ACRoMa - An Architecture of Cooperative Routing Management in Wireless Mesh Networks

Vinicius C. M. Borges[1], **Marilia Curado**[2], **Edmundo Monteiro**[2]

[1]Institute of Informatic (INF)
Federal University of Goiás (UFG)
Goiânia – Brazil

[2]Laboratory of Communications and Telematics, Center for Informatics and Systems
University of Coimbra
Polo II, 3030-290 Coimbra, Portugal

`{vinicius}@inf.ufg.br, {marilia,edmundo}@dei.uc.pt`

Abstract. *Wireless Mesh Networks (WMN) provide a wireless backbone for ubiquitous Internet access and are being challenged to improve their management to support various kinds of scalable multimedia applications. This paper sets out an Architecture of Cooperative Routing Management (ACRoMa) for scalable triple play service in WMN. A simulation study has been carried out to assess the performance of ACRoMA with different configuration matrices. The results provide evidence that by combining an efficient clustering and load balancing mechanism with a cross-layer routing metric aware of link quality, ACRoMa improves the traffic performance in the presence of challenging traffic patterns, such as triple play services.*

Resumo. *Redes em Malha Sem Fio (RMSF) fornecem um backbone sem fio para acesso ubíquo à Internet e estão sendo desafiados a melhorar seu gerenciamento para suportar vários tipos de aplicações multimídia de forma escalável. Este trabalho apresenta um arquitetura de gerenciamento chamada Architecture of Cooperative Routing Management (ACRoMa) para serviços triple play em RMSF. Um estudo de simulação foi realizado para avaliar o desempenho dos ACRoMa com diferentes matrizes de configuração. Os resultados fornecem evidência de que através da combinação de um agrupamento eficiente, um mecanismo de balanceamento de carga e uma métrica de roteamento cross-layer ciente da qualidade do link, ACRoMa melhora o desempenho do tráfego na presença de padrões de tráfego desafiadores, tais como serviços triple play.*

1. Introduction

Although Wireless Mesh Networks (WMN) [Akyldiz et al. 2005] are not subject to the traditional restrictions of more traditional ad hoc networks (e.g. energy and processing capacity), they have mainly employed the IEEE 802.11a/b/g standards as a form of wireless technology which results in a restricted wireless link capacity (e.g. a limited number of non-overlapping channels). The wireless backbone constitutes the main component of the WMN structure, as it comprises mesh routers and gateways and multi-hop paths are formed through the mesh routers towards the gateways. Access to and from the Internet is processed through the gateways, which can become bottlenecks. Moreover, WMN seeks to support services that requires suitable Quality of Service (QoS) levels, e.g. triple play services. Triple play services [Ekling et al. 2007] combines voice, video and data applications in a single access subscription (service providers). The provision of these services is a challenging task for WMN, since it is difficult to manage the limited resources effectively, so that they can support the service assurance needed by these kinds of services.

Scalability is a critical management issue in WMN, as it seeks to handle growing amounts of traffic load, as well as an increasing number of network elements, while providing suitable QoS levels. In this context, setting up a routing path in a large wireless network may take a long time, and the end-to-end delay may be unsuitable for delay-sensitive applications. Furthermore, it should be pointed out that the low-cost solutions provided by these networks make it easier to increase the size of the WMN and enable it to cover larger areas. In light of this, the dynamic routing process has become one of the most useful mechanisms to complement the current wireless technologies (e.g. cognitive radio, Multiple-Input Multiple-Output (MIMO) and Long Term Evolution (LTE)) and thus support the requirements of multimedia applications. The routing process comprises routing algorithms, protocols and metrics that allow computation of the best routes in the network. Moreover, it offers a more complete performance optimization of the wireless medium without additional deployment costs and thus results in an autonomic and synergetic management of WMN.

The main purpose of this paper is to present a research work which has been undertaken by means of an architecture for cooperative routing management, called Architecture of Cooperative Routing Management (ACRoMa), that can address the scalability issue of the application traffic in WMN; it comprises a clustering load balancing routing schema and a cross-layer routing metric. The specific goals of this paper are twofold. First, the paper outlines the main components of the architecture as well as their interactions and synergies. Second, there is an analysis of the performance evaluation results of the different mechanisms described, which takes account of tangible scenarios where the configuration matrix is composed of triple play applications in WMN. To the best of our knowledge, this paper is the first example of a triple play services performance where the different load balancing routing methods are compared with different network size and topologies.

The remainder of the paper is structured as follows: Section 2 discusses the main open issues to provide a scalable solution in WMN. Section 3 sets out the proposed architecture. Section 4 presents the performance evaluation of the main component of the architecture. Finally, Section 5 describes the conclusions and makes recommendations for further studies.

2. Open Issues

There has been considerable discussion about ways to improve scalability through the routing process in emerging network management architectures [Azcorra et al. 2009, Zhu et al. 2011, Ashraf et al. 2011]. Usually these approaches combine the routing process with spectrum management [Azcorra et al. 2009], channel assignment [Zhu et al. 2011] and link breakage assessment [Ashraf et al. 2011]. In other words, most of the solutions found in the literature cause a large overhead and delay in large wireless networks. To tackle this recognized problem, clustering schemes have been employed in WMN [Langar et al. 2009, Wu et al. 2014] to improve the management of the routing decision making process. This is because they increase the scalability of the current routing protocols in large wireless networks by reducing the routing overhead. This type of scheme divides the WMN into different kinds of virtual groups, where the nodes are allocated geographically so that they are adjacent to the same cluster and conform to specific rules. This means that WMN become self-organized in a modular and virtual topology, where a cluster consists of a gateway (i.e. clusterhead) and a set of mesh routers in WMN.

Although the clustering schemes improve the performance of routing protocols in WMN and make easy the WMN management, clustering is not sufficient to achieve a truly scalable solution when the traffic load increases in the network. This means that routing decisions that focus on load balancing, play an important role in WMN, both at the intra-cluster and inter-cluster levels. Intra-cluster load balancing schemes [Hsiao et al. 2001, Dai and Han 2003, Gálvez and Ruiz 2013] handle the load balancing locally (i.e. inside a single cluster), by distributing the traffic load among the routing sub-trees in which the gateway is the root. Nevertheless, intra-cluster routing load balancing can not distribute the traffic load uniformly throughout the whole network, since the intra-cluster load balancing is restricted by the capacity of the gateway.

The inter-cluster load-balancing deals with load balancing by reducing the cluster congestion in a holistic perspective, and directing the mesh router traffic towards lightly-loaded gateways. It thus, improves the overall capacity of the network by cooperating with each other. Hence, inter-cluster load balancing routing between multiple gateways is a necessary mechanism to manage the traffic load in WMN in a scalable way. The inter-cluster load balancing routing is an efficient solution to provide a horizontal cooperation in the network layer between all the mesh nodes that improve the traffic scalability, where these nodes must have a collective awareness of the traffic load in the adjacent clusters (i.e. the nodes share the information about the cluster traffic load with each other). There are some proposed approaches that have been established for the mesh router migration method, where the Load-Balancing Approach (LBA) [Xie et al. 2008], Partition-based Load Balancing (PLB) [Choi and Han 2010] and DIffusion Load Balancing (DILB) [He et al. 2009] are the most widely accepted.

LBA, PLB and DILB are based on a load threshold that enables them to decide whether the inter-cluster routing will occur or not and also to select the lightest adjacent cluster. LBA is primarily based on the hop count metric to make an inter-cluster decision, and it does not consider intra-cluster load-balancing. On the other hand, PLB and DILB take into account a more accurate load metric than LBA, i.e. the number of flows for each mesh router. Moreover, PLB and DILB perform both intra-cluster and inter-cluster load balancing routing. DILB extends PLB by taking into account nodes with different number of wireless interfaces. However, the mesh router migration results in a very slow traffic migration.

Since the wireless medium is shared, the interference and traffic load are the main factors that influence the link quality in the wireless links. It is also worth noting that the traffic load also causes interference (i.e. self-interference) and increases the congestion in the wireless links. For these reasons, network layer routing awareness of interference and traffic load is a key enabling tool to optimize the wireless resource, since it avoids paths with a high interference level and traffic load. In view of this, cross-layer routing metrics play a key role in measuring interference levels and traffic load using local information to make a routing decision in a distributed way, while avoiding introducing the excessive overhead that is caused by the measurement and distribution of this information. The cross-layer design has been employed in WMN to exchange information between different layers; for instance interference and traffic load are picked up from the MAC and physical layers to support the routing decision. In this way, the cross-layer design enables a vertical cooperation in WMN where information from different layers is combined. On the basis of an extensive state-of-the-art analysis, it was pointed out that the accuracy of the existing cross-layer routing metrics is not sufficiently accurate to depict the interference and traffic load precisely [Borges et al. 2011], such as Interference-Load Aware (ILA) [Manikantan Shila and Anjali 2008], Contention-Aware Transmission Time (CATT) [Genetzakis and Siris 2008] and Contention Window Based (CWB) [Nguyen et al. 2008].

It is important to stress that, in attempting to improve the scalability of the traffic applications for WMN without additional costs, previous work has failed to take account of the integration of a clustering scheme, load balancing routing and cross-layer routing metrics. This integration improves the overall network performance through the routing process, by achieving a greater degree of traffic scalability and hence, enabling paths to be selected that can satisfy the requirements of applications such as VoIP and video, as well as more complex configuration matrixes including triple play services.

3. ACRoMa - An Architecture of Cooperative Routing Management

ACRoMa integrates the most significant means of managing the routing process in order to improve the scalability of WMN, allowing a higher degree of traffic performance to be achieved. Figure 1 describes the architectural model that combines the three components and allows them to cooperate.

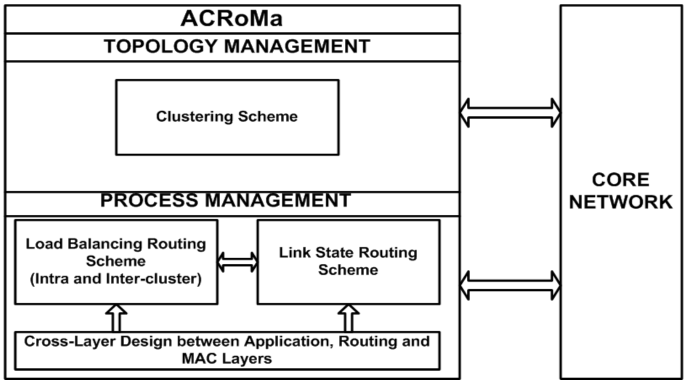

Figure 1. ACRoMa - Architectural Model

It combines the following components: a clustering scheme, load balancing routing algorithms, and a cross-layer routing metric. When clustering routing scheme, cross-layer routing metrics and load balancing routing are combined, they can collaborate to support features such as self-configuration, self-healing and self-optimization and this reduces the need for human intervention in network management. The specific goals of the architecture are as follows:

- to enable the best paths to be selected by depicting accurate measures of the link quality through a cross-layer routing metric.
- to reduce the routing overhead of the traditional routing protocols by using a clustering scheme.
- to avoid overload situations in the gateways through an inter-cluster load balancing routing algorithm.

ACRoMa employs a bottom-up approach which involves integration and testing; the components were integrated in an incremental way from lowest level components to highest level components. In the light of this, each component was tested separately and then aggregated incrementally. The main synergies between the components are as follows: the cross-layer routing metric provides information which helps to make link state routing decisions, the clustering scheme provides a virtual structure that allows an efficent load balancing, while reducing the routing overhead. Furthermore, each component seeks to overcome any limitations found in its respective related work. The components and their interactions will be discussed in the next sub-sections.

3.1. MIND - Cross-layer Routing Metric

The Metric for INterference and channel Diversity (MIND) [Borges et al. 2011] combines measurement that take into account interference and traffic load through accurate and passive measurements. To reach this, MIND employs a relation between Signal Noise Ratio (SNR) and Signal to Interference plus Noise Ratio (SINR) as well Channel Busy Time (CBT) to depict interference and traffic load, respectively. These measurements can be obtained from MAC and physical layers. MIND regards Channel Busy Time as a smooth function of multiple weighting through measurement of interference. For this reason, MIND strikes a combination between interference and load, in which interference has a higher weight than traffic load. MIND also uses smoothing out functions to avoid routing instability. For instance, the SINR and CBT measurements are smoothed out through their respective averages of a set of packets. In addition, they are passive measurements and thus, no additional overhead of the active monitoring mechanisms (e.g. AdHoc Probing) are required to obtain them. As a result, it cooperates with the clustering scheme to mitigate the routing overhead and to improve the load balancing. Furthermore, as was made clear in [Borges et al. 2010], MIND improves the traffic performance of triple play services in WMN.

3.2. Collaborative Clustering Scheme

The main purpose of the proposed clustering scheme, called Collaborative Clustering Scheme (CoCluS) [Borges et al. 2010] is to provide a clustering structure that enables more efficient inter-cluster load balancing routing than mesh router migration method in PLB [Choi and Han 2010]. In view of this, the drawback in PLB is demonstrated first and after that, CoCluS is described. Moreover, with clustering, routing decisions become more precise, due to the smaller scale where cross-layer routing metrics are employed. CoCluS is described in the next paragraphs.

Figure 2 shows the network model used in the clustering scheme that has been set out. Mesh routers form a tree-like structure that is used to communicate with the gateway.

In this way, the network is partitioned into clusters in which the root is a gateway. Each mesh router is characterized by its weight which depicts the load level and is usually represented by the number of active flows. These flows are derived from mesh clients who attach themselves to the mesh router. The Cumulative Load (CL) is the sum of the weights of all the nodes in the sub-tree, including the weight of the root. Thus, the CL of a node is the amount of uplink traffic present on the node. The links between a gateway and its neighbor nodes are called Top Sub-Links (TSL), and the neighbors that are one hop from the gateways are called Top Sub-Nodes (TSN). A TSN of an adjacent cluster is called an adjacent TSN. The overload condition occurs when the CL of TSN exceeds the defined maximum load threshold. The network is indicated as a matrix m(x, y), where x is the x-axis index and y is the y-axis index. The numbers in the squares correspond to the weight of each node. The numbers which are alongside the TSL are the CL in the TSN sub-tree.

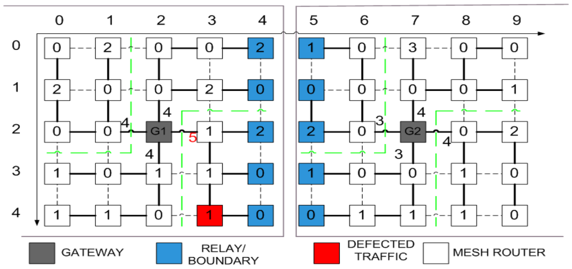

Figure 2. CoCLuS Network Model

First, *m(4,4)* is migrated (Figure 3), since it is a border node and has one of the smallest CL. Next, *m(4,3)* is also migrated (Figure 4) since it is the next border node which has one of the smallest CL. However, they do not help to improve the load balancing of the network, since it actually has no traffic load. It should be noted that *m(3,4)* is a better candidate to make the load balancing more efficient, but is not yet a border node in Figure 3. Hence, *m(3,4)* has to wait to become a border node with smallest CL, which occurs when *m(4,4*) and *m(4,3*) migrate to the adjacent cluster (Figure 5). Figure 6 shows the balanced clusters *G1* and *G2* after the migration of three mesh routers. It is important to point out that the clustering structure was modified by the migration process.

It is also important to take note of the messages required by this method. The messages required by this method are illustrated in Figure 3. The *G1* gateway sends the defection request message (blue arrow) to *m(4,4)* which then forwards it to *m(7,3),* the adjacent TSN. When *m(7,3)* receives this message, it sends back a defection response message (red arrow) to the *G1* gateway and *m(4,4)* to confirm the acceptance status of *m(4,4)*. The defection decision could have been made locally at *m(4,4)*, if the nodes had had the information about the CL of the TSN in the adjacent clusters. In this case, *m(4,4)* would not need to forward the defection request message to *m(7,3)* and thus, could reduce the time needed to make the inter-cluster routing decision.

The relay and boundary nodes are new elements of the proposed clustering scheme which enable an exchange of information (e.g. CL of adjacent clusters) to occur with the mesh routers belonging to the adjacent clusters, since they are within each other's transmission range. As a result, the relay and boundary nodes provide information that can support the inter-cluster routing decision. Even though the boundary and relay nodes

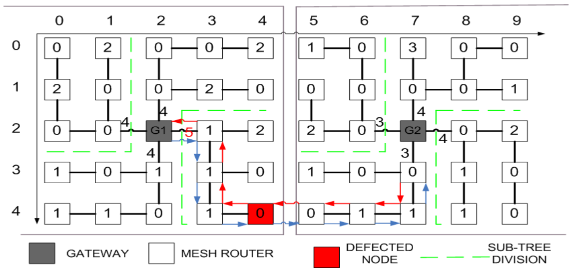

Figure 3. Mesh router migration: Step 1

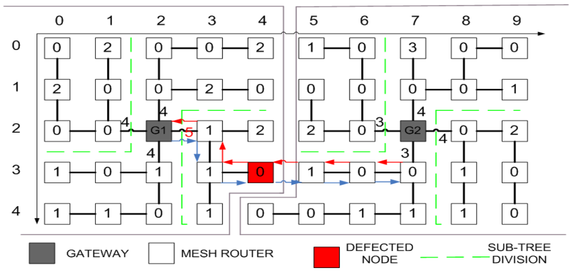

Figure 4. Mesh router migration: Step 2

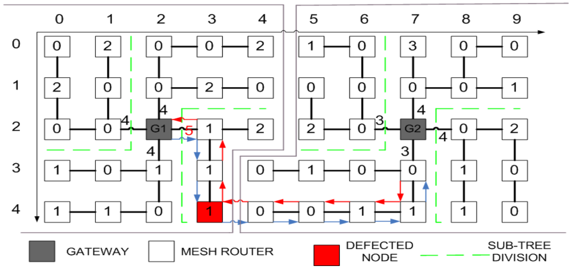

Figure 5. Mesh router migration: Step 3

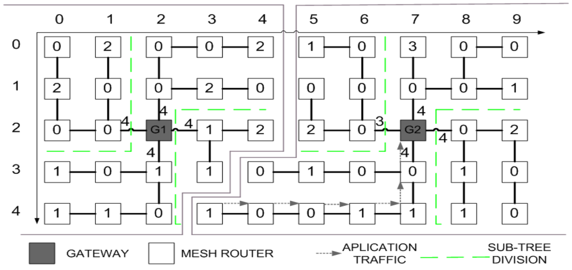

Figure 6. Mesh router migration: Step 4

play a similar role in the traffic migration process, they are described in distinct ways, depending on the cluster in which the mesh routers are found. For instance, m(4,4) is a relay node for all the mesh routers in the G1 cluster and a boundary node for all the mesh routers in the G2 cluster. In other words, a boundary node does not belong to the cluster, whereas a relay node does.

Although the relay node and its respective boundary node are not in the same cluster, the relay node receives the CL of the adjacent TSN because the boundary nodes disseminate this information to their neighbors inside the cluster, as well to the relay nodes. In this way, the candidate is able to select the lighter adjacent cluster locally. Hence, the candidate nodes do not need to send a defection request to the adjacent TSN, since the clustering scheme allows a more proactive migration strategy to start the traffic migration, which further reduces the time required to start the traffic migration.

CoCluS employs a new hybrid routing scheme which combines two different routing structures which cooperate with each other to improve the overall network performance. First, the spanning tree structure (solid line) is used to communicate with the gateway (i.e. intra-cluster load balancing routing scheme). The Inner Domain Load Balancing (IDLB) algorithm, proposed in [Choi and Han 2010], is employed as the intra-cluster load balancing routing algorithm to form the spanning trees rooted at the gateways. Afterwards, the nodes calculate the routes to every neighbor (specially for relay and boundary nodes) inside the cluster by means of the Dijkstra routing algorithm and the MIND cross-layer routing metric (i.e. the link state routing scheme). This latter routing scheme (dotted line) is necessary to forward data from the defected traffic to the selected relay nodes (intra-cluster path).

3.3. RAILoB - Inter-cluster Load Balancing Routing

The Routing Algorithm for Inter-cluster Load Balancing (RAILoB) [Borges et al. 2012] approach employs a new traffic migration method, called mesh traffic migration, that enables the main limitation of the mesh router migration method [Choi and Han 2010] to be overcome, which is its slowness.

The mesh traffic migration method allows the traffic migration of the selected mesh routers without the need for mesh router migration (i.e. only traffic application is migrated). This new method enables candidate nodes to be selected for the traffic migration in which the candidate is not required to be a border node. The main purpose of this method is to find a flexible means of reducing the traffic load in the nodes which are close to the TSN, while keeping control of the number of hops required to reach the destination. As a result, it is better if the criteria for selecting the candidate nodes are based on the nodes which are farther away from the gateway in the sub-tree of the TSN overload. This method has two advantages. First, it avoids links close to the congested gateway. Second, it increases the likelihood that nodes will be selected that are closer to the adjacent cluster. By adopting this flexible method, RAILoB can provide agility to the process of traffic migration and thus, reduce the time needed to carry out the inter-cluster traffic routing. The mesh traffic migration requires a more complex clustering structure which is provided by the clustering scheme explained in the previous sub-section.

The complete path consists in the mesh traffic migration of two main sub-paths, namely, intra-path (the path between the selected node and the relay node, using the link state routing with MIND) and inter-path (the path between the relay node and the lighter gateway, using the spanning tree). Figure 7 also shows an example of traffic migration when RAILoB is employed for the same case.

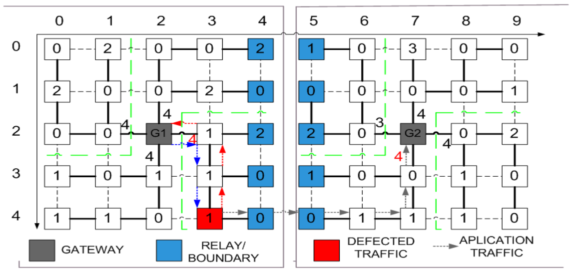

Figure 7. Mesh traffic migration - Example

There is an overload condition in *m(3,2)* in Figure 7 (CL with a value of 5), the *G1* gateway chooses *m(3,4)* for traffic migration and sends it a defection request message (blue dotted arrow). Next, *m(3,4)* checks in its routing database and finds *m(7,3)* (i.e. TSN that can accept the traffic in *m(3,4)* without overloading it). Then, *m(3,4)* sends back a defection response message (red dotted arrow) and starts to allow the traffic to migrate (dotted gray arrow) using *m(4,4)* and *m(5,4)* as relay and boundary nodes, respectively.

ACRoMa seeks solutions for each of the open issues previously discussed, such as a clustering solution to reduce the routing overhead, a load balancing routing algorithm to avoid overload situations at the gateways, and a cross-layer routing metric to improve the accuracy of the route selection process. It should be pointed out that these solutions are coordinated to increase network scalability (e.g., greater degree of traffic performance and greater number of nodes) and thus, improve the overall capacity of WMN.

4. Simulation Study

The simulation study outlined in this section aims at throwing light on the ability of ACRoMa to confirm the hypothesis that it has the potential to achieve a greater degree of traffic performance when a more efficient routing solution is used. For this reason, a comparison was drawn between RAILoB and the most effective inter-cluster load balancing routing (i.e. Partition Load Balancing (PLB) [Choi and Han 2010]), since PLB is the most effective related work on routing and RAILoB represents the ACRoMa architecture conceptually by combining all the components. In this section, the performance evaluation within the NS2 simulator will examine a mixed traffic comprising the VoIP, video and FTP applications which configure triple play services. In this way, we will be able to evaluate the impact of load balancing methods on each application of these services. Particularly in this paper, we investigate the impact of the routing approaches on different aspects of WMN, such as network size and different types of network topology. Such evaluation was not taken into consideration in [Choi and Han 2010, Borges et al. 2010, Borges et al. 2011, Borges et al. 2012].

4.1. Effects of Network Size

The performance evaluation will also assess a triple play service configuration when varying the network size and the impact of the inter-cluster routing methods on this factor will be analysed. The scenario configuration and traffic model are outlined in sub-section 4.1.1. The simulation results are examined in sub-section 4.1.2.

4.1.1. Simulation Configuration

Each data point in the graphical results is computed as the average of 10 different simulations and the graphs also show the confidence intervals of the performance parameters which have a confidence level of 95%. The inter-cluster load-balancing approaches were implemented in an extended version of the OLSR routing protocol [Ros and Ruiz 2007] by means of the NS-2, which supports the clustering. All of the nodes have the same physical configuration. Table 1 shows the configuration of both scenarios used in this sub-section.

The traffic combination of each application was based on [Quintero et al. 2004][Kim et al. 2008]. That is, the percentage of flows for VoIP, FTP and video are 60%, 30% and 10% of the total load, respectively. Thus, a set of four combinations (two for each network size) of mixed traffic were formed, as shown in Table 2. Both scenarios have the same traffic proportion by gateway, which makes it possible to analyze the impact of the network and cluster size on the inter-cluster routing methods.

The scenario consists of gateway (one for each cluster) and static mesh routers with multi-channel multi-radio capability, which is typical of outdoor city-wide deployments. There are two channels and two network interfaces. On each node, one particular channel is combined with one particular network interface, and no channel assignment

Table 1. Scenario Setup

Parameter	Value
Simulation Time	300s
Flow Lifetime	275s
Network Sizes	50 and 100
Cluster Sizes	17 and 20
Number of Gateways	3 and 5
Grid Topology Sizes	2000m x 2000m, 2500m x 2500m
Transmission Range	250m
Interference Range	550m
Propagation Model	TwoRayGround
Network Interface Cards	2
MAC/PHY Specification	IEEE 802.11 b/g
Antenna	Omnidirectional

Table 2. Traffic Combination for Triple Play Services for Different Network Sizes

Combinations of Number of Flows	**Video**	**FTP**	**VoIP**
A for 50 Nodes (Low load)	1	4	12
D for 50 Nodes (High load)	4	12	24
A for 100 Nodes (Low load)	3	8	20
D for 100 Nodes (High load)	10	20	40

algorithm has been employed. Furthermore, grid topology is used to limit the maximum number of neighbours of a mesh router (i.e. four at maximum). The scenario uses a typical WMN backbone traffic pattern feature, where several flows originated from the source nodes (i.e. mesh routers) to a destination node (i.e., gateway), and the source nodes were chosen at random. The gateway is located in the central position [Bejerano et al. 2007].

4.1.2. Simulation Results

Figures 8 and 9 show that the network size has little impact on throughput of video and VoIP applications, whereas these factors have a signifcant effect on FTP application. For example, FTP achieves 408,78 Kb/s and 263,84 Kb/s in the highest load D in a network size of 50 and 100 nodes respectively, when using RAILoB. This can be explained by the fact that an increase of the network size tends to raise the interference level and traffic load. The transmission rate control policy of the TCP protocol is very sensitive to the packet loss rate which rises to the same extent that the interference and traffic load increase.

As expected, figures 10 and 11 shows that network size has greater impact on delay than throughput for all aplications. The reason being that increasing the network size also increase the average path length which increases delay in wireless networks. Nevertheless, the impact is smaller when RAILoB is used for distinct network sizes. Furthermore, when comparing throughput and delay for all load configurations and network sizes, RAILoB achieves better results. This can be explained by the fact that RAILoB is more agile and fexible than PLB, while keeping the same cluster structure and therefore,

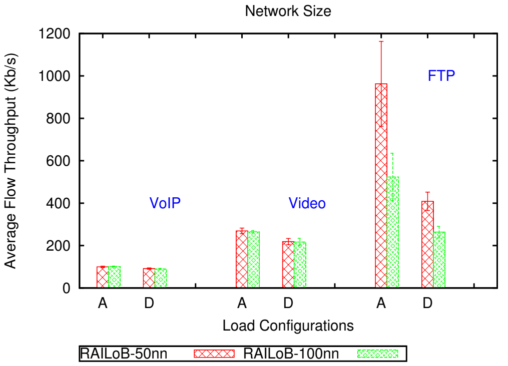

Figure 8. Average Flow Throughput of RAILoB

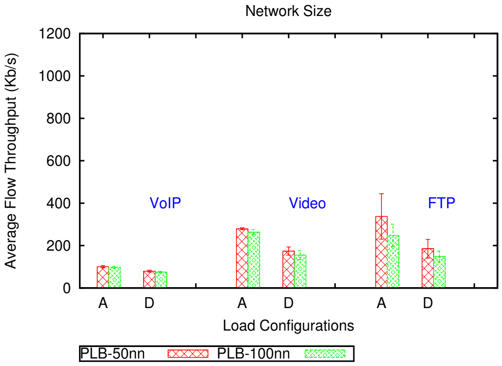

Figure 9. Average Flow Throughput of PLB

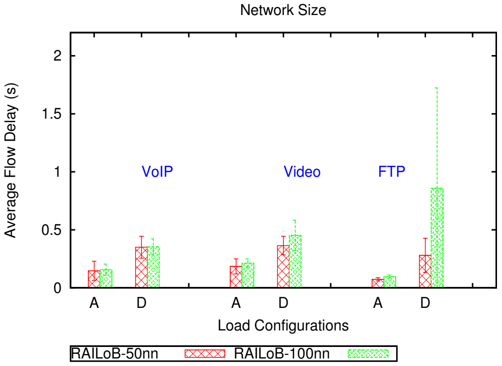

Figure 10. Average Flow Delay of RAILoB

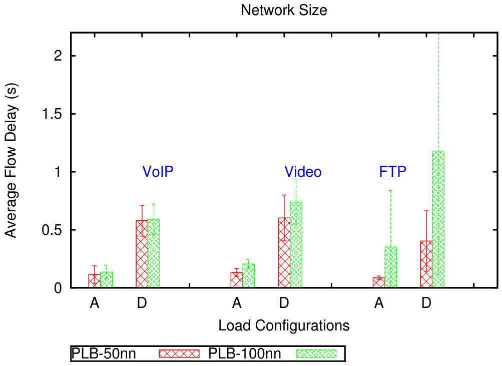

Figure 11. Average Flow Delay of PLB

the triple play services are able to reach lighter adjacent clusters more quickly and the overloaded gateways are lightened at a faster rate. As a result, the overall network capacity is improved since RAILoB deals with growing amounts of traffic load and nodes better than PLB.

4.2. Effects of Topology Scenario

The effects of topology types on the routing approaches will be investigated in this sub-section where a triple play service configuration is employed. This sub-section is structured as follows sub-section 4.2.1 shows the scenario configuration and traffic model. The simulation results are described in sub-section 4.2.2.

4.2.1. Simulation Configuration

The scenario configuration is very similar to subsection 4.1.1. In addition, the traffic model is equivalent to that used in subsection 4.1.1. These tests are also used here to compare random and grid topologies. The amount of traffic is the same for both topology types. The network parameters of network size, topology size and number of gateways used from the previous scenario are 50, 2000m x 2000m and 3, respectively.

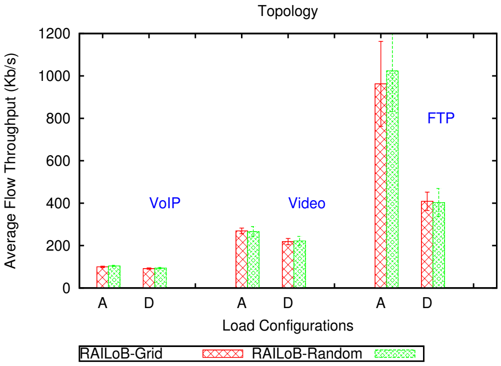

Figure 12. Average Flow Throughput of RAILoB

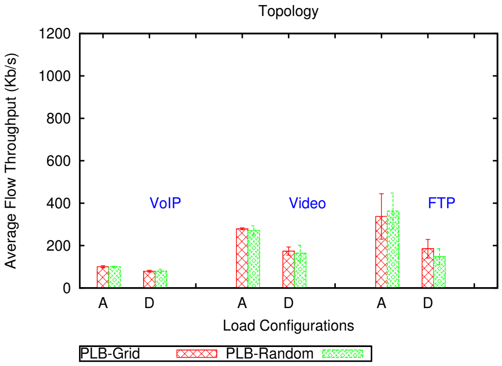

Figure 13. Average Flow Throughput of PLB

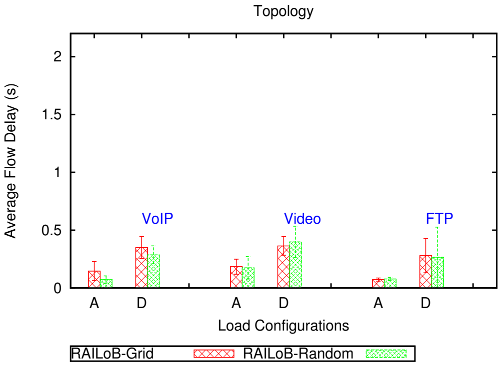

Figure 14. Average Flow Delay of RAILoB

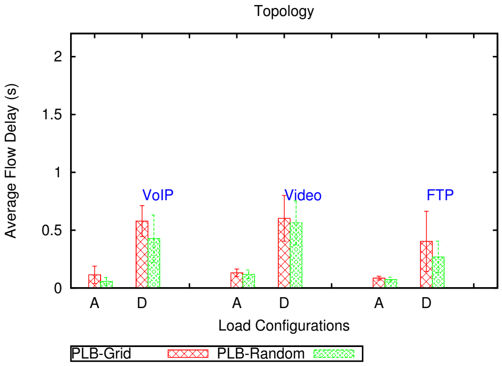

Figure 15. Average Flow Delay of PLB

4.2.2. Simulation Results

Figures 12 to 15 show that the topology type does not have a significant effect on the triple play service. Nevertheless, there are some cases where the traffic performance slightly increases or decreases in a random topology that depends on the inter-cluster routing approach. For example, RAILoB achieves a higher improvement of throughput than PLB for FTP application in low loads, FTP achieves 962,70 Kb/s and 1023,75 Kb/s for grid and random topologies respectively when RAILoB is used, whereas FTP achieves 337,55 Kb/s and 362,93 Kb/s for grid and random topologies respectively, when PLB is used. The reason for this is that the random topology can have a varied number of border nodes for the mesh router migration method (i.e. PLB), including no single border node, since the node placement is not regular. This means that the traffic performance can be affected by slow and inflexible load balancing approaches in this specific case. Nonetheless, RAILoB results in the best traffic performance for most cases, as well as for both of the topology types.

5. Conclusions and Comments on Future Work

In this article, we have outlined an architecture of cooperative routing management called ACRoMa, which is mainly concerned with scalability for triple play services. This proposed architecture is able to handle the scalability issue arising from the most relevant routing approaches by combining a cross-layer routing metric, which proved to improve the performance of the triple play service when making the routing decision, and an inter-cluster routing method for load balancing that enables the traffic migration to occur between multiple gateways in a more efficient way. In other words, ACRoMa integrates solutions to provide traffic scalability in a collaborative fashion for WMN, which are the cross-layer design metrics, clustering scheme and load balancing routing. Furthermore, it was evidenced by the results that the chosen approaches for each solution and their synergies result in a better scalability for triple play services than the concorrent approaches. For instance, ACRoMa speeds up the load balancing procedure by employing a proactive and flexible strategy of traffic migration for inter-cluster routing decision-making. Hence, it was confirmed that the proposed architecture lays down mechanisms that provide scalable triple play service in WMN with multiple gateways. As a means of advancing this research, it is expected that ACRoMa will extend by integrating a cognitive radio solution.

Acknowledgement

This work was partially funded by the Portuguese Ministry of Science (scholarship contract SFRH/BD/44378/2008), supported by the iCIS project (CENTRO-07-ST24-FEDER-002003), co-financed by QREN, in the scope of the Mais Centro Program and European Union's FEDER.

References

Akyldiz, I. F., Wang, X., and Wang, W. (2005). Wireless mesh networks: a survey. *Computer Networks*, 47(4):445–487.

Ashraf, U., Abdellatif, S., and Juanole, G. (2011). Route maintenance in ieee 802.11 wireless mesh networks. *Computer Communications*, 34(13):1604 – 1621.

Azcorra, A., Banniza, T., Chieng, D., Fitzpatrick, J., Von-Hugo, D., Natkaniec, M., Robitzsch, S., and Zdarsky, F. (2009). Supporting carrier grade services over wireless mesh networks: the approach of the european fp-7 strep carmen. *Communications Magazine, IEEE*, 47(4):14–16.

Bejerano, Y., Han, S.-J., and Kumar, A. (2007). Efficient load-balancing routing for wireless mesh networks. *Computer Networks*, 51(10):2450–2466.

Borges, V. C. M., Curado, M., and Monteiro, E. (2010). A Cross-Layer Routing Scheme for Scalable Triple Play Service in Wireless Mesh Networks. In *Proceedings of the 19th IEEE ICCCN 2010, Zürich, Switzerland, August 2-5, 2010*, pages 1–6.

Borges, V. C. M., Curado, M., and Monteiro, E. (2011). Cross-Layer Routing Metrics for Mesh Networks: Current Status and Research Directions. *Computer Communications Elsevier*, 34(6):681 – 703.

Borges, V. C. M., Dimitrov, E., Curado, M., and Monteiro, E. (2012). Performance Assessment of Cluster Load Balancing Routing Methods for Triple Play Services in Wireless Mesh Networks. In *Proceedings of the 13th IEEE/IFIP NOMS 2012*, pages 974–980.

Choi, H.-G. and Han, S.-J. (2010). Domain load balancing routing for multi-gateway wireless mesh networks. *Journal of Wireless Networks. Springer Kluwer Academic Publishers*, 16:2105–2122.

Dai, H. and Han, R. (2003). A node-centric load balancing algorithm for wireless sensor networks. In *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, volume 1, pages 548 – 552 Vol.1.

Ekling, J., Gidlund, M., and Flodin, R. (2007). Performance of triple play services in wireless meshed networks. ieee isce. pages 1 –5.

Gálvez, J. J. and Ruiz, P. M. (2013). Efficient Rate allocation, Routing and Channel Assignment in Wireless Mesh Networks supporting Dynamic Traffic Flows. *Ad Hoc Networks*, 11(6):1765–1781.

Genetzakis, M. and Siris, V. (2008). A contention-aware routing metric for multi-rate multi-radio mesh networks. In *5th IEEE SECON '08*, pages 242–250.

He, B., Sun, D., and Agrawal, D. (2009). Diffusion based distributed internet gateway load balancing in a wireless mesh network. In *GLOBECOM 2009. IEEE*, pages 1 –6.

Hsiao, P.-H., Hwang, A., Kung, H., and Vlah, D. (2001). Load-balancing routing for wireless access networks. In *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 986 –995 vol.2.

Kim, H., Claffy, K., Fomenkov, M., Barman, D., Faloutsos, M., and Lee, K. (2008). Internet traffic classification demystified: myths, caveats, and the best practices. In *CONEXT '08: Proceedings of the 2008 ACM CoNEXT Conference*, pages 1–12, New York, NY, USA. ACM.

Langar, R., Bouabdallah, N., and Boutaba, R. (2009). Mobility-aware clustering algorithms with interference constraints in wireless mesh networks. *Comput. Netw.*, 53(1):25–44.

Manikantan Shila, D. and Anjali, T. (2008). Load aware traffic engineering for mesh networks. *Comput. Commun.*, 31(7):1460–1469.

Nguyen, L. T., Beuran, R., and Shinoda, Y. (2008). A load-aware routing metric for wireless mesh networks. In *Computers and Communications, 2008. ISCC 2008. IEEE Symposium on*, pages 429–435.

Quintero, A., Elalamy, Y., and Pierre, S. (2004). Performance evaluation of a broadband wireless access system subjected to heavy load. *Computer Communications*, 27(9):781 – 791.

Ros, F. J. and Ruiz, P. M. (2007). Cluster-based olsr extensions to reduce control overhead in mobile ad hoc networks. In *Proceedings of IWCMC*, pages 202–207, New York, NY, USA. ACM.

Wu, D., Yang, S.-H., Bao, L., and Liu, C. H. (2014). Joint Multi-radio Multi-channel Assignment, Scheduling, and Routing in Wireless Mesh Networks. *Journal of Wireless Networks (Springer)*, 20(1):11 – 24.

Xie, B., Yu, Y., Kumar, A., and Agrawal, D. P. (2008). Load-balanced mesh router migration for wireless mesh networks. *Elsevier Journal of Parallel and Distributed Computing*, 68(6):825 – 839.

Zhu, Y., Ma, Q., Bisdikian, C., and Ying, C. (2011). User-centric management of wireless lans. *Network and Service Management, IEEE Transactions on*, 8(3):165–175.