OET BULLETIN No. 71

Guidelines for Testing and Verifying the Accuracy

of

Wireless E911 Location Systems

April 12, 2000

Table of Contents

Guidelines for Testing and Verifying the Accuracy of E911 Location Systems

I. INTRODUCTION

This Bulletin has been prepared to provide assistance in determining whether systems for identifying the location of wireless handsets comply with the accuracy standards set by the Federal Communications Commission (FCC) for 911 Automatic Location Identification (ALI). The bulletin offers guidelines and suggestions for evaluating compliance. ***However, it is not intended to establish mandatory procedures. Other methods and procedures may be acceptable if based on sound engineering and statistical practice.***

The methods outlined in the present document are intended to produce results meaningful to public safety personnel and persons responsible for wireless system implementation and maintenance. The methods were developed in consultation with interested parties, many with experience in the design and conduct of ALI accuracy tests.

This document will help encourage E911 deployment by providing guidance to assist wireless equipment manufacturers, wireless carriers, and public safety personnel in designing, selecting, testing, and operating ALI systems. Some wireless carriers, for example, have expressed concern that particular ALI systems might not be considered to be compliant with the Rules unless the methods for testing and verifying compliance are specified. Compliance with these guidelines will establish a strong presumption that appropriate means have been applied to ensure that an ALI system complies with the Commission's Rules.

This document will also be helpful to groups and organizations working to establish standard test conditions and protocols. It is likely that testing and verification practices will evolve over time as representative models are developed and gain acceptance. In addition, it should be possible to establish standardized test procedures that include accuracy measurements for calls made from moving vehicles. It should also be possible to establish standard testing and verification methods that apply the same conventions to all location technologies and systems. This document is intended as a contribution to such efforts.

It is anticipated that methods for testing and verifying the accuracy of E911 location systems will continue to advance in a number of ways, in two areas in particular. First, it is likely that better information will be developed regarding locations from which wireless 911 calls are made. This document does not establish whether testing should treat all locations in a service area as equally important sources of possible 911 calls, or instead whether testing should concentrate on locations where 911 calls are most likely to be made. However, the text of this document states a preference for using data on wireless 911 call location information, if available.

Second, it is also anticipated that methods are likely to be developed in the future by which particular ALI technologies can be certified as compliant independent of the operating areas where they may be implemented. It is important to note that the approaches for predicting performance set out in these guidelines require local information and must be validated by experience. These approaches do not result in certification of an ALI technology as compliant independent of the operating area where it may be implemented. In the future, as warranted by advances in ALI technology and/or evaluation techniques, appropriate guidance may be provided for certifying ALI technologies for compliance independent of the area in which they are implemented.

II. BACKGROUND

The FCC adopted accuracy and reliability requirements for ALI as part of its rules for wireless carrier enhanced 911 (E911) service in CC Docket No. 94-102, *Revision of the Commission's Rules to Ensure Compatibility with Enhanced 911 Emergency Calling Systems.* Those rules were adopted in 1996 and revised in the Third Report and Order in that docket (released October 6, 1999). The revised rules set the following accuracy and reliability requirements for E911 Phase II operations:

- For network-based solutions: 100 meters for 67 percent of calls, 300 meters for 95 percent of calls;
- For handset-based solutions: 50 meters for 67 percent of calls, 150 meters for 95 percent of calls.

The obligation of carriers to deploy such wireless E911 technologies is scheduled to begin next year, in 2001. The standards and the phase-in schedule are incorporated in Section 20.18(g) of the Commission's Rules (47 C.F.R. § 20.18(g)).

III. GENERAL PRINCIPLES

<u>Default Fallback</u>: The wireless E911 rules require that location identification attempts be made and a location estimate provided to the Public Safety Answering Point (PSAP) for each 911 call. However, the Commission also recognized that, despite the importance of accurate location information to rapid emergency response, the inherent uncertainty of the radio technology used for wireless ALI may mean that location cannot always be reported accurately and determination of Phase II position fixes may not even be possible in some instances. The Commission accordingly adopted Phase II location accuracy requirements that apply to 95 percent of calls. Where Phase II location can not be provided, Phase I information, *i.e.*, the cell site or sector where the call is received, should be reported.

<u>Uncompleted Calls</u>: The failure to complete a call should be reported in testing results to ensure completeness and consistency, but only completed calls should be included in test statistics. In other words, ALI accuracy data should be calculated only for completed 911 calls. For purposes of assessing whether a 911 call can be completed, a test of analog cellular calls should include 911 calls which can be completed by either of the cellular carriers, consistent with Section 22.921 of the Commission's rules, 47 C.F.R. § 22.921.

<u>Timing</u>: Location information must be delivered to PSAPs within a reasonable time to permit its effective use by emergency response teams. This presents at least two separate issues. First, location information should be available as soon as possible, with little or no delay in normal call

delivery, to assist in routing the call to the correct PSAP and to provide rapid location information to the dispatcher. Second, location information is needed by emergency response teams responding to the call, who will benefit from more accurate location information. To accommodate both of these objectives, available location information should be delivered with call completion, but verification of the accuracy of the information may take place shortly after call completion. Any test protocol should identify the time to first fix (including fixes from Phase I or other location methods), which will be used to route calls to the proper PSAP, and should also employ a reasonable time limit for tests of location accuracy. An acceptable time limit for such testing is 30 seconds after the call is sent. Multiple attempts to determine location may be made within that period and the latest location data based upon these attempts within the period may be used in calculating accuracy. In evaluating compliance, recommendations by the National Emergency Number Association and standards committees regarding time limits for location accuracy measurement should be considered.

Motion: Any ALI system should operate effectively in conditions where 911 calls are made – for example, from within vehicles moving at highway speeds. Certain system developers have tested ALI delivery for calls from moving vehicles. From the information available, however, it appears that such calls present significant additional testing issues for verification purposes. Resolution of these issues should not delay completion of initial testing. Verification procedures should proceed and calculations of accuracy should be prepared even though it may not be possible initially to include information regarding calls in motion. The industry should work to further develop methods of verifying the accuracy of calls in motion, and as such methods are refined, they should be incorporated into E911 verification and testing protocols. This bulletin may be modified to accomplish such incorporation.

Coverage areas: Reports of compliance testing should clearly define the subject geographical areas. Accuracy tests may be based on the coverage areas of local PSAPs that request Phase II deployment. It may be appropriate to subject a wireless service provider's entire advertised coverage area within a metropolitan area or similar region to testing, as suggested by the National Emergency Number Association, but these are typically large areas and initial ALI deployment may proceed more gradually. Thus, testing may initially cover an urban core and later extend to the response area of a local PSAP. Compliance may be verified for these sub-areas separately or in combination. However, the areas delineated for compliance testing should not overlap. It is unacceptable to include the same geographic sub-area in two or more test areas, especially if the sub-area is relatively undemanding for the location technology.

Starting conditions of handset-based systems: For purposes of these guidelines, a reasonable test would begin with the handset in its normal off condition for at least 15 minutes before use in a test. "Normal off" means the condition in which wireless customers wishing to preserve battery life would carry the unit. However, it may be presumed that the customer will retain the unit in a stay-warm, low-current drain state if that feature is available.

Technical and competitive neutrality: Testing and validation methods should reflect the Commission's policy of technical and competitive neutrality. In general, any method should be applied in the same manner to all ALI technologies being used, except where some variation is clearly necessary and fully explained.

Post-installation monitoring and testing: Any empirical testing methodology should include protocols sufficient to test and verify that the system will continue to provide the required level of accuracy in normal operation. This methodology should reflect foreseeable variables that may affect ALI accuracy for the methodology subject to the test. These protocols may be performed in association with other carrier performance tests to the extent feasible. A biannual validation of accuracy should ordinarily be sufficient.

IV. POSSIBLE TESTING AND VERIFICATION METHODS

Any ALI testing and verification method should reasonably and accurately represent real world performance. This may be achievable through methods such as actual empirical testing of ALI equipment of systems in operation, or by predictive methods such as laboratory simulation or predictive models, if such predictive methods can be demonstrated to have the necessary robustness and accuracy.

It has been suggested by a number of parties that the Commission should adopt a "type-acceptance" approach, under which handsets would be subject to validation trials which, if successfully met, would lead to Commission approval of the handsets as satisfying the E911 ALI rules. It would be premature to adopt such an approach at this time. The relevant performance parameters that would justify type-acceptance have not been identified, especially for the varied environments in which ALI technologies must perform. In the future, an approval process might be possible as the technical and operational issues are clarified and relevant performance parameters are identified based on empirical data.

It has also been suggested that testing and verification issues be submitted to organizations such as the Telecommunications Industry Association (TIA) or that these issues be treated by the technology evaluation procedure under development for the CDMA Development Group (CDG), the CDG Test Plan. CDG submitted a draft test plan in nearly final form immediately prior to the publication of this Bulletin and that plan has not been reviewed. It is unclear how long it might take for TIA or other groups to develop acceptable standards.

Wireless manufacturers and carriers, in cooperation with public safety organizations and other interested parties, are encouraged to pursue and complete projects such as the CDG Test Plan, and to seek to develop standards under the auspices of organizations such as TIA, efforts which may obviate the need for extensive Commission guidelines and permit a type-acceptance approach through consensual standards or Commission approval processes. In general, the development of efficient, reliable, simple, cost-effective methods to test and verify ALI accuracy and reliability is encouraged.

V. EMPIRICAL TESTING METHODS

The empirical approach is a direct way to verify compliance with the accuracy standards. This approach may also be appropriate in situations where performance cannot be adequately modeled, for example, when adequate predictive models have not yet been developed, or where the models do not adequately represent an ALI system's performance in a particular operating area, or where the ALI system has undergone major modifications.

In outline, the empirical approach may proceed as follows: An accuracy measurement is made at each point of a sample set of locations randomly selected, preferably a data set representing 911 call locations. If 911 call location data is not available, other available, relevant data are also acceptable, such as the locations of ordinary wireless calls or population density. Tests are then performed at each of these sample locations to determine the distance between the actual location and the location reported by the ALI system. A sufficient number of observations should be included to establish compliance with the FCC accuracy requirements with a statistical confidence of at least 90 percent. See Appendix A for a statistical approach for demonstrating compliance for empirical testing.

Random Selection of Sample Locations

Because the selection of independent test locations in empirical testing is so important for meaningful determination of confidence, this document describes a possible procedure.

Assume that the outline of the geographic service area to be tested is given in a form for use by computer. Proceed as follows: Using one of the random number generator algorithms found on computers applied to scientific/engineering problems, generate a number of locations from random latitude-longitude pairs. The latitude and longitude numbers generated should at first be uniformly distributed inside a coordinate rectangle containing the operating area, and points out of the operating area itself subsequently dropped. If it is later found that a point is inaccessible, replace it by one no more than three meters beyond the nearest point on the boundary with an accessible area. If a point is found to be inside a building, the floor of the building may be selected when the testing crew arrives on site (see "vertical dimension" under the heading "conventions" below).

Models for Weighting

The relationships among radio environment, geographical features, and the inherent variability of location systems are complex and are only now becoming better appreciated and drawing attention. Wireless carriers, the public safety community and the companies involved in location technology have expressed the desire to take into consideration such factors as the likelihood that a wireless 911 call (or any wireless call) will be made from a particular location. Unfortunately, there is currently little factual information from which to establish or evaluate weighting factors. For purposes of these guidelines, reasonable assumptions may have to be made in order to begin to develop models. These assumptions might be based on consultation with public safety officials with experience in wireless 911 call patterns. Accordingly, if estimates of such weighting factors become available, they could be used as the likelihood, *e.g.*, of a wireless 911 call, from any particular point in the area. In the selection of test sites, these probabilities could be used in conjunction with the random selection of latitude and longitude of test sites to decide whether to test at a particular site, based on a Monte Carlo simulation of whether a wireless 911 call would be made from that site. As more experience is gained, more accurate weighting factors may be developed.

Conventions

Results will be most useful to others if all parties conducting empirical tests follow the same fixed set of conventions. The following are suggested to ensure reliability and comparability.

- Measurement precision: Latitude and longitude coordinates should be expressed in tenths of arc-seconds and distance errors from "ground truth" sample locations should be computed to the nearest meter. (This convention does not imply that the ALI system is able to determine position with this degree of resolution.) Conventional DGPS location equipment, which typically provides an accuracy of 3 to 5 meters, may be used to establish the coordinates of the sample locations if no more accurate means is available.

- Vertical dimension: The ALI system is not required to report a vertical dimension, *e.g.*, for tests made inside buildings, but sample locations in buildings should reflect the likelihood that calls will placed from different floors. The vertical dimension may be included in testing but this information should not be included in reported statistics for purposes of evaluating Phase II compliance.

- Statistical confidence: Conclusions derived from test results should be stated for at least a 90 percent level of confidence. A 90 percent level of confidence means that performance can be expected to be at least as good in 9 out of 10 test areas with the same relevant characteristics.

- Systematic errors: Instances of large errors or environments or locations where no Phase II fix is obtained shall be reported separately to help identify faulty test equipment, demanding environments, and areas for improvement. Separate treatment of outliers is not required for determining compliance with the Commission's accuracy and reliability requirements.

- Handset environment: Empirical testing should reasonably reflect the expected use of the handset, for example the effects of holding the handset next to the head and the effects of calling from within vehicles.

- Portable and mobile phones: Empirical testing should reasonably reflect the current distribution of handset types. However, because it appears likely that higher power mobile phones should perform at least as well as portable phones, tests may be performed using only portable phones.

- Analog and digital phone service: Separate data should be collected for each air interface addressed by the location system under test to identify any differences in performance. This data may be combined, with appropriate weighting, to compute overall performance.

VI. PREDICTIVE TESTING METHODS

Since predictive methods may be desirable or even essential for initial evaluation or selection of ALI systems and components, this document suggests the form predictive models might take and how they might be used. It is expected that predictive models will be based on a classification of different types of environments within which ALI technologies may be expected to vary in performance. The CDMA Development Group, for example, is constructing a catalog of environments including urban settings such as "urban canyons" (locations between high-rise buildings); suburban settings such as shopping plazas, residential houses and streets; a number of rural settings distinguished by terrain and vegetation coverage; and others. Environments should be further distinguished by the geometry of base station locations where this is a factor in the performance of the system being modeled.

General Form of Prediction Models

Because of the way the FCC standards are stated, models of ALI performance will presumably involve cumulative distribution functions representing expected location errors. To be useful, each such model will provide a prediction of the probabilities P_1 and P_2 corresponding respectively to location errors within the 67% and 95% FCC accuracy standards. For handset-based systems, for example, P_1 will be the predicted probability of location errors not exceeding 50 meters, P_2 the predicted probability of location errors not exceeding 150 meters. These values are 100 meters and 300 meters, respectively, for network-based solutions. In demanding environments, these probabilities may be less than 67% and 95%; in environments ideally suited to the ALI technology, the probabilities will exceed the standard values.

Ideally, the models should be clearly defined in a manner appropriate for publication in technical journals. A measure of statistical confidence should be established for each type of environment where a particular model is to be applied. The confidence measure should be established by analysis of the results of applying the model in a significant number of different locations.

Application of Prediction Models

If the probabilities P_1 and P_2 are less than the target values of 67% and 95% in some environments, it will be necessary to form a weighted average based on the possibility or likelihood of E911 calls from various locations in the service area under consideration. Compliance with FCC accuracy would then be demonstrated in terms of the weighted sums. The total service area should be divided into smaller areas with probabilities P_1 and P_2 in each predicted by model parameters specifically applicable to the respective sub-area. The weights appropriate for these sub-areas will have to be established using locally available data such as maps, calling statistics from various cell sites, and the experience of public safety personnel. Although no required level of statistical confidence has been established, a method should be established to show how closely the prediction model results approach the FCC accuracy criteria and its significance.

GLOSSARY OF TERMS

Automatic Location Identification (ALI) - Delivery of the location of a wireless handset to a PSAP without the need for inquiry by the dispatcher

Differential GPS (DGPS) – A method for correcting inaccuracies in GPS location calculations by use of signals from a terrestrial reference station.

Enhanced 911 (E911) – An emergency telephone system using the digits 9-1-1 that provides additional information to the emergency dispatcher, such as Automatic Number Identification and Automatic Location Identification.

Global Positioning System (GPS) – A network of 24 U.S. government satellites, supported by ground control systems, transmitting radio signals that can be decoded to compute precise locations.

Handset-based Location Technology - A method of providing the location of wireless 911 callers that requires the use of special location-determining hardware and/or software in a portable or mobile phone. Handset-based location technology may also employ additional location-determining hardware and/or software in the wireless network and/or another fixed infrastructure.

Network-based Location technology - A method of providing the location of wireless 911 callers that employs hardware and/or software in the wireless network and/or another fixed infrastructure, and does not require the use of special location determining hardware and/or software in the caller's portable or mobile phone.

Public Safety Answering Point (PSAP) – A 911 answering station designated to receive 911 calls from a specific geographic area.

Phase I E911 – The first step in implementing wireless E911. Under Phase I, as of April 1, 1998, licensees subject to the E911 rules must provide the telephone number of the originator of the 911 call and the location of the cell site or base station receiving the call from any mobile handset accessing their systems to the designated PSAP. This requirement applies only if certain conditions are met: that the PSAP has requested the service and is capable of receiving and utilizing the data, and that a mechanism for recovery of the PSAP's costs is in place.

Phase II E911 – The second step in implementing wireless E911. Under Phase II, as of October 1, 2001, licensees subject to the E911 rules must provide to the PSAP the location of all 911 calls by longitude and latitude in conformance with specified accuracy requirements, subject to the same conditions that apply to Phase I. Wireless carriers are required to report their plans for implementing Phase II, including the technology they plan to use to provide caller location, by October 1, 2000.

APPENDIX A

A Statistical Approach for Demonstrating Compliance for Empirical Testing

Summary: A method for determining whether or not a set of location errors resulting from empirical testing demonstrates compliance may be obtained from order statistics.[1] Confidence intervals for the FCC's accuracy standards may be selected based on the 90% confidence level and the number of samples. These confidence intervals are not based on any knowledge of the actual probability distribution function of the location errors. They are expressed in terms of the subscripts of the list of location errors after ordering these errors from smallest to largest. A specific set of accuracy measurements is said to show compliance if the confidence intervals contain the location error thresholds established by the FCC. These are 100 meters for 67% and 300 meters for 95% for network-based solutions, or 50 meters and 150 meters, respectively for handset-based solutions. Otherwise, compliance would be rejected.

Approach:

In general, when

- the number of measurements is n,
- the r^{th} and s^{th} largest measurements are x_r and y_s respectively, and
- x and y are the percentile points associated with probabilities p_1 and p_2 respectively, then –

the probability that x is less than x_r while simultaneously y is less than y_s is given by the formula

$$confidence(x \le x_r,\ y \le y_s;\ n, r, s,\ p_1,\ p_2) = \sum_{i=0}^{r-1}\sum_{j=i}^{s-1}\binom{n}{i}\binom{n-i}{n-j}p_1^i\left(p_2-p_1\right)^{j-i}\left(1-p_2\right)^{n-j}.$$

In the present application, p_1 is 0.67, and p_2 is 0.95.

Upper bounds on the percentile points can be determined from this expression by finding pairs of values (r,s) such that the desired 90% confidence level is achieved. The resulting pair of ordered samples (x_r,y_s) forms one-sided confidence intervals for the two sample percentile points associated with 67% and 95%, respectively. See Table 1. The r^{th} sample x_r and s^{th} sample y_s of n location errors are then compared with 100 meters and 300 meters for the networked-based solutions or with 50 meters and 150 meters for the handset-based solutions. If the r^{th} ordered sample is less than 100 meters and the s^{th} ordered sample is less than 300 meters, then the confidence intervals are found to cover the desired values and compliance would be established, for networked-based solutions. A similar approach would establish compliance for a set of location errors obtained from a test of a handset-based solution.

The confidence level of 90% is suggested here as a threshold, and the value calculated from the actual data may be greater. Table 1 is derived from the above confidence expression and shows for several sample sizes which ordered samples of errors should be compared with the FCC

[1] Barry C. Arnold, N. Balakrishnan and H. N. Nagaraja, *A First Course in Order Statistics*, Wiley Series in Probability and Mathematical Statistics, July 1992.

criteria. For higher numbers of sample sizes such as 500 or 1000, the confidence expression should be re-calculated with the higher value of n.

Confidence in the compliance assessment is important and will depend on randomness and independence in the selection of test locations. Reports of compliance testing should describe the method used to guarantee random and independent accuracy measurements.

Table 1

IDENTIFICATION OF LOCATION-ERROR SAMPLES FOR COMPARISON WITH FCC REQUIRED THRESHOLDS FOR 67% AND 95%
(At the 90% confidence level)

Sample Size	Pairs of Test Samples
45	(x_{40}, y_{45})
50	(x_{41}, y_{50})
55	(x_{44}, y_{55})
60	(x_{47}, y_{60})
65	(x_{50}, y_{65})
70	(x_{53}, y_{70})
75	(x_{57}, y_{75})
80	(x_{60}, y_{80}) or (x_{63}, y_{79})
85	(x_{64}, y_{85}) or (x_{66}, y_{84})
90	(x_{67}, y_{90}) or (x_{68}, y_{89})
95	(x_{71}, y_{95}) or (x_{72}, y_{94})
100	(x_{74}, y_{100}) or (x_{75}, y_{99})

Illustration: A handset-based solution would be found in compliance, if, in a test of 75 accuracy measurements, the 57th largest location error is less that 50 meters and the 75th largest error is less than 150 meters. Note that for larger sample sizes the pair of test samples is not unique, because of the statistical dependence of the 67% and 95% levels. For example, for a sample size of 80, two pairs are shown; the 67% level could be increased from the 60th to the 63rd sample, (*i.e.*, made more difficult), if the 95% level test were relaxed to the 79th largest sample. Either (x_{60}, y_{80}) or (x_{63}, y_{79}) is an acceptable pair to test against the FCC-required thresholds.