Traffic Performance Analysis of Handover in GMPCS

Jaeweon Cho and Daehyoung Hong
Dept. of Electronic Engineering, Sogang University
C.P.O. Box 1142, Seoul, 100-611, Korea
E-mail: jcho@eecom1.sogang.ac.kr, dhong@ccs.sogang.ac.kr

Abstract—This paper presents an analysis of the handover process and its effect upon the traffic performance in GMPCS. With the non-stationary satellite used for the system, the handover scheme needs to be applied to make calls completed without any interruption. An analytical model has been developed for the analysis of the handover process. We have derived mean number of handovers and handover delay by using the developed model. A suitable traffic model of the whole system has also been derived after due considerations of the handover process. The system performance measures include new call blocking probability, call dropping probability and mean number of handovers per call. A computer simulation has been developed and used. We also analyzed the system performance with a number of handover priority schemes applied. Based on the study results, handover parameters have been selected to maximize the traffic performance. It is shown that we can make overall traffic performance of GMPCS system improved by setting handover parameters properly.

I. INTRODUCTION

One of the main objectives of the next generation mobile communication services is to provide communications between handheld personal terminals located at diverse locations around the world at any time. It does not seem possible to achieve this objective by means of only terrestrial mobile networks because of its limitation of service coverage. The global mobile personal communications by satellite (GMPCS) is the system that can provide the personal communication service anywhere at any time. A number of GMPCS systems are already under construction.

GMPCS systems use a constellation of low or medium earth orbit (LEO or MEO) satellites. The non-stationary satellites in LEO or MEO rotate the earth at very high speed. Whenever the satellite passes over the serving mobile terminal, a handover scheme needs to be applied to make calls completed without any interruption.

The quality of service (QoS) of the whole system is affected by the handover process. If unnecessary handover occurs frequently in fringe of the satellite footprint, the signaling load can become too much. On the other hand, if the decision for handover is delayed too long, calls can be dropped before handed over successfully. Such degradation of QoS can influence negatively on the system capacity. In order to provide high quality services to maximum number of subscribers, we need to apply the handover process very carefully and optimally.

Previous studies analyzed the traffic performance of GMPCS but without considering the handover process in real environment [1]–[3]. Paper [4] analyzed the handover process itself but not the system traffic performance.

In this paper, the traffic performance of GMPCS is analyzed with the handover process. In Section II, we analyze the handover process in GMPCS propagation environments. An analytical model is developed and used. A computer simulation is also developed to derive the handover effect upon traffic performance. The performance analysis results are described in Section III. In Section IV, we present the conclusion and topics for future work.

II. HANDOVER PERFORMANCE

A. Received Signal Level Model

The decision to initiate a handover can be made based on several quantities such as the received signal level from communicating and neighboring satellites, and the distance from satellites. Inter-satellite handover occurs at the coverage edge of a satellite. At the coverage edge the elevation angle is low and the received signal level can be influenced by shadow fading. Therefore, we can not performed the handover process reliably based only on distances. The received signal levels need to be measured and used for inter-satellite handover process.

The received signal level model to analyze inter-satellite handover performance is shown in Fig. 1. When each satellite has a number of beams, inter-satellite handover occurs between boundary beams. The signal level received from satellite 1 and 2 at time $t$, $r_1(t)$ and $r_2(t)$ (in dB) are given by

$$r_1(t) = K_1 - K_2 \log d_1(t) + G_1(t) + \zeta_1(t),$$

$$r_2(t) = K_1 - K_2 \log d_2(t) + G_2(t) + \zeta_2(t),$$

where $K_1$ (in dB) is the received signal level at a unit distance from satellite, and $K_2$ (in dB/dec) is the slope of the path loss (e.g. $K_2 = 20$ in free space). Satellites are moving in fixed directions and the speed is much higher than that of mobiles. Hence, the mobile’s position is represented by the distances from each satellite $d_1$ and $d_1$ (in km) [5]. The distances are functions of time and represented as

$$d_1(t) = \sqrt{R_c^2 + R_o^2 - 2R_cR_o \cos(\delta_1 + tw)},$$

$$d_2(t) = \sqrt{R_c^2 + R_o^2 - 2R_cR_o \cos(\delta_2 - tw)},$$

where $R_c$ is the radius of earth (6378 km), $R_o$ is the sum of the orbit altitude and $R_c$, and $w$ is the angular velocity of satellites (deg/sec).

We have assumed in this work the reference mask for the antenna radiation pattern as shown in Fig. 2 [6]. Region a corresponds to the part of main lobe that is out of coverage. In this region, the typical gain variation versus off-axis angle $\psi$ is expressed as

$$G(\psi) = G_{max} - 3(\psi/\psi_0)^a,$$
where $G_{\text{max}}$ (in dBi) is maximum antenna gain of the satellite and $\psi_0$ (in degree) is the half of the 3dB-beamwidth. The parameter $s$ indicates the sharpness of the main lobe. From equation (5), the antenna gains of each satellite to the direction of mobile, $G_1(t)$ and $G_2(t)$ (in dBi) are derived as

$$G_1(t) = G_{\text{max}} - 3 \left[ \left( T_1(t) - \tau \right) / \psi_0 \right]^s,$$

$$G_2(t) = G_{\text{max}} - 3 \left[ \left( T_2(t) - \tau \right) / \psi_0 \right]^s,$$

where

$$T_1(t) = \arcsin \left[ R_e \sin \left( \delta_1 + t \psi_{\text{ang}} \right) / d_1(t) \right],$$

$$T_2(t) = \arcsin \left[ R_e \sin \left( \delta_1 - t \psi_{\text{ang}} \right) / d_2(t) \right],$$

$$\tau = \arctan \left[ \sin \delta_1 / (1 + R_e / R_e - \cos \delta_1) \right].$$

In equations (1) and (2), $\zeta_1(t)$ and $\zeta_2(t)$ (in dBi) are the shadow fading of the signal level from each satellite. They are assumed to be Gaussian process and independent of each other. Mean $\mu_\zeta(t)$ and standard deviation $\sigma_\zeta(t)$ of the fading process depend on the satellite’s elevation angle $\theta(t)$ as shown below [7]. Each shadowing process is also assumed to be correlated in time like the case of terrestrial cellular systems.

$$\mu_\zeta(t) = (20 / \ln 10)(-2.331 + 0.1142 \theta(t))$$

$$- 1.939 \times 10^{-3} \theta^3(t) + 1.004 \times 10^{-5} \theta^5(t),$$

$$\sigma_\zeta(t) = 4.5 - 0.05 \theta(t),$$

for $20^\circ < \theta(t) < 80^\circ$.

**B. Mathematical Analysis**

The performance of handover process in terrestrial cellular system is usually measured by mean number of handovers and handover delay [8]. The same parameters can also be used for handovers in GMPCS. When a mobile served by a satellite enters the service area of a new satellite, it is desirable that handover should be performed only once at the boundary of two service areas. One of the ways to prevent unnecessary handovers, is to initiate the handover procedure based on the averaged signal levels. The measured signal levels at the mobile are averaged using window of duration $T_w$. Another way is to use a hysteresis margin $h$. Handover process is initiated if the averaged signal level of a new satellite exceeds that of current satellite by the hysteresis margin $h$. The difference between the averaged signal levels of two satellites will be expressed as $x(t)$.

We divided $t_{\text{du}}$ into $N$ equal intervals for evaluation. $t_{\text{du}}$ is the period from the time when beam coverage of a satellite is centered on the mobile to the time that of the next satellite moves over the mobile. In this case, the sampling interval $t_k$ is represented by $t_{\text{du}}/N$ and the $k$th sampling point $t_k$ by $k t_k$. Let $P_{\text{ho}}(k)$ be the probability of one handover occurring in the $k$th interval, $P_{\text{ho}21}(k)$ the probability of handover from satellite 1 to satellite 2, and $P_{\text{ho}21}(k)$ the probability of handover from satellite 2 to satellite 1. Probabilities of a mobile being connected to each satellite, $P_{\text{ho}1}(k)$ and $P_{\text{ho}21}(k)$, can be recursively computed as follows [9].

$$P_{\text{ho}}(k) = P_{\text{ho}1}(k-1)P_{\text{ho}21}(k) + P_{\text{ho}2}(k-1)P_{\text{ho}12}(k),$$

$$P_{\text{ho}1}(k) = P_{\text{ho}1}(k-1) (1 - P_{\text{ho}21}(k)) + P_{\text{ho}2}(k-1)P_{\text{ho}12}(k),$$

$$P_{\text{ho}2}(k) = P_{\text{ho}2}(k-1) (1 - P_{\text{ho}12}(k)) + P_{\text{ho}1}(k-1)P_{\text{ho}21}(k),$$

where $k = 0, 1, \ldots, N$, and $P_{\text{ho}1}(0) = 1$, $P_{\text{ho}2}(0) = 0$ as the initial values. Since $x(t_k)$ is a Gaussian process, $P_{\text{ho}12}(k)$ and $P_{\text{ho}21}(k)$ can be expressed as

$$P_{\text{ho}12}(k) \approx \Pr \left[ x(t_k) > h \right| x(t_{k-1}) \leq h]$$

$$\approx \int_{h}^{\infty} \left[ 1 - Q \left( \frac{h}{\sigma_x(t_k)} \sqrt{1 - \gamma^2} + A(k) \right) \right] P_x(t_k)(\alpha) \, d\alpha,$$

$$1 - Q \left( \frac{h - \mu_x(t_{k-1})}{\sigma_x(t_k)} \right),$$

where

$$Q(\xi) = \int_{\xi}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \, dy.$$
\[ P_{s2|s1}(k) \approx Pr \{ x(t_k) < -h | x(t_{k-1}) \geq -h \} \]
\[ = \int_{-\infty}^{-h} Q \left( \frac{-h - \mu_x(t_{k-1})}{\sigma_x(t_k)} \right) P_x(t_k)(\alpha) d\alpha + A(k), \]
\[ Q \left( \frac{-h - \mu_x(t_{k-1})}{\sigma_x(t_k)} \right) \]
where
\[ A(k) \triangleq \frac{-\mu_x(t_{k-1}) - \gamma(\alpha - \mu_x(t_{k-1}))}{\sigma_x(t_k) \sqrt{1 - \gamma^2}}. \]

The number of handovers occurring during \( t_{dw} \) is equal to the number of intervals in which handovers occur. Therefore,

\[ \text{Mean number of handovers} = \sum_{k=1}^{N} P_{ho}(k). \]

**Handover delay** is defined from the time of the halfway between the beams of two satellites to the time of handover initiation. The handover delay can be represented as

\[ \text{Handover delay} = k_{cp} t_s - t_{dw}/2, \]

where the crossover point, \( k_{cp} t_s \), is defined as the point where \( P_{s1}(k_{cp}) \) is equal to 0.5.

**C. Results**

In order to analyze the handover performance using the proposed analytical method, we have selected Iridium as the system model. The system parameters according to the characteristics of Iridium are chosen such that the orbit altitude is 780 km, the angular velocity of satellite about 0.06 deg/sec, and the radius of the outer beam about 643.6 km [10]. Therefore, if a mobile is located on the center of satellite 1’s beam at 0 sec, that mobile will be on the center of satellite 2’s beam at 83.8 sec. In this period, the elevation angles of two satellites are below 20 degrees, and then the standard deviation of shadowing is set to 3.5 dB according to equation (12). Decorrelation distance of shadowing is assumed 20 m and the mobile speed 60 km/h.

The results are shown in Fig. 3 and Fig. 4 with various handover parameter settings. Fig. 3 shows that mean number of handovers is decreased as the hysteresis margin or the period of averaging window is increased. On the other hand, the Fig. 4 shows that handover delay is increased as two handover parameters are increased. The results show that it is not easy to minimize both mean number of handovers and handover delay at the same time.

The tradeoff of handover delay against mean number of handovers is illustrated as shown in Fig. 5. In the ideal case, mean number of handovers will be one and handover delay 0 sec. The region close to the ideal point is marked with a circle in the figure. From the figure, it is desirable to decide the handover parameters around the circle marked.

We have also developed computer simulations to validate the analytical results. Simulation results are shown in the figures above. It is confirmed by the simulation results that the analytical method proposed and used in this section is reliable.
III. TRAFFIC PERFORMANCE WITH HANDOVER

A. Service Coverage Model

Satellites of GMPCS use the multibeam antenna in order to reduce the size of the mobile terminal, to increase the link capacity, and to improve the spectrum reuse efficiency. If the multibeam antenna is designed such that the beam coverages on the ground are identical, the footprint of each beam will form a honeycomb pattern similar to the cell coverage of the terrestrial cellular systems. The footprint can therefore be represented as shown in Fig. 6. Each satellite of Iridium is designed to have 48 beams [10]. In this figure, the thick solid line is the coverage boundary between adjacent satellites.

Because the multibeam foot print in Fig. 6 moves in the direction of satellite orbit, we can analyze the traffic performance of each polar orbit individually. Besides, the signal level received from the satellite at the left and right side of the orbits can be symmetrical. Thus, each side can also be analyzed respectively. We have analyzed the traffic performance of the left side area that is marked with gray color in the figure.

B. System Performance Parameters

The handover occurs frequently during a call duration because satellites pass over the mobile at very high speed. If no channel is immediately available in the target satellite beam, handover fails and the call can be forced to be terminated. The call dropping due to handover failure can have much worse impact to the user than the occurrence of a new call blocking. We must therefore apply the priority schemes to handover attempts. In this paper, we have performed the analysis with two priority schemes, one with channels reserved for handovers and the other with queue for the handover [11].

The signal level received from satellite can vary rapidly due to shadowing. A call drop timer can be used to prevent an ongoing call from dropping due to short duration of signal level degradation. If the signal level stays below a certain threshold until the call drop timer expires, the call is forced to be terminated. If the signal level recovers above the threshold before the call drop timer expires, the timer is reset.

We have chosen the measures for the system performance as follows.
- **New call blocking probability** — The probability that a new call is blocked because there is no channel available or the received signal level is below the drop threshold.
- **Call dropping probability** — The probability that a call is forced to be terminated prematurely because the mobile experiences unsuccessful handover or received signal level is below the drop threshold until the call drop timer expires.
- **Mean number of handovers per call** — The mean number of handovers that the mobile experiences prior to completion or dropping of a call.

C. Simulation Results

A computer simulation has been developed and used to analyze the traffic performance. Fig. 7 shows an orbital plane chosen for the performance analysis. As described previously, we have considered the half side of the orbital plane (110°E ~ 125°E and 40°W ~ 55°W) for the analysis. We have applied the following assumptions. Calls generated mostly on the continent. Also calls generated on the continents are distributed uniformly in this analysis. Traffic generated by each mobile is 0.02 erlang with the average call duration of 120 sec. To each beam 10 channels are assigned. Antenna gain, $G_{\text{max}}$ is 23 dBi [12]. The drop threshold is $K_1 = 63.8$ dBW [13].

Even though a uniform traffic distribution has been assumed in this analysis, any form of traffic distribution can be applied to the developed simulation. For the generation of shadowing process, the elevation angle of the satellites has been calculated and considered. We have divided the orbital plane into the unit area of 5 degrees in latitude and 5 degrees in longitude, and then analyzed traffic performance for every unit area.

System performance results are shown at various settings of handover parameter in Fig. 8. Call dropping probability is plotted as functions of **mean number of handovers per call** with settings of $h = 0, 3, \cdots, 12$ dB and $T_{\text{drop}} = 0, 2$ sec. The results are based on 18,000 subscribers in the analyzed area (110°E ~ 125°E and 40°W ~ 55°W). Every point illustrated in the figure represents the maximum value among results of unit areas in the analyzed region. The figure shows the tradeoff between **call dropping probability and mean number of handovers per call** similar to that in Fig. 5. The optimal results can be obtained with no shadowing applied. This result is also displayed as the ideal point in the fig.
ure. In this case, call dropping probability and mean number of handovers per call can be minimum because the handover occurs once at the boundary between two beams. In real environment with shadowing, the handover would not occur ideally. Thus, system performance varies with the handover parameter settings as shown in the figure. Handover parameters should therefore be selected properly in order to maximize the system performance. The desirable operating point is inside the region close to the ideal point. Based on the results, we select $h = 4 \text{ dB}$ and $T_w = 2 \text{ sec}$. The numbers selected are used for the continuing analysis of system capacity.

Fig. 9 shows the traffic performance analyzed with priority schemes for handover calls. In this figure, Ch.HO denotes the number of channels reserved only for handover. The figure shows that call dropping probability is decreased but new call blocking probability is increased as Ch.HO is increased. As an example, we have analyzed the system which has the required QOS of new call blocking probability below 2% and call dropping probability below 0.01%. With this QOS, subscribers for the system can be maximum at 16,000 when one channel is reserved for handover.

IV. CONCLUSION

An analytical model has been developed and used to analyze the handover process of the GMPCS system. Mean number of handovers as well as handover delay are plotted for different settings of handover parameters such as the averaging window duration $T_w$ and the hysteresis margin $h$. Results show that it is not easy to minimize mean number of handovers and handover delay at the same time. We note that handover parameters can be selected properly from the tradeoff between mean number of handovers and handover delay.

A simulation model has been developed and used to analyze the overall traffic performance of the GMPCS system with handover process. Call dropping probability and mean number of handovers per call are influenced by handover parameter settings. Results show that those performance parameters have the tradeoff with each other. The system performance could be maximized by setting the handover parameters properly as shown in this paper. A number of handover priority schemes can be also applied to increase the system capacity. The study results show that the overall traffic performance of the GMPCS system can be improved by setting system parameters properly and using the handover priority scheme.

REFERENCES