Multi-Radio Access Selection to Reduce Reorder Buffer Requirements in Heterogeneous Wireless Networks with Multi-homed Mobile Terminals

Allen L. Ramaboli, Olabisi E. Falowo, Member IEEE, and Anthony H. Chan, Fellow IEEE
Department of Electrical Engineering, University of Cape Town, Rondebosch, 7701, South Africa
Email: {allen, bisi}@crg.ee.uct.ac.za, h.a.chan@ieee.org

Abstract—Aggregating bandwidth from multiple radio access technologies (RATs) provides an opportunity for network providers to enhance network capacity. This is important for serving applications that require more bandwidth than a single RAT can offer. One of the important elements of bandwidth aggregation is the selection of a suitable set of RATs to achieve efficient aggregation. In this paper, we propose a multi-radio access selection scheme for selecting an appropriate set of RATs to achieve the desired bandwidth aggregation for a connection request initiated by a multi-homed terminal in a heterogeneous wireless network. The objective of the proposed scheme is to reduce the need for packet reordering at the receiver whenever possible. The operation of the proposed scheme is explained in details.

Keywords—bandwidth aggregation; radio access technology; heterogeneous wireless network; multi-mode terminal; packet reordering.

I. INTRODUCTION

Despite the increase in the number of radio access technologies (RATs), radio resources in wireless networks are still scarce. Therefore, network operators find it challenging to keep up with the ever increasing bandwidth demands posed by emerging applications, such as video streaming, online gaming and high definition television, which require high bandwidth. Bandwidth aggregation provides an opportunity for network operators to scale up network capacity in order to serve applications with high bandwidth demands. Bandwidth aggregation takes advantage of the coexistence of multiple RATs in a heterogeneous wireless network; the RATs can be accessed using a multimode terminal (a terminal with multiple radio interfaces). To utilize the aggregated bandwidth capacity, packets belonging to the same application (session) can be transmitted simultaneously across the RATs to the same destination. This can increase application throughput and improve quality of service support [1].

While bandwidth aggregation can boost network operators’ network capacity for high bandwidth demands, the use of multiple RATs to serve an application introduces challenges in the form of packet re-ordering at the receiver and increased battery power consumption cost due to activation of multiple radio interfaces. To address these challenges, efficient bandwidth aggregation architecture needs to be developed. The architecture should execute the following functions: RAT (interface) selection, traffic scheduling, and packet re-sequenceing [2]. RAT selection ensures that the appropriate set of RATs is chosen to aggregate bandwidth for demanding applications; traffic scheduling decides how to split and schedule an application for transmission over multiple RATs. Packet re-sequenceing restores the original order of arriving packets at the receiver, ensuring that multi-path transmission is transparent to higher layers. In this paper, we focus on RAT selection.

The use of multiple radio interfaces to achieve bandwidth aggregation has been studied before. For instance, in [3], a bandwidth aggregation scheme was presented to transmit packets of the same application flow over multiple RATs, ensuring that packet reordering is minimized while efficiently balancing the load across the RATs. The proposed bandwidth aggregation mechanism attempts to reduce the need for packet re-sequence by using transmission buffers, which artificially throttle transmission rate on faster radio links in order to minimize the difference between the heterogeneous links’ latencies. The authors in [2] presented a scheduling policy called earliest delivery path first (EDPF), which minimizes packet reordering by sending packet pairs on the low latency RAT; the scheme deals with residual reordering using the re-order (re-sequenceing) buffer at the receiver. However, the proposal does not explain how the paths (RATs) are selected to achieve the desired bandwidth aggregation; in fact, the use of all available interfaces is assumed. In [4], a bandwidth aggregation mechanism, which uses Round Robin and expected EDPF (EEDPF) scheduling to distribute packets in the form of segments over multiple RATs, has been proposed. The proposed scheduling mechanism increases throughput and, to some extent, reduces packet reordering; but it does not detail how the considered set of RATs is chosen. A time-slotted EDPF algorithm, whose main objective is to minimize reordering, has also been introduced in [5]. The algorithm is a variation of the EDPF in [2], and it has been developed specifically for time-slotted allocation of the aggregated bandwidth capacity; details of how the RATs in use are selected have not been fully explained. Another bandwidth aggregation approach has been presented in [6] to transmit packets of a single flow over two RATs. The proposed approach uses a scheduling technique that delays sending packets on the faster RAT in order to reduce the difference

Abstract—Aggregating bandwidth from multiple radio access technologies (RATs) provides an opportunity for network providers to enhance network capacity. This is important for serving applications that require more bandwidth than a single RAT can offer. One of the important elements of bandwidth aggregation is the selection of a suitable set of RATs to achieve efficient aggregation. In this paper, we propose a multi-radio access selection scheme for selecting an appropriate set of RATs to achieve the desired bandwidth aggregation for a connection request initiated by a multi-homed terminal in a heterogeneous wireless network. The objective of the proposed scheme is to reduce the need for packet reordering at the receiver whenever possible. The operation of the proposed scheme is explained in details.

Keywords—bandwidth aggregation; radio access technology; heterogeneous wireless network; multi-mode terminal; packet reordering.
between estimated delays of the RATs; hence, packet reordering is minimized. In [7] another scheme has been presented to split traffic over two RATs using allocation ratios determined using feedback information from the receiver. The scheme uses the feedback information to balance link bandwidths as a way to reduce packet reordering at the receiver and efficiently balanced the load between the RATs.

While the current bandwidth aggregation schemes, some of which have been highlighted above, can achieve the desired aggregation of the different bandwidth capacities, they seldom address the problem of selecting the best set of RATs for efficient bandwidth aggregation. In fact, most of the schemes assume the use of all the RATs at the terminal’s disposal. This has implications on packet reordering, which increases with the number of parallel paths used [8]. Using a larger number of paths can also affect the battery power consumption cost of the terminal’s radio interfaces, as activating a large number of the terminal’s interfaces increases the battery power consumption cost, resulting in shorter battery lifetime, which can cause premature transmission termination. In this paper, we focus on how to select an appropriate set of RATs to minimize the reorder buffer requirements at the receiver while achieving efficient bandwidth aggregation.

The rest of the paper is organized as follows. Section II presents a brief overview of bandwidth aggregation. System model and assumptions are outlined in section III. Section IV illustrates the operation of the proposed selection scheme. In section VI, we conclude the paper.

II. OVERVIEW OF BANDWIDTH AGGREGATION

Bandwidth aggregation refers to simultaneous use of multiple interfaces for transmission and reception of data. Usually, the data stream is split into smaller streams, which are transmitted through different paths to the same destination. This can enhance application’s throughput and improve QoS provisioning. It is also possible to duplicate the data stream and send the resulting duplicates over multiple paths to the same destination. This is mostly suited for improving delivery of data traffic over error prone and unreliable paths [9]. However, achieving efficient bandwidth aggregation to serve high bandwidth demands is difficult due to challenges in the form of out-of-order packet delivery and unbalanced traffic load, which are a result of data traffic being transmitted over links with heterogeneous and varying path delays. Another important challenge is power consumption cost of the interfaces, which will undoubtedly increase with the number of active interfaces. This is especially important for mobile terminals running on limited battery power. To deal with these challenges, an efficient bandwidth aggregation scheme must be developed. The functions of bandwidth aggregation architecture are [2]: interface selection to identify a proper set of paths to use, traffic splitting and distribution to decide how to split the data stream for transmission over the selected interfaces, and packet reordering at the receiver to restore the original order of arriving packets. A sample configuration of the functional components that form bandwidth aggregation architecture is illustrated in Fig. 1.

As seen in Fig. 1, there are N paths that can be selected to serve an application with high bandwidth demand. The bandwidth aggregation architecture implements traffic splitting and/or duplication to either split or duplicate traffic for transmission over a subset or all of the N paths. Information on path or channel status is required by the scheduler to make accurate traffic splitting and distribution decisions. To get the information, existing traffic and channel estimation techniques found in the literature can be used. Due to the varying delay characteristics of the paths, it is likely for packets to arrive at the receiver out of their original order. Therefore, it is important, as shown in Fig. 1, for efficient bandwidth aggregation architecture to include packet resequencing functionality to correct out-of-order arrival, as some applications and protocols, such as transmission control protocol (TCP), can be affected adversely by packets that arrive out of order.

![Fig. 1. Bandwidth aggregation configuration](image)

A. Packet reordering and Buffer Requirements

Packet reordering occurs when the order of packets of a flow at the receiver is different from the order of the same packets at the sender. That is, the sequence number of the arriving packet of a flow is lower than the sequence number of the consecutive packet that has already arrived at the receiver. The example in Fig. 2 illustrates packet reordering at the receiver. The reordered packets are shown in bold.

![Fig. 2. Packet reordering illustration](image)

Packet reordering in heterogeneous wireless networks is caused by parallel transmission of packets across wireless links that have different end-to-end delays and transmission rates [10], resulting into consecutive packets of a flow arriving at the receiver out of the intended order. An increase in the number of parallel paths and huge delay difference between them can exacerbate packet reordering [8]. Packet reordering can adversely affect the performance of any real-time applications. The time taken to put the received packets in correct order increases the packets’ end-to-end delays, thus resulting into some packets of real-time applications missing
their respective deadlines and being discarded. Transmission Control Protocol (TCP) can also be affected by packet reordering. TCP can allow packet reordering by a maximum of two positions, and this can be corrected by its inherent resequencing mechanism. However, packet reordering beyond two positions will be interpreted as loss, and the transmission window will be reduced. When this happens, the aggregated capacity will be underutilized, and the application throughput may drop drastically. An efficient bandwidth aggregation solution is, therefore, one that includes mechanisms to minimize packet reordering to alleviate the effects of packet reordering so that the associated benefits can be realized.

Reduction of packet reordering can be achieved at the sender by intelligently scheduling packets over multiple paths. It can also be handled using reorder buffer at the receiver. The reorder buffer holds out-of-order packets and puts them in correct order before delivering them to higher layers. The size of the reorder buffer depends on the number of out-of-order packets [8]. The larger the number of out-of-order packets, the larger the size of the reorder buffer required to correct packet reordering. The problem with a large reorder buffer is that it is not a practical measure to reduce packet reordering for handheld devices, which are characterized by limited resources; therefore, its size must be controlled. In this paper, we show how to minimize the size of a reorder buffer by selecting appropriate set paths to achieve efficient bandwidth aggregation.

III. SYSTEM MODEL, ASSUMPTIONS, AND DEFINITIONS

We study bandwidth aggregation in the context of wireless networks. Our envisioned network architectural model consists of a heterogeneous wireless network with tight-coupled RATs as shown in Fig. 3. The RATs are controlled by a radio network controller (RNC) as shown. The RATs belong to a single network operator. We assume that the functions necessary to efficient achieve bandwidth aggregation are executed in the RNC, which monitors and controls the activities of the interworked RATs. The heterogeneous wireless network consists of N RATs, serving a mobile terminal, which can connect to all the RATs simultaneously. The terminal, as illustrated in the figure, can access content from the content server through multiple RATs. Even though the envisioned architecture shows the downlink configuration, the proposed access selection scheme is also relevant for the uplink.

As it was stated earlier, a subset or all of the N paths (RATs) can be used to aggregate bandwidth in order to optimize performance. In this paper, we assume that bandwidth from m (m = 1, ... , N; m ≤ N) of the N RATs is aggregated, forming a larger virtual link that is used to deliver the requested content. Let b_i (i=1, ... , N) denote the amount of bandwidth that the i th RAT can offer for the request. Assuming that m RATs are selected for the request, the total amount of bandwidth aggregated from the m RATs is, therefore, given by:

\[ B = \sum_{i=1}^{m} b_i \]  

Let \( l_j \) be the end-to-end delay of the j th RAT. Let’s assume that the RATs are sorted in descending order of delay. Relative to the slowest RAT, we denote by \( T_s \) the delay difference for the m-1 RATs. According to [8], the total required buffering is given by

\[ Buffering = \sum_{j=1}^{m-1} b_j T_j \]  

For the i th RAT, the required buffering is therefore given as:

\[ b_i T_i \]  

Equations (2) and (3) tell us that the buffering requirement increases with the RAT bandwidth and delay difference relative to the slowest RAT. It has also been reported that the buffering requirement increases with the number of parallel paths (RATs). In this work, we seek to minimize the delay difference and the number of RATs used in order to reduce buffering requirement.

For delay sensitive applications, the maximum delay tolerance is denoted by \( d \). For such applications, the RATs that have been selected for bandwidth aggregation must have estimated delay that is less than or equal to the application’s maximum tolerable delay; otherwise, the application’s data packets will reach the destination late and will be discarded.

Fig. 3. Envisioned Architecture for Bandwidth Aggregation in Heterogeneous Wireless Networks
IV. MULTI-RADIO ACCESS SELECTION

In this section, we present the proposed access selection scheme. The objective of the multi-radio access selection scheme is to select a set of RATs that can result in minimum packet reordering at the receiver. To achieve this, we attempt to use the least number of RATs whenever possible. We also make a selection of the RATs such that the difference in delay relative to the high-delay RAT is kept as small as possible. A small number of parallel paths used for bandwidth aggregation and small delay difference between them have been reported to significantly lower the reorder buffer requirements [8].

Our access selection scheme works by composing possible combinations of available RATs, noting the delay differences relative to the slowest RAT will be determined. Then, a set that has the smallest delay difference will be selected, provided the set satisfies the application bandwidth and delay requirements. It should be noted that for delay insensitive applications, only the bandwidth criterion needs to be met. 

A. Problem Formulation

The multi-radio access selection problem considered in our work can be stated as follows. Given an application requesting R bandwidth units and a set of N RATs in heterogeneous wireless network according to the configuration in Fig. 1, the problem is how to select a set of m (m ≤ N) RATs that minimizes $T_i$, $T_i$ is delay difference between the fastest RAT and the slowest on in the selected set. The set must be selected carefully, ensuring that it consists of the least number of RATs whenever possible. It should also be ensured that the application bandwidth and delay requirements are met. The access selection problem is stated as:

$$\text{RAT} = \text{argmin } T_i$$ \hspace{1cm} (4)

Subject to the following constraints

1. $B \geq R$
2. $l_i \leq d_i$

The variables: $T_i$, $B$, $R$, $l_i$, and $d_i$ are as described above. The solution to eq. 4 returns RAT, which is the optimal vector of RATs selected from the N RATs in the heterogeneous wireless network. The first constraint ensures that the selected RATs have enough aggregated bandwidth capacity to meet the application’s bandwidth requirements. The second constraint seeks to satisfy applications that put strict delay bounds.

B. Access Selection Algorithm

The access selection algorithm deals with how to find the solution to eq. 4 above—how to find a set (vector) of RATs that minimizes the RATs’ delay difference subject to the stated bandwidth and delay constraints. In this paper, we solve the problem by generating all possible RAT combinations and then finding the appropriate combination. Contrary to some of the aforementioned proposals, which use all the available RATs for bandwidth aggregation, we attempt to use the least number of RATs possible. This is crucial for lowering the buffering requirement and packet reordering, which escalate with an increase in the number of active radio interfaces.

Let $k$ be the number of RATs per combination. Then, we start with $k = 2$ and generate a set of all possible $N$ combination $k$ elements. Let this set be denoted by $A_k$. Based on the current and envisioned heterogeneous wireless networks implementations, only a handful of RATs is likely to be used to enhance network capacity; so, from an optimization point of view, the cardinality of $A_k$ is expected to be small. For instance, consider a heterogeneous wireless network with 10 overlapping RATs that a mobile terminal can simultaneously connect to. For $k = 2$, we get a search space of 45 combinations to traverse, and the space gets smaller as $k$ is increased. Therefore, we may not really need advanced search optimization techniques, such as genetic algorithms, particle swarm optimization, and Cuckoo Search, to pick the desired combination—even simple linear search would suffice.

After generating $A_k$, we search for a combination of RATs that solves eq. 4. If we fail to get the right combination for $k = 2$, we increase $k$ and perform a search until we reach $k = N$. The proposed multi-radio access selection algorithm is depicted in Fig. 4.

C. Illustration

Consider a heterogeneous wireless network with a set of integrated RATs, $A = \{\text{RAT1, RAT2, RAT3, RAT4}\}$ and bandwidth request of $R$ bandwidth units. Assume the corresponding RAT latencies are: 5ms, 7ms, 10ms, and 15ms respectively. The respective offered bandwidth is given by the vector $b = \{b_1, b_2, b_3, b_4\}$. We assume that the individual RATs do not have enough bandwidth to serve the request; hence, bandwidth aggregation is necessary. For simplicity, we assume that $b_1=b_2=b_3=b_4$, and the application delay requirement can be met by all the RATs, i.e., $l_1< d_i$. According to the proposed multi-radio access selection algorithm, the RATs that should be used to achieve the required bandwidth aggregation are selected as follows.

First, we generate a set of all possible RAT combinations, starting with $k = 2$ RATs per combination. Let us denote the generated set by $A_k$. Therefore, $A_k = \{(\text{RAT1, RAT2}), (\text{RAT1, RAT3}), (\text{RAT1, RAT4}), (\text{RAT2, RAT3}), (\text{RAT2, RAT4}), (\text{RAT3, RAT4})\}$. The corresponding set of delay differences is $\{2, 5, 10, 3, 8, 5\}$. Suppose that any combination we pick can satisfy the bandwidth requirement. Then, it is clear that (RAT1, RAT2) would be an optimal choice since it is the combination that minimizes $T_i$, resulting in the lowest buffer requirement of $2b_i$.

Now, let us assume that none of the combinations with $k = 2$ can meet the bandwidth requirement. Then, we set $k = 3$ and generate 4 combination 3 RAT sets. The result is $\{(\text{RAT1, RAT2, RAT3}), (\text{RAT1, RAT2, RAT4}), (\text{RAT2, RAT3, RAT4})\}$ with corresponding delay differences $= \{5, 10, 8\}$. So, according to equation 4, (RAT1, RAT2, RAT3) will be selected to serve the application. The total buffer requirement for this combination is $5b_i$. If $k = 3$ had produced RAT combinations that could not meet the bandwidth requirement, then we would have to use all the 4 RATs; in which case, the total buffer requirement would be $10b_i$. 


From the given illustration, we notice that reordering buffer requirements increase with the number of RATs that have different end-to-end delays. Therefore, it makes sense to attempt to use the least number of RATs possible to efficiently achieve bandwidth aggregation. As it has been demonstrated, it is even more important to ensure that the delay difference of the fastest RAT in the selected set, relative to the slowest RAT, is minimal.

V. CONCLUSION AND FUTURE WORK

In this paper, we discussed bandwidth aggregation to enhance performance of applications that require large bandwidth in a heterogeneous wireless network belonging to a single operator. We highlighted challenges of transmitting packets of the same application over the aggregated bandwidth capacity; most importantly, we described how packet reordering can adversely affect performance of TCP applications. We also pointed out that packet reordering can be effectively solved by installing a large buffer at the receiver. However, this may not be practical for handheld devices with limited resources. Other ways to deal with packet reordering is through intelligent packet distribution and interface (RAT) selection at the sender.

We have, therefore, proposed a multi-radio access selection algorithm to select a suitable set of RATs to use for aggregating bandwidth. The algorithm attempts to minimize the reordering buffer requirement by selecting the least number of RATs with the smallest delay difference relative to the slowest RAT. We have illustrated using simple analysis that the proposed access selection scheme can reduce the need for a large reorder buffer. In future, we intend to analyze the behavior of the proposed access selection algorithm using simulations. We also intend to develop our solution into a multi-criteria access selection, which will incorporate other important performance parameters, such as power consumption cost.

REFERENCES


Allen Ramaboli received his master of engineering in telecommunications at the University of Cape Town. He is currently pursuing a PhD in Electrical Engineering at the same institution. His research interests include: radio resource management in wireless networks and mobile application development for next generation networks.