol{\nu}} \triangleq \begin{bmatrix} \mathbf{J}_{x\boldsymbol{\nu}} \\ \mathbf{J}_{y\boldsymbol{\nu}} \end{bmatrix} \quad (56)$$

The FIM is given by,

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\mathbf{pp}} & \mathbf{J}_{\mathbf{p}\boldsymbol{\nu}} \\ \mathbf{J}_{\mathbf{p}\boldsymbol{\nu}}^T & \mathbf{J}_{\boldsymbol{\nu}\boldsymbol{\nu}} \end{bmatrix} \quad (57)$$

The CRLB is obtained by inverting the complete FIM.

$$\mathbf{J}^{-1} = \begin{bmatrix} \left(\mathbf{J}_{\mathbf{pp}} - \mathbf{J}_{\mathbf{p}\boldsymbol{\nu}}\mathbf{J}_{\boldsymbol{\nu}\boldsymbol{\nu}}^{-1}\mathbf{J}_{\mathbf{p}\boldsymbol{\nu}}^T\right)^{-1} & () \\ () & () \end{bmatrix} \quad (58)$$

The bound on the position coordinates is given by the top-left block of $\mathbf{J}^{-1}$,

$$\mathbf{C}_{\mathbf{pp}}^{1^{\text{st}}\text{Fix}} = \left[\mathbf{J}_{\mathbf{pp}} - \mathbf{J}_{\mathbf{p}\boldsymbol{\nu}}\mathbf{J}_{\boldsymbol{\nu}\boldsymbol{\nu}}^{-1}\mathbf{J}_{\mathbf{p}\boldsymbol{\nu}}^T\right]^{-1} \quad (59)$$

2) Approximate CRLB for CToA client in "Tracking" Mode: When the EKF is converged and the bSTA clock offsets are known (up to some residual error), and are being continuously tracked, then $\boldsymbol{\mu} \simeq c^{-1}\mathbf{d}$, and $\boldsymbol{\psi} \triangleq \mathbf{p}$.

Consequently,

$$\mathbf{J}_{ij} \simeq \frac{2}{c^2(\bar{\sigma}^2 + \sigma_r^2)}\left\{ \frac{\partial \boldsymbol{\mu}^T}{\partial \psi_i}\frac{\partial \boldsymbol{\mu}}{\partial \psi_j} \right\} \quad (60)$$

The partial derivatives are obtained using (51), and the CRLB on the position coordinates estimation error is obtained by,

$$\begin{aligned} \mathbf{C}_{\mathbf{pp}}^{\text{Tracking}} &\simeq \frac{1}{J_{xx}J_{yy} - J_{xy}^2}\begin{bmatrix} J_{yy} & -J_{xy} \\ -J_{xy} & J_{xx} \end{bmatrix} \\ &= \begin{bmatrix} \sigma_{xx}^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy}^2 \end{bmatrix} \end{aligned} \quad (61)$$

B. Proof of Proposition 1

Assume that every CToA broadcast includes N replicas of the timing measurements collected by the bSTA. If the clock offsets were time-invariant then one could define,

$$\check{\mathbf{E}} \triangleq \begin{bmatrix} \bar{\mathbf{E}} \\ \hdashline \mathbf{1}_N \otimes \tilde{\mathbf{E}} \end{bmatrix}, \ \check{\mathbf{W}} \triangleq \begin{bmatrix} \bar{\sigma}^2\mathbf{I}_{\mathcal{L}} & \mathbf{0} \\ \mathbf{0} & \tilde{\sigma}^2\mathbf{I}_{N\times L} \end{bmatrix} \tag{62}$$

Next, from (54) we have,

$$J_{\nu_i\nu_j} \quad = \quad 2\mathbf{e}_i^T\mathbf{E}^T\mathbf{W}^{-1}\mathbf{E}\mathbf{e}_j \tag{63}$$

Then, under (62), $J_{\nu_i\nu_j}$ becomes,

$$\begin{aligned} \check{J}_{\nu_i\nu_j} &= 2\mathbf{e}_i^T\check{\mathbf{E}}^T\check{\mathbf{W}}^{-1}\check{\mathbf{E}}\mathbf{e}_j \\ &= 2\mathbf{e}_i^T \left[\begin{array}{c:c} \bar{\sigma}^{-2}\bar{\mathbf{E}}^T & \tilde{\sigma}^{-2}\mathbf{1}_N^T \otimes \tilde{\mathbf{E}}^T \end{array} \right] \begin{bmatrix} \bar{\mathbf{E}} \\ \hdashline \mathbf{1}_N \otimes \tilde{\mathbf{E}} \end{bmatrix} \mathbf{E}\mathbf{e}_j \\ &= 2\mathbf{e}_i^T \left(\bar{\sigma}^{-2}\bar{\mathbf{E}}^T\bar{\mathbf{E}} + N\cdot\tilde{\sigma}^{-2}\tilde{\mathbf{E}}^T\tilde{\mathbf{E}} \right) \mathbf{e}_j \\ &\underset{N\to\infty}{\approx} 2N\tilde{\sigma}^{-2}\mathbf{e}_i^T\tilde{\mathbf{E}}\tilde{\mathbf{E}}^T\mathbf{e}_j. \end{aligned} \tag{64}$$

Recall that from (59) we have,

$$\mathbf{C}_{\mathbf{pp}}^{1^{\text{st}}\text{Fix}} = \left[\mathbf{J}_{\mathbf{pp}} - \mathbf{J}_{\mathbf{p}\boldsymbol{\nu}}\mathbf{J}_{\boldsymbol{\nu}\boldsymbol{\nu}}^{-1}\mathbf{J}_{\mathbf{p}\boldsymbol{\nu}}^T\right]^{-1} \tag{65}$$

Hence, under $N\to\infty$, $\check{\mathbf{J}}_{\mathbf{p}\boldsymbol{\nu}}\check{\mathbf{J}}_{\boldsymbol{\nu}\boldsymbol{\nu}}^{-1}\check{\mathbf{J}}_{\mathbf{p}\boldsymbol{\nu}}^T \to 0$.
Thus, given enough bSTA→bSTA measurements (equivalent to an EKF in "tracking" mode), $\mathbf{C}_{\mathbf{pp}}^{1^{\text{st}}\text{Fix}} \approx \mathbf{C}_{\mathbf{pp}}^{\text{Tracking}}$ (up to additive noise level scaling).
This concludes the proof.

C. On the Relation between the Concentration Ellipse and the CRLB

In [27], Torrieri derived the theory for bounding Gaussian-distributed estimation errors. In the geolocation problem considered, the parameters of interests include the 2-D coordinates vectors, which are defined as: $\mathbf{p} \in \mathbb{R}^2$, $\mathbf{p} = [x, y]^T$. The following appendix is based on the derivation in [27], and outlines the relation between the concentration ellipse, which provides a measure of accuracy for an unbiased estimator of a 2-D Gaussian vector, and the CRLB.

Let $\hat{\mathbf{p}}$ denote an unbiased estimate of $\mathbf{p}$, then given that the additive noise can be modeled as Gaussian, the probability density function (pdf) of the estimation error is given by,

$$f(\hat{\mathbf{p}}|\mathbf{p}) = \frac{\exp\left[-\frac{1}{2}(\hat{\mathbf{p}}-\mathbf{p})^T\boldsymbol{\Sigma}^{-1}(\hat{\mathbf{p}}-\mathbf{p})\right]}{(2\pi)^{\frac{n}{2}}\sqrt{\det(\boldsymbol{\Sigma})}} \tag{66}$$

where,

$$\boldsymbol{\Sigma} \triangleq E\left\{(\hat{\mathbf{p}}-\mathbf{p})(\hat{\mathbf{p}}-\mathbf{p})^T\right\} \tag{67}$$

The loci of constant density is defined as,

$$(\hat{\mathbf{p}}-\mathbf{p})^T\boldsymbol{\Sigma}^{-1}(\hat{\mathbf{p}}-\mathbf{p}) \triangleq \kappa \tag{68}$$

The scalar κ is a constant that determines the size of the n-dimensional region enclosed by the surface, which in the 2-D case is an ellipse. The probability that $\hat{\mathbf{p}}$ is contained inside that region is given by,

$$P_{in} = \int\limits_R \int \cdots \int f(\hat{\mathbf{p}}|\mathbf{p})d\hat{\mathbf{p}} \tag{69}$$

where,

$$R = \left\{\hat{\mathbf{p}} : (\hat{\mathbf{p}}-\mathbf{p})^T\boldsymbol{\Sigma}^{-1}(\hat{\mathbf{p}}-\mathbf{p}) \le \kappa\right\} \tag{70}$$

As shown in [27, eq. (39)], for the 2-D case,

$$P_{in} = 1 - \exp(-\kappa/2) \tag{71}$$

Further we have,

$$\boldsymbol{\Sigma} = E\left\{(\hat{\mathbf{p}}-\mathbf{p})(\hat{\mathbf{p}}-\mathbf{p})^T\right\} = \begin{bmatrix} \sigma_{xx}^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy}^2 \end{bmatrix} \tag{72}$$

The eigenvalues of $\boldsymbol{\Sigma}$ can be found by solving: $\det\{\boldsymbol{\Sigma}-\lambda\mathbf{I}\} = 0$.

$$\lambda_1 = \frac{1}{2}\left[\sigma_{xx}^2 + \sigma_{yy}^2 + \sqrt{(\sigma_{xx}^2-\sigma_{yy}^2)^2 + 4\sigma_{xy}^2}\right] \tag{73}$$

$$\lambda_2 = \frac{1}{2}\left[\sigma_{xx}^2 - \sigma_{yy}^2 + \sqrt{(\sigma_{xx}^2-\sigma_{yy}^2)^2 + 4\sigma_{xy}^2}\right] \tag{74}$$

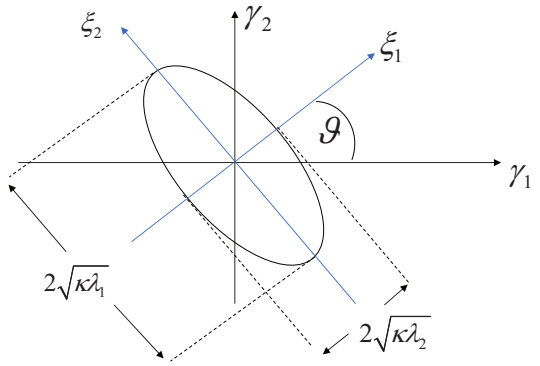

Fig. 13. Concentration ellipse and coordinate axes

As depicted in Fig. 13, assuming that the principle axes of the concentration ellipse lie on the axes ξ_1, ξ_2, which are counterclockwise rotated w.r.t. to axis system γ_1, γ_2 by an angle ϑ, then:

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \mathbf{A}^T \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} \tag{75}$$

where,

$$\mathbf{A} = \begin{bmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{bmatrix} \tag{76}$$

is an orthogonal matrix (with eigenvectors as columns). Since $\boldsymbol{\Sigma}$ is symmetric positive-definite matrix, and thus, so is $\boldsymbol{\Sigma}^{-1}$, the matrix $\mathbf{A}^T\boldsymbol{\Sigma}^{-1}\mathbf{A}$ is diagonal, provided that $|\vartheta| \le \frac{\pi}{4}$.

It can be shown that for a consecration ellipse defined by $\boldsymbol{\gamma}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}$, the principle axes of the concentration ellipse are given by $2\sqrt{\kappa\lambda_1}$ and $2\sqrt{\kappa\lambda_2}$. In order to obtain a bound on the maximal positioning error with a given probability, P_{in}, at a specific position, one needs to evaluate the size of half of the ellipse's major axis that is given by $\sqrt{\kappa\lambda_{max}}$, where λ_{max} is obtained by (73) and,

$$\kappa = -2\ln(1-P_{in}) \tag{77}$$

To summarize, the concentration ellipse is defined by 3 parameters: its major axis, its minor axis and its rotation angle - all of which can be obtained from the theoretical covariance matrix defined by the CRLB. This concludes the appendix.

REFERENCES

[1] "Wireless E911 Location Accuracy Requirements," *Federal Communications Commission*, PS Docket No. 07-114, Feb. 3, 2015.

[2] "Mobile Device Feature Attach Rate and Penetration," *ABI Research*, August 14, 2014

[3] J. Zheng, C. Wu, H. Chu and P. Ji, "Localization Algorithm Based on RSSI and Distance Geometry Constrain for Wireless Sensor Network," *2010 International Conference on Electrical and Control Engineering*, Wuhan, 2010, pp. 2836-2839.

[4] J. Torres-Sospedra *et al.*, "Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems," *Expert Systems with Applications*, vol. 42, no. 23, pp. 9263-9278, Dec. 2015.

[5] A. T. Parameswaran, M. I. Husain, S. Upadhyaya, "Is RSSI a Reliable Parameter in Sensor Localization Algorithms: An Experimental Study," 2009

[6] E. Elnahrawy, X. Li and R. P. Martin, "The limits of localization using signal strength: a comparative study," *2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks, IEEE SECON 2004.*, 2004, pp. 406-414.

[7] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo and L. M. Ni, "CSI-Based Indoor Localization," *IEEE Trans. on Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1300-1309, July 2013.

[8] Y. Shu et al., "Gradient-Based Fingerprinting for Indoor Localization and Tracking," I*EEE Trans. on Industrial Electronics*, vol. 63, no. 4, pp. 2424-2433, April 2016.

[9] X. Li and K. Pahlavan, "Super-resolution TOA estimation with diversity for indoor geolocation," *IEEE Trans. on Wireless Communications*, vol. 3, no. 1, pp. 224-234, Jan. 2004.

[10] F. X. Ge, D. Shen, Y. Peng and V. O. K. Li, "Super-Resolution Time Delay Estimation in Multipath Environments," *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 54, no. 9, pp. 1977-1986, Sept. 2007.

[11] L. Banin, U. Schatzberg, Y. Amizur, "Next Generation Indoor Positioning System Based on WiFi Time of Flight," *26th International Technical Meeting of the Satellite Division of The Institute of Navigation*, Nashville TN, Sept. 16-20, 2013.

[12] U. Schatzberg, L. Banin and Y. Amizur, "Enhanced WiFi ToF indoor positioning system with MEMS-based INS and pedometric information," *2014 IEEE/ION Position, Location and Navigation Symposium - PLANS 2014*, Monterey, CA, 2014, pp. 185–192.

[13] L. Banin, U. Schatzberg and Y. Amizur, "WiFi FTM and Map Information Fusion for Accurate Positioning," *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Alcalá de Henares, Spain, 4-7 October 2016.

[14] H. Kim, X. Ma and B. R. Hamilton, "Tracking Low-Precision Clocks With Time-Varying Drifts Using Kalman Filtering," *IEEE/ACM Trans. on Networking*, vol. 20, no. 1, pp. 257–270, Feb. 2012.

[15] X. Cao *et al.*, "Joint Estimation of Clock Skew and Offset in Pairwise Broadcast Synchronization Mechanism," *IEEE Trans. on Comm.*, vol. 61, no. 6, pp. 2508–2521, June 2013.

[16] T. S. Rappaport, *Wireless Communications: Principles and Practice,* Prentice Hall PTR, 1996.

[17] E. Perahia and R. Stacey, *Next Generation Wireless LANs: 802.11n and 802.11ac, and Wi-Fi Direct,* Cambridge University Press, 2nd Ed., 2013.

[18] A. J. Weiss, "On the Accuracy of a Cellular Location System Based on RSS Measurements," *IEEE Trans. on Vehicular Technology*, vol. 52, no. 6, pp. 1508–1518, Nov. 2003.

[19] S. Mazuelas et al., "Robust Indoor Positioning Provided by Real-Time RSSI Values in Unmodified WLAN Networks," *IEEE Jour. of Selected Topics in Signal Processing*, vol. 3, no. 5, pp. 821–831, Oct. 2009.

[20] J. Elson, L. Girod and D. Estrin, "Fine-Grained Network Time Synchronization using Reference Broadcasts," ACM SIGOPS Operating Systems Review - OSDI '02, vol. 36, no. SI, pp. 147-163, Winter 2002.

[21] M. A. Lombardi, T. P. Heavner and S. R. Jefferts, "NIST Primary Frequency Standards and the Realization of the SI Second," *The Journal of Measurement Science*, vol. 2, no. 4, pp. 74-89, Dec. 2007.

[22] P. D. Groves, *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*, Artech House Boston-London, 2008.

[23] *IEEE Std 802.11™-2016 (Revision of IEEE Std 802.11-2012) - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,* IEEE 802.11 Working Group, December 7th, 2016.

[24] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. of the IEEE*, vol. 92, no. 3, pp. 401-422, Mar. 2004.

[25] G. A. Terejanu, "Extended Kalman Filter Tutorial," available online: https://www.cse.sc.edu/ terejanu/files/tutorialEKF.pdf

[26] D. Taniuchi, X. Liu, D. Nakai and T. Maekawa, "Spring Model Based Collaborative Indoor Position Estimation With Neighbor Mobile Devices," in *IEEE Jour. of Selected Topics in Signal Processing*, vol. 9, no. 2, pp. 268–277, Mar. 2015.

[27] D. J. Torrieri, "Statistical Theory of Passive Location Systems," *IEEE Trans. on Aerospace Systems*, vol. AES-20, no. 2, pp. 183-198, Mar. 1984.

[28] Q. M. Chaudhari, E. Serpedin and K. Qaraqe, "On Maximum Likelihood Estimation of Clock Offset and Skew in Networks With Exponential Delays," *IEEE Trans. on Signal Processing*, vol. 56, no. 4, pp. 1685-1697, April 2008.

[29] H. L. Van Trees, *Optimum Array Processing. Part IV of Detection, Estimation, and Modulation Theory*, John Wiley & Sons, New York, 2002.