A NEW PRIORITY BASED DYNAMIC HANDOFF ALGORITHM MINIMIZING UNNECESSARY HANDOFFS IN CELLULAR SYSTEMS

Anurag Chandra; Deepak Bansal*, Anand! Rajeev Shorey

IBM Solutions Research Center,
Block 1, Indian Institute of Technology,
New Delhi - 110016, India
email: srajeev@in.ibm.com

Abstract - We propose a dynamic handoff algorithm that, as compared to the static algorithm, yields reduced number of unnecessary handoffs while maintaining the Quality of Service constraints in a cellular mobile system. The fraction of unnecessary handoffs in the dynamic algorithm is found to be 25% to 30% lower than that in static algorithm. The proposed scheme takes decisions regarding initiating handoffs at fixed sampling intervals depending upon the state of the system at that instance. The main idea behind our algorithm is to defer handoff initiation until the handoff failure probability at the next sampling interval exceeds the specified QoS guarantee. We give a mathematical formulation to determine the appropriate time at which a mobile should initiate a handoff.

I. INTRODUCTION

Processing handoffs is an extremely important task in any cellular system. Handoffs must be performed successfully and as infrequently as possible, and be imperceptible to the users. In order to meet these requirements, system engineers must specify a particular signal level as the minimum usable signal level for acceptable voice quality at the base station receiver. In this paper, we refer to the minimum usable signal level as the "threshold". Once the threshold has been specified, a slightly stronger signal level is used at which a handoff is initiated. The margin, \( \delta \) (also referred to as the threshold margin in the paper) given by the difference between these two signal levels, cannot be too large or too small. If \( \delta \) is large, unnecessary handoffs which burden the Mobile Switching Center may occur, and if \( \delta \) is small, there may be insufficient time to complete a handoff before a call is terminated due to weak signal conditions. Therefore, \( \delta \) has to be chosen carefully to meet these conflicting requirements. In Figure 1, we illustrate the handoff scenario at the intersection of two adjacent cells. For a mobile moving from cell 1 (source cell) to cell 2 (destination cell), \( L_i \) is the point at which a handoff is initiated and \( L_f \) is the point corresponding to the minimum acceptable signal level to maintain a call. Therefore, the base station at cell 1 can support a call (originating at cell 1) only up to the point \( L_f \); after this threshold, if a call does not get a channel in the destination cell 2, it is forcefully terminated. It is important to note that \( \delta \) can be determined by considering the mobility pattern and the velocity

*The authors are in the Computer Science Department, Indian Institute of Technology, New Delhi. The work was done while the authors were at IBM Research Center, New Delhi during summer 1998.

Mathematics Department, Indian Institute of Technology, New Delhi.
distribution of mobiles in the cell.

In the literature, many efficient handover strategies have been proposed that minimize the number of handoffs. Minimizing the number of handoffs is important as each handoff increases the signal and processing load, and causes traffic management problems. This problem can be alleviated by identifying necessary and unnecessary handoffs. Note that a handoff is considered unnecessary if the required quality can be sustained without that handoff. Various methods of limiting unnecessary handoffs include: usage of a hysteresis margin, usage of long term prediction scheme when the signal strength is sufficiently high.

An algorithm that initiates handoffs for all mobiles at a fixed RSSI (Radio Signal Strength Indicator) within the handover region, will be referred to as the static algorithm. In contrast to the static algorithm, a dynamic handoff algorithm takes decisions regarding initiating handoffs at fixed decision intervals, depending upon the state of the cellular system at that instance.

In this paper, we propose a dynamic handoff algorithm that reduces the number of unnecessary handoffs while maintaining the Quality of Service (QoS) constraint. The QoS constraint that is of interest in this paper is the handoff failure probability, also referred to as the forced termination probability of handoff calls. The state variables that influence the handoff decision are (i) the time available for the mobiles in the potential handoff region to cross the threshold, (ii) the number of requests in the handoff queue awaiting a channel in the destination cell, and (iii) the number of new handoff calls that arrive in the handover region in a decision interval.

The main idea behind our algorithm is to defer handoff initiation until the handoff failure probability (at the next decision instance) exceeds the specified QoS guarantee. We give a mathematical formulation to determine the appropriate time at which a mobile should initiate a handoff. In a recent paper, the authors have proposed a new handoff ordering method that can be used to provide rapid handovers with a smaller percentage of dropped calls than other methods. They propose a signal prediction priority queuing that uses both the RSSI and the change in RSSI to determine the priority ordering of a mobile unit. The techniques that we have proposed are close to that in [1].

The paper is organized as follows. In Section 2, we explain the notation used in the paper. In Section 3, we discuss the proposed algorithm followed by the mathematical formulation of the problem. Section 4 describes the simulation model and parameters. This is followed by the simulation results in Section 5. We conclude in Section 6.

11. NOTATION

The following notation is used in the proposed algorithm.

- \( P_{m}(t) \): Radio signal strength indicator (RSSI) for mobile \( m \) at time \( t \)
- \( P_{\text{threshold}} \): Radio signal strength indicator for a mobile at the "Threshold"
- \( M: \{ m : \text{mobile } m \text{ is in the potential handoff region and } \% \leq 0 \} \)
- \( R_{m}(t): \left| \frac{P_{m}(t) - P_{\text{threshold}}}{1 - e^{-t}} \right| \)
- \( Q: \text{Number of queued handoff requests in the FIFO sampled at a decision instance} \)
- \( JV: \text{Number of new handoff calls that arrive in the handover region in a decision interval} \)
- \( S: \text{Number of service completions (departures) from the handover FIFO in a decision interval} \)
- \( P_{FT}: \text{Probability of forced handoff termination (or the handoff failure probability)} \)

From the notation above, it is clear that \( M \) is the total number of active mobiles in the potential handoff region at time \( t \) that are moving away from the source (reference) base station and towards the destination base station. \( \% \) is an indicator of the remaining time for a mobile to reach the threshold (i.e., the minimum usable signal level with respect to the source base station). Finally, note that \( P_{FT} \) is the fraction of handoff calls that are forcefully terminated due to non-availability of channels at the destination cell. \( Z+T \) is a critical measure of performance in a cellular system.

111. PROPOSED ALGORITHM

We sample the state of the cellular system at periodic intervals of time, denoted by \( At \) (henceforth referred to as the sampling interval or the decision interval). The \( i \)th decision instance corresponds to time \( iAt \).

To begin with, we describe the ordering used for prioritization of the mobiles in the handover region at each decision instance. The basis of prioritization are the \( R_{m}(t) \)s, that, in turn, depend on the mobile velocity and the direction of motion. The decision
to initiate handoff for a mobile is taken only if handoffs have been initiated for all mobiles with a higher priority than the mobile under consideration. Note that in our algorithm, handoff initiation occurs when a mobile is queued in the handover FIFO. The decision regarding handoff initiation is taken starting from the mobile with the highest priority in the ordered list.

We now describe the proposed algorithm; the pseudocode is given below.

At each decision instance do:

(1) Sort mobiles in the set \( M_t \) in increasing order of \( R_i(t) \). Denote the ith mobile in the ordered list by \( m_i \).

(2) Set \( i \leftarrow 1 \), continue \( t \) true

while ((\( i \leq IMt \)) and (continue = true))

if \( \left( \sum_{j=1}^{Q+i+N-S} X_j \geq R_{m_i}(t) - \Delta t \right) \leq P_{FT} \)
{ Initiate handoff for \( m_i \); \( i \leftarrow i + 1 \);
}
else continue \( t \) false; /* stop as no more mobiles need handover (see Lemma 1) */

In the algorithm, \( X_j \) is the service time (i.e., the channel holding time) of the \( j \)th queued request in the FIFO. From step (1) of the algorithm, it can be inferred that mobile \( m_i \) in the ordered list has the highest priority, since the remaining time for \( m_i \) to reach the threshold (denoted by \&. \( t \)) is the least amongst all the mis in the list. It follows that for all \( i \).

The state variables that influence the decision to initiate handoffs are (i) the time available for the mobiles in the potential handoff region to cross the threshold (\( R_m(t) \)), (ii) the number of requests in the handoff queue awaiting a channel in the destination cell (\( Q \)), (iii) the number of new handoff calls that arrive in the handover region in a decision interval (\( N \)), and, (iv) the number of service completions from the handover queue in a decision interval (\( S \)). The proposed algorithm tries to defer handoff initiation until the handoff failure probability (at the next decision instance) exceeds the specified \&OS guarantee, \( P_{FT} \). The algorithm always meets the desired quality of service, since, at the time of handoff initiation (step (2) in the algorithm), the probability of handover failure does not exceed \( P_{FT} \).

The following Lemma ensures that we stop scanning the ordered list of mobiles when we come across a mobile that does not need handover at the current instance.

**Lemma 1:** Let \( m_i \), \( i \geq 1 \), correspond to a mobile in the ordered list for which handoff can be deferred. Then, handoff for mobile \( m_{i+1} \), \( i \geq 1 \) can also be deferred.

**Proof:** Suppose at time \( t \), the handoff for mobile \( m_i \) is deferred. This happens when,

\[
\sum_{j=1}^{Q+i+N-S} X_j \geq R_{m_i}(t) - \Delta t < P_{FT}
\]

The above equation implies that mobile \( m_i \) does not get queued to the handover FIFO. Since \( m_i \) does not result in an increase in the queue length, the waiting time of mobile \( m_{i+1} \) for handoff initiation is the same as in the above expression, i.e., \( x_2 \cdot X_j \).

Now, since \( R_m(t) \leq R_{m_i}(t) \), it follows that

\[
\sum_{j=1}^{Q+i+N-S} X_j \geq R_{m_{i+1}}(t) - \Delta t \leq \sum_{j=1}^{Q+i+N-S} X_j \geq R_{m_i}(t) - \Delta t < P_{FT}
\]

**IV. SIMULATION MODEL**

We simulate a two cell system; \( A \) is the area of overlap of two adjacent cells, and is therefore the handover region in the system.

In the simulation, the fraction of unnecessary handoffs (denoted by \( \text{PUH} \)) is calculated as follows. Denote by \( n \), the number of active mobiles that are allotted a destination channel (in cell 2), and, whose RSSI with respect to the source cell (i.e., cell 1) increases and crosses the \( L_i \) level. The reader will observe that this serves the purpose of a hysteresis margin. Let \( n \) be the total number of calls that are allotted a channel in the destination cell. Then, the ratio \( \% \) is the fraction of unnecessary handoffs.

The mobility parameter, denoted by \( p \), captures the velocity and the direction of the mobile in the handover region. A high \( p \) would imply that a mobile in the area \( A \) has a high probability of moving towards the destination cell (i.e., crossing the threshold); a low \( p \) would imply that a mobile in the handover region returns to the source cell with a high probability. An interpretation of \( p \) is the number of steps (i.e., the distance) travelled by a mobile before it changes its direction. It is easy to see that
Figure 2: Fraction of Unnecessary Handoffs vs Mobility Parameter $b$. Mobile Velocity is Normally Distributed.

Figure 3: Fraction of Unnecessary Handoffs vs Mobility Parameter $(p)$. Mobile Velocity follows a Uniform (1,20) Distribution.

Figure 4: Handoff Failure Probability vs Threshold Margin $(6)$ for Static Algorithm; $p = 0.88$

Figure 5: Handoff Failure Probability vs Threshold Margin $(6)$ for Dynamic Algorithm; $p = 0.88$

V. SIMULATION RESULTS

In Figure 2 and Figure 3, we plot the fraction of unnecessary handoffs (PUH) versus the mobility parameter $p$. We observe from the two figures that as $p$ increases, the fraction of unnecessary handoffs, for both the static and the dynamic algorithm, decrease. This follows since a high value of $p$ implies that a mobile moving towards the destination cell crosses the threshold with a high probability, and therefore does not significantly contribute to the number of unnecessary handoffs. What is to be noted is that PUH for the dynamic algorithm is substantially lower than FUH for the static algorithm. The difference in the PUH is large (between 25% to 30%) for almost all values of $p$. For very high $p$ (greater than 0.98), this difference reduces and this is as expected. Note that the performance improvement results since the dynamic algorithm defers the handoff initiation process as much as possible so as to reduce the number of unnecessary handoffs in the system while ensuring that the PFT constraint is not violated.

The probability of forced termination of handoff calls or the handoff failure probability ($F_{1-}$), versus the threshold margin $(6)$ is plotted in Figure 4 and Figure 5. Figure 4 shows the static case, whereas the dynamic handoff algorithm is shown in Figure 5. From
both these figures, we observe that as \( b \) increases, \( FL \) decreases. For small values of \( b \), there may be insufficient time to complete a handoff before a call is terminated due to weak signal conditions, thus resulting in a high handoff failure probability.

We plot \( H \) versus \( b \) for different values of \( p \) (mobility parameter) in Figure 6 and Figure 7. For both the static algorithm (Figure 6) and the dynamic algorithm (Figure 7), \( RH \) increases with \( b \); and with increasing \( p \), \( PUH \) decreases. Finally, in Figure 8, the handoff failure probability is plotted versus the sampling interval, \( At \). From the figure, for two distinct values of \( p \), we find the appropriate value of \( At \) that corresponds to \( &T = 0.01 \). Thus, for example, with \( p = 0.96 \), a good choice of \( At \) is between 0.003 and 0.004. Note that for small values of the handoff failure probability, we need to sample the cellular system more frequently and this value of the sampling interval is used in the discrete event simulations.

VI. CONCLUSION

In this paper we have proposed a new dynamic handoff algorithm that attempts to minimize the unnecessary handoffs in a cellular radio system. Since the computational effort required in this algorithm at each step is minimal, it is easily implementable. The algorithm defers the handoff initiation process as much as possible so as to reduce the number of unnecessary handoffs in the system while ensuring that the handoff failure probability constraint is not violated. The algorithm yields significant performance improvement over the static handoff algorithm. More work is required to develop a propagation model which translates path loss into distance in the simulations. A useful extension of the current work can be to find an expression for the mobility parameter \( p \) that captures the velocity and the direction of the mobiles in the handover region of a cellular system.

REFERENCES