Characterizing User Behavior and Network Load on a Large-Scale Wireless Mesh Network

Michele Vincenzi[1], Roberto Tomasi[1], David Tacconi[1], Dzmitry Kliazovich[2], and Fabrizio Granelli[2]

[1] Futur3 srl,
via Antonio Abondi, 37, 38100 Trento, Italy
{m.vincenzi,r.tomasi,d.tacconi}@futur3.it
[2] DISI – University of Trento,
via Sommarive, 14, 38123 Trento, Italy
{kliazovich,granelli}@disi.unitn.it

Abstract. Wireless mesh networks represent a promising paradigm to provide a scalable infrastructure for Internet access in metropolitan areas. In this paper, a large-scale wireless mesh testbed deployed in three cities in the Trentino region is described and experimentation results obtained from the public use of the testbed are reported and analyzed. The large-scale of the deployment and high number of users ensure to have proper traces which can capture the trends in user traffic based on the applications used and realistic mobility patterns.

Keywords: mesh testbed, large-scale, wireless mesh networks.

1 Introduction

Wireless Mesh Networks (WMN) represent a hybrid solution between infrastructure and ad-hoc networking paradigms, where data forwarding is enabled by all the nodes in the network. Potentially, all the nodes act as hosts and as routers, forwarding packets generated by other nodes. Mesh networking offers several advantages: (i) it allows the combination of different wireless technologies, such as cellular networks, WiFi, WiMAX, etc.; (ii) WMNs can be incrementally deployed, in order to gradually extend connectivity and capacity, avoiding massive investments. Moreover, WMNs autonomously set up and maintain the connectivity: if a node fails, another route to the gateway is found through another path, improving robustness, resilience, preservation and providing self-healing properties.

WMNs provide a technological bridge between mobile ad hoc networks (MANETs) and traditional wireless local area networks (WLANs), such as the ones based on the IEEE 802.11 family of standards. A typical WMN consist of several nodes (routers and gateways) which exploit multi-hopping in order to build and maintain a wireless backhaul. WMNs enhance traditional star-shaped network architectures by providing increased robustness (e.g. no single points of failure are present and broken/congested links are encompassed), scalability and flexibility (without the need for deploying cables, connectivity may be provided only where and when

T. Magedanz et al. (Eds.): TridentCom 2010, LNICST 46, pp. 255–264, 2011.

needed/economically attractive), and incremental deployment. Moreover, WMNs can support heterogeneous transmission technologies. WMN currently represent a promising paradigm for cost-effective deployments in several metropolitan area scenarios, including community networks, digital divide affected areas, mobile internet infrastructure, etc. Such scenarios can fully capitalize on the scalability, incremental deployment and robustness of WMNs.

Nevertheless, as WMNs become a service infrastructure to deliver high end services, effective design is needed to provide the required levels of service to the plethora of applications demanded by the users. To this aim, some theoretical or mostly simulation studies are available, but only a few works address realistic user behaviour characterization (both in terms of mobility and connection, but also application preferences) and provide extended analyses of the network load on a WMN infrastructure.

In this framework, performance evaluation and testing using results obtained from real testbed experiments become a vital requirement on the way of wide deployment and public offering. Most of the mesh testbed studies available in the literature are performed in a small- or medium-scale testbeds deployed inside a single building or a campus and are mostly focused on data transfer performance of individual flows analyzing underlining protocol semantics. Examples of such testbeds include Roofnet [5], UCSB Meshnet [4], as well as mesh testbeds at Georgia Tech. [6] and Carleton University [7].

The main contribution of this paper is in the analysing of traffic traces derived from a large-scale wireless mesh network testbed deployed in three cities in the Trentino region (Trento, Rovereto and Riva del Garda). Being used by citizen and tourists in those cities, such WMN is able to capture the trends in user traffic based on the applications used and realistic mobility patterns – and thus provide useful models for further development and optimization of WMNs.

2 Wireless Mesh Testbed Setup and Components

2.1 Network Setup

Futur3 manages a wireless mesh testbed deployed in the province of Trento (Italy). It covers the main areas of the city of Trento, the city of Rovereto and the northern region of the lake Garda. At the core of each city, one internet connection point is deployed which forms the basis for a 5 GHz network deployment using HYPERLAN and HYPERLAN2 protocols. Such Internet connection points create a so-called first layer network. A second layer of the mesh network is built on 2.4 GHz Access Points (APs) proving WiFi access to the end users. These APs form an extension of the first layer infrastructure and operate according to the IEEE 802.11b/g standard.

The coverage of the second layer network APs overlap, thus creating redundancy for the multihop paths of the mesh network and guaranteeing improved connectivity.

Network Access Servers (NAS) are installed in each region for assigning IP addresses configured via DHCP functionality. An IP address assigned to a user will remain the same for the entire session duration. Routing is handled at the MAC layer and the hand-over between neighbouring APs is supported.

Furthermore, authentication to Futur3 network is provided according to the Italian anti-terrorism law: a user needs to register once by giving either his Italian valid mobile phone number or a credit card as a unique identifier. Once a user is registered, he or she can use his account all over the network. In addition, users are able to select the preferred level of privacy by updating their personal profile, changing the visibility of personal data, position or deciding to be not visible at all.

At the current state (February 2010), Futur3 network is composed by more than three hundred Access Points (APs) covering around 40 Km of streets continuously and around 16000 users are registered.

2.2 Localization

Futur3 localization is performed at three different levels of accuracy:

1. *Zone-based Localization:* A user is associated, given his IP address, to the zone (i.e. the corresponding NAS) he is connected to. In such a way, it is possible to associate an approximate position to every user.
2. *WiFi server localization:* From the user MAC address it is possible to know the AP a user is associated to. No particular software is installed on the user device.
3. *WiFi client localization:* For those users who have installed a client-side software, it is possible to acquire all the beacons sensed in the surrounding by the user and apply a triangulation technique to define a precise location.

Client localization works in combination with a localization server. Each client sends his list of sensed beacons to the server which computes the corresponding position estimates and sends it back to the client. The localization algorithm works in two subsequent steps: the first one provides a coarse position estimation, while the second one adjusts it.

Three different algorithms are evaluated for the first step: a simple centroid, a weighted centroid and a reduced weighted centroid. Centroid formula is an average of each APs' coordinates, while weighted centroid formula computes a position by averaging APs' coordinates after weighting them with their sensed RSSI. In order to define the last algorithm, the weighted centroid algorithm was applied to a subset of beacons obtained by filtering the sensed beacons according to their RSSI and distance from the previous client's position. RSSI filtering avoids location error propagation caused by multipath losses on the wireless channel. In a similar way, the distance filtering is due to the testbed position: apparently, water presence amplifies lake-side APs coverage area, thus affecting the overall results.

Location estimation is performed by both APs belonging to the Futur3 network as well as private APs. This requires an a-priori knowledge of the APs' positions which can be found using multiple wardriving sessions and a subsequent data analysis made by wigle.net technology [8].

In order to validate performance and precision characteristics of the performed localization, the testbed area is logically divided into a number of regions selected by isolating border areas. In each area the average error metric is calculated between each position computed by the location algorithms and a real position recorded in the system. Location error per region is reported in Fig. 1. The reduced weighted centroid algorithm appears to be better than the others because it has the lowest average error. Despite a large error in region 2 which has been accepted, average errors are much lower in other regions such as 4, 6, 8 (where the other algorithms have large errors).

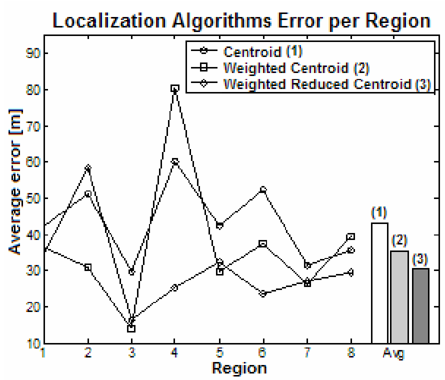

Fig. 1. Average error in meters computed in each region

Once a position of the user is estimated, the second step begins: the final latitude-longitude pair is selected from the set of available positions determined taking road topology of the area into account. At first, all positions within a range of 100 meters[1] from the coarse position are selected as candidate positions. Then, the final position is chosen by finding the position which minimizes both distances with the coarse position and with the previous valid one. Furthermore the selected position must be on the same road of the previous one; this constraint is ignored (and the second step repeated), either when the previous position is in proximity of street cross or if the distance from the coarse position is higher than 30 meters[2].

This client-server localization architecture doesn't charge clients with high computational load, enabling easy update of the positioning algorithms. Moreover, it reduces network overhead because the client doesn't have to query the database to fetch the available positions and the list of APs position used in the centroid algorithm.

3 The Measurements Campaign

3.1 Futur3 Network as a Data Collection Test-Bed

Futur3 network covers a wide area, with the purpose to offer both Internet connection and services to the largest number of people.

Typical users are residents of the area and university students, who connect during week days. A consistent fraction belongs also to the tourists coming to the area. Roughly 5-6% of the area population has joined the network, contributing to a growing penetration rate. This enables to have a suitable user base for performing a meaningful statistical analysis on the application and traffic traces.

---

[1] It is a value used to reduce the computational load because it avoids the server to process positions which are too distant from the previous valid one and therefore can't be selected.

[2] This represents the algorithm average error. When the distance between coarse position and position chosen accordingly to previous position is higher, it means that the user is in another street.

An additional advantage is that the network owner is Futur3; therefore it is possible to have simultaneous access to data about both users and network usage. On the user side, it is possible to perform accurate statistical analysis due to the knowledge of users' personal information such as age, sex and job [3], while on the network side it is possible to aggregate and process server activity patterns as well as APs load and links status. In detail, personal information about users is registered, their connection data in form of session time and traffic and AP usage in term of connection time and traffic; such data are periodically collected and aggregated in a data warehouse. Proper anonymization processes are used to guarantee privacy of the network users.

Moreover, Futur3 maintains a record of intranet activities related to its applications: users can interact with each other by exchanging messages and adding contacts. Analysing those data allows to study users behaviour and interaction in the social context. Furthermore, by using the provided positioning system, it is possible to study users' movement patterns.

The network covered area can be extended with on-the-fly installations, like in the case of the Blogfest event described in the following sub-section.

3.2 A Scenario of Interest: The BlogFest Event

Blogfest [9] is the second edition of an event hosted in Riva del Garda, a small city in northern Italy. The event is focused on the web community and its interaction within the Net; the aim is to allow people to speak about blog, social network and communities. It mainly consists of BarCamps organized and held by Italian bloggers all around the city. BarCamps are meetings held in different streets and squares within the old town, organized by bloggers on topics they usually write about, in which everyone can share his thoughts with the participants. BlogFest was held in October 2nd-4th, with more than 200 people attending 25 BarCamps. Moreover, there were a relax zone with radio entertainment and some stands of Internet companies and facilities such as food and kid areas.

In this context, the Futur3 WMN offered several services: (i) free Internet access, by improving the coverage of every area involved in Blogfest activities and most of the city area, (ii) introducing three beta applications based on the wireless positioning system. Those applications made Blogfest an interesting testbed, since they added the possibility to study users intranet interaction as well as their location and moving patterns. The applications provide information about other connected users, events and points of interest such as bars and restaurants which are nearby the user location, highlighting them on a map.

The data capture started the second day after the conference, when location services were deployed, at 11am. In order to be able to get access to Futur3's network and to use its applications, each user had to register to the network. 140 registrations were performed during the event: 20 the first day, 96 the second and 24 the third. The launch of the applications also explains the registration growth on the second day.

4 Experimental Data Analysis and Discussion

In this section an analysis of the available data is presented, with focus on users' behavior and network performance. The first sub-section introduces network usage and load history. Then, users' behaviour is considered, focusing on their interaction with each other and with respect to Blogfest event. A relevant number of people were observed (259 single users connected during Blogfest's event).

4.1 Network Usage History

Futur3 network is a growing reality: starting from 1000 users of Jan.'09 there have been more than 3000 users joining the network with a constant trend in 10 months.

Analyzing in more detail October, it is possible to observe an average of 875 unique users, with a maximum of about 1000 along workdays and a minimum of 750 during weekends. This difference is due to the fact that many users are university students who live in the cities only during workdays.

The relevant number of people connecting every day allows performing a reasonable statistical analysis on the users' surfing behavior and habits. Traffic supported by the network is estimated in roughly 2500 GB per month and peer-to-peer traffic is not allowed. The average downloaded data in October is around 78250 MB, while the maximum is 92600 MB and the minimum is 51600. Lower loads are experienced nearby week ends, probably because of a lower number of people connecting.

Figure 2 shows APs' positions in the Riva del Garda area, with different markers representing the number of connections which have been registered in October. There are clearly hot areas, which mainly correspond to the old town.

However, such distribution is not directly correlated to the traffic load distribution, since there are some APs which had few connections but many downloaded data.

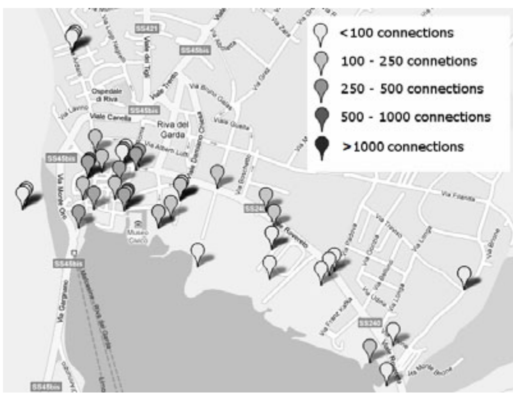

Fig. 2. Number of connections per AP in Riva del Garda

4.2 iPhone/iPod Application Usage

One of the most versatile applications was developed for iPod Touch and iPhone, which allows users to get access to all services and to surf the internet while moving. The application is interesting as iPhone can represent the reference for future mobile devices.

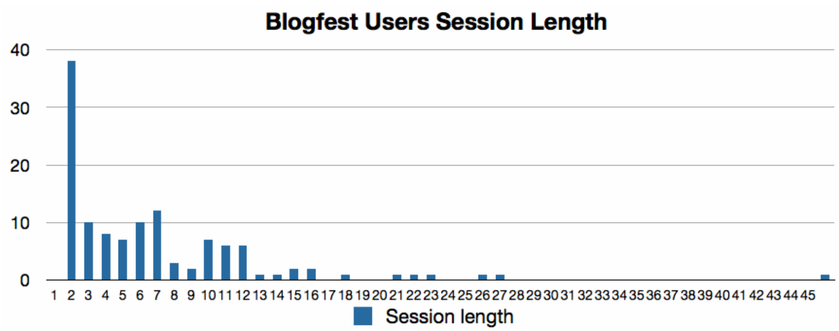

Fig. 3. Session length in minutes

Figure 3 describes the average session length in minutes. Several users had very short connections. Each user has had more than one session of different length; we measured around 5 sessions on average per user.

As far as movement is concerned, a mobility index is defined which estimates the level of movement of a user within one session. The index is defined as the ratio (covered distance)/(session length) measured in [meters/minutes]. Fig. 4 shows the indexes measured for each session. Sessions with mobility index lower than 60 are classified as very slow walk, or a session in which user didn't move while using the application. 72% of the overall sessions are characterized by such mobility index. However, 28% of the total sessions have a mobility index larger than 60, which means that users have used the provided application while moving around the area. Table 1 underlines the interaction patterns among users, captured by the Futur3 application enabled users to add contacts and to chat among the WMN users.

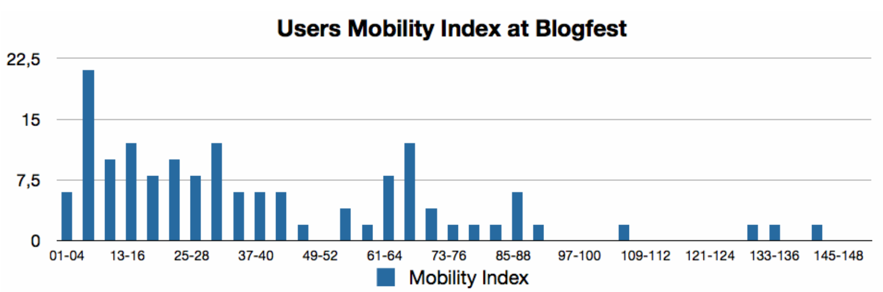

Fig. 4. Mobility index histogram

In the observation period, 35 new contacts were added by 14 over 31 overall users and 261 chat messages were registered. However, users didn't have long chat conversations. Most of the activities were done by users between 25 and 34 years old; those percentiles grow, by taking into account people younger than 24, to 80% for chat messages and to 78% for contact activities, as young people are keener to build social networks.

Table 1. Usage per age interval of features provided by the application

	<18	18-24	25-34	35-44	44<
Adding contacts	0%	21%	58%	21%	0%
Chatting	0%	13%	67%	20%	0%

4.3 User Distribution vs. BlogFest Activities

In this section, measurements are used to check the correlation between users' positions and Blogfest activities. Barcamps attendance is reported by highlighting on the map Barcamps' locations and indicating the users positioned there.

Figure 5.a shows a summary of users' participation at Barcamps held during saturday, Oct. 3rd in the afternoon. The purple area identifies both relaxing zone and a Barcamp. That area shows the largest number of people mainly due to the relaxing zone. The overall number of users do not match the number of Blogfest users because many of them have attended more than one event and have been in the relax zone.

Figures 5.b and 5.c show Barcamps attendance between 2pm and 3pm and between 5pm and 6pm, respectively. In those time intervals there were several active Barcamps. It is possible to observe a growing number of users connecting to the Internet while approaching evening. This is explained by taking into account the start of new Barcamps, which attracts more people.

Figure 5.d is a summary of Barcamps participation on Sunday morning. There were a few Barcamps active in the city, which causes most of the people to be in the purple area – i.e. the relaxing zone plus a Barcamp.

(a) October, 3rd, afternoon

(b) October, 3rd, 2-3 pm

(c) October, 3rd, 5-6 pm

(d) October, 4th, morning

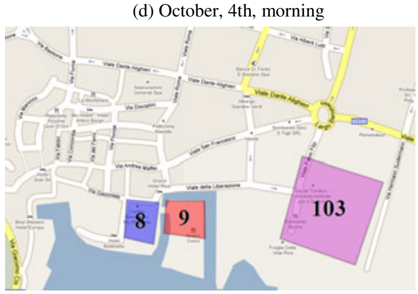

Fig. 5. Users' distribution at Barcamps and relax zone

4.4 Comment on the Results

Three different aspects, i.e. network usage, social behaviour and users' movement were analyzed, by taking advantage of a large WMN testbed and a special event, the Blogfest. The WMN has a sufficient dimension in term of connected users and traffic to allow conducting accurate statistical analyses. The possibility to estimate the position of the users identified areas highly loaded in term of connections and network traffic.

Basic social behaviour analysis is performed measuring the interaction among users allowed by intranet applications. Even tough most of the users were between 35 and 45 years old, younger users had most interaction in term of new contacts and chat traffic.

Moreover, a mobility index was defined, which allowed to discover that most of the users used the provided application while standing or for short movements, while 28% had relevant mobility.

During Blogfest events, it was possible to analyze the correlation between event and Internet connection (i.e. WMN resources usage). Knowledge of the users' location would be beneficial in the future for improving the coverage of the areas (in terms of transmission capacity, too). Based on the preliminary results obtained during the experiments, it seems reasonable to study methods for self-organization or self-resource allocation of the WMN possibly jointly with additional external information.

In addition, location-based traffic load information enabled to verify that some underutilized APs during regular days can become more crowded during a special event, but also that a WMN can be employed as a distributed locationing service to provide interesting insights to the events' organizers.

5 Conclusions and Future Work

The paper presented an analysis of measurements performed on a wireless mesh network deployed in three cities in Trentino. The presented test-bed is composed by a 350 APs network that provides access to around 16K users. Users can surf the web for free and receive also added-value applications to be used only as intranet services. Information and presented results are generated by analysing data gathered from users' normal behavior. In order to enable data information sharing within the scientific community, Futur3 is currently designing a testbed, working with the local university, in order to create a network subset with a similar behaviour.

The possibility of gathering data from a large testbed used by real users represents an important step in the investigation of large-scale access WMNs.

Results presented in the paper underline that most often the network usage is highly dependant on the APs' position and overall scenario (normal life or a special event). Moreover, the embedded locationing feature of the considered WMN enabled to check the percentage of moving users, which demonstrated to be relevant (28%).

Future works are aimed at using the gathered information in order to derive useful design guidelines for optimization and extension of the considered WMN.

References

1. Balanchandran, A., Voelker, G.M., Bahl, P., Rangan, P.V.: Characterizing user behavior and network performance in a public wireless LAN. In: Proceedings of the 2002 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, Marina Del Rey, California, June 15-19 (2002)
2. Afanasyev, M., Chen, T., Voelker, G.M., Snoeren, A.C.: Analysis of a mixed-use urban wifi network: when metropolitan becomes Neapolitan. In: Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement, Vouliagmeni, Greece, October 20-22 (2008)
3. Mulder, S., Yaar, Z.: The user is always right: a pratical guide to creating and using personas for the web. New Riders Publishing, Indianapolis (2006)
4. Lundgren, H., Ramachandran, K., Belding-Royer, E., Almeroth, K., Benny, M., Hewatt, A., Touma, A., Jardosh, A.: Experiences from the design, deployment, and usage of the UCSB MeshNet testbed. IEEE Wireless Communications 13(2), 18–29 (2006)
5. Bicket, J., Aguayo, D., Biswas, S., Morris, R.: Architecture and evaluation of an unplanned 802.11b mesh network. In: ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom) (August 2005)
6. Akyildiz, I.F., Wang, X., Wang, W.: Wireless mesh networks: a survey. Computer Networks 47(4), 445–487 (2005)
7. Wireless mesh networking, Carleton University, http://kunz-pc.sce.carleton.ca/MESH/index.htm
8. Wireless Geographic Logging Engine: Making maps of wireless networks since 2001 (2001), http://wigle.net/
9. Blogfest (2009), http://www.blogfest.it/
10. HotSpot Management without the Pain, http://www.coova.org