ACCURACY ENHANCEMENTS FOR ROBUST TOA ESTIMATION ON RESOURCE CONSTRAINED MOBILE PLATFORMS

By

Kumar Gaurav Chhokra

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Electrical Engineering

August, 2004

Nashville, Tennessee

Approved:

Gabor Karsai

Theodore Bapty

D. Mitchell Wilkes

DEDICATIONS

To my parents, for never letting me give up.

ACKNOWLEDGEMENTS

This work was sponsored by the Defense Advanced Research Projects Agency's (DARPA), Information Exploitation Office (IXO) under the Widely Adaptive Signal Processing (WASP) contract F30602-02-2-0206.

As I finish writing this thesis, I realize that there are many people to thank for my achievements but two come to mind first: Dr. Gabor Karsai, my advisor, for taking me under his wing and giving me the freedom and encouragement to pursue my research interests; Dr. Theodore Bapty, for supporting me when I was right, cautioning me when I wasn't and guiding me when I needed help. Without his friendly guidance and encouragement, this work would've been impossible.

I also wish to thank Dr. Wilkes for supporting my ideas and encouraging me to pursue my interests. The thought that I could always turn to him for an explanation was always soothing. Over the two years that I've been at ISIS, I have come to admire and respect Dr. Jason Scott for his grit and dedication. I am grateful to him for showing me that the best way needn't always be the most elegant.

I am also indebted to Dr. Ben Abbot and Dr. Don van Rheeden from SwRI for suggesting this direction of research, and to (soon to be Dr., no seriously!) Brandon Eames for introducing me to ISIS and motivating me to finish this thesis.

Finally, I wish to acknowledge the role my family has played in all my life: my father for teaching me to learn from my mistakes, my mom for helping me focus on the important things in life, and Vrinda, my sister, for being there when I needed her the most.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

LIST OF ACRONYMS

A/D:	Analog to Digital convertor
A-GPS:	Assisted – Global Positioning System
AOA:	Angle of Arrival
BLOS:	Beyond Line of Sight
BS:	Base Station
BW:	Band-width
CDMA:	Code Division Multiple Access
COMBAT-Q:	Raytheon's Combat Cueing system
COTS:	Commercial Off-The-Shelf components
EEPROM:	Electrically Erasable Programmable Read Only Memory
FRS:	Family Radio Service
GCC:	Generalized Cross-correlation
GPS:	Global Positioning System
GSM:	General System for Mobile communications
HUMINT:	Human Intelligence
IS:	Input Stream
JSTARS:	Joint Surveillance Target Attack Radar System
LOS:	Line of Sight
LPI:	Low Probability of Intercept
MS:	Mobile Station
ns:	nano-seconds

OAV:	Organic Aerial Vehicle
PCS:	Personal Communication System
PPM:	Parts per Million
PR:	Pseudo-random
PRS:	Pseudo-random-sequence
PSAP:	Public Safety Answering Point
RADAR:	Radio Detection And Ranging
RF:	Radio-Frequency
RF ID:	Radio Frequency Identification
RPS:	Remote Positioning System
RTC:	Relative Time Companding
SIGINT:	Signals Intelligence
SNR:	Signal to Noise Ratio
SOI:	Signal of Interest
SWPC:	Size, weight and power constraints
TA:	GSM Timing Advance
TDOA:	Time Difference of Arrival
TOA:	Time of Arrival
TTF:	Time to fix
UAV:	Unmanned Aerial Vehicle
UGS:	Unattended Ground Sensor
WASP:	Widely Adaptive Signal Processing
WU:	WASP Unit

CHAPTER I

INTRODUCTION

Recent advances in MEMS technology, embedded processing, and wireless communication are enabling the deployment of mobile networks, and location-aware systems and services. Some notable examples include the federal E911 program which mandates that the location of a distressed mobile user be made available to the public safety access points (PSAPs) with an accuracy of 50-100m [4]. Several industrial and marketing efforts are based on customization: Prada, a fashion company is currently fielding a marketing system that suggests matching clothes to a customer trying out a particular item of clothing [1]. The system communicates with RF IDs, or labels, attached to the merchandize to identify the user selection and searches a database to suggest complimentary fashion accessories . RF IDs or infra-red tags may also be used to build an in-building location network [6]: such a system may be used in a health-provider facility to keep track of doctors and other healthcare professionals within its premises. Location information may be used to route timely help to the patients in need or to contact the professionals without a need for broadcasting messages over the intercom. Also, for law-enforcement agencies, real-time updates on the location of a suspect or a distressed caller can be invaluable.

A similar trend is also observed in the strategic and commercial surveillance sectors. Knowledge of troop location and movement, both friendly and hostile, improves the effectiveness of the armed forces while significantly reducing friendly fire and

collateral damage [7]. Industrial customers interested in facility security may also employ similar systems.

While location intelligence may be gathered using a variety of methods, the solution offered by a network of distributed, low-power, low-cost sensors has several advantages including ubiquitous-presence, longevity and low-cost. Autonomous sensors may be deployed to fill gaps in information collection left by other higher cost surveillance mechanisms such as human intelligence (HUMINT) [7]. Sensor systems designed for low-power consumption can provide reliable monitoring capabilities with low false alarm rates from a few hours to several months. OmniSense ® [3] and Sparton IDS ® [4] are two commercially available systems that provide such capabilities. Typically maintaining surveillance capabilities using automated sensors is cheaper than maintaining a team of dedicated surveillance personnel [12]. In the recent past DARPA programs such as SensIT [8], [9], and Smart modules [2] have been instituted to investigate such networks.

Prior endeavors, such as Smart-dust [22], have primarily concentrated on remote information gathering using acoustic, infra-red, magnetic or seismic sensors, with little or no effort expended towards monitoring the radio frequency (RF) spectrum. Presence in the RF spectrum can not only help provide location fixes, but can also potentially provide valuable communication associated with a location. Should communication be lost, a history of transmission locations and the associated conversation can not only help first response personnel estimate the distressed caller's last location, and can help emergency response personnel better address the emergency. Similarly, in military operations, cues about the quarry's motion and intent can be invaluable in planning a course of action.

Widely Adaptive Signal Processing [11] was a DARPA effort to enhance situation awareness and geolocation capabilities in the RF spectrum. Widely Adaptive Signal Processing (WASP) aimed at producing a fleet of unmanned aerial vehicles (UAVs) capable of performing radio geolocation. Since manufacturing cost and physical size limitations were primary concerns, the geolocation sensor electronics platform was highly constrained in size, cost, weight and processing power.

Problem Statement

This thesis develops three techniques for accurately estimating the time-of-arrival (TOA) on resource constrained sensor nodes of a distributed radio geolocation system. Firstly, a multi-resolution approach for discriminating between signals of interest and spurious transmissions is presented. It is shown that this technique reduces the computational costs involved in the discrimination operation, and consequently, also reduces the overall time-required by the system to obtain a fix on the transmitter. Secondly, the problem of drifting sample rate clocks on different units is posed as a time-scaling problem. Instead of the conventional, but computationally expensive resampling, frequency shifting (known in the RADAR community as Doppler frequency shifting) is proposed as a solution. To circumnavigate the computational costs associated with the Doppler frequency shifting solution, a related time-based shifting technique is developed and analyzed. The duality between the two approximation techniques is highlighted. Finally, the problems of estimating a node's operating frequency and GPS jitter problem are posed together as a linear regression problem. Results are presented to support the techniques proposed.

Chapter II examines the existing radio-geolocation techniques and lists the challenges in designing a distributed, resource-constrained, mobile geolocation system. It also translates the real-world physical and operational requirements into overall system and lower level hardware, software design constraints. In Chapter III the multi-resolution signal detection algorithm is developed and analyzed mathematically. The computational savings achieved with this method are highlighted. In Chapter IV, two related methods of Doppler shift correction and time-shift (or equivalently, phase-shift) correction are developed to compensate for sampling frequency differences between received signals and their ideal expected versions. In Chapter V, the GPS jitter correction solution is mathematically developed and analyzed. Chapter VI showcases the actual system performance and examines the overall contribution of the compensation algorithms. Finally, Chapter VII envisions the possible future improvements and developments on the current work.

CHAPTER II

BACKGROUNDS

Depending on where the position measurements are made and how the location information is used, location determination systems can be broadly classified as *self-positioning systems* and *remote-positioning systems* [32], [16]. This classification is useful when comparing solutions for a given geolocation problem.

Self-positioning systems

In self-positioning systems, the positioning receiver makes appropriate signal measurements from geographically separated transmitters and uses these measurements to compute its position [32], [16]. Applications associated with the receiver may then use this information as needed. The most notable examples of self-positioning systems are the GPS receivers. A GPS receiver receives synchronized transmissions from a constellation of 24 low-orbit satellites. By measuring the differences in the time of arrival (TOA) of the signals from the different satellites, the GPS receiver is able to compute its location on the globe [17]. Figure 1 shows a schematic of this operation. Another common example is autonomous vehicle self-localization via triangulation in a bounded environment using active beacons [18].

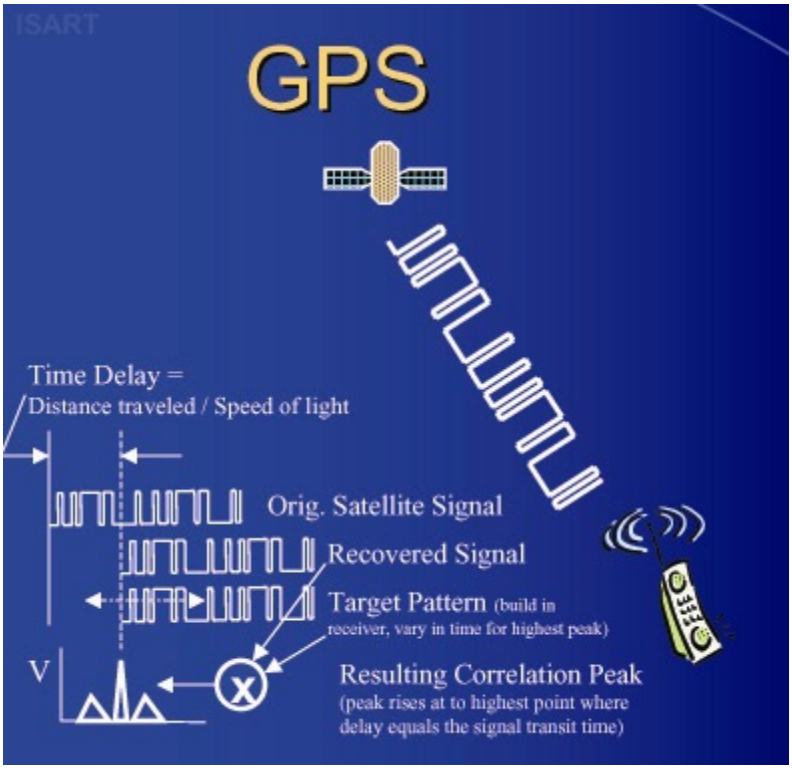


Figure 1. Schematic of the GPS self-positioning system. The mobile GPS-enabled receiver generates a pseudo-random sequence that is identical to the one transmitted by the GPS satellites. Once carrier phase-lock is achieved, a simple correlation operation estimates the time of flight of the transmitted signal. Using 3 to 8 such measurements, the receiver can determine its position on the globe. Figure adapted from [9]

Remote positioning systems

In remote positioning systems (RPS), receivers at one or more locations measure a signal originating from, or reflecting off, the source to be located. These measurements are communicated to a central location where they are combined to estimate the location of the transmitter. The onus of running the computationally intensive location estimation algorithms is thus relegated from the remote device to the central location. Depending on the source of the electromagnetic transmissions, such remote positioning systems may be further sub-classified as *direct* or *ambient illumination* systems. In direct illumination systems, the receivers "hear" a signal emanating from transmitter to be positioned. The

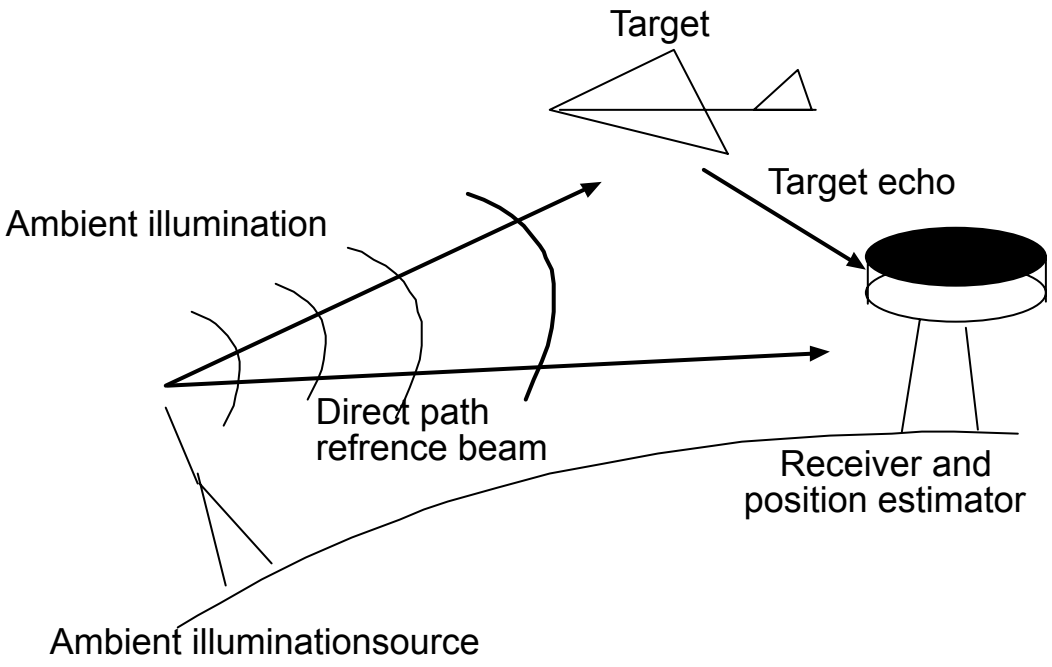


Figure 2. Schematic of a passive RADAR system. A powerful emitter, such as a TV or FM radio broadcast station provides ambient illumination. These signals when bounced off a potential target, such as an enemy aircraft, produce multi-path effects at the receiver site, which can be analyzed for Doppler shifts to produce an estimate of the range and bearing of the target aircraft. Figure adapted from [15]

JSTARS radar-jammer detection system [13] and cell phone location systems fall into this category. In ambient illumination systems [15], the receivers compare signals from ambient, non-cooperative sources of illumination, such as a television or radio broadcast stations, against echoes from the target being tracked. Figure 2 and Figure 3 illustrate the difference between these two sub-classes.

In practice, however, the remote positioning systems have several advantages over self-positioning systems. With RPS, the computationally expensive operations of determining location are relegated to base stations which have access to greater resources. The remote stations are usually severely constrained in terms of computational capabilities, form factor, and operating power, making computationally intensive tasks such as correlation (as performed by the GPS receivers) infeasible. Incorporating a GPS

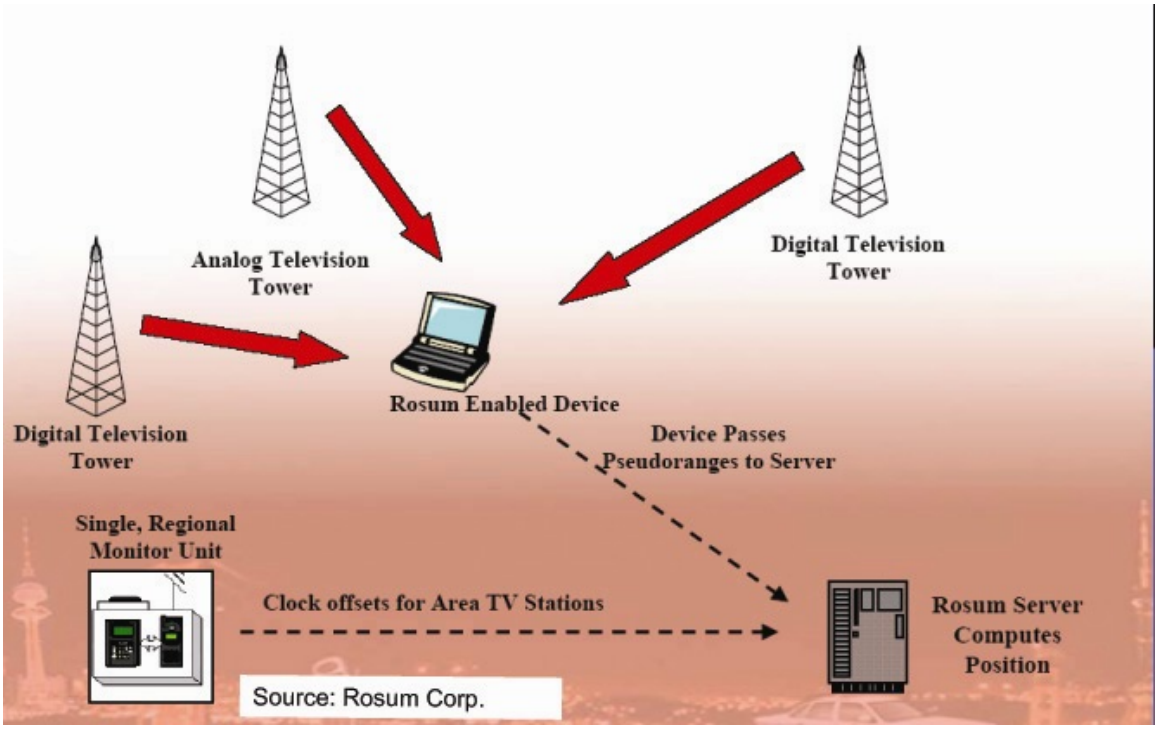


Figure 3. Schematic of passive self-location system. A Rosum enabled device "hears" transmissions from different ambient transmitters, such as radio and TV broadcast stations, and transmits the observed time of arrival characteristics to a base-station. Knowledge of the different clock offsets for each of the ambient sources heard enables the base-station to compute the location of the remote station. This information may be communicated back to the remote station or used by the base-station to provide location specific services [9]

receiver in a commercial cell-phone increases its weight, size, and cost. This extra hardware negatively impacts end-user convenience and hence mars usability. When dealing with a non-cooperative quarry, positioning systems such as GPS and assisted-GPS (A-GPS) generally fail and one must rely on passive methods of locus computation such as Angle of arrival (AOA), time difference of arrival (TDOA) or carrier phase. Remote positioning systems which use the uplink time difference of arrival (UTDOA) are extensively used by North-American cell phone providers as a solution for the E911 mandate [13], [19]. These techniques are briefly discussed in the following section.

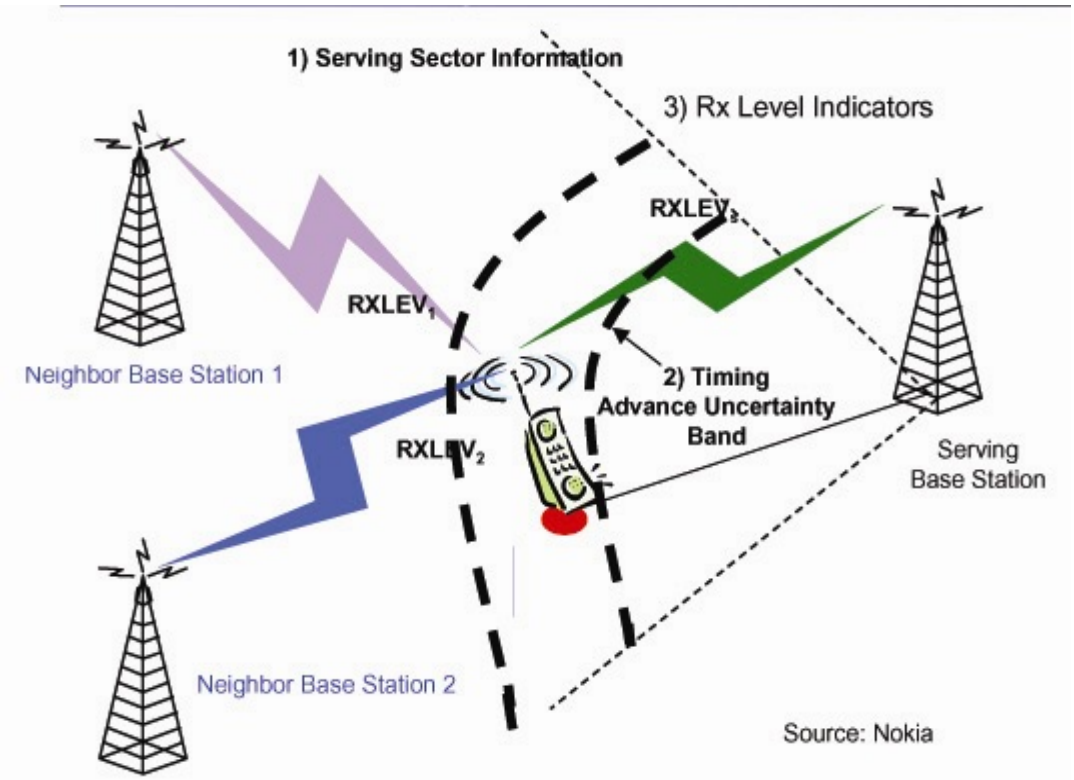


Figure 4. Schematic of active positioning system. When requested by the mobile-station, the base-stations "listen" to the signals emanating from the MS. For remote-positioning systems, the BSs fuse the measurements to produce a location estimate for the mobile station. For self-positioning systems, the roles are reversed. Adapted from [9]

Positioning techniques

Table 1 lists the various positioning techniques and their applications. We discuss each of the techniques below.

Table 1. Various geolocation techniques and their current applications

	Techniques	Applications
1.	Propagation time	Cell-phone geolocation, RADAR, commercial laser range measuring devices
2.	Angle of arrival (AOA)	RADAR
3.	Signal strength	Cell-phone geolocation (GSM)
4.	Time difference of arrival (TDOA)	Cell-phones (GSM and CDMA), WASP, COMBAT-Q
5.	Carrier phase	GPS
6.	GPS / A – GPS	CDMA Cell-phone geolocation, vehicle navigation, PCS based “friend finder”

Propagation time

In this technique, the time required for the signal from a transmitter to reach a measuring station is measured. Alternatively, an artificial echo approach may be used, wherein the receiver echoes back a signal transmitted from the receiver, giving a result twice that of the one-way measurement. For self-positioning systems, the mobile station (MS) usually initiates the positioning protocol and "listens" for transmissions or echoes from one or more fixed base-stations (BS) [9], [32], [16]. In remote positioning systems, one or more base-stations listen for transmissions from the remote mobile transmitter and fuse measurements to compute the transmitter's location. The one-way measurement approach assumes a good synchronization of the transmitter and receiver clocks, while the echo approach depends on small and accurately known response times from the remote station.

For systems using the Global Systems for Mobile Communication (GSM) technology, the range measurements may also be available as a consequence of the timing advance (TA) requirement [9].

Each propagation time measurement constrains the locus of the mobile station to a circle. An intersection of two such loci produces two possible locations in a 2D space. While using 3 such measurements theoretically resolves this ambiguity, system noise introduces errors and uncertainty; the locus of the transmitter is transformed from a circle to a circular band or ring. The thickness of the band is a function of the accuracy with which the propagation time can be measured: the lower the accuracy, the wider the band. By using 4 or more propagation readings, one may construct an overly constrained system, which may be solved by searching for an optimal least-squared solution.

Time difference of arrival (TDOA)

A mobile station can "listen" to a series of base-stations and measure the time difference between each pair of arrivals. If, for example, there are n base-stations, then n choose two (${}^{n}C_{2}$) independent difference measurements can be made. Each TDOA measurement defines a hyperbolic locus on which the mobile station must lie [27]. The intersection of two such hyperbolas produces a location estimate. Since the intersection of two hyperbolas can produce ambiguous estimates, in practice three or more are used. This also has the advantage of reducing the influence of noise. As in the propagation time measurement technique, the accuracy and synchronization of the receiver clocks is critical. However, TDOA scores over propagation time measurement because it doesn't require a protocol between the mobile and base-stations.

There also exists a trade off between the computational complexities of the propagation time (or range measurement) and TDOA techniques. While the former produces a set of linear equations amenable to a closed form solution, the TDOA measurements produce a non-linear system, which must either be solved iteratively or be cast as linear approximation [20].

Angle of arrival (AOA)

This involves measuring the direction of approach of signal transmissions from the mobile station to the plane of a directional antenna installed at the base-station. For a single measurement, the locus of the mobile station is computed as a straight line along the estimated angle of arrival. Combining two measurements produces an estimate of the location, since two lines can intersect at only one point. In practical situations noise

corrupts the location estimates. Hence, as before, a greater number of estimates are used to construct an overly-constrained system, which is solved using a least mean squared error approach. Location estimates are optimal when the transmitter within the convex hull of the receiver constellation [25]. AOA determination also needs a special array antenna and beam-forming algorithms for best results [23].

Signal strength

The strength of a received signal can serve as an indicator the distance traversed, provided that the initial transmission strength and the channel characteristics are well known. It is not difficult to see that this system suffers from many practical difficulties. RF shadowing and multi-path effects can change significantly over relatively small fractions of the carrier wavelength, altering perceived signal strengths. [24], [16]. Also, in GSM and GPRS systems remote handsets frequently change their transmission power levels to optimize battery life and communication quality. To ensure reliable measurements, the network must not only possess an accurate model of RF environment, but also update the model as the MS moves through then environment and encounters interferes and effects not known *a priori*. Also, the network must be cognizant of the changes in power levels of the communicating MS. Nonetheless, in conjunction with other techniques such as AOA or propagation time, the strength of a signal may be used to reduce ambiguity of position estimates. Figure 4 illustrates a cell-phone geolocation system that uses both the received signal strength (RXLEV) and the GSM timing advance delay to estimate transmitter location.

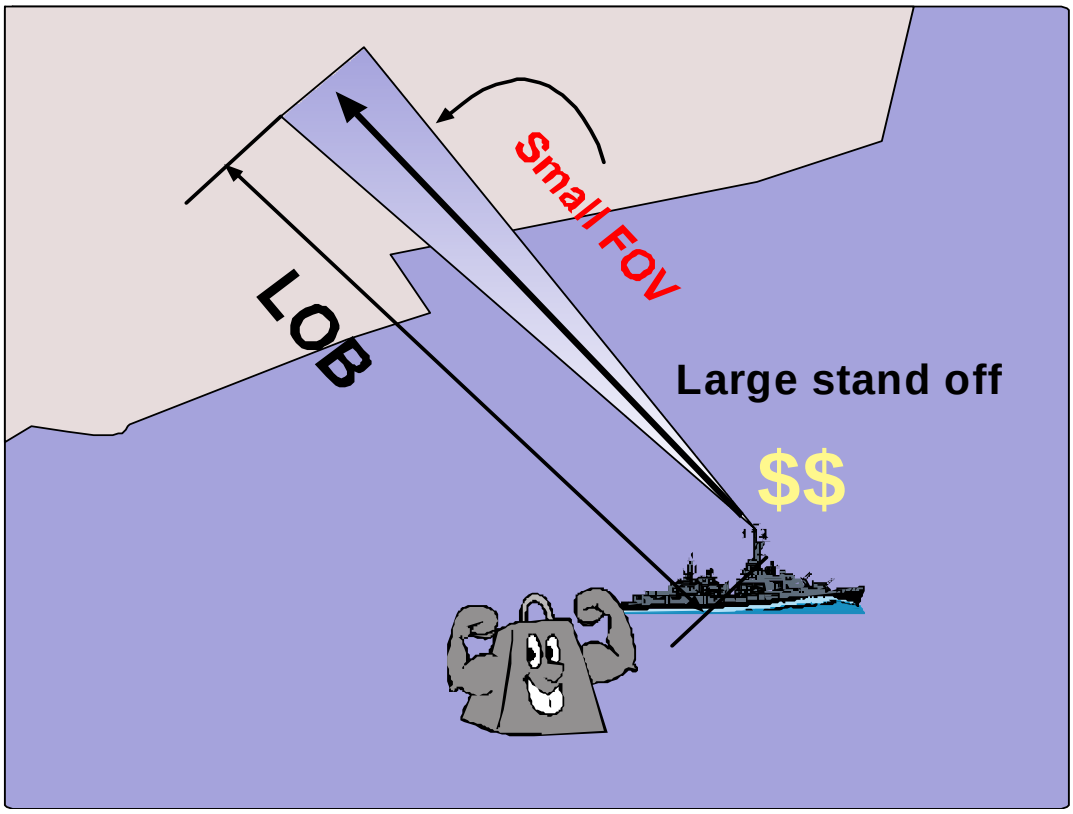


Figure 5 Characteristics of existing radio geolocation systems. Typically such systems are expensive and installed on defense vehicles enforcing large stand-off ranges. Consequently, they tend to have a small field of view (FOV) leading to a greater geometric dilution of precision (GDOP).

Carrier phase

The phase of a carrier has the potential to provide position estimates with an error considerably less than the carrier wavelength [32]. However, with this technique, a large number of ambiguities arise in the positioning solution: the technique only determines the phase differences, but cannot resolve the absolute path difference between signals received at two different sites. Another challenge is maintaining a continuous phase-lock on the carrier. Though challenging, this technique has been successfully employed in GPS receivers [27].

Disadvantages of existing geolocation systems

Given the strong need for position awareness, several geolocation systems (using one of the technologies listed above or their combinations) exist. For e.g., the armed forces have had access to such devices for over a decade. Commercial geolocation solutions for North-American cell-phone service providers have been mandated by the E911 program. The FCC mandate [28] requires that cellular and broadband personal communication system (PCS) providers be able to provide the PSAP attendants with the location of the mobile station with an estimate of the callers location within a radius of 125 meters in 67 percent of all cases. The FCC further classifies accuracy requirements based on the nature of the geolocation solution deployed: for handset based solutions (GPS, A-GPS etc.) the location of the mobile station must be accurate to 50 meters in 67 % of all calls and 150 meters for 97 % of all calls; for network based solutions (U-TDOA, propagation time etc.) the reported location must be accurate within 100 meters for 67% of all calls and 150 meters for 95% of all calls.

Existing long stand-off solutions suffer because they assume a point-source, while in a cluttered urban environment, a low-power cell phone or FRS radio behaves more like a spatially distributed source [29]; the MS may be obstructed or shadowed by large structures or be in a location where multi-path effects may distort available measurements. Such effects lead to a dilution of precision of location estimates. Also, with fixed measurement sites, tracking a MS in a cluttered urban environment is difficult.

Existing systems are usually expensive: cell-phone geolocation requires either expensive A-GPS enabled handsets. For network based approaches, such as uplink TDOA (UTDOA) [13], expensive (~$25,000) location measurement units must be

installed at each cell site. Deployed military systems cost several hundreds of thousands of dollars. All these systems also suffer from large weight and form factors and are either not mobile or, at best, have limited mobility. Moreover, existing commercial solutions fail for close range tracking of non-cooperative entities beyond the established network range. Thus there is a strong, justified need for a low-cost, mobile and robust geolocation system that is easily deployable and employable. Figure 5 demonstrates the characteristics of existing geolocation solutions.

Challenges in distributed UAV based system

Fleets of UAVs and UGSs have long been the appropriate means of providing reconnaissance at close quarters, with low risk of life to the scouting mission [7]. These can also be designed for stealth, so that they may aid in greater situation awareness without alerting the target under surveillance [3]. A fleet of mobile measurement sites can also overcome the shortcomings of conventional geolocation systems: the receiver site configuration may be dynamically adapted to account for the transmitter's motion. Given knowledge of the RF environment, the receivers may be housed in locations with minimum parasitic effects, resulting in optimal measurements.

Equipping UAVs with the technological components necessary to achieve this aim is fraught with challenges. These mobile units must be economical to mass produce. Should such a device fail, its replacement should be relatively cheap and easy. While this imposes a strong constraint on the production and maintenance costs, it also implies that the components used be economical, and hence neither as accurate nor powerful as in a high precision system. The mobility and stealth requirement of such devices suggests that

they be constrained in size and weight. To be of any practical use, these devices must have long field lives, implying extensive operating cycles. The constraint of a small weight limits the battery size and, hence, bounds the operating life. To ensure longevity, the devices must therefore consume as little power as possible. To be usable, the final system must provide the necessary accuracy: for the WASP project, the desired system accuracy is 75 meters over an area of a few square kilometers. The accuracy constraints and physical requirements of mobility and cost are generally non-confluent and hence produce a challenging problem. Table 2 illustrates the implication of these constraints.

Translating Constraints into Design Objectives

The logistical and fiscal challenges, adumbrated in prior sections, form a set of interacting and cross-cutting constraints. This prohibits a highly-structured, quantitative analysis to produce the best hardware / software design. Instead, we adopt a more qualitative approach based on heuristics. The constraints may be mapped into the following criteria: economic feasibility, distributed processing, goal-oriented on-demand processing, and multi-resolution processing [56]. The translation of the higher level requirements into design objectives are presented below.

Table 2. Desired characteristics of a UAV based radio-geolocation system and implications on system design

	Qualities	Implications
1.	Accuracy	**75m – 100 m over a few sq. km.** To be usable, the final system accuracy must match existing solutions. Since the system being developed is a passive RF geolocation network, the accuracy of existing network based geolocation systems is used as a starting point. The system is expected to perform better as it capable of relocation to compensate for transmitter motion.
2.	Low cost	**Lack of high-precision components.** Components have greater error tolerances, which subsequent processing must compensate for. This also precludes the use of algorithms which work only under ideal or near ideal conditions; i.e. the system must account for no-high SNRs, no common global clock synchronization, varying operational characteristics etc. **Lack of large memory.** While memory may no longer be expensive, the total system cost provides a strong upper bound for the financial that may be allocated for more memory, providing a strong constraint: cannot have arbitrarily large data rates; limit on length of FFT, or equivalently, limited tempo-frequency resolution.
3.	Longevity	**Low-power consumption.** To be usable over a finite deployment period, the units must consume as little power as possible. Since the same on-board energy sources (battery packs and/or solar cells) would power the units' control and functional components, the available power must be used judiciously. This precludes long-duration power hungry algorithms. **Reduced radio communication**. Since each radio transmission consumes pico-joules to milli-joules per transmitted bit [30] constantly synchronizing or extremely collaborative system designs must be eschewed.
4.	Stealth	**Mechanically design.** The UAV based system must be designed to closely track and monitor without alerting the target being tracked. Mechanical operational noise (such as the whirring of rotors) might alert the quarry to the presence of these tracking devices. **Radio silent design**: Frequent radio communications in or around the target may cue it to take evasive or aggressive action against the tracking devices. Also, a capable target may use a similar triangulation technique to locate and eliminate one or more of these devices.
5.	Mobile / UAV mountable	**Small form factor** **Light weight** **Low power**
6.	"Attritable"	**Easy to replace.** **Low cost.**
7.	Adaptable	**Modular/ Configurable design**. The system should be independent of the modulation technique or the frequency spectrum range to enable agility.
8.	Real time	**Finite and deterministic time to fix**. Throughput must be maintained. Real time constraints must be satisfied. The system cannot take an arbitrarily long time to compute the results. Typically a few seconds after sensing activity in the RF spectrum of interest is acceptable delay.

Economical feasibility

The WASP units (WUs) must be "attritable", viz. the loss of a WASP unit is favored to loss of trained military personnel. Should a WU be damaged or lost during a mission, it must be easy and affordable to procure a replacement. This implies an ease of production, maintenance and, if necessary, augmentation. This is best achieved by using cheap, mass-produced, and commercial-off-the-shelf (COTS) components. The overall system cost is minimized by realizing a minimal system that delivers acceptable performance.

Distributed processing

To enable stealth, a passive geolocation strategy is adopted. The deployed units serve as remote, distributed sensors that gather target signal characteristics, while the computationally intensive tasks of information-fusion, higher level objective definition and path planning are relegated to a remote base-station (BS).

Low operating power, low communication bandwidth and stealth behavior require that the signals sampled at each node be processed locally. With current technology, heuristics for power consumption vary between picojoules to nanojoules per instruction for processing, while radio communication consumes microjoules to millijoules per bit, depending on desired range and link geometry [30]. When communicating at a low elevation and in an environment cluttered with foliage and human-made structures, significant losses in transmitted power may arise due to shadowing and channeling effects [31] necessitating significantly higher power levels for reliable RF links.

To ensure adequate temporal resolution the RF signals must be sampled at a fairly high rate (~2 MHz), while a sufficiently large number (100,000 to 1,000,000) of sample points must be generated to enable reliable "discrimination" between signals of interest and static. The conventional sensor-array technique of correlating data sampled by the different sensors would require repeated cross-communications of the entire time-series data, implying prohibitive power and time costs. Also, sustained transmission by the sensor nodes would render them susceptible to detection.

The above arguments make a strong case for reducing the number and duration of transmissions over the wireless network, obviating common time-series processing at the base station. Instead, the nodes must use some form of *a priori* knowledge about the signal of interest (SOI) to extract the relevant discriminators. Figure 6 depicts the distributed nature of the WASP solution.

Table 3. Template (pseudo-random sequence) design parameters

Sampling frequency	2 MHz
Duration	200 m-sec
Bandwidth	3 KHz
Center frequency	1.5 KHz

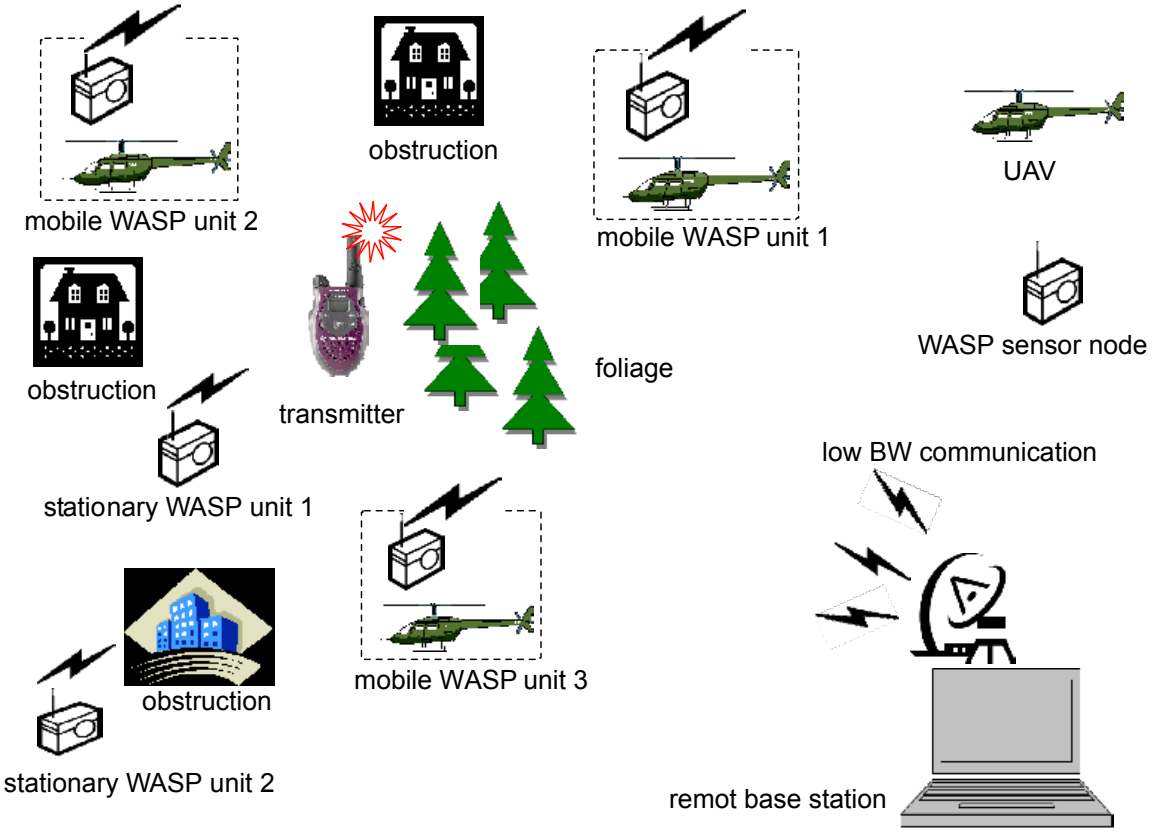


Figure 6. Distributed processing in the WASP geolocation system. Each deployed sensor node "watches" for the RF emissions of the transmitter. The nodes match received transmissions against an on-board version of the SOI. Once found, the time of detection, denoted time of arrival (TOA), is noted and transmitted to the remote-base station. Knowledge of geographical location of the nodes with the observed time-differences enables the base-station to estimate the location of the transmitter. The base-station may also initiate a repositioning of the mobile nodes to enhance future measurements or compensate for the transmitter's motion. The inter-node communication is completely eliminated using this approach. Since the geographic location and TOA may be compactly represented, communications between the nodes and base-station are limited to a few bytes.

The "optimal" set of signal features that constitutes a SOI is heavily dependent on the application as well as the nature of device being tracked. For a digital communication system, such as a GSM cell-phone, a relevant SOI may be the synchronization sequence in each GSM frame [24] [32], while for an analog system, it may be the frequency content or a certain keying pattern, such as the calling tone or "auto-over" pulse common to most voice carrying radio units [33]. For passive operation, the SOI is identified and

stored on board the WUs as a *template*, which is compared against the received signals. For the current application, a band-limited (base-band) pseudo-random sequence sampled at 2 MHz is used as the template. The template parameters are enumerated in Table 3.

The target emitter is assumed to be a 500 mW family radio service (FRS) band unit. The FRS unit modulates audio frequencies in the range of about 500 Hz – 3 KHz onto a specified FM channel in the frequency range 462.5625 MHz - 467.7125 MHz [34] with a bandwidth of approximately 12.5 kHz [35]. This designated radio channel is continuously monitored by the nodes. When radio activity (above background noise) is observed, time-series data are recorded for further processing. The intermediate frequency (IF) FM data are gathered by a 12-bit A/D at a sampling frequency of 2 MHz, barring oscillator clock drift. The time-series data are demodulated and compared with the on-board template to discriminate between relevant signals and spurious transmissions.

Since communicating the time-series data between the nodes or the base-station is infeasible, the TDOA approach is adopted: each node communicates its location and the time at which it detects the SOI on the radio channel to the remote base-station, which uses this knowledge to estimate the location of the transmitter. Other passive techniques such as angle of arrival (AOA) and beam-forming are ruled-out because they either require expensive or bulky specialized equipment (such as a directional antenna) [23] and are computationally expensive.

Goal-oriented on-demand processing

The nodes must expend power only when an interesting event occurs. The nodes communicate their location and TOA estimates only when requested by the base station. This aids in stealth because the nodes operate in a passive mode between transmission requests, and even when they transmit, the transmissions are low-power, small duration, frequency hopping, low-bandwidth (9600 BPS) transmissions and thus have a lower probability of intercept (LPI). Since each radio communication consumes micro-joules to milli-joules per transmitted bit (depending on the desired range and link geometry of the radio network), power is conserved by transmitting data on demand or when an signal of interest is detected [30].

Multi-resolution processing

The signal processing algorithms required to estimate the time of arrival typically are computation and power intensive. A higher temporal resolution demands the time series data be gathered and processed at a high rate. A higher bandwidth, however, implies larger power consumption. The nodes may thus save power by searching for an event of interest at a coarser temporal resolution, and then switch to a more computation intensive higher resolution when a SOI has been detected. This also aids in reducing the time-to-fix (TTF) as the initial search is performed on a pruned signal space. This is examined in greater detail in chapter III.

CHAPTER III

MULTI-RESOLUTION SIGNAL DETECTION

Introduction

Traditionally, as explained in Appendix A, the "searching" of the template in the sampled data stream is typically done using frame-based cross-correlation using windowed Fourier transforms. However, one may also use a direct convolution approach when the template length is much smaller than the length of the stream. In this case, a time-reversed and delayed (by the filter duration, T) version of the template is used as a filter that operates over the IS. For continuous systems, the output of the filter is the maximum exactly T seconds after the start of the SOI in the input stream [42]. For signals with temporal characteristics different from the template, the response of the matched filter is generally small. The matched filter receiver has traditionally been the system of choice for analog radar systems due to the convenience of implementation [43]. In discrete time systems, this technique obviates the need for the forward transform, reverse transform and searching required in the frame based approach described in Appendix A. It also eliminates complex additions and multiplications and thus has a natural, simplistic appeal from an implementation perspective. However, the computational burden associated with the discrete time version of this technique is proportional to the product of the lengths (measured in number of samples) of the template and the IS, which for large template lengths may be significantly greater than the cost of searching the IS using a frame based approach.

For the current case, the SOI has a temporal footprint of about 0.2 seconds, which amounts to approximately 400,000 samples given a sampling rate of 2 MHz. Clearly, using a matched-filtering approach with such a large filter is computationally prohibitive. Even for a windowed approach FFT which has $O(N\log N)$ complexity, where N is the length of each window (or, equivalently the template), the cost associated with such a long template represents an enormous burden.

Considerable savings may be achieved during the searching phase by recognizing that the high sampling rate (2 MHz) introduces enormous redundancy in the sampled input stream. While the high sampling rate is justified for satisfying the Nyquist criterion (for sampling the FM data on the 450 kHz intermediate frequency), it represents an oversampling of the demodulated base-band data by a factor of nearly 200. The redundancy due to oversampling may be reduced by resampling the demodulated input stream at a much lower sampling frequency (say 15 KHz). Consequently, the lengths of the template and the input stream are also proportionately reduced. Searching on these reduced length sequences is several orders of magnitude faster (as shown in the subsequent sections) leading to enormous savings in computational effort and time. E.g, downsampling by a factor of 128 reduces the sampling rate to 15.625 KHz, and the template length to 3125 samples, which is small enough to be implemented using "pre-packaged" optimized DSP functions.

Once the SOI has been located in the downsampled input stream, temporal resolution may be regained by processing only the relevant section of the original input stream. The following sections develop these ideas formally.

Definitions

Let P be the pseudo-random template to be detected in the sampled signal stream S. Let lengths of P, S be L_p, L_s samples respectively. It is assumed that $L_s >> L_p$ and that there may be more than one occurrence of P in S. Let R be the number of resolution levels. Let the subscript $\{*\}_i$ denote a signal at resolution i, with a higher number denoting a signal sampled at a higher resolution, viz. at a higher sampling frequency. Let superscript $\{*\}^{(k)}$ denote the k^{th} iteration or step through a stream. Since the incoming stream S is processed in overlapping blocks or frames, let α_i, β_i, respectively, denote the starting, ending sample numbers of the block being processed at resolution i. Let N_i denote the number of samples in each block, at each resolution i. Let M_i denote the number of samples of overlap between two consecutive blocks, and m_i be the number of samples skipped over, i.e. $m_i = L_{Pi} - M_i$. Let $\prod_{\alpha,\beta}(n)$ denote the discrete "gate" or "rectangle" function given by

$$\prod_{\alpha,\beta}(n) = \begin{cases} 1; \alpha \le n \le \beta \\ 0, otherwise \end{cases} \tag{5}$$

Let, $\circ$ denote the cross-correlation operation between two signals, e.g.

$$R_{xy}(\tau) = x \circ y(\tau) \tag{6}$$

denotes the cross-correlation between signals $x(t)$ and $y(t)$. Finally, let D be the downsample factor, i.e. the ratio of the length of signal at resolution, i, to the length of its downsampled version at the next coarser resolution, *i-1,* i.e.,

$$D = \frac{L_{P_i}}{L_{P_{i-1}}} \tag{7}$$

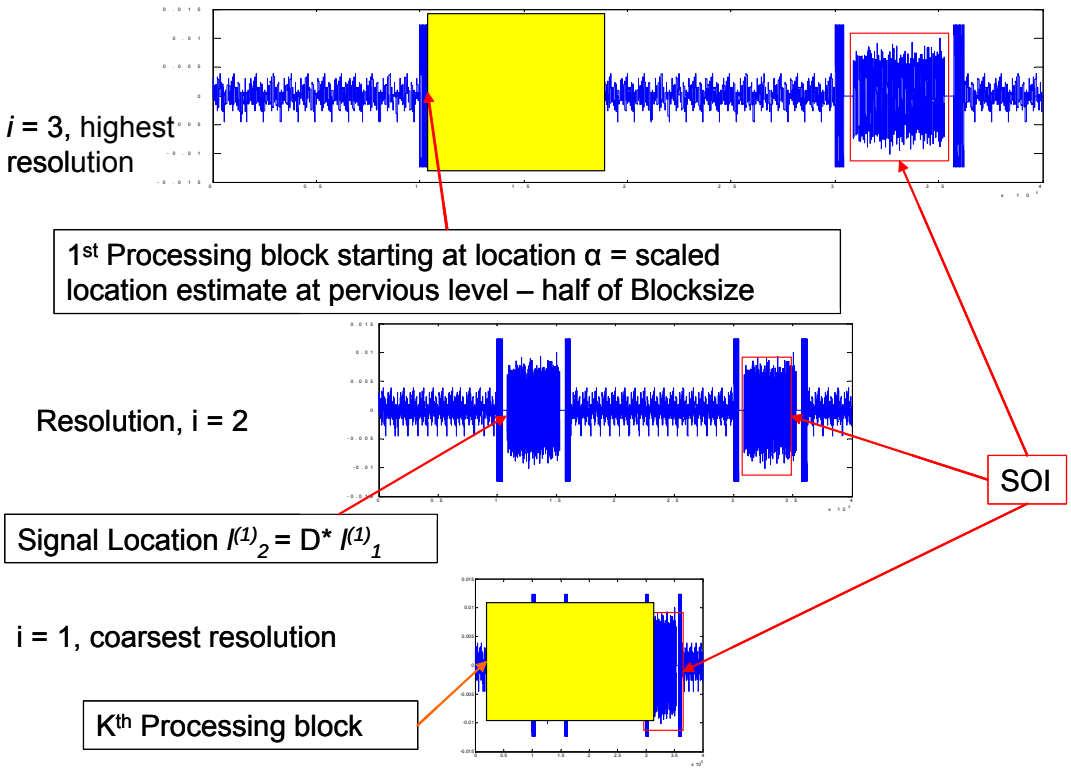


Figure 7. Detection of the SOI using a multi-resolution scheme. The input stream is downsampled after low pass filtering and searched for the template. Once found, location estimates are subsequently refined at higher resolutions by concentrating searches around expected locations

Multi-resolution searching

Table 4 summarizes the notations and variables used. At each resolution level, the location for the pulse may be estimated by detecting the lag, τ, corresponding to the maximum of the correlation function. The maximum is accepted as a valid indicator of the occurrence of P if its magnitude is greater than a specified threshold. Let this threshold be γ_i at each resolution level i. The location at the coarsest resolution, i, may be expressed mathematically as,

$$l_i^{(k)} = \arg\max_n \left[P_i(n) \circ \left(\prod_{\alpha_i^{(k)} \beta_i^{(k)}} S_i(n) \right) \right] + \alpha_i^{(k)} \qquad (8)$$

Table 4. List of notations and definitions

1.	P	Pseudo-random template to be detected
2.	S	Sampled signal stream to be searched
3.	$L_{()}$	"Length of" operator. E.g., L_s = length of signal S, L_{P_i} = length of signal P_i
4.	R	Number of resolution levels
5.	$\{*\}_i$	"at resolution i" operator. E.g., P_i means signal P at resolution i
6.	$\{*\}^{(k)}$	"At iteration" operator. E.g., β^k means β at iteration k
7.	α_i	Starting sample number of processing block at resolution i
8.	β_i	Ending sample number of processing block at resolution i
9.	N_i	Number of samples in a processing block at resolution i
10.	M_i	Step size in number of samples at resolution i
11.	m_i	Number of samples of overlap between two consecutive blocks
12.	$\prod_{\alpha,\beta}(n)$	Discrete rectangle function. For values of n between the integers α and β the function has a value of unity. It is zero otherwise. $\prod_{\alpha,\beta}(n) = \begin{cases} 1; \alpha \le n \le \beta \\ 0, otherwise \end{cases}$
13.	$R_{xy}(\tau) = x \circ y(\tau)$	Cross correlation between signals $x(t)$ and $y(t)$
14.	D	Downsample factor. Ratio of length in number of samples of a signal at one resolution to its length at the next coarser resolution.

subject to the constraint

$$\max\left(\left|P_i(n)\circ\left(\prod_{\alpha_i^{(k)}\beta_i^{(k)}} S_i(n)\right)\right|\right)\geq\gamma_i\max\left(P_i(n)\circ P_i(n)\right) \tag{9}$$

where, at the lowest resolution, i = 1,

$$\alpha_1^{(k)}=\alpha_1^{(k-1)}+M_1 \tag{10}$$

$$\beta_1^{(k)}=\alpha_1^{(k)}+N_1;\ \ 0\leq k\leq\left\lfloor\frac{L_{s_1}}{M_1}\right\rfloor \tag{11}$$

and for higher resolutions, viz. $2\leq i\leq R$,

$$\alpha_i^{(q)}=Dl_{i-1}^{(q)} \tag{12}$$

$$\beta_i^{(q)}=\alpha_i^{(q)}+N_i;\ 1\leq q\leq Q \tag{13}$$

where Q is the number of locations at which P was detected at resolution i = 1.

Equations (10) and (11) indicate that at the lowest resolution, the algorithm divides the signal stream S into blocks of size M_1 and correlates each block with the corresponding low-resolution template P_1. The constraint of equation (9) ensures that any block is considered to contain P only if the magnitude of the maximum of the correlation function is greater than a specified threshold at that level, γ_i. Once all such locations for a given resolution level have been detected, the search moves to the next higher resolution. At the higher resolutions, equations (12) and (13) specify that instead of searching through the whole time series data, the correlations are "centered" around the scaled location estimate. Figure 7 illustrates these ideas. For given signal stream *S* containing 2 occurrences of the SOI (*P*), we construct the 2 other lower resolution versions of *S*. The downsample factor, for the sake of exposition, is 2. Thus if $L_s = L_{s3}$ = 40K samples, then

L_{s2} = 20K and L_{s1} = 10K samples. At the lowest level, the algorithm steps through the downsampled signal stream with overlapping blocks till it locates the SOI at locations (say) 2000 and 7000. At the next higher resolution level, i = 2, the first block searched begins at sample number 4000, which is 2000 scaled by the downsample factor (2). The location estimate is then refined to sample number 4002 at resolution 2. At the highest resolution, i =3, the search block starts at location 8004 and refines the estimate to location 8011. Thus the multi-resolution search refines the location (sample number) estimate of the SOI within the stream *S*. At each resolution level, a portion of the signal stream to be searched is culled resulting in computational savings. This algorithm as implemented on the WUs is shown in Figure 8.

- The EPSON crystal oscillator generates a 48 MHz clock, which is divided by factor of 24 generates a nominal 2 MHz sampling clock. Since the clock frequency is specified to be accurate to ± 100 PPM, the sampling clock may be assumed to be accurate to approximately ±5 Hz.

- The signal stream from the FRS receiver after the relevant software filtering and demodulation operations is sampled at this frequency and stored in memory to produce the time series data for the signal stream *S*. *S* is filtered using a polyphase band-pass filter and decimated to generate S_1, the downsampled version of the signal *S*.

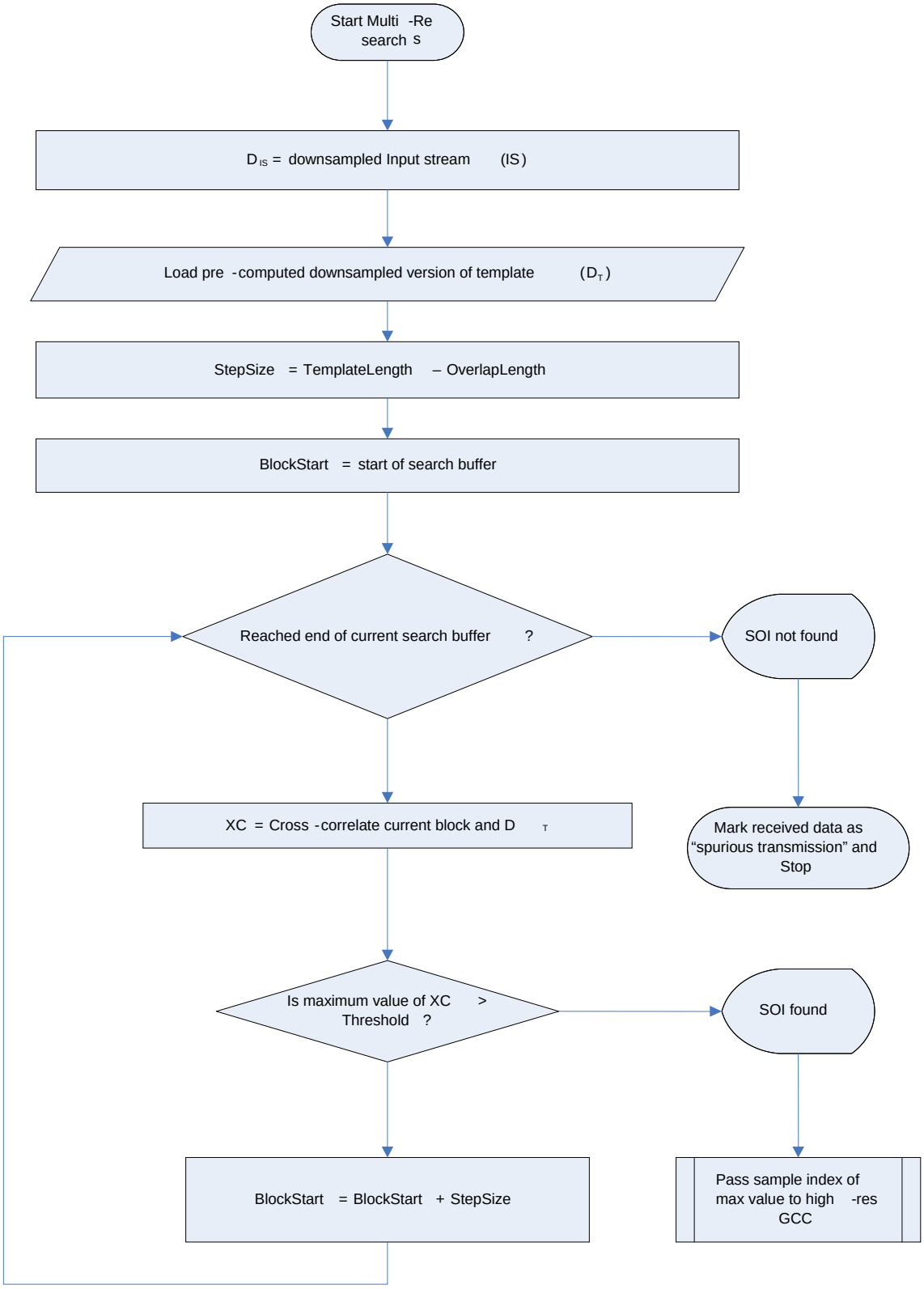


Figure 8 . Control flow for multi-resolution search

- The template P also passes through the same filtering and decimation operation to generate P_1. As an optimization, P_1 is pre-computed and stored in the non-volatile memory of each WU. Thus we trade memory for savings in computational effort.
- S_1 is searched for occurrences of P_1 using the block correlation (frame based matched filtering) approach. The maximum for each block is compared against a threshold (see constraint specified by equation (9)).
- If a valid maximum, i.e. if the maximum for a block is greater than the threshold, P_1 is considered found and the estimated signal location, $\hat{\eta}_\tau$, is passed to the GCC block, else the estimates are discarded as invalid and the system simply keeps processing more data till a valid detection is obtained.
- When activated, the GCC "block" performs a high resolution cross correlation of the template P against an a section of S length L_p, starting at location $\hat{\eta}_\tau$. The refined estimate gives the TOA for the SOI at the node under consideration. The GCC block uses knowledge of the current operating / sampling frequency of the node to compensate for the affects of relative time scaling (due to incongruent sampling frequencies) between S and P.

Computational Efficiency

To understand the computational gains using the multi-resolution approach, we must compare the computational efforts required for estimating the signal location using the single resolution approach with those required for the multi-resolution approach. The location of the SOI at the highest resolution may be obtained using either the matched filter or block correlation approach.

Matched filter approach

As explained earlier, in the matched filter approach, a time reversed and delayed version of the SOI is used to define a filter, P, which operates on the incoming data stream, S. Since each filter output sample requires L_p multiplications and L_p additions, the cost for computing one filtered output sample is $2L_p$ operations (additions and multiplications), and since the desired output has $L_s + L_p - 1$ samples, the total computational cost is $2L_p\ (L_s + L_p - 1)$ operations. Since $L_s >> L_p$, this may be written as $O(L_sL_p)$.

Fast overlap-add convolutions

For a frame or block approach, the computation effort for each block is $O(L_s log(L_p))$ assuming that the block sizes are the same as the template length. Such techniques are also referred to as "fast-overlap-add convolutions" in literature. [44]. The signal S is split into L_s/L_p having L_p non-zero samples:

$$S[n] = \sum_{r=0}^{L_s/L_p - 1} S_r[n] \text{ with } S_r[n] = S[n + rL_p] \prod_{[0, L_p - 1]}[n] \tag{14}$$

For each $0 \le r \le \left\lfloor \frac{L_S}{L_P} \right\rfloor$, the $2L_p$ non-zero samples of $g_r = S_r * P$ (where the '*' indicates convolution) is computed using the FFT convolution algorithm. Each such step requires $O(L_p \ logL_p)$ operations, and since there are L_s/L_p such steps, the total cost of computing the individual filtered blocks is $O(L_s logL_p)$. The results are then combined to produce the entire filtered sequence as

$$S * P[n] = \sum_{r=0}^{L_s/L_p - 1} g_r[n - rL_p] \tag{15}$$

The addition of these L_s/L_p blocks of size $2L_p$ is done with $2L_s$ operations. The overall convolution is thus performed in $O(L_s logL_p)$, which is significantly less than the computational burden in the matched filter approach.

Multiresolution approach

For the multi-resolution case at the lowest resolution, the operation of the algorithm is identical to the fast overlap-add algorithm presented above. Since the signal and template lengths are reduced, the location computation procedure at the lowest resolution takes C_{coarse} steps, where

$$C_{coarse} = O(L_{S1} \log(L_{P1})) \tag{16}$$

At the higher resolutions each estimate is refined in $O(\log(L_{Pi}))$. Given Q estimated locations and R, the total cost is then

$$C_{refine} = O(\sum_{i=2}^{R} Q \log(L_{Pi})) \tag{17}$$

where for $2 \le i \le R$,

$$L_{Pi} = D^{i-1} L_{P1} \tag{18}$$

Thus,

$$\sum_{i=2}^{R} \log(L_{Pi}) = \sum_{i=2}^{R} \left(\log(L_{P1}) + \log(D^{i-1}) \right) \tag{19}$$

$$= \sum_{i=2}^{R} \log(L_{P1}) + \log D \sum_{i=1}^{R-1} i \tag{20}$$

Substituting equation (20) in equation (17) and exploiting laws of arithmetic under the order notation, we get

$$C_{refine} = O(QR \log(L_{P1}) + QR^2 \log D) \tag{21}$$

$$= O(QR \log(L_{P1} D^R)) \tag{22}$$

where D is the downsample factor as defined earlier. The overall cost to obtain a final refined estimate is the sum of the efforts at all levels, given by

$$Ctotal = C_{coarse} + C_{refine} \tag{23}$$

$$= O(QR \log(L_{P1}) + QR^2 \log D + L_{S1} \log(L_{P1})) \tag{24}$$

$$= O(L_{S1} \log(L_{P1})) \tag{25}$$

To enable a fair assessment, we must factor in the costs of creating multi-resolution representations of the sampled signal and template. Since the template is know *a priori*, its multi-resolutions may be pre-computed. This affords computational savings at a moderate cost to non-volatile system memory. Before the input stream can be decimated, it must be low-pass filtered to avoid aliasing. The low-pass filter is of a length M (usually 32 or less), which is much smaller than Ls. The downsampling and creation of the multi-resolution approach may be carried out in an efficient manner by recognizing that the downsampled version of the signal is smaller by a factor of D^R and because of the small length of the filter, each lowest resolution signal sample depends only on M samples in the highest resolution. Thus the entire S_1 signal stream may be computed in $O(L_s D^{-R} M)$, i.e. in $O(L_{s1} M)$steps, yielding an overall computation expense of $C_{mulitres}$, given by

$$C_{mulitres} = O(L_{S1} M + L_{S1} \log(L_{P1})) \tag{26}$$

$$= O(L_{S1} \log(L_{P1})) \tag{27}$$

When compared to the computational expense for a direct fast overlap-add convolution implementation of the high-resolution search, $O(L_s log L_p)$, the multi-resolution approach leads to a reduction in the computational effort by a factor of L_s/L_{s1}, i.e. by a the downsample factor between the highest and lowest resolutions (D^R).

For the WASP system, we have a two level ($R = 2$) and a downsample factor of $D = 128$, implying theoretical savings by a factor 128, i.e., the multi-resolution approach is computationally over *two orders of magnitude* better than a direct search.

CHAPTER IV

SAMPLE RATE COMPENSATION

Due to device inconsistencies and changing environmental conditions the operating frequency of a node usually drifts from the ideal. The difference in sampling frequencies of the template and the WUs produce correlation artifacts. For relatively narrow-band signals, these artifacts may be approximated by Doppler shifts [47], [48]. Narrow band signals may be defined as those with small fractional bandwidths, i.e. when the ratio of the bandwidth of the SOI to the sampling frequency is fairly small. For the WUs, the baseband demodulated data is sampled at approximately 2 MHz, while the bandwidth of interest is less than 10 KHz implying a fractional bandwidth of less than 0.005, justifying the narrowband assumption.

Effects of sample rate mismatch

When estimating the time delay between two signals using a spectrum estimation method such the GCC, the difference in the sampling frequencies of the two signals can introduce significant errors. Figure 9 shows the effects of sampling a 20 Hz cosine wave at two different frequencies. When the sampled signal streams are compared on a sample to sample basis, the common features appear skewed in time. In particular, for the signal sampled at a higher frequency (1000 Hz) instead of at the intended frequency (500 Hz), the trough appears to be delayed.

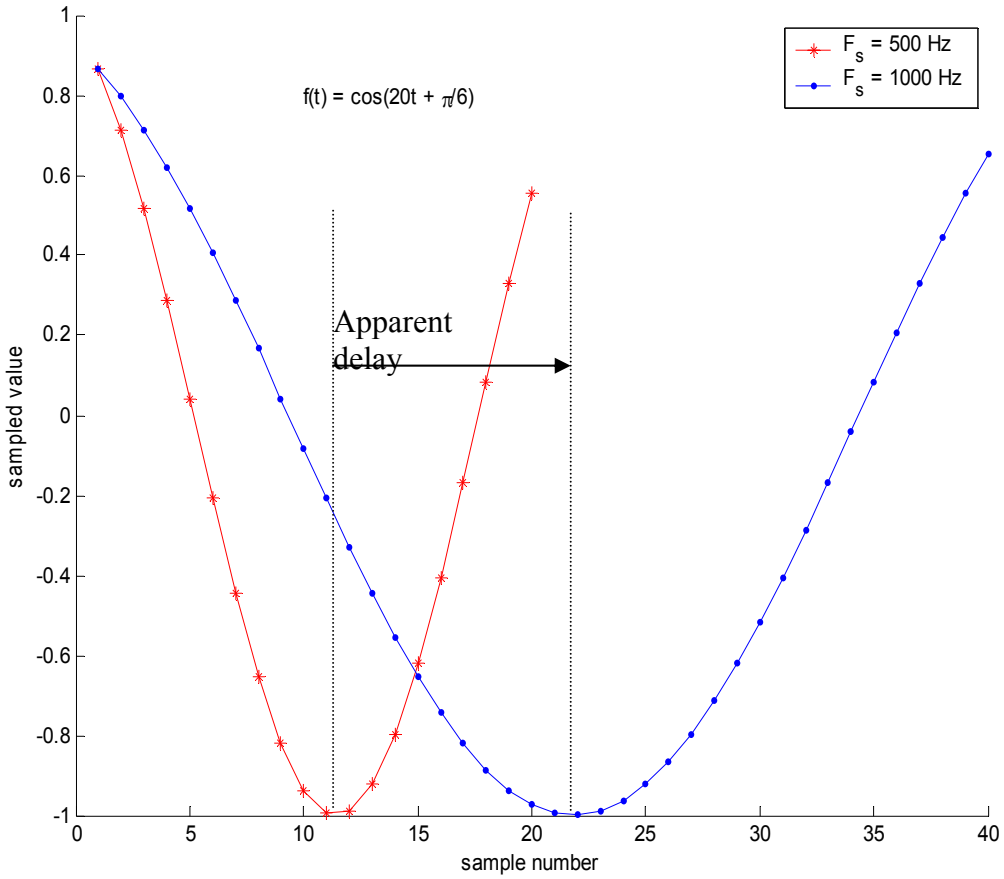


Figure 9. Effects of differing sample rates. The red curve indicates a 20 Hz cosine wave sampled at 500 Hz, while the blue curve shows the same wave sampled at 1000 Hz. Notice that the trough appears delayed by the by a factor equal to the ratio of the sampling frequencies, i.e. by a factor of 2. The situation is reversed when the actual sampling rate is slower than intended

As shown in Figure 10, a similar effect is observed when these two signals are correlated: the expected peak appears at location 32 instead of 20. Also observe that the correlation peak is smaller and flatter, representing a dilution in precision of the location of the peak. When correlating two signals with relative time scaling due to dissimilar sampling rates using GCC, the problem becomes more acute. Since GCC effectively computes the time average of successive correlograms, the resulting shape (depending on the basic correlation curve) suffers a similar dilution of precision.

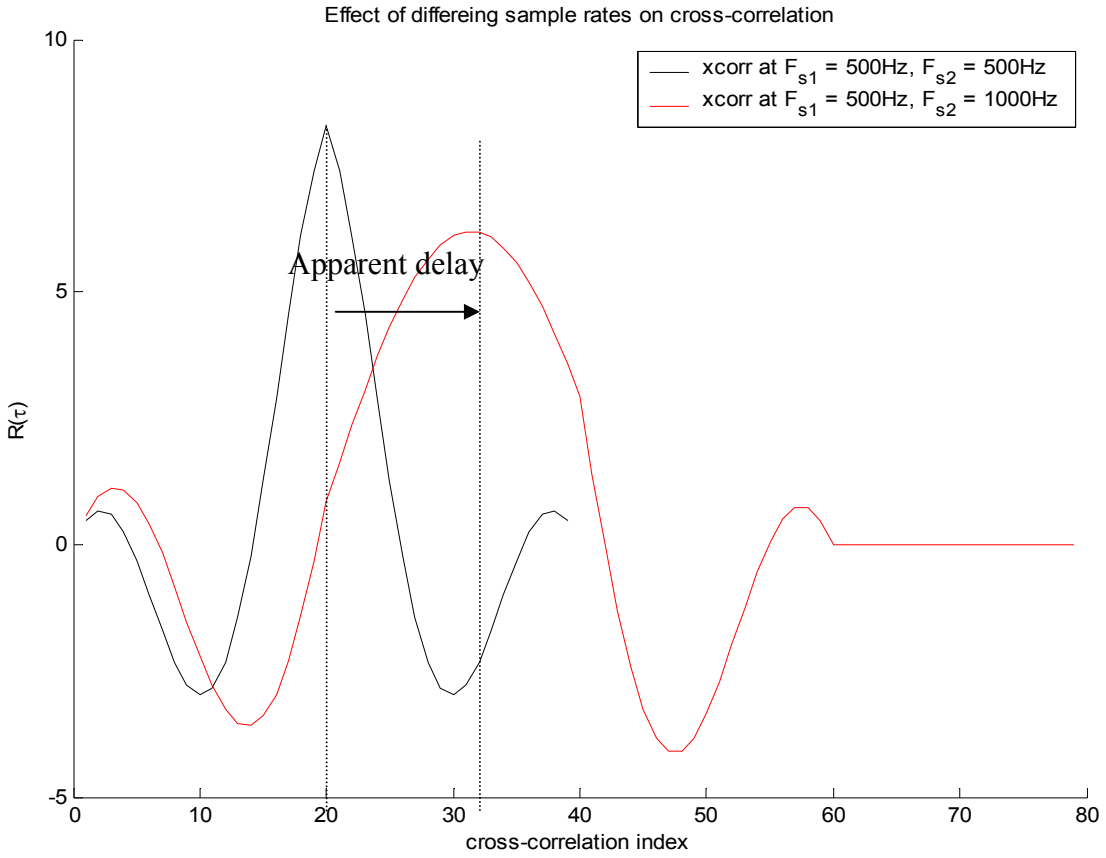


Figure 10. Comparing effects of sample rate mismatch on direct-correlation of two signals. The black curve denotes the correlation of two discrete sequences generated by sampling a 20 Hz cosine wave at 500 Hz. The red curve denotes the same when one of the signals is sampled at 1000 Hz instead. Notice the apparent delay, the increase in spread and reduction in magnitude of the peak in the black curve

Figure 11 illustrates this effect by superimposing two correlograms of signals (sampled at 2 MHz and 2.2MHz) and their mean. For clarity, the intermediate correlograms are ignored. The mean curve (shown in dashed black) exhibits a distinct deviation from its original location (shown in red), indicating a spurious delay. The clocks on the WASP units have been observed to vary as much as 350 Hz for a nominal operating frequency of 48 MHz. Since the sampling clock is obtained by dividing the system (48 MHz) clock by a factor of 24, the sampling frequencies differ by up to 15 Hz. For a template length of 500 ms, this was observed to degrade the delay estimate by

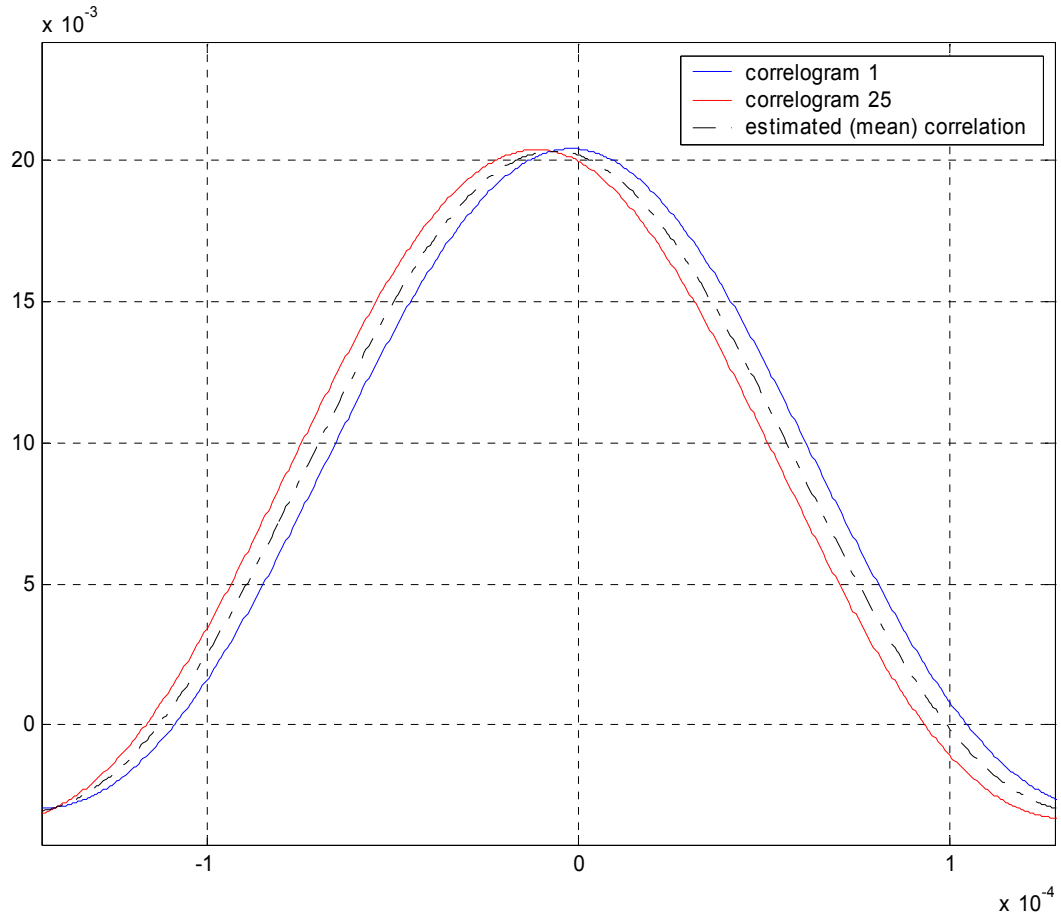


Figure 11. Estimating the cross-correlation spectrum of time companded signals using GCC. The Welch spectrum estimation method is applied to the samples of a (1 KHz – 5 KHz) band-limited pseudo random sequence, sampled at 2 MHz and 2.000020 Hz, i.e. for a difference of 20 Hz. The red curve (block #1) shows the estimated correlation curve for the first block in the time-series, while the blue curve shows the same for the very last block (block #25). The dashed black curve denotes the final estimate of the true correlation. Notice that effect of RTC for block 1 (red curve) is minimal while it has a significant delay for block 25 (blue curve). Also note the spread and apparent delay (few μs) in the final estimate (dashed black curve).

approximately 700 ns. Thus TDOAs between different units exhibit inaccuracies of a few micro-seconds. While these errors are nearly an order of magnitude lesser than depicted in Figure 11 they significantly degrade the final geolocation accuracy.

Ideally, resampling the signals to ensure a consistent sampling rate would eliminate the errors due to RTC. However, resampling is temporally and computationally expensive, and hence, an infeasible option for a resource constrained system. Instead we approximate the effects of the apparent delay in each correlogram by time shifts. Alternatively, one may attack the relative time scaling problem in the frequency domain using Doppler shifts. The following sections mathematically model these problems and develop the RTC compensation techniques.

Let the template be represented by $f(t)$ and the received signal by $g(t)$. We shall analyze relative time companding in the continuous domain as it lends itself better to analysis. Since differences in the sampling frequency manifest appears as a time scaling, we may express the SOI within received signal as a dilated and shifted version of the template, i.e.,

$$g(t) = f(s(t-\tau)) \tag{28}$$

Taking the fourier transform on both sides, we have,

$$G(\omega) = s^{-1}F(\omega / s)e^{-j\omega s\tau} \tag{29}$$

let s = 1 – a, then

$$G(\omega) = 1/(1-a)F(\omega/(1-a))e^{-j\omega(1-a)\tau} \tag{30}$$

By the binomial theorem, we have,

$$\frac{1}{1-a} = (1-a)^{-1} = 1 + a - \frac{1}{2}a^2 + \ldots \tag{31}$$

since a << 1, we have $\frac{1}{1-a} \approx 1 + a$ and hence,

$$G(\omega) = (1+a)F(\omega(1+a))e^{-j\omega(1-a)\tau} \tag{32}$$

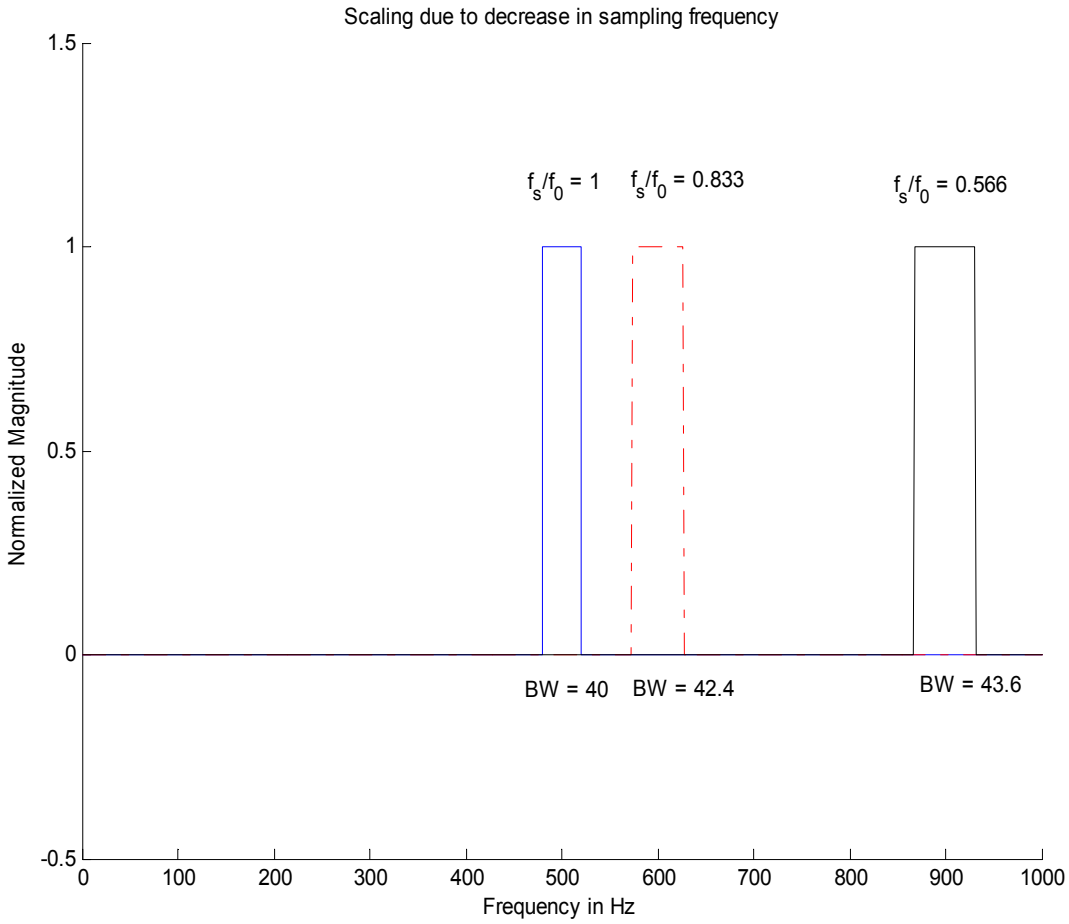


Figure 12. Approximating frequency scaling with Doppler shifts for small fractional bandwidth signals. The blue rectangle represents the spectrum of a narrow bandwidth pseudorandom sequence when sampled at the intended frequency, f_0. As the sampling frequency is reduced, the width of the rectangle scales and its position shifts closer to the Nyquist frequency. The dashed red, solid black curves indicate the spectrum when the sampling frequency is reduced by approximately 17%, 44% respectively. Note that the shift is more pronounced than the change in the width of the spectrum, suggesting that the scaling may be approximated by merely shifting the original (blue) spectrum

For narrow band signals, $F(\omega_0 + \delta\omega) = 0, |\delta\omega| > \frac{\omega_b}{2}$ where ω_0 is the center or the rms radian frequency and ω_b is the bandwidth of the signal. Hence the above expression for G(ω) may be approximated by

$$G(\omega) \approx (1+a)F(\omega + a\omega_0))e^{-j\omega\tau}e^{j\omega a\tau} \tag{33}$$

noting that both a and τ are small, so that the second exponential maybe approximated to unity, we get,

$$G(w) \approx (1+a)F(\omega + a\omega_0))e^{-j\omega\tau} \tag{34}$$

taking the inverse fourier transform, we obtain,

$$g(t) \approx f(t-\tau)e^{-j\omega_d(t-\tau)} \tag{35}$$

where $\omega_d = a\omega_o = (1\text{-}s)\ \omega_o$. The above expression indicates that a Doppler shift results whenever the two signals are companded with respect to each other. Refer Figure 12 for an intuitive explanation.

Compensating for Doppler effects due to relative time companding

Observe that the form of the time scaled narrow band signal closely resembles the kernel of the cross-ambiguity function (CAF) [45]. For our case, s is the ratio of the sampling frequencies of the two signals. Fortunately, this is always known: the sampling frequency of the template is known *a priori*, and can be accurately controlled through correct construction. The instantaneous sampling frequency for the data obtained at a WU may be estimated using the PPS alignment algorithms. Also, the center frequency of the transmitted signal may be known *a priori* by construction. For arbitrary signals, it may also be estimated as the mean frequency as defined in [45] [46] by

$$\omega_0 = \frac{\int_{-\infty}^{\infty} \omega F(\omega)\, d\omega}{\int_{-\infty}^{\infty} F(\omega)\, d\omega} \tag{36}$$

The effect of the differing sampling frequency may be modeled as a modulation with a complex exponential, as indicated above. To compute the time of arrival then, a generalized cross correlator may be used to correct for the frequency effects. As before, if $f(t)$ represents the template and $g(t)$ the received signal, then we create a new signal, $f_1(t)$ which is a "shifted" version of the template to compensate for the current operating frequency of the node.

$$f_1(t) = f(t)e^{-j\omega_d t} \tag{37}$$

signals f1(t) and g(t) are then fed through a generalized cross-correlator (GCC) to obtain an estimate for the time delay, τ.

$$\tau = \arg\max \left| \int_{-\infty}^{\infty} f_1^*(u) g(u+t) du \right| \tag{38}$$

i.e.,

$$\tau = \arg\max \left| \int_{-\infty}^{\infty} f^*(u) g(u+t) e^{j\omega_d t} du \right| \tag{39}$$

The above expression may be viewed as a cross-ambiguity function evaluated at a given shift ω_d.

Implementation issues in Doppler based relative time companding

As explained earlier, the effects of relative time companding may be ameliorated by modulating the template with a complex exponential. The angular frequency of this complex exponential ω_d, is given by

$$\omega_d = a\omega_o = (1\text{-}s)\ \omega_o \tag{40}$$

Usually, the complex exponential may be computed using Euler's expansion as shown below

$$\exp(j\omega_d t) = \cos(\omega_d t) + j\sin(\omega_d t) \tag{41}$$

For a discrete time system, the above evaluation must be performed for each sampling instant. Given a template of length N, this produces an N x 1 vector of samples of the modulating complex exponential

$$\mathbf{e} = \cos(\omega_d \mathbf{t}) + j\sin(\omega_d \mathbf{t}) \tag{42}$$

where, $\mathbf{t}[k] = k / \omega_s$ is the vector of sampling instants, with $0 \le k \le N, k \in \mathrm{Z}$ and sampling frequency ω_s. Computing the elements of this complex exponential in terms of trigonometric functions is both computationally and temporally intensive. Furthermore, once the vector has been computed, its application to the real template vector requires N complex multiplications. Examining the structure and nature of this calculation and applying a few engineering assumptions, we can significantly reduce this burden.

For the present case, from the local oscillator crystal specifications [37], it is known that the maximum deviation of the oscillator from its nominal frequency is ±100 ppm (upto ±200 ppm have been observed in lab-experiments). Thus the value of s is fairly close to 1, or equivalently, α is fairly small. Also, the time period over which the Doppler correction must be employed is limited a small fraction of a second (If it is large, then the time-bandwidth product condition for narrow band signals is violated, making the Doppler shift assumption invalid). Thus the product $\omega_d t$ is fairly small (nominal value of 0.02). Using the basic trigonometric approximations for small radian angle θ,

$\cos(\theta) \approx 1$ and

$$\sin(\theta) \approx \theta$$

we can make the following approximation

$$\exp(j\omega_d t) = 1 + j(\omega_d t) \tag{43}$$

Expressing this vectorially, we have for the discrete case

$$\mathbf{e}[k] = 1 + j\omega_d \mathbf{t}[k] \tag{44}$$

This computation represents enormous savings in creating the correction vector, e. The structure of each **e**[k] also enables a very efficient computation routine. The Doppler shifting of the template may be accomplished in a single pass, leading to an O(N) computation routine, as shown in Appendix A.

Alternative time shift based correction

While the Doppler correction has the advantage of offering a very intuitive and simple technique of compensating for the different sampling frequency between the nodes and the templates, it suffers from a computational perspective. The TI C67x DSPs used to provide the signal processing have a limited internal cache. When implementing computationally and temporally intensive operations such as a generalized cross correlation, it is imperative that the memory access requirements be as limited as possible to meet performance constraints. Ensuring that the data and result vectors fit in the fast internal memory is a proven and routinely employed means towards this end.

Both the Doppler correction and approximation techniques suffer in this regard because they convert normally real data vectors into complex vectors via the Doppler frequency correction. This doubles the initial memory requirement and simultaneously

causes performance degradation which, accrued over several iterations, can mean the violation of a real-time constraint. Such a situation occurs in the computation of the cross spectral density (CSD) of the received signal with the template.

To search for an alternative means of ameliorating the effects of relative companding due to incongruent sampling rates, let us examine the relative companding problem again. As before, let

$$g(t) = f(s(t-\tau)) \tag{45}$$

where, $g(t)$, $f(t)$ are, respectively, the received and template signals, s is the scaling factor and τ is the time delay introduced. As explained earlier, the presence and location of the template (FOI) in the input stream is estimated using a matched filter. The equation below recalls the matched filtering operation. The * denotes complex conjugation, which for real-signals is inconsequential. However, for the sake of generality, we shall assume complex data envelopes.

$$R_{gf}(u) = \int_{-\infty}^{\infty} g^*(t) f(t+u) dt \tag{46}$$

τ is estimated as

$$\tau = \arg\max \left| \int_{-\infty}^{\infty} g^*(t) f(t+u) dt \right| \tag{47}$$

when $s = 1$, a value of $t = \tau$ maximizes R_{gf}. For the present case (s>0), the value of t that maximizes R is given by,

$$\tau_m(t) = \arg\max \left| \int_{-\infty}^{\infty} f^*(st - s\tau) f(t+u) dt \right| \tag{48}$$

$$\text{i.e. } \tau_m(t) = u \quad \text{such that } s(t+\tau) = (u+t) \tag{49}$$

$$\therefore \tau_m(t) = (s-1)t + s\tau \tag{50}$$

Notice that this delay $\tau_m(t)$ is not fixed as in the previous case, but is "smeared" in time.

We wish to approximate this time varying delay $\tau_m(t)$ by a constant τ_d. We also wish to choose τ_d according to a least squared error criterion, i.e., the sum of the squared errors, $\in(\tau)$, given by

$$\in(\tau_d) = \int_{-\infty}^{\infty} e^2(t)dt = \int_{-\infty}^{\infty} \left(\tau_m(t) - \tau_d\right)^2 dt \tag{51}$$

is minimized. For finite length sequences, the above may be rewritten as

$$\in(\tau_d) = \int_{T_1}^{T_2} e^2(t)dt = \int_{T_1}^{T_2} \left(\tau_m(t) - \tau_d\right)^2 dt \tag{52}$$

where T is the interval over which the two sequences are compared. This is typically the case when a data stream is processed in chunks or blocks of finite length. Proceeding in the usual manner, we search for a value of τ_d that minimizes the above error function. Using the Leibniz integral rule [50], we have

$$\frac{\partial}{\partial \tau_d} \in(\tau_d) = \int_{T_1}^{T_2} \frac{\partial}{\partial \tau_d} \left(\tau_m(t) - \tau_d\right)^2 dt + \left(\tau_m(T_1) - \tau_d\right)^2 \frac{\partial}{\partial \tau_d}(T_1) - \left(\tau_m(T_2) - \tau_d\right)^2 \frac{\partial}{\partial \tau_d}(T_2) \tag{53}$$

which vanishes for the optimal value, $\tau_d{}^o$. i.e.

$$\frac{\partial}{\partial \tau_d} \in(\tau_d)\Big|_{\tau_d = \tau_d{}^o} = \int_{T_1}^{T_2} 2\left(\tau_m(t) - \tau_d\right)dt\Big|_{\tau_d = \tau_d{}^o} = 0\,, \tag{54}$$

$$\therefore \tau_d{}^o = \frac{1}{T_2 - T_1}\int_0^T \tau_m(t)dt = (s-1)\frac{T_2 - T_1}{2} + s\tau \tag{55}$$

Over a given interval T, the effects of relative companding between $f(t)$ and $g(t)$ may be minimized in the least squares sense by constructing a time-shifted version of $f(t)$, denoted $\widetilde{f}(t)$, as

$$\widetilde{f}(t) = f(t - \tau_d^{\,o}) = f(t - (s-1)\frac{T}{2} - s\tau) \tag{56}$$

In the discrete domain, the above equation may be expressed as

$$\widetilde{f}[n] = \widetilde{f}(\frac{2\pi n}{\omega_s}) = f(\frac{2\pi n}{\omega_s} - (s-1)\frac{T}{2} - s\tau), \tag{57}$$

where ω_s is the angular sampling frequency.

For most DSP applications, the analog signal $f(t)$ is usually not available to permit arbitrary time-shifting of the template signal. $\widetilde{f}$ would then need to be evaluated by interpolating at "arbitrary" time instants. While this is achievable, the interpolation in the time domain usually comes with a computationally expensive price tag.

An easier alternative may be found by migrating to the frequency domain. Taking the fourier transform of the preceding equation, we obtain

$$\widetilde{F}(\omega) = F(\omega)e^{-j\tau_d\omega} \tag{58}$$

where $F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt$ is the Fourier transform of $f(t)$. Also note that we have dropped the superscript of τ_d in favor of brevity. The above expression is computationally more convenient because it allows for a pre-computed version of the Fourier transform of the template F(w) to be used. As in the Doppler correction, it can be argued that the nature of τ_d and ω allow the correction factor $e^{-j\tau_d\omega}$ to be computed as an approximation

$$e^{-j\tau_d\omega} = 1 - j\tau_d\omega \tag{59}$$

This time-shift based approach has the following notable advantages

a) Smaller memory foot print for $F(\omega)$ computation and reduced memory latency: For real signals, the Fourier transform exhibits Hermitian symmetry. This may be exploited to reduce the number of computations required to compute the entire FFT. It also means that the input vector to the FFT routine is a real-vector, requiring half as much memory as its complex counterpart. As explained earlier, a smaller memory footprint facilitates the containment of the entire signal in the fast internal memory (64KB) of the DSP chip, reducing memory access times.

b) Ability to precompute $F(\omega)$: since the template is known by design, its frequency content $F(\omega)$ is also predetermined. $F(\omega)$ may be computed offline during the design, thus obviating its computation at run-time entirely. This can, needless to say, introduce huge savings in the number of computation cycles.

c) Facility to efficiently implement correction in the frequency domain: As with the Doppler shift correction, we can utilize the nature of the correction to devise an efficient method of applying the correction in the frequency domain. This has the added advantage of allowing an *in situ* computation, bringing further reduction in memory access latencies and computation time. Appendix C lists the C functions that implement these corrections.

Results

The sample rate compensation techniques were implemented and tested in MATLAB. The delays estimated by the "standard" uncompensated GCC (UG) were

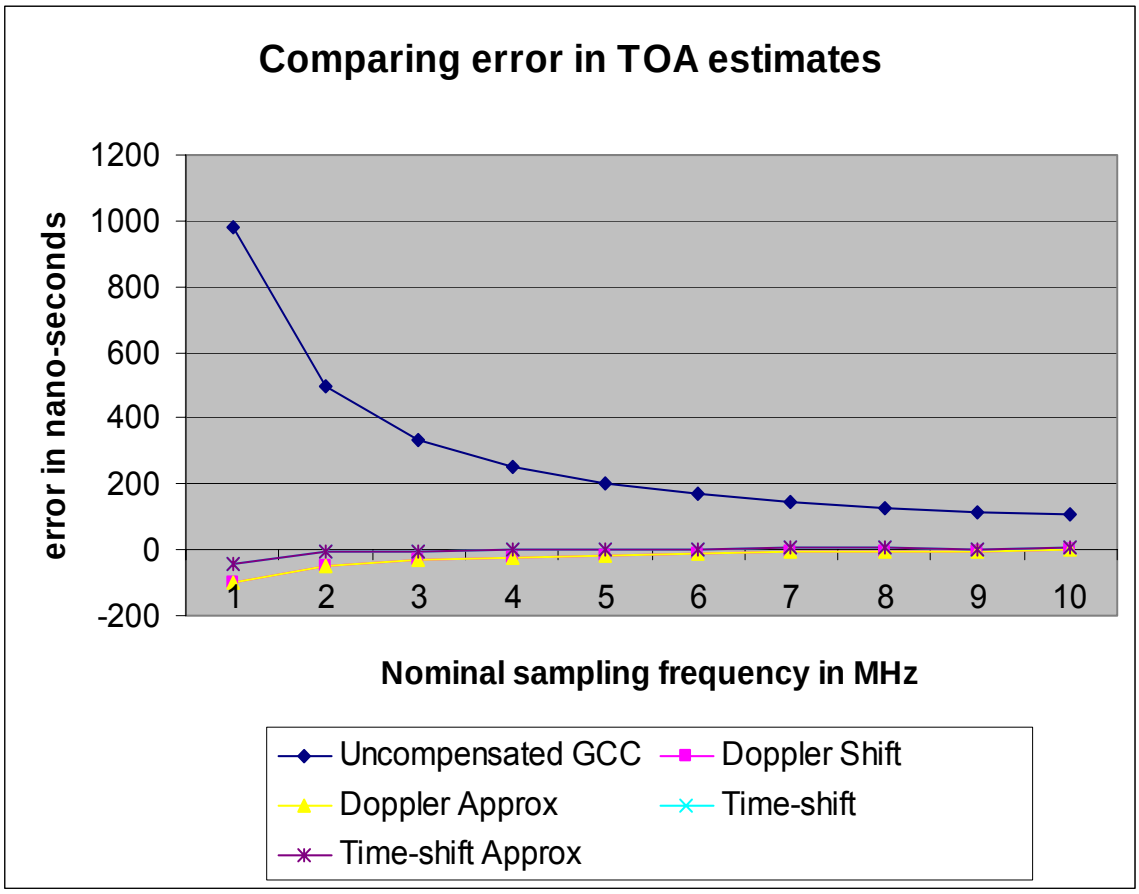


Figure 13 Comparing effects of different sampling frequencies on errors in predicted delays. A pseudorandom sequence was sampled at different nominal frequencies. Notice the decrease in the uncompensated GCC prediction errors as the sampling frequency is increased.

compared against the delays reported by the Doppler shift (DS), Doppler shift approximation (DA), time-shift (TS) and time-shift approximation (TA) compensated GCC implementations. The following factors were recognized as important to the performance of the algorithm:

- Nominal sampling frequency (FS)
- Difference between sampling frequencies of the two signals being compared (DELTA)
- Bandwidth of the SOI (BW)

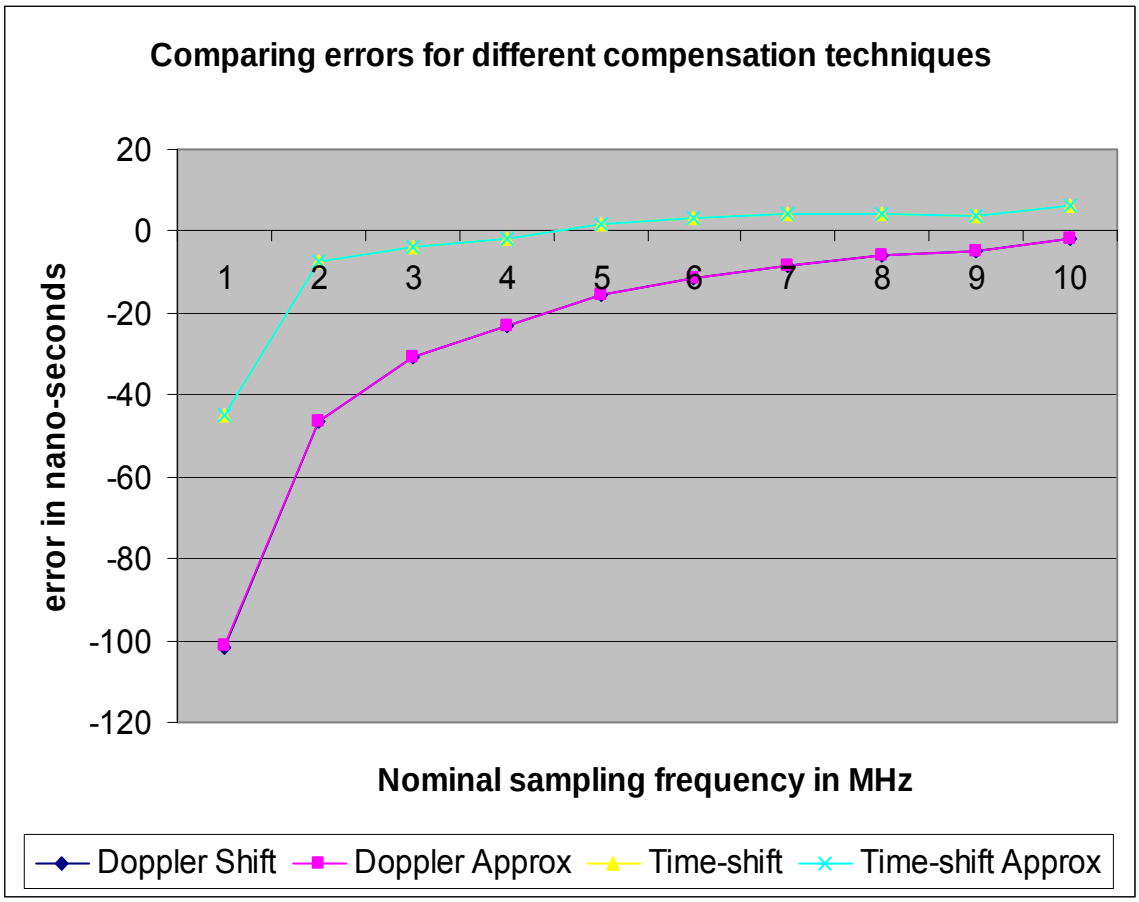


Figure 14 Comparing effects of different sampling frequencies on the different compensation techniques. Notice that the errors in the Doppler shifting techniques reduce as the sampling frequency is increased.

Table 5. Default parameters used for generating PR test sequences. Only one of these parameters was varied at a time to study its effect on the errors in delay estimation

1.	Nominal sampling frequency (FS)	2 MHz
2.	Difference between sampling frequencies of the two signals being compared (DELTA)	10 Hz
3.	Inherent time lag between the two signals being compared (DELAY)	4000 ns
4.	Bandwidth of the SOI (BW)	5 KHz (1 KHz – 6 KHz)
5.	Duration of the signals being compared (DURATION)	0.4 seconds
6.	Number of sinusoids in PR sequence	42
7.	Signal to noise ration (SNR)	20 dB

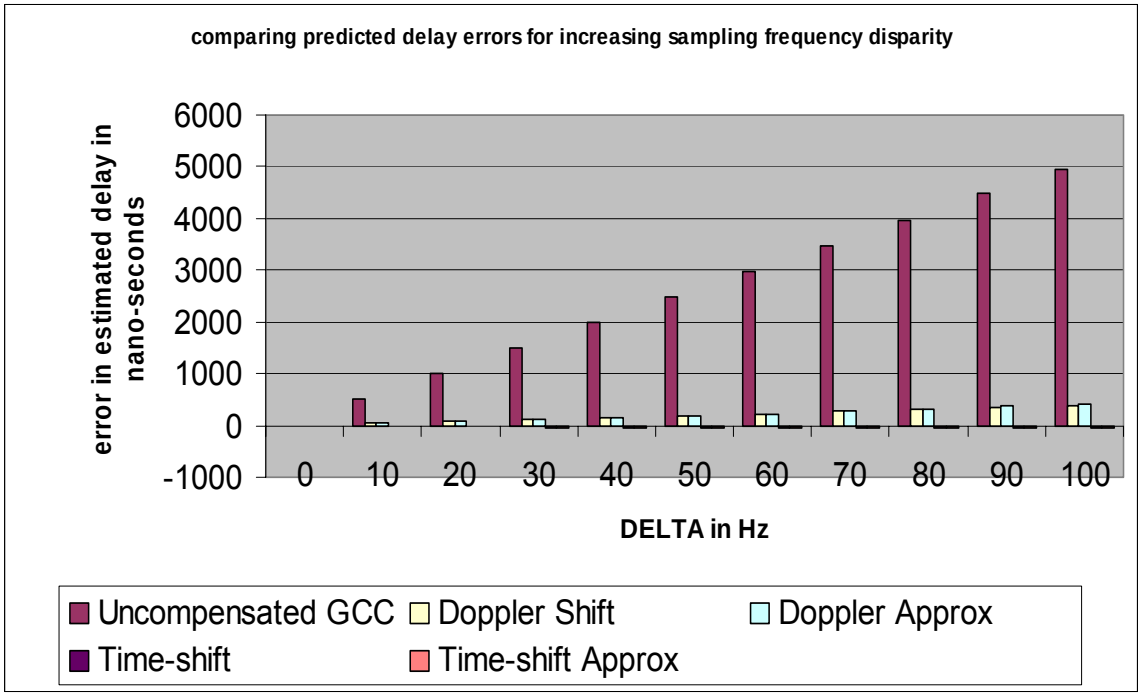


Figure 15 Comparing error in delay estimates when the relative companding between two sequences is increased. Note that the magnitude of the error in the uncompensated GCC version increases almost linearly. Overall, the compensation techniques deliver atleat an order of magnitude better performance than the uncompensated technique

Since the effects of these parameters are strongly coupled, the effects were best observed by varying the parameters in accordance with the "response surface" [51] testing technique. The tests were conducted using pseudo-random sequences generated a specified nominal sampling frequency (FS), bandwidth (BW) and duration (DURATION). The two generated sequences (y_1 and y_2) were identical except for the one controlled characteristic. The results for the different tests are shown in Figure 13 through Figure 20. The default values for the different parameters are listed in Table 5.

From Figure 13 it is evident that the errors due to relative companding decrease as the nominal sampling frequency increases provided the difference in the two sampling frequencies remains constant. This is expected because with an increase in the nominal

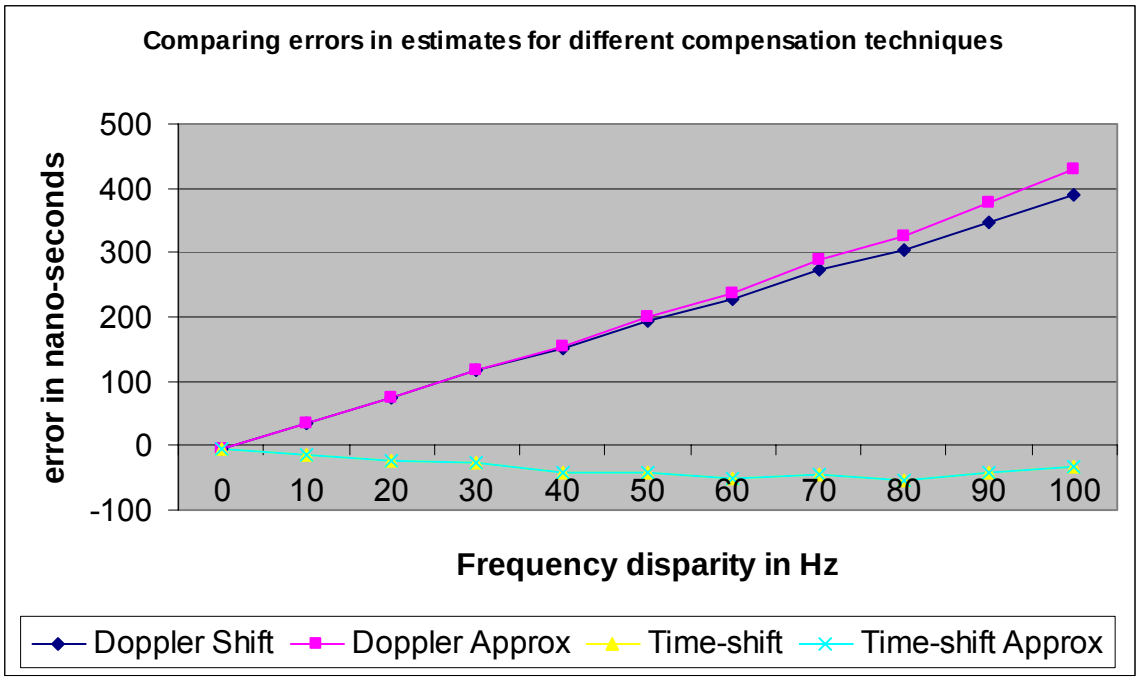


Figure 16 Errors in compensation techniques due to frequency disparity. The Doppler shifting and its approximation deteriorate almost linearly with increasing frequency disparity. The time-shifting and time-shift approximation techniques also suffer a similar degradation in performance albeit at a much lower rate in the opposite direction.

sampling frequency, the ratio of the two sampling frequencies, (the scale factor) $s = fs_2/fs_1$, is closer to unity and hence the optimal time delay, τ_d° in equation (55) is closer to the actual delay between the two signals. From another perspective, the signal streams must be longer for effects of the relative scaling to become pronounced. Notice that the errors due to the uncompensated GCC are nearly an order of magnitude greater than with any of the compensation techniques. Figure 14 compares the errors of just the compensation techniques.

Figure 15 compares the errors in the estimated delay when the difference in the sampling frequencies (DELTA) increases. From the bar charts, it is evident that the performance of the uncompensated GCC degrades almost linearly with an increase in DELTA. This degradation is as per expectation because as DELTA increases, the effects

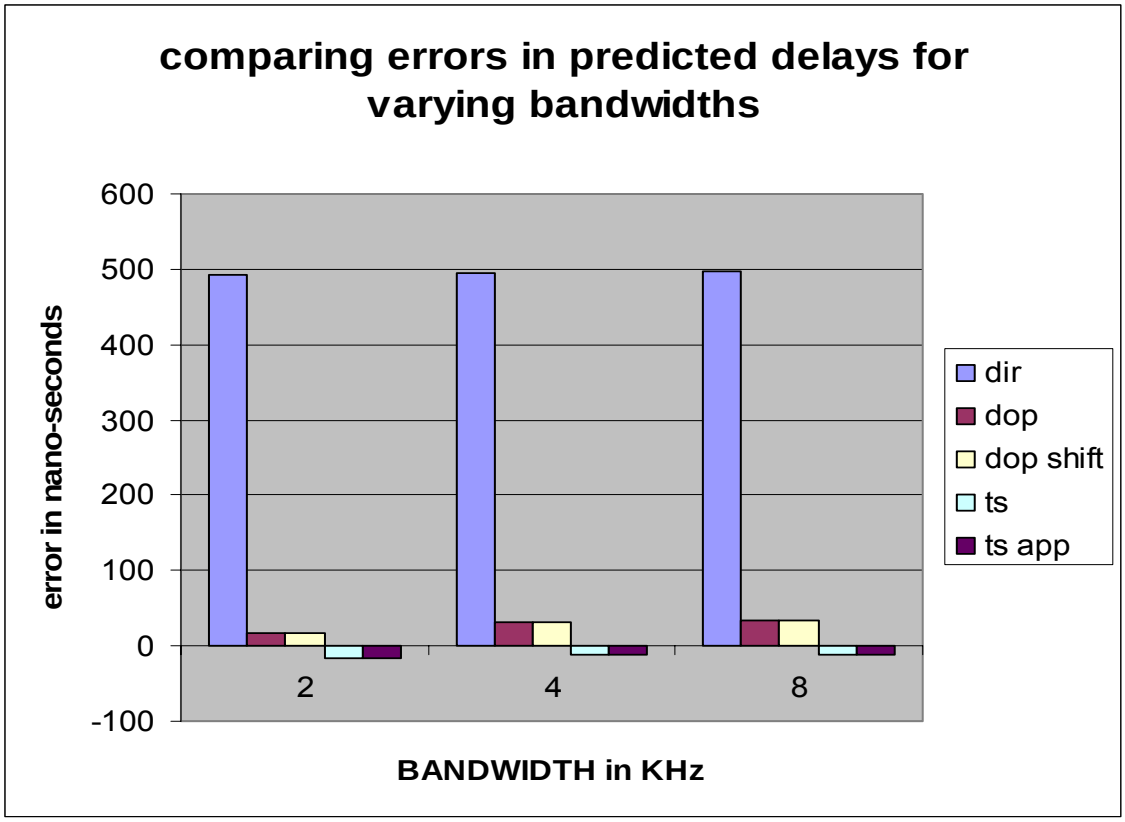


Figure 17. Effect of bandwidth on delay estimation accuracy. The bandwidth of the generated PR sequence is varied, while the other parameters are held at their default values. Notice that uncompensated GCC error is nearly independent of signal bandwidth, indicating that the errors are more sensitive to the frequency disparity than the signal bandwidth.

of RTC become more visible. The compensation techniques too demonstrate such a trend though with a much smaller magnitude. This effect examined better in Figure 16

For the case of increasing bandwidths, it was observed that performance of the uncompensated GCC (UG) remained consistently poor (around 500 ns), while that for the Doppler shift (DS) and Doppler approximation (DA) techniques worsened slightly. This is not unexpected because as the bandwidth increases, the effects of frequency scaling become more dominant and the essential narrow bandwidth assumption is violated. The Doppler shift technique performs best when the signal is a single sinusoid. Figure 17 illustrates this point. The performance of the TS and TA techniques on the other hand

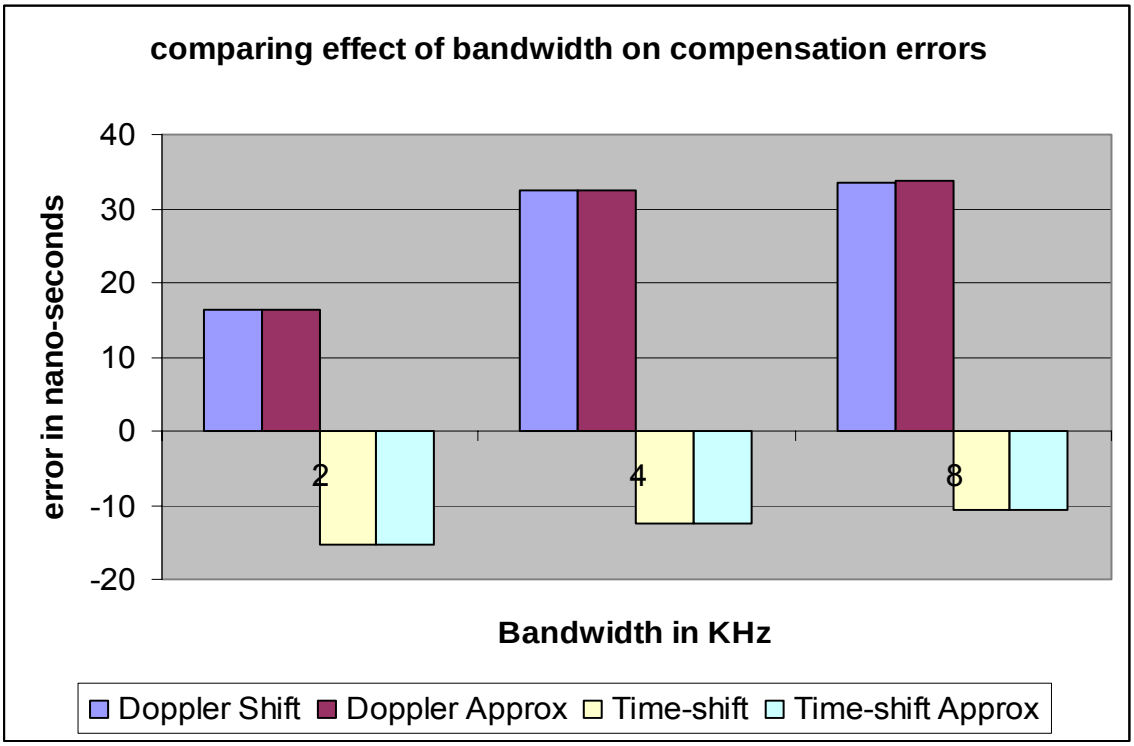


Figure 18 Effect of bandwidth on delay estimation accuracy of the compensation techniques. Notice that the Doppler compensation techniques suffer while the time-shifting techniques improve with increasing bandwidth.

show a marginal improvement as the bandwidth of the SOI is increased, as shown in Figure 18. This may be explained by recognizing that due to the uncertainty principle, the individual correlograms (correlation of individual blocks) within the GCC become sharper and narrower as the BW is increased. Thus when compensated for RTC, they "align" better, producing a lower variance estimate of the true peak.

Figure 19 compares the performance of the algorithms as the frequency for a single sinusoid with fixed delay is varied. The graph shows that the UG implementation produces a consistent error of around 500ns for a DELTA of 10 Hz. The TS and TA algorithms show reasonable performance in the 1 kHz – 30 KHz region and start to demonstrate a slow, but nearly linear deviation with increasing signal frequency. The DS

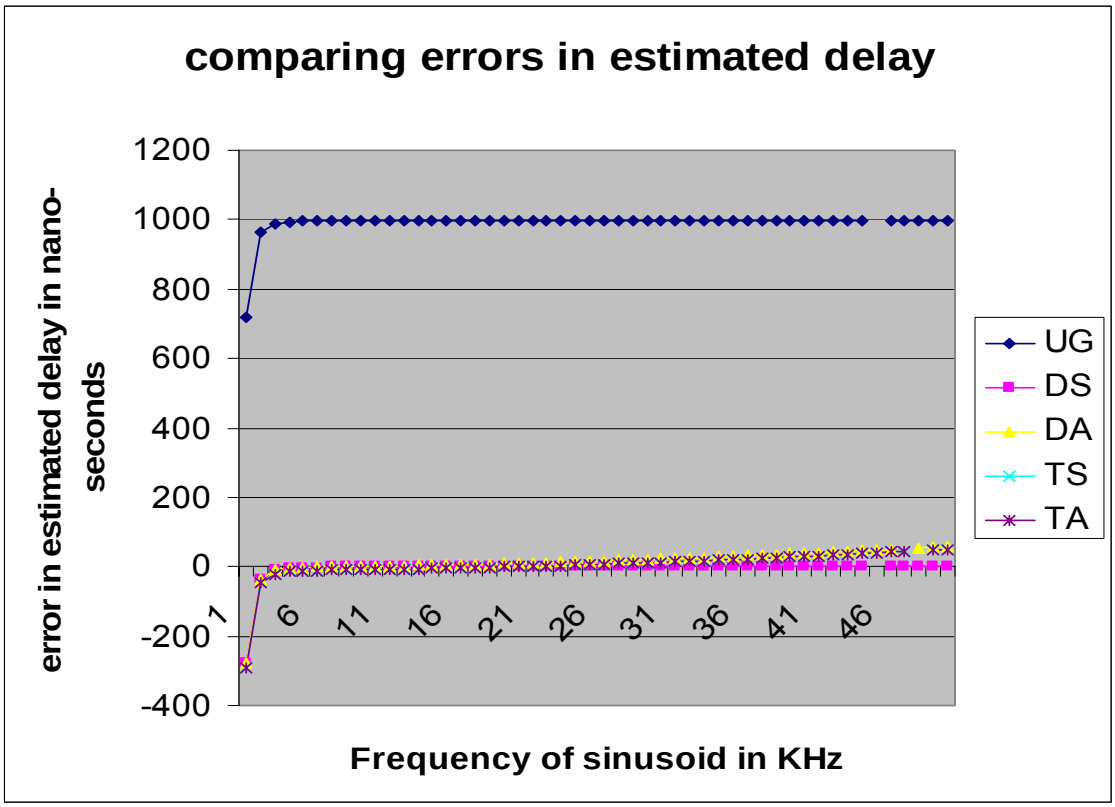


Figure 19. Comparing errors in delay estimation in the presence of RTC when the SOI is a simple sinusoid. Note that the Doppler compensation techniques outperform the others. This is a consequence of the SOI being a simple sinusoid, which in the frequency domain represents a delta function, i.e., a function with vanishing support. This case is the dual of the previous test in which the increasing bandwidth improved the performance of the time-shifting techniques

and DA techniques however perform much better because the signal bandwidth is the narrowest possible. Figure 20 compares just the compensation techniques.

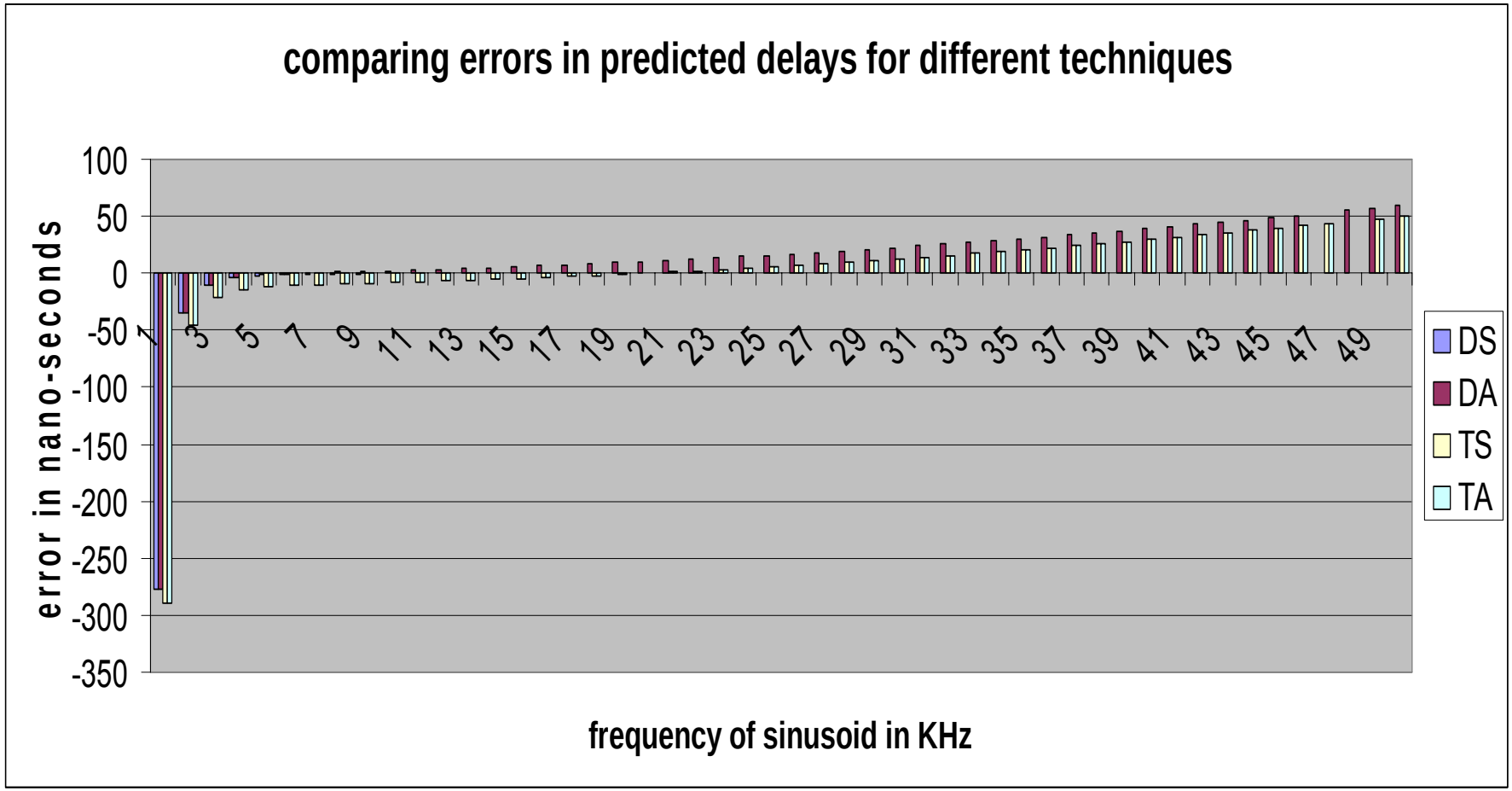


Figure 20. Comparing only the compensation techniques in the presence of RTC when the SOI is a simple sinusoid. Note that the Doppler correction (DS) and its approximation (DA) consistently outperform the time-shifting techniques. This is a consequence of the true narrow-band nature of the SOI. A single sinusoid represents a delta function in the frequency domain; a function with vanishing support. Hence the frequency shifting simply boils down to a heterodyning operation. The DA technique begins to suffer as the frequency of the sinusoid is increased because the approximation used for avoiding trigonometric computations for the complex exponential begins to deviate

CHAPTER V

GPS JITTER COMPENSATION

Since the WUs are by design spatially separate, they cannot have a common physical clock. A high degree of external synchronization [52] may be obtained by using the GPS clocks. The 24 satellites of the GPS constellation transmit a synchronization pulse every second. Since each of these satellites carries within it cesium atomic clocks, and are in external synchronization with each other, the pulse per second (PPS) signal provides a reliable mechanism for synchronization. However, as explained in the previous sections, the PPS signal as received from the GPS receivers on-board the WUs suffer from jitter, which must be eliminated to ensure reliable timing measurements between the different units.

Figure 21 illustrates the GPS jitter between the PPS signals received by two GPS receivers. Figure 21 was generated by feeding the PPS of receiver 1 to the X channel, which served as the reference (yellow lines). The PPS from receiver two was fed to channel Y (blue lines) and plotted with persistence (i.e. the each newer plot was superimposed over the older ones). A jitter in the range of 250 ns (±150ns) was observed.

Achieving global clock synchronization

A commonly used technique for correcting timing jitter is the least mean squared (LMS) fit. A “history” of the time difference between consecutive PPS pulses is created using a 32-bit counter. Each WU has a high-resolution (hi-res) clock that operates at a

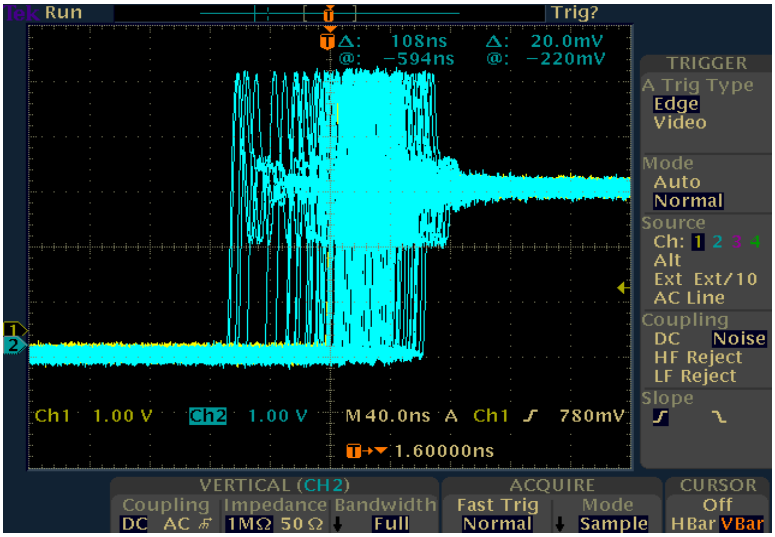


Figure 21. Screen capture of an oscilloscope showing the extent of the PPS jitter. PPS trains from independent GPS receivers were plotted against each other. The channel one (yellow) was used as the trigger, against which channel two (blue) was plotted with persistence, i.e., the plot updates were overlayed with the ones already displayed. The horizontal spread (~150 ns) of the blue plot gives an estimate of the maximal jitter possible

nominal frequency of 48 MHz. This provides the triggering clock for the various on-board electronics after necessary division and shaping. The counter counts the number of clock pulses between subsequent GPS synchronization pulses. The Figure 23 shows the schematic of the PPS correction operation.

With a clock frequency of 48 MHz, the counter wraps around roughly every 89 seconds. The wrap-around-corrected hi-res counts produce a monotonically increasing set of values. We perform a LMS straight-line fit to this data to compute the current operating frequency as well as compensate for the GPS jitter. The parameters for the straight line fit, viz. the operating frequency (C_t) and the initial counter value, or offset (C_0) may be computed from the closed form solutions of the LMS fit. In the following sections, the equations for the LMS fit are derived and expressions are presented for the corrected PPS counts. Also, the solution for the last PPS pulse is treated as a random variable and an analysis of its variance is presented.

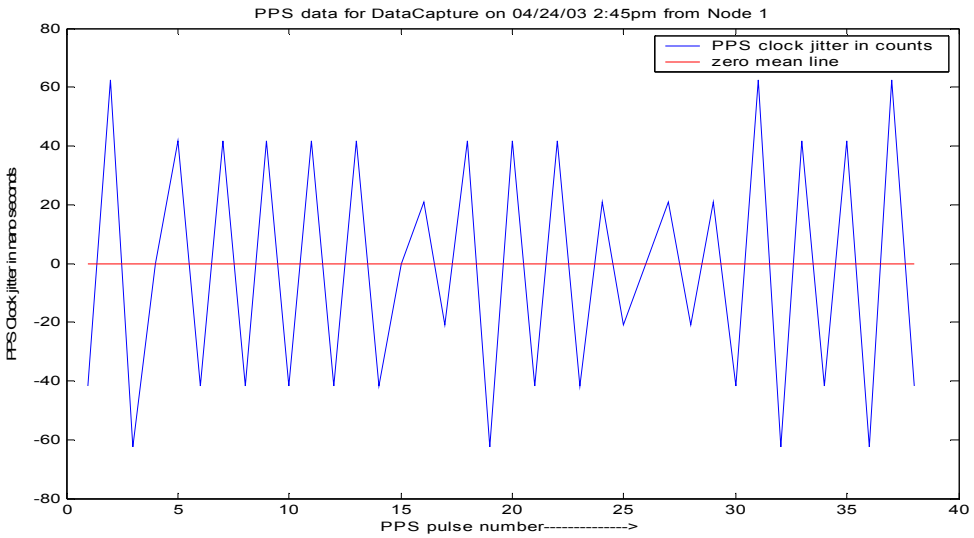


Figure 22. Time series data showing the variation in the interval between two consecutive PPS pulses

Let C_i represent the high resolution count for the i^{th} second, C_t be the actual operating frequency (high-resolution), and C_0 the offset or the initial counter value when the first PPS pulse arrives. Ideally,

$$C_i = C_0 + iC_t \tag{60}$$

where $i = 0, 1, 2, \ldots, n$.

Let the observed C_is be perturbed from the ideal by independent identically distributed (iid) zero mean Gaussian jitter, δ_i, with variance σ^2, i.e.,

$$C_i = C_0 + iC_t + \delta_i \tag{61}$$

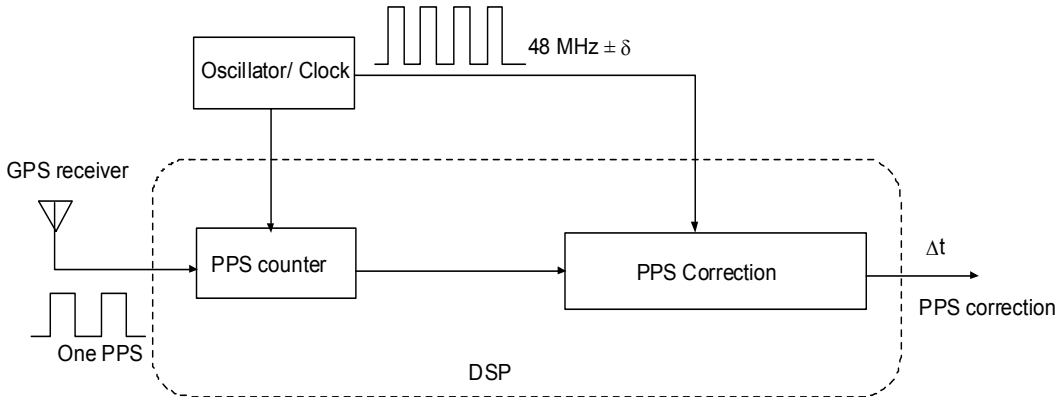


Figure 23. Schematic of PPS jitter compensation infrastructure. The number of ticks of a 48 MHz clock are counted between each consecutive PPS pulse. This builds a history of the durations of each second in terms of the local oscillator frequency. A straight line is fitted to this history. The slope of the line provides an estimate of the local frequency, while the points along the line provide the de-jittered estimates of the actual events. The final time of arrival is extrapolated from the last recorded, de-jittered PPS

Assuming a history of N pervious hi-res counter counts, and $n = N\text{-}1$ with C_i is the his-res count for the i^{th} second, we have the sum of squared errors, L is given by,

$$L = \sum_{i=0}^{n} \delta_i^{\,2} = \sum_{i=0}^{n} \left(C_i - iC_t - C_0\right)^2 \tag{63}$$

The least squared error estimators of C_t and C_i must satisfy

$$\left.\frac{\partial L}{\partial C_t}\right|_{\hat{C}_t,\hat{C}_0} = -2\sum_{i=0}^{n}\left(C_i - i\hat{C}_t - \hat{C}_0\right)i = 0 \tag{63}$$

$$\left.\frac{\partial L}{\partial C_0}\right|_{\hat{C}_t,\hat{C}_0} = -2\sum_{i=0}^{n}\left(C_i - i\hat{C}_t - \hat{C}_0\right) = 0 \tag{64}$$

Simplifying these equations yields, the following *normal equations* [51]

$$n\hat{C}_0 + \hat{C}_t \sum_{i=0}^{n} i = \sum_{i=0}^{n} C_i \tag{65}$$

$$\hat{C}_0 \sum_{i=0}^{n} i + \hat{C}_t \sum_{i=0}^{n} i^2 = \sum_{i=0}^{n} iC_i \tag{66}$$

Solving the above normal equations, we obtain the following estimators for the slope (current operating frequency) and offset

$$\hat{C}_t = \frac{12}{n(n+1)(n+2)}\left[\sum_{i=1}^{n} iC_i - \frac{n}{2}\sum_{i=0}^{n} C_i\right] \tag{67}$$

$$\hat{C}_0 = \frac{1}{(n+1)}\left[\sum_{i=1}^{n} C_i - \frac{n(n+1)}{2} C_t\right] \tag{68}$$

The fitted LS linear model is

$$\hat{C}_i = \hat{C}_0 + i\hat{C}_t \tag{69}$$

In order to ascertain the adequacy of the model, we wish to investigate the statistical properties of the above estimators. Consider first the slope (operating frequency), $\hat{C}_t$. The expected value of $\hat{C}_t$ is

$$E\left(\hat{C}_t\right) = E\left(\frac{12}{n(n+1)(n+2)}\left[\sum_{i=1}^{n} iC_i - \frac{n}{2}\sum_{i=0}^{n} C_i\right]\right) \tag{70}$$

Using the linearity of the expectation operator, we have,

$$E\left(\hat{C}_t\right)=\frac{12}{n(n+1)(n+2)}\left[\sum_{i=1}^{n}E\left(iC_i\right)-\frac{n}{2}\sum_{i=0}^{n}E\left(C_i\right)\right] \tag{71}$$

We note that the expected value of each PPS count is

$$E(C_i)=E(C_0+iC_t+\delta_i)=C_0+iC_t \tag{72}$$

which follows directly from our assumption $E(\delta_i)=0$. Hence,

$$E\left(\hat{C}_t\right)=\frac{12}{n(n+1)(n+2)}\left[\sum_{i=1}^{n}\left(iC_0+i^2C_t\right)-\frac{n}{2}\sum_{i=0}^{n}\left(C_0+iC_t\right)\right] \tag{73}$$

$$E\left(\hat{C}_t\right)=\mu_{C_t}=C_t \tag{74}$$

Thus (67) represents an *unbiased* estimator of the actual operating frequency under the given assumptions. Now consider the variance of $\hat{C}_t$.

$$\sigma^2_{\hat{C}_t}=E\left[\left(\hat{C}_t-C_t\right)^2\right] \tag{75}$$

$$=E\left[\left(\hat{C}_t\right)^2\right]-C_t^2 \tag{76}$$

After some algebra, we get the result

$$\sigma^2_{\hat{C}_t}=\frac{12\sigma^2}{n(n+1)(n+2)} \tag{77}$$

Following a similar approach, we get

$$E\left(\hat{C}_0\right)=\mu_{C_0}=C_0 \tag{78}$$

$$\sigma^2_{\hat{C}_0}=\sigma^2\left[\frac{1}{n}+\frac{3n}{(n+1)(n+2)}\right] \tag{79}$$

Note that the estimators $\hat{C}_t$ and $\hat{C}_0$ are not uncorrelated; their covariance is given by

$$Cov(\hat{C}_0, \hat{C}_t) = -\frac{6\sigma^2}{(n+1)(n+2)} \tag{80}$$

Substituting nominal values of σ = 30 nano-sec and n = 22, we get $\sigma^2_{C_t} = 0.889$ and $\sigma^2_{C_0} = 148.518$ (nano-sec)^2. The value of $\sigma^2_{C_t}$ suggests that we can track the slope fairly accurately, while the standard deviation in the estimation of the offset is approximately 12 ns. The values of C_t and C_0 are used to estimate the actual time of occurrence of the latest GPS pulse.

Given a history of N PPS counts, the location of the Nth PPS pulse is estimated as

$$\hat{C}_N = \hat{C}_0 + (N-1)\hat{C}_t \tag{81}$$

Since C_t and C_0 are computed as a function of the observed PPS counts, which are themselves independent random variables, C_n is also a random variable. The mean and the variance of C_n may be computed as below. Since the observed PPS counts are assumed to be i.i.d. Gaussian random variables, the mean value of $\hat{C}_N$ is given by

$$\mu_{\hat{C}_n} = E\left(\hat{C}_N\right) = E\left(\hat{C}_0\right) + E\left[(N-1)\hat{C}_t\right] \tag{82}$$

$$= \mu_{C_0} + (N-1)\mu_{C_t} \tag{83}$$

$$= C_0 + (N-1)C_t \tag{84}$$

Thus, $\hat{C}_N$ is an unbiased estimator. The "goodness" of the estimator may be judged by its variance, $\sigma^2_{\hat{C}_N}$

$$\sigma^2_{\hat{C}_N} = E\left[\left(\hat{C}_N - \mu_{\hat{C}_N}\right)^2\right] \tag{85}$$

$$= E(\hat{C}_N^2) - \mu^2_{\hat{C}_N} \tag{86}$$

After some algebraic manipulation, we get

$$\sigma^2_{\hat{C}_N} = \sigma^2_{\hat{C}_0} + 2(N-1)Cov(\hat{C}_t, \hat{C}_0) + (N-1)^2\sigma^2_{\hat{C}_t} \tag{87}$$

$$= \sigma^2\left[\frac{1}{n} + \frac{3n}{(n+1)(n+2)}\right] + 2n\left(-\frac{6\sigma^2}{(n+1)(n+2)}\right) + n^2\left(\frac{12\sigma^2}{n(n+1)(n+2)}\right) \tag{88}$$

$$= \sigma^2\left[\frac{1}{n} + \frac{3n}{(n+1)(n+2)}\right] \tag{89}$$

which is the same as the variance of the offset. A standard deviation of σ = 30 ns corresponds to a maximum deviation of ± 90 ns (within a 99% confidence interval) in the uncompensated case. Substituting nominal values of σ = 30 nano-sec and n = 22 in (89), we obtain $\sigma_{\hat{C}_N}$ = 12.44 ns. Thus, the estimated value of the Nth PPS pulse would be within ± 37 ns of the true value, which represents a significant improvement over the uncompensated case.

To visualize the effectiveness of the jitter compensation algorithm, the LMS correction can be viewed as an FIR filter and the variation in the "output" of the filter as a metric for comparison. Figure 24 shows variation of the filter output vs. the real, raw PPS variation for a history length of 64 seconds. Note the significant reduction in the variation of the filtered output.

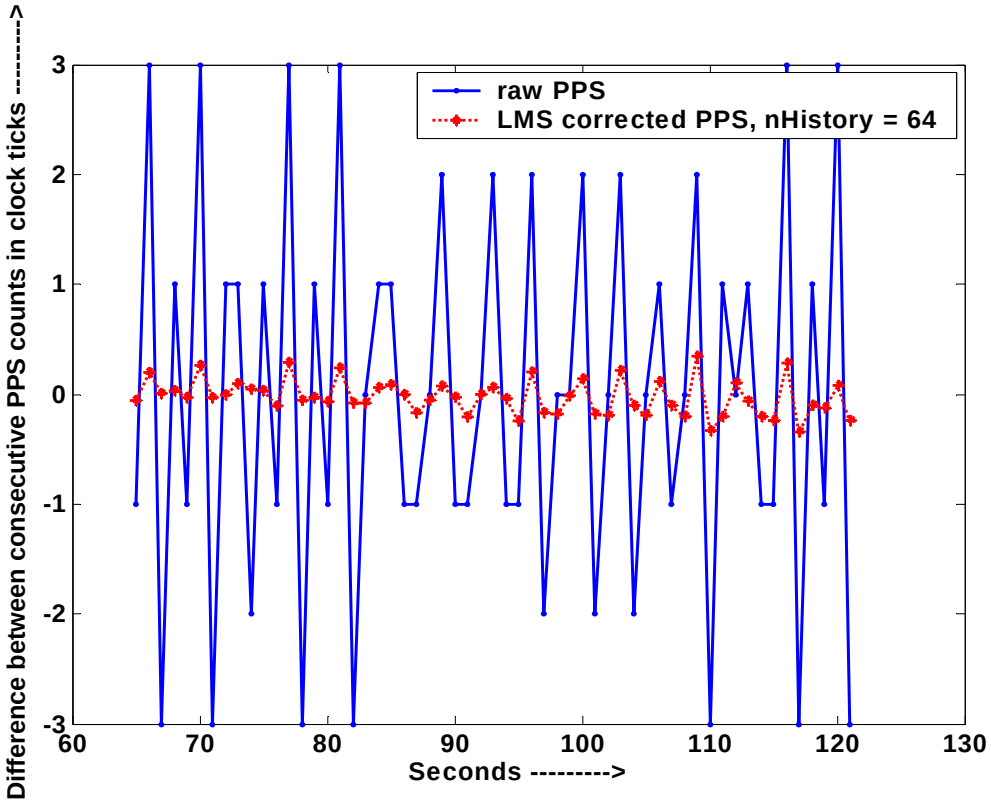


Figure 24. Effect of PPS correction. The LMS fit may be viewed as an FIR filter. the past 64 values are used to predict the next value. The dotted line with x's show the result of the "filtering" operation. Note that the maximum deviation is less than one hi-res clock count (~20 ns)

CHAPTER VI

ACTUAL SYSTEM PERFORMANCE

Field Experiments

The mobile, distributed RF geolocation system described herein is still under development. A prototype, however, was used to conduct some preliminary field tests to ascertain the feasibility of the project. Up to 8 nodes were used in several tests at Percy Warner park (PWP), Bellevue, TN. The PWP site was selected because of the large area and low ambient FRS interference levels. It also offered an elevated area in the middle of the field where the target transmitter could be placed ensuring line-of-sight to all the test nodes dispersed in the field. Two sets of tests were conducted: simple linear range estimation, and distributed geolocation.

Linear range estimation

To test the basic functionality of the system, 6 nodes were placed at increasing distances from the transmitter. The transmitter was positioned at the highest point in the field to ensure LOS to each node. This set up is depicted in Figure 25. A number of initial transmissions were made to enable a calibration of the system, so that the biases inherent to a radio-node combination could be measured and removed from the final reading. As depicted in Figure 25, the measured TOAs at each node bear a nearly linear relationship to the distance of the node from the transmitter, in accordance with theory.

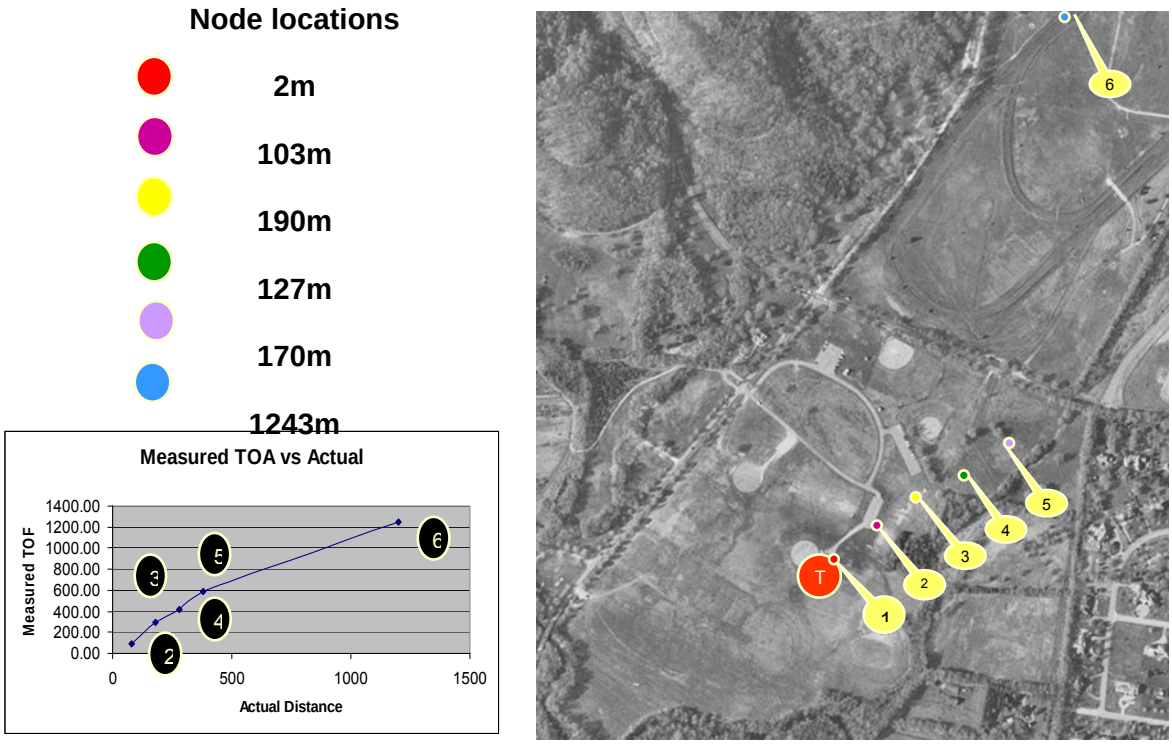


Figure 25. Linear range measurement at PWP field. The nodes (indicated by different colored dots) were placed at intervals of 100 m from the transmitter (red circle labeled 'T'). Node 6, positioned in the top right corner of the picture was placed about 1.2 Km away from the transmitter. The SOI was transmitted and the TOAs at each node were recorded and relayed back to the base-station. As shown in the adjoining graph, the measured TOAs show a nearly linear relationship with distance.

Distributed geolocation

For this test, the nodes were spread around the field in a non-collinear configuration, close to ground level. The transmitter was initially fixed at one location, where several calibration measurements were made. Once the system was calibrated, the transmitter was moved to 3 different locations within the convex hull of the node locations. The transmitter was raised several feet about the ground level to prevent ground effects. Though the nodes were actually immobile, changes in the node configuration were effected by changing the locations of the transmitter. These

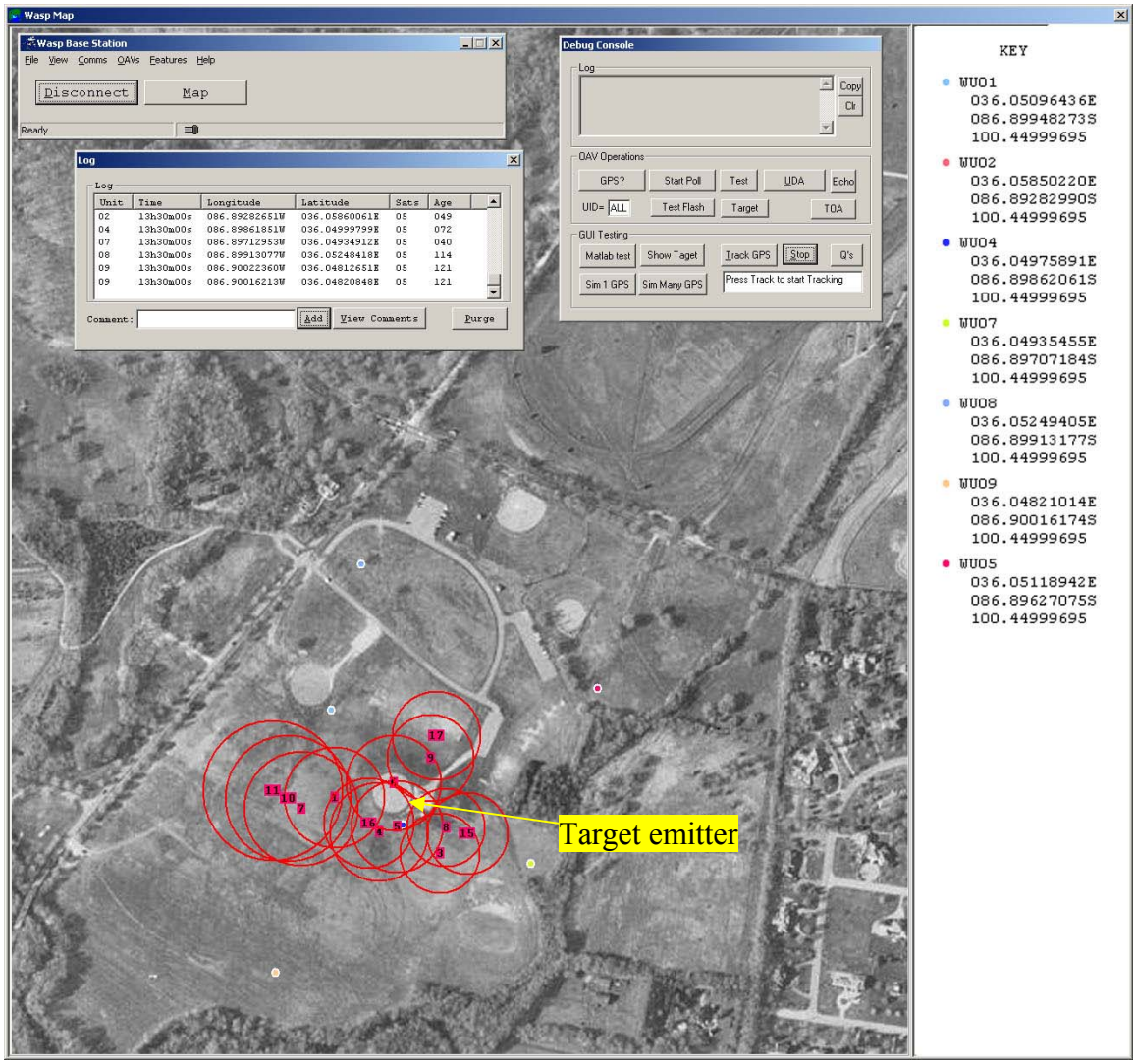


Figure 26 WASP GUI showing the calibration node configuration. The location of the units used in the test is indicated by small colored dots. The key on the right of the aerial photograph lists the latitude, longitude and elevation of the color-coded WUs.

experiments were performed to simulate the changing configuration of the nodes in response to the transmitter motion

.Figure 26 shows a snap shot of the WASP user interface (GUI) during the calibration phase. The node locations are marked by colored dots while the position of the transmitter is indicated by a number (the number indicates the transmission or test number for which the target was located). The circles around the estimated position

Figure 27 Geolocation in a low lying area. As before, the node locations are coded by the colors given on the right of the aerial photograph. The estimated transmitter locations are given by the numbers (black over yellow) at the center of the CEP circles. Notice the large errors in estimation due to non-line of sight position of the transmitter. Also notice the larger radii of the CEPs indicating low confidence in the estimates.

signify circular error probability (CEP) or the level of confidence in the estimate: the smaller the circle, the greater the confidence.

The confidence levels are obtained as a byproduct of the location estimation using the Taylor series approximation of the hyperbolic location equations [20]. Note that

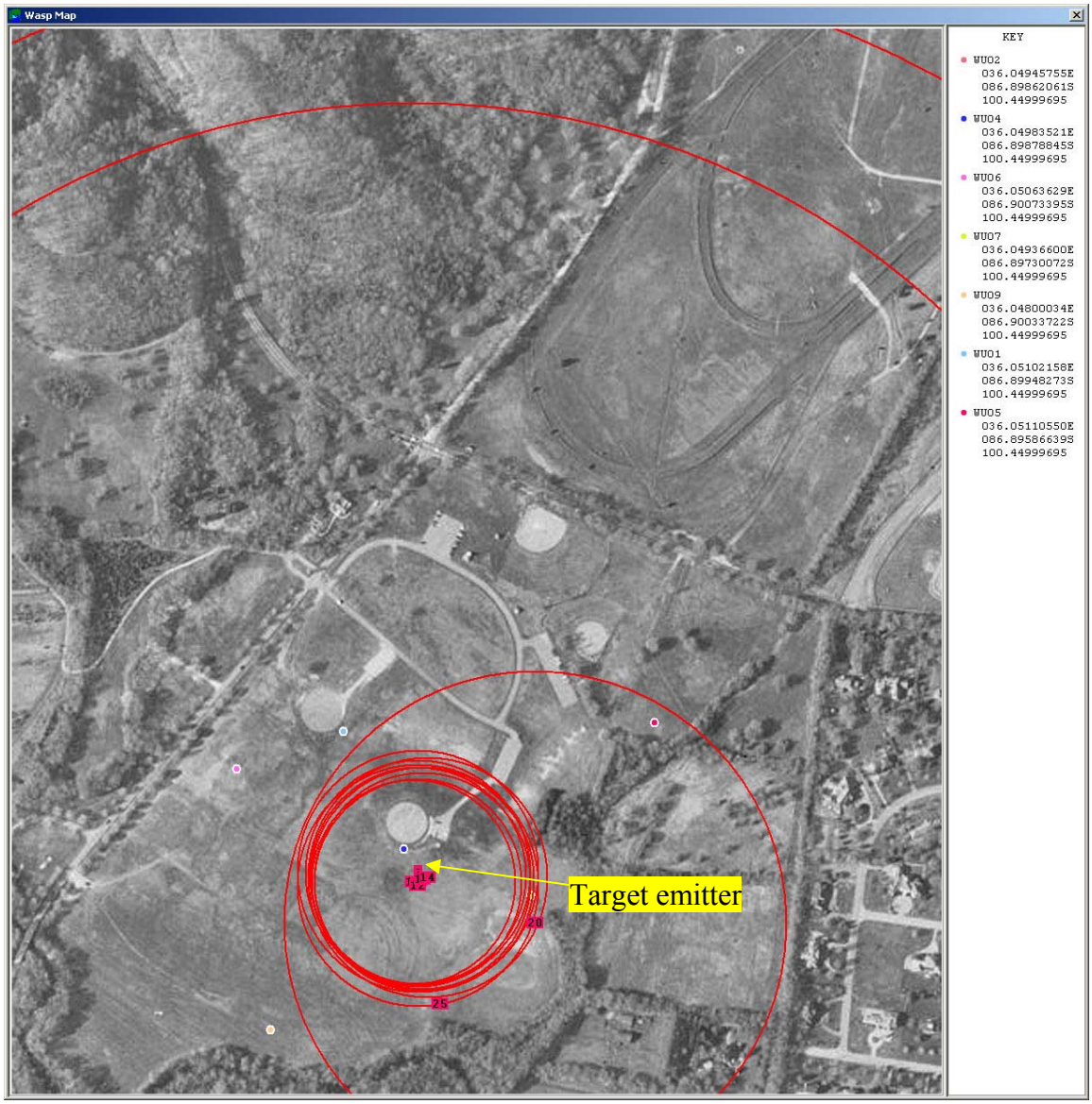


Figure 28. Geolocation at location 3

estimated positions vary quite a bit due to the inherent receiver/node biases. The actual target position is indicated by the arrowhead.

Figure 27 shows the results when the transmitter was located in a low lying area in the field. The curvature of the ground shaded the transmitter from the nodes closer to the left edge of the field. Consequently, most of the location estimates had a poor

confidence level, as evinced by the larger diameters of the CEP circles. Figure 28 shows the results when the transmitter was moved to a higher elevation. Note the improvement in the results as shown by the greater number of "concentric" circles with smaller radii.

Error contribution of compensation techniques

Since the objective of this thesis is to develop techniques to enable accurate TOA estimation for the mobile nodes, it is instructive to examine the effect of the corrections relative to the other error sources and estimate the potential performance of the system.

While there exist multiple sources of error, the significant sources are sample rate drift, time-base jitter (GPS), and the various errors in the radio reception (analog receiver front-end phase error, A/D conversion effects, and demodulation errors). Since the radio reception effects are highly complex and handset-specific, we lump these together and defer to the measured values rather than an analytical calculation. The field experiments also suggest that other factors such as varying RF channel characteristics and radio-stability can be significant contributors to the final error. Table 6 summarizes the contributions of this thesis in terms of an error-budget.

Table 6: Contribution of various error sources with and without compensation

Sources of error	Error magnitude before compensation	Error magnitude after compensation
Drifting sample rates	+/- 400 ns	+/- 45 ns
GPS Jitter	+/- 125 ns	+/- 40 ns
Radio variances (measured)	+/- 100 – 1000 ns	+/- 100-1000 ns
Total	**+/- 625-1525 ns**	**+/- 185-1085 ns**

As is evident from the table, the corrections contribute significantly to the accuracy of the system. It should be noted that the improvements in the radio variances are brought about by careful design of the receiver and demodulation sections. With a high quality demodulation, the compensation techniques developed here can reduce error from a barely-usable 625 ns to a promising 185 ns resolution.

CHAPTER VII

CONCLUSIONS AND FUTURE WORK

Conclusions

The distributed mobile radio-geolocation proposed by the WASP project was demonstrated to be feasible. Estimating the TOA at each node incorporates several challenges, which were identified and techniques were proposed to address them. Firstly, a multi-resolution approach for discriminating between signals of interest and spurious transmissions was presented. It was shown that, depending on the downsample factor, this technique reduces the computational costs involved in the discrimination operation by at least two orders of magnitude.

Secondly, the problem of drifting sample rate clocks on different units was posed as a time-scaling problem. To avoid the computational costs associated with resampling, the Doppler frequency shifting solution and a related time-based shifting technique was developed. The duality between the two approximation techniques was highlighted. It was also shown that the time-shifting techniques had a smaller memory footprint, and afforded approximations that made them computationally superior to the Doppler shifting techniques. Finally, the problems of estimating a node's operating frequency and GPS jitter problem are posed together as a linear regression problem.

The computed error budget shows that the algorithms developed here can reduce TOA errors by approximately 500 ns for a typical set of system parameters (e.g. those of the WASP system), greatly improving the geolocation accuracy.

Future work

Though the WASP system is feasible, several developments need to be made, both in the signal processing software and hardware, for it to be a robust, deployable system. The field experiments have shown that many assumptions, such as lack of multi-path or common-band interference, high SNR, deterministic radio/transmission channel transfer functions etc., required for an amenable theoretical solution are frequently violated. Consequently, the system design must evolve to address these concerns. The following problems and potential solutions have been recognized.

Multi-path interference

Multipath interference is a significant error contributor even when there exists a LOS path between the transmitter and the receiver. The cross-correlation detector is particularly sensitive to the "amplitude" of the received signals. If the delayed signals incident on the radio antenna have amplitude greater than the original, the results of the cross-correlation detector will be biased towards the bigger signal, i.e. the multi-path signal biases the delay estimates. Though the effects of multi-path may be reduced by using higher frequency and higher bandwidth signals, the inherent channel capacity and the characteristics of the FRS radios make these enhancements impractical. Super-resolution techniques such as TLS - ESPIRIT (Total Least Squares version of Estimation of Signal Parameters via Rotational Invariance Technique) [53] and Root-MUSIC

(Multiple Signal Classification) may be employed to determine and counteract the effects of multi-path [54].

Radio / RF environment characterization

Laboratory and field experiments indicate that the radios tend to introduce arbitrary phase distortions on the received signals, which distort the TOA estimates by several micro-seconds. While on-line calibration and averaging of results tends to reduce the influence of these errors, the non-deterministic nature of these distortions makes for a challenging problem. Also, the radios seem to perform better for certain audio-frequencies than for others. Thus a template designed to target the "favored" frequency ranges can provide performance improvements. However, the transfer functions of the radios are unknown and can be determined only empirically. Furthermore, the RF environment itself may change, e.g. when the transmitter or the nodes reposition themselves, leading to a different configuration of multi-path and channel noise characteristics. Knowing the channel and radio transfer functions, we may employ homomorphic and blind-deconvolution processing to mitigate the related errors [55].

Target radio

To make the system more usable, the target transmitter may be changed to a commercial GSM cell phone, implying a change in not only the frequencies of operation, but also the template. Alternatively, GSM being a digital communication system, we could completely eliminate the need for a template and instead use the symbol transitions endemic to the GMSK (Gaussian minimum shift keying) digital bitstream as features of

interest. In this scheme, each node would estimate the "global" time (TOA) of each symbol transition (a '0' changing to a '1' or vice versa) and relay the sequence of TOAs to the base-station. The base-station would then use these timing sequences to compute the "phase differences" or equivalently the time differences for each pair of nodes. The system would also need to accommodate the frequency and power-level hopping nature of the GSM system.

APPENDIX A

System Architecture: UAV Hardware Platform And Signal Processing Framework

A schematic of the resulting architecture is presented in Figure 29. As explained in the previous chapter, the target emitter is assumed to be an FRS radio transmitting frequency modulated audio-frequency data. Consequently, a commodity FRS band radio is used as a receiver.

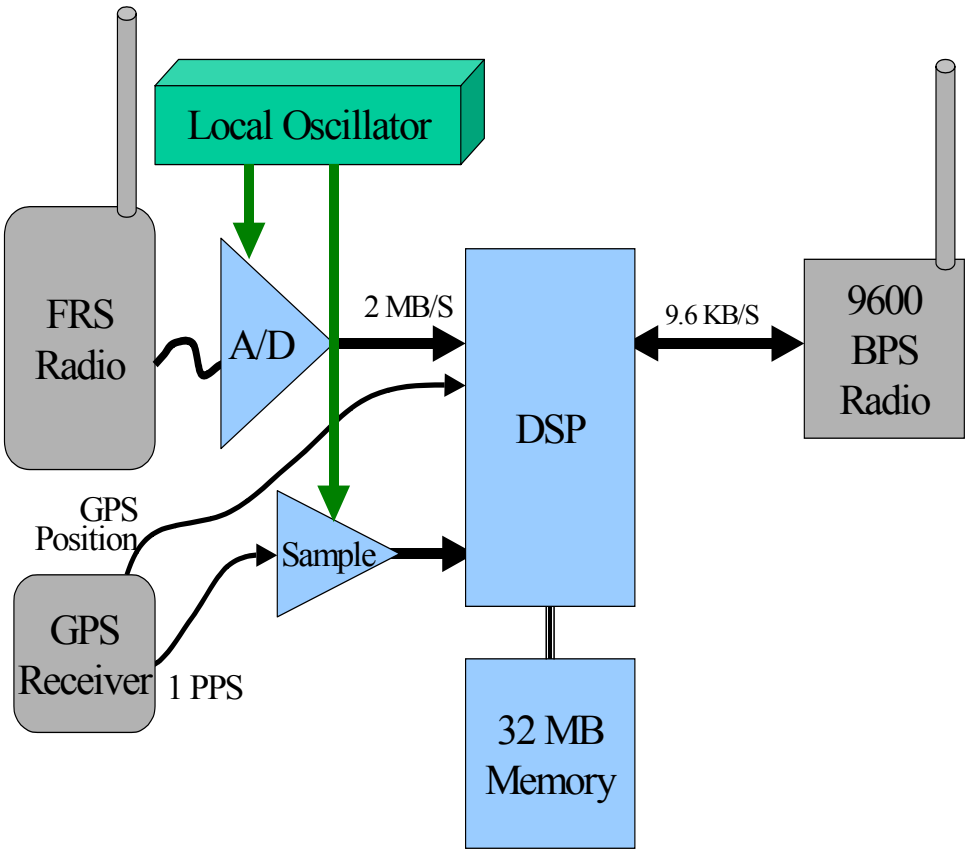
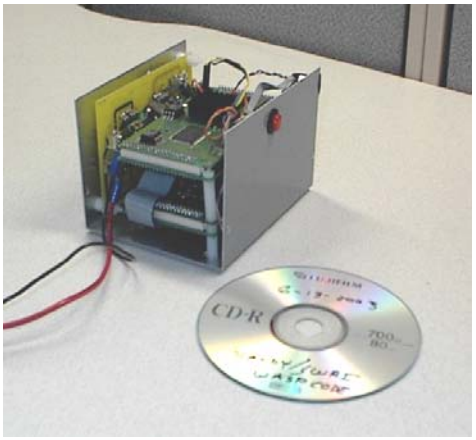
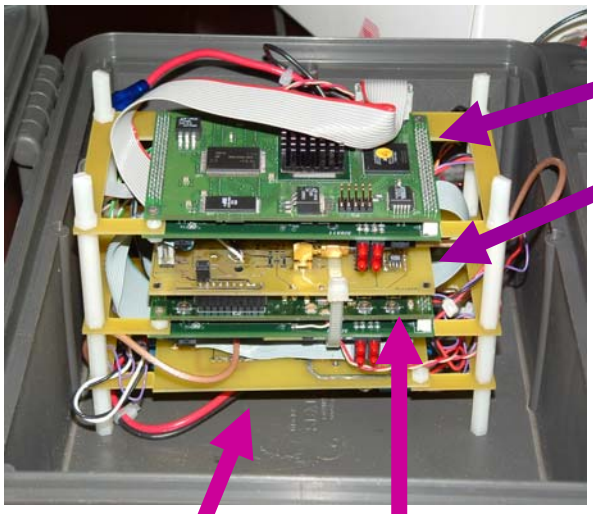


Figure 29. Schematic of a WASP unit processing hardware

Weight :
under 1 lb
4”x5”x6”





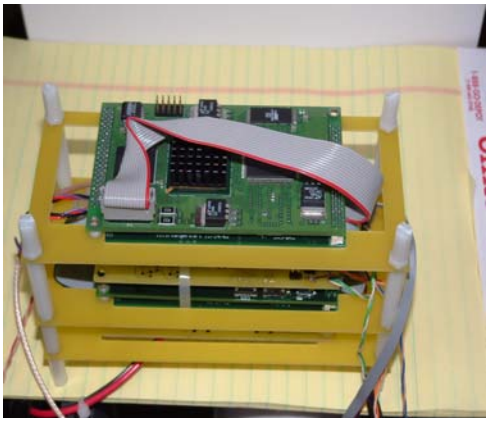


Figure 30. Different views of assembled WASP unit processing payload. The final system costs a little under $1000/node and weighs less than 1 lb. The entire unit can be assembled into a 4"x5"x6" metallic chassis

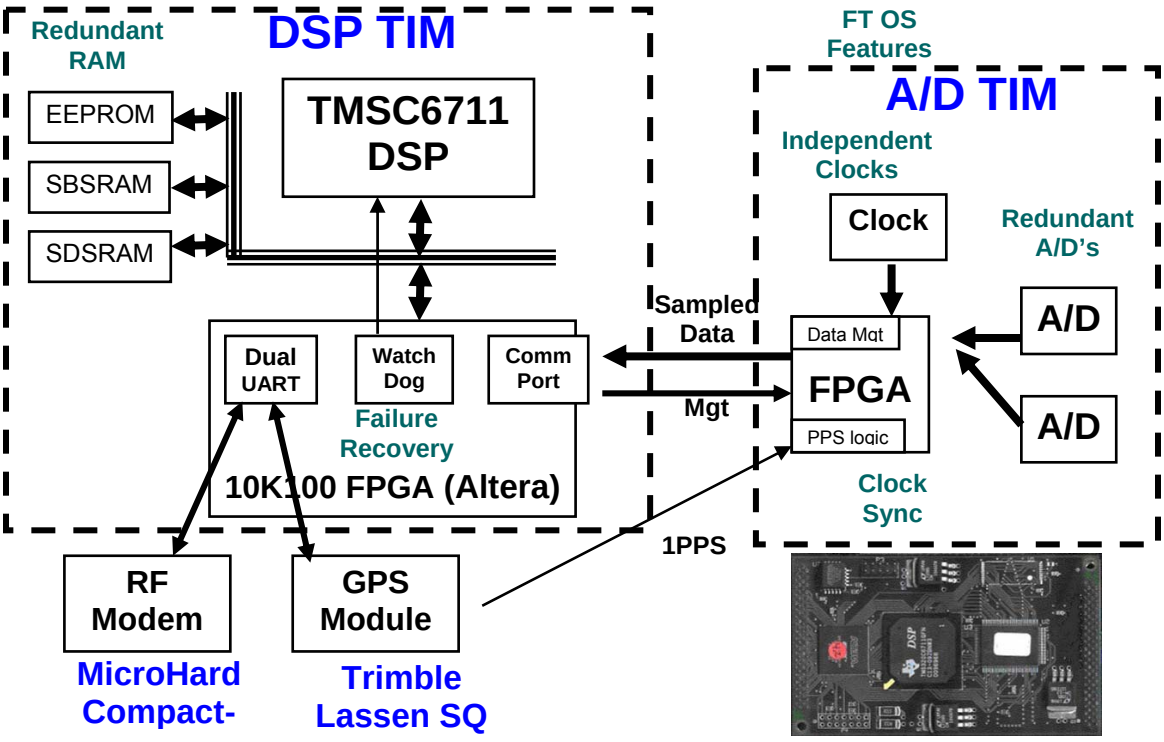


Figure 31. Functional schematic of the WASP unit hardware

Initial experiments showed that the demodulated base-band output of these radios demonstrated poor phase-response characteristics and, hence, resulted in poor TDOA estimates. Further analysis revealed the analog demodulation stages of the FRS receivers as the largest contributors to these errors. Consequently, the WASP nodes bypass the analog demodulation circuitry of the radios and perform the demodulation in software. The receiver front-end, however, is still used to select the correct channel and down-mix the FM data from the FRS frequencies to an intermediate frequency (IF) at 450 KHz. The FM data are gathered by sampling the IF using a 12-bit A/D at a nominal sampling frequency of 2 MHz. Gathering the data at the IF stage gives the system the flexibility of tracking any of the 14 channels in the FRS band while circumventing the error prone additional analog demodulation stages of the receivers.

A Trimble Lassen SQ GPS module provides both location and timing information. The module also provides a 1 pulse-per-second (PPS) signal accurate to ±100 nano-seconds (ns) [36]. A Texas Instruments TMS320C67 family DSP with 32 MB of RAM serves as the main processing element (PE). An 37 SG-531 series crystal oscillator provides a nominal 48 MHz clock that drives the PE and the peripheral hardware [37].

The 48MHz clock is divided by a factor of 24 and provides a 2 MHz clock that drives the 12-bit Analog to digital (A/D) hardware. A 100K transistor Altera FPGA implements a watchdog timer and handles the communication protocols for the PE. A MicroHard Compact-RF provides a 9600 baud data-link between the remote WUs and the base-station. The entire hardware payload can be packaged in a metallic chassis as shown in Figure 30 and measures approximately 4”x5”x6” and weighs less than 1 lb (excluding the battery). The operational code along with other relevant meta-data (e.g. template parameters, threshold values to accept or reject received data, node ID etc.) is stored in the non-volatile EEPROM. A compressed version of the target or template is also stored in the EEPROM, which on first boot up is uncompressed and later used to compare against received signals to detect the SOI. Figure 31 depicts the functional organization of a WU.

Challenges imposed by hardware design

While the system and hardware designs satisfy the some of the constraints presented earlier, they spawn a set of secondary, yet significant, challenges, which are enumerated below:

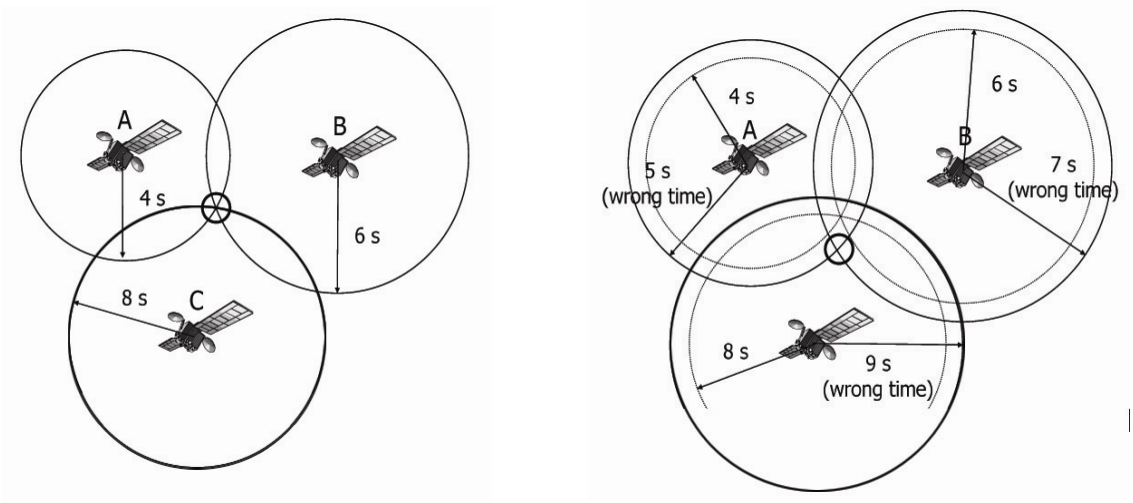


Figure 32. Effect of errors in measuring time of arrivals from different satellites. The figure on the left indicates the "ideal" case when there is no noise in the TOA measurements. The receiver's position is computed accurately, as shown by the unique intersection. The figure on the right depicts the situation in the presence of timing jitters. The estimate locations are "spread" around the actual location. The true location must thus be estimated as the centroid or the least means squared fit of the individual intersections. Figure adapted from [DIAMLER]

Low quality radio units

Analog processing of the radio signal impedes high precision measurement. Circuit inconsistencies between different radio units, temperature drift, interference, varying stray capacitance and inductance etc. all affect phase performance. Since the FRS radios were designed primarily as a cost-effective enabler for human-to-human voice interactions, their inherent performance metrics while satisfactory for normal voice communication are far below the necessary limits for an accurate geolocation system. In particular, the radio base-band output contains spurious harmonics, which for normal operation are either rejected by the final transduction elements (such as the output amplifier and the speaker diaphragm) or simply ignored by the human psycho-auditory

system. The radios also exhibit, for the current geolocation application, poor phase linearity, inadequate base-band performance and inconsistent near-field effects.

Drifting sample rate clocks

Each WU has an independent clock, accurate to ±100 PPM [37]. The frequencies and/or phases of the clocks on different nodes will generally not align. Also, varying environmental and operating conditions, such as ambient temperature, operating voltage, age etc., will influence the operating frequency and contribute to phase and frequency drifts. The TOA algorithms must correct for frequency and phase drift when estimating the time of arrival. Figure 32 illustrate the effects of timing errors for a GPS receiver (self-positioning). The left half of Figure 32 depicts the case when there are no timing errors in the receiver. The TOAs or, equivalently, the times of flight from the different space vehicles are measured correctly and hence the self-positioning accurately computes its location. In the adjacent figure, the perceived TOAs are different from the actual times of flight, leading to incorrect position estimates as indicated by the misplaced circle. Similar effects are observed when trying to triangulate the location of the target emitter form the TOA estimates from the different WU locations (remote-positioning).

Lack of common global clock

To ensure usability, the TOA estimates from each unit must be referenced against a common time. Since the WUs are geographically separated and cannot synchronize frequently or effectively using the RF network, they are coerced to use a local clock correction algorithm. The only available "global clock" is the GPS 1-PPS, specified accurate to ±100 ns. The GPS based synchronization achieves several aims concurrently:

1. Being a passive method of synchronization, the number of messages between participating nodes is reduced thus providing greater stealth and savings in transmission power.

2. In most cases, it also improves the area coverage as the nodes needn't be in line of sight (LOS) and within each other's transmission radius.

3. The spread spectrum encoding of the GPS signal also affords some level of security against malicious jamming.

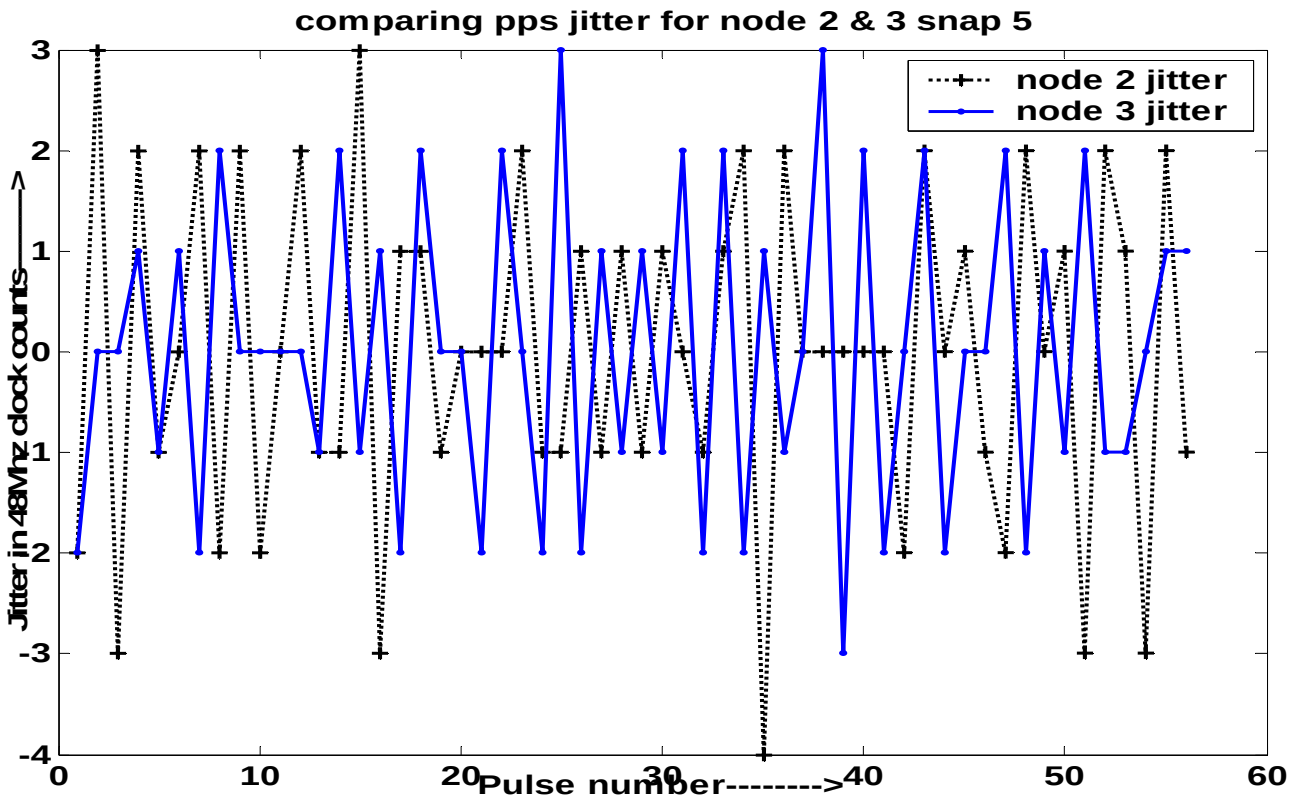


Figure 33 Comparing the inter-pulse durations (jitter) in PPS pulse trains from two different GPS units. Notice the visual non-conformity of the two pulse trains. The two nodes were situated within 5 feet of each other and were allowed to build a PPS pulse history for 60 seconds. To build the history, each node counts the number of high-resolution clock ticks (48 MHz) between consecutive PPS pulses, i.e. it counts the length each inter-pulse interval (IPI) in terms of clock ticks. The first order difference of these inter-pulse intervals provides a measure of the deviation in the observed duration of a second, and is plotted above for each node.

However, PPS signals received by WASP units (WUs) in geographically distinct areas may contain jitters. These timing jitters may arise due to several factors including atmospheric effects (such as varying ionospheric charged particle densities) or echoes or obstructions in the GPS signal path. Moreover, the GPS devices themselves possess some tolerances that may introduce significant errors. Also, some GPS devices create time glitches when they switch between satellite clusters. Some are unable to keep a time lock with fewer then three satellites in view. Some consider their serial output a “low priority”

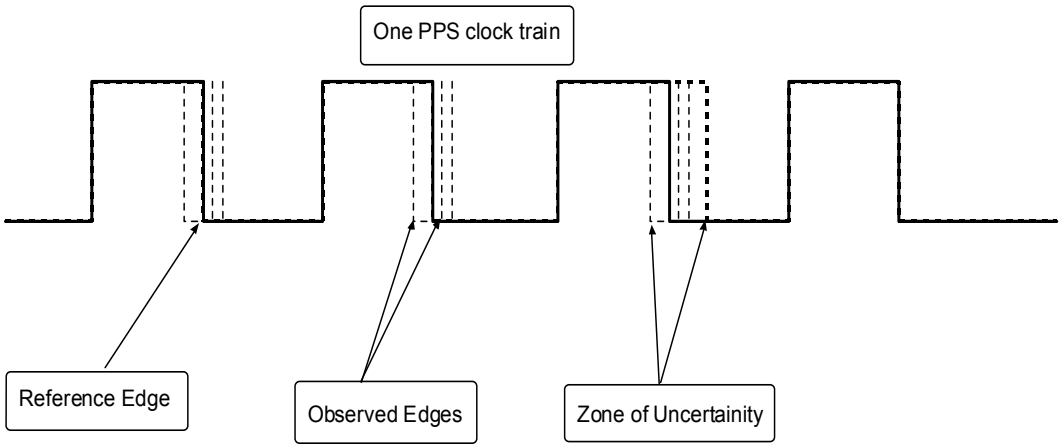


Figure 34. Schematic of PPS jitter. The actual edge occurs within an interval of uncertainty around the expected time. Since the length of a second is measured by the "distance" between two consecutive low-to-high edges, the duration of a second appears to change between consecutive measurements. The jitter correction algorithm aims to reduce the effects of these timing artifacts

task and can delay time code outputs by randomly varying amounts if they are busy computing. Figure 33 illustrates the jitter in the PPS trains of two co-located GPS devices. Note the mismatch between the crests and troughs for the two devices.

While most COTS GPS devices, such as the Trimble Lassen SQ, have tolerances of few 100s of nano-seconds, for the current application these are prohibitive. Since the WASP units use radio signals that travel at the speed of light (~ 3 x 10^8 m/s) every nano-second of error in the TOA estimate corresponds to roughly a foot of uncertainty in the position estimate of the source. Consequently, the device tolerance may add potentially 250 feet (~80m) of uncertainty in the final position estimate. Figure 34 shows the effect of clock jitter: the actual clock edge happens not at the expected instant in time, but within some region of uncertainty around the expected edge. Since the randomness of the edge is caused by a variety of reasons, the central limit theorem may be invoked to model the overall distribution of the clock jitter as Gaussian. Figure 35 shows experimental data corroborating this assumption.

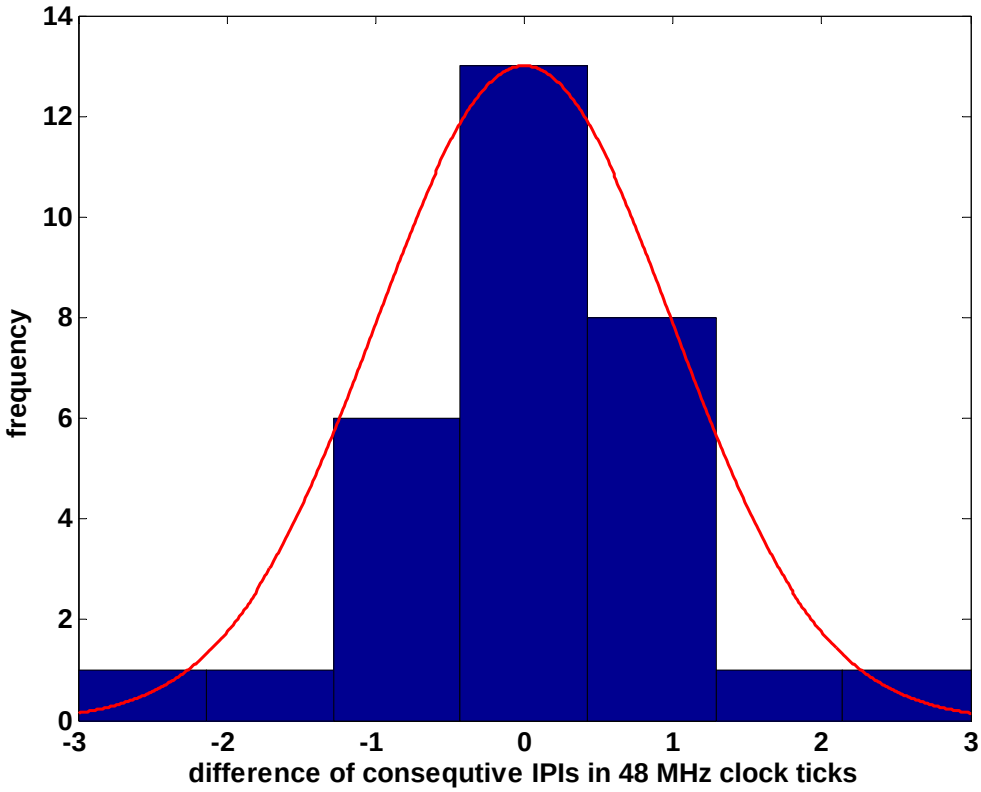


Figure 35. Histogram of differences between consecutive IPIs measured in high-resolution (48 MHz) clock ticks. A Gaussian curve (solid line) is superimposed over the histogram data to highlight the similarity of the two distributions. The above data suggests that the PPS variation is ±150 ns

Moreover, the local clocks at different nodes generally start at different times and thus have an inherent offset. They also suffer from thermal drifts, leading to varying operating frequencies depending on environmental conditions. While there are methods for reducing thermal drifts (e.g. by using a temperature controlled environment for the crystal oscillator) and GPS jitters (better GPS receivers and antennas, differential GPS etc.) the SWPC constraints preclude the use of any such means.

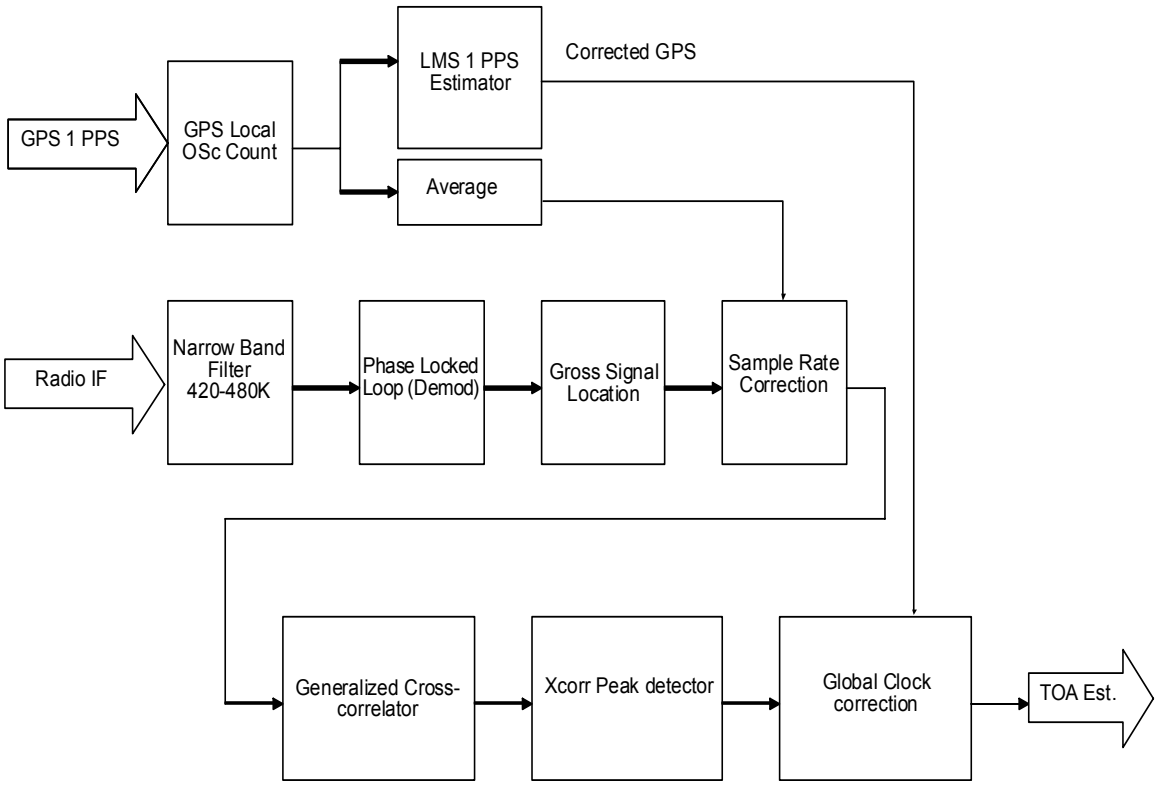


Figure 36. Block diagram of the signal processing architecture on a WASP unit. The design is motivated by physical/ operational constraints and the challenges imposed by the hardware platform. The GPS one pulse-per-second history is used to account for long-term drifts in the node sampling clocks, while the short term stability of the clocks is used to correct GPS jitters and refine the final TOA estimates. Demodulation in software eliminates errors introduced by the analog demodulation circuitry of the radio.

Signal processing system architecture

Figure 36 shows the schematic of the signal processing architecture designed for addressing the challenges recognized in the foregoing text. Digital processing removes many of the variables that impact radio phase performance. Ideally, a fully digital radio could be employed, but the size, weight, power and cost (SWPC) parameters prohibit this approach. Instead, we compromise and sample the intermediate frequency of the FRS radios. The IF is a 450 KHz FM signal, with a 25 KHz channel bandwidth, sampled at 2

MHz. A narrowband filter is applied to the IF. Demodulation is performed using a software based phase-locked loop.

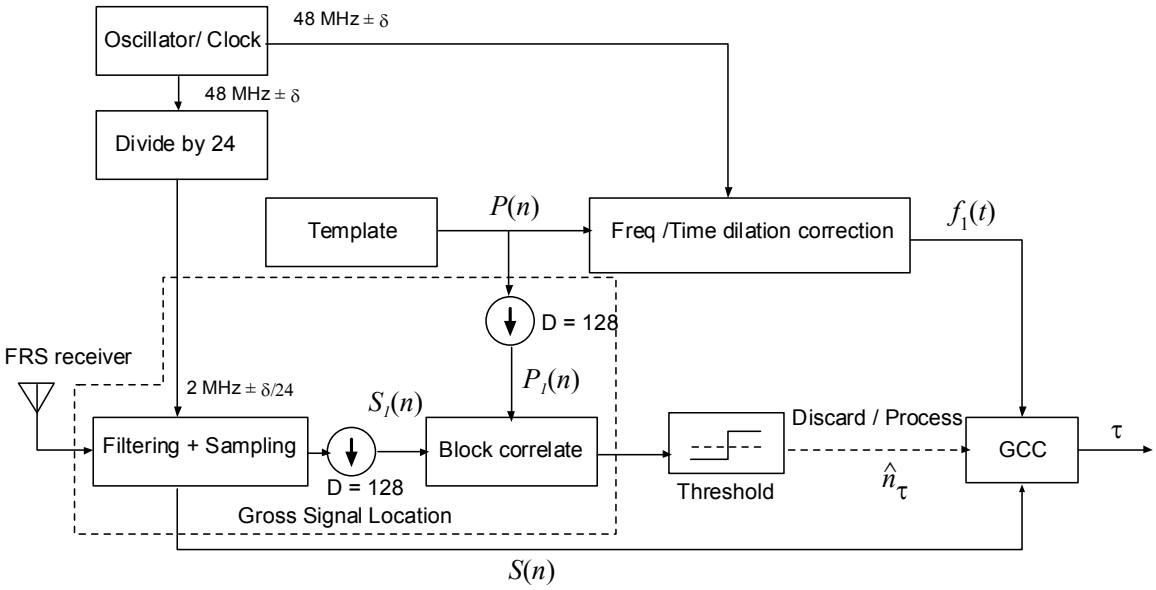


Figure 37. Filtering and downsampling in the GSL block. The elements enclosed by the dashed lines constitute the GSL block. Both the received signal and the template are filtered using a poly-phase filter to prevent aliasing. The filtered signals are then down-sampled by a factor (D) of 128, resulting in a signal stream sampled at approximately 15 KHz. If the output of the GSL block is greater than a heuristically chosen threshold, then the subsequent blocks are activated, else, the results are discarded and the node continues the radio-channel monitoring operation.

Once the IF data has been demodulated, it is subjected to a multi-resolution search. Decimated, low resolution versions of the sampled signal and the template (SOI) are created using an anti-aliasing polyphase filter in the Gross Signal Location (GSL) block. A search for the SOI using a frame based correlator is performed at this lower resolution. The effective sampling rate at this juncture is approximately 15 KHz. The signal decimation reduces the number of samples that must be processed to detect the occurrence of the SOI, effectively decreasing the computational effort and time required to search for the SOI. This is beneficial in terms of throughput and power-consumption.

Figure 37 highlights the operation of the GSL block and shows its interaction with subsequent stages. A detailed discussion of the multi-resolution approach is presented in chapter III.

The frame or block correlator essentially factors the signal stream to be searched into multiple overlapping blocks. Each block is correlated with the template and the maximum for the correlation function is compared against a heuristically selected threshold. If the value of the maximum is above the threshold, the template is said to have been detected and the following processing blocks are enabled. Figure 38 illustrates the block correlation operation.

When using a block or frame based approach, the SOI may be spread over two adjacent blocks, leading to "partial matches" with the template. For low SNRs or high rejection thresholds, it is possible that one or more occurrences of the SOI may be missed. The simplest technique for avoiding this is to over-lap a significant section of the subsequent blocks. This automatically brings up a trade-off between accuracy and computational efficiency: the greater the over-lap, the better the chances of not missing a feature, at a greater computational cost. The extent of the over-lap is chosen heuristically; it is generally between a quarter to 3 quarters of the window length. For the low-resolution search, an over-lap of one half the window length was found to be adequate, while for the high resolution GCC operation, the overlap was set at three-fourths the window length.

After the SOI has been detected at the lowest resolution, the estimate is refined by a high-resolution cross correlation operation in the generalized cross correlation (GCC)

block. The GCC block estimates the cross-correlation of the detected section against a version of the template sampled at a higher resolution (2 MHz). The peak detector determines the location of the cross-correlation maxima and interpolates the cross correlation curve, using a quadratic polynomial fit, to determine precise location of the maximum. Using the knowledge of the current time (as provided by the GPS receiver) and current operating frequency the refined peak location is translated into a time of arrival (TOA) estimate. These estimates may be communicated immediately to the remote base-station or stored on board until requested.

As explained earlier, the Lassen SQ GPS receiver provides a 1 Hz synchronization sequence called the pulse-per-second or the PPS. A free running 32-bit counter counts the number of system clock ticks (48 MHz) between each subsequent low to high transition of the PPS train. At each low-to-high transition of the PPS, the counter value is appended to a history list of counter counts. Since the counter length is finite, it wraps around approximately every 80 seconds. Once corrected for wrap-arounds, the history of counter counts is subject to an LMS regression analysis to eliminate the effects of GPS jitter. The LMS fit produces a jitter compensated estimate of the actual time of arrival of the last PPS pulse. The current operating frequency of the node is also obtained as a by-product of the LMS fit. The PPS correction technique is developed and analyzed in greater detail in chapter III.

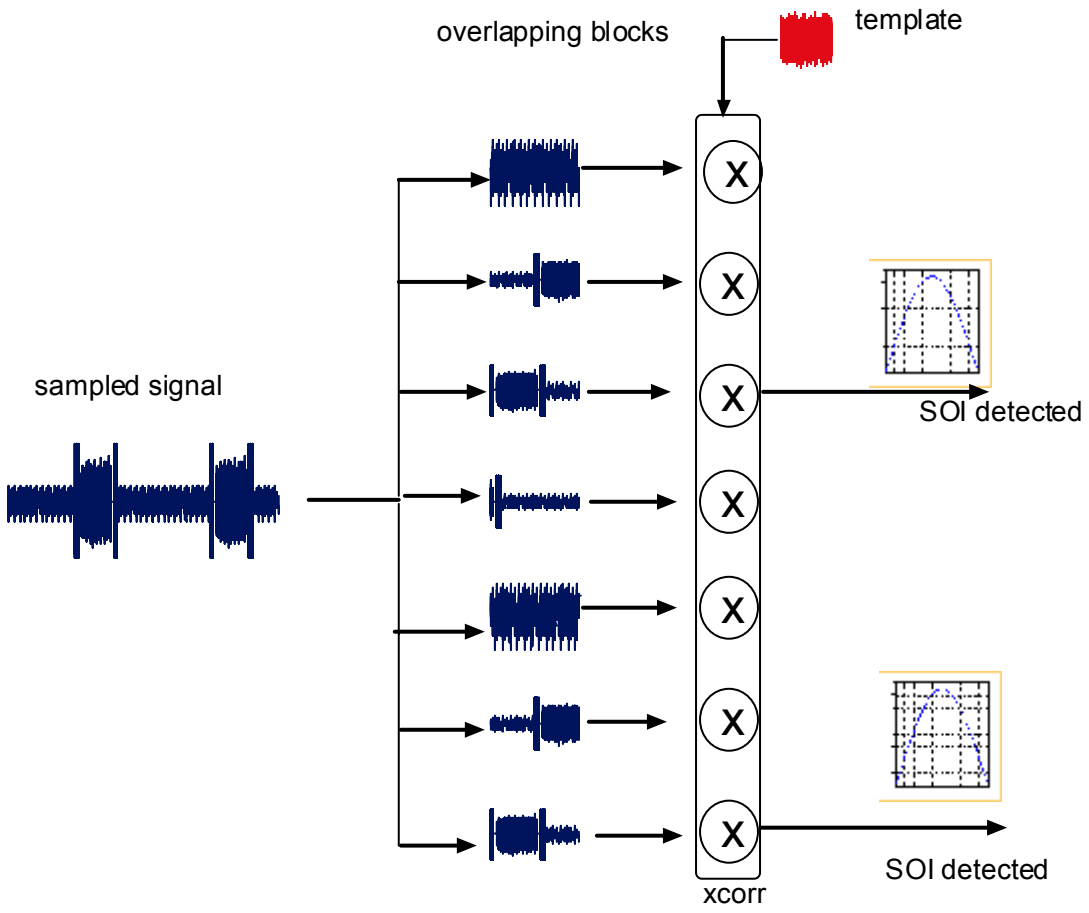


Figure 38. Schematic of block correlation algorithm. The downsampled signal stream is divided into overlapping blocks. Each block is then cross-correlated with the downsampled template. If the magnitude of the cross-correlation peak is greater than a specified threshold, the SOI is considered found

The software radio functions of filtering and demodulation are standard DSP operations and will not be discussed further here. We concentrate instead on the processing and implementation of the PPS correction, Gross Signal Location (GSL), Generalized Cross Correlation (GCC) [38], and Sample Rate Corrections (SRC) blocks.

Generalized Cross Correlation (GCC)

Cross-correlation or matched filtering is widely accepted as the best method for detecting the presence of a signal stream corrupted by additive Gaussian noise. The cross correlation of a signal x(t) with another signal y(t) is mathematically expressed as

$$R_{xy}(\tau) = x \circ y(\tau) = \int_{-\infty}^{\infty} x^*(t) y(t+\tau) dt \tag{1}$$

or equivalently in the causal form as

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x^*(t-\tau) y(t) dt \tag{2}$$

where the * indicates complex conjugation. For causal discrete time sequences of length N, the correlation may be expressed as

$$R_{xy}[m] = \sum_{n=0}^{2N-1} x^*[n-m] y[n] \tag{3}$$

When $y(t)$ is a scaled and delayed version of $x(t)$ corrupted by additive Gaussian noise, the correlation function is maximum when the argument τ is equal to the delay. The cross correlation may also be computed in the frequency domain via the cross-correlation property of the Fourier transform (also known as the Wiener-Khinchin theorem when $x(t) = y(t)$) [39], [40]

$$R_{xy}[n] = \sum_{k=0}^{2N-1} X^*[k] Y[k] e^{j 2\pi k n / N} \tag{4}$$

where N denotes the lengths of the two sequences. For discrete, finite length signals, the correlation may be computed efficiently in the frequency domain using FFTs:

- Zero pad each signal to the correct length ($2N$-1)

- Compute forward transforms of each signal using the FFT in $O(N\log N)$ steps
- Evaluate the conjugate product, also called as the *cross-spectral density* (CSD), of the two sequences in $O(N)$ steps
- Inverse transform the CSD to produce the desired correlation sequence in $O(N\log N)$ steps

While this is very convenient for record lengths of up to 2^{14}, for longer records computing the FFT on embedded resource constrained systems is currently expensive. The template on each WU at the highest resolution is roughly 400,000 ($\sim 2^{19}$) samples long. Since each sample is represented by a 32-bit floating point value, processing such a long record entails processing over mega-bytes of data. Moreover, the added record length only affords greater resolution in frequency; the narrow bandwidths of interest obviate the need for this extra resolution. Instead of evaluating the spectrum directly for the each record, we may simply use the estimate the spectrum computed for a convenient length [49]. Welch's [41] method of modified periodograms offers an efficient solution.

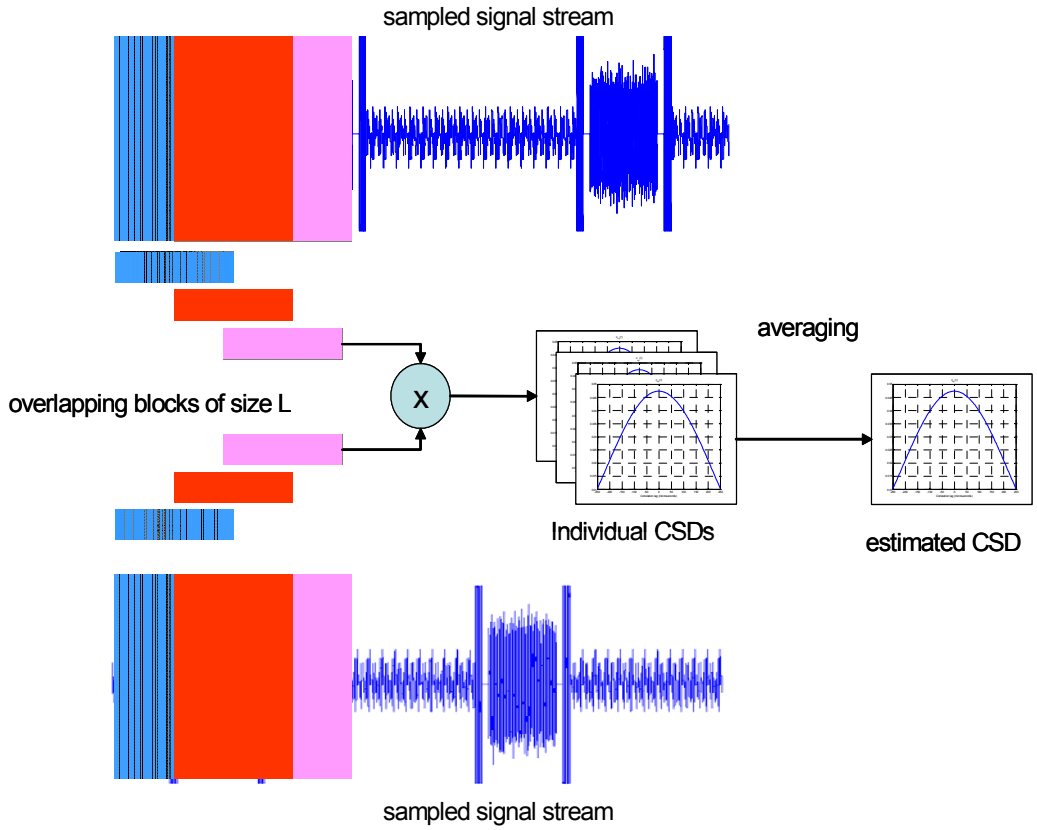


Figure 39. Cross-correlation via Welch's method of spectrum estimation. The two high-resolution signal streams to be cross-correlated are factored into pairs of overlapping blocks of equal length. The cross spectral density (CSD) for corresponding blocks is computed in the frequency domain (using FFTs), and the true cross-correlation is estimated as the average of the individual CSDs

Consider the signal stream presented in Figure 39. Similar to the block correlation method, Welch's method of modified periodograms splits each N-sample long data stream into overlapping blocks of size L. The cross spectral densities for each of these blocks are computed and the final estimate is obtained by averaging the individual CSDs. Consequently, the CSD of the two N sample long sequences is estimated using just 2*N/L complex FFT computations. The inverse Fourier transform the CSD then estimates the cross-correlation of the two sequences, as per the Wiener-Khinchin theorem [40].

Choosing the right window for estimating the PSD using Welch's method

A finite length fourier transform of a discrete time signal may be viewed as the least squares projection of that signal vector on to the space spanned a group of complex exponentials. If the signal vector has a length N, i.e. it has N discrete samples for a given period of time, T, then the Fourier transform is the LMS projection on to a space spanned by N complex exponentials. For real-world signals, this may be reduced to a space of N/2 complex exponentials. If the signal contains finite singularities, the fourier transform converges to the mean of the function.

Because complex exponentials have infinite support and are periodic, the discrete Fourier transform is also periodic. The Fourier coefficients may thus also be viewed as Fourier series coefficients of a periodic signals constructed by repeating the N samples of the given discrete signal [39], [55].

Typically the extremities of a finite-support real-world signal are dissimilar. This is more so for a finite window from a discrete time signal stream. Since the Fourier transform tries to represent discontinuities with a large number of higher frequency components, the discontinuity due to windowing adds higher-frequency components to the original signal. These spurious higher-frequency components are not part of the original signal and may be viewed as windowing artifacts. If a shaping window, such as a hanning window, is used, these artifacts may be reduced. The DCT is widely used in the field of image compaction for similar reasons. Its application, however, to the current problem is rather cumbersome. Hence we employ a shaping window such as the Hanning

window to estimate the power spectral density. Also, experimental results have shown that the use of a Hanning window in spectral estimation produces acceptable results [49]

APPENDIX B

Justification for compensating demodulated base-band data instead of original FM signal

Let the continuous time base-band modulating signal be represented by m(t). The frequency modulated carrier wave is then given by

$$s(t) = A\cos(\omega_c t + \theta + k_f \int_0^t m(\tau)d\tau) \qquad (1)$$

where A is the carrier amplitude, ω_c is the carrier angular frequency, θ is the default carrier phase, k_f is the modulation index (between 0 and 1). Since the FM signal is heterodyned prior to demodulation, the carrier may be assumed to be at the heterodyned intermediate frequency without loss of generality.

Let $s(t)$ be sampled independently to generate two sampled sequences $s_1[n]$, $s_2[n]$ sampled at $fs_1 = 1/Ts_1, fs_2 = 1/Ts_2$ respectively. i.e. for $i = 1,2$, we have

$$s_1[n] = A\cos(\omega_c n T_{s1} + \theta + k_f \int_0^{nT_{s1}} m(\tau)d\tau) \qquad (2)$$

$$s_2[n] = A\cos(\omega_c n T_{s2} + \theta + k_f \int_0^{nT_{s2}} m(\tau)d\tau) \qquad (3)$$

Demodulating FM signal sampled at fs_1.

Suppose we have a method to reliably demodulate $s_1[n]$ at f_{s1} to recover $M[n]$, where

$$M[n] = k_f \int_0^{nT_{s1}} m(\tau) d\tau \tag{4}$$

from which, m[n] can be recovered as

$$m[n] = \frac{M[n] - M[n-1]}{T_{s1}} \tag{5}$$

$$\approx \frac{d}{dt} M(t)\Big|_{t=nTs1} \tag{6}$$

$$= m(t)\Big|_{t=nT_{s1}} = m(nT_{s1}) \tag{7}$$

which is just m(t) sampled with a sampling period of T_{s1}.

We now want to apply the same modulation technique to *s2*[*n*]. To do so, we must express *s2*[*n*] in terms of $s_1[n]$.

$$s_2[n] = A\cos(\phi(nTs_2)) \tag{8}$$

where $\phi(nTs_2) = \omega_c nT_{s2} + \theta + k_f \int_0^{nT_{s2}} m(\tau) d\tau$. Now to express phi in terms of nTs1, we denote $\Delta = Ts_2 - Ts_1$, which gives

$$\phi(nTs_2) = \omega_c n(T_{s1} + \Delta) + \theta + k_f \int_0^{nT_{s2}} m(\tau) d\tau \tag{9}$$

$$= \omega_c n(T_{s1}) + \theta + k_f \int_0^{nT_{s2}} [m(\tau)d\tau] + n\Delta\omega_c \frac{T_{s2}k_f}{T_{s2}k_f} \quad (10)$$

$$= \omega_c n(T_{s1}) + \theta + k_f \int_0^{nT_{s2}} \left[m(\tau) + \frac{\Delta\omega_c}{T_{s2}k_f} \right] d\tau \quad (11)$$

Changing the variable of integration in the above equation to *u*, such that, when $\tau=nTs_2$, $u = nTs_1$, and when $\tau=0$, $u = 0$. Thus $\mathrm{u} = \tau \frac{\mathrm{Ts}_1}{\mathrm{Ts}_2}$, and $d\tau = \frac{Ts_2}{Ts_1} du$, whereby,

$$\phi(nTs_2) = \omega_c nT_{s1} + \theta + k_f \int_0^{nTs_1} \left[m(\frac{Ts_2}{Ts_1}u) + \frac{\Delta\omega_c}{T_{s2}k_f} \right] \frac{Ts_2}{Ts_1} du \quad (12)$$

letting $\alpha = \frac{Ts_2}{Ts_1}$ and $\tilde{m}(u) = \alpha \left[m(\alpha u) + \frac{\Delta\omega_c}{T_{s2}k_f} \right]$, we get

$$\phi(nTs_2) = \omega_c nT_{s1} + \theta + k_f \int_0^{nTs_1} \tilde{m}(u) du \quad (13)$$

Thus,

$$s_2[n] = A\cos(\omega_c nT_{s1} + \theta + k_f \int_0^{nT_{s1}} \tilde{m}(\tau)d\tau) \quad (14)$$

and (eqn no for m[n]), this on passing through our demodulation technique will generate the output,

$$\tilde{m}[n] = \tilde{m}(t)\big|_{t=nT_{s1}} = \tilde{m}(nT_{s1}) \quad (15)$$

$$= \alpha m(\alpha nTs_1) + \frac{\alpha\Delta\omega_c}{T_{s2}k_f} \quad (16)$$

$$= \alpha m(nTs_2) + (\alpha - 1)\frac{\omega_c}{k_f} \qquad (17)$$

Which is clearly the original base-band data sampled at Fs2 scaled by α and DC-level-shifted by $(\alpha - 1)\frac{\omega_c}{k_f}$. Thus the effects of RTC due to disparate sampling frequencies may be compensating by treating the base-band data directly. This is advantageous because this permits the correction to be affected by shifting only a small number of frequency components in the Fourier domain for the base-band data.

APPENDIX C

C Function implementation of Doppler shift approximation.

The input data vector is modulated with the appropriate complex sinusoid to effect a frequency shift. The output complex data is returned in the variable fOut. The fairly decoupled nature of the correction technique can be easily exploited by a VLIW compiler.

```
void   ApplyDoppler(FLOAT_TYPE*   fOut,   FLOAT_TYPE   *fX,
FLOAT_TYPE fWd, int nSize, FLOAT_TYPE fSamplingFreq)
{
     int nIndex;
     FLOAT_TYPE fTheta = 0;
for(nIndex  =  fTheta  =  0;  nIndex  <  nSize;  ++nIndex,
fTheta+=fWd/fSamplingFreq)
                                          {
          fOut[2*nIndex] = fX[nIndex],
          fOut[2*nIndex+1] = -(fX[nIndex] * fTheta);
     }
}
```

C Function implementation of the time-shift approximation.

The input to the function is the current data block being processed, *x*, the current step number, *nStep*, the sampling frequency of the data, *fFreqSignal*, and the sampling frequency of the template, *fFreqTemplate*.

```
//Parameters:
      //x : the current block of the radio data to be operated on.      //The
template (target) is loaded from EEPROM
      //nStep : the current step number through the input stream. (the nth-
step)
      //fFreqTemplate : sampling frequency of the template
      //fFreqSignal : sampling frequency of the signal

//Globals:
      //Gxy: the frequency corrected CSD for the nth Step is returned in this
global memory block
      //working_buf : scratch pad area. Allocated and initialized during WASP
unit start-up initialization
      //fft_return_buf: scratch pad area where results of large_fft are stored
      //LOW_BIN_OF_INTEREST: FFT bin corresponding to smallest significant
frequency in base-band data
```

```
	//HIGH_BIN_OF_INTEREST: FFT bin corresponding to largest significant
frequency in base-band data
	//TargetFftBins     : matrix of precomputed FFT blocks of the template


void gcc_step_template(float *x, int nStep, float fFreqTemplate, float
fFreqSignal)
{
	int i;
	float fDELTA_T, fScale, fTd;
	float ; //these need values from some place
	int nLengthStep = FFT_SIZE;
	float fCorrectionReal, fCorrectionImag;

	memcpy(working_buf,x,FFT_SIZE*sizeof(float));
    large_fft(working_buf, hanning_window, FFT_SIZE); // Take complex FFT
	//max = NUM_TARG_BLOCKS;
	if(nStep >= num_target_fft_bins)
	{
		return;
	}


	fScale = fFreqTemplate / fFreqSignal;
	fDELTA_T = (1/fFreqTemplate - 1/fFreqSignal);

	//td = (s*i + 0.5) * BLOCK_SIZE * DELTA_T;
			//matlab equivalent
	fTd = (fScale * (float)nStep + 0.5) * nLengthStep * fDELTA_T;


	for(i=LOW_BIN_OF_INTEREST; i < HIGH_BIN_OF_INTEREST; i++)
    {
		//computing Xy2 = Yy.*conj(Xx);
		fCorrectionReal = 1;
		fCorrectionImag = -2.0 * M_PI * i/(float)FFT_SIZE * fFreqSignal *
fTd ;

		//(a + ib) * (1 + iw) = (a-bw) + i(aw + b)
		//fCorectionReal = (a-bw) ; fCorrectionImag = (aw + b)
		fCorrectionReal = TargetFftBins[nStep][2*i] - fCorrectionImag *
TargetFftBins[nStep][2*i+1]; //(a - bw)
		fCorrectionImag = TargetFftBins[nStep][2*i] * fCorrectionImag +
TargetFftBins[nStep][2*i + 1]; //(aw + b)

		//conjugate is taken into account when multiplying... the
conjugation is on the signal (x) as desired: KG
		Gxy[2*i]   += ( fft_return_buf[2*i]*fCorrectionReal +
fft_return_buf[2*i+1]*fCorrectionImag);
		Gxy[2*i+1] += ( fft_return_buf[2*i]*fCorrectionImag -
fft_return_buf[2*i+1]*fCorrectionReal);
    }

	if(nStep == 0) {
		memcpy(gcc_x_Debug, x, FFT_SIZE*sizeof(float));
		memcpy(GxyDebug, Gxy, FFT_SIZE*2*sizeof(float));
	}

}
```

REFERENCES

[1] Warrior, J.; McHenry, E.; McGee, K. "They know where you are [location detection]", Spectrum, IEEE, Vol.40, Iss.7, July 2003, Pages 20- 25

[2] DARPA smart modules program http://www.darpa.mil/MTO/SmartMod/Factsheets/CBT.html

[3] http://www.mcqassociates.com/products/3.php

[4] http://www.sparton.com/service/sti/intrusion.pdf

[5] http://www.fcc.gov/911/enhanced/

[6] Bahl P; Padmanabhan, V.N., "RADAR an in-building RF-based user location and tracking system", INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, Vol.2, Iss., 2000, Pages775-784 vol.2

[7] Stotts, L.B.; "Unattended ground sensor related technologies; an army perspective", Proc. SPIE. Vol. 4040. (2000) 2–10

[8] Kumar, S., Shepherd, D.; "SensIT: Sensor information technology for the warfighter", Proc. 2001 International Conf. on Information Fusion. Volume 1., Montreal, (2001) TuC1–3

[9] "Collaborative signal and information processing in microsensor networks", Kumar, S.; Feng Zhao; Shepherd, D.; Signal Processing Magazine, IEEE, Vol.19, Iss.2, Mar 2002, Pages13-14

[10] Lee, R. "Mobile geolocation tutorial", International Symposium on Advanced Radio Technologies, Boulder, CO 2004 (10 04)

[11] Sztipanovits, J. et. al, "Widely Adaptive Signal Processing, project proposal documentation"

[12] Capt. 12, L., et. al., "Networked unattended ground sensor fields: tradeoff study and configuration rules methodology", DSE-TR-02-10, http://www.dean.usma.edu/math/research/msce/10th_AUTS/papers/12_02.pdf

[13] "USAF Intelligence Targeting Guide", Air Force Pamphlet 14- 210 Intelligence, 1 FEBRUARY 1998, http://www.fas.org/irp/doddir/usaf/afpam14-210/part16.htm

[14] http://www.trueposition.com/

[15] "Doppler compensation for passive coherent location", Tang 15; He You; Dong Shijia; Ni jinlin , Signal Processing, 2002 6th International Conference on, Vol.2, Iss., 26-30 Aug. 2002, Pages 1457-1460 vol.2

[16] "Overview of radiolocation in CDMA cellular systems", 16, J.J.; Stuber, G.L., Communications Magazine, IEEE, Vol.36, Iss.4, Apr 1998, Pages38-45

[17] "Geolocation and assisted GPS", 17, G.M.; Richton, R.E. , Computer, Vol.34, Iss.2, Feb 2001, Pages123-125

[18] "Generalized geometric triangulation algorithm for mobile robot absolute self-localization", 18, J.S.; Carvalho, A.; Couto, C, Industrial Electronics, 2003. ISIE '03. 2003 IEEE International Symposium on, Vol.1, Iss., 9-11 June 2003, Pages: 346- 351 vol. 1

[19] http://www.cell-loc.com/

[20] W. 20, "Position-location solutions by Taylor-series estimation," IEEE Trans. Aerosp. Electron. Syst., vol. ASE-12, no. 3, pp. 187-194, 1976.

[21] "Source localization and beamforming", Chen, J.C.; Kung Yao; Hudson, R.E. , Signal Processing Magazine, IEEE, Vol.19, Iss.2, Mar 2002, Pages30-39

[22] J.M. 22, R.H. Katx, and K.S. J. Pister, "Next century challenges: Mobile, networking for 'Smart Dust'," in Proc. Mobicom, 1999, pp. 483-492.

[23] "Beamforming: a versatile approach to spatial filtering", Van Veen, B.D.; Buckley, K.M., ASSP Magazine, IEEE [see also IEEE Signal Processing Magazine], Vol.5, Iss.2, Apr 1988, Pages:4-24

[24] "GSM Evolution Towards 3rd Generation Systems", Author Zvonar, Zoran.; Jung, Peter.; Kammerlander, Karl, Publication Boston Kluwer Academic Publishers, 1999.

[25] "A maximum-likelihood parametric approach to source localizations", Chen, J.C.; Hudson, R.E.; Kung Yao, Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on, Vol.5, Iss., 2001, pages3013-3016 vol.5

[26] "A simple and efficient estimator for hyperbolic location, Chan", Y.T.; Ho, K.C. , Signal Processing, IEEE Transactions on, Vol.42, Iss.8, Aug 1994, Pages:1905-1915

[27] "An Orthogonal Transformation Algorithm For GPS Positioning", Xiao-Wen Chang and Christopher C. Paige, SIAM Journal of Computer Science. vol. 24, no. 5, pp. 1710–1732, 2003 Society For Industrial And Applied Mathematics

[28] Federal Communications commission docket 94-102 http://www.fcc.gov/Bureaus/Wireless/News_Releases/1996/nrwl6026.txt

[29] "Detection of distributed sources using sensor arrays", Yuanwei Jin; Friedlander, B, Signal Processing, IEEE Transactions on Vol.52, Iss.6, June 2004, Pages: 1537- 1548

[30] "Energy-Constrained Collaborative Processing for Target Detection", Tracking, and Geolocation, Peter W. 30 and Gary A Shaw, Lecture notes in computer science, Information Processing in Sensor Networks, LNCS, Springer 2003, Second International workshop, IPSN 2003

[31] "Progress In Electromagnetics Research", PIER 40, 131–153, 2003, Radio Propagation In Rural Residential Areas With Vegetation, N. 31, D. Censor, and D. Katz

[32] "Positioning GSM telephones", Drane, C.; Macnaughtan, M.; Scott, C. , Communications Magazine, IEEE, Vol.36, Iss.4, Apr 1998, Pages46-54, 59

[33] Motorola Talk-about T5320 product sheet

[34] http://wireless.fcc.gov/services/personal/family/

[35] http://frwebgate.access.gpo.gov/cgi-bin/get-cfr.cgi?TITLE=47&PART=95&SECTION=627&YEAR=2000&TYPE=TEXT

[36] http://trl.trimble.com/dscgi/ds.py/Get/File-22183/

[37] EPSON SG 531 crystal oscillator data sheet http://www.37device.com/www/PDFS/epdoc.nsf/308956542197017449256a9c001d58b4/ee97b916bce00f0c49256a7e00399245/$FILE/SG-51531series.pdf

[38] Knapp, C., Carter, G.; "The generalized correlation method for estimation of time delay"; IEE. Trans. Acous., Speech. Sig. Proc. Vol.24, Aug 1976 pp. 320- 327

[39] Proakis J; Manolakis, D; "Digital Signal Processing, Principles Algorithms, and Applications", Prentice Hall, Third Ed., 1996.

[40] Eric W. Weisstein. "Wiener-Khinchin Theorem." From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/Wiener-KhinchinTheorem.html

[41] "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms", Welch, P, Audio and Electroacoustics, IEEE Transactions on, Vol.15, Iss.2, Jun 1967, Pages: 70- 73

[42] Haykin, S., "Communication systems", 4th ed., John Wiley and Sons, 2003. pp 248 – 253 (matched filters)

[43] Skolnik, M., "Introduction to RADAR Systems", 3rd Ed, Tata McGraw-Hill Publications, 2003, pp. 281

[44] "A Wavelet Tour of Signal Processing (Wavelet Analysis & Its Applications)", Stephane 44, Academic Press; 2 edition , September 1999, pp. 59

[45] Stein, S; "Algorithms for Ambiguity function processing"; IEE. Trans Acou., Spch. & Sig. Proc. Vol.29, Iss.3, Jun 1981 pp 588- 599

[46] Ulman, R.; Geraniotis, E; "Wideband TDOA FDOA processing using summation of short-time CAF's", IEE. Trans. Sig. Proc. Vol.47, no 12, Dec 99

[47] "On digital replica correlation algorithms with applications to active sonar", Glisson, T.; Black, C.; Sage, A., Audio and Electroacoustics, IEEE Transactions on, Vol.17, Iss.3, Sep 1969, Pages 190- 197

[48] "Wavelets and wideband correlation processing", Weiss, L.G.,Signal Processing Magazine, IEEE, Vol.11, Iss.1, Jan 1994, Pages13-32

[49] "Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing", Carter, G.; Knapp, C.; Nuttall, A, Audio and Electroacoustics, IEEE Transactions on, Vol.21, Iss.4, Aug 1973, Pages: 337- 344

[50] Eric W. Weisstein. "50 Integral Rule." From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/50IntegralRule.html

[51] "Applied Statistics and Probability for Engineers", Douglas C. Montgomery, George C. Runger, Wiley Text Books; 3 edition (August, 2002)

[52] "Integrating External and Internal Clock Synchronization", Fetzer, C., Cristian, F., Real Time Systems, Vol. 12, no. 2, 123-171, 1997

[53] "TLS-ESPRIT in a time delay estimation", Saarnisaari, H. Vehicular Technology Conference, 1997 IEEE 47th, Vol.3, Iss., 4-7 May 1997, Pages:1619-1623 vol.3

[54] “Super-resolution TOA estimation with diversity for indoor geolocation”, Xinrong Li; Pahlavan, K., Wireless Communications, IEEE Transactions on, Vol.3, Iss.1, Jan. 2004, Pages: 224- 234

[55] “Digital Signal Processing”, 55, A.V., Schafer, R. W., Pearson Education; 1 edition (January 2, 1975)

[56] "Detection, Classification, and Tracking of Targets",Dan Li, Kerry D. Wong, Yu Hen Hiu, Akbar M. Sayeed IEEE Signal Processing Magazine, Mar 2002, pp 17-29