Mobile Phone Location Determination and Its Impact on Intelligent Transportation Systems

Yilin Zhao, *Senior Member, IEEE*

***Abstract*—Research and development on the technologies of locating the mobile (wireless) phone caller have been rapidly gaining momentum around the world. Once these technologies are mature enough to be deployed, they will have significant impact on automotive telematics and modern public transit systems. In this paper, we will discuss why locating mobile phones becomes a hot topic among telecommunications giants, what technologies are being studied and standardized, when we are going to see the actual deployment, and what services they may provide. We will then consider its potential impact on future intelligent transportation systems (ITS), including telematics and public transit systems. Many of us have already recognized how important a role the communications systems play in modern transportation. In the near future, if every mobile phone is able to determine its location, advances in our current transportation systems become inevitable.**

***Index Terms*—A-FLT, AOA, assisted-GPS (A-GPS), CDMA, E911, E-OTD, fusion, GSM, intelligent transportation systems (ITS), public transit systems, standards, TDOA, telematics, TOA, wire- less location.**

I. Introduction

Many transportation applications can be supported by centralized location and navigation systems, which utilize communications networks, host facilities, and other infrastructures together with the on-board vehicle equipment to locate and navigate [1]. These applications include automotive telematics and modern public transit systems. An automotive telematics system typically communicates over a telecommunications network while a modern public transit system is an application of the automatic vehicle location (AVL) system.

One important application that can be supported by centralized systems is a wireless-enhanced 911 (E911) system. The 911 telephone number is used in the United States to provide emergency assistance for the caller, initially for the line-based (wireline) telephone system only. Other countries have similar systems, but may use different numbers such as 110 for China, 999 for the United Kingdon, and 17 for France. The key advantage of an enhanced 911 over the basic 911 is that the public safety officer knows the caller's location and phone number. "Wireless E911" is a system evolved from the original wireline-based system to the wireless system. Ideally, it would provide the same functionality as the wireline-based implementation.

Manuscript received July 1, 1999; revised March 21, 2000. The Associate Editor for this paper was Dr. Shoichi Washino. This paper was presented in part at the Modern Transport in Hong Kong for the 21st Century Conference, May 1, 1999.

The author is with PCS Research Labs, Motorola, Inc., Libertyville, IL 60048-5343 USA (e-mail: yilin.zhao@motorola.com).

Publisher Item Identifier S 1524-9050(00)07345-2.

TABLE I
Accuracy Required for Locating Mobile Phones

Solutions	67% of Calls	95% of Calls
Handset-Based	50 meters	150 meters
Network-Based	100 meters	300 meters

The U.S. Federal Communications Commission (FCC) has recently made E911 a mandatory requirement for wireless communications services such as cellular telephone, wideband (broadband) personal communications services (PCS), and geographic area specialized mobile radio (SMR). This ruling and upcoming service is called wireless E911. The FCC requires that by October 1, 2001, public safety answering point (PSAP) attendants of wireless communications networks must be able to know a 911 caller's phone number for return calls and the location of the caller, so that calls can be routed to an appropriate PSAP and related emergency assistance attendants [2]. On September 15, 1999, the FCC decided to tighten the location accuracy requirement for Phase II implementation from 125 m in 67% of all cases to the new numbers shown in Table I [3]. In addition, the FCC requests manufacturers to begin selling and activating location-capable handsets no later than March 1, 2001. Although certain companies are still petitioning the FCC to relax this new edict, FCC's action could facilitate the development of many vehicle location and navigation applications that use communications infrastructures similar to those used for wireless E911. Besides emergency assistance, it will certainly trigger many location-based services with the mobile phone or wireless network. Therefore, it is not difficult to understand why telecommunications manufacturers and operators have been actively pursuing the technologies to locate the mobile phone.

II. Location Technologies

There are three most commonly used location technologies: stand-alone, satellite-based, and terrestrial radio-based [4]. As examples, a typical stand-alone technology is dead reckoning. A typical satellite-based technology is global positioning system (GPS). A typical terrestrial radio-based technology is the "C" configuration of the Long Range Navigation (LORAN-C) system. For wireless E911, the radio-based (satellite and terrestrial) technologies are the most popular ones. Cellular networks are terrestrial-based communications systems. It is natural to utilize the signals of the network to determine the mobile phone location or to assist the location determination. Research in this area has been very active recently as evidenced

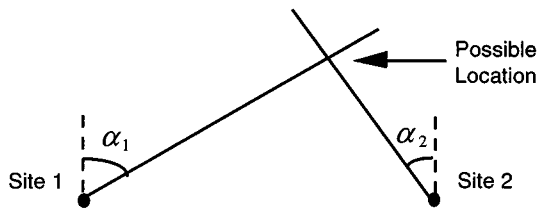

Fig. 1. Location determination by angle of arrival (AOA).

by the new round of publications [1], [4]–[20]. In this paper, we will address these radio-based technologies only. The principles behind them are discussed below. For information on technology developers, refer to [21].

Radio-based technology typically uses base stations, satellites, or devices emitting radio signals to the mobile receiver to determine the position of its user. Signals can also be emitted from the mobile device to the base. Commonly studied techniques are angle-of-arrival (AOA) positioning, time-of-arrival (TOA) positioning, and time-difference-of-arrival (TDOA) positioning. All these methods require radio transmitters, receivers, or transceivers. In other words, they depend on emitting and receiving radio signals to determine the location of an object on which a radio receiver, or a transceiver is attached. To make the position determination, these methods generally have the assumption that one end of the positioning system is fixed and the other end is moveable such as a mobile phone. However, the location determination capability can be either at the fixed end or at the mobile end. Generally, it is up to the system designer to decide where the final location determination capability should reside. For performance improvement, hybrid methods (various combinations of the techniques discussed or with additional techniques) are possible.

The *angle-of-arrival (AOA) system* determines the mobile phone position based on triangulation (Fig. 1). It is also called direction of arrival in some literature. The intersection of two directional lines of bearing defines a unique position, each formed by a radial from a base station to the mobile phone in a two-dimensional space. This technique requires a minimum of two stations (or one pair) to determine a position. If available, more than one pair can be used in practice. Because directional antennas or antenna arrays are required, it is difficult to realize AOA at the mobile phone.

The *time-of-arrival (TOA) system* determines the mobile phone position based on the intersection of the distance (or range) circles (Fig. 2). Since the propagation time of the radio wave is directly proportional to its traversed range, multiplying the speed of light to the time obtains the range from the mobile phone to the communicating base station. Two range measurements provide an ambiguous fix and three measurements determine a unique position. The same principle is used by GPS, where the circle becomes the sphere in space and the fourth measurement is required to solve the receiver-clock bias for a three-dimensional solution. The bias is caused by the unsynchronized clocks between the receiver and the satellite. Similarly, for terrestrial-based systems, it also requires precisely synchronized clocks for all transmitters and receivers.

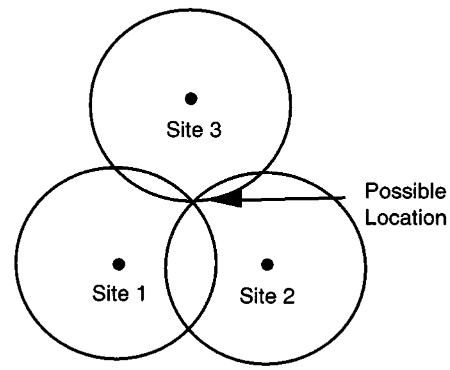

Fig. 2. Location determination by time of arrival (TOA).

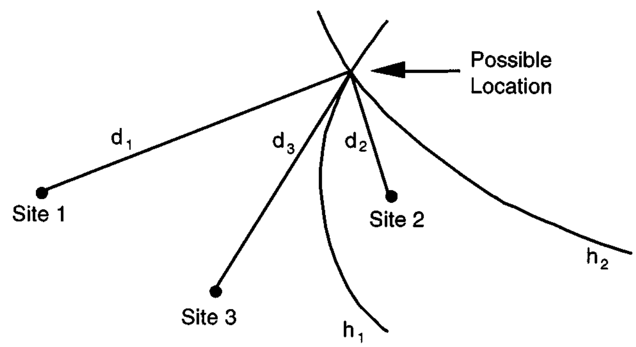

Fig. 3. Location determination by time difference of arrival (TDOA).

Otherwise, a 1-μs timing error could lead to a 300-m position error.

The *time-difference-of-arrival (TDOA) system* determines the mobile phone position based on trilateration (Fig. 3). This system uses time difference measurements rather than absolute time measurements as TOA does. It is often referred to as the hyperbolic system because the time difference is converted to a constant distance difference to two base stations (as foci) to define a hyperbolic curve. The intersection of two hyperbolas determines the position. Therefore, it utilizes two pairs of base stations (at least three for the two-dimensional case as shown in Fig. 3) for positioning. The accuracy of the system is a function of the relative base station geometric locations.

Other location methods can also be used. One simple method for mobile phone location is to use the cell area (or *cell ID*) of the caller, assisted with other coarse estimates, as the approximate location of the mobile phone. This results in the position error as large as the cell area. For instance, a picocell could be 150 m in radius while a large cell could be more than 30 000 m in radius. Therefore, this method has not demonstrated that it can achieve 100-m accuracy reliably even under the best of conditions. Another method is to use *short-range beacons* (or signposts in some literature) installed in the coverage area to provide location-specific information to the mobile receiver [1], [4]. Due to its limited communication zones, discontinuous communication, high system installation and maintenance costs, it has not been considered for the mobile phone location. Other methods are based on measuring the *signal strength* or measuring the *signal characteristic patterns* and multipath characteristics of radio signals arriving at a cell site from a caller. For measuring the signal strength, it employs multiple cell sites to find the location. For measuring the signal characteristic patterns, it identifies the unique radio frequency pattern or "signature" of the call and matches it to a similar pattern stored in its central database.

Because AOA requires the installation of directional antennas or antenna arrays, TOA and TDOA have been chosen as the current standardization choices. Of course, this may change if the next-generation systems can be equipped with these antennas. Both TOA and TDOA are time-based measurement technologies. They can be implemented either based on the forward (down) link signal or reserved (up) link signal. In addition, the location determination capability can reside either at the network side or at the mobile phone. In order to "hear" several base stations or cell sites, the sensitivity of the mobile phone may need to be increased. For better location accuracy, certain phones may require higher chip or bit resolution such as $1/8$ or $1/16$. These methods also require software modification on the mobile phone and additional location determination units and related software in the network. As discussed above, the mobile phone needs to listen to the signals of at least three base stations or cell sites. The hearability and geographical locations of these base stations will affect the availability and the accuracy of the location determination.

Since the performance of the satellite-based GPS receiver is getting better and better while the receiver size and price keep going down, it becomes feasible to develop an *assisted GPS* (A-GPS) solution for the mobile phone, which requires software and hardware modifications of both the mobile phone and its communications network. To understand this popular technology better, we will spend a little bit more space below to discuss it.

GPS provides an affordable means to determine position, velocity, and time around the globe [1], [22], [23]. The satellite constellation is developed and maintained by the U.S. Department of Defense. Civilian access is guaranteed through an agreement with the Department of Transportation. GPS satellites transmit two carrier frequencies. Typically, only one is used by civilian receivers. From the perspective of these civilian receivers on the ground, GPS satellites transmit at 1575.42 MHz using code-division multiple-access (CDMA) technique, which uses a direct-sequence spread-spectrum (DS-SS) signal at 1.023 MHz (Mchips/s) with a code period of 1 ms. Each satellite's DS-SS signal is modulated by a 50-bit/s navigation message that includes accurate time and coefficients (ephemeris) to an equation that describes the satellite's position as a function of time. The receiver (more precisely, its antenna) position determination is based on TOA.

The four main conventional GPS receiver functions are as follows.

1) Measuring distance from the satellites to the receiver by determining the pseudoranges (code phases).
2) Extracting the time of arrival of the signal from the contents of the satellite transmitted message.
3) Computing the position of the satellites by evaluating the ephemeris data at the indicated time of arrival.

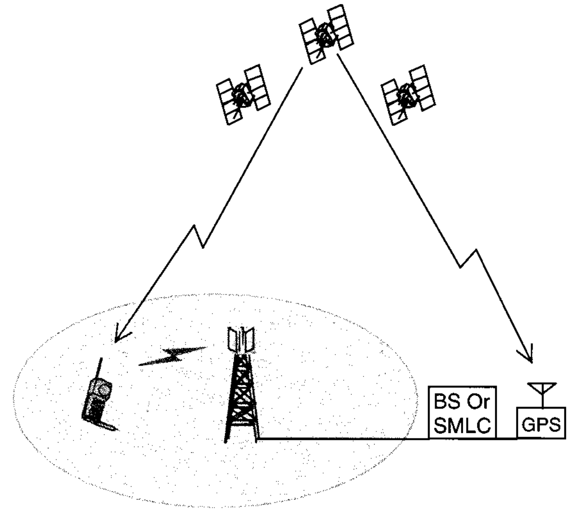

Fig. 4. Assisted GPS positioning (BS stands for base station and SMLC stands for serving mobile location center).

4) Calculating the position of the receiving antenna and the clock bias of the receiver by using the above data items.

Position errors at the receiver are contributed by the satellite clock, satellite orbit, ephemeris prediction, ionospheric delay, tropospheric delay, and selective availability (SA). SA is an accuracy-degradation scheme to reduce the accuracy available to civilian users to a level within the national security requirements of the United States. It decreases the accuracy capability of autonomous GPS to the 100-m (2D-RMS) level, where RMS stands for root-mean square. To reduce these errors, range and range-rate corrections can be applied to the raw pseudorange measurements in order to create a position solution that is accurate to a few meters in open environments. The most important correction technique is differential GPS (DGPS) [1], [22], [23]. It uses a reference receiver at a surveyed position to send correcting information to a mobile receiver over a communications link. Note that SA has been turned off since May 2000.

In addition to the task of shrinking the GPS antenna to fit a typical mobile phone, a traditional autonomous GPS receiver chipset is difficult to embed in the mobile phone for three main reasons. First, its startup time (from turning on to the initial position fix) is relatively long due to its long acquisition time of the navigation message (at least 60 s to a few minutes). Second, it is unable to detect weak signals that result from indoor and urban canyon operations, especially with small cellular sized antennas. Third, its power dissipation is relatively high per fix, primarily due to the long signal acquisition time in an unaided application. To deal with these problems, the assisted GPS method was proposed (Fig. 4).

The basic idea of assisted GPS is to establish a GPS reference network (or a wide-area DGPS network) whose receivers have clear views of the sky and can operate continuously. This reference network is also connected with the cellular infrastructure, and continuously monitors the real-time constellation status and provides precise data such as satellite visibility, ephemeris and clock correction, Doppler, and even the pseudorandom noise code phase for each satellite at a particular epoch time. At the request of the mobile phone or location-based application, the assist data derived from the GPS reference network are transmitted to the mobile phone GPS receiver (or sensor) to aid fast startup and to increase the sensor sensitivity. Acquisition time is reduced because the Doppler versus code phase uncertainty space is much smaller than in conventional GPS due to the fact that the search space has been predicted by the reference receiver and network. This allows for rapid search speed and for a much narrower signal search bandwidth which enhances sensitivity and reduces mobile receiver power consumption. Once the embedded GPS receiver acquires the available satellite signals, the pseudo-range measurements can be delivered to network-based position determination entity (PDE) for position calculation or used internally to compute position in the handset.

Additional assisted data, such as DGPS corrections, approximate handset location or cell base station (BS) location, and other information such as the satellite almanac, ionospheric delay, universal time coordinated (UTC) offset can be transmitted to improve the location accuracy, decrease acquisition time, and allow for handset-based position computation. Several schemes have been proposed in the standards which reduce the number of bits necessary to be exchanged between the handset and the network by using compression techniques such as transmitting only the nonredundant or the changes to parameters instead of the raw parameters themselves. Other satellite systems could be used, such as the Russian GLONASS system, but none of the standards have made provision for anything except GPS and the future GPS Wide Area Augmentation System (WAAS) signals. Besides adding a GPS reference network and additional location determination units in the network, the mobile phone must embed, at a minimum, a GPS antenna and RF down-converter circuits, as well as make provision for some form of digital signal processing software or dedicated hardware.

Recent field trials of the assisted GPS system have shown the feasibility of this technology. However, the current implementation has not demonstrated that it can cover every location where voice communication is available. In addition, this solution will not work for legacy phones.

In general, all the radio-based technologies discussed can be affected by interference, blockage, and multipath. It is a great challenge to solve these adverse effects caused by the environment in which we live.

III. Location Technology Being Standarized

Telecommunication standards organizations are busy incorporating the new location technologies into their standards, whether it is GSM, TDMA, cdmaOne, cdma2000, W-CDMA, UMTS, UWC-136, and even Analog. Three main standard organizations involved in second-generation (2G) systems are the European Telecommunications Standards Institute (ETSI), Telecommunications Industry Association (TIA), and the T1 Committee. T1 is sponsored by the Alliance for Telecommunications Industry Solutions (ATIS), which is accredited by the American National Standards Institute (ANSI). For third-generation (3G) systems, the work has been handled by the Third Generation Partnership Project (3GPP) and 3GPP2, respectively.

TABLE II
Standardized Location Technologies

MA	Technology
Analog Mode (TIA's TR45.1)	A-GPS
CDMA (TIA's TR45.5)	A-GPS
	A-FLT
GSM (T1P1.5's LCS SWG for ETSI)	A-GPS
	E-OTD
	TOA (Network-based)
TDMA (TIA's TR45.3)	A-GPS (Will be approved soon)

Listed in Table II are technologies being standardized by the above organizations. However, finding a technology which can achieve the accuracy requirements set by the FCC presents a great challenge. cdma2000, W-CDMA, UMTS, and UWC-136 (3G systems) may adopt and further evolve the location technologies being developed for CDMA and GSM, and are not included. Interested readers can refer to 3GPP and 3GPP2 Web sites for new development information [24], [25].

Note that A-FLT and E-OTD are described in the following section. LCS SWG stands for LoCation Services Sub-Working Group. Provided in Table II is a sampling of the technologies being considered by different standards organizations based on different multiple-access (MA) techniques. In the following sections, we will discuss these technologies in detail.

A. Time Difference of Arrival (TDOA)

The main TDOA location technologies considered for GSM and CDMA are E-OTD (enhanced observed time difference) and A-FLT (advanced forward link trilateration). In the following subsections, we examine each of them more closely.

Same as the assisted GPS discussed shortly, when the position is calculated at the network, we call it a network-based MS-assisted TDOA solution. When the position is calculated at the handset, we call it a network-assisted MS-based TDOA solution. Note that the "network-based" term used here may not have the same meaning as the one used in the recent rulings of the FCC. In the telecommunication standards literature, handset is often referred to as mobile station (MS).

1) Enhanced Observed Time Difference (E-OTD): E-OTD has been finalized by the GSM standard committees (T1P1.5 and ETSI) in LCS Release 98 and Release 99. Future releases will be handled by 3GPP. E-OTD is a TDOA positioning method based on the OTD feature already existing in GSM [26]. The MS measures relative time of arrival of the signals (bursts) from several BTSs (Base Transceiver Stations). The position of the MS is determined by trilateration (Fig. 5).

There are three basic timing quantities associated with this method:

1) Observed Time Difference (OTD) is the time interval observed by an MS between the reception of signals (bursts) from two different Base Transceiver Stations (BTSs). If we denote t_1 as the moment that a burst from the BTS 1 is received and t_2 as the moment that a burst from the

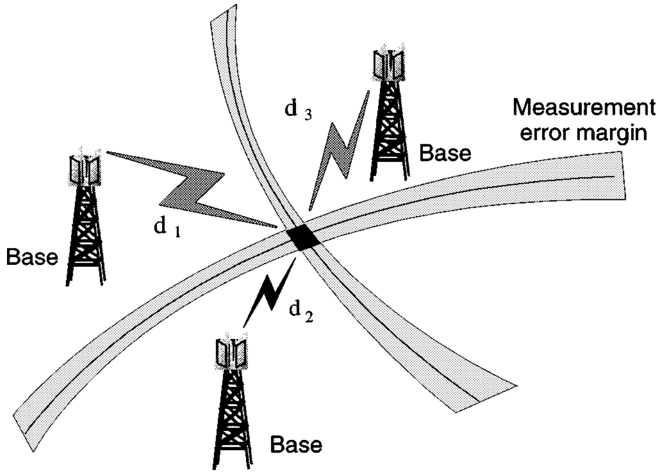

Fig. 5. E-OTD: Hyperbolic Positioning.

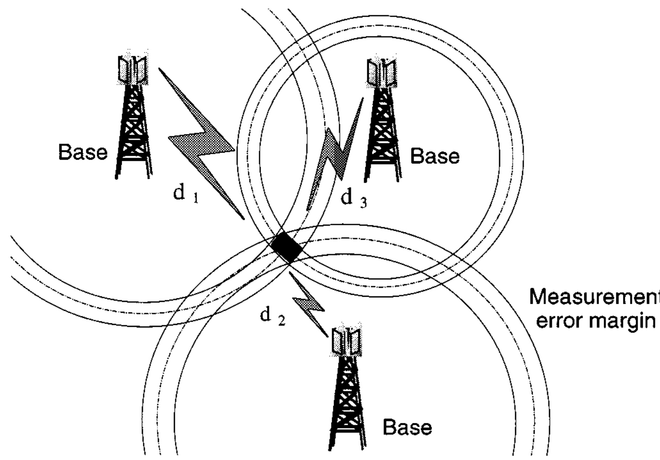

Fig. 6. E-OTD: Circular positioning.

BTS 2 is received, the OTD value is the time difference, i.e., OTD $= t_2 - t_1$. If the two bursts arrive exactly at the same moment, the difference is zero, i.e., OTD $= 0$.

2) Real-Time Difference (RTD) is the relative synchronization interval in the network between two BTSs. If we denote t_3 as the moment that the BTS 1 sends a burst and t_4 as the moment that the BTS 2 sends a burst, the RTD value is the difference of these moments, i.e., RTD $= t_4 - t_3$. If the BTSs transmit exactly at the same moment, the difference is zero, i.e., RTD $= 0$. This implies that we have a synchronized network.
3) Geometric-Time Difference (GTD) is the time interval measured at the MS between bursts from two BTSs due to geometry. If we denote that d_1 as the length of the propagation path between the BTS 1 and the MS, and d_2 as the length of the path between the BTS 2 and the MS, the GTD value can be calculated as GTD $= (d_2 - d_1)/c$, where c is the speed of light. If the distances to the MS are the same for both BTSs, GTD $= 0$.

These quantities are related by

$$\text{GTD} = \text{OTD} - \text{RTD}.$$

Since the MS knows OTD, and RTD can be measured by an additional location measurement unit (LMU) in the infrastructure, we can calculate GTD as shown in the above equation. A constant GTD value between two BTSs defines a hyperbola. Intersection of two hyperbolas determines the location of the MS.

Another method classified under E-OTD is a mixed TOA and TDOA approach. It measures the time of arrival of the signals from a BTS to the MS and to the network node LMU and uses the equation described below to derive the MS position.

There are five quantities associated with this method.

1) The observed time from a BTS to the MS (MOT) is a time measured against the internal clock of the MS.
2) The observed time from a BTS to the LMU (LOT) is a time measured against the internal clock of the LMU.
3) Time offset ε is the bias between the two internal clocks of the MS and LMU.
4) The distance from MS to BTS (DMB).
5) The distance from LMU to BTS (DLB).

These quantities are related by

$$\text{DMB} - \text{DLB} = c(\text{MOT} - \text{LOT} + \varepsilon),$$

where c is the speed of light.

There will be one such equation for each BTS. Since there are three unknowns (MS position x, y, and clock offset ε), at least three BTSs are required to solve for the MS location x and y and the unknown clock offset ε. The position of the MS is determined by the intersection of circles centered on the BTSs common to observations made by the MS and LMUs (see Fig. 6).

The E-OTD method requires a minimum of three spatially distinct BTSs. All these BTSs must be detectable by the MS. More than three measurements generally produce better location accuracy. An implementation of the E-OTD method may require an LMU to BTS ratio between $1:3$ and $1:5$.

2) Advanced Forward Link Trilateration (A-FLT): A-FLT has been standardized by the CDMA standard committee (TR45.5) [27]. The next LCS release will be handled by 3GPP2. Unlike GSM, CDMA (IS-95) is a time-synchronized system. Therefore, time-difference measurement is easier than GSM. The basic idea of this method is to measure the time difference (phase delay) between CDMA pilot signal pairs. Each pair consists of the serving cell pilot and a neighboring pilot. The time difference is converted to the range information. Finally, the range data is used to form certain curves at which an intersection is defined for the MS location.

Although the name of this method implies that A-FLT is a handset-based solution, the location can be determined either at the MS or at the network. For an MS-based solution, the MS must determine the time difference of arrival among multiple pilot signals through its searcher. For an MS-assisted solution, Pilot signal measurement message (PSMM) along with the round-trip delay can be used to determine the time difference. Since the basic principle of this method is not much different than TDOA (or E-OTD), we will not discuss it in detail in this paper.

3) Challenges: For TDOA to work properly, additional location-determination software and hardware must be added to the network. In GSM standards literature, the hardware and associated element are named as location measurement unit (LMU) and serving mobile location center (SMLC). In CDMA standards literature, the software and associated device are called as position determination entity (PDE).

Many challenges exist. For GSM, one challenge is the unsynchronized nature of the network. Unsynchronized clocks in the network can make the accurate measurements of TDOA very difficult since 1-μs (one millionth of a second) time error could lead to a 300-m position error. The primary purpose of the LMU is to make radio measurements for unsynchronized systems. For CDMA, GSM, and TDMA, the main challenge is the power control implemented for the smooth operation of the network. Power control means that the power of each individual MS must be carefully controlled so that no one is unnecessarily interfering with others who are sharing the same frequency or spreading code. To make TDOA-based methods work, at least three "hearable" BSs with good geometry are required. In some areas, this condition may not be satisfied easily. Even in an area that enough BSs are available, the power control mechanism may make the detection of the neighboring BSs more difficult. One choice is to utilize the power-up function (PUF), i.e., waking up the idle MS periodically for a very short period of time so the BSs can detect the signal from it. Recently, members of the Third Generation Partnership Project (3GPP) have proposed to pause the transmission of the serving BS for a very short moment in order to make sure that the MS closest to it can hear the neighboring BSs, i.e., idle period downlink (IPDL). This could solve the problem caused by the power control mechanism.

B. Time of Arrival (TOA) in GSM

The actual time-of-arrival (TOA) system determines the mobile phone position based on the intersection of the distance (or range) circle [1], [4]. For brevity, we concentrate here on the network-based TOA (or uplink TOA) which has been standardized in GSM [26].

This method is based on measuring the TOA of a known signal from the handset at three or more LMUs in the infrastructure. The known signal is the access burst generated by having the handset perform an asynchronous handover.

After signal measurement, the TDOA principle is used to determine the position of the MS. In other words, it calculates the time difference of at least two pairs of TOA signals and derives the MS position by hyperbolic trilateration. Therefore, it is a hybrid of TOA and TDOA methods. Its position calculation technique is very similar to E-OTD. The main difference is that the network-based TOA does its calculations at the infrastructure, while the MS-based E-OTD does its calculations at the MS.

One analysis has compared the uplink TOA and E-OTD methods [28]. It shows that the uplink TOA method is more effective at reducing noise and interference through correlation and burst averaging than E-OTD. On the other hand, higher deployment density ($1:1$ to $1:2$) is expected of the E-OTD LMU due to the impact of RTD error on the MS location accuracy. We are led to believe that the second-generation LMUs cannot be reused for W-CDMA. Listed in Table III is a summary of the main factors degrading the performance of both the uplink TOA and E-OTD.

TABLE III
ERROR SOURCES IN TOA AND E-OTD

Source	TOA	E-OTD
Multipath	x	x
Noise and Interference	x	x
Clock Instabilities	x	x
Implementation Errors	x	x
Base Station Geometry	x	x
RTD Errors		x
No Benefit from Antenna Diversity		x
No Benefit from Frequency Hopping		x
No Benefit from Radio Motion in RTD Link		x
Does Not Function in Areas with Repeaters		x
Limited Signal Processing Capability in the Handset		x

C. Assisted-GPS (A-GPS)

For classification, as mentioned above when the position is calculated at the network, we call it a network-based MS-assisted GPS solution. When the position is calculated at the handset, we call it a network-assisted MS-based GPS solution.

Despite the above classification of two assisted GPS solutions, their principles are the same. If the GPS receiver does not know its approximate location, it will not be able to determine the visible satellites or estimate the range and Doppler frequency of these satellites. It has to search the entire code phase and frequency spaces to locate the visible satellites. For the code phase space, it spans from 0 to 1023 chips. For the frequency space, it spans from -4.2 to $+4.2$ kHz. The relative movements between the satellites and receiver make the search even more time-consuming. Therefore, the time-to-first-fix (TTFF) is one important parameter to evaluate the quality of a receiver. For autonomous GPS, the present state-of-the-art fix time for an uninitialized GPS sensor is approximately 60 s. Clearly, this is unacceptable for certain applications such as E911. By transmitting assistance data over the cellular network, we can reduce the TTFF of a receiver to a few seconds. It is achieved by significantly reducing the search window of the code phase and frequency space by sending precise measurements of these parameters to the handset from the network. The reduction in search space allows the receiver to spend its search time focusing on where the signal is expected to be.

1) MS-Assisted GPS: The network-based MS-assisted solution shifts the majority of the traditional GPS receiver functions to the network processor. This method requires an antenna, RF section, and digital processor in the handset for making measurements by generating replica codes and correlating them with the received GPS signals. The network transmits a short assistance message to the mobile station (MS), consisting of time, visible satellite list, satellite signal Doppler and code phase, and their search windows. These parameters help the embedded GPS sensor reduce the GPS acquisition time considerably. These assistance data are valid for a few minutes. The sensor returns to the network from the MS the pseudo-range data processed by the GPS sensor in the handset.

After receiving the pseudorange data, the corresponding network processor or location server estimates the position of the MS. The differential correction (DGPS) can be applied to the pseudo-range data or final result at the network side to improve the position accuracy.

2) MS-Based GPS: The network-assisted MS-based solution maintains a fully functional GPS receiver in the handset. This requires the same functionality as described in MS-assisted GPS, plus additional means for computing the positions of the satellites and ultimately the position of the MS. This additional handset function generally adds to the handset's total memory (RAM, ROM) requirements in addition to the extra computing capability such as million instructions per second (MIPS). In the initial startup scenario, more data in the form of the precise satellite orbital elements (ephemeris) must be provided to the MS than for the network-based MS-assisted case. For the case of ephemeris data transmitted to the handset, this data is valid for 2–4 h or more and can be updated as necessary over time, thus, once the handset has the data, subsequent updates are rare. Besides point-to-point transmission, it also includes using a broadcast channel to distribute this data efficiently to all handsets in a network. If better position accuracy is required for certain applications, differential correction (DGPS) data must be transmitted to the MS approximately every 30 s while SA is on. The final position of the MS is generated at the MS itself. The calculated MS location can then be sent to an application outside of the MS if required.

D. Fusion of A-GPS and TDOA/TOA

Many trial results have been reported for the location technologies discussed. As expected, each technology has its own advantages and disadvantages. In general, A-GPS has better accuracy, but worse coverage than TDOA/TOA in buildings and urban canyon areas. For instance, buried inside high-rise buildings, GPS could fail to produce a fix while TDOA/TOA may have very good coverage because more cell towers are available in dense areas. The main problem for TDOA/TOA is that its solution quality depends heavily on the geometric locations of the contributing BSs. For instance, on a long, straight, open-view highway, GPS may have very good accuracy and coverage, but TDOA may fail to produce required solutions because its BSs may simply lie along the same highway. This leads to the idea of integrating these two technologies or another time-based method, TOA, to complement each other.

To determine a two-dimentional MS location, the current A-GPS and TDOA/TOA technologies must have at least three satellites and at least three BSs visible all the time, respectively. In reality, this is not always the case. A logical improvement is to integrate the measurements of the A-GPS and TDOA/TOA. In other words, we can view each BS as a pseudo-satellite to supplement the real satellites when the visible satellites are insufficient in number to provide a position fix. As a result, any combination of three of these real/pseudo-satellites will likely generate the location for the MS. Data taken in severe in-building and very dense downtown environments with simultaneous A-GPS and cellular signal measurements indicate the promise of the fusion of the two technologies.

Fusion of assisted-GPS and TDOA/TOA methods not only can deal with the situations where less than three satellites or less than three BSs are available, it can also improve position fix quality and can be used as additional aiding information for the GPS receiver (or sensor). For instance, three visible real/pseudo-satellites may not give us the best geometry. Additional measurements will improve the geometry coverage of the MS so the location accuracy will be improved.

IV. Location-Based Services

It is generally believed that location-capable cellular phone and network will be available by October 1, 2001. Although U.S. FCC's original intention was to fulfill the role of the mobile phone as a part of the emergency call and assistance system, its rulings will clearly have positive impact on many existing services and will certainly generate more new services which were not available to the general public before.

One market study report has categorized the mobile location services into safety, information, tracking, remote, and billing services, respectively [29]. Safety services, especially personal security, are very critical to many countries as evidenced by the U.S.-led mandatory wireless E911 ruling. Similar services will also be available to the other counties early in the 21st century. Information services include weather, traffic, navigation, and directory assistance. They can dramatically improve the quality of peoples' lives. Tracking services can monitor continuously the location of the vehicle, asset, and people. Through these services, companies could increase their productivity while minimizing the cost for tracking down their goods and properties. People would have less concern over whereabouts of their loved ones. Remote services can provide further convenience as unlocking the car, monitoring the engine, collecting tolls, and guiding precision surveying and farming equipment, etc. Finally, billing services will be able to differentiate a variety of customer services. For instance, home-zone billing could encourage low-mobility subscribers to migrate traffic from conventional wireline-based networks to wireless networks. All these services will generate revenues for many old and new businesses. By the year 2005, as estimated by Ovum, there will be a $20 billion market for network-based location services and a $2.5 billion market for vehicle-based location services worldwide. Similar forecasts have been available from other companies for the world-wide or individual regions, and specific segments of the market.

The Universal Mobile Telecommunication System (UMTS) forum expects about 400 million mobile subscribers worldwide in the year 2000 (actual number will be higher) and nearly 1800 million subscribers in the year 2010 [30]. For the third-generation mobile radio systems, it plans to support a wide range of services from voice and low-rate to high-rate data services up to at least 144 kb/s in vehicular, 384 kb/s in outdoor-to-indoor, and 2 Mb/s in indoor and picocell environments. Both circuit-switched and packet-oriented services for symmetric and asymmetric traffic will be supported. As estimated by Ovum (Fig. 7), the cellular data capacity and services will be expanded very rapidly in the next few years. This will in turn generate many location-based services.

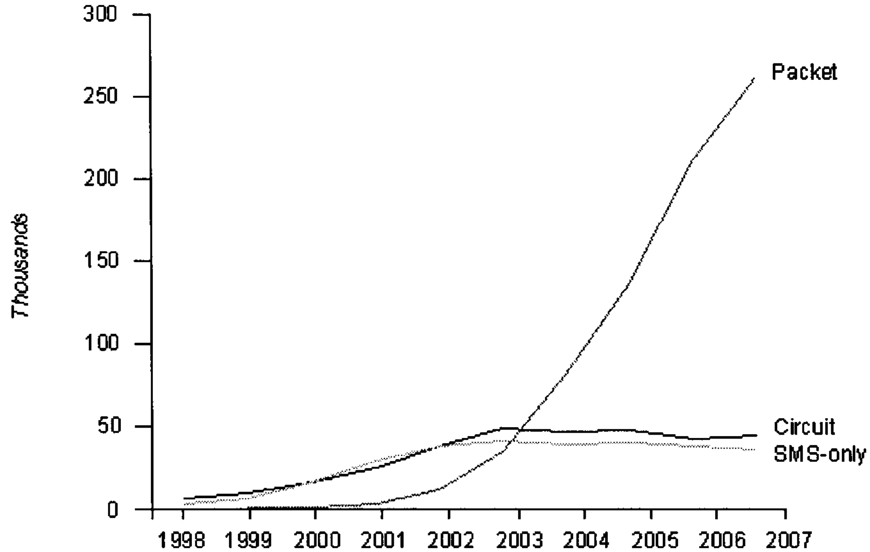

Fig. 7. United States and Europe data service subscribers (SMS stands for short message service).

TABLE IV
TELEMATICS ATTITUDE SURVEY

Attitude	Roadside Assistance	Routing/ Location Assistance	Convenience Services	Remote Door Unlock	Remote Vehicle Diagnosis
Extremely Satisfied	65%	71%	61%	75%	41%
Satisfied	24%	21%	29%	20%	39%
Somewhat Satisfied	0%	5%	4%	2%	5%
Not Very Satisfied	0%	3%	2%	0%	10%
Not at all Satisfied	7%	1%	4%	3%	3%

V. IMPACT ON TELEMATICS AND PUBLIC TRANSIT SYSTEMS

Location-capable phones and networks will have significant impact on intelligent transportation systems (ITS). In this section, we discuss two specific applications of ITS: automotive telematics and modern public transit systems.

A typical example of an automotive *telematics* system is a mayday system [1]. It provides vehicle occupants instant connection with a service center for emergency assistance or roadside services while automatically reporting the vehicle position. Many people in the United States view this system as their top priority when adding new equipment to their vehicles. It can be expanded to include many other services such as remote door unlocking, remote engine diagnosis, theft detection, notification and stolen-vehicle tracking, airbag deployment notification, automatic route guidance, travel information, and hands-free and voice-activated mobile phone or pager. A recent survey of 300 current customers by The Dohring Company has confirmed its popularity (Table IV). Because of this popularity, many automobile manufacturers have been and are bundling it as an original equipment manufacturer (OEM) unit for new model cars. In the future, this system will be able to add even more safety, security, and fun features, including connecting to the Internet, controlling by enhanced voice recognition, and combining with entertainment equipment.

A *mayday system* uses a cellular phone for voice and data communications and a global positioning system (GPS) receiver for positioning [31]. The key features of the mayday system are its human-centered design with a cost-effective location capability and its on-demand wireless communications [1]. With a human-centered design, the system can be activated either by the user with a push button or by an emergency event detected by one of the vehicular safety sensors. After the communications channel is established, the user can keep in voice contact with a human operator at the service center. With on-demand communications, the system does not need to communicate with the remote host on a regular basis as most automatic vehicle location (AVL) systems do, so there is a drastic reduction in silent air time and its associated expenses. Future systems will be offered in a single platform [32], and will include many customized services such as information, entertainment, and wireless Web connection [33].

With the current mayday system, the location device and communication device are separate items integrated into one system. Generally, the cellular phone and its transceiver are attached to the vehicle as nonremovable devices. So is the location device. Once the location-capable cellular phone is available, there may be no need for the fixed location and communication devices in the car. All we need could be a portable phone which integrates location and communication functions into one device. This will keep the same telematics functionality while reducing the number of phones for many users and could leave room for other in-vehicle devices.

A typical modern *public transit* system has the automatic vehicle location (AVL) capability [1]. An *AVL system* tracks the locations of a fleet of vehicles in a particular area and reports the information to a centralized server via a communications network [34]. This server can take different forms, such as a dispatch center, a traffic information center, or a transportation management center. For such a system, the location sensors keep updating the dispatch center on the route the vehicle is traversing. The communications network for information transmission can be dedicated radios, satellites channels, or short-range beacons installed along the road. Additional functions could be added such as route-by-route transit schedule, en-route information (on-board and at the bus station), transfer management, fare collection registration, passenger counting, vehicle diagnosis, emergency alert, paratransit management, and on-board video surveillance [35]–[37]. Due to the availability of centralized communications and its management center, a commonly used location technology is differential GPS (DGPS). Some public transit systems are further assisted by dead-reckoning sensors to complement GPS weakness often encountered in urban canyons, where tall buildings and other human-made landmarks cause satellite signal blockage and reflection.

With the advent of the location-capable phone, the complexity of the on-board equipment for the modern public transit system will be further reduced. If the cellular phone location determination could be as accurate as 15 m or less, there would be no need for any on-board dead-reckoning sensors. If the cellular service could be less costly, there would be no need for any specialized communications network. Furthermore, we could imagine how such a phone could assist the passenger. For instance, before even reaching the bus stop, the phone would be able to display when the next bus would be available based on the approximate time of arrival or distance to the bus stop. If the passenger wishes, it could also remind them individually when they have arrived at a predetermined destination or inform them of attractions and services along the bus route. Many more possibilities exist to serve the passenger better than before.

Despite technological advances on the mobile phone location, there are many other issues that require specific attention before it can be deployed in every cellular network around the world. These include cost recovery, handset development, system overload, and privacy issues. For cost recovery, wireless carriers and their partners need to find out how to make money for providing the location-specific ITS services. For handset development, manufacturers need to determine what kind of location solutions need to be implemented and how to minimize the handset cost. For system overload, system developers need to figure out whether these location-specific ITS services can be handled smoothly in the nearly full-loaded network. For privacy issues, it is always a concern when a system can monitor each move of the owner of a handset, whether it is a vehicle used as a location and traffic probe or a cellular phone used as a location-monitoring and call-tracking device.

In spite of the existence of the above concerns, the mobile phone location activities have been carried forward because of the many benefits that location service can provide to ITS. For instance, it can improve the response time of the emergency services to accident victims and the severely ill. It can aid in the development of vehicle crash avoidance and antitheft systems. It can improve the intelligent traffic management and control systems potentially reducing traffic congestion and air pollution.

VI. Conclusions

Mobile phone location determination activities have been intensified recently due to the October 1, 2001 deadline. Telecommunications standard organizations are busy incorporating the new location technologies into their standards, whether it is GSM, UMTS, cdmaOne, cdma2000, W-CDMA, TDMA, or UWC-136. Among the technologies discussed above, TOA, TDOA, and assisted-GPS solutions are the leading contenders for the current communication systems. Once these technologies are finalized in various standards organizations, the location-capable phone will hit the market soon. The location-based services will certainly follow. Besides wireless E911, these services may include location-sensitive billing, location tracking, location-based advertising, and information services such as navigation, weather, and points of interest. Similarly, automotive telematics and public transit systems will benefit, as will the other intelligent transportation systems. As we learned, these systems will be less and less complex while providing more convenient and attractive services. We certainly hope that this will in turn make our transportation systems operate more safely and efficiently, with less congestion, pollution, and environmental impact.

References

[1] Y. Zhao, *Vehicle Location and Navigation Systems*. Norwood, MA: Artech House, 1997.

[2] FCC, "Revision of the commission's rules to ensure compatibility with enhanced 911 emergency calling systems," in *Report and Order and Further Notice of Proposed Rulemaking*. Washington, DC: Fed. Commun. Comm., June 1996.

[3] FCC, "FCC acts to promote competition and public safety in enhanced wireless 911 services," Washington, DC: WT Rep. 99-27, Sept. 15, 1999.

[4] Y. Zhao, "Vehicle navigation and information systems," in *Encyclopedia of Electrical and Electronics Engineering*, J. G. Webster, Ed. New York: Wiley, 1999, vol. 23, pp. 106–118.

[5] C. Drane and C. Rizons, *Positioning Systems in Intelligent Transportation Systems*. Norwood, MA: Artech House, 1998.

[6] J. C. Liberti Jr. and T. S. Rappaport, *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*. Upper Saddle River, NJ: Prentice Hall, 1999, ch. 9, 10.

[7] J. J. Caffery, *Wireless Location in CDMA Cellular Radio Systems*. Norwell, MA: Kluwer, 1999.

[8] R. Klukas and M. Fattouche, "Line-of-sight angle of arrival estimation in the outdoor multipath environment," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 342–351, Feb. 1998.

[9] L. Xiong, "A selective model to suppress NLOS signals in angle-of-arrival (AOA) location estimation," in *Proc. 9th IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications*, vol. 1, Sept. 1998, pp. 461–465.

[10] R. Owen and L. Lopes, "Experimental analysis of the use of angle of arrival at an adaptive antenna array for location estimation," in *Proc. 9th IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications*, vol. 2, Sept. 1998, pp. 607–611.

[11] L. A. Stilp, "Time difference of arrival technology for locating narrowband cellular signals," in *Proc. SPIE Conf. Wireless Technologies and Services for Cellular and Personal Communication Services*. Philadelphia, PA: SPIE, Jan. 1996, vol. 2602, pp. 134–144.
[12] R. Klukas, G. Lachapelle, and M. Fattouche, "Field tests of a cellular telephone positioning system," *Proc. 47th IEEE Vehicular Technologies Conf.*, vol. 2, no. 1, pp. 470–474, May 3–5, 1997.
[13] T. S. Rappaport, J. H. Reed, and B. D. Woerner, "Position location using wireless communications on highways of the future," *IEEE Commun. Mag.*, vol. 34, pp. 33–41, Oct. 1996.
[14] S. Tekinay, Ed., "Wireless geolocation systems and services," in *IEEE Commun. Mag.*, Apr. 1998, vol. 36, pp. 28–76.
[15] J. O'Connor, B. Alexander, and E. Schorman, "CDMA infrastructure-based location finding for E911," in *Proc. 49th IEEE Vehicular Technology Conf.*, vol. 3, May 1999, pp. 1973–1978.
[16] G. Gutowski, L. Jalloul, E. Golovin, M. Nakhjiri, N. Yousef, and P. DeClerck, "Simulation results of CDMA location finding systems," in *Proc. 49th IEEE Vehicular Technology Conf.*, vol. 3, May 1999, pp. 2124–2128.
[17] R. Klukas, G. Lachapelle, and M. Fattouche, "Cellular telephone positioning using GPS time synchronization," *GPS World*, vol. 9, no. 4, pp. 49–54, Apr. 1998.
[18] M. Moeglein and N. Krasner, "An introduction to SnapTrack server-aided GPS technology," in *Proc. ION GPS-98*, Sept. 1998, pp. 333–342.
[19] L. Sheynblat and N. Krasner, "Description of a wireless integrated smart server/client system architecture," in *Proc. 55th ION Annu. Meet.*, June 1999, pp. 667–676.
[20] L. J. Garin, M. Chansarkar, S. Miocinovic, C. Norman, and D. Hilgenberg, "Wireless assisted GPS—SiRF architecture and field test results," in *Proc. ION GPS-99*, Sept. 1999, pp. 489–498.
[21] Y. Zhao. (2000) Wireless location: Companies and technologies. [Online] http://birch.dlut.edu.cn/~yzhao/locate.html
[22] B. W. Parkinson, J. J. Spilker Jr., P. Axelrad, and P. Enge, Eds., *Global Positioning System: Theory and Applications*. Washington, DC: Amer. Inst. Aeron. Astron., 1996.
[23] E. Kaplan, Ed., *Understanding GPS: Principles and Application*. Norwood, MA: Artech House, 1996.
[24] 3GPP. (2000) Third generation partnership project. [Online]. Available: http://www. 3gpp.org/
[25] 3GPP2. (2000) Third generation partnership project 2. [Online]. Available: http://www.3gpp2.org/
[26] "Digital Cellular Telecommunications System (Phase 2+); Location Services (LCS); (Functional Description)—Stage 2," ETSI, Sophia Antipolis, France, GSM 03.71, V7.3.0, Feb. 2000.
[27] "Position determination service standard for dual-mode spread spectrum systems," TIA, Arlington, VA, IS-801, Oct. 1999.
[28] W. Lindsay, M. Bilgic, G. Davis, B. Fox, R. Jensen, T. Lunn, M. McDonald, and W.-C. Peng, "GSM mobile location systems," Omnipoint, T1P1.5/99-411r1, July 1999.
[29] T. Blonz and C. McCarthy, *Mobile Location Services*. London, U.K.: Ovum Ltd, 1998.
[30] UMTS Forum, "A regulatory framework for UMTS," UMTS Forum, London, UK, Rep. 1, June 1997.
[31] Y. Zhao, "Efficient and reliable data transmission for cellular-and-GPS-based mayday system," in *Proc. IEEE Intelligent Transportation Systems Conf.*. Boston, MA: IEEE, 1997, pp. 555–559.
[32] M. Maes and S. Buytaert, "Telematics: The third generation," *Traffic Technology Int.*, pp. 66–68, Dec. 1999–Jan. 2000.
[33] M. Zenios, "Telematics, the future is in the car," *ITS World*, pp. 26–29, Jan./Feb. 2000.
[34] Y. Zhao, A. M. Kirson, and L. G. Seymour, "A configurable automatic vehicle location system," in *Proc. 1st World Congr. Applications of Transport Telematics and Intelligent Vehicle-Highway Systems*. Paris, France: Artech House, Nov./Dec 1994, pp. 1569–1576.
[35] L. Nowland-Margolis and B. Hiller, "On the move in Ann Arbor," *ITS World*, vol. 3, no. 2, pp. 22–26, Mar. 1998.
[36] B. Pantall, M. Stewart, M. Tsakiri, and J. Walker, "CATS on the prowl," *GPS World*, vol. 10, no. 4, pp. 32–36, Apr. 1999.
[37] A. Ampelas and M. Daguerregaray, "Paris public transit: The GPS difference," *GPS World*, vol. 10, no. 10, pp. 36–40, Oct. 1999.

Yilin Zhao (S'89–M'92–SM'97) received the B.E. degree in electrical engineering in 1982 from Dalian University of Technology, Dalian, China, and the M.S.E. degree in 1986 and the Ph.D. degree in 1992, both from the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor.

From 1982 to 1984, he was an Instructor and since 1995, has been an Adjunct Professor in the Department of Computer Science and Engineering, Dalian University of Technology. From 1987 to 1991, he was a Teaching Assistant and Research Assistant at the University of Michigan. In 1992, he joined Motorola, Inc. as a Senior Research and Development Engineer. His research interests include intelligent transportation systems (ITS), mobile phone location systems, vehicle location and navigation systems, integrated circuit place-and-route systems, and real-time computer systems. He has delivered ITS tutorials and seminars at many universities, IEEE, SAE, and other international conferences.

Dr. Zhao is an Associate Editor of the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS (ITS), and a Senior Representative of the IEEE Robotics and Automation Society to the IEEE ITS Council.