Adaptive Algorithms for Reducing PCS Network Authentication Traffic

Yi-Bing Lin, *Senior Member, IEEE*, Seshadri Mohan, *Member, IEEE*, Nelson Sollenberger, *Fellow, IEEE*, and Howard Sherry, *Member, IEEE*

***Abstract*— Two authentication schemes (AS's) have been proposed in the Electronic Industry Association/Telecommunications Industry Association's (EIA/TIA) Telecommunications Systems Bulletins (TSB's) 51 for incorporation in the cellular industry Interim Standard IS 41 Revision C. In the first scheme, which we refer to as the WS scheme, a unique secret key [also known as shared secret data (SSD)] is shared only between the authentication center (AC) and handset. In the second scheme, referred to as the S scheme, the SSD is also shared with the visited system. The performance of the two schemes can be compared by using the expected number of call originations/terminations between two consecutive registrations or call-to-mobility ratio (CMR). Based on the message flow (accesses to databases), it is apparent that the S scheme outperforms the WS schemes if the CMR ratio is large. The CMR's of users will likely span a wide range and, even for the same user, will perhaps vary from time to time. It is therefore desirable to switch between the AS's based on the users' CMR to reduce the signaling network traffic. We propose two adaptive algorithms to determine how and when to switch between the AS's for a given user. Our performance study indicates that as the CMR of a user changes, the adaptive algorithms automatically select the best AS in real time.**

***Index Terms*— Authentication, IS-41, mobility management, personal communications services.**

I. Introduction

THE STANDARD commonly used in North America for locating users is the Electronic Industry Association/Telecommunications Industry Association (EIA/TIA) Interim Standard 41 (IS 41) [2]. It employs a two-level hierarchical strategy, which maintains a system of home and visited databases [home-location registers (HLR's) and visitor-location registers (VLR's)]. In [5], we have analyzed the network impact of IS 41 for some scenarios such as call origination, call delivery, and registration/deregistration. Since then, some additions have been proposed in EIA/TIA Telecommunications Systems Bulletins (TSB's) 51 [3], defining protocols for authentication, voice privacy, and signaling message encryption, which are expected to be incorporated into IS 41 Revision C. The algorithm for authentication and generation of voice privacy mask and signaling message encryption keys employed by the standard is based on private key cryptographic techniques in which a secret key [also known as shared secret data (SSD)] is shared between the handset and authentication center (AC) and is known only to these two parties.

Two authentication schemes (AS's) have been proposed in TSB 51. In the first scheme, which we call the WS scheme, the SSD is shared only between the AC and handset. In the second scheme, or the S scheme, the SSD or some aspect of SSD may be shared with the visited system as well. In the S scheme, since the visited system has the SSD, it can authenticate the handset at call origination or delivery and thereby considerably reduce message flow and call setup time. However, as will be seen later, during registration, this scheme requires additional message flow as compared to the WS scheme. Thus, there exists a tradeoff between the two schemes based on the call-to-mobility (CMR) ratio of a user (i.e., the expected number of calls to/from the user between two consecutive registrations). For a large CMR, sharing the SSD with the visited system is beneficial and, consequently, the S scheme is preferable. For a small CMR, sharing SSD with the visited system results in considerably increased traffic and, consequently, the WS scheme is preferable. For a given user, the CMR may vary from time to time, and, therefore, it is desirable to switch between the two AS's as the user's behavior changes. We propose two adaptive algorithms to select the AS for a user. Our performance study indicates that as the CMR of a user changes, the adaptive algorithms automatically select the appropriate AS in real time.

II. Privacy and Authentication in IS-41

This Section describes the two EIA/TIA TSB 51 schemes for privacy and authentication. It is helpful to first describe the message flow due to the WS scheme and then to show how the flow differs for the S scheme. To facilitate the discussion that follows, we introduce the following notions and terminologies. The AC is solely responsible for maintaining and updating the SSD's. Associated with each terminal is a terminal identifier mobile identification number (MIN) and an electronic serial number (ESN). A MIN is a North American numbering plan (NANP) number that is stored in the terminal at the time of manufacture and cannot be changed. Users move about registration areas (RA's) belonging to one or more personal communications system (PCS) service providers (PSP's). Each PSP may provide some combination of base stations (BS's) for providing wireless access to the PCS handsets, mobile switching centers (MSC's) that control the BS's associated with it. See [1] for a discussion of how the advanced intelli-

Manuscript received July 27, 1995; revised August 6, 1996.

Y.-B. Lin is with the Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan, R.O.C. (e-mail: liny@csie.nctu.edu.tw).

S. Mohan and H. Sherry are with Bellcore, Morristown, NJ 07960 USA.

N. Sollenberger is with the Wireless Systems Research Department, AT&T.

Publisher Item Identifier S 0018-9545(97)04632-X.

gent network-based public telephone network may be used to provide access services to PSP's. Associated with each registration area is a VLR. The VLR's may be part of the PSP network. It is likely that the VLR's are collocated with MSC's. For demonstration purposes, we assume that VLR's are separated from the PSP network. Note that the results of this paper are not affected by the locations of VLR's. One or more HLR's maintained by the local-exchange-carrier (LEC) network exist that maintain user profiles, current registration area, etc.

A. The WS Scheme

Message flow for privacy and authentication for the WS scheme is described below.

1) Message Flow Due to Registration: Refer to Fig. 1 for the message flow due to registration.

1) A handset determines, based on the signal transmitted by the base station, that it has entered a new RA and that authentication is required to access PSP services. It executes the cellular authentication and voice encryption (CAVE) algorithm using the SSD, its ESN, MIN, and a random number (RAND) obtained from the PSP system at that time. The algorithm produces a registration authentication result (AUTHR).
2) The handset requests registration with the PSP by supplying the authentication result AUTHR, its ESN, MIN, the most significant 8 b of RAND denoted as RANDC, and a certain COUNT giving an account of the most significant events such as registration, call origination, termination, etc., initiated by the handset. This call history count is also maintained by the AC.
3) The PSP system forwards the authentication request in a AUTHRQST message to the VLR serving the PSP registration area.
4) The VLR forwards the request to the HLR along with all the parameters it received.
5) In turn, the HLR forwards the authentication request to the AC.
6) The AC retrieves from its database the SSD associated with the MIN, and using the retrieved SSD, executes the CAVE algorithm, with the additional parameters (MIN, ESN, and RAND) it received from the HLR, to produce the authentication result.
7) After verifying that the generated result matches AUTHR it received from the handset via HLR and the COUNT value it has stored matches the COUNT value supplied by the handset, the AC provides its response to the authentication request in an AUTHRQST message, which, eventually, gets forwarded to the visited PSP system.

Once the handset is authenticated, the serving PSP system will initiate a registration notification message. See [2] and [8] for details.

2) Message Flow Due to Call Origination: Fig. 2 illustrates message flow generated when a handset originates a call in the visited PSP system. The steps below describe the activities that take place.

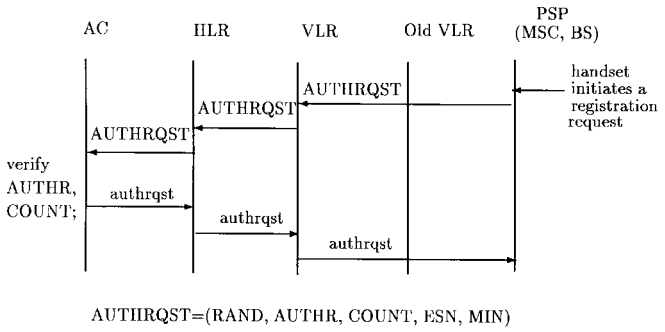

Fig. 1. Message flow due to registration for the WS scheme.

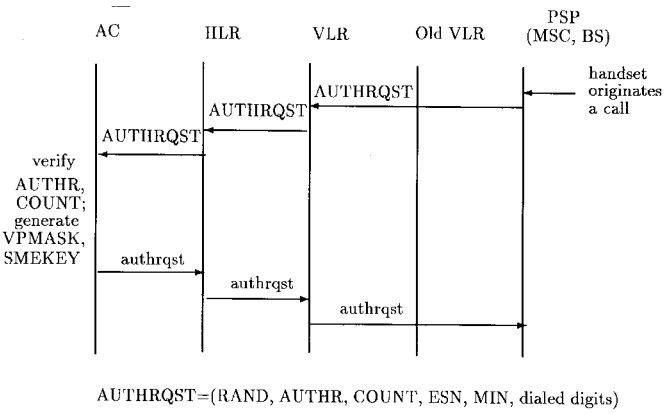

Fig. 2. Message flow due to call origination for the WS scheme.

1) The handset executes the CAVE algorithm with the digits dialed by the user as input along with the same input parameters as in Step 1) of registration message flow to produce an authentication request result AUTHR, a voice privacy mask VPMASK,[1] and a signaling message encryption key SMEKEY.[2]
2) The handset then forwards to the serving PSP system all the parameters as before in Step 2) of the registration message flow along with the dialed digits.
3) The remainder of the steps are as before with the addition that the AC now generates the VPMASK and SMEKEY, which get forwarded to the serving PSP system.

Once the handset is authenticated, the serving PSP system will initiate an IS 41 LOCREQ (location request) message to determine the called party's current serving PSP (see [2] for details).

3) Message Flow Due to Call Termination: Following the completion of IS 41 LOCREQ, the PSP system serving the called handset pages it. The handset responds to the page by executing the CAVE algorithm and sending AUTHR, COUNT, ESN, MIN, and RANDC. The message flow that follows is similar to that of the previous two cases and will not

[1] VPMASK is applied to voice transmission over the air interface between the handset and serving PSP system.

[2] SMEKEY is used to encrypt certain fields within the signaling message sent by the handset to the serving PSP system. VPMASK and SMEKEY are generated by the AC and forwarded to the serving PSP system.

be repeated here. Once the called handset is successfully authenticated, a voice channel is established for the call between the two handsets.

B. The S Scheme

When SSD is shared with the visited PSP system, authentication of the handset during registration requires additional steps compared to that in Section II-A. This is due to the fact that the VLR at the previous system has the current value of COUNT, and, therefore, the AC needs to request COUNT from the previously visited PSP system's VLR. Once the handset is registered with the new VLR, for all other system accesses, such as call origination, termination, and flash request (three party call), the VLR can authenticate the handset without having to forward the authentication request to the AC. The message flow is, therefore, considerably reduced when compared with the WS scheme.

1) Message Flow Due to Registration: Fig. 3 illustrates message flow during registration when SSD is shared with the visited system.

1) The handset determines that it has entered a new RA and that authentication is required to access PSP services. It executes the CAVE algorithm using the SSD, its ESN, MIN, and RAND obtained from the PSP system at that time. The algorithm produces a registration authentication result (AUTHR).
2) Further steps proceed as in Steps 2)–6) of message flow due to registration in Section II-A at which point the AC has completed verification of the authentication result AUTHR generated by the handset. However, it does not have the current value of COUNT, and, therefore, it cannot verify what is supplied by the handset in its registration request. It needs to get the current value from the VLR of the previously visited PSP system.
3) The AC sends a COUNTREQ message to the HLR intended for the previously visited VLR, requesting the current value of COUNT.
4) The HLR forwards the request to the VLR for the handset it is pointing to, which is actually the one serving the RA the handset just moved out of.
5) The VLR responds to the request by sending COUNT in a countreq message.
6) The HLR forwards it to the AC.
7) The AC verifies the COUNT and sends an authentication result in an AUTHRQST message to the requesting system.

2) Message Flow Due to Call Origination: Fig. 4 illustrates the message flow when a handset in the visited system initiates a call.

1) The handset executes the CAVE algorithm with input parameters as in Step 1) of call origination for the WS scheme of Section II-A to generate AUTHR, SMEKEY, and VPMASK and sends RANDC, AUTHR, COUNT, ESN, MIN, and the dialed digits to the visited PSP system.
2) The visited PSP system sends an authentication request AUTHRQST message to the serving VLR.

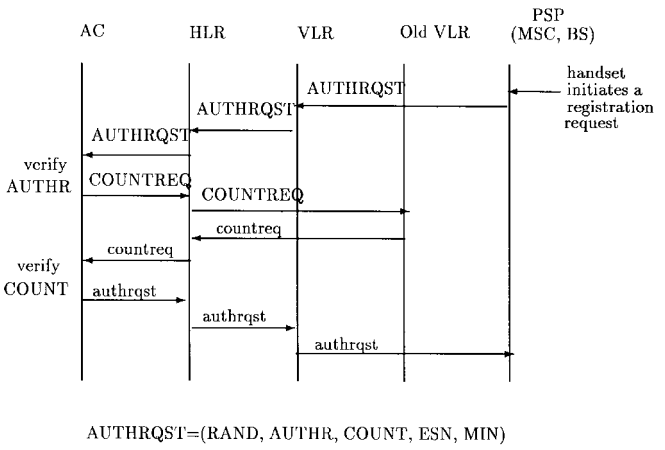

Fig. 3. Message flow due to registration for the S scheme.

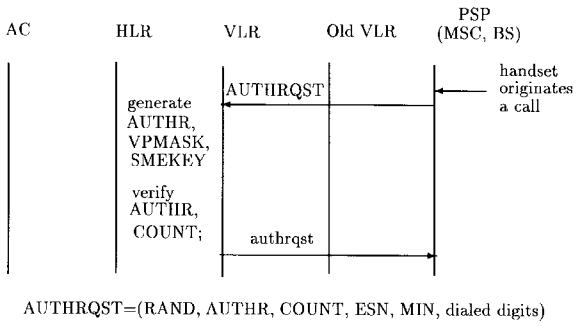

Fig. 4. Message flow due to call origination for the S scheme.

3) The VLR executes the CAVE algorithm and generates AUTHR, VPMASK, and SMEKEY.
4) After verifying AUTHR and COUNT, it includes the result of the verification along with VPMASK and SMEKEY in its AUTHRQST response message to the visited system.

The message flow for flash request (e.g., adding a third party after a call is established) and an SSD update as dictated by administrative needs are not described in this paper (see [3]).

III. The Adaptive Algorithm—AA 1

The WS scheme is appropriate when the number of registration operations is larger than the call origination/termination operations. On the other hand, the S scheme is appropriate in the opposite situation. For simplicity, we regard events such as flash requests and COUNT updates as being infrequent in comparison to call originations/terminations and registrations and, therefore, ignore them in the following analysis. These events can be easily accommodated in our analysis. Note that the "cost" of sending a message is determined by the length of the message and the "distance" between the sender and receiver locations. To simplify and strengthen our results, we assume that the costs of sending messages from one location (e.g., AC, HLR, or VLR) to another are the same. Thus, the number of the delivered messages can be considered as the cost of the AS. To reduce the number of the database (AC/HLR/VLR) accesses, it is important to choose an appropriate AS for a user. This section describes our first adaptive

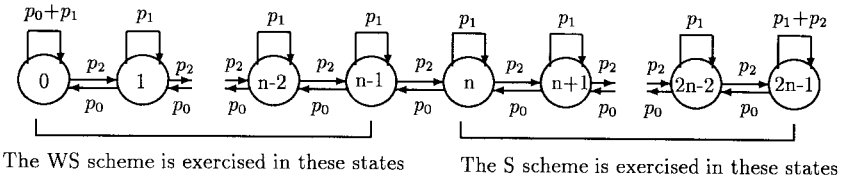

Fig. 5. The state diagram for AA 1.

algorithm (AA 1), which automatically selects an appropriate AS for any given user in real time based on the user behavior (in terms of registration, call origination, and call termination). The concept of adaption used in AA 1 is similar to the one in the cache scheme (for mobility management) proposed in [7]. However, the heuristics used to determine the adaption are very different in both schemes.

Define a *cycle* as the period between two consecutive registrations for a user. Define the *cost* of a cycle as the number of messages sent to access AC/HLR/VLR due to the call originations/terminations in the cycle and the registration at the beginning of the cycle. (Note that the cost does not include the registration at the end of the cycle.) The costs of the S scheme and WS scheme are computed as follows.

Let λ be the call-arrival rate and μ be the mobility (the frequency that a user changes RA's). Then, the CMR ρ (the expected number of call arrivals during a cycle) is

$$\rho = \frac{\lambda}{\mu}.$$

In the WS scheme, five database accesses are required to authenticate a registration operation (see Fig. 1), and five database accesses are required to authenticate a call origination operation or a call termination operation (see Fig. 2). Our cost analysis excludes the messages sent from the VLR's to the PSP. Note that the number of messages sent from the VLR's to the PSP as well as the number of messages sent within the PSP system (i.e., between MSC and BS) are the same for both AS's [3]. Thus, including these messages in our cost analysis will not change the results. In the S scheme, nine database accesses are required to authenticate a registration operation (see Fig. 3) and one database access is required to authenticate a call origination operation or a call termination operation (see Fig. 4). The expected numbers of database (AC/HLR/VLR) accesses during a cycle for the WS scheme and S scheme are $C_{ws} = 5 + 5\rho$ and $C_s = 9 + \rho$, respectively. It is interesting to note that

$$C_{ws} = C_s \Leftrightarrow \rho = 1.$$

Also, the S scheme outperforms the WS scheme (i.e., $C_s < C_{ws}$) if and only if $\rho > 1$.

AA 1 is a $2n$-state finite automaton with the state diagram shown in Fig. 5. The WS scheme is exercised if AA 1 is in state i for $0 \leq i \leq n-1$. The S scheme is exercised if the algorithm is in state j for $n \leq j \leq 2n-1$. The transition probabilities for the finite automaton are

$$p_0 = \Pr[L = 0]$$
$$p_1 = \Pr[L = 1]$$

and

$$p_2 = \Pr[L > 1]$$

where L is the number of call arrivals during the previous cycle. Note that if the steady state of the algorithm exists, then

$$\lim_{t\to\infty} p_0 = \Pr[\rho = 0]$$
$$\lim_{t\to\infty} p_1 = \Pr[\rho = 1]$$

and

$$\lim_{t\to\infty} p_2 = \Pr[\rho > 1].$$

AA 1 is easily implemented by incorporating the rules (to be described) for state transition within the AC. The AC needs to maintain $\lfloor \log_2 n \rfloor + 1$ AS bits per user.

The VLR will need to maintain an AS bit per user to indicate the type of AS applicable to that user.

When the AC is accessed for a registration operation, the AC checks the following. Suppose that the algorithm is in state i.

1) If no call arrived during the previous cycle, then the algorithm moves to state $i-1$ for $i > 0$ and remains in the same state i for $i = 0$.
2) If exactly one call arrived during the previous cycle, then the algorithm remains in the same state i.
3) If more than one call arrived during the previous cycle, then the algorithm moves to state $i+1$ for $i < 2n-1$ and remains in the same state i for $i = 2n-1$.

When the algorithm moves from state $n-1$ to state n, the authentication procedure switches from the WS scheme to the S scheme. Similarly, when the algorithm moves from state n to state $n-1$, the authentication procedure switches from the S scheme to the WS scheme. If $\rho < 1$, then the algorithm tends to move to state 0 and the WS scheme is exercised. On the other hand, if $\rho > 1$, then the algorithm tends to move to state $2n-1$ and the S scheme is exercised.

The cost C_{AA1} of AA 1 is given in (13) in Appendix A.

IV. The Adaptive Algorithm—AA 2

Our second adaptive algorithm (AA 2) requires an AS bit in the AC and VLR's to indicate whether the S or WS schemes are exercised. At the beginning of a cycle, AA 2 always exercises the WS scheme (i.e., the AS bit is "WS"). When a call arrives (either a call origination or a call termination), the S scheme is exercised (i.e., the AS bit is switched to "S"). The switching from the WS scheme to the S scheme occurs as follows. Without loss of generality, assume that all calls to the user are call originations. Consider a cycle.

1) When the first call arrives, the authentication message flow follows Fig. 2, except that when the AC receives

AUTHRQST, the AS bit at the AC is switched to "S" and the SSD is sent to the VLR in the AUTHRQST message. When the VLR receives the SSD, its AS bit is set to "S." At this moment, the S scheme is exercised.

2) For subsequent call arrivals in this cycle, the message flow in Fig. 4 is followed.
3) At the end of the cycle (i.e., the handset moves to a new RA, and the authentication/registration occurs), the AUTHRQST messages are sent to the AC.
   a) If the AS bit at the AC is "WS" (i.e., no call origination/termination occurs during the cycle), then the message flow follows (see Fig. 1).
   b) If the AS bit at the AC is "S" (i.e., there are call originations/terminations during the cycle), then the message flow follows Fig. 3. The AS bit at the AC is set to "WS." When the VLR receives the AUTHRQST message, its AS bit is set to "WS."

   At the end of Step 3), the WS scheme is exercised.

Let $E[N|N \geq 1]$ be the expected number of call arrivals during a cycle under the condition that at least one call arrives in that cycle. The cost C_{AA2} of a cycle (the costs of the call authentications in this cycle and the cost of the registration authentication at the end of this cycle) for this algorithm is expressed as

$$C_{\mathrm{AA2}} = 5p_0 + (4 + E[N|N \geq 1] + 9)(1 - p_0). \quad (1)$$

The cost (1) consists of two parts. If no call arrives during the cycle (with probability p_0), the number of authentication messages sent in the cycle is five. If at least one call arrives (with probability $1 - p_0$), the first call requires five authentication messages [see Step 1)], and a subsequent call requires one authentication message [see Step 2)]. The total number of the authentication messages sent in a cycle is $4 + E[N|N \geq 1]$. At the end of the cycle, nine authentication messages are required to switch from the S scheme to the WS scheme [see Step 3)]. Since

$$E[N|N \geq 1] = \frac{\rho}{1 - p_0}$$

(1) is rewritten as

$$C_{\mathrm{AA2}} = 5p_0 + \left[4 + \frac{\rho}{1 - p_0} + 9\right](1 - p_0) = 13 - 8p_0 + \rho. \quad (2)$$

The probability p_0 is derived in (3) in Appendix A. If the RA residence time has a Gamma distribution, p_0 is expressed as (7) in Appendix A.

V. Discussions

This Section discusses the performance of the adaptive algorithms by assuming that the RA residence-time distribution is Gamma with mean $1/\eta$ and shape parameter γ. As we mentioned in Appendix A, the Gamma distribution is selected because it can be used to approximate many other distributions.

Fig. 6 compares C_{ws}, C_s, C_{AA1} (two-state AA 1), and C_{AA2}, where the RA residence-time distribution is exponential (i.e., Gamma distribution with $\gamma = 1$). The figure indicates

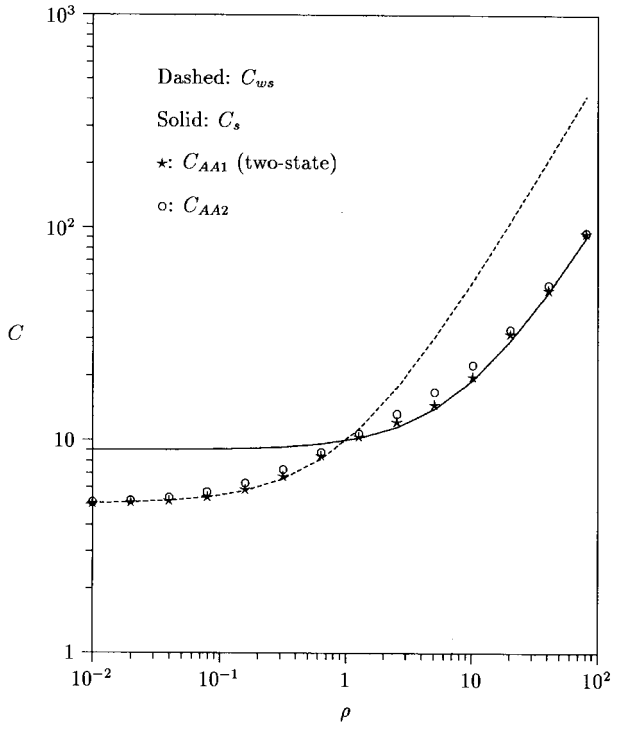

Fig. 6. The comparison of C_{ws}, C_s, $C_{\mathrm{AA}\,1}$ (two-state AA 1), and $C_{\mathrm{AA}\,2}$ ($\gamma = 1$).

that for $\rho < 1$, $C_{\mathrm{AA2}} \simeq C_{\mathrm{AA1}} \simeq C_{ws} < C_s$. For $\rho > 1$, $C_{\mathrm{AA2}} \simeq C_{\mathrm{AA1}} \simeq C_s < C_{ws}$. In other words, two-state AA 1 and AA 2 automatically switch to the appropriate AS (either the S or WS scheme). Note that the performance of AA 1 is slightly better than AA 2.

Fig. 7(a) shows the effect of the shape of the RA residence-time distribution on C_{AA1} for two-state AA 1. The figure indicates that the performance of two-state AA 1 worsens as γ decreases. This effect is easily explained as follows.

Note that for the Gamma RA residence-time distribution, the variance increases as γ decreases. Suppose that $\rho > 1$. As the variance of f_m increases, the number of cycles with call arrivals exceeding ρ and the number of cycles with no call arrivals will increase. Due to the increased zero-call-arrival cycles, the probability that two-state version of AA 1 moves into state 0 increases (and, thus, the WS scheme is likely to be exercised). Similarly, when $\rho < 1$, the probability that the AA 1 moves into state 1 increases. Consequently, the cost of the AA 1 C_{AA1} increases.

Fig. 7(b) shows the effect of the shape of the RA residence-time distribution on C_{AA2} for AA 2. It is interesting to note that the effect of γ on C_{AA2} is opposite to that on C_{AA1}. For a fixed ρ, C_{AA2} increases as p_0 decreases [see (2)]. Since p_0 decreases as the variance of the RA residence-time distribution decreases (i.e., γ increases), C_{AA2} increases as γ increases. Fig. 7(a) and (b) indicate that C_{AA2} is less sensitive to γ than C_{AA1} is.

Fig. 8(a) compares C_{AA1} (two-state) and C_{AA2} for $\gamma = 10$ (low-variance case). The figure indicates that AA 1 (two-state) is better than AA 2 when the RA residence-time distribution has a small variance. However, the improvement is not very

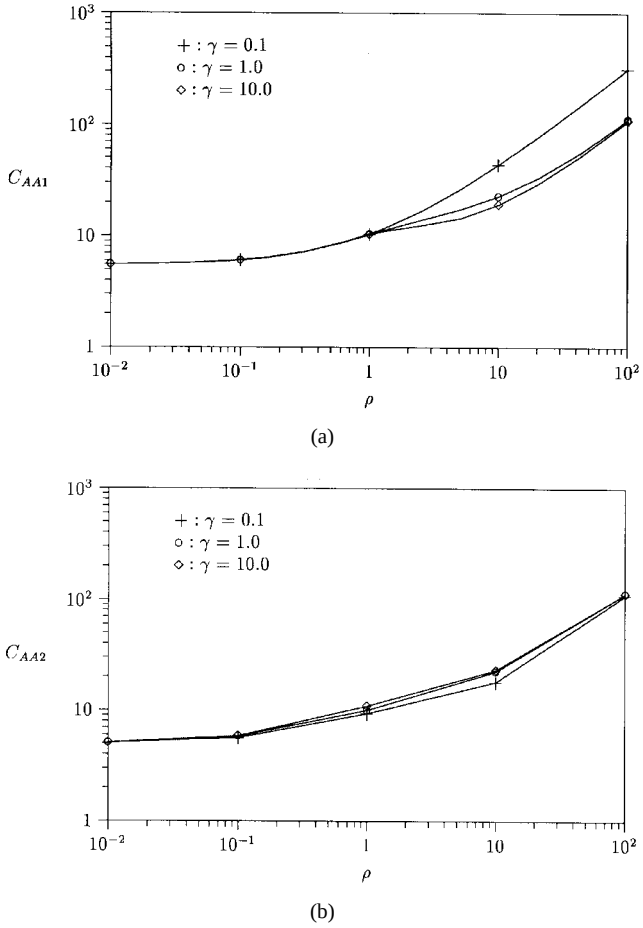

Fig. 7. The effect of γ on C_m for two-state AA 1 and AA 2: (a) algorithm 1 and (b) algorithm 2.

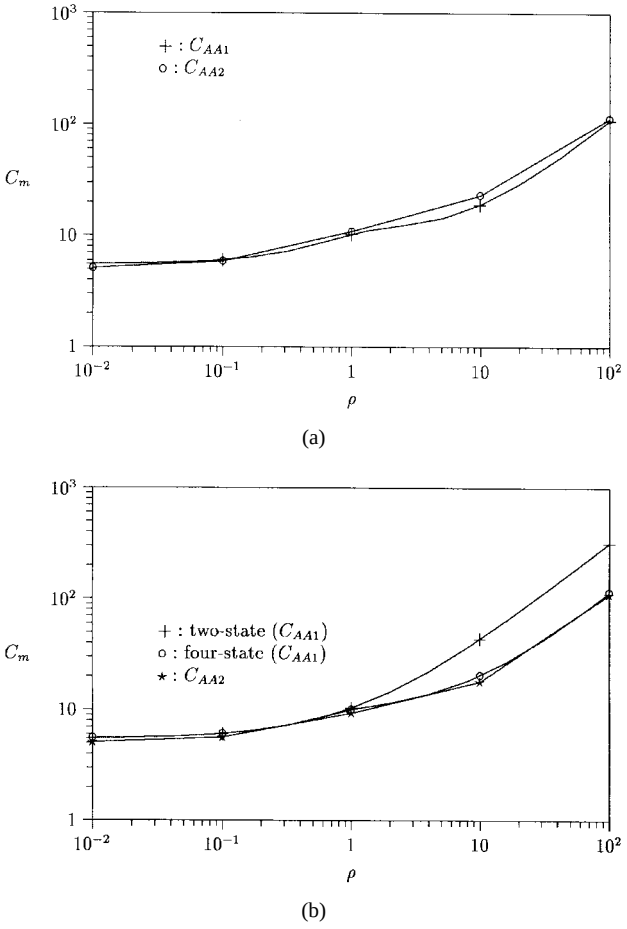

Fig. 8. The effect of γ on C_m: (a) $\gamma = 10$ and (b) $\gamma = 0.1$.

significant. Fig. 8(b) compares C_{AA1} (two-state), C_{AA1} (four-state), and C_{AA2} for $\gamma = 0.1$ (high-variance case). The figure indicates that AA 2 is better than two-state AA 1 when the RA residence-time distribution has a small variance.

The performance of two-state AA 1 can be improved by increasing the number of states. The problem of the two-state algorithm is that it does not have enough memory to count the number of cycles of large ρ (and small ρ). The memory of AA 1 can be increased by increasing the number of states. For large n, cycles with $\rho > 1$ tend to move the automation to state $2n - 1$. When cycles with $\rho = 0$ are encountered, the automation may move to states closer to state $n - 1$. By increasing the number of states, the chances of the automation entering a state smaller than n are reduced. Fig. 8(b) indicates that the effect of small γ (or high variance) is reduced by increasing the number of states from two to four. For example, at $\rho = 10$ four-state AA 1 achieves more than 50% reduction in cost as compared with two-state AA 1, which is almost identical to the performance of AA 2.

Note that the number n of the states in AA 1 cannot be arbitrarily increased for the following reason. If the user's CMR changes, the algorithm reflects to the change slower if n is larger. Let $E[N]$ be the expected number of the cycles for the AA 1 to move from state 0 to state n. Appendix B provides a derivation of $E[N]$. Fig. 9 plots $E[N]$ when $\rho \geq 1$. The figure indicates that it takes more cycles for a four-state

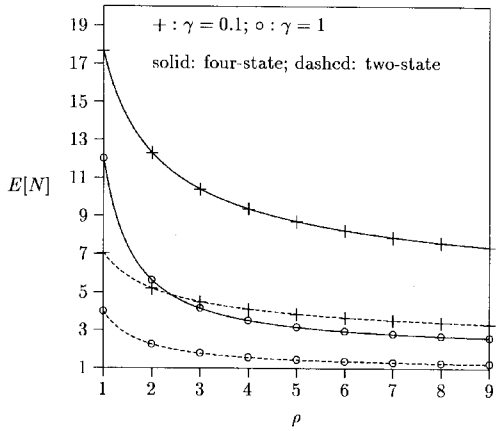

Fig. 9. The expected number $E[N]$ for AA 1 to move from state 0 to state n when $\rho \geq 1$.

algorithm to switch to the appropriate algorithm (e.g., the S scheme in this case) than a two-state algorithm.

VI. Conclusions

We described two AS's proposed for inclusion in the cellular industry Interim Standard IS 41. The performance of the two schemes was evaluated and parameterized by the CMR. For a given PCS user, the WS scheme is appropriate when the CMR is small, while the S scheme is appropriate when the CMR is large. Then, we proposed two adaptive algorithms to

select the appropriate AS for a user based on the user's CMR. The adaptive algorithms do not require any modifications to the handsets HLR's and VLR's. The selection of the AS is determined by the AC.

AA 1 is implemented as a $2n$ state-finite automation (for $n \geq 1$). When the automation is in states $0, 1, \cdots, n-1$, the WS scheme is exercised. Otherwise, the S scheme is exercised. The state of the automation changes based on the number of the call arrivals in a registration cycle. AA 2 switches from the WS scheme to the S scheme if a call arrives in a registration cycle. We showed that both algorithms are able to capture the CMR value to select the appropriate AS in real time.

The differences between the two adaptive algorithms are described below.

1) AA 2 is easier to implement than AA 1. The implementation complexity of two-state AA 1 is about the same as that of AA 2. In some cases, it may be desirable (see the next item) to increase the number of states in AA 1, which somewhat increases the implementation complexity.
2) AA 2 is better than two-state AA 1 if the variance of the RA residence-time distribution for a user is large. In this case, The performance of AA 1 can be improved by increasing the number of states of the automation. The performance of four-state AA 1 is about the same as AA 2.
3) AA 1 is better than AA 2 if the variance of the RA residence-time distribution is small.

APPENDIX A
DERIVATION OF C_{AA1}

This section studies the performance of AA 1. Let t_m be the interval that the user resides in an RA. Let f_m be the density function of t_m, $E[t_m] = 1/\mu$, and the Laplace transform be

$$f^*_{\mathbf{m}}(s) = \int_{t_m=0}^{\infty} f_m(t_m) e^{-st_m}\, dt_m.$$

Since the call arrivals form a Poisson process and because of the excess life property [9] of the exponential distribution, the probability that L calls arrive during the time period $X = t_m$ is

$$\Pr[L = l | X = t_m] = \frac{(\lambda t_m)^l}{l!}\, e^{-\lambda t_m}.$$

We have

$$p_0 = \Pr[L=0] = \int_{t_m=0}^{\infty} \Pr[L=0|X=t_m] f_m(t_m)\, dt_m = f^*_{\mathbf{m}}(\lambda) \tag{3}$$

$$p_1 = \Pr[L=1] = \int_{t_m=0}^{\infty} \Pr[L=1|X=t_m] f_m(t_m)\, dt_m = \lambda \int_{t_m=0}^{\infty} t_m f_m(t_m) e^{-\lambda t_m}\, dt_m = (-\lambda)\left[\frac{df^*_{\mathbf{m}}(s)}{ds}\right]\Big|_{s=\lambda}$$

$$p_2 = \Pr[L \geq 2] = \sum_{l=2}^{\infty} \int_{t_m=0}^{\infty} \Pr[L=l|X=t_m] f_m(t_m)\, dt_m = \sum_{l=2}^{\infty} (-1)^l \lambda \left[\frac{d^l f^*_{\mathbf{m}}(s)}{ds^l}\right]\Big|_{s=\lambda} = 1 - p_0 - p_1. \tag{4}$$

We assume that the RA residence times have a Gamma distribution [6]. The Gamma distribution is considered because it can be shaped to represent many distributions as well as measured data by selecting appropriate values for the two parameters of the distribution. Thus, one may measure the user residence times in a real PCS network, and the measured data can be approximated by a Gamma distribution as the input to our model. The density function of the Gamma distribution is

$$f_m(t_m) = \frac{\gamma\mu^\gamma t^{\gamma-1} e^{-\gamma\mu t_m}}{\Gamma(\gamma)}$$

where

$$\Gamma(\gamma) = \int_{\tau=0}^{\infty} e^{-\tau}\tau^{\gamma-1}\, d\tau$$

and

$$\gamma > 0 \tag{5}$$

where $\gamma\mu$ is called the *scale* parameter, and γ is called the *shape* parameter. The Laplace transform for the Gamma residence-time distribution is

$$f^*_{\mathbf{m}}(s) = \left(\frac{\gamma\mu}{s+\gamma\mu}\right)^{\gamma}. \tag{6}$$

Thus, p_0 and p_1 are rewritten as

$$p_0 = \left(\frac{\gamma}{\rho+\gamma}\right)^{\gamma} \tag{7}$$

$$p_1 = p_0\left(\frac{\gamma\rho}{\rho+\gamma}\right). \tag{8}$$

The *transition probability matrix* $\mathbf{P} = [p_{ij}]$ of the state diagram in Fig. 5 is given in (8a), shown at the bottom of the page, for $n \geq 1$. Let π_i be the stationary probability that AA 1 is in-state i. The transition probability matrix $\mathbf{P}$ has the following properties.

1) For every pair of states i, j there is a path $k_1, \cdots, k_r$ for which $p_{ik_1}p_{k_1k_2}\cdots p_{k_rj} > 0$.
2) There is at least one state i for which $p_{ii} > 0$.

$$\mathbf{P} = \begin{pmatrix} p_0+p_1 & p_2 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ p_0 & p_1 & p_2 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & p_0 & p_1 & p_2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & p_0 & p_1 & p_2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & p_0 & p_1 & p_2 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & p_0 & p_1+p_2 \end{pmatrix}_{2n\times 2n}. \tag{8a}$$

From [4, Theorem 4.1], we have

$$\pi_j = \sum_{k=0}^{2n} \pi_k p_{kj}, \qquad j = 0, 1, \cdots, 2n-1$$
$$1 = \sum_{k=0}^{2n} \pi_k. \tag{9}$$

Equation (9) can be solved to yield

$$\pi_i = \frac{p_0^{2n-(i+1)} p_2^i}{\sum_{j=0}^{2n-1} p_0^{2n-(j+1)} p_2^j}. \tag{10}$$

The probability that the WS scheme is exercised in AA 1 is

$$p_{ws} = \sum_{i=0}^{n-1} \pi_i = \frac{\sum_{i=0}^{n-1} p_0^{2n-(i+1)} p_2^i}{\sum_{j=0}^{2n-1} p_0^{2n-(j+1)} p_2^j} \tag{11}$$

and the probability that the S scheme is exercised is

$$p_s = 1 - p_{ws} = \frac{\sum_{i=n}^{2n-1} p_0^{2n-(i+1)} p_2^i}{\sum_{j=0}^{2n-1} p_0^{2n-(j+1)} p_2^j}. \tag{12}$$

The expected number of database transactions for authentication during a cycle is

$$C_{\mathrm{AA1}} = p_{ws} C_{ws} + p_s C_s. \tag{13}$$

For $n = 1$ (two-state AA 1), we have [see (10)–(12)]

$$p_{ws} = \pi_0 = \frac{p_0}{p_0 + p_2}$$

and

$$p_s = \pi_1 = \frac{p_2}{p_0 + p_2}. \tag{14}$$

For $n = 2$ (four-state AA 1), we have

$$\pi_0 = \frac{p_0^3}{p_0^3 + p_0^2 p_2 + p_0 p_2^2 + p_2^3}$$
$$\pi_1 = \frac{p_0^2 p_2}{p_0^3 + p_0^2 p_2 + p_0 p_2^2 + p_2^3}$$
$$\pi_2 = \frac{p_0 p_2^2}{p_0^3 + p_0^2 p_2 + p_0 p_2^2 + p_2^3}$$
$$\pi_3 = \frac{p_2^3}{p_0^3 + p_0^2 p_2 + p_0 p_2^2 + p_2^3}$$
$$p_{ws} = \frac{p_0^3 + p_0^2 p_2}{p_0^3 + p_0^2 p_2 + p_0 p_2^2 + p_2^3}$$

and

$$p_s = \frac{p_0 p_2^2 + p_2^3}{p_0^3 + p_0^2 p_2 + p_0 p_2^2 + p_2^3}. \tag{15}$$

Appendix B
Derivation of $E[N]$

Consider two-state state diagram (i.e., $n = 1$ in Fig. 5). To compute the expected number $E[N]$ of cycles, the automation is replaced by a random-walk model, where the state diagram is modified by marking state 1 as the absorbing state [see Fig. 10(a)]. To compute $E[N]$ is equivalent to computing the expected *hitting time* of the absorbing state 1 from the transient state 0. The probability that AA 1 moves from state 0 to state 1 with i cycles is

$$\Pr[X_i = 1] = (1 - p_2)^{i-1} p_2$$

and

$$E[N] = \sum_{i=1}^{\infty} i \Pr[X_i = 1] = \frac{1}{p_2}.$$

Consider the four-state state diagram. To compute $E[N]$, the state diagram is modified by marking states 0 and two as the absorbing states [see Fig. 10(b)]. Starting at state 0, let $X_i = j$ (for $j = 0, 2$) be the random variable that state j is visited at the ith step, and within the $i - 1$ steps, neither states 0 nor 2 are visited. Then, from Fig. 10(b)

$$\Pr[X_i = 0] = \begin{cases} 1 - p_2, & i = 1 \\ p_2 p_0 p_1^{i-2}, & i \geq 2 \end{cases}.$$

The probability that the algorithm visits state 0 (again) before visiting state 2 is

$$\Pr[X_I = 0] = \sum_{i=1}^{\infty} \Pr[X_i = 0] = (1 - p_2) + \sum_{i=2}^{\infty} p_2 p_0 p_1^{i-2}$$
$$= 1 - \frac{p_2^2}{1 - p_1}.$$

Suppose that $X_I = 0$. Then, the expected number of I before revisiting state 0 is

$$E[I | X_I = 0] = \frac{\sum_{i=1}^{\infty} i \Pr[X_i = 0]}{\Pr[I]} = \frac{1 - p_2 + p_2 p_0 \sum_{i=2}^{\infty} i p_1^{i-2}}{\Pr[I]}$$
$$= \left[1 - p_2 + \frac{p_2 p_0 (2 - p_1)}{(1 - p_1)^2}\right] \frac{1 - p_1}{1 - p_1 - p_2^2}$$
$$= \frac{(1 - p_2)(1 - p_1)^2 + p_2 p_0 (2 - p_1)}{(1 - p_1)(1 - p_1 - p_2^2)}.$$

From Fig. 10(b)

$$\Pr[X_i = 2] = p_2^2 p_1^{i-2}, \qquad \text{for} \quad i \geq 2.$$

The probability that the algorithm moves from state 0 to state 2 without revisiting state 0 is

$$\Pr[X_I = 2] = \sum_{i=2}^{\infty} \Pr[X_i = 2] = p_2^2 \sum_{i=2}^{\infty} p_1^{i-2} = \frac{p_2^2}{1 - p_1}.$$

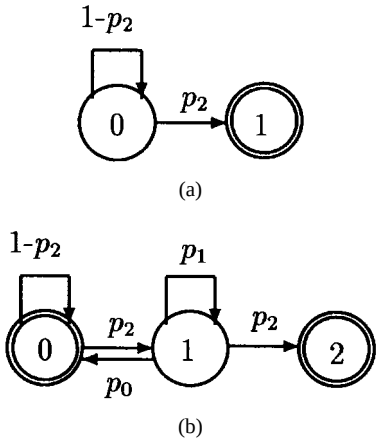

Fig. 10. Computing $E[N]$ with the absorbing states.

Suppose that $X_I = 2$. Then, the expected number of I before entering state 2 is

$$\Pr[I|X_I=2] = \frac{\sum_{i=2}^{\infty} i \Pr[X_i=2]}{\Pr[X_I=2]} = \frac{p_2^2 \sum_{i=2}^{\infty} i p_1^{i-2}}{\Pr[I_X=2]} = \frac{p_2^2(2-p_1)}{(1-p_1)^2 \Pr[X_I=2]} = \frac{2-p_1}{1-p_1}.$$

Suppose that AA 1 revisited state 0 i times before it moves to state 2. Then, the expected number of cycles before visiting state 2 is $iE[I|X_I=0] + E[I|X_I=2]$. Thus

$$\begin{aligned} E[N] &= \sum_{i=0}^{\infty} (iE[I|X_I=0] + E[I|X_I=2]) \cdot \Pr[X_I=0]^i \Pr[X_I=2] \\ &= E[I|X_I=2] + \frac{E[I|X_I=1]\Pr[X_I=0]}{\Pr[X_I=2]} \\ &= \frac{2-p_1}{1-p_1} + \frac{p_1[(1-p_2)+p_0p_2(2-p_1)]}{(1-p_1)^2p_2^2}. \end{aligned}$$

REFERENCES

[1] "Network and operations plan for access services to personal communications services (PCS) providers," Bellcore, Tech. Rep. SR-TSV-002402, 1993.

[2] "Cellular radio-telecommunications intersystem operations: Automatic roaming," EIA/TIA, Tech. Rep. IS-41. 3-B, 1991.

[3] "Cellular radio-telecommunications intersystem operations: Authentication, signaling message encryption and voice privacy," EIA/TIA, Tech. Rep. TSB-51, 1993.

[4] H. M. Taylor and S. Karlin, *An Introduction to Stochastic Modeling.* New York: Academic, 1984.

[5] R. Jain, Y.-B. Lin, C. N. Lo, and S. Mohan, "A caching strategy to reduce network impacts of PCS," *IEEE J. Select. Areas Commun.*, vol. 12, no. 8, pp. 1434–1445, 1994.

[6] N. L. Johnson, *Continuous Univariate Distributions–1.* New York: Wiley, 1970.

[7] Y.-B. Lin, "Determining the user locations for personal communications networks," *IEEE Trans. Veh. Technol.*, vol. 43, no. 3, pp. 466–473, 1994.

[8] Y.-B. Lin and S. K. DeVries, "PCS network signaling using SS7," *IEEE Personal Commun. Mag.*, vol. 2, no. 1, pp. 44–55, 1995.

[9] S. M. Ross, *Stochastic Processes.* New York: Wiley, 1983.

Yi-Bing Lin (SM'95) received the B.S.E.E. degree from National Cheng Kung University, Taiwan, in 1983 and the Ph.D. degree in computer science from the University of Washington, Seattle, in 1990.

From 1990 to 1995, he was with the Applied Research Area at Bell Communications Research (Bellcore), Morristown, NJ. Since 1995, he has been a Full Professor of the Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan, R.O.C. His current research interests include design and analysis of personal communications services networks, mobile computing, distributed simulation, and performance modeling.

Dr. Lin is an Associate Editor of the *ACM Transactions on Modeling and Computer Simulation*, Subject-Area Editor of the *Journal of Parallel and Distributed Computing*, Associate Editor of the*International Journal in Computer Simulation*, Associate Editor of *IEEENetworks*, Associate Editor of*SIMULATION* magazine, Columnist of *ACM Simulation Digest*, Member of the Editorial Board of *International Journal of Communications*, Member of the Editorial Board of*Computer Simulation Modeling and Analysis*, Program Chair for the 8th Workshop on Distributed and Parallel Simulation, General Chair for the 9th Workshop on Distributed and Parallel Simulation, Program Chair for the 2nd International Mobile Computing Conference, Publicity Chair of ACM Sigmobile, Guest Editor for the ACM/Baltzer Wireless Networks special issue on personal communications, and Guest Editor for the IEEE TRANSACTIONS ON COMPUTERS special issue on mobile computing.

Seshadri Mohan (S'76–M'79) received the Ph.D. degree in electrical and computer engineering from McMaster University in 1980.

He is a Research Scientist in the Information and Networking Research Laboratory in Bellcore, Morristown, NJ. He joined Bellcore in 1990 and has worked on the Information Networking Architecture and Personal Communications Services and Protocols. His interests are in the areas of protocols for real-time services for mixed IP/ATM networks, applications of distributed computing principles to telecommunications network architectures, impact of PCS on signaling networks, security, and network resources. Prior to joining Bellcore, he taught at Clarkson and Wayne State Universities. He is the coauthor of the textbook *Source and Channel Coding: An Algorithmic Approach* and has contributed to several books, including *Mobile Communications Handbook*.

Nelson Sollenberger (S'78–M'81–SM'90–F'96) heads the Wireless Systems Research Department at AT&T. His department performs research on next-generation wireless systems concepts and technologies including high-speed transmission methods, smart antennas and adaptive signaling processing, system architectures, and radio-link techniques to support wireless multimedia and advanced voice services.

Howard Sherry (S'63–M'71) received the B.S. degree in electrical engineering from Worcester Polytechnic Institute and the M.S. and Ph.D. degrees in electrical engineering from the University of Pennsylvania, Philadelphia.

He is a Chief Scientist and Director in the Wireless Research Department at Bellcore, where he is responsible for architecture, mobility management, economic assessment, and protocol definition efforts associated with wide-area wireless-access networks. Prior to assuming his current position, he managed efforts involved with planning and laboratory implementation of high-speed-data networks, high-definition television services, and broadband multimedia services. Previously, he had responsibility for managing advanced network architecture planning and signaling system #7 standardization efforts. He has been active in PCS standardization efforts in the United States since their inception.

Dr. Sherry served on the executive committee that organized and conducted the Joint Experts Meeting on PCS air interface standards and chaired several of the ad hoc committees in the Joint Technical Committee on wireless-access standards.