PAGING COMMUNICATION FOR LOCATING MOBILE USERS

The increased demand for wireless personal communication systems (PCS) coupled with limited spectrum has motivated many researchers to look into the techniques that minimize the radio traffic needed to keep track of the mobile stations (MSs) and deliver information (voice, data, video, etc.) to them. The process of keeping track of the MSs by updating the location information and paging for them when they receive a call is known as *location management*. In order to keep track of the MS, the whole geographical area is divided into location areas (LAs) which are nothing but a logical group of cells. Whenever an MS changes its LA, location information about the MS, stored in certain registers in the fixed network, will be updated. This process is described in detail in the following section. *Location updating* will generate wireless traffic in the form of access traffic and will sometimes generate wireline traffic in the form of signaling traffic of Signaling System No. 7 (SS7) network (1). The act of seeking an MS when an incoming call has been placed to it is called *paging*. This reference to paging is specific to a function within cellular telephony, as opposed to paging systems (2,3). When a call arrives at the MS, its current location area is determined by interrogating the registers in the fixed network. Exact locating of mobile stations will reduce the time for call setup and increase the paging efficiency. Hence, the paging process generates wireless traffic in the form of paging traffic and wireline signaling traffic on the SS7 network when the registers are interrogated.

Keeping track of the current LAs associated with each MS, as well as locating the user when a call is received, requires frequent interactions between the mobile stations and the system. Reducing such interactions is one of the goals of efficient location management (4–11). Several strategies have been proposed in the literature that deal with the reduction of signaling traffic and database loads imposed by the need to locate and track users (12–15). Some researchers have looked into methods of reducing the access traffic (16–18). For example, in Ref. 16, instead of the LA-based method of updating the current user location, the authors investigate time-based, distance-based, and movement-based methods of updating. However, in this article, we define the overall cost of location management in terms of the three components, namely, paging cost, access cost, and signaling cost, with emphasis on the reduction of the wireless paging cost.

SYSTEM LAYOUT

The main idea behind location management is to locate the MS with minimum overhead, whenever there is an MS-terminated call. Note that the entire process of locating users is carried out between calls when the mobile stations are in stand-by mode. In the European Global System for Mobile Communications (GSMs), SS7 interconnects Mobile Switching Centers (MSCs), Home Location Registers (HLRs) and Vehicle Location Registers (VLRs). An MSC is connected to several base stations, other MSCs and the public switched telephone network (PSTN). The permanent records of the mobile users are stored in the HLR. It also contains the pointers to the MS's current VLR. The VLR is a database associated with a particular MSC. When an MS crosses into an LA under the MSC, the VLR associated with that MSC is updated. On the other hand, when the MS crosses into an LA under a different MSC, the old and new VLRs as well as the HLR is updated (1, 14). These processes generate traffic on reverse control (access) channels. Signaling System No. 7 (SS7) traffic is generated whenever there is an HLR update.

When an MS receives a call, it needs to be paged for in a certain area. Its current position (LA) is found out by interrogating the HLR. This generates SS7 traffic. The size of the area that is paged depends on the paging technique used by that particular system. In any case, there is traffic generated

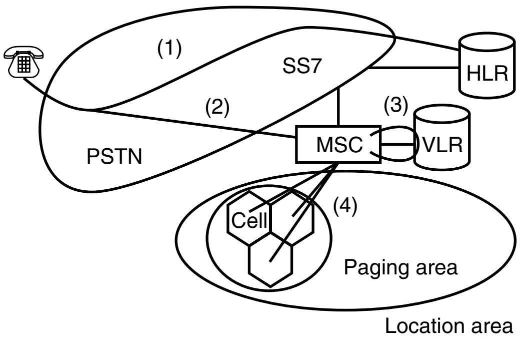

Figure 1. Cellular communications system configuration with visitor location register (VLR), home location register (HLR), paging area, and location area. The call setup procedure is shown by numbers (1) to (4). (1) Interrogation with HLR, (2) wireline call setup, (3) interrogation with VLR, (4) paging and radio call setup. MSC: mobile switching center, SS7: signaling system no. 7.

on the forward control (paging) channels. The various steps involved in a call setup are depicted in Fig. 1.

The above-mentioned traffic components (i.e., access traffic, paging traffic, and SS7 traffic) are the primary factors that govern the cost of location management. Since this is just an overhead cost for the service providers, the lower the cost incurred for locating the MS, the better it is. Hence, there is a need for better techniques that would lower the overall cost.

Location area (LA) and paging area (PA) are differentiated in the following discussion (4,5). A location area is a *logical* group of multiple paging areas, and a paging area is a *logical* group of cells. To find out the relationship between the size of location area, paging traffic, radio access traffic, signaling traffic, and their dependence on call arrival rate and the boundary crossing rate, as well as to determine the optimal paging techniques, the following cases of location management are considered in this article.

Case 1: Simultaneous Paging. One location area is composed of n cells. All the cells inside an LA will be paged *simultaneously* when there is a call to any one cell in the LA. The VLR(s) and, in cases described above, the HLR are updated every time the MS changes LAs and performs location update using the access channel. It is noted that if the size of LA is changed, the paging and access traffics generated operate in a counteractive fashion (4,5). In other words, small LAs generate less paging traffic and more access traffic compared with large LAs. Therefore, it is not possible to decrease the cost generated by the simultaneous paging technique by varying the size of the LA.

Case 2: Sequential Paging. One location area is composed of n cells. Paging will be done in k-cell ($k = n$) PAs *sequentially* until the called MS is found. The PAs are selected in a random fashion for this preliminary analysis. The probability of finding an MS is assumed to be the same in all the PAs of the LA. The VLR(s) and, in cases described above, the HLR are updated every time the MS changes LAs and performs location update using the access channel.

Case 3: Intelligent Paging. One location area is composed of n cells. Paging will be done in k-cell ($k \leq n$) PAs *sequentially* until the called MS is found. The PAs are selected using some kind of user information (the most recent interaction area, speed, movement pattern etc.), so that the probability of finding the MS in a certain PA or PAs is more than that for the rest of the LA. The VLR(s) and, in the cases described above, the HLR are updated every time the MS changes LAs and performs location update using the access channel.

PAGING TRAFFIC AND NORMALIZED PAGING DELAY

When simultaneous paging is used, for an n-cell LA, the paging traffic generated for each cell is given by $t_{\mathrm{p}} = n\lambda$, where λ is the call-arrival rate (incoming calls/s) for each cell. Here, we assume that the call-arrival rate is uniform for all the cells in the LA.

In the case of sequential paging, LAs are composed of multiple PAs. The PAs are paged one by one until the called MS is located. The paging traffic will depend on the average number of paged cells until the MS is located. When the probability of locating an MS is assumed to be the same in each PA, k/n, the probability of finding a called MS after paging i number of PAs, remains the same; that is, $p = k/n$. Therefore, the average number of paged PAs (ξ_k), which is the same as normalized paging delay (normalized with respect to simultaneous paging), will be as follows (4):

$$\xi_k = 0.5\left(\frac{n}{k} + 1\right) \tag{1}$$

Here we assume that n is an integer multiple of k. The equation for the general case is also derived in Ref. 4. The paging traffic for each cell with k-cells PA and n-cells LA with the same sequential paging technique will be

$$\begin{aligned} t_{\mathrm{p}} &= k\xi_k\lambda \\ &= 0.5(n+k)\lambda \qquad \text{(pages/s)} \end{aligned} \tag{2}$$

As a numerical example, when 10 cells make one LA, the paging traffic compared to one-cell LA increases by a factor of 10 with simultaneous paging, and by a factor of 5.5 with sequential paging with $k = 1$. In this case, the paging delay is 5.5 times longer with sequential paging compared to simultaneous paging. However, when $k = 5$, the paging traffic for sequential paging is 7.5 times that of a one-cell LA, but its paging delay is reduced to just 1.5 times that of simultaneous paging. Hence, when multiple-cell PAs are used in the case of sequential paging, there is a considerable decrease in the paging delay. The price paid is a slight increase in the paging traffic. Results on paging delay decrease and paging traffic increase normalized to the case of n being 1 with four different values of k (1, 3, 5, and 10) are shown in Figs. 2 and 3, respectively.

In Eq. (2), it seems that in order to minimize the paging cost, k can take any value from 1 to n; in other words, the paging delay is completely ignored. This is not true for real systems. There is a finite value of the acceptable average delay ($\tau_{\max}$). $\tau_{\max}$ is normalized with respect to the delay of simultaneous paging in the remainder of this article. This $\tau_{\max}$ can be used to determine the maximum ratio of the size of LA and the size of PA [see Eq. (1)]. Since $\xi_k \leq \tau_{\max}$, using Eq. (1), the size of PA (k) can be expressed in terms of the acceptable average paging delay ($\tau_{\max}$) and the size of LA (n) as follows:

$$k \geq \frac{n}{2\tau_{\max} - 1} \tag{3}$$

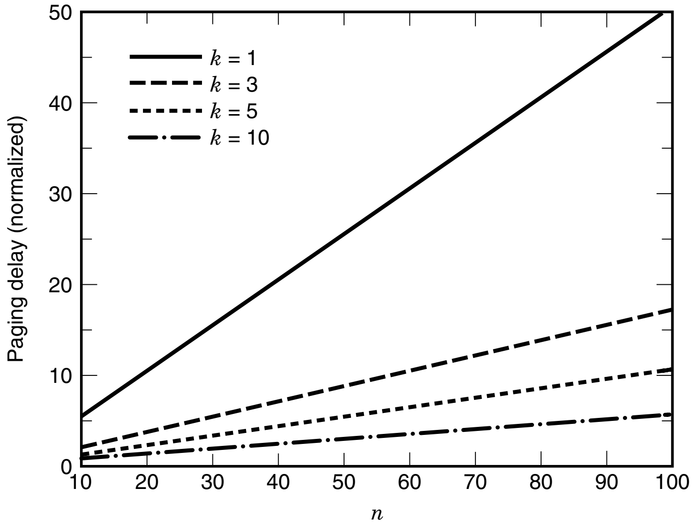

Figure 2. Normalized paging delay (normalized with respect to the delay for simultaneous paging) of the sequential paging scheme for different paging area sizes.

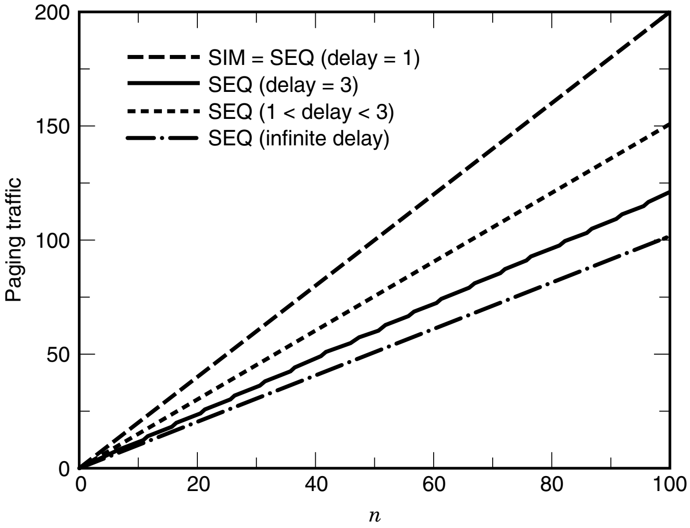

Figure 4. Normalized paging traffic [normalized with respect to the traffic for sequential paging (ignoring delay) for $n = 1$] of all schemes.

Hence,

$$k_{\min} = \left\lceil \frac{n}{2\tau_{\max} - 1} \right\rceil \tag{4}$$

where ⌈⌉ is the ceiling function, or the "rounding-up" function. Substituting this $k_{\min}$ for k in Eq. (2), we obtain the paging traffic generated for each cell while taking the paging delay into account. This paging traffic is given by

$$t_p = 0.5\left(n + \left\lceil \frac{n}{2\tau_{\max} - 1} \right\rceil\right)\lambda \qquad \text{(pages/s)} \tag{5}$$

Note that Eq. (4) gives the minimum k that would satisfy the delay requirement. So, k can be larger than $k_{\min}$ and still satisfy the requirement. Of course, as k increases, the paging traffic increases. Hence, it is desirable to operate with PA size as $k_{\min}$.

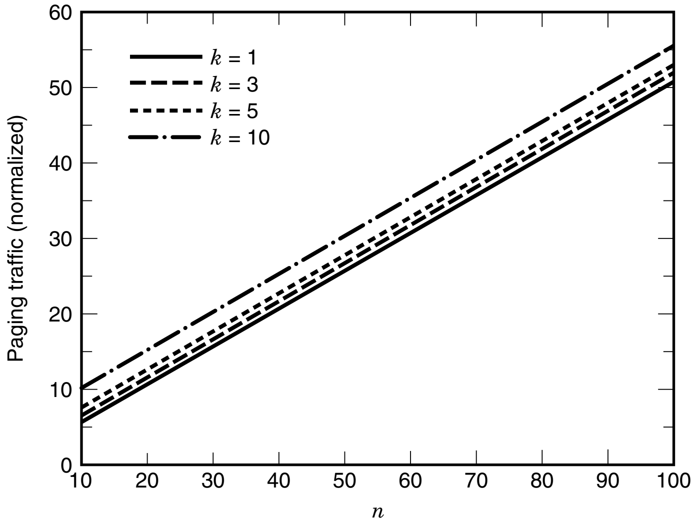

Figure 3. Normalized paging traffic (normalized with respect to the traffic for simultaneous paging with $n = 1$) of the sequential paging scheme for different paging area sizes.

In Fig. 4, the normalized paging traffic for simultaneous paging, sequential paging (without considering delay), and sequential paging (with paging delay consideration) are plotted versus the LA size. Observe that the traffic for sequential paging (without considering delay, i.e., infinite delay) has a normalized slope of 1 and the simultaneous paging traffic has a normalized slope of 2. The paging traffic due to sequential paging (with paging delay consideration) has a curve which lies between the other two. In other words, the enforcement of the delay requirement on sequential paging increases the gradient of the paging traffic curve. In fact, as the delay requirement becomes more and more stringent (i.e., as $\tau_{\max}$ decreases), the curve of sequential paging traffic tends toward that of simultaneous paging traffic, which is the case when $\tau_{\max} = 1$.

The behavior of paging traffic in Fig. 4 for different paging delay requirements confirms the physically intuitive result that to minimize the delay in sequential paging, the logical distinction between paging area and location area should be abolished (i.e., $k = n$), in which case we note that sequential paging converges to simultaneous paging. In practice, the permissible paging delay in a PCS environment will dictate the slope of the paging traffic curve. If, for instance, the normalized acceptable average paging delay is 3, then the paging traffic will be the solid line shown in Fig. 4. During the busy hours, if the network chooses to operate with less delay (suppose $1 <$ normalized paging delay < 3), then one will have a paging traffic curve which will lie between that of simultaneous paging and the curve for delay $= 3$, as shown in Fig. 4. Also note that when the required paging delay becomes more stringent (i.e., decreases), the value of paging traffic for a fixed n increases. This implies that the effective weight of paging traffic and, consequently, the cost of location management goes up.

In order to decrease this paging traffic for sequential paging, we need some more information (intelligence) about the users in the system. In this article, the forms of intelligence that are studied are the speed information and the most re-

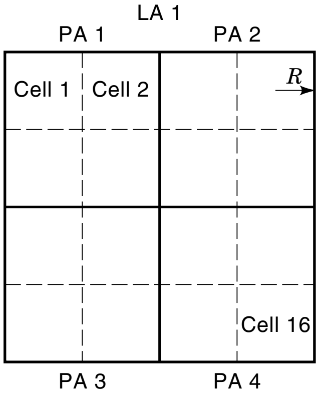

Figure 5. Cellular layout showing the cells and the paging areas within a location area.

cent interaction information about the MS. For simplicity, instead of finding the exact speed, the users are classified into two categories, namely, pedestrians and cars. The average speed of pedestrians is considered to be 5 km/h and the average speed of fast-moving MSs is assumed to be 30 km/h. The determination of the category is done by the MS, and this information is passed on to the system in the form of an extra bit when the LA update process is being carried out. This is done so that no additional radio traffic is generated. The information about the MS speed is stored in the VLR along with the other records. The speed is also updated when the MS uses the access channel to answer a paging message or to make a call.

Suppose an MS is in PA1 of LA1 when its most recent interaction with the fixed network takes place (see Fig. 5). This interaction could be in the form of a location update or a call origination or termination, and so on. We consider square-shaped cells in this work. If the MS receives a call *within* a certain time duration after update, then the probability (p) of finding the MS in PA1 is greater than that of the other PAs in the LA, that is, $p > k/n$. The expected number of paged PAs in terms of the probability of finding the MS in the first PA paged (p), which is referred to as *Probability of Successful First Paging Step* (PSFS) in Ref. 11, is derived in Refs. 6 and 7, and is given by

$$\xi_k = p + (1 - p)\left(1 + \frac{n}{2k}\right) \tag{6}$$

Hence, the paging traffic generated per MS, t_{pu}, is given by

$$\begin{aligned} t_{\mathrm{pu}} &= k\lambda_{\mathrm{u}}\xi_k \\ &= k\lambda_{\mathrm{u}}\left[p + (1 - p)\left(1 + \frac{n}{2k}\right)\right] \qquad \text{(pages/s)} \end{aligned} \tag{7}$$

where $\lambda_{\mathrm{u}} = \lambda/N$ is the average call-arrival rate per user per cell. Here, N is the average number of in-use and standby MSs in a cell. Note that k is to be determined considering the acceptable average delay ($\tau_{\max}$) for sequential paging.

Since p is dependent on the MS speed, it is different for pedestrians (p_{ped}) and cars (p_{car}) at a certain instance of time. Hence, the total paging traffic generated for each cell by intelligent paging (pages/second) is given by

$$\begin{aligned} t_{\mathrm{p}} = \Big[&F_{\mathrm{ped}}\left\{p_{\mathrm{ped}} + (1 - p_{\mathrm{ped}})\left(1 + \frac{n}{2k}\right)\right\} \\ &+ F_{\mathrm{car}}\left\{p_{\mathrm{car}} + (1 - p_{\mathrm{car}})\left(1 + \frac{n}{2k}\right)\right\}\Big] k\lambda \end{aligned} \tag{8}$$

where F_{ped} is the fraction of the total number of MSs in a cell that are pedestrians and F_{car} is the corresponding fraction of cars.

DETERMINATION OF *p*

It is obvious that the probability of successfully finding the MS in the first area paged (i.e., p) is inversely proportional to the time lapse between the last update and received call (T). Assume that $p = 1$ up to $T = t_{\mathrm{i}}$. Also, for $T > t_{\mathrm{m}}$ [where t_{m} is a certain time duration that depends on the MS speed, the cell radius, the number of cells in a PA (k), etc.], $p = k/n$—that is, the case of nonintelligent sequential paging. Hence, this simple version of intelligent paging is different from the sequential paging technique only when $T < t_{\mathrm{m}}$.

If we assume a linear relationship between p and T for $t_{\mathrm{i}} < T < t_{\mathrm{m}}$, then

$$p = \begin{cases} 1 & \text{for } 0 \le T < t_{\mathrm{i}} \\ \dfrac{\left(t_{\mathrm{m}} - \dfrac{k}{n}t_{\mathrm{i}}\right) - T\left(1 - \dfrac{k}{n}\right)}{t_{\mathrm{m}} - t_{\mathrm{i}}} & \text{for } t_{\mathrm{i}} \le T < t_{\mathrm{m}} \\ \dfrac{k}{n} & \text{for } T \ge t_{\mathrm{m}} \end{cases} \tag{9}$$

Intuitively, t_{i} and t_{m} is inversely proportional to the MS speed (v). However, they are directly proportional to the cell radius (R) and the number of cells in a PA (k). Of course, the exact relationship between them depends on the trajectory of the MS. For example, for a straight-line motion of the MS across a square-shaped PA, $t_{\mathrm{i}} = 2\sqrt{k}R/v$. On the other hand, for diagonal motion across the PA, $t_{\mathrm{i}} = 2\sqrt{2k}R/v$. It can be shown that this is a good approximation for non-square-shaped PAs also, provided that the PAs are chosen in a manner that is a good approximation to the regular square shape. If we have a square-shaped LA also, then $t_{\mathrm{m}} = (\sqrt{(n/k)} - 1)t_{\mathrm{i}}$ for $\sqrt{n/k} > 2$ in each case. For, $1 \le \sqrt{n/k} \le 2$, $t_{\mathrm{m}} = t_{\mathrm{i}}$. Again, this can be used as an approximation for non-square-shaped cases also. We consider straight-line motion for square-shaped cells in this article. So, other things being equal, since $v_{\mathrm{ped}} < v_{\mathrm{car}}$, the p versus T curve for pedestrians is different from that for fast-moving users, as shown in Fig. 6.

Note that the above discussion is for the scenario when we assume that the MS is always on the move. If it stays in the same location after speed update until it receives its next call, it will *always* be found in the first paging attempt because we start paging from the *most recent interaction area.*

LOCATION UPDATE AND SIGNALING TRAFFIC

The location management schemes considered in this article vary only in terms of the paging technique used. Since access and SS7 traffics are generated exactly in the same manner for all three cases, they are discussed in common in this section. Whenever an MS crosses an LA, access traffic is generated. This traffic is dependent on the boundary crossing rate (BCR) of the MS, which is defined as the number of MSs crossing into a cell-boundary per second. This BCR is denoted by μ, and the average access traffic per cell is $t_{\mathrm{a}} = \mu/\sqrt{n}$ accesses/second. This means that the average access traffic per cell, for location updating, will decrease as n increases. Here, we

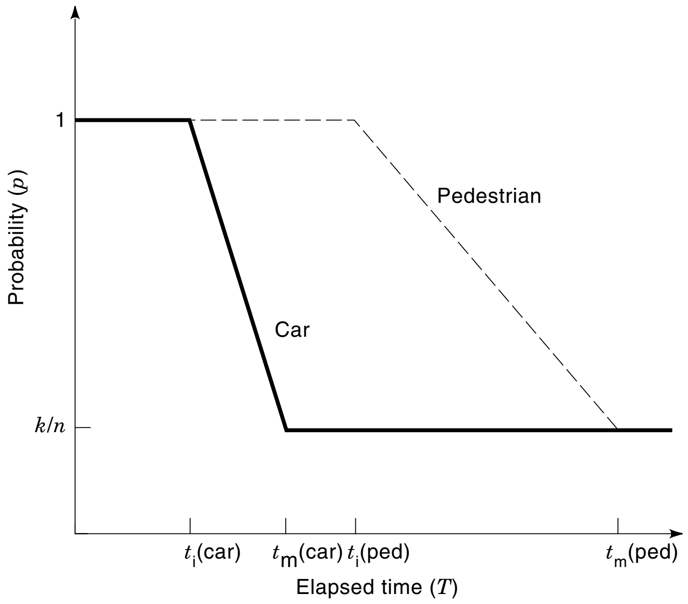

Figure 6. Plot of the probability of first successful paging step versus the elapsed time from last update for pedestrians and cars.

average out the access traffic over all the cells of an LA to see the overall picture. However, it is clear that the border cells of an LA will encounter updating traffic whereas the center cells have no access traffic due to updating. This implies that more control channels should be assigned to the border cells of an LA compared to the center cells. Note that μ is dependent on the number of MSs in a cell (N), the cell radius (R), and the speed of the MS (v) and is given by $2Nv/\pi R$ crossings/second (4,18).

SS7 traffic is generated whenever the MS crosses LAs under different MSCs, since this requires the HLR to be updated. Also, when there is an MS-terminated call, we need to interrogate the HLR for the current position of the MS, thus generating SS7 traffic. If n_{t} is the total number of cells controlled by an MSC, then the SS7 traffic generated per cell is $t_{\mathrm{s}} = \mu/\sqrt{n_{\mathrm{t}}} + \lambda$ signaling-messages/second.

COST FUNCTION

As mentioned before, the key components of location management cost are the costs generated by paging traffic (t_{p}), access traffic (t_{a}), and SS7 signaling traffic (t_{s}). These costs depend on the BCR (μ), call arrival rate (λ), the number of cells in a location area (n), the number of cells in a paging area (k), and the number of cells controlled by an MSC (n_{t}).

Comparison of the three schemes is done via the cost function defined as follows:

$$\begin{aligned}\mathrm{Cost} &= w_{\mathrm{p}}t_{\mathrm{p}} + w_{\mathrm{a}}t_{\mathrm{a}} + w_{\mathrm{s}}t_{\mathrm{s}} \qquad \text{(bytes/s)}\\ &= f(\lambda,\ \mu,\ n,\ k,\ n_{\mathrm{t}})\end{aligned} \tag{10}$$

where w_{p} (*bytes/page*), w_{a} (bytes/access-message), and w_{s} (bytes/signaling-message) are weights assigned to each traffic component. These weights may be different depending on the access method used, the availability of channels, and the cost of constructing the signaling network. Depending on the access method, the number of bytes required per second (i.e., the cost of location management) will be different. Also, if the SS7 network is constructed using the not-so-expensive fiber-optic technology, then w_{s} may be significantly less than w_{p} and w_{a}.

Hence, the location management cost for the three schemes is as follows:

$$\mathrm{Cost}_{\mathrm{sim}} = w_{\mathrm{p}}n\lambda + w_{\mathrm{a}}\frac{\mu}{\sqrt{n}} + w_{\mathrm{s}}\left[\frac{\mu}{\sqrt{n_{\mathrm{t}}}} + \lambda\right] \tag{11}$$

$$\mathrm{Cost}_{\mathrm{seq}} = \begin{cases} w_{\mathrm{p}}\dfrac{\lambda}{2}(n+k) + w_{\mathrm{a}}\dfrac{\mu}{\sqrt{n}} + w_{\mathrm{s}}\left[\dfrac{\mu}{\sqrt{n_{\mathrm{t}}}} + \lambda\right] \\ \hfill \text{(delay ignored)} \\ w_{\mathrm{p}}\dfrac{\lambda}{2}\left(n + \left\lceil\dfrac{n}{2\tau_{\max}-1}\right\rceil\right) + w_{\mathrm{a}}\dfrac{\mu}{\sqrt{n}} + w_{\mathrm{s}}\left[\dfrac{\mu}{\sqrt{n_{\mathrm{t}}}} + \lambda\right] \\ \hfill \text{(considering delay)} \end{cases} \tag{12}$$

$$\begin{aligned}\mathrm{Cost}_{\mathrm{int}} = w_{\mathrm{p}}&\Big[F_{\mathrm{ped}}\Big\{p_{\mathrm{ped}} + (1-p_{\mathrm{ped}})\left(1+\frac{n}{2k}\right)\Big\} \\ &+ F_{\mathrm{car}}\Big\{p_{\mathrm{car}} + (1-p_{\mathrm{car}})\left(1+\frac{n}{2k}\right)\Big\}\Big]\,k\lambda \\ &+ w_{\mathrm{a}}\frac{\mu}{\sqrt{n}} + w_{\mathrm{s}}\left[\frac{\mu}{\sqrt{n_{\mathrm{t}}}} + \lambda\right]\end{aligned} \tag{13}$$

Note that for the intelligent scheme, the delay requirement is taken into account, and so $k = [n/(2\tau_{\max} - 1)]$.

RESULTS

Costs of the three cases are plotted versus the location area size n for two values of BCR μ. In the first scenario, we consider the case of an area where the pedestrian population is more than the fast-moving MSs. For example, this could be the situation in a downtown city area. The plots are computed for $n_{\mathrm{t}} = 400$, $\tau_{\max} = 3$, $N = 200$, $F_{\mathrm{ped}} = 0.9$, $F_{\mathrm{car}} = 0.1$, $R = 200$ m. For this case, $\lambda = 0.75$ calls/s and $\mu = 1.33$ crossings/s.

In Fig. 7, we consider the case when all weights are equal. We find that the intelligent scheme is always the most cost-

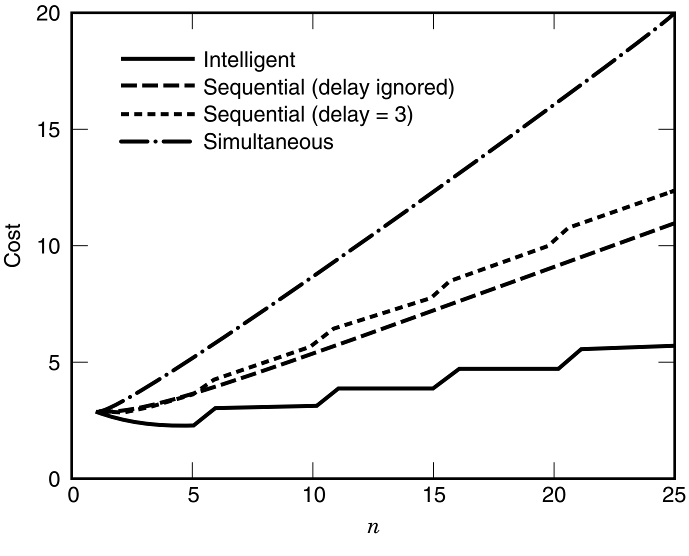

Figure 7. Overall cost of the four cases versus the size of location area (n) for equal weights assigned to paging, access, and signaling costs and dominant pedestrian population.

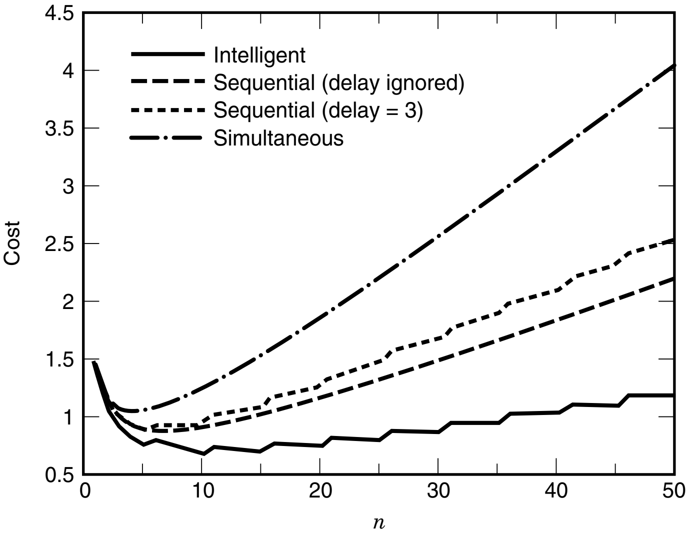

Figure 8. Overall cost of the four cases versus the size of location area (n) when access cost is assigned more weight than the paging and the signaling costs.

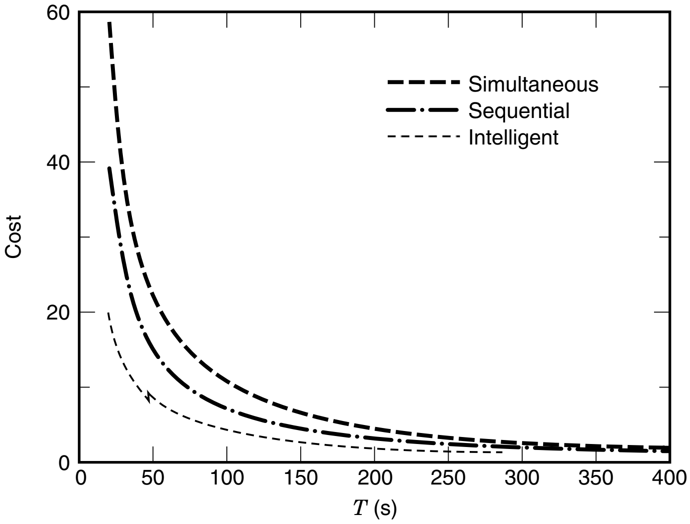

Figure 10. Overall cost of the three schemes versus the elapsed time between the last interaction of the mobile with the system and the call received. $\mu = 1.33$.

effective scheme for this situation. Also, there is an optimal value of n (n_{opt}) in each case, for which the overall cost is minimum. We find that $n_{\text{opt(sim)}} = 1$, $n_{\text{opt(seq)}} = 2$ and $n_{\text{opt(int)}} = 5$. Hence, larger location areas are possible in the case of the intelligent scheme. The PA size k is 1 for both sequential and intelligent schemes, since it is dictated by the acceptable average delay with respect to simultaneous paging, i.e., τ_{max}.

In Fig. 8, we consider the case when the weight of access traffic is 10 times more than each of the other components. It is worth noting that the general trends of all the schemes remain the same, although a much larger size of optimal LA is obtained when the access traffic is the dominating factor. For dominant paging traffic and dominant SS7 traffic scenarios, it is found that the trends are the same as the "all weights equal" case.

In Fig. 9, we consider the situation when the fast-moving population is dominant. This could be a situation in the suburban area. Here $N = 200$, $F_{\text{ped}} = 0.2$, $F_{\text{car}} = 0.8$ so that $\mu = 4.42$ crossings/s. All the other parameters are the same as the previous case. Again the intelligent scheme is a winner. However, the margin of difference between the costs of the sequential and intelligent paging techniques becomes smaller. This shows that the intelligent scheme is more beneficial in situations where the pedestrian population is dominant. This is because the probability of finding a slow-moving MS during the first paging attempt is higher than that of finding a car at a given instance of time.

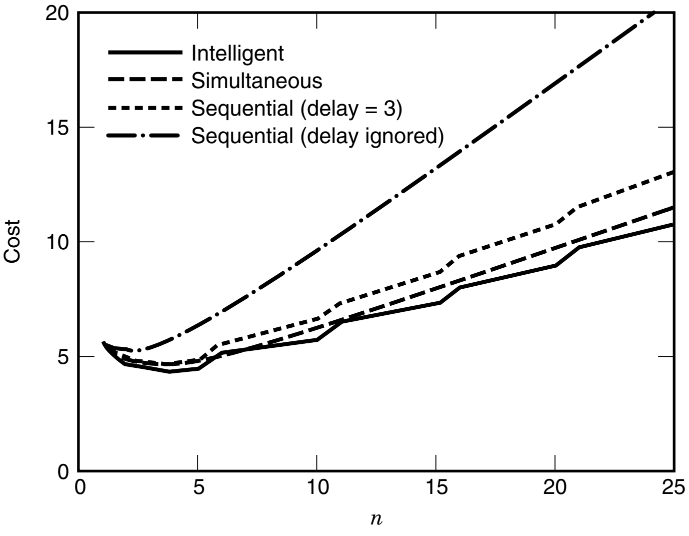

Figure 9. Overall cost of the four cases versus the size of location area (n) for equal weights assigned to paging, access, and signaling costs and dominant fast-moving population.

Figure 10 is plotted to show that the intelligent scheme fares better than the other schemes only up to a certain value of T for a certain scenario. Beyond this value of T, it is the same as the sequential paging scheme. For $\mu = 1.33$ crossings/s, $n = 5$, and $k = 1$, this particular value of T is 288 s. Note that this analysis is for the worst-case scenario when it is assumed that the MS is constantly on the move. For a dominant slow-moving population in office-buildings in a downtown area, intelligent paging fares better than sequential paging up to a much larger T. The costs of the simultaneous and the sequential paging schemes change with T because of changing λ. Also, since $k = 1$, sequential paging has just one curve.

CASE STUDY: LIMITING THE NUMBER OF PAGING AREAS TO TWO

The above results are obtained on the basis of the theoretical framework provided in the article. However, in practice, it may not be desirable to have more than two paging areas (8). In other words, if the MS is not located in one attempt, it should be located in the second try (9). If that is the case, then $n/k = 2$ so that $k = \lceil n/2 \rceil$ and Eq. 5 becomes

$$t_{\text{p}} = 0.5\left[n + \left\lceil \frac{n}{2} \right\rceil\right]\lambda \tag{14}$$

For intelligent paging, the expected number of paged PAs becomes $2 - p$ so that the paging traffic given in Eq. (8) becomes

$$t_{\text{p}} = [F_{\text{ped}}(2 - p_{\text{ped}}) + F_{\text{car}}(2 - p_{\text{car}})]k\lambda \tag{15}$$

In this case the expression for p becomes

$$p = \begin{cases} 1 & \text{for } 0 \le T < t_i \\ 0.5 & \text{for } T \ge t_i \end{cases} \tag{16}$$

The costs of sequential and intelligent paging schemes are now given by

$$\text{Cost}_{\text{seq}} = w_{\text{p}} \frac{\lambda}{2} \left(n + \left\lceil \frac{n}{2} \right\rceil \right) + w_{\text{a}} \frac{\mu}{\sqrt{n}} + w_{\text{s}} \left[\frac{\mu}{\sqrt{n_{\text{t}}}} + \lambda \right] \tag{17}$$

$$\begin{aligned} \text{Cost}_{\text{int}} = {} & w_p [F_{\text{ped}}(2 - p_{\text{ped}}) + F_{\text{car}}(2 - p_{\text{car}})] k\lambda \\ & + w_{\text{a}} \frac{\mu}{\sqrt{n}} + w_{\text{s}} \left[\frac{\mu}{\sqrt{n_{\text{t}}}} + \lambda \right] \end{aligned} \tag{18}$$

Figure 11 shows the cost curves under this constraint for a dominant pedestrian population. The parameter values are the same as those for Fig. 7. It is found that the optimum cost of intelligent paging is 82% of the optimum cost of sequential paging when the SS7 traffic is negligible. On the other hand, when the SS7 traffic is considered to have the same weight as the other two components, there is a 12% reduction in the cost of intelligent paging, compared to sequential paging.

DISCUSSION

In this article, location management in PCS is studied via three different paging schemes. It is also shown that if a simple form of intelligence (i.e., additional information about MS speed and recent interaction) is added to the sequential paging scheme, the overall cost *does* come down under certain circumstances. The performance of this intelligent scheme is highly dependent on the time lapse between the speed update and call-arrival. In any case, the performance of this intelligent scheme is never worse than the sequential paging scheme. This performance (overall cost of location management) can be further improved if more information about the MSs is incorporated (11).

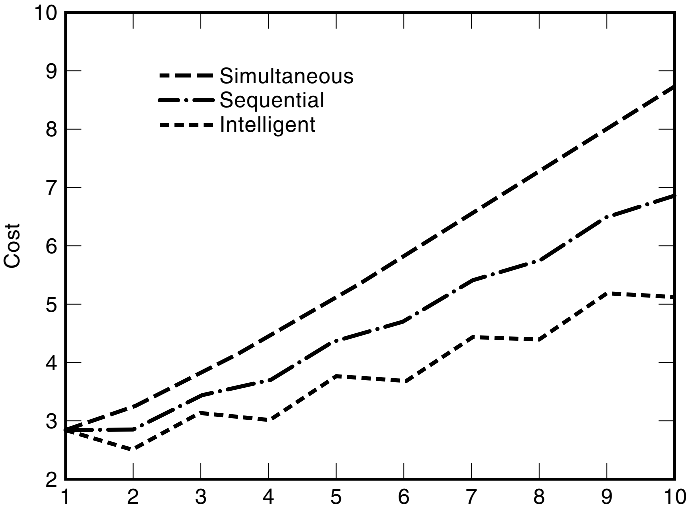

Figure 11. Overall cost of the three schemes versus the size of location area (n) for two paging areas.

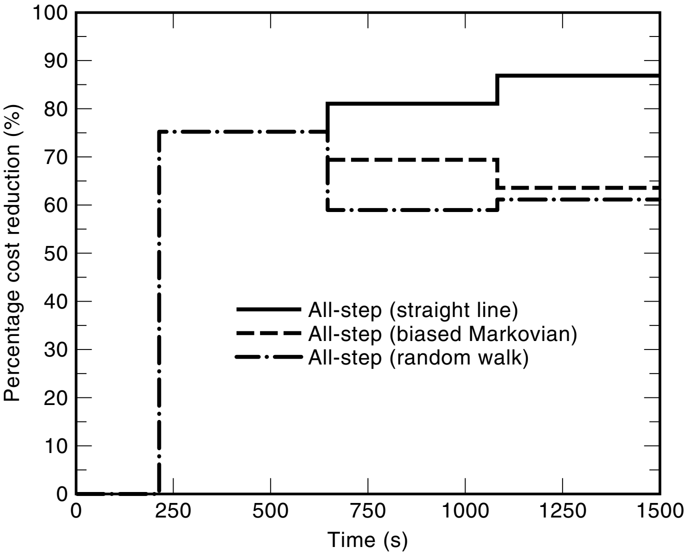

Figure 12. Percentage cost reduction with respect to One-Step Intelligent paging in All-Step Intelligent paging strategies (for straight-line, biased Markovian and Random Walk motion models).

To assess the total cost of location management, a simple cost function is defined and used whereby the weights assigned to each of these traffic components would also depend on other factors, such as (1) the specific realization of the PCS environment and (2) the architecture of the signaling network. This analytical framework can be used to analyze the performance of simultaneous, sequential, and "intelligent" paging schemes in a unified manner. While this framework leads to similar results with those obtained via the more rigorous approaches (19,20), it also provides valuable physical insight into the impact of main system parameters such as maximum paging delay, call-arrival rate, and boundary crossing rate on the overall cost of location management.

In this article, we have provided results for simplistic cases such as deterministic call-arrival rate and boundary crossing rate, straight-line motion for the mobile users, and so forth. However, our preliminary investigations show that the analytical framework provided in this chapter holds for time-varying call-arrival and boundary crossing rates and for random user-movement patterns (Markovian, Random Walk, etc.) as well (21). Also, the concept of "intelligence" has so far been investigated only for the first paging step. It has been verified that this can be extended to the second and subsequent steps under the same framework (21). Figure 12 compares the All-Step strategy with One-Step Intelligent paging, for straight-line, Markovian and Random Walk movement models. Our goal is to come up with a generalized theory for analyzing *any* location management scheme for the future PCS.

BIBLIOGRAPHY

1. CCITT Blue Book Recommendations Q.1000, "Public Land Mobile Network Interworking With ISDN and PSTN," *The Interna-*

tional Telegraph and Telephone Consultative Committee, Nov. 1988.

2. D. Baker, *A Comprehensive Guide to Paging,* BIA Publications Inc., 1992.
3. K. Siwiak, *Radiowave Propagation and Antennas for Personal Communications,* 2nd ed., Norwood, MA: ArtechHouse, 1998.
4. O. K. Tonguz, S. Mishra, and H. Jung, A simple analytical framework for location management in personal communication systems, *IEEE Trans. Veh. Technol.,* **42** (2): 428–439, 1998.
5. D. Plassmann, Location management strategies for mobile cellular networks of 3rd generation, *Proc. 44th IEEE Vehicular Technology Conference,* 1994, pp. 649–653.
6. S. Mishra and O. K. Tonguz, Most recent interaction area and speed-based intelligent paging in personal communication systems, *Proc. 47th IEEE Vehicular Technology Conference,* May 1997, pp. 505–509.
7. S. Mishra and O. K. Tonguz, Analysis of intelligent paging in personal communication systems, *Electronics Letters,* **34** (1): 12–13, Jan. 1998.
8. S. Madhavapeddy, K. Basu, and A. Roberts, Adaptive paging algorithms for cellular systems, *Proc. 45th IEEE Vehicular Technology Conference,* 1995, pp. 976–980.
9. S. Mishra and O. K. Tonguz, Location management with intelligent paging in personal communication systems, *IEEE Proc. Communications,* submitted for publication.
10. H. Xie, S. Tabbane, and D. J. Goodman, Dynamic location area management and performance analysis, *Proc. 43rd IEEE Vehicular Technology Conference,* 1993, pp. 536–539.
11. G. L. Lyberopoulos et al., Intelligent paging strategies for third generation mobile telecommunication systems, *IEEE Trans. Veh. Technol.* **44** (3): 543–554, August 1995.
12. R. Jain et al., A caching strategy to reduce network impacts of PCS, *IEEE J. Selected Areas in Communications,* **12** (8): 1434–1445, 1994.
13. Y.-B. Lin and A. Noerpel, Implicit deregistration in a PCS network, *IEEE Trans. Veh. Technol.,* **43** (4): 1006–1009, Nov. 1994.
14. K. S. Meier-Hellstern and E. Alonso, The use of SS7 and GSM to support high density personal communications, *Proc. ICC '92,* 1992, pp. 1698–1702.
15. G. P. Pollini, The intelligent network signalling and switching costs of an alternate location strategy using memory, *Proc. 43rd IEEE Vehicular Technology Conference,* 1993, pp. 931–934.
16. A. Bar-Noy, I. Kessler, and M. Sidi, Mobile users: to update or not to update?, *Proc. IEEE INFOCOM '94,* 1994, pp. 570–576.
17. S. Okasaka et al., A new location updating method for digital cellular systems, *Proc. 41st IEEE Vehicular Technology Conference,* 345–350, 1991.
18. I. Seskar et al., Rate of location area updates in cellular systems, *Proc. 42nd IEEE Vehicular Technology Conference,* 1992, pp. 694–697.
19. C. Rose and R. Yates, Minimizing the average cost of paging cost under delay constraints, *Wireless Networks, I,* 211–220, 1995.
20. J. S. M. Ho and I. Akyldiz, Mobile user location update and paging under delay constraints, *Wireless Networks, I,* 413–425, 1995.
21. O. K. Tonguz et al., All-step intelligent paging: a location strategy for mobile networks, submitted for presentation at *IEEE INFOCOM '99.*

SUMITA MISHRA
OZAN K. TONGUZ
State University of New York at Buffalo