Survey of Experimental Evaluation Studies for Wireless Mesh Network Deployments in Urban Areas Towards Ubiquitous Internet

Serdar Vural*, Member, IEEE, Dali Wei, Member, IEEE, and Klaus Moessner, Member, IEEE

Abstract—Establishing wireless networks in urban areas that can provide ubiquitous Internet access to end-users is a central part of the efforts towards defining the Internet of the future. In recent years, wireless mesh network (WMN) backbone infrastructures are proposed as a cost effective technology to provide city-wide Internet access. Studies that evaluate the performance of city-wide mesh network deployments via experiments provide essential information on various challenges of building them. In this survey, we particularly focus on such studies and provide brief conclusions on the problems, benefits, and future research directions of city-wide WMNs.

Index Terms—Deployment, city-wide wireless mesh networks, experiments.

I. INTRODUCTION

Recent advances in mobile devices, e.g. web-browsing capability, powerful processing units, and the availability of larger storage spaces, have transformed simple communication devices into highly sophisticated but powerful handheld computers. Combined with the popularity of web-based applications, the focus of the efforts in emerging wireless technologies is on building highly dynamic and robust networks of wireless mobile personal devices that can exchange, distribute, and collaboratively process data. The emphasis of these efforts is on social networking, context and data sharing, and multimedia data delivery.

Internet connectivity is now a core feature of wireless devices due to an increasing demand on any-time and anywhere web access. During our day-to-day activities, online web pages and applications are now in more demand either for business or for entertainment and social networking purposes. Network and financial sectors [1]–[4] forecast that in the coming decade an average broadband mobile user will consume around 5 – 6 times more than what an average user consumes today [5]. The decreasing cost of portable computers with wireless connection capability, the emergence of smart phones, and the massive popularity of Internet-based applications gave rise to this demand, particularly in urban areas where wireless fidelity (Wi-Fi) hotspots are readily available at public locations. Hence, “out-of-office” Internet browsing is now quite popular with the availability of handheld devices having more intelligent features and better hardware capabilities.

In order to meet the ever increasing popularity of mobile wireless Internet, cellular network providers have embedded an additional support for data delivery in cellular infrastructures. As a result, the third generation (3G) cellular data networks, developed and enhanced under the 3GPP [6] consortium, are now widely deployed in city areas and provide today’s commercial solution for wireless Internet access to mobile users. Pioneered by 3GPP, 3G has recently been evolving towards 4G; technologies such as High Speed Packet Access (HSPA) [7], HSPA+, High Speed Downlink Packet Access (HSDPA), and finally Long Term Evolution (LTE) [8] have emerged to provide better data rates, reduced latency, and better cell-edge performance. Supported by most telecommunication companies, 4G LTE-Advanced [9] networks are expected to play a strong part in delivering city-wide mobile wireless Internet service.

In parallel to the research on cellular network technologies, alternative solutions towards establishing a city-wide mobile wireless Internet service have been investigated by a number of research groups. The main motivation is to provide an effective network architecture that incurs minimal building and maintenance costs. An obvious solution is to use available public Wi-Fi hotspots that are easy and inexpensive to deploy [5], [10]. Experiments in rural areas like that in a Californian desert [11] and that on an infrequently travelled road in Canada [12] report that significant amount of data can be transferred from Wi-Fi along the road. Even in the case of an urban setting, the experimental study in a residential area of Pittsburgh/Pennsylvania [10] shows that intermittent Wi-Fi connectivity can yield equivalent or even higher throughput than what an always-connected 3G network can provide. Similarly, although a metro-scale Wi-Fi experiences frequent disconnections in a mobility scenario, it delivers significantly higher throughput (median 1500 Kbps higher) than 3G and around 90% of the time outperforms 3G [13]. This potential of Wi-Fi motivates the efforts for providing more Wi-Fi connectivity in urban areas.

Despite Wi-Fi’s potential benefits, its connectivity and coverage properties as well as the eventual user experience depend on the range and position of individual hotspots, which vary between different urban areas. According to [13], the instantaneous throughput of Wi-Fi can be zero roughly around 1/3 of the time. Moreover, higher variability is observed in

The authors are with Center for Communication Systems Research (CCSR), University of Surrey, UK.
e-mail: {S.Vural, Dali.Wei, K.Moessner}@surrey.ac.uk
* Corresponding author
throughput relative to 3G. In [5], it is reported that only about 8.3% spatial Wi-Fi coverage is observed by mobile users, although temporal coverage is higher due to the type of user mobility patterns. On the other hand, 3G provides better coverage and has less throughput fluctuations [13]. Furthermore, frequent switches from one hotspot to another makes it difficult to maintain long-lasting Wi-Fi connections, not to mention the fact that RSSI (Received Signal Strength Indication) quickly diminishes at locations in between neighboring hotspots. [14] confirms this scalability problem of a single hotspot area caused by inefficient multihop Wi-Fi relays, which deteriorates the performance of delay-sensitive applications like voice over IP (VoIP). Since Wi-Fi is normally designed for relatively stationary users [10] and for indoor settings [15], public Wi-Fi hotspots are often unsuitable for mobile users [16]. An extensive list of these issues with wireless local area network (WLAN) deployments and hotspot designs is explained in [17] and [18].

A. Wireless Mesh Networks (WMN)

The inexpensive and promisingly high-throughput Wi-Fi technology can provide ubiquitous Internet connectivity. However, the observed coverage and connectivity drawbacks of fixed Wi-Fi hotspots should be eliminated. To achieve this, access points (AP) are developed with the capability of supporting standard low-power Wi-Fi clients at 10 times the range of conventional indoor APs. However, even with better hardware, around 25 to 30 APs are required per square mile in a suburban setting [15]. Deploying a T1 [19] line or fiber connection to each AP is quite costly. Fortunately, wireless links can inter-connect APs, which is a cost-effective and practical solution for a city area. This solution links all routers to a single backbone and deploys them in high density across an entire city area, forming a city-wide WMN that provides clients with Internet access with sufficiently high coverage and throughput.

The mesh of routers provide multihop paths to Internet gateways. It is shown in [20] that multihop routing has advantages over a single-hop architecture. With only single-hop forwarding, a number of well-placed gateways would be required to cover all mesh nodes, and even more would be required to match the average throughput of a multihop mesh backbone.

A WMN also enables end-users to connect to a single network as opposed to switching between local hotspots. Simply, a wireless router acts as an AP to a user device. Seamless connectivity is provided via transfer of user-related data and session settings between different APs through the backbone while the user moves. This architecture is shown in Figure 1.

Fig. 1. WMN structure: router backbone and user devices.

B. Cost

Wide-area mesh networking solutions consisting of cheap router equipments provide a cost-effective alternative to expensive [10] [22] centralized 3G networks. Besides providing cheap and almost free service to customers, Wi-Fi is also an affordable solution for service providers since customer equipment, e.g. laptops, handheld, mobile phones, already has built-in Wi-Fi cards at no incremental cost [15]. Via increased coverage and more efficient node deployment strategies, a WMN is also envisioned as a better solution for Internet access compared to conventional Wi-Fi hotspots.

To make a market comparison of 3G and Wi-Fi, the authors in [16] first recognise Wi-Fi networks and cellular 3G networks as major competitors and then aim to determine the market positions of the two technologies in the coming decades. Using a game theoretical approach, a competition model demonstrates that Wi-Fi can have a better market position in the future due to its low costs, whereas 3G is likely to make higher profits with higher pricing policies. Although Wi-Fi in general is projected as a more attractive service [16], a WMN deployment should be carefully planned, as results in [23] suggest that the deployment cost would be significantly raised to guarantee sufficient coverage if the network is not pre-planned.

C. City-wide WMNs: present and future

Various deployment efforts have recently been made by companies and private organizations [24], and many companies now provide outdoor mesh router equipment [25]–[30]. Cities around the world are today in a constant motive to build up their own wireless network to address the digital divide [31], [32], to ensure public safety via video surveillance, or to provide better community services. A few examples are the networks deployed in Las Palmas/Spain, and Corpus Christi [33], Oulu/Finland [34], and Madison/Wisconsin [35] with many others being added to the list each year [15], [24]. Therefore, city-wide mesh network deployments are now getting more popular since municipalities require cost-effective broadband solutions and the public demands free
outdoor wireless access. To realize this vision around the world, mesh networking companies are making more and more efforts for developing better mesh hardware. This is only possible by understanding the current performance of WMNs, which can truly be achieved via detailed site surveys and in-field experiments on currently deployed networks. Seeing this need for a better understanding of city-wide WMNs, research groups have been investigating these deployments by onsite experiments. These efforts will shed light on the necessary next steps to improve this technology.

D. The focus of the survey

In the current literature, there is a vast number of studies on WMNs with topics ranging from analysis [36] and modeling to architectural design [37]. The evaluation method of these studies is usually an analytical framework or a simulation platform. Recognizing the fact that experimental evaluation of networks provides precious know-how and deep insights to researchers [38], this survey specifically focuses on studies that utilize field experiments. Although many papers address indoor and outdoor settings usually within a limited geographical area [39], to understand the true potential of a city-wide deployment, we particularly select papers on large-scale urban-area WMNs, e.g. city-wide or campus-area deployments, which are built to deliver Internet access to end-users. Furthermore, the survey deliberately limits the focus to published research papers that have goals towards critically assessing specific WMN deployments via experiments. In doing so, we aim at providing researchers with a categorization of the research issues commonly encountered in city-wide WMNs.

The structure of this paper is as follows. First, in Section II, we briefly explain the properties of a set of deployed networks that are frequently mentioned in research papers on experiments of urban-area WMNs. Then, in the rest of the survey, we present the evaluation of these networks from different perspectives. Specifically, Section III presents deployment related issues, such as connectivity, coverage, and node placement. Topics related with physical propagation and link-level aspects are then explained in Section IV. Section V investigates the observed data traffic properties, whereas Section VI briefly describes mobility-related issues. Last but not the least, Section VII provides a categorization of different types of measurements made on these networks and their corresponding findings. Each of these sections are finalized with a table summarizing our major conclusions on the subject matter. Finally, Section VIII concludes the paper.

II. BRIEF DESCRIPTIONS OF THE NETWORK DEPLOYMENTS EVALUATED BY RESEARCH STUDIES

A number of flagship deployments are evaluated via experiments. These are Google’s Wi-Fi network in Mountain View/California, TFA (Technology For All) in Houston/Texas, MadMesh in Madison/Wisconsin, Roofnet in Cambridge/Massachusetts, and Metromic’s Ricochet network in San Francisco/California. In this section, the specific properties of these networks are summarized. After presenting a brief introduction of these networks, a summary of campus-wide network deployments is also provided since their physical medium characteristics and deployment challenges are similar to those of urban-area networks. For example, Dartmouth College’s wireless network is studied in a couple of studies and hence included in this survey. A short description of University of Surrey’s (UoS) network is also included since we were able to have first-person interviews with the staff members who deployed the network. Table I compares the basic properties of these networks.

A. Google Wi-Fi

The Google Wi-Fi network [40], deployed in Mountain View/California, provides a free outdoor wireless Internet service to public users and has been operational since 2006. Its architecture consists of 500 Tropos MetroMesh [33] pole-top APs, which are approximately 100 m apart on average, covering an area of around 31 km² with some 72,000 inhabitants. The APs have distinct identifiers based on their well-known locations, while the topology is dynamically constructed with the proprietary Tropos routing algorithm [41]. Each Tropos AP has a 7.4 dBi antenna and a single 802.11 g wireless interface.

The Tropos nodes are hierarchically clustered around 70 point-to-point radio uplinks that build up a fixed long-haul backbone. Traffic is eventually routed to one of three distinct wired gateways spread across the city and connected to the Google campus using GigaBeam equipment. With this structure, the network is more densely deployed than MadMesh [42], but not as dense as Roofnet [43]. The relatively dense deployment of the APs and the hierarchical structure ensure that most clients have short paths to the gateway, e.g. two hops. To extend the indoor network coverage, Google suggests two particular Wi-Fi modems: the PePLink Surf and the Ruckus MetroFlex.

B. TFA-Houston

The TFA network in Houston is part of the TFA project [44] that aims to develop new wireless network architectures and protocols that enable high-performance ultra-low cost wireless access to low-income users in city areas. The network is deployed in an urban area in Houston where 20000 residents live, mostly in one-story homes of average 5 m height. The network includes 17 mesh nodes, providing coverage to a 3 km² area, with high gain 15 dBi omni-directional antenna, placed approximately 10 m above the ground. NEMA 4 waterproof enclosure is used to protect the mesh node hardware since the mesh nodes are mounted externally on residences, schools, libraries, and other properties.

TFA has a two-tier structure: (i) an access tier for connecting client devices and (ii) a backhaul tier for interconnecting mesh nodes and forwarding traffic to/from wireline entry points. A single radio and a single channel is used at each mesh node for both tiers. Traffic management techniques (rate
limiting) are employed for ensuring proper division of these resources between access and backhaul tiers.

C. MadMesh

MadMesh consists of 250 Cisco 1510 Mesh Access Points (MAPs) distributed in 25.9 km² of Madison/Wisconsin. The network has a tree structure rooted at Root Access Points (RAP) with wireline access to the Internet. For each RAP, a separate tree of Mesh Access Points (MAP) is configured. These MAPs are deployed on utility poles, creating an expected coverage of around 0.3 to 0.4 km. Client devices associate with a MAP to join the network. Based on changing channel conditions, MAPs can change their parents in the tree, or even switch to a different tree.

In MadMesh, links between different MAPs on a tree form a mesh backbone. Cisco 1510 MAPs are equipped with two interfaces: (i) Backbone interface, for communication on the mesh backbone and (ii) Access Interface, configured to act as a regular Access Point (AP) for client devices. The backbone interface operates on 802.11a 5 GHz band, while different trees operate on different 802.11a channels. The access interface uses 802.11b/g in the 2.4 GHz band with different MAPs operating on different channels.

D. Roofnet

The Roofnet network in Cambridge/Massachusetts consists of 38 mesh nodes distributed over an area of around 6 km² that is characterized with its urban environment and dense population. Each node is a PC with 802.11b Intersil Prism 2.5 chip-set card and a roof-mounted omni-directional antenna. The PCs belong to volunteer users living in houses with antennas mounted on the roof or placed on the exterior walls of the buildings.

Roofnet radios transmit at 200 mW, operate with RTS/CTS disabled, and share the 2.422 GHz 802.11b Channel 3. The omni-directional antennas provide 8 dBi of gain with a 20° –3 dB vertical beam-width. These Roofnet nodes, which are completely self-configuring, run identical turn-key Linux-based software.

At start-up time, a node advertises itself to Roofnet as an Internet gateway in case it is able to reach the Internet directly; otherwise, it acts as a DHCP server and a default router for the hosts on its Ethernet, and connects to the Internet via Roofnet, which has four Internet gateways.

Roofnet has its own routing protocol Srcr, which aims to find the highest-throughput route between any pair of Roofnet nodes. To achieve this, Srcr chooses a suitable bit-rate among the 802.11b transmit bit-rates using an algorithm called SampleRate. Srcr carries IP packets inside its own header format. The routing software is implemented into a DHCP server and a web server so that users can monitor the network status.

E. Metricom Ricochet

Metricom’s Ricochet packet radio network [45] is installed in three major metropolitan areas (San Francisco, Washington D.C., and Seattle) as well as airports, hotels, and college campuses scattered across the US. Poletop repeaters are used as either a wireless AP with a wired and a radio interface, or as a wireless repeater that forwards packets to another poletop via the radio interface. Subscribers register at poletops which they choose according to RSSI values.

F. Campus-Wide Networks

Many college and university campuses around the world today have medium to large scale wireless networks already established to provide Internet access to their students and faculty members. As deployed in an urban area over a number of departmental buildings and recreational areas, studies on such networks also provide valuable information on similar deployments.

The Dartmouth College network is one of the campus-wide networks studied in some recent papers. At the time of these studies, the university had around 5,500 students and 1,200 faculty members using the network. The network serves 190 buildings in the 0.81 km² campus area and consists of a total of 557 Cisco Aironet 350 APs, running VxWorks. The APs operate at 11 Mbps with a range of about 39 – 106 m indoors, while there is no provided specific intentional outdoor coverage. The network APs share the same SSID, while APs of each building are connected to one of the 81 wired subnets through a switch or hub.

The UNC network in University of North Carolina (UNC) covers a 2.95 km² campus and a number of off-campus
administrative offices. There are totally 488 APs that belong to three different series of the Cisco Aironet Platform: 269 APs are the 1200 Series, 188 APs are the 350 Series and 31 APs are the 340 Series, where the 1200 Series and 350 Series run Cisco IOS and the 340 Series runs VxWorks.

University of Surrey (UoS) Wireless Network: The UoS wireless network (UoS-Net) covers multiple buildings in the UoS campus area. The APs in UoS-Net are deployed indoors due to the fact that student usage outside the residence halls is mostly confined in some popular buildings. Among others, the university library and the lecture hall are the most popular locations for wireless users.

The initial design of the UoS network in 2002 originally consisted of APs each having a direct connection to the Internet. However, due to maintenance difficulties and scalability problems, this design was replaced by the current architecture, which is based on Central Controller (CC) units. In UoS-Net, MX-200 type CCs manufactured by Trapeze Networks are deployed. Coverage is provided by 444 APs (MP-422B type) that are wirelessly connected to 6 CCs. CCs have direct connection to the wired Internet and provide easy maintenance and configuration of APs. CCs are also essential since they assign frequency channels to individual APs. All APs are powered by Power-over Ethernet (PoE), which makes it easy to simultaneously supply and configure the devices. Additional APs are planned to be deployed since a significant increase in the number of smart phones and portable computers has been observed during the past two years.

### III. NETWORK DEPLOYMENT STRATEGY AND NODE LOCATIONS

In this section, issues related with the deployment strategy of a network are discussed. This includes choice of node locations, how to organise the network hierarchically and determine AP locations, connectivity and coverage, as well as the use of advanced antennae to improve deployment performance.

#### A. Provisioning

An obvious question when deploying a large-scale WMN is whether to apply a detailed provisioning or to simply consider a static deployment strategy. Little deployment planning and operational management is clearly attractive as a cost and time saving solution. For instance, the evaluation in [20] of Roofnet shows that its unplanned mesh architecture achieves an average end-to-end throughput of 627 kbps, requires few Internet gateways to serve the entire network, and has better connectivity and throughput over a hypothetical single-hop network.

On the other hand, the sophisticated structure of urban environments may lead to unexpected performance deficiencies [21]. Hence, having an unplanned network runs the risk of degraded performance during operation time. Motivated by this, the deployment of MadMesh [42] is well-structured and was preceded by detailed site surveys. The structure is also managed via a continuous change of AP locations based on performance requirements.

In short, preplanning a network is a safe but time-consuming precautionary step. Although some networks may be lucky enough to not require it, provisioning is generally suggested [53], [54], especially for a city-wide WMN. Specifically, placement of nodes affects various characteristics of a WMN, e.g. node associations to APs, network connectivity and coverage. In the following, the observations on these characteristics are explained.

#### B. Node locations: Grid vs. Random

Among others, perhaps the most interesting issue is how to place nodes in the network such that optimum performance can be obtained. This is a clear point of consideration to have a successful transition from a setting where there is a number of sparse and non-coordinated Wi-Fi hotspots [10] to another in which a WMN enables seamless connectivity.

Various perspectives on how to achieve an optimal node placement are suggested in wireless networking. One obvious design choice is whether to build a grid or a random network structure. To understand the effect of this choice on city-wide WMNs, inter-node distances in a grid deployment are varied and compared with random deployments [47]. It is found that a regular grid topology achieves up to 50% higher throughput than a random node placement. Furthermore, when small perturbations in node locations are made, regular node placement experiences no performance degradation for location perturbations of 1/6 of inter-node distances. In fact, small perturbations show some improvement in average throughput as the routing protocol may take advantage of better links. For higher perturbations, coverage gaps start to occur. For instance, in the Google Wi-Fi network, 20% loss in client throughput is due to uneven spacing of mesh nodes [40].
Despite these particular experiences with a specific network, there is no sufficient evidence on how much node locations should be randomized so that a better connectivity and coverage is obtained as compared to a pure grid. The effect is expected to be different in different deployments, hence a further extensive research is required to derive some useful models and more specific guidelines. It is clear that such a study should consider local topological constraints and provide suitable strategies to overcome each encountered problem.

C. Interference

Node deployment, especially the choice of AP locations, should obviously consider possible interference sources. The study in [14] provides a deployment procedure and has specific emphasis on how to deploy a high-speed backbone so as to avoid interference sources as much as possible. However, despite the planning efforts, unexpected performance degradation is still observed at some routers due to topological settings. This fact suggests that interference avoidance is possibly one of the hardest challenges of a wireless network deployment in an urban area.

D. Indoor locations

The indoor location of the APs is also a critical issue; a suitably located AP can cover most of the indoor users and extend the network’s coverage. Regarding this, network deployers are strongly suggested to carefully choose AP locations according to device installation manuals, taking into account the signal reception and transmission capabilities [21]. The use of software tools can be particularly helpful in indoor locations. For instance, the deployment of APs in UoS-Net is planned by the RingMaster software [55], which determines how many APs are needed at a particular floor plan as well as where they need to be placed. The software uses the floor plan and the types of obstacles in making these decisions. Then, deployment engineers make use of this initial provision as a guide for similar floors and locations. Moreover, the APs are mostly deployed in corridors so as to avoid interrupting the staff’s office work.

E. Node Associations

The efficiency of a multi-tier network depends on the interaction between the lower-tier user devices and the upper-tier APs. User devices must first connect to individual APs to retrieve Internet access. The study in [48] signifies the importance of AP positioning since the distribution of clients, their roaming experience, and the eventual wireless user experience is affected by the deployment strategy. A couple of remarks are made on AP associations.

One particularly important issue for a user device is how to choose among a number of possible APs, in the event that multiple are available. The traditional approach is a comparison of RSSI, however [48] suggests otherwise. Considering the node association policy, it is found in [48] that a 15 – 20% gain in average client throughput can be obtained when a load-balanced policy is chosen instead of a simple SNR-comparison based association policy. With a SNR-based policy, a single card may associate with multiple APs residing at different buildings, although the card is stationary. Furthermore, when a user approaches an AP (although it has higher RSSI from another) it may be better to connect to this new AP so that the usable connection duration can be maximized [10]. This observation calls for better AP selection algorithms for mobile users. Faster AP association techniques and setup procedures similar to QuickWiFi [56] can also improve the experiences of mobile clients [10].

The availability of multiple APs is an advantage towards achieving high connectivity and coverage. However, problems may emerge if user devices make frequent AP switches. In [57], many such authentication events are observed. In fact, more than half of the users authenticate more than 50 times during a week. This is due to frequent re-authentication and switching events between different APs caused by fluctuating RSSI levels. However, the average number of packets generated by each authentication event is low, compared to the number of the authentications in some parts of the network. This is attributed to repeated disconnections and reauthentications. A similar observation is provided in [50], [51].

Besides unnecessary AP switches, another problem is the existence of “dummy” AP associations. Examination of sample sessions [18] show many instances when a card connects to an AP although the card is already associated. This is attributed to the fact that the AP and the card states are not properly synchronized, suggesting the need for possible improvements of card firmware. Furthermore, most results demonstrate that a large number of users associate to the network but not actually use it and are so-named “pending” users. In short, there is considerable redundancy in user association events, which should be dealt with to improve network efficiency.

F. Connectivity

For a seamless connection experience, one fundamental requirement is high connectivity. To have Internet access, user devices should have a complete path to the wired mesh points. This probability that a node has at least one access path to a wired Internet mesh point is defined as the access network’s reliability [47]. An ideal deployment should provide a high reliability to reach Internet gateways no matter where and when the connection is made.

1) Node density: Deployment node density is one factor affecting connectivity and reliability of connections. To achieve a reasonable reliability, it is stated in [47] that a node spacing of approximately 200 m is sufficient for the TFA network. Obviously, this value is different for different networks and what the optimal value should be is a potential topic of research for city-wide WMNs. Availability of more wired mesh points certainly increases reliability, yet the number of wired mesh points is constrained, i.e. not necessarily all mesh...
nodes have a direct connection to the Internet. Therefore, when this number is smaller, a higher wireless mesh density is required to achieve a certain level of reliability.

2) Throughput as a measure of connectivity: Connectivity performance can be evaluated via measuring the aggregate throughput when random user devices try to reach the Internet and retrieve files. It has been observed that node density comes into play and affects the achieved throughput. For instance, a consistent decline in average throughput is observed in [47] for increasing inter-node distances, whereas adding an additional wired location to the network increases the average throughput by a factor of 2.75. A similar observation in [20] shows that when the number of nodes increases, the average throughput also increases. This is because a denser network offers short high-quality links, despite the fact that using them causes longer multihop routes and higher time delays. The tradeoff is that overprovisioning the network density may have diverse effects, especially a high deployment cost. If one is to build a network with optimal number of nodes and obtain suitable local node density at different locations for a set of given budget constraints, a typical on-site survey should take into account how to choose proper inter-node distances.

3) Node degree: The number of neighbours that a mesh node has, defined as the node degree, is an indication of how dense a network is. Various networks reportedly have different node degrees. For instance, in Google Wi-Fi, it is observed that an AP can communicate with at least 4 neighbours and also that the most well-connected 10% APs each have more than 8 neighbours [40]. Connectivity properties of the MadMesh network studied in [42] show that around 20% of APs have node degrees less than 2, while 60% have degrees more than 3, demonstrating a sufficiently dense deployment, yet not as dense as Roofnet [43]. The measurements in [20] illustrate that Roofnet takes advantage of a highly-connected mesh architecture, with the majority of nodes having many neighbours. In spite of the fact that these networks demonstrate mostly a high node degree, their backbone tree structure is of importance when possible node failure events are considered. In other words, when an AP fails, this may disconnect its sub-trees from the rest of the network. In such cases, the routing protocol of the WMN should quickly adapt to such changes in the topology and modify the tree structure. Node deployment strategy should also consider this situation and ensure that a number of APs are available to serve the same network region. Hence, deployment density should be adjusted according to the possibility of node loss events in the backbone as well. Site surveys can provide useful information on local connectivity properties and help define the set of conditions for such properties to emerge.

4) Sharing broadband connections: Towards enhancing Internet connectivity, the possibility of sharing broadband access in urban areas is investigated in different scenarios in [21]. The results indicate that it would be attractive for users to share their connections, which dramatically increases wireless connectivity in urban areas offering around 10 Mbps in most cases. Specifically, on upper floors of buildings, the achieved gains are more significant: a service area of at least 35 m and with an offered bandwidth of 4 Mbps was obtained. Similar to [21], the integration of community networks to form large-scale connectivity is considered in [14]. Although such integration efforts suggest a slightly better performance, inter-networking performance of the mesh networks with the existing platforms still needs to be studied in detail.

G. Coverage

One of the fundamental requirements of wireless network deployments is to provide a sufficient coverage area. As a user moves, a disconnection occurs if the user’s mobility path crosses through holes in network coverage. Such holes are caused by either insufficient number of mesh nodes deployed in localities or due to misconfiguration of transceivers causing sub-optimal use of transmission power.

One interesting area of research is how to deploy the mesh nodes so that maximum coverage is ensured. This is certainly a very common networking problem, especially for networks with resource constraints. This is addresses by a few papers for city-wide WMNs. In [46], three disjoint regions of the Google Wi-Fi network are considered: one residential, one commercial, and one transit. It is observed that although clients are distributed widely among the APs, all APs contribute substantially to network coverage. Since this is not the case for the initial deployment for the Dartmouth network, additional nodes are used and an average percentage increase from 66.4% to 76.4% in the number active APs is achieved. This is attributed to the fact that with the availability of new APs, additional locations can be covered [52]. Likewise, the results collected in a corporate campus area in [23] show that with an increased number of APs better coverage is achievable.

On the other hand, high deployment density does not necessarily lead to increasing coverage and network performance in every deployment. For instance, in [48], it is found that deployment density in Google Wi-Fi has little impact on coverage hole probability above $\frac{\text{nodes}}{\text{km}^2}$. It is also questionable as to how many more nodes are needed to provide a sufficient coverage probability. The current deployment in Google Wi-Fi, for instance, has nearly 25% hole probability. To have 90% coverage, the density should be as high as $\frac{77 \text{nodes}}{\text{km}^2}$, dramatically higher than the current $\frac{17 \text{nodes}}{\text{km}^2}$. A similar issue is observed in the TFA network. Results in [48] indicate that the TFA network is deployed with sufficient density. Since most coverage holes are located at network edges, they are removed easily in TFA by adding more nodes. However, in other networks, the problem is challenging as holes are small and spread out. Moreover, when mesh nodes are deployed to minimize the client-mesh distances, this leads to a three times over-deployment in TFA, since coverage may be provided by mesh nodes not necessarily closest to clients.

Another issue is that having additional nodes does not necessarily lead to high RSSI, although it generally improves the coverage area. For example, in [23], it is reported that only a few areas with excellent signal conditions of signal-to-noise ratio (SNR) greater than 25 dB are available. Therefore,
when a coverage test is made, the minimum RSSI required by everyday devices should be considered and the RSSI across a network should be evaluated accordingly to have a better picture of coverage in that network.

Another potential topic for the case of city-wide WMNs is how to detect the existing coverage holes. Detecting coverage holes, however, is not straightforward since some sort of continuous feedback is needed from every single part of the network. The easiest way to attack this problem is a fail-and-improve procedure, simply put: client feedback. This can be done when periodic data, such as RSSI and location, are collected from clients [42], [48]. For instance, a monitoring structure is suggested in [42] to be put in place in order to identify potential coverage holes, despite the successful connectivity results of MadMesh.

In contrast to other studies, [57] suggests that a complete coverage that supports an excellent on-the-go connectivity is not necessary for practical reasons. In fact, it is argued that mobile computing should focus on user location rather than user mobility patterns. The “usefulness” of the network is observed to be proportional to the probability that users will be able to get connection at locations that they normally visit, leading to *islands of connectivity*. Therefore, it is argued in [57] that wireless access should be given initially (with higher priority) to locations where users actually need to connect to wireless resources frequently. The mobility study on connectivity to public Wi-Fi hotspots in [10] suggests a similar finding: intermittent throughput achieved by hotspots is sufficient for the mobility patterns observed at client devices. This argument is however controversial. Mobility patterns of users may change over time and there is no widely accepted model that generalises user mobility preferences. Hence, a city-wide WMN is supposed to provide seamless connectivity to most types of mobile users, exception being only high-speed vehicles, at least until more stable WMNs are provided. Hence, coverage studies should also consider user mobility.

### H. Directional antennas and MIMO

The use of directional and multiple antennas is useful to improve connectivity, coverage, and the overall performance of a network. Although, this topic has not been extensively studied for city-wide WMNs, studies frequently recommend the use of MIMO and directional antennas.

Routers with multi-antenna transceivers have higher gain and provide better signal quality [58]. Especially for a sophisticated environment like urban areas that suffer from multi-path effects and signal distortions, multi-antenna systems help amplify and rebuild signal levels. For a city environment, a multi-radio based architecture is likely to improve network performance [23]. Similarly, the use of higher-gain antennas may a better cost-effective solution rather than a denser deployment [48]. For instance, in [40], Wi-Fi modems and bridges with highly capable antennas are recommended to be deployed to improve coverage and connectivity of the Google Wi-Fi network.

Directional antenna systems have a higher gain in a certain angular direction, which improves SNR and also helps improve transmission range towards that direction for a given transmission power level [59]. Since the transmission beam is concentrated to a particular direction with a certain beam width, the remaining angles get significantly reduced portion of the transmission power. This property is exploited to reduce interference to other receivers. The study in [60] has the observation that when directional antennas are deployed with various orientations in a metropolitan environment, the interference between neighbouring channels is greatly reduced. On the other hand, since the beams do not cover a 360° range, connectivity of a network with directional antennas is weaker than that of a network of omni-directional ones. Despite this, it is shown in [42] that good connectivity in WMN backhaul is achieved, although more than 70% of APs use 11 dBi sectored antennas at backhaul interface to achieve high RSSI performance.

Table III provides our conclusions on deployment-related topics.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planned vs. unplanned</td>
<td>Pre-planning a deployment is required to avoid unexpected performance degradation due to poorly chosen router/gateway locations.</td>
</tr>
<tr>
<td>Deployment strategy</td>
<td>Grid deployment is a better choice as compared to a completely random node placement strategy.</td>
</tr>
<tr>
<td>Mesh node locations</td>
<td>Small perturbations in node locations may be helpful to improve the overall performance.</td>
</tr>
<tr>
<td>AP load</td>
<td>Mesh node positioning affects user association load on the APs.</td>
</tr>
<tr>
<td>Deployment modeling</td>
<td>An intelligent configuration of indoor mesh node locations using simulation software can boost network performance.</td>
</tr>
<tr>
<td>Interference</td>
<td>Site surveys should carefully consider interference sources.</td>
</tr>
<tr>
<td>Client association policy</td>
<td>A load-balanced policy can provide improvement over a simple comparison of SNR levels.</td>
</tr>
<tr>
<td>Card associations</td>
<td>Card firmware improvement is required to avoid redundancy in user-association events</td>
</tr>
<tr>
<td>Mesh node density</td>
<td>Suitable node density values for network regions should be determined via site surveys considering connectivity, QoS guarantees, and deployment cost. Systematic methods are required e.g. models and prediction software</td>
</tr>
<tr>
<td>Broadband sharing</td>
<td>Sharing of broadband connections may improve wireless connectivity.</td>
</tr>
<tr>
<td>Internet access reliability</td>
<td>Site surveys should determine ideal spacing between routers considering terrain features.</td>
</tr>
<tr>
<td>Mobility support</td>
<td>Users with different mobility patterns should be considered.</td>
</tr>
<tr>
<td>Advanced antennas</td>
<td>Coverage and connectivity improvements require further research.</td>
</tr>
</tbody>
</table>
IV. Physical Medium and Link Properties

The wireless medium in urban settings has high variability due to the diverse environment consisting of buildings of various heights, public parks, and vehicles. This complexity necessitates accurate characterization of signal propagation properties. Among others, pathloss exponent and shadowing parameters are the most fundamental metrics that need to be estimated at different parts of the city. Furthermore, multi-path effects as well as individual link properties, (in terms of SNR and bit error rate (BER) values) are precious pieces of prior information for deployment designs. In this section, the focus is on these parameters of physical signal propagation in the wireless medium.

A. Pathloss characterization

One characteristic of a wireless propagation medium is the loss in signal strength along the path that the signal traverses. According to the medium characteristics, a certain amount of pathloss is observed, which is quite dependent on topological properties. Hence, having some prior information on the pathloss characteristics of a deployment area is highly beneficial to better provision necessary network resources.

Accurate prediction of the exponent is crucial, because a too high estimate of the pathloss exponent may cause an order of magnitude increase in deployment cost, as it may be wrongly predicted that more routers are required to provide sufficient coverage. For instance, using a pathloss exponent of 5 as in [61] yields a highly over-provisioned network (over 9 times the cost in the number of deployed nodes). On the contrary, low pathloss estimates may lead to even completely disconnected networks.

Unfortunately, conventional models of the pathloss exponent are found to be inaccurate and unfit in urban settings. For instance, in MadMesh [42], significant variation between different MAPs is observed, suggesting that the conventional pathloss model may not be appropriate for a city environment. In fact, the metric should be determined separately in different parts of the city; variations are likely. A pathloss of 3.3 and a shadowing standard deviation of 5.9 dBm are observed for an environment of small, wood-framed houses, and dense foliage in the TFA network [47], out of 50 random measurements made at various locations and link distances. However, this is in contradiction with [62] that predicts a pathloss exponent of 4 and 8 dBm deviation in an environment of concrete and steel highrise buildings.

B. Propagation model

The propagation model used to plan a network deployment should also be sufficiently accurate. Unfortunately, the classical uniform propagation model is often unsuitable for urban settings. For instance, the estimation framework proposed in [48] reportedly removes more than half of the errors made by the uniform propagation model. Estimation methods are suitable candidates for provisioning a network deployment. However, a sophisticated area, such as a city that may have high variations in terrain types, requires that the topology be divided into manageable sub-sections, a process called sectorization. This helps better model different parts of a city.

The study in [48] is a good example of sectorization. The network is first divided into feature areas whose boundaries are represented by polygons determined by a map processing algorithm. Afterwards, attenuation weights are assigned to each feature type, where a weight value represents a feature’s impact on pathloss and shadowing. This sectorization technique is performed for Google Wi-Fi and TFA networks [48] and is found to have less effect on Google Wi-Fi compared to TFA, indicating a larger variation in terrain features in TFA.

C. Link loss rate

Link-level measurements are crucial to understand the pairwise communication quality in a local area. Before proceeding to an end-to-end test of larger scale, such measurements reveal what individual links can deliver as part of a multihop route. Although this topic is extensively studied for indoor settings and WMNs covering small areas, such measurements appear in a few papers for backbone links of a WMN in urban environments [14], [20], [47], [60], [63].

Characterising packet loss events is a suitable way to assess link quality. Deploying mesh routers while ensuring that they have proper link connectivity is central to deployment strategies since it determines the required node density. A weak mesh link can be an indication of a need for an additional mesh router, a higher transmission power, a better equipment, or a more suitable router position. For an urban environment, [43] is an example study, focusing on packet loss events. The aim is to characterize when and how packet losses occur. As a result of extensive measurements at the Roofnet network, it is found that most node pairs have intermediate levels of loss rates, indicating no sharp transition between high and low packet loss rates. This is in contradiction to an assumption-basis called the neighbour abstraction that defines the set of nodes a particular node can communicate with. The abstraction assumes a rapid transition from an essentially zero BER to a BER high enough to corrupt every packet. Although some studies claim that abstraction holds [64], [65], some others say otherwise [66], [67].

Currently there appears to be a need for more research on link-level measurements, especially to understand causes of packet loss and link disconnection events. Preliminary observations in [43] have some simple conclusions so far. First, a significant set of links in Roofnet show no discernible relationship between distance and link-level delivery probability, hence the effect of distance is not considered to be a major cause of losses. Secondly, although links with very high signal strengths are likely to have low loss rates, signal strength and SNR also have little predictive value in general. In contrast, a link is likely to have a significant loss rate at its optimum 802.11 bit-rate.

Observations may also differ among different networks, possibly caused by a different selection of the individual links that the measurements are made on. Although, most links are
found to have non-bursty loss patterns in [43], observations in [60] and [63] find many asymmetric links with daily-varying qualities. Hence, a generic globally-accepted method of assessing link quality is required, which can be used as a common guideline for every deployment.

D. Multi-path effects

Signal interference caused by multiples of the same signal traversing different paths is quite likely for 802.11 systems in urban areas since these systems are normally designed for indoor environments. Experiments carried out using a single transmitter-receiver pair and measuring throughput and packet loss rate are useful to observe how much multi-path fading affects network performance. Multi-path interference cancellation techniques need to be employed at receiver devices as studies clearly show a need [43], [47].

In [47], throughput is measured to be lower than expected, which is attributed to multipath effects. 235 trials of 60 second intervals performed around 4 fixed mesh points indicate noticeable separation between throughput measurement data and its linear approximation. This high deviation is likely due to a wide range of delay spread, which can dramatically increase packet loss rate. UDP throughput is found to be “7 dB” less than nominal values reported by wireless card specifications. (Throughput is approximated as a piecewise function of RSSI, due to the lack of theoretical throughput prediction models that can use environmental parameters as input.)

Tests are performed in using a channel emulator in Roofnet [43] that emulates an original signal following a line-of-sight path and a single reflective signal. It is observed that some delay values at multiples of modulation symbol boundaries cause more loss than others. However, since the results are not obtained in the Roofnet environment, which would produce many reflective rays, the emulation results are a lower bound on the losses caused by reflections, signifying the strength of multi-path effects in cities.

E. Interference effects

Interference is one of the most challenging issues in a wireless network. The low transmit power of client devices and high levels of interference often cause high packet error rates at access links [42]. Therefore, in-depth monitoring of possible interference sources from inside and outside of a network should be performed.

External interference should be clearly investigated in deployment scenarios, although some networks may be lucky to have little outsider impact. For instance, in [43], packets originating from foreign 802.11b sources seem to have no correlation with the number and frequency of packet losses experienced by Roofnet nodes. However, this is getting more and more unusual as 802.11-based networks are now commonplace and public hotspots are major sources of interference. In fact, external interference has indeed an impact on the client throughput [42].

A sub-optimal channel allocation scheme may cause significant internal interference, especially when wireless interfaces are assigned to neighboring channels [60]. However, if some channel separation is applied, each flow’s throughput can be improved. In fact, when there is only a single channel separation, no significant interference is observed [60] and the throughput is essentially equal to the UDP sending rate. Furthermore, if there are only a few channels available or the channel assignment scheme aggressively assigns a fixed small set of channels to transmitters, interference is likely to occur, e.g., 83% of the time links use the same three channels in [60]. This is also observed in [42]; low signal quality is found to be partially caused by the fact that APs are not operating at the best available channels that minimize interference.

We derive a set of conclusions on the physical and link-level properties of city-wide WMNs, as listed in Table IV.

### TABLE IV

<table>
<thead>
<tr>
<th>Topic</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propagation models</td>
<td>Models and parameters should be found for different network locations.</td>
</tr>
<tr>
<td>Link-level loss rates</td>
<td>Predominantly intermediate and non-bursty. SNR and signal level have little predictive value.</td>
</tr>
<tr>
<td>Multi-path effects</td>
<td>Possible cause of intermediate loss rates. Methods are required to avoid them.</td>
</tr>
<tr>
<td>External interference</td>
<td>Channel selection mechanisms are required to find best channels for APs.</td>
</tr>
</tbody>
</table>

V. NETWORK TRAFFIC

Understanding traffic sources and flow characteristics is required for suitable provisioning of network resources. Key points to investigate are the properties of existing traffic sources and flows and types of application traffic. Affected by these properties, a network’s successful operation is limited by how much load balancing is achieved among APs and how efficiently the existing traffic flows are serviced, i.e. packet routing. In the following, some issues about traffic sources is first mentioned, then the load by currently dominant application types are briefly explained, followed by a discussion on routing-related topics.

A. Traffic sources and flows

There are a number of interesting observations on network traffic flows found in city-wide WMNs. In summary, these are on: the network usage by mobile terminals, daily traffic variations, the effect of management traffic on performance, proportions of inbound and outbound traffic, and flow modeling.

1) Large share to few users: In most networks, traffic is dominantly generated by a small fraction of the user population. This is also found to be the case for 3G networks as reported in [5]; about 3% of all users generate around 40% of the network traffic.
In [50] and [51], a striking observation is that more than half of the total traffic is caused by only 5% of all cards. The busiest one transfers 117 GB, while the median card uses only 350 MB. The same observation in [18] reports that most of the traffic is generated by fewer than 5% of the users. Similarly, according to [68], most (66.34%) of the traffic is generated by a small fraction (10%) of the sessions. The work in [69] also presents similar results captured at a corporate wireless network. It is found that the most active 10% of users are responsible for more than 40% of data.

These results show that user profiling can be quite handy for resource allocation purposes. Certain user classes according to their generated/collected data volume can be formed. According to each classes other properties, such as temporal usage or location and mobility, resources can be dynamically allocated to avoid possible congestion or starvation events.

2) Outbound vs inbound traffic: The proportion of inbound traffic seems to dominate the total traffic. This is likely due to high tendency of users to download/stream multimedia from Internet servers. There is always less outbound traffic than inbound traffic in the Dartmouth network with varying proportions around 89% [50], [51]. The study in [68] has the same conclusion on this ratio: inbound session traffic is mostly larger than outbound session traffic. However, the reverse is true for 6.62% of the sessions in [68]. The mean ratio of daily outbound to inbound traffic is 1 : 5.82.

3) Daily traffic variations: In [50], [51], the daily traffic varies considerably. During the busiest day the network moves over 240 GB, whereas the median daily traffic is only 53 MB. The busiest day of the week also varies dramatically at different buildings. In the study of [52], the average daily traffic per active card is found to have risen from 27 MB to 71.2 MB between 2001 to 2004, indicating more active users. An overall increase in daily traffic is observed with a rise from peaks of 150 – 250 GB in 2001 to over 400 GB in 2004.

4) Control traffic overhead: Unnecessary maintenance traffic may cause performance degradation in a city-wide WMN. Some causes of such overhead are duplicate packets sent due to continuous reception errors, broadcast packets, and RTS/CTS packets. Management traffic also produces additional control traffic. For instance, in [57], it is detected that approximately 38% of the packets are generated by automated network management and maintenance traffic, with a continuous arrival rate of 15 packets per second. Some networks may enjoy small packet sizes and not really be adversely affected by control messages, yet this is not the case in general for all networks. In [70], while the majority of APs send and receive packets of relatively small sizes, a significant number of APs show rather very asymmetric packet sizes, which may possibly lead to starvation in access links.

5) Flow modelling: Modelling network flows helps understand how to deploy a hierarchical network structure and how to distribute network resources. However, well-known non-competing network flow models for mesh networks are found to overestimate throughput. Instead of models, measurements can be helpful in performance prediction. Measurements of single and non-contending flows can capture the effect of reduced throughput with increased path length. For instance, according to [47], when deployment decisions are made with the help of measurements in a multi-flow environment, a large fraction of starving and disconnected nodes are located. As a remedy, it is proposed that the rate of each flow must be controlled and set to a fixed value. Using this idea, when around 9 flows go through the links, a rate limiting of 1/9 of the downlink capacity is applied to each flow and better fairness is obtained.

6) Transmission gaps: Rate controlling is less effective for uplink traffic due multiple sources producing traffic towards the AP simultaneously. With each additional flow, there is some probability of lost because nodes give equal priority to their own traffic and the forwarded traffic. This backlogged traffic leads to starvation, yet if there are transmission gaps applied between different sources, congestion at upper links towards the AP can be removed. For instance, web traffic permits sufficient spaces in time if statistical multiplexing is used. As an example, web-emulated traffic patterns show better uplink performance in [47]. The aggregate throughput is more than double the throughput of a fully backlogged download scenario and with much greater fairness.

7) Data offloading: Application of data-offloading can lead to substantial gains in client throughput when a high number of users are simultaneously active. In fact, many data transfers can be delayed which reduces network congestion [5]. This is possible by giving incentives to users to tolerate transfer delays and enjoy decrease in service costs whenever their needs is not urgent (movie or music file downloads, non-urgent email or file transfers). In fact, it is found that the average completion time of data transfers is much shorter than the deadlines users currently choose [5]. When applied, data-offloading has been shown to provide more than 50% cost reduction to service providers [5].

B. Application traffic

To have an effective provisioning of a network and allocate sufficient software and hardware resources, it is important to be aware of the types of applications using the network. Studies show various results, yet the common point is that multimedia and p2p traffic are getting higher shares of the total traffic. Multimedia streaming traffic and especially voice data is quite popular in today’s networks. This is one observation in [23], in which VoIP (e.g. originating calls with Skype software to another Skype user and to a cellular subscriber) traffic is monitored in the Google Wi-Fi network. Good voice quality is found to be possible in the case of Skype to Skype calls, yet the quality is very poor or literally at an unusable level when calls are made to cellular phones. VoIP traffic is analyzed also in [52] and the hourly number of calls shows a diurnal pattern. The number of devices that makes a VoIP call each day is found to be almost static, with users making short calls around 41 seconds.
In [50], it is found that http is the most popular protocol occupying about 53% of the total traffic. In contrast, a later study [52] provides detailed classification of application types in the same network and reports significant changes in the type of applications. For instance, thanks to the new embedded devices with peer-to-peer (p2p) services, streaming data is more prevalent in the network, with an increase from 5.2% to 19.3%. The same study shows a www-traffic decrease from 62.9% to 28.6%, hence a shift from a web-dominated workload to a p2p dominated workload. The proportion of streaming traffic also increases, an almost 400% change from 0.9% to 4.6%, with over 129 GB of streaming traffic observed.

C. AP load

When it comes to a hierarchical structure, one way to evaluate its performance is to understand whether the total network load is balanced among APs. Since APs connect user devices to the network’s backbone, overloading the APs leads to local, or even more serious cases, a sectional loss of services due to congestion. However, as networks get larger with more clients, AP usage also gets an incline. Due to this increase in the number of users, an overall increase in the density of users per AP is unavoidable. For instance, the number of network clients in [52] has steadily grown since the authors’ prior study [50].

1) Increase in clients per AP: Besides additional clients being added to a network each year, the usage per each client is also increasing in recent years, mostly due to the increasing popularity of web-based streaming applications. This leads to a high activity of APs. There are many observations supporting this trend: In [18], most APs are found to be active everyday, while in [52] 429 out of the total 566 APs are found to be active at a peak time. In [46], results show that even at night, over 80% of the APs are servicing at least one client. A gateway node in [20] forwards an average of 160 kbps between Roofnet and the Internet and its radio is busy for about 70% of the monitoring period.

2) Unbalanced AP load: Despite the general increase in average client load on APs in recent years, this load is not always evenly distributed in every network. For instance, in [18], the majority of cards associate with a single AP and almost all spend most of their time at a single AP, while some other APs may serve only a few clients. Similarly, in [42], many APs in MadMesh are lightly loaded with 110 APs having an average of less than 1 client connected. In contrast, some APs may be highly popular: for instance, around 50% of the clients are connected to 20% of the APs in [42]. This load imbalance is also found in [52] and [49]. In [52], 91 cards logged on a single AP in Dartmouth network in its peak time. In [49], poletop APs may handle up to 182 distinct radios an hour, with most poletops handling 500 or fewer radios events per hour. Even some busy ones may have a load of 1600 events in an hour. In contrast, some poletops are found to be inactive, processing only a few events during the entire trace, while others may handle thousands per hour.

3) Load depends on location: Load on APs is found to be greatly dependent on AP locations; popular APs are often concentrated at certain geographic areas [42]. As a result, several small clusters of clients can be formed in a network resulting in an overall, uneven distribution of clients. For instance, according to [50] and [51], in Google Wi-Fi, the APs with the most active cards in their busiest hour are those located near large halls, with the busiest AP having around 71 active cards. In contrast, due to their locations and traffic patterns, most APs do not engage in data relaying, and are the leaves of the backbone [40]. Similar results are reported in [69]. Load distribution among APs is highly variable with APs located at large auditoriums being largely used while others are usually idle. 30% of active user traffic is affected by user location.

D. Routing

The choice of routing metric is central to the success of any routing protocol. For instance, it is ideal to avoid formation of long routes as long routes may have diverse effects on performance. Especially in download traffic scenarios, unfairness occurs due to forwarding overhead of multihop flows proportional to hop count [47]. Even greater effects are observed for upload traffic with pronounced falloff in bandwidth when hop count increases. The metric called ease [42], which is a weighted sum of SNR and link qualities, causes around 8% of the APs to have hop counts of more than 5 to RAPs.

Link quality is definitely a determining factor on routing decisions. The sensitivity of ease [42] to link-quality and SNR fluctuations is a contributing factor in the unsteady routing paths of about 10% APs in MadMesh. Furthermore, frequent parent node changes at the backbone routing tree can occur if the routing protocol is too sensitive to SNR changes, largely caused by existence of links with similar qualities. Dampening mechanisms can be applied to avoid frequent route oscillations.

Routing algorithms should avoid weak links as much as possible, yet there needs to be a load balancing mechanism in place as well. For instance, in [20], the routing protocol Srcr largely ignores most links that are 500-1500 m long and have throughputs of less than 500 kbps. However, if multiple flows are assigned the same favourable path simultaneously, congestion can easily emerge. As an example, in Google Wi-Fi [40], very few backbone links are found to be actively used, with a large portion of the traffic concentrated on specific paths, which may be problematic as more flows are added to the network and the existing flows start to carry larger data rates. Hence, scalable solutions are required considering the rapid increase in Internet traffic.

In conclusion, our key derivations on network traffic are summarized in Table V.

VI. MOBILITY

Mobility support is one crucial feature of a city-wide network. Although some users have habits of logging on the
Internet at stationary stations, the proportion of users on-the-go is on the rise. This facilitates the need for functionalities that enable seamless mobility.

A. Roaming

Roaming among different APs is a critical issue. A frequent roaming pattern leads to a high overall volume of management data transferred between APs, which not only disrupts network performance, but also is a situation prone to loss of user-related data, temporary disconnection of sessions, and eventual user dissatisfaction. Unfortunately, roaming issues are frequently reported in research papers [70]. For instance, the proportion of mobile and roaming cards increased from $\frac{1}{4}$ [50] to $\frac{1}{2}$ [52] of all cards.

Roaming can be caused by various factors. For instance, in [46], the observed frequent roaming events are attributed to the movement patterns of user devices; i.e., by changing its location, a device may associate to different APs. Another factor is a sub-optimal network deployment causing coverage overlaps. A device may fluctuate back and forth between two APs of similar signal strength. Hence, this is one area when the network performance is affected by deployment decisions. A third reason is the fact that wireless device cards may follow aggressive search procedures to catch the strongest signal as much as possible. This leads to frequent changes in AP-associations, when multiple APs are in the range. Card firmware should be improved to avoid this overhead. Solutions such as MobileIP [71] can also help reduce roaming.

Sessions may demonstrate varying degrees of roaming frequency. For instance, most roaming sessions in [50] and [51] are found to be short and infrequent, while some sessions roam extremely frequently. Furthermore, different devices may have different roaming behaviour; while smartphones frequently associate with a large number of APs, modems and hotspot clients may associate with fewer APs [46].

B. User mobility

User mobility patterns can provide useful insights on how APs adjust their resources to avoid session disconnections. For instance, using usual clustering algorithms, various mobility patterns can be found [49], [52]. These clusters can be formed as different combinations of minimal, moderate, and high mobility and/or network usage. This sort of classification is quite beneficial; network service providers can understand how much mobility support is needed and should be delivered to what percentage of the user population and in which occasions.

Characterising and classifying [69] user mobility provides useful information on user preferences in connection locations. For instance, users associate with APs for longer periods of time at home locations rather than guest locations [69]. Similarly in [49], the distance that users move between two registration/query events is found to be mostly confined within 161 km. The more locations users visit in a day, the closer those locations are. Although users move with their mobile devices, this movement is generally limited and show close relation to the hour of the day, with 64% of all users appearing in only one location. Such findings tell network managers exactly what they need to know about user mobility: the preference of connection location.

In contradiction to usual predictions, some studies notice that user mobility may not be a very critical issue and that mobility-related events occur scarcely. For instance, in [50] [51], user mobility results show that few cards move around frequently, with nearly 18% of the cards spending all their time in a single building. Most users limit their activity to a few key sites during their daily routine and are found to be stationary within each session. Hence, the authors suggest that providing wireless access to a wide range of popular locations is favourable rather than ensuring seamless mobility at all locations (which requires a more difficult and costly deployment). Similarly, it is shown in [52] that almost all users (95%) almost all the time (98.7%) spend their time in their home AP which they associate to for more than 50% of the time. Despite these observations, it is still a controversial issue whether to design a network’s mobility support feature solely based on user mobility patterns, although they provide useful information in the short term. In fact, when a city-wide mesh network is considered, the system should support service to all mobility classes no matter how often they are seen in these patterns. Users may not even try to use the Internet while they are mobile simply because they anticipate that the network would probably not support the service or with very low quality. Therefore, trends of user mobility sessions is not a perfect indication of the future trends.

We have two brief conclusions on mobility in city-wide WMNs, given in Table VI.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roaming</td>
<td>Leads to frequent association and authentication events. Its frequency varies among different sessions.</td>
</tr>
<tr>
<td>Seamless mobility</td>
<td>May not be required in most cases. Should be supported for different mobility classes.</td>
</tr>
</tbody>
</table>
VII. MEASUREMENTS

This section is purely focused on two topics: (i) in Section VII-A, measurements techniques to analyze the performance and characteristics of the networks mentioned in Section II, (ii) observations on network usage patterns in Section VII-B. By mentioning the measurement techniques, the survey aims to provide an overall idea on how research groups evaluated their networks. The part on usage patterns show the resulting classification of network usage as retrieved by these measurements.

A. Methods used in experiments

Various methods are used to collect measurement data, while most studies share a few popular techniques. Here, a brief information on each technique is provided along with some examples from research papers.

1) SNMP polls: SNMP\(^1\) polling is a usual method of data collection [18], [70]. As an example, around 193 million records are collected in [69] to monitor traffic as well as usage profiles.

SNMP polls generate a large database including an extensive number of parameters. For instance, in [42], each record has more than 150 parameters; some examples are: total number of clients, MAC address of the parent AP, MAC address of neighbours, channel number, number of failed transmissions, and noise floor level at the AP. To achieve this, the study uses a two-week excerpt of a measurement database of 1.7 million SNMP log entries, which is around 100 hours of active measurements. Similarly in [52], 16 million SNMP polls are performed to collect data on inbound/outbound bytes, packets, errors at APs, client IDs at APs, MAC/IP addresses of clients, and signal quality and strength at clients.

SNMP logs are collected in certain time intervals resulting in periodic user data. For instance, in [42] polls from all active APs are retrieved once every 3 minutes. In [50] and [51], APs are periodically polled every 5 minutes. Each poll returns the MAC address of recently associated client stations, inbound and outbound bytes, and a list of cards associated with an AP.

2) Device-type detection: Device-type detection schemes are used to classify users. For instance, in [52], the OS fingerprinting tool pOf, which identifies the operating system of devices, is used on tcpdump traces to determine the type of wireless devices. For each card (MAC address) seen in syslog and SNMP traces, pOf is run on all its TCP flows recorded by the sniffer\(^2\). The unidentified cards are checked for their Organizationally Unique Identifier (OUI) at their MAC addresses. An approximately 150 embedded 802.11 PDA and VoIP devices are detected in this way.

3) System logs: Several papers use syslog messages to collect data from APs, related with client associations and their sessions. This helps understand traffic load on APs as well as statistical information on traffic flows collected by the APs.

Some examples are as follows: In [46], overall network statistics are collected by the central Google Wi-Fi RADIUS server for 28 days. Periodic updates are generated by all Tropos nodes for each associated client every 15 minutes. In [68], data are collected within a period of about 5 months, generating a total of 14273 sessions and 174.75 GB of traffic. Each entry in the log consists of 5 fields: the user name, the login timestamp, the number of bytes inbound and outbound, and the session duration. In [57], a software package, called EtherPeek, is used to collect trace data, to record MAC addresses, and to analyze usage patterns. Cisco’s LEAP authentication system along with a RADIUS server keeps track of wireless users connected to the system. The information includes each authentication instance, its date and time, the associated username and client MAC address, and the associated AP’s IP address.

A common goal of system logs is to detect how frequently AP-node associations take place. To achieve this, APs are configured to transmit syslog messages each time a client card authenticates, associates, reassociates, disassociates, or de authenticates, [50]–[52]. In [51], among the 476 installed APs, 430 are monitored by syslog. Similarly, in [49], a trace of log entries at the nameserver is collected, with each log entry consisting of registration and query events. A log entry is associated with a single radio and includes a poletop ID. Poletops correspond to certain localities and are used to determine registration and data transmission events that occur in different locations of the network. Queries are made when the radios make a new connection. The idea is to determine how much activity takes place between consecutive connections.

4) Active measurements: Besides automatic updates, such as SNMP polls and system logs, making active measurements is another popular method to measure specific performance parameters, such as throughput, or link related metrics like error rate. Such measurements are usually performed using a laptop dedicated to experimentation or sometimes by setting up measurement scenarios with network nodes as active traffic sources.

Extensive field measurements of physical and application layer performance for access and backhaul links are performed in [47]. First, access links are evaluated via 138 trials of 60 sec intervals over distances greater than 100 m. A laptop is connected to an AP to send backlogged traffic to a single client device. For various distance values between the AP and the client, the RSSI power is recorded. Similarly, in [14], different source and destination pairs, bit-rate, packet size, and inter departure time values are selected.

In [43], network nodes generate experimental traffic for active measurements. Each node sends 1500-byte 802.11 broadcast packets for 90 sec at each of the 802.11b rates, while the rest passively listen. Senders record the time of transmissions and each receiver records the arrival time and the RSSI right before packet reception.

In [20], network nodes are configured to conduct experiments. The performance of the Roofnet is evaluated via multihop and link-based TCP measurements. For multi-hop

\(^{1}\)SNMP: Simple Management Network Protocol

\(^{2}\)sniffer: a software program installed to monitor network traffic.
evaluations, the TCP throughput between a fixed set of nodes is measured. For link-level settings, the loss rate between each pair of nodes and the TCP throughput on the direct radio link between each node-pair are found.

5) Passive monitoring: Passive monitoring is more commonly performed than active measurements since most studies focus on understanding real usage patterns rather than experimental settings involving a device that sends artificial data. However, for some particular goals, such as application level traffic analysis, measurements are not possible since all traffic is encrypted [42]. Despite this, sniffer applications usually provide valuable information, such as link-level metrics like signal strength and noise, without interrupting network traffic.

In [52], 18 Ethernet sniffers located at 14 different buildings (dormitories, libraries, department halls, and social areas) obtain detailed network-level traces and collect a total of 4.6 TB of sniffing data. One interface of the sniffer box is used for device maintenance, while the other is used for collecting data by the sniffer. To monitor the usage of the campus VoIP system, every voice log details parameters like time, duration, caller’s and callee’s IP Addresses, and call termination reason (hang up or diverted).

The wireless measuring tool in [21] is WiNet Explorer, which uses Linux wireless extensions to collect link layer statistics such as signal quality, signal level, noise level, and AP configuration. Network layer statistics, such as average RTT, packet loss ratio, and delay are also calculated.

In [50] and [51], in order to understand the usage patterns, tcpdump is used to capture all packet headers at specific APs of the campus area. Among the 476 installed APs, 22 are monitored by tcpdump. A computer in promiscuous mode connected to an AP through a common hub is used to monitor the traffic. However, only part of the wireless traffic is successfully captured due to some factors, such as the network structure consisting of many subnets, the high number of APs, and AP geographic locations.

B. Usage Patterns

This section presents the observations on network usage based on different device types and user classes as well as the temporal changes in usage patterns.

1) Temporal Usage Patterns: Considerable amount of measurement work is devoted to time-based network usage patterns. The common goal is to derive useful conclusions from user behaviours towards developing better provisioned networks. Like many others, it is reported in [18] that hourly, daily, and weekly trends emerge in network usage.

Daily usage patterns:

Daily usage patterns are observed in each network. Such information can be exploited to estimate the load on the network, which can assist management tasks. In [50] and [51], results on the variation of daily traffic reveal that traffic is steady throughout the afternoon and evening and tails off through the night. Hence, the activity level varies widely from hour to hour. The diurnal usage pattern is similar in [52]; i.e. the traffic peaks in the afternoon, is almost steady during the evening, and drops from midnight to early morning. In [68], the highest user activity appears between 8 pm and midnight, while in [42], traffic volumes are reported to peak in the evenings, at around 10 pm, and night hours, rather than daytime, consistent with the traffic patterns of residential users.

Connected clients:

Patterns of the number of active clients are also observed in some studies. The information on connected, active, passive, and idle devices can help provision network management resources, such as number clients associated to each AP and number of authentications to network servers.

Regular daytime patterns may emerge in some networks. For example, the number of active clients in [40], [46] peaks in user activity roughly twice early in the morning. The study shows that the Google Wi-Fi network has a substantial daily user population, peaking around 2 pm and 500 simultaneous users in any 15 min interval. Daily number of users is also derived in [69] and demonstrates a regular pattern. However, the number of days each user appears on the network is found to be highly variable among different users.

On the other hand, radios in the Metricom network are used mostly during non-work hours, due to users connecting to faster networks during work hours [49]. Similarly, the users of a commercial Wi-Fi hotspot network in [68] access the network for casual use rather than everyday activity, e.g. most sessions are for after work activity and not part of the users daily routine.

Weekly usage patterns:

Weekly usage patterns are also frequently reported in various networks. In [49], it is shown that weekday usage of the network is higher than the weekend usage, especially in the middle of each week. Similarly, in [40] and [46], weekend use is lower than on weekdays, with roughly 15% fewer users during peak times at the weekends. On the other hand, a roughly uniform daily activity is reported in [68] for all days of the week, where Friday is observed as the busiest day.

Weekly usage patterns can be a good indicator of who is actively using the network. For instance, in [50] and [51], card activation events reflect a typical student pattern of activity. This pattern is similar to that of an office worker’s pattern of activity. A reasonably strong weekly pattern is seen with Monday being the busiest day during the week with a dramatic burst in both inbound and outbound traffic at around 10 am.

2) Device Types: Matching usage patterns with user device types gives some clues about what sort of services users will be needing according to estimations on how much these devices will be popular among wireless network users. Recognizing this fact, insightful results are provided in [40] and [46].

Three distinctive classes are determined: (i) local residents and businesses that make static connections to the Wi-Fi access network, as a substitute to DSL (Digital Subscriber Line) or cable modem, (ii) laptop users with mobility and workload patterns typical in public Wi-Fi hotspots, and (iii) smartphone users with a higher mobility pattern.

The three device types show markedly different application
usage: smartphones generate very few connections and almost all their bytes are Web or TCP applications; both modem and hotspot users show a significant amount of peer-to-peer, http, and non-TCP traffic.

Regarding the amount of active clients, it is observed that modem activity overall has the highest transmission rates, hotspot activity is roughly uniformly distributed across the range, whereas smartphone activity varies.

Clients with different devices exhibit different distributions of session length: most modem clients have sessions that span the entire trace, many hotspot clients have sessions shorter than an hour, and smartphone clients have the shortest session length.

Different types of clients show dramatically different usage profiles as well. The number of modem clients is constant throughout the day; hotspot users show peak usage in later afternoon; smartphone users are highly correlated with commute and travel times and present three distinct peaks during the day, i.e. 9 am, 1 pm, and 6 pm.

3) User Classes: User classes emerge according to user locality and session lengths. For instance, users may be classified according to session lengths [40], [46]; three distinct user classes (clusters) are determined: (i) a cluster of short and light sessions, (ii) a cluster of extremely long and heavy sessions, and (iii) a cluster that spans a wide range of session lengths.

Another classification is based on geographic locality. Significantly distinct properties are detected with respect to residential, commercial, and transportational areas of the city. Each user class is observed for their presence in these distinct local areas. Furthermore, different devices show different usage in the distinct areas. For instance, clients in commercial areas are the most active ones, whereas smart phones are far less prevalent in residential areas. In contrast, residential wireless activity is reported to dominate all usage [51], while in [52] the busiest and most popular types of building are academic and library buildings. Hence, usage locality depends on user preferences and where the network is deployed.

4) User privacy: To ensure the security and privacy of network clients, all studies have measures of removing/avoiding any data on user identities. For instance, [40] is focused on aggregated client data rather than data pertaining to individual users. Hence, no client side information is collected or stored. Similarly, in [57], user anonymity is central and collected data are cleared of any user-related signatures.

Measures to protect user identity may lead to limited control of collected data and an incomplete set of results. For instance, data in [49] are collected from the Metricom network under a number of restrictions to protect user privacy and network security, which makes it hard to infer more insightful results specific to individual traffic flows.

Main conclusions on experiences in measurements carried out for existing deployments can be listed as in Table VII.

VIII. CONCLUDING REMARKS

The fast emergence and development of mobile wireless user devices have led to an increasing demand for Internet-based applications. Urban dwellers now request ubiquitous Internet access, which gives a strong motivation to the wireless research community to provide a service that is low-cost and efficient for providers, and also affordable and sufficient for clients. Towards this objective, a recent perspective is to build a wide-area mesh network covering an entire city area consisting of a backbone of routers that act as wireless APs. Typical mesh networking know-how is applicable to such networks, however insights are needed to see how physical properties of an urban environment and those of a laboratory setting differ. Hence, experimental evaluation of actual WMN deployments in city areas provides valuable information on the research issues that need to be tackled in order to realize cost-efficient Internet service to mobile users.

In this survey, we have evaluated the observations on experimental tests carried out for a number of city-wide mesh network deployments that appear in published research papers. Several topics are identified as potential research directions in this field. These topics range from deployment-related issues, e.g. coverage, connectivity, and planning, to issues pertaining to the physical characteristics, e.g. pathloss measurements, multipath effects, and interference. Apart from such usual wireless networking problems, understanding what kind of usage patterns emerge, how users behave, and the nature of the existing traffic flows is of fundamental importance as a guideline for future deployment efforts. We believe that the survey draws key conclusions on the research in citywide mesh deployments and extracts valuable information on provisioning such networks towards ubiquitous Internet access in urban areas.

ACKNOWLEDGMENT

The authors would like to thank to Mr. Leigh Jones and the rest of the IT Services staff in University of Surrey for the information they provided about UoS-Net.

REFERENCES


<table>
<thead>
<tr>
<th>Topic</th>
<th>Measurement methods</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Use of SNMP polls, systems logs, and Ethernet sniffers provides valuable data.</td>
<td></td>
</tr>
<tr>
<td>Usage Patterns</td>
<td>Hourly, daily, and weekly usage patterns exist according to traffic volume and number of clients.</td>
<td></td>
</tr>
<tr>
<td>Device Classes</td>
<td>Different classes emerge according to application usage, session lengths, transmission rates.</td>
<td></td>
</tr>
<tr>
<td>User Classes</td>
<td>Distinctive classes emerge according to session lengths and geographic locality.</td>
<td></td>
</tr>
</tbody>
</table>
the Development of Networks & Communities (TridentCom), March 2008.


