On the Exploitation of User Aggregation Strategies in Heterogeneous Wireless Networks

Leonardo Badia*†, Nicola Bui*, Marco Miozzo*, Michele Rossi*‡, Michele Zorzi*‡,
* Consorzio Ferrara Ricerche (CFR), via Saragat 1, 44100 Ferrara, Italy
† IMT Lucca Institute for Advanced Studies, via San Micheletto 3, 55100 Lucca, Italy
‡ University of Padova, Dept. of Engineering (DEI), via Gradenigo 6B, 35131 Padova, Italy
E-mail: `{lbadia, nbui, mmiozzo, mrossi, zorzi}@ing.unife.it`

Abstract— **In this paper we discuss the exploitation of aggregated mobility patterns and physical proximity of nodes in a so-called ambient network, i.e., a wireless network with heterogeneous nodes and access techniques. We advocate to use the knowledge about node movements and geographical positions to create routing groups of adjacent nodes, which might be beneficial in order to decrease signaling overhead and increase transmission efficiency. Basically, routing groups (RGs) consist of aggregated logical structures which are built and maintained at the application layer. Their aim is to decrease the signaling overhead between group of nodes and access points and, at the same time, to improve connectivity by exploiting technology diversity and relaying techniques. On this matter, we describe a validation through simulation of a previously developed analytical work which is useful to evaluate the effectiveness of RG structures. Finally, we show the validity of the RG approach in terms of throughput and connectivity performance.**

I. INTRODUCTION

IN future generation networks, a strong synergy of heterogeneous radio resources is expected. This is true in particular for the novel concept of Ambient Networks [1], where environmental information, as well as the co-existence of different radio access technology and network management entities, are exploited to achieve a high level of integration. In particular, in this paper we address the routing problem in strongly differentiated mobility scenarios, where the geographical information about physical proximity and correlated node mobility patterns can be exploited by network protocols, similarly to what discussed in [2].

We discuss the creation of aggregated structures that we name Routing Groups (RGs), which are composed of nodes that have similar mobility patterns and other mobility-related parameters. In fact, the aggregation operation might be beneficial in taking advantage of the existing mobility structure and in improving the efficiency in transmitting data and/or handling network related procedures such as the handover between different access points (APs). As an example, multiple users moving together and handing over at the same time between the same pair of access points, may be aggregated in a routing group so that a single message (to the RG leader) needs to be exchanged to successfully accomplish the handover procedure, instead of using a dedicated transmission for every terminal. In general, this is true every time the information can be shared among users, that is, for all applications where some sort of multicast messaging is required or can be supported. In other cases, we may detect similarities in user characteristics and "aggregate" them in order to increase the transmission efficiency [3]–[5].

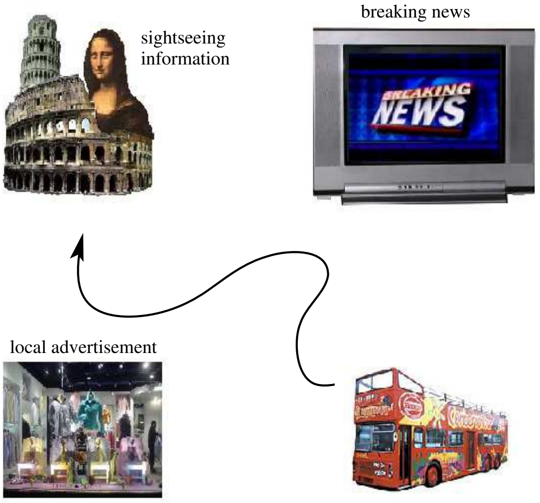

Fig. 1. A moving network receiving different context-related information.

As a further example, consider a shuttle moving into a city (see Fig. 1). In such a scenario it would potentially be more efficient to elect a RG leader such as the on-board shuttle multimedia system, which would then transmit the information related to, e.g., nearby tourist attractions, route information, TV programs, to all routing group members on board, in a multicast fashion. Hence, the shuttle will retrieve the wanted information from the external network through dedicated access points, and the information will be more efficiently distributed to the users in the shuttle by exploiting their inherent physical proximity. This simple example illustrates the opportunities and advantages offered by grouping network entities when they exhibit a group mobility behavior. More complicated reference cases can be thought of, where we may exchange contextual information (related to the user mobility behavior) among users and exploit it to improve network protocol performance [6].

Such an application-oriented information exchange also reflects on the lower layers. In fact, one of the most important advantages of aggregating users consists of the reduction of interfering and colliding messages exchanged by neighboring

0-7803-9536-0/06/$20.00 ©2006 IEEE.

nodes. This is possible as RGs can be exploited to increase the degree of coordination in the data transmission, i.e., by properly scheduling the transmissions to and from the RG leader. On the other hand, these benefits do not come for free, as the aggregated structures also need signaling [7] to be realized and maintained, thus increasing the message exchange of the network.

For these reasons, a quantitative evaluation of whether grouping network terminals in aggregated structures is beneficial, and how much, is a very important subject for heterogeneous network management. In a previous contribution [8] we presented an analytical study about the effectiveness of the grouping approach, which was investigated as a function of various parameters including the density of RGs, the number and type of radio technologies and the related energy consumption. In particular, such a model makes it possible to quantify the aggregation benefits and weigh them against the costs incurred in creating and maintaining RG structures. The goal of the present work is to validate these results through a simulation approach, which aims at modeling a realistic scenario in order to relax all the restrictive assumptions made about physical propagation and radio access technology details. To this end, we first present a thorough description of the simulator. Subsequently, we report numerical results to validate the analysis.

The results reported in this paper show good agreement between theory and simulation, in spite of the simplifications introduced in the analytical framework. Quantitative differences are present, but they will be discussed and explained in light of physical effects which are accounted for in the simulator and that, for complexity reasons, are difficult to model exactly by analysis.

The rest of the paper is organized as follows: in Section II we show the key aspects of the simulator, by giving a detailed description of its structure, the design philosophy and its capabilities. Section III summarizes the analysis that we intend to validate by simulation considering a detailed and realistic environment. In Section IV we present numerical results about this evaluation and in Section V we draw our conclusions.

II. DESCRIPTION OF THE SIMULATION ENVIRONMENT

This section gives a detailed description of our event-driven network simulator for heterogeneous wireless system called ANEMURAS (Ambient NEtworks MUlti-Radio Architecture Simulator). This tool has been developed within the Ambient Networks project [1]. It has been specifically designed to model a multi-technology mobile and wireless communication scenario, where both mobile users and fixed access points (APs) coexist and communicate through the wireless medium. Node mobility, wireless channel variability and inter-user/inter-system interference are explicitly and accurately accounted for. In Fig. 2 we report the structure of the communication node implemented in the simulator. This node may be either mobile or static, and behave as a user or an AP. We stress that the main goal in designing the simulation structure was to derive a truly modular and scalable tool, where physical layer and upper layers are fully decoupled and interchangeable. This allows, for example, to exploit the same physical channel module for different wireless technologies while leaving all the calculations related to timing, power levels, interference and synchronization to the physical layer. As a consequence, the design of a multi-technology platform becomes very easy as the addition of a new technology requires to simply add the technology dependent aspects in additional modules, which can easily be interconnected with the rest of the simulator. The simulator already implements several radio technologies and supports per packet scheduling over them. In the following subsections, we examine in some detail the node structure by providing a short description for each of its parts.

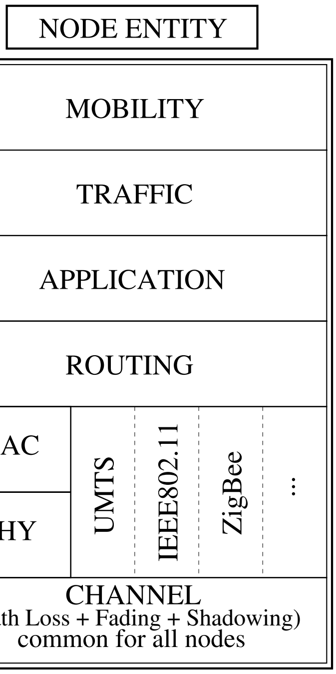

Fig. 2. Communication node structure.

A. Topology

Access points can be either deterministically or randomly placed at the beginning of the simulation and can be either static or on the move. The same applies for mobile users. The only difference between APs and user terminals is that APs may have a high speed path connecting them. That is, we may assume them to be connected to a wired backbone. As an example, we exploit this in the simulator for the UMTS system, where base stations (BSs) exploit wired links to exchange the information needed to maintain synchronization, handle handover and re-initialize link layer mechanisms as a user connection is switched to a new access point.

B. Channel Module

The channel is modeled accounting for path loss, shadowing and multi-path fading phenomena and using their product as the link gain which is subsequently associated with each transmission link (a transmission link exists between each pair of nodes in the simulation). Path loss is implemented according to the well known Hata model [9]. Specifically, if P_{tx} is the transmitted power, the path loss (in linear units) is calculated as $P_{tx}/P_{rx} = Kd^{\beta}$, where d is the distance separating the two communicating entities, whereas K and β are proper constants. Shadowing is accounted for according to the Gudmunson model [10] and multi-path fading is implemented for each link through a Jakes simulator with programmable number of oscillators [11]. The channel module is only responsible for the calculation of link gains as the simulation time evolves, whereas a further entity, called *physical module*, processes the channel gain matrix and the user transmission powers in order to derive intra- and inter-system interference metrics. Further details about this are reported below.

C. Physical Module

The physical layer entity takes as input the channel gain matrix created and maintained by the channel module and the transmission powers selected by each user and returns Signal to Interference plus Noise Ratio (SINR) metrics for each active link. Let us clarify this with an example by referring to the SINR calculation for a generic non-CDMA system. Let $\mathcal{N}$ be the set of nodes in the network and consider a specific user j, which is interested in receiving a given data flow from a specific user i. The SINR at user j at the generic time t is calculated as:

$$\gamma_{ij}(t) = \frac{P^{i}_{tx}(t) g_{ij}(t)}{\sum_{k \in \mathcal{N}, k \notin \{i,j\}} \alpha_{kj} P^{k}_{tx}(t) g_{kj}(t) + N_o B} \qquad (1)$$

where $\gamma_{ij}(t)$ is the SINR experienced by user j at time t and associated with the communication $i \rightarrow j$, $g_{kj}(t)$, $k, j \in \mathcal{N}, k \notin \{i,j\}$ is the link gain associated with the wireless link connecting node k to node j, $P^{i}_{tx}(t)$ is the transmission power used at time t by user i, N_o is the white noise power spectral density (psd) and B the transmission bandwidth. α_{kj} is a parameter that we use, at user j, to model the fraction of the received power from user k which interferes with the useful transmission $i \rightarrow j$. For instance, in UMTS systems this parameter corresponds to the commonly used orthogonality factor, which is usually applied to the interfering users for the same serving BS. However, α_{kj} can also be used to model the *inter-technology interference*. Consider, for instance, that the technology exploited for the transmission over link $i \rightarrow j$ is different from the technology selected by user k. In this case, α_{kj} has to be understood as the fraction of power transmitted by user k that interferes with the transmission over link $i \rightarrow j$ when the two transmissions exploit different radio technologies.

The current version of the simulator implements physical layer modules for IEEE802.11b, IEEE802.15.4 (ZigBee) and the UMTS radio levels [12], [13]. However, due to the modular structure of the simulator, new technologies can be easily added. For the UMTS standard, we fully implemented uplink dedicated channels (DCH), including power control features and Radio Link Control (RLC) mechanisms, as specified in the 3GPP standard [12]. To this end, we also implemented special nodes acting as base stations (node Bs in the UMTS terminology). Finally, the UMTS physical level calculates intra- and inter-cell interference with an equation similar to Eq. (1), where we additionally account for the spreading gains which are typical of the wideband CDMA (W-CDMA) technique used in UMTS [14]. The errors on the transmitted data streams can be tracked at the bit level and coding can also be accounted for through pre-computed coding gain curves.

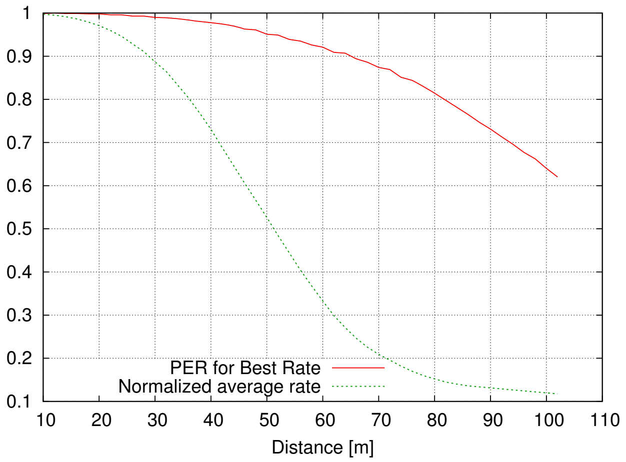

Fig. 3. Implemented mechanism of IEEE802.11 rate selection: signaling rate, normalized to 11 Mbps (dashed line) and error probability obtained with it (solid line) as functions of the distance. These curves have been derived by following the approach of [15].

The physical modules of IEEE802.11b and IEEE802.15.4 account for channel interference using Eq. (1) and tracking SINR values at the symbol level. These are subsequently translated to bit errors according to the selected transmission mode (bit rate and modulation). The rate selection mechanism has been implemented by considering the analytical expressions presented in [15]. In Fig. 3 we plot both the normalized rate selected for the connection as a function of the distance and the error probability obtained by using the best physical rate among the ones available according to the standard. Both the IEEE802.11b and the IEEE802.15.4 radio levels implement a receiving model which accounts for possible interfering transmissions during the reception of a packet. Therefore, simultaneous transmissions do not necessarily lead to a certain collision at the receiver. In this way, we can effectively account for the *capture effect* at the physical layer. Finally, in order to speed up the simulation time several optimizations have been implemented. In particular, we can set the frequency at which channel and data errors are tracked, which means that we can update the error probability for every bit in the data sequence or use the same error rate for several subsequent bits. In this last case, both channel gains and SINR values are updated less frequently according to a pre-selected *granularity* parameter. If correctly configured, this procedure can substantially reduce the simulation time without affecting the accuracy of the results.

D. MAC Level

The MAC level of IEEE802.11b implements the *Distributed Coordination Function* (DCF) technique as specified by the standard [13]. For what concerns the Collision Avoidance feature, both the *Basic Access Mode* and the *RTS-CTS* mode are available. For the IEEE802.15.4 technology we implemented the beaconless peer-to-peer mode. This mode allows to use this technology to transmit in the same way as IEEE802.11 does, i.e., every node can transmit after winning the contention for the shared channel. The MAC level of UMTS implements the fragment-unfragment function, while a selective repeat algorithm is implemented as the error recovery algorithm at the link layer.

E. Routing Level

This level plays an important role in the simulator as routing in heterogeneous networks is a new and interesting research topic. Currently, the level includes a variant of the *Dynamic Source Routing* (DSR) algorithm [16] with the capability of routing packets across different technologies. Additional routing algorithms are currently under study. The aim is to exploit the presence of multiple radio technologies to improve performance, and to include this aspect directly at the routing level according to a cross-layer approach. Moreover, the simulator allows to set Routing Groups which may influence both MAC and path selection procedures. In fact, under specific circumstances nodes may be forced to route their packets through Routing Group leaders, which might be selected among them according to various criteria [3].

F. Mobility Level

In the simulator we implemented both independent and group mobility behaviors, according to the model described in [17]. Every simulated entity can be static or mobile and, in this last case, can either move independently or in a group fashion [6]. While independent mobility behaviors are well described in the literature, it is useful to give here a short description of the group mobility feature: this type of mobility is realized superimposing an attraction force between a number of followers and a group leader. This force is basically added to the force governing independent mobility behaviors and, according to its strength, we can control the impact of the attraction towards the group leader. If this force is much stronger than the one related to the independent mobility component, the resulting trajectory will be similar to that of the leader. In the opposite case (the attraction force is weaker) the independent component dominates and the user trajectories tend to become uncorrelated. In addition, the mobility model is composed by a wide range of parameters which can be set to simulate very different mobility patterns, ranging from people walking in a square (very uncorrelated mobility) to a military parade (correlated trajectories).

G. Traffic Level

This level is in charge of generating data traffic. Currently, both Poisson, continuous and periodic packet generations are implemented. Future extensions include bursty and Pareto shaped traffic patterns.

III. ANALYTICAL MODEL

In this section we briefly outline the analytical framework adopted in [8] and its key assumptions. For brevity, only an outline of the obtained results, without proofs or detailed calculations, is reported here. Interested readers are referred to [8] for further details about these points.

A. General assumptions

We consider an ambient network where a number of access points (APs) and a number of users coexist. We assume that each user requires a separate flow and all flows have the same bit-rate B_U.[1] Both APs and users support a number of different radio technologies which can be described by the indices $1, 2, \ldots, J$, where technologies are indicated with integer numbers and sorted according to the required transmission energies. That is, $E_i^{tx} \leq E_j^{tx}$ if $i < j$. Accordingly, we define three further vectors $\mathbf{E}^{tx} = \{E_1^{tx}, E_2^{tx}, \ldots, E_J^{tx}\}$, $\mathbf{E}^{rx} = \{E_1^{rx}, E_2^{rx}, \ldots, E_J^{rx}\}$ and $\mathbf{r} = \{r_1, r_2, \ldots, r_J\}$ tracking the energies required to transmit and receive a single bit and the maximum transmission ranges for every technology, respectively. Not all nodes offer all radio interfaces and, in general, the set of offered interfaces may differ in different nodes. Here, we assume that a generic node has an interface of type i with a given probability p_i and that interfaces are assigned independently to each user in the network. For the topology, we consider that users are independently placed according to a planar Poisson process of intensity ρ, i.e., the average number of users within an area $\mathcal{A}$ is given by $\rho\mathcal{A}$, whereas the probability to have exactly n nodes in this area is derived as $\mathcal{P}(n, \mathcal{A}) = ((\rho\mathcal{A})^n/n!) \exp(-\rho\mathcal{A})$. We further consider that access points are placed according to a uniform distribution with density ρ_{AP} and are equipped with all the technologies present in the network.

Within the analysis, all nodes are always considered to be stationary. Moreover, only slow fading is considered, which is appropriate for stationary networks; nevertheless, adding the nodes' movement in the simulator will also relax this hypothesis. Finally, we consider an idealized MAC by neglecting interference aspects. In fact, in the analysis our only interest is to capture the network connectivity, i.e., the availability of a path between any two nodes, rather than investigating the performance of simultaneous parallel transmissions.

For this reason, we model the physical layer as follows. We assume that every transceiver device has the same receiver sensitivity which depends on the considered radio interface $i \in \{1, 2, \ldots, J\}$. We assume that packets are correctly decoded when the received power is above the respective technology-dependent sensitivity. The propagation loss $L(d)$ (in decibel) at a distance d is given by $L(d) = K_0 + K_1 \ln d + s$, where K_0 and K_1 are proper constants, while s is a shadowing sample which is assumed to be log-normally distributed with zero mean and standard deviation σ_{shad}. Thus, the received power (decibel) at the generic interface i of a given node is $P_{rx,i} = P_{tx,i} - L(d)$, where d is the distance between the source (S) and the node itself and $P_{tx,i}$ is the power used by

---

[1] The extension to different bit-rates are possible with minimal modifications to the approach.

S to transmit. Hence, the vector $\mathbf{r}$ may be found depending on transmission power levels and radio sensitivities according to the propagation model presented above and to the quality requirements. That is, we can translate the requirements on received powers (minimum signal quality) into maximum transmission distances. The analytical formulation that we present in the sequel will make direct use of the vector $\mathbf{r}$. Given the network topology and the radio interface models, we can easily find the density ρ_i of nodes with an interface of type i. Formally:

$$\rho_i = \frac{\sum_{n=0}^{\infty} \mathcal{P}(n, \mathcal{A}) \sum_{k=0}^{n} k p_i(k|n)}{\mathcal{A}}, \tag{2}$$

where $\mathcal{A} \in \mathbb{R}^+$, $p_i(k|n) = \binom{n}{k} p_i^k (1 - p_i)^{n-k}$.

B. RG formation

Routing groups can be formed exploiting a distributed approach. That is, users cooperate and exchange data in order to gain information about their physical proximity and, at the same time, to measure the worthiness of grouping with other network entities. This involves the periodic exchange of the so called HELLO messages between mobile nodes [18]. In each HELLO, each node can for example include the list of its “stable neighbors”, i.e., the nodes that have been in its close proximity for a long enough period of time. In fact, if movements are correlated, stable nodes are likely going to stay in close proximity of the sending device and are therefore good candidates to be grouped with it.

The goal of the present paper is to keep the evaluation of algorithms for RG formation as general as possible, thus we do not investigate a specific aggregation strategy but we simply assume that the RG formation is possible and can either be activated or not, determining different network behaviors. However, we claim that every terminal aggregation strategy can be framed in our evaluation. The objective of our investigation is to quantify the potential benefits in terms of connectivity which are offered by having RGs in place and weigh them against the energy expenditure required to maintain RGs. We further assume that a leader is elected within each RG; this device has the special role of handling the data traffic so as to optimize the transmission and the access to the channel for RG members. This can be seen, as in standard clustering algorithms [4], [5] for ad hoc networks, as a way to partially centralize the transmission control thereby enhancing the performance. In order to abstract from the specific clustering techniques that might be used to form RGs, and also from the network topology, we adopt a randomized approach for the RG leader selection, i.e., we simply elect RG leaders arbitrarily with probability p_L. This holds both for the analysis and for the simulation results that will be shown later in this paper. After the RG leader selection, the group is assumed to be created by neighboring nodes which join the leader. Note that this happens in an ideal way in the analysis (nodes always join their closest RG leader), whereas in the simulation we actually implement the HELLO message exchange (it might happen that a HELLO packet is lost and, in turn, a node may not detect all terminals in its physical proximity).

We assume that every interface $i \in \{1, 2, \ldots, J\}$ sends HELLO messages with an interface-specific period T_i^h and we refer to b_i^h as the number of bits composing HELLO packets sent by an interface of type i. Moreover, we consider that all T_i^hs are multiple of a reference time period ΔT such that $T_i^h = \xi_i \Delta T$, $\xi_i \in \mathbb{N}^+$, $i \in \{1, 2, \ldots, J\}$, $\Delta T \in \mathbb{R}^+$. If we define the least common multiple (LCM) of all ξ^hs as ξ, then we have that $n_i^h = \frac{\xi}{\xi_i}$ is the number of HELLOs sent by the i-th interface in a time period equal to $\xi \Delta T$.

C. Connectivity evaluation

We evaluate the average distance between two APs as $\overline{d}_{AP} = 1/(2\sqrt{\rho_{AP}})$. Hence, on average each AP is in charge of delivering data to all users placed within a circle of radius $\overline{r}_{AP} = \overline{d}_{AP}/2$ (we neglect the overlap in circular regions, i.e., the terminals which can be identically covered by more than one AP by arbitrarily assigning them to one AP).

The density of nodes with technology j is still given by ρ_j as derived in Eq. (2). The approach to evaluating the network coverage proceeds by slicing the area covered by an AP into J different annuli, where the jth annulus is of area $\mathcal{A}_j = \pi(r_j^2 - r_{j-1}^2)$ and contains the region covered by technology j but not by any of the technologies with lower index (i.e., with lower coverage radius). The average number of users that have to be reached in the jth annulus, $\overline{n}_j$, is found according to $\overline{n}_j = \rho \mathcal{A}_j$. The average number of users $\overline{n}_{j,h}$ in the jth annulus that can be optimally covered by exploiting interface h is therefore found as:

$$\overline{n}_{j,h} = \begin{cases} \rho \tilde{p}_{hj} \mathcal{A}_j & h \geq j \\ 0 & h < j \,, \end{cases} \tag{3}$$

where $\tilde{p}_{hj} = p_h[\prod_{j \leq \ell < h}(1 - p_\ell)]$, and p_h, p_ℓ are the probabilities for a generic user of having interface of type h and ℓ, respectively.

The average number of users which can not be reached by any technology depends on whether the RG formation is active or not. In the case RGs are present, in fact, the coverage radius required is smaller, so that the coverage generally increases. However, an additional connection between the AP and the RG leader is required. The analytical evaluation of these two quantities leads to [8]:

$$\overline{n}_u = \pi r_{AP}^2 \rho - \sum_{j=1}^{J} \sum_{h=1}^{J} \overline{n}_{j,h} \,, \tag{4}$$

for what concerns the case where RG are inactive, and:

$$\overline{n}_u = \pi r_{RG}^2 \rho - \left(1 - \exp\left(-\rho_{AP} \pi r_J^2\right)\right) \sum_{j=1}^{J} \sum_{h=1}^{J} \overline{n}_{j,h} \,, \tag{5}$$

when they are active. In both cases, the number of uncovered users per unit area is evaluated as $\overline{n}_u/(\pi \overline{r}_{AP}^2)$.

D. Energy consumption evaluation

To evaluate the energy spent for covering the network, we compare the normalized values per unit area. It is possible to obtain the total value multiplying by the value $\pi\overline{r}_{AP}^2$ where $\overline{r}_{AP} = 1/(4\sqrt{\rho_{AP}})$. In the following, the asterisk indicates normalization to the unit area.

We obtained in [8] the following expression for the case where RGs are not activated:

$$\overline{E}^*_{noRG} = \frac{\sum_{h=1}^{J}\sum_{j=1}^{J}\overline{n}_{j,h}(E_j^{tx}+E_j^{rx})B_U}{\pi\overline{r}_{AP}^2}\,. \qquad (6)$$

For the case where RGs are present, we have instead that the total energy expenditure consists of three terms:

$$\overline{E}^*_{RG} = \overline{E}^*_{(a)RG} + \overline{E}^*_{(b)RG} + \overline{E}^*_{(m)RG}\,, \qquad (7)$$

where $\overline{E}^*_{(a)RG}$ is the energy required to transmit the flow from the AP to the RG leader, $\overline{E}^*_{(b)RG}$ is required to deliver the flow from the leader to the nodes, and finally $\overline{E}^*_{(m)RG}$ is the energy to maintain the RG structure.

The normalized value $\overline{E}^*_{(b)RG}$ is identical to $\overline{E}^*_{noRG}$ (the difference is in the fact that the area of a RG is usually smaller than the one covered by an AP, but their evaluation is identical). Thus, with respect to the case where RGs are not present, the energy expenditure has two more terms.

The first one is:

$$\overline{E}^*_{(a)RG} = \frac{(E_1^{tx}+E_1^{rx})B_U}{\pi\overline{r}_{AP}^2}\,, \qquad (8)$$

which corresponds to the energy consumed to transfer the control to the RG leaders. The energy spent to *maintain* the RG structures over an area $\mathcal{A}$ in a time period of $\xi\Delta T$ seconds can instead be derived as:

$$\overline{E}_{(m)RG}(\mathcal{A}, T=\xi\Delta T) = \sum_{j=1}^{J}\sum_{n=1}^{\infty}\mathcal{P}(n,\mathcal{A}) \times \times \sum_{k=1}^{n}\left\{kp_j(k|n)b_jH_j[E_j^{tx}+E_j^{rx}\varepsilon_j]\right\}, \qquad (9)$$

where ε_j is the mean number of nodes receiving the HELLO message sent by a given sending node and using interface j (assuming that HELLO packets are only decoded by the node neighbors whose distance is less than or equal to $\overline{r}_{RG}$), i.e.:

$$\varepsilon_j = \sum_{n=2}^{\infty}\mathcal{P}(n,\pi\min(r_j,\overline{r}_{RG})^2)\sum_{k=1}^{n}(k-1)p_j(k|n)\,. \qquad (10)$$

The normalized value of the cost to create and maintain RG structures is therefore derived as:

$$\overline{E}^*_{(m)RG} = \frac{\overline{E}_{RG}(\mathcal{A},\xi\Delta T)}{\mathcal{A}\xi\Delta T}\,. \qquad (11)$$

IV. VALIDATION OF THE ANALYSIS

In this section, we present accurate simulation results aimed at validating and further investigating some of the above facts.

A. Validation scenario

We consider a network scenario composed by two radio access technologies: IEEE802.11b and UMTS. User devices move within a simulation area of 160×160 m^2, with speeds uniformly distributed in the range $[0.5, 2]$ m/s, so as to mimic a typical pedestrian scenario. The density ρ of the mobile nodes is chosen in $[0.001, 0.01]$. Mobility patterns are generated according to a random way point mobility model. We consider a single AP, placed at the center of the simulation area and owning both technologies. Exactly 20% of the mobile devices own both wireless technologies, whereas the remaining 80% of the population is equipped with the IEEE802.11b technology only. As above, we consider two different access strategies: *with* and *without* RGs. In the former case (RGs), each user can access the AP either relaying its data to an in range RG leader or by direct transmission to an in range AP. Moreover, each node decides between the two previous options by minimizing the number of hops required to deliver its data to the selected access point. In case the user has to decide between two alternative paths with the same number of hops the choice is driven by the available bandwidth. RG leaders are the only nodes which are allowed to aggregate traffic and, act as relays, by therefore providing coverage extension through multi-hop routing. RG leaders are elected at random at the beginning of the simulation with probability p_L and among the users having both technologies. We stress again that this is an artifice to keep our investigations as general as possible. General RG creation strategies can in fact be included in this framework by using the corresponding value for p_L. Note that, in the no RG case relaying is not permitted and a mobile device is connected to the AP if and only if the AP is directly reachable through at least one of the radio technologies owned by the user. On the other hand, in the presence of RGs, a user which does not have a direct connection with any AP can still exploit the relaying functionalities of an in range RG leader, if present. As will be shown next, this improves connectivity performance. Finally, the UMTS network covers the whole simulation area, whereas the IEEE802.11b technology roughly covers half of the simulation area. All users generate uplink traffic (i.e., users → AP) at the rate of one packet per second. Packets are 512 bytes long. Users' traffic is exploited, in part, for the establishment and maintenance of the routes to get to the AP. To this end, we modified the DSR protocol in such a way that only RG leaders and APs can relay data traffic.

B. Results

Next, we report accurate simulation results obtained with the multi-technology simulator presented in section II. In Fig. 4 we plot the density of unconnected users for both scenarios (with and without RGs). As expected, and in accordance with the analysis, the case without RGs gives the worst performance in terms of connectivity. However, note that the theoretical gain is only partially exploited. In fact, while analytical curves show a probability of being disconnected which is around 5 times lower in the case with RGs (even though it is not shown, a similar gain is observed for $p_L = 0.05$), simulation results show a higher amount of disconnected nodes. This is due to a

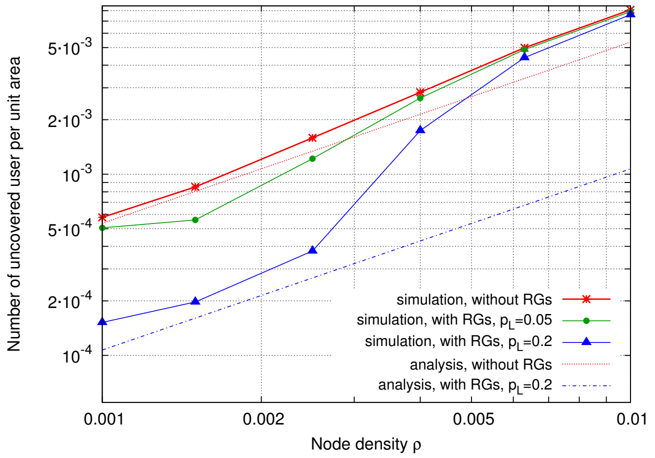

Fig. 4. Average number of uncovered users per unit area as a function of the node density ρ.

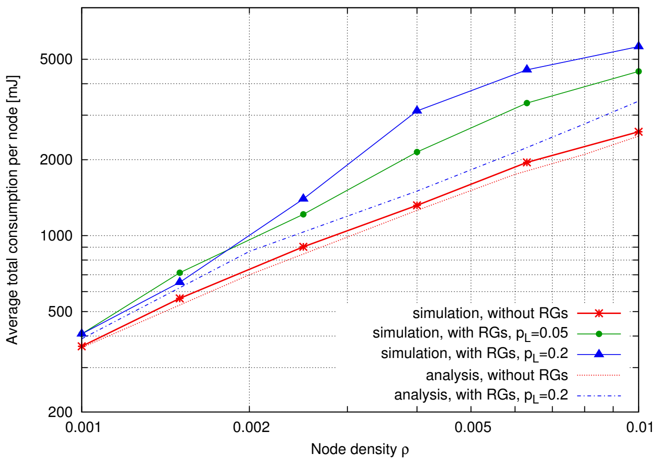

Fig. 5. Average energy consumption per node as a function of the node density ρ.

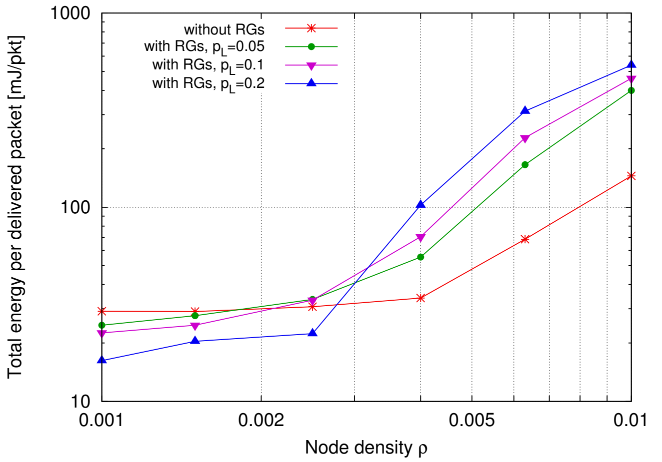

Fig. 6. Energy consumed to delivery a packet as a function of the node density ρ.

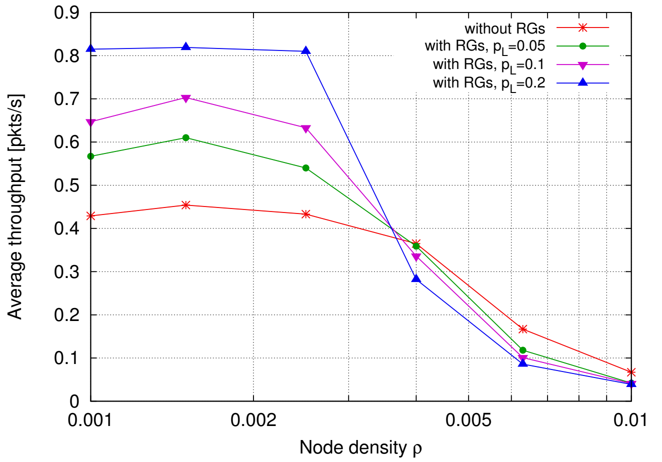

Fig. 7. Network throughput as a function of the node density ρ.

more precise modeling of the propagation scenario, which no longer consists of circular areas.

As ρ increases, the performance of the RG case saturates to the scenario without RGs. This is basically due to the following two facts: 1) the capacity of the AP is limited, and 2) an increasing ρ leads to an increasing user interference that, in turn, limits the maximum number of communicating users that can be simultaneously supported by the system. Note also that the probability of being disconnected is always *higher* for the simulations than for the analytical results. For $\rho \approx 0.01$ the gains offered by RGs are almost vanished, since at this point the number of simultaneous transmissions overcomes the maximum network capacity. The interference here is substantial and the RG solution can not help any longer due to the limited transmission capacity of both APs and RG leaders. In other words, the interference at this point becomes the dominant factor, which dominates over the adopted relaying strategy. As a side remark, this operating region should be avoided as performance is dominated by interference and is usually too bad to be considered acceptable. However, apart from the quantitative evaluation, the qualitative behavior is similar and also the simulation evaluations confirm the benefit of RGs in terms of increased network coverage. This is true especially at low densities, which is also the case where the coverage provisioning might be more problematic.

Fig. 5 analyzes the energy expenditure of the nodes. The theoretical results are satisfactorily confirmed, though there is always an additional term due to the interference between parallel transmissions, which is not considered in the analysis. Also, in the RG case not only is the consumption higher in absolute terms, but the increase due to taking interference into consideration is also higher, because in the RG case there are more transmitting sources and therefore more interference.

The last two figures show instead how the investigation through simulation results can be useful to analyze in more depth some aspects which were *not captured* by the analysis as they were not modeled. In Figs. 6 and 7, we report the energy expenditure per correctly delivered packet and the average throughput per node, respectively. It is important to observe that for small to moderate densities ($\rho \leq 0.003$ in the reported graphs) the RG solution leads to better performance in terms of both energy expenditure and throughput. For these densities the aggregation of terminals in close proximity and the exploitation of RG leaders as the relay nodes towards the APs is actually a good strategy. In fact, this makes it possible to exploit the technology diversity offered by the presence of RG leaders and, at the same time, to better distribute the traffic among the available wireless technologies. However, for very dense networks (all terminals in our scenario are active and constantly transmitting) relaying is less useful as in such a case the high interference causes congestion and RG leaders become unable to forward data traffic. In this case relaying appears to be useless and the users should instead run rate control algorithms to decrease the network load.

Again, from Fig. 7 we can observe that the maximum improvement is reached for $p_L = 0.2$, where the throughput is almost doubled with respect to the no RG case. Clearly, this is due to the exploitation of RG leaders which provide coverage to nodes that would be otherwise disconnected by all APs. As said above, when the node density grows larger than 0.004 users per unit area (i.e., per square meter) the relaying feature causes too high an interference level between parallel transmissions. In other words, when the network density overcomes this value, the interference effect dominates over connectivity improvements.

C. Lessons learned

The obtained results verify the correctness of the analytical method, which well approaches the real case. Even though the analysis does not take into consideration some constraints or implementation aspects, the qualitative behavior is similar. However, the simulation approach is also useful as it is able to correctly model the interference among users, which would cause too high mathematical complexity. Interference issues also justify the disagreement between the results obtained with the two evaluation methods when the node density exceeds a certain level: simulation results show the congestion due to interference, while this is not reflected by the analytical framework. Finally, simulation allows a number of additional evaluations which are difficult to obtain by analysis, such as packet-based metrics (energy expenditure per correctly delivered packet).

V. CONCLUSIONS

This paper has presented several concepts related to the energy consumption connected with the Routing Group approach in Ambient Networks. We investigate whether the RG establishment is advantageous, in particular by exploring simulation techniques in order to validate an already existing analytical framework. The theoretical approach correctly evaluates the intrinsic trade-off in RG establishment. However, accurate quantitative evaluation of these concepts can be obtained only through a detailed investigation by means of simulation. In fact, computer-aided investigations make it possible us to account for interference phenomena, which in general lead to a less optimistic upper bound than provided by the analysis. Moreover, the simulation framework proposed in this paper provides additional features, which can be exploited to evaluate other metrics besides the simple network connectivity and overall energy consumption. In particular, the simulation tool proposed here can be used to compute per-packet metrics and, more in general, to shed new light on the implementation aspects in real network scenarios.

ACKNOWLEDGMENT

This work has been carried out in the framework of the Ambient Networks project that is partially funded by the Commission of the European Union. The views expressed in this paper are solely those of the authors and should not be interpreted as necessarily representing the views of their employers, this project or the European Commission.

REFERENCES

[1] The Ambient Networks project, `http://www.ambient-networks.org/`.

[2] C. K. Toh, "Associativity-based routing for ad-hoc mobile networks," *Kluwer Wireless Personal Communications*, vol. 4, pp. 103–139, 1997.

[3] P. Popovski, F.H.P. Fitzek, H. Yomo, T.K. Madsen, R. Prasad, and N.J. Vej, "MAC-layer approach for cluster-based aggregation in sensor networks," in *Proceedings of International Workshop on Wireless Ad-Hoc Networks (IWWAN)*, 2004, pp. 89–93.

[4] M. Chatterjee, S. K. Das, and D. Turgut, "WCA: a weighted clustering algorithm for mobile ad hoc networks," *Kluwer Cluster Computing*, vol. 5, pp. 193–204, 2002.

[5] S. Basagni, "Distributed and mobility-adaptive clustering for multimedia support in multi-hop wireless networks," in *Proceedings of IEEE VTC Fall*, vol. 2, 1999, pp. 889–893.

[6] M. Zonoozi and P. Dassanayake, "User mobility modeling and characterization of mobility patterns," *IEEE Journal on Selected Areas in Communications*, vol. 15, pp. 1239–1252, 1997.

[7] C. Prehofer and B. Souville, "Synchronized reconfiguration of a group of mobile nodes in ad-hoc networks," in *Proceedings IEEE ICT 2003*, vol. 1, 2003, pp. 400–405.

[8] M. Rossi, L. Badia, P. Giacon, and M. Zorzi, "On the effectiveness of logical device aggregation in multi-radio multi-hop networks," in *Proceedings IEEE MobiWac*, 2005, an extended version has been accepted for publication in *Wiley's Wireless Communications and Mobile Computing Journal*.

[9] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[10] G. L. Stüber, *Principles of Mobile Communication*. Kluwer Academic Publishers, 2003.

[11] W. C. Jakes, *Microwave Mobile Communications*. Wiley–IEEE Press, 1994.

[12] Third Generation Partnership Project, `http://www.3gpp.org/`.

[13] IEEE 802 LAN/MAN Standards Commitee documentation, `http://www.ieee802.org`.

[14] H. Kaaranen, A. Ahtiainen, L. Laitinien, S. Naghian, and V. Niemi, *UMTS Networks, Architecture, Mobility and Services*. John Wiley & Sons, 2001.

[15] V. Atanasovski and L. Gavrilovska, "Influence of header compression on link layer adaptation in IEEE 802.11b," in *Proceedings IWCMC*, vol. 2, 2005, pp. 1551–1556.

[16] D. B. Johnson, D. A. Maltz, and Y.-C. Hu, "The dynamic source routing protocol for mobile ad hoc networks (DSR)," 2003, available as Internet draft, `draft-ietf-manet-dsr-09.txt`.

[17] M. Rossi, N. Bui, L. Badia, and M. Zorzi, "On group mobility patterns and their exploitation to logically aggregate terminals in wireless networks," in *Proceedings IEEE VTC Fall*, Dallas, Texas, US, Sep 2005.

[18] B. McDonald and T. Znati, "Comparative analysis of neighbor greeting protocols: ARP versus ES-IS," in *Proceedings of IEEE SIMULATION*, 1996, pp. 71–80.