2009 International Conference on Machine Learning and Computing
IPCSIT vol.3 (2011) © (2011) IACSIT Press, Singapore

Bandwidth Reservation and Admission Control Based on Mobility Predictions in Cellular Mobile Networks

Esmat Shoja [1+], Vahid Tabataba Vakili [2]

Electrical Engineering Department
University of Science and Technology, Iran

Abstract. The scarcity of available bandwidth is one of the limitations in mobile networks and frequency spectrum must be efficiently utilized. In this paper, a call admission control (CAC) and resource reservation mechanism (RR) is described to provide a consistent QoS guarantee in cellular mobile networks and we propose an idea to improve it. This mechanism is based on user mobility profile which is effectively used for path prediction. The previous behavior of mobile users are collected and analyzed for generating mobility profiles. To make a balance between the call blocking and call dropping probabilities, non-real time new calls are queued and also advantage of variable bandwidth allocation is used with mobility prediction to improve evaluation metrics.

Keywords: Mobility prediction, resource reservation, call admission control, cellular mobile network, variable bandwidth.

1. Introduction

Recently there has been a rapid growth of need in mobile communications and not only the number of users has been increased but also future mobile communication networks support services such as video, multimedia and data which require a higher guarantee from the networks. In order to support such wide range of traffic, the network must be capable of satisfying various quality of service requirements. Since bandwidth is limited, it is necessary to develop mechanisms that can provide effective bandwidth management. Two important characteristics of cellular networks that are important in radio resource management are user mobility and limited bandwidth. From a service provider's point of view, user mobility has a significant overload in bandwidth management. Due to user mobility, some part of bandwidth must be reserved in neighboring cells to avoid call dropping. If the available bandwidth is not sufficient to provide the call service, the new call is blocked and the handoff call is dropped. Good CAC and RR schemes have to make a balance between the call blocking and call dropping. Call admission control is considered as a provisioning strategy to limit the number of calls into the network in order to reduce call dropping. New call may not be accepted if there is not enough bandwidth to support the call. Bandwidth reservation is responsible for previous admitted calls. Call blocking and call dropping probabilities and bandwidth utilization are the three important metrics for QoS. Keeping CDP and CBP as low as possible while maximizing the bandwidth utilization is one of the greatest challenge in mobile networks. Existing bandwidth reservation schemes can be classified into two main categories: static-reservation and dynamic-reservation schemes, in static scheme a fixed portion of the bandwidth capacity is permanently reserved for handoff. This approach is unable to handle traffic load and mobility and is not suitable for real time networks. This scheme decreases CDP and increases CBP and its drawback is degradation in BWU. Dynamic-reservation schemes classified into two main categories: predictive and non-predictive schemes. Predictive techniques can be classified into three main sub categories. The first one is the schemes that are based on the

[+] Corresponding author. Tel.: +(9821) 7808022-3; fax: +(9821) 7454055.
E-mail address: (Esmat_shoja@vu.iust.ac.ir).

mobile‘s past mobility history. The second one is the schemes that use the current mobility factors (velocity, direction, angle or distance) to estimate the next cell, the movement history of the user is not used in these schemes. The third one is a combination of both of these subcategories.

There are two approaches by which handoff calls access the reserved channels: non-sharing and sharing approaches. In non-sharing approach, each reserved channel is strictly mapped to the mobile that made the corresponding channel pre-reservation request. A handoff call may be dropped when all channels are either being used or are being reserved by other calls (not actually being used). In sharing approach, the reserved channels are pooled for being allocated to any handoff call. A reserved channel is randomly assigned for use by any handoff call.

The remainder of this paper is organized as follows: Section 2 gives an overview of CAC and RR schemes and their advantages and disadvantages. Section 3 introduces existing PCACRR scheme. In section 4 we propose two ideas that improve PCACRR scheme. Section 5 describes the simulation environment, its parameters, and the experimental results. Finally, conclusion is given in Section 6.

2. Related Works

It has been some research to predict path based on user mobility profile. In [1] a predictive and adaptive resource reservation technique was proposed. In this technique, the mobility of the user is estimated using the history of user‘s movements recorded in each cell and handoff history. Using this information, the scheme predicts the handoff time and the amount of required bandwidth to be reserved. The scheme predicts user movement based on collected history of all users and may not be precise for each user. It considers only the one dimensional case of the cells, while the two dimensional case is a more realistic. Also if they used two dimensional case, the recorded information that must be saved increased rapidly. In addition, it does not consider different types of users’ movement patterns. In [2] the algorithm is motivated by computational learning theory, which has shown that prediction is synonymous with data compression. The proposed mobility prediction scheme is based on the character-based version of Ziv-Lempel algorithm. The sequence of events (new call, handoff or end of call) during the lifetime of a call corresponds to a substring in the Ziv-Lempel algorithm. A generalized graph model is used to represent the actual cellular network. It recorded both the locations and the handoff times of the mobile users. Also, the prediction is done for each individual user, which makes it more accurate. There are some limitations in this approach, the delay sensitive applications are not considered, which is not suitable for real time multimedia applications in mobile networks. Besides, the approach requires from every BS to reserve the information about all users. In [3] PR_CAT4 is proposed which is suitable for multimedia and real time application. This algorithm predicts the next cell based on the mobility graph. It is also based on the 2-tire structure to determine the amount of bandwidth to be reserved. The prediction of next cell is based on the current cell and previous cell and the number of times the user has moved from the current cell to the next cell. One of the disadvantages in this approach is that only short paths (previous, current and next cells) are considered, which limits the prediction to the next cell only. Also the scheme did not consider any information related to the time at which these movements were performed or for how long the user stayed at each cell. In [4] a road topology based (RTB) mobility prediction is used. The RTB approach assumes that BSs have road-map information and the mobile sets have a positioning system. RTB uses the previous history of handoffs for the mobile users passed through a BS to estimate the time of handoff for the current user in that BS. One of the main advantages of the RTB approach is that it considers also the irregular shapes of the mobile cells which make it more realistic. Also, the authors consider both estimated incoming and outgoing handoff calls to estimate the available bandwidth for reservation. This approach can be enhanced if they consider also the different periods of time during the day. The proposed mobility Prediction algorithm in [5] makes use of the user‘s movement history. Movements are considered to be a combination of random and regular movements matched using a Markov chain model. However it is assumed in the model that most users have regular movement patterns. The algorithm is dependent on the available history of the user’s movement pattern. As a result the accuracy of prediction decreases linearly as random movement increases. In [6] the BS considers not only the number of incoming handoff calls, but also the number of outgoing handoff calls. Then, the BS predicts the total required bandwidth in its cell for both handoff and active calls during a certain time interval.

This mechanism requires complete information about user‘s mobility and duration of their calls. Furthermore, only a single traffic type is considered in this mechanism. Unfortunately, if the network should support diverse types of traffic with various bandwidth requirements, this mechanism can not work. In [7] The hierarchical position prediction Algorithm is proposed makes use of the user's history of movements in addition to the instantaneous RSSI measurements of surrounding cells and cell geometric assumption. While agreeing the next mobility of the user is governed directly by the movement pattern of the user within the current cell. In [8] the user's handoff behavior is modeled based on an aggregate history of handoffs observed in each cell. In reality, mobiles move according to the presence of high ways, streets, and roads. The mobile do not move randomly and follow patterns that are somewhat predictable. The user's mobility patterns are classified into four different classes. User mobility prediction contains sequence of passed cells and also information of the rate of occurrences. These databases are stored in the base station.

3. Introduction of PCACRR Methodology

The PCACRR scheme is mainly based on generation of two types of user mobility profiles [9]. User's behavior will be used to analyze mobility profiles. These profiles contain frequent paths and their associated probabilities and the information related to user's behavior, visited paths, time of day, estimated time for each path. The local mobility profile is generated individually for each user by his mobile. The mobile will have the responsibility to collect the mobility data of a user that uses this mobile. The global mobility profile is generated by each BS. The BS has the responsibility to collect the mobility data of the users passing through the cell. Every time a user switches to a new BS, the mobile will start to predict the next set of BSs according to the current local profile of this user. Once a call has started, this information about the next cells will be sent to the current BS (with additional information about the expected handoff time) before the handoff process. The current BS will use its own global profile if the mobile could not predict the next cell. In this scheme the time of service is considered as an important factor in prediction. This technique improves accuracy of prediction and predicts the future path of the user (not only the next cell). User behaviors and therefore user profiles are not the same for different periods of the day. The day is divided into several intervals, these intervals will be determined generally for the users according to the known rush hours, general work hours, etc. A path may have different probabilities for different intervals in a day. For each of the generated paths two types of probabilities are calculated. The first one is the overall probability for each path during the whole day the other type is the probability of the path in a certain time interval. This probability determines the amount of the required bandwidth to be reserved in every cell belonging to the path. A set of frequent paths are determined and stored in a table called LP(Ti) (local profile at time interval Ti) which will contain the predicted local paths as well as the corresponding associated probabilities. LP table corresponding to each time interval is stored locally in the mobile handset. Global mobility profiles are generated and maintained by the BS. The global profiles contain a set of paths that are most likely to be visited by the users that have passed through the BS. As for the local profile there are two types of probabilities for each generated path in the global profile. BS will have a table called GP(Ti) (global profiles at time interval Ti) that will contain all frequent paths as well as the corresponding associated probabilities.

It is assumed that there are two types of calls: Class-I and Class-II. The calls form Class-I are real time calls (such as video and audio). It means these calls are delay-sensitive. The calls from Class-II are non-real time calls (such as text and image).

In order to make a reservation, the cells have to monitor the available bandwidth at any time. Start to look at the local mobility profile of that user at time interval Ti to predict the frequent paths. If there is at least one predicted frequent path in the local profile, then the bandwidth reservation is performed based on this local profile. If there is not sufficient bandwidth available and the call is from Class-I, then this call can borrow some amount form the Class-II calls in that cells. If there is not any frequent local path then global profile is used.

The admitted call may be a new call or a handoff call. Each of these categories of calls will be handled in a different way taking into account that the handoff calls will have a priority over the new calls. For new calls from class I according to time and day type (may be weekend day), start to look at sequential patterns stored in the mobile of the user. If no match pattern could be find then start to look at all sequential patterns stored in the mobile of the user then if no frequent sequence could be predicted then sequential pattern stored in the BS is used. If each of the next expected cells has available bandwidth to cover this new call then the call is accepted and required bandwidth is allocated, otherwise the call is blocked. For new calls from class II, check only available bandwidth at the current cell at which the admission is performed. If required bandwidth is available then the call is accepted and required bandwidth is allocated otherwise if the available bandwidth is greater than zero then the call is accepted and this bandwidth readjusted in future when there is bandwidth available, otherwise the call is blocked. For handoff calls from class I, handoff calls can allocate bandwidth not only from the reserved bandwidth, but also from the available bandwidth (while the new calls can allocate bandwidth from the available bandwidth only). Check the reserved bandwidth for the handoff calls added to the available bandwidth in the cell at that time interval. If there is required bandwidth the call is accepted and required bandwidth is allocated otherwise the call is dropped. For handoff calls from class II, check the reserved bandwidth added to the available bandwidth in the cell. If bandwidth is greater than zero the call is accepted, otherwise the call is dropped.

4. Proposed Methods

4.1. Queuing New Calls from Class II

As blocking probability and dropping probability are the two important metrics of quality-of-service, we are considering both in order to improve the performance of the system. PCACRR decrease call dropping probability. We added capability of using queue for new calls from class II to decrease call blocking probability. Unlike real-time traffic, non real-time traffic is delay-tolerant. It can be put into the queue. Because real-time traffic is delay-sensitive, the call is blocked immediately if available bandwidth is not enough. If the system is without queue and available bandwidth is not enough, the originating calls arriving at the base station are blocked immediately. If the system is providing queues, then the calls arrived at base station are queued. The calls in the queue will wait for a particular time period. If any channel is released by some call before the call in queue is deleted from the queue, the call is served with the channel. The time period for which the call waits in the queue depends on the threshold parameter. We adjust maximum time in which a call can wait in a queue with a threshold parameter. During this time period if any bandwidth is released and the user doesn't leave the cell, the call is omitted from queue. In this way the calls will obtain some waiting time during which there is a chance of getting a channel if all the channels are busy at the time of its arrival. As non real-time traffic is delay-tolerant, if sufficient bandwidth is not available then new calls from class II arrived at BS are queued but handoff calls are dropped. We do not use queue for real-time traffic. In each BS there is a queue, we use FIFO queuing strategy and we assume the size of queue is unlimited. A class II call waiting in the queue is transferred to the target cell when the mobile user moves out of the current cell before it gets a channel. The calls are served from the beginning of the queue. If available bandwidth is greater than zero and there is any bandwidth available the call is accepted and bandwidth will be allocated. Flow diagram of CAC and queue processing are illustrated in Fig. 1 and Fig. 2.

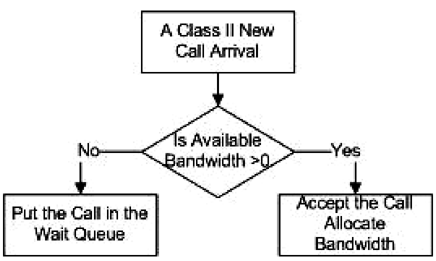

Fig. 1: Flow diagram of CAC

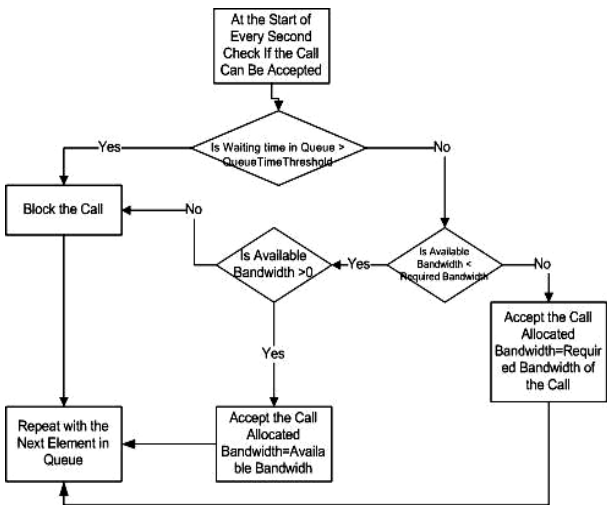

Fig. 2: Flow diagram of Queue processing in the cell

4.2. Combination of Variable Bandwidth and Mobility Prediction

The next generation mobile networks are expected to provide multiple services with diverse Quality of Service requirements to users on the move. Requested services may differ for each individual user. The quality of service offered to the user depends on the price paid for the service and can range over a wide range affecting the quality of the service. For data transmission such as file transfer and text messaging, packet transmission is usually used (compared to circuit or virtual circuit for interactive audio/video transmission). That allows us to arbitrarily control the bandwidth according to network situation. In PCACRR technique they assume fix bandwidth for calls from class I during the lifetime of a call. However many applications such as video and audio can support variable bandwidth. We assume calls from class I have minimum bandwidth requirement and maximum bandwidth requirement. If minimum and maximum bandwidth requirement are equal it means of constant bandwidth. If BW_{req} is requested bandwidth of a call and $BW_{\min}$ is minimum bandwidth requirement, then allocated bandwidth must be between BW_{req} and $BW_{\min}$.

The idea is that when a real-time call arrives at or handoff to a congested cell, the channels allocated be reduced to the minimum channel requirement of the arriving call. The dynamic QoS approach centers on the notion of providing QoS support at some point within a range requested by applications. To utilize dynamic QoS, applications must be capable of adapting to the level of QoS provided by the network, which may vary during the course of a connection. This approach provides the necessary flexibility that allows network elements and end systems to adapt to varying network conditions. The simulation result shows that the performance is superior that of existing scheme. The reason for better resource utilization is due to the adaptive nature of these services that allows the network to offer services whenever there is sufficient amount of resources by intelligently adjusting resource allocation. Flow diagrams of variable bandwidth processing are illustrated in Fig. 3 and Fig. 4.

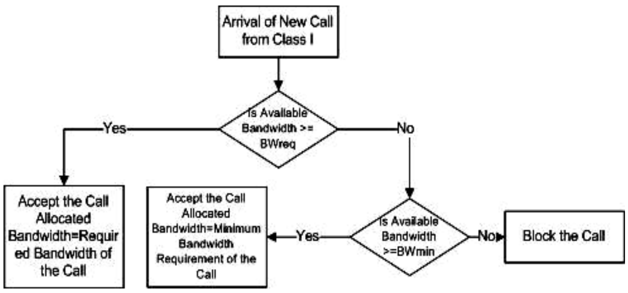

Fig.1: Flow diagram for variable bandwidth processing of new call

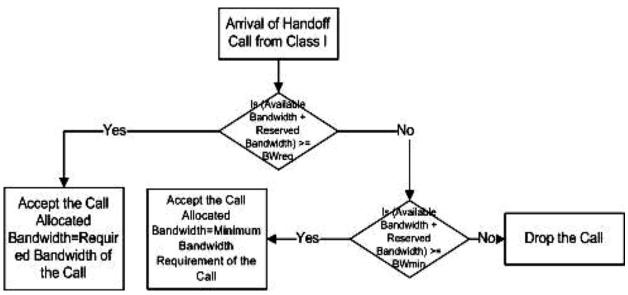

Fig. 2: Flow diagram for variable bandwidth processing of handoff call

5. Experimental Results

We present simulation and experimental results in this section. All profile based approach has two phases: The first phase is the initialization phase. The main objective during this phase is to collect the movement data, which will be used to generate the local and global profiles. Every move of the user is recorded during the call. In this initial period of simulation, we focus only on collecting the data of movements between the cells. Hence a simple CAC technique based on the availability of the resources at the time of request is used in this phase. Also, during this period, a fixed bandwidth is reserved for handoff calls. Second phase is handoff prediction, bandwidth reservation depends on traffic condition and is based on movement prediction. No matter which multiple access technology (FDMA, TDMA, or CDMA) is used, we assume discrete amount of bandwidth is used. In order to evaluate the performance we have built a discrete event simulation tool using C++ language.

The simulation starts by building the topology of the cells. Each cell has a BS. BSs are connected together through wireless or wired networks. Then a graph of paths is produced and mapped on the cell grid. Some of these cells are marked as used cells and the others are marked as unused cells. The unused cells simulate the natural structure of the roads where there are some constraints to deploy BSs in locations (like river, mountain, etc) The used cells are then divided into home cells, work cells or ordinary cells[10]. Each cell has an ID and a flag that determine cell‘s type. These cells will be associated for the users to travel through according to the moving statistics shown in Table I [11].These statistics show the main trips for each user, the percentage of each trip. The user movements in our simulation are based on these statistics. The CAC in BS is responsible for accepting or rejecting a call and is responsible for borrowing management.

Simulation model must show user‘s personal and common habits. The movement path for each user will be set according to home cell, work cell and personal and common habits. User movements are ruled based on cells and road topology in addition to personal and common habits and random events. There are many different parameters that affect user mobility model such as geographical topology, transport way, personal habit and random events. In fact transport way affects the number of handoffs. Knowledge about mobility modes, Movement Observation, Cell Geometry is essential for mobility prediction.

Movement Observation + Movement Model + Cell Geometry = Prediction Algorithm

Having constructed the simulation geography, the simulator enters a cycle of event generation. The event generator generates call and move events. Each event is enqueued at its proper time. At the end of simulation, results are collected in out put files and we use them to analyze. The simulation parameters are shown in Table II.

Table I. Movement Statistics

Purpose of trip	Percentages of trip (%)
To/from work	20.2
Work-related business	1.4
Personal business	52.9
Social/recreation	25.3

Table II. Simulation Parameters

Parameter	Value
Number Of Cells	128
Type Of Cells	Hexagonal
Diameter of each Cell	1 mile
Call arrival rate	Poisson distribution with mean (varies from 0.01 to 0.1 calls/s)
Call duration time	Exponentially distribution with mean = 180 s
Maximum speed	60 mi/h
Channel capacity of each cell	40 bandwidth units (Bu)
BW required by a voice call	1 Bu
BW required by a text call	2 Bus
BW required by a audio call	5 Bus
BW required by a video call	10 Bus

In this simulation we used four different activity periods:

Morning rush hours are between 6 a.m. and 9 a.m. that commuters go to work cells. Work activities are done between 9 a.m. and 4 p.m. Evening rush hours are between 4 p.m. and 7 p.m. that commuters go to home cells. From 7 p.m. to midnight is considered as evening activities.

We show the effect of using queue for new calls from class II in performance metrics and compare it with PCACRR scheme in Fig. 5–7.

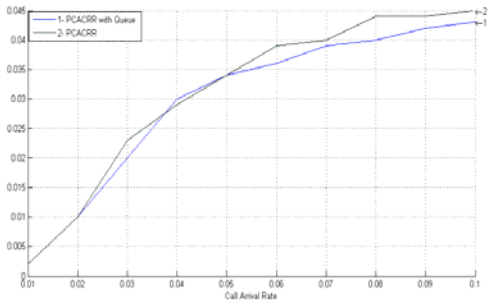

Fig. 5: Effect of Using Queue in CDP

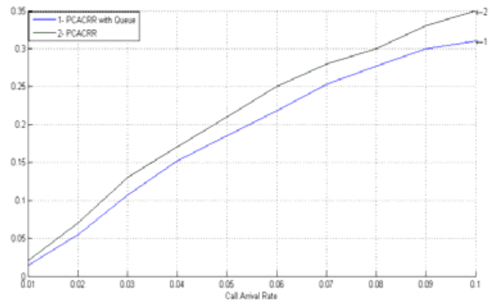

Fig. 6: Effect of Using Queue in CBP

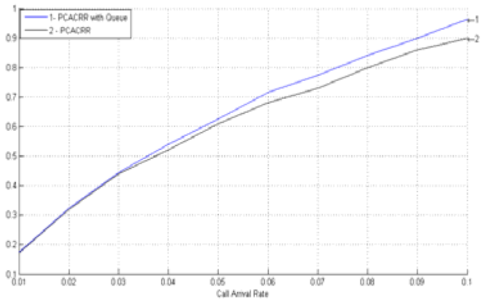

Fig. 7: Effect of Using Queue in BWU

In our scheme the minimum bandwidth for calls from class II is equal to 1. Now for considering effect of minimum and maximum bandwidth, the minimum bandwidth for video calls is set to 7. So a video call could be established with 7, 8, 9 or 10 BU and audio call could be established with 3, 4 or 5 BU. In Fig.8–10 we observe effect of variable bandwidth in performance metrics and compare it with constant bandwidth (PCACRR).

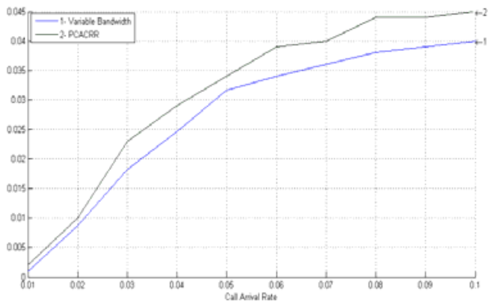

Fig. 8: Effect of Variable bandwidth in CDP

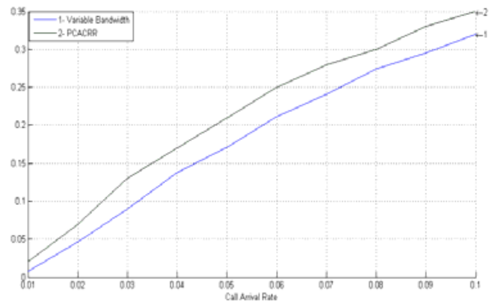

Fig. 9: Effect of Variable bandwidth in CBP

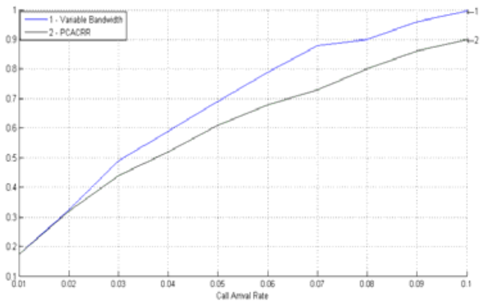

Fig. 3: Effect of Variable bandwidth in BWU

Utilization improvement is due to the adaptive nature of these services that allows the network to offer services whenever there is sufficient amount of resources by intelligently adjusting resource allocation. As we can see by increasing call arrival rate the improvement is more obvious.

6. Conclusions

Call Admission Control plays a very important role in performance of networks. In this paper, we survey a call admission control and resource reservation mechanism for cellular networks which is based on user movement profiles. The required amount of bandwidth to be reserved is calculated according to network situation. Unlike real-time calls, non real-time calls are delay-tolerant and can be put into the queue, so we evaluate effect of queuing new calls from class II.

The next generation mobile networks are expected to provide multiple services with diverse Quality of Service requirements. QoS of requested services may differ for each individual user. We assume calls from class I have minimum bandwidth requirement and maximum bandwidth requirement. Allocated bandwidth is variable and a value between minimum and maximum instead of being a constant value. The simulation results demonstrated that performance metrics improved.

7. References

[1] S. Choi and K.G. Shin, "Predictive and Adaptive Bandwidth Reservation for Handoffs in QoS-Sensitive Cellular Networks," *Proc. ACM SIGCOMM* '98, vol. 27, pp. 155-166, Oct. 1998.

[2] Fei Yu, Victor C. M. Leung: Mobility-based predictive call admission control and bandwidth reservation in wireless cellular networks. *Computer Networks 38*(5): 577-589 (2002)

[3] C. H. Choi, M. I. Kim, S. J. Kim, "Call Admission Control Using the Moving Pattern of Mobile User for Mobile Multimedia Networks," lcn,pp.0431, *27th Annual IEEE International Conference on Local Computer Networks (LCN'02), 2002*

[4] Wee-Seng Soh and Hyong S. Kim, Dynamic Bandwidth Reservation in Cellular Networks using Road Topology based Mobility Predictions, *Proc. of IEEE infocom. 2004*

[5] Liu, G. and Maguire, G. Jr., A class of mobile motion prediction algorithms for wireless mobile computing and communications, *ACM/Baltzer MONET*, 1 (2), 1996

[6] M. Naghshineh and M. Schwartz, "Distributed Call Admission Control in Mobile/Wireless Networks," *IEEE J. Selected Areas in Comm.*, vol. 14, no. 4, pp. 711-717, May 1996.

[7] T. Liu, P. Bahl, I. Chlamtac, Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks, *IEEE Journal on Selected Areas in Communications (JSAC)* 16 (6) (1998) 389-400

[8] Prihandoko,"User Mobility Prediction in Bandwidth Reservation Schemes for Supporting QoS Provisioning in Multimedia Wireless Networks",*ejournal gunadarma*,2004

[9] Mehmed M.kantardzic, SherifRashad, , Anup Kumar: User mobility oriented predictive call admission control and resource reservation for next-generation mobile networks. *J. Parallel Distrib. Comput.*66(7): 971-988 (2006)

[10] K. Al Agha, T. Al-Meshhadany, L. Perato, Resource allocation based on handoff prediction in WCDMA, in: *Proceedings of IEEE 56th VehicularTechnology Conference*, vol. 1, September 2002, pp. 127–131.

[11] D. Lam, D. Cox, J. Widom, Teletraffic modeling for personalcommunications services, *IEEE Comm. Magazine* 35 (2) (February 1997)79–87.