A Feasibility Study of an In-the-Wild Experimental Public Access WiFi Network

Arjuna Sathiaseelan
Computer Laboratory
University of Cambridge
Cambridge, UK
first.last@cl.cam.ac.uk

Richard Mortier,
Murray Goulden
Horizon Digital Economy Research
University of Nottingham, UK
first.last@nottingham.ac.uk

Christian Greiffenhagen
Department of Social Sciences
Loughborough University, UK
first.last@lboro.ac.uk

Milena Radenkovic
School of Computer Science
University of Nottingham, UK
first.last@nottingham.ac.uk

Jon Crowcroft
Computer Laboratory
University of Cambridge
Cambridge, UK
first.last@cl.cam.ac.uk

Derek McAuley
Horizon Digital Economy Research
University of Nottingham, UK
first.last@nottingham.ac.uk

ABSTRACT

Universal Internet access has become critical to modern life, leading to many explorations of approaches to increase its availability. In this paper we report on a study of one such approach, PAWS, that seeks to understand the technical and social constraints of providing Internet access, free at the point of use, by sharing existing broadband subscribers' connections. We elaborate the technical and social context of our deployment, a deprived neighbourhood in a medium-sized British city, and discuss the constraints on and resulting architecture of this system, including the authentication and security mechanisms necessary for a service of this kind. We then report on the use of our deployment over a period of seven months from July 2013 to February 2014, including analyses of the performance and usage of the network. Our data show that PAWS is socially and technically feasible and has the potential to provide Internet access economically to many who are currently digitally disenfranchised. However, doing so requires overcoming numerous challenges, both technical and social.

Categories and Subject Descriptors

C.2.3 [**Network Operations**]: Network Monitoring

Keywords

Free Internet; Socio-economic; Wireless

1. INTRODUCTION

The Internet is commonly held to be a ubiquitous part of everyday life, and with worldwide penetration standing at around 34% [1] this is increasingly the case. Indeed, in July 2012, the United Nations unanimously backed a notion stating that "all people should be allowed to connect to and express themselves freely on the Internet". Internet access is increasingly presumed in the developed world by everyone from commercial organisations to governments. For e.g., the UK Government's *Digital-by-Default* programme aims to make essential services, e.g., access to social housing[1] and benefits, available only online to realise claimed cost savings and service improvements [2]. Unfortunately at the same time the UK's telecommunications regulator, OFCOM, reports that fixed-line broadband penetration in the UK is only 75% and data from the Internet World Stats survey of October 2012 reports that 10% of the UK population do not have Internet access at home [3]. Given this trend for services to move online, the negative consequences of digital exclusion are only going to increase [4].

In many cases (particularly among the elderly) cultural factors (particularly perceived lack of value) are given as the critical reasons for remaining digitally excluded, but cost is also often a significant factor. For e.g., 2011 data from OFCOM shows cost is the primary barrier in younger demographics, e.g., for 40% of 16–44 year olds. Similarly, cost is a key barrier in lower socio-economic groups, e.g., broadband Internet takeup is just 56% in working class and non-working households but reaches 90% in upper and middle class households; and of those lacking Internet access in working class and non-working households, 17% cite cost as the main reason for their lack of an Internet connection [3]. Current efforts by the UK Government to address digital inclusion have focused primarily on subsidising industrial deployment of broadband: 'superfast' in urban areas and 'standard' in rural areas [5]. Crucially, this approach addresses infrastructure barriers without addressing economic or social: pricing is left to the market.

One means of addressing the cost of access is to share the WiFi of existing broadband connections utilising their unused capacity [6]. Systems that share home broadband in this way already exist, e.g., Fon [7], where subscribers' existing broadband connections are shared by making available a wireless network accessible to anyone with a Fon account. These networks are typically provided either as a paid service or incorporated as part of an existing broadband subscription, enabling subscribers to

ACM DEV-5 (2014), December 5–6, 2014, San Jose, CA, USA.
Copyright 2014 ACM 978-1-4503-2936-1/14/12 ...$15.00.
http://dx.doi.org/10.1145/2674377.2674383.

[1] UK *social housing* is equivalent to American *public housing*.

access their network's broadband service via other subscribers' access points while roaming. However, when considering whether to *mandate* such sharing, there are a number of questions that remain unanswered: (*i*) is there spare capacity on existing domestic broadband links to do this without impacting service? (*ii*) how would this capability be used? (*iii*) would such a service be useful to the target demographic? (*iv*) are existing subscribers willing to share their capacity?

In this paper we report on a limited-scale feasibility study of a Public Access WiFi System (PAWS) [8] that answers these questions. The intent of PAWS is to explore provision of a *restricted service* that is *free at the point of use*, targeting demographics that want and need but cannot afford Internet access. As we are interested in the use of such a system in spatial and social context, we apply the "in the wild" method from HCI [9]: lab-based studies fail to show appropriation and use in context; surveys/questionnaires reveal only stated preference rather than actual behaviour [10]; and measurement studies fail to uncover the social challenges that are at least as important as the technical.

The intent behind this study is not to make broad statements about the use of a system like PAWS, but to uncover a rich understanding of a small number of users so as to sensitise us to the underlying challenges in deploying such a system. Thus we do not collect the data necessary to comment on broader questions such as how to incentivise sharers or what the commercial or operational impact would be on ISPs. Instead, we make the following contributions: (*i*) there are a number of technical challenges, principally signal reach and the variation in both home router availability and ISP configuration, which suggest that the density of deployment required for a system like PAWS to provide free Internet access for the digitally excluded is quite high and probably requires regulation to be effective; (*ii*) perhaps surprisingly, free access to the Internet is not the instant success one might expect, although some citizens do find even relatively limited access of considerable use; (*iii*) many existing broadband subscribers appear quite willing to share their bandwidth locally for the common good without requiring significant financial incentive.

The rest of the paper is organised as follows: we first present the social and technical context in which this study was carried out in (§2), we then describe the study setup, including the technical architecture and the recruitment process (§3). We then present analysis of data collected from the deployment (§4), followed by discussion of the social and technical challenges we uncovered through this study (§5). Finally we conclude in (§7).

2. STUDY CONTEXT

In-the-wild-research study enables a rich understanding of a complex, poorly understood context with many interdependent and unconstrained factors, rather than data to support a statistically significant hypothesis-based test, hence the relatively small sample sizes used in such work [11]. The results allow more effective, targeted design of larger scale trials due to an understanding of the important factors that must be taken into consideration in this context. In this case the novel technology is our community WiFi system, where existing broadband subscribers (*sharers*) make available an amount of their bandwidth to other local residents (*citizens*) via VPN connectivity over an open WiFi network. As a study of a socio-technical system there are, perhaps obviously, two elements to the context in which this study occurred: the social and the technical. Both have a critical bearing on the problems faced by PAWS and so on our design and deployment.

2.1 Social Context

Our deployment took place in Aspley, a moderate size council ward in Nottingham, a medium sized British city with a population of around 300,000. Aspley is a council estate and ward of the city, with a 2004 population estimate of just over 16,000 in around 6,280 households. It contains three large housing estates and was originally developed in the 1920s as a location to move people out of Nottingham's notorious inner city slums. Housing is primarily social with a small proportion in private ownership, and is predominantly multiple occupancy terraced or semi-detached. Aspley is one of the more deprived council wards in Nottingham and the UK, ranking in the bottom 10% nationally. It has one of the highest teenage pregnancy rates in Europe, dependent children in 45% households, and an unemployment rate of over 10% (around three times the national average and twice that of the city). The population age is skewed towards children, particularly aged 0–4 years, and away from people of retirement age.

Several relevant characteristics make Aspley a very suitable site for study of PAWS: (*i*) it has the highest levels of digital exclusion in Nottingham, and some of the highest in the country (28.5%, compared to 8% in the more affluent neighbouring ward); (*ii*) it is one of the 10% most deprived wards nationally on several measures (e.g., skills, employment, income and crime); and (*iii*) it has a disproportionately young population, the most likely to cite affordability as a barrier to Internet access. At the same time, most citizens have access to WiFi-enabled devices: it is known that 52% of 16–24 year olds and 23% of the skilled working class, working class and non-working class now have smartphones. This proportion is growing rapidly: in both groups 65% obtained their phone in the last 12 months [3].

To summarise then, our participants are in an economically deprived urban area with some existing penetration of broadband but a relatively high proportion of non-adoption. Based on the socio-economic classification of the population in Aspley we are primarily concerned with, using terminology from Horrigan [4], the 'near converts' and 'digital hopefuls' where cost is the main barrier to adoption, rather than the 'digitally uncomfortable' and 'digitally distant' where skills and perceived relevance are the main barriers.

2.2 Technical Context

The access technology provided by PAWS is standard 802.11b/a/g/n WiFi – its ubiquity among home broadband deployments (providing sufficient opportunity and capacity for sharing) and client devices (providing sufficient opportunity for it to be used) makes it the only reasonable choice. Again, Nottingham City Council survey data from 2012 indicates that 20% of the population of Aspley is without access to a device that supports Internet access, including smartphones, tablets, televisions and games consoles.

Before beginning the deployment we sought up-to-date data concerning the broadband penetration visible over WiFi in Aspley. Survey data from 2012 suggests penetration of 78.8% but this could not tell us about the geographic distribution of visible access points (APs), nor how strong was the signal from those APs. We performed a 'war-drive' to sample WiFi coverage in Aspley to determine whether or not we needed to focus recruiting in particular areas.

The war-drive sampled approximately 40% of the streets within Aspley, and found 1,067 unique APs (by MAC address).

Density of APs was not concentrated in any particular area, and was sufficiently high that we felt there was no need to focus recruitment in particular areas. Using observed Broadcast Service Set Identifiers (B-SSIDs), around 23% customers were with Sky (an ADSL and fibre provider), 23% were with BT (both a retail and a wholesale ADSL provider), and 21% were with Virgin (a fibre provider). Of the BT customers 61% were advertising the Fon service and the remainder were not. The remaining 33% customers appeared to be using user-specified B-SSIDs and so we could not easily determine their ISP.

A more thorough sampling of B-SSIDs (e.g., different periods of the week, different times of the day, greater coverage throughout Aspley) would have given a more complete picture, but the sample we have indicates that, even in a relatively deprived area, a reasonably large number of customers are now on contracts likely to have spare capacity.

3. PAWS

We developed and deployed PAWS to explore the requirements of and challenges inherent to provision of a free-at-the-point-of-use restricted Internet service. We now present the study configuration: the design of PAWS, based on the requirements this context engendered, and the recruitment process we used to access participants.

3.1 Recruitment

We recruited members of this community through door-to-door initial contact and by placing advertisement posters in public areas such as the local community centre. Initial contact by either means was followed up by a face-to-face visit (in most cases) or phone call (in cases where the engagement did not require on-site visit, and where on-site visit was inconvenient for the participant). Participants who continued with the project were contacted by phone periodically to check on their engagement with PAWS, and all were interviewed on completion. For their time and the inconvenience, participants were offered compensation in the form of shopping vouchers in line with standard ethical practice when studying human behaviour; the scale of such compensation was limited so as to avoid providing a direct monetary incentive to participate. The size of compensation was scaled in line with the inconvenience suffered: sharers who made their broadband available to others, requiring installation of a device in their homes and more extensive follow-up interviews, received higher compensation (£100) than those citizens who simply used the PAWS service (£50) and only had to participate in a small number of in-person and telephone interviews.

We recruited 98 (13.4%) sharers and 36 (4.9%) citizens of 730 respondents from just over 2000 houses approached. However, some were unsuitable for deployment due to spatial considerations (the need to locate sharers and citizens within WiFi range of each other), and others dropped out for a range of reasons including going silent, acquiring broadband through other means, or simply changing their minds with no reason given. We deployed two PAWS nodes in public spaces (a local community centre and a local church, though placement of the node in the church meant it had very poor range and so was never actually used). We also implemented a self-signup website enabling citizens to register without requiring a site visit, though a follow-up phone call was made within a few days, resulting in a further 54 citizens signing up. Of these, 38 never activated their account and so never actually made use of PAWS. The end result was that we deployed 18 domestic sharer nodes plus 2 sharer nodes in public spaces, and observed 18 active citizens, of which just 2 were recruited and the others were self-signups.

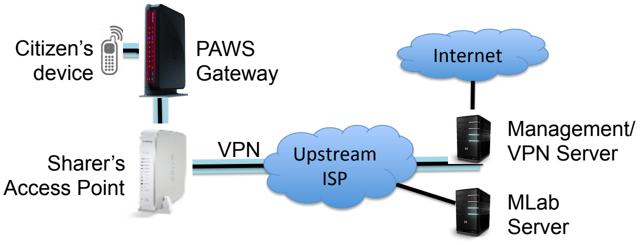


Figure 1: PAWS Architecture.

3.2 Technical Requirements

This context, a deprived area with corresponding comparatively low levels of education and low familiarity with the Internet, imposed several technical constraints on PAWS.

Accessibility. The service had to be accessible to all without imposing any additional costs in terms of devices. Only standard, commonly deployed protocols could be used.

Simplicity. The entire system had to be simple to use for both citizens and sharers, requiring little effort to sign-up and minimal effort to maintain engagement.

Security. Deployment required both *privacy* and mutual *authentication* (i.e., authentication of the citizen by the system, and of the system by the citizen).

3.3 Technical Architecture

In combination these requirements led us to implement PAWS using VPN technology accessed via an open B-SSID. WiFi is virtually ubiquitous in client devices (phones, tablets, laptops, desktops) and most of these devices have built-in support for VPN protocols (albeit with varying degrees of success and simplicity). Use of a VPN technology meant that we could meet the security requirements, authenticating citizens when accessing the network and enforcing per-citizen accounting. Figure 1 depicts the overall organisation of PAWS.

Hardware. We deployed Netgear WNDR3800 routers as gateways in sharers' homes. These contain Atheros AR7161 rev 2 680MHz chipset with 128MB RAM and 16MB Flash running a custom build of OpenWRT and were remotely managed by a device management server hosted in a local datacenter. Each gateway is connected via a wired Ethernet port to the sharer's home broadband router, and advertises an unsecured WiFi network with SSID PAWS on its own Virtual LAN (VLAN) at 2.4GHZ with auto-channel selection. Each access point is configured to bridge only the wireless interface to the wired interface, and offers IP addresses to associated devices via DHCP from RFC1918 address space.

Infrastructure. To support deployed gateways we use a single server hosted in a local datacenter. This provides both device management functionality, where each gateway registers with the management service and is controlled via a reverse SSH tunnel, as well as the VPN server endpoint and citizen authentication and signup service, including per-device connection instructions. The signup service allows users to self-signup and start using PAWS straight away, as well as providing password reset and reminder features. We also rely on the Bismark [12] infrastructure which involves measurement probes in the OpenWRT image running on the gateways and the M-Lab service [13] that acts as an endpoint measurement server and measurement data collection service.

Bandwidth Management. We apply a rate limit to total citizen use on a per-gateway basis to ensure we can give simple guarantees to sharers about the possible impact on their own Internet service. The PAWS gateways throttle traffic between the PAWS WLAN and the VPN server using Linux's *tc qdisc* hierarchical token bucket queuing discipline: all PAWS traffic is assigned to the default class throttled to 2Mb/s download and 512kb/s upload.

Traffic Management. Each gateway implements firewall policies that permit control protocols like DHCP, ICMP and DNS, and block all other traffic except to the device management server, the VPN server and the measurement server. Any web (HTTP/HTTPS) requests that are blocked are redirected to a page on the project web site giving information about PAWS and signup instructions. A key aim for this study was to uncover the uses that a system like PAWS might be put to and so we chose not to restrict access to external websites or services. We discuss this issue further in §5.

Security. Given the nature of the service PAWS provides and its target user group, it was incumbent upon us to address several potential security threats as we could not assume the citizens would know how to protect their traffic:

(*i*) *Protecting the citizen from nefarious activity by the sharer*. We deliberately target the digitally excluded, an audience likely to be particularly unfamiliar with the details of wireless networking and thus potentially vulnerable. Specifically, it would be inappropriate (and arguably unethical) for us to teach them that it was generally safe to connect to an open WiFi network and carry out private and personally sensitive activity on the basis of an advertised SSID alone. We were thus concerned to protect citizens from simple attacks such as a sharer presenting an open PAWS SSID and thus intercepting their traffic. To mitigate this threat, we enforce use of an encrypted, authenticated end-to-end VPN connection between the citizen's device and the PAWS-provided Internet. This ensures privacy for the citizen's traffic as well as providing at least some surety for the citizen that they are genuinely connected to the PAWS network.

(*ii*) *Protecting the sharer from nefarious activity by the citizen*. The PAWS system permits free access to the Internet, albeit at restricted rates. As we could not be sure how the network would be used we needed to protect the sharer from the charge of providing open network access without the required auditing and control. Using a VPN in this way prevents the sharer being aware of citizens' use of their network, and mitigates the risk of them being charged because their broadband connection is used inappropriately.

(*iii*) *Protecting ourselves from nefarious activity by the citizen*. Much the same threat applies to us, the researchers, as providers of this service. The VPN served a third purpose which is to prevent citizens from using the network anonymously – all access is via a session with the VPN server, so we can log precisely who is using the PAWS network, what they are accessing, at what time. This mitigates the risk to us of citizens using the network inappropriately, as well as allowing us to collect data for analysis.

3.4 Deployment

We initially deployed only a PPTP VPN service as PPTP is widely supported by end devices, is supported by most middleboxes, and is relatively simple to configure at the client device. However, early deployments indicated that some devices, particularly older Android releases, have flawed implementations of PPTP preventing them from establishing connections with the server. We thus later added support for L2TP as well: although slightly more complex to configure, it appears to be more reliably implemented.

ISP	# Gateways	
	Total	Measured
Sky	6	5
Virgin	10	8
Orange	1	1
TalkTalk	1	1
Griffin	1	0
Tiscali/PIPEX	1	0

Table 1: ISPs of all participating sharers, and of those that permitted us to enable performance measurement. The two public space nodes were with Tiscali/PIPEX and Griffin.

User Management. We built a simple centralised web application to manage user accounts and provide credentials for VPN authentication via RADIUS. This enabled citizens to sign-up without direct intervention from us (self-signup), an option many chose particularly in the public access areas; in the end all bar one of the active citizens using PAWS were self-signups. To maintain the accountability of the use of PAWS, we required a contact number as part of the process and followed up each self-signup with a phone call within a few days to verify their details. Any accounts found not to be genuine were cancelled at that point.

Measurement. Using the *whois* service on the public IP address used by each PAWS gateway, we mapped PAWS gateways to ISPs. Most sharers allowed us to carry out daily measurements of the throughput, latency and loss experienced by their broadband network, enabling us to characterise their home broadband performance. Table 1 indicates the ISPs of all sharers, along with those who permitted us to carry out performance measurements which are part of the standard Bismark measurements suite [12]:

- *Throughput*. We calculated upload and download throughput (Mb/s) using *netperf* every six hours using three parallel TCP threads to provide an accurate estimate of access link capacity [14].
- *End-to-End RTT (e2ertt)*. We collected round-trip-time (RTT) measurements (ms) to a set of servers in the UK once every 10 minutes using *ping*.
- *Loss*. We measured loss rates at both upload and download using D-ITG [15] at 15 min intervals.
- *Availability*. We measured the availability of the PAWS service (i.e., of the combination of the PAWS gateway and the sharers' access links) every minute by sending a 60 byte UDP probe/minute to our management server.

We also measured citizens' usage of PAWS using RADIUS accounting records and packet captures for each VPN session. We present analysis of the various data we collected in the following section.

Finally, there were two phases to our deployment. The first, Jul–Nov/2013 involved all sharers plus citizens recruited and self-signed-up during that period. The second, from Nov/2013–Feb/2014, involved only the active citizens already participating plus any further citizens who self-signed-up. Due to a lack of sufficiently nearby citizens, the bulk of the sharers simply provided measurements of their ISP performance. As the data

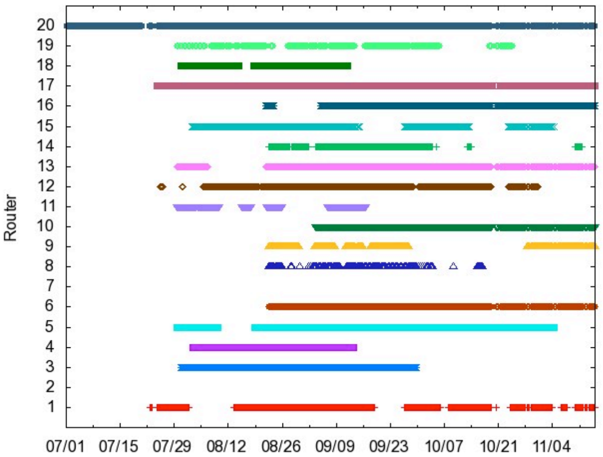


Figure 2: Availability of each PAWS gateway. Gateways 5, 6, 8, 12 ,13, 14, 17, 20 are in use by citizens. Gateways 2 and 7 were off for almost the entire period.

gathered about ISP performance during the first phase was so stable, there seemed little need to extend the participation of those sharers acting solely as measurement nodes. In contrast, the success of the self-signup process at the community centre meant that we could continue to observe new citizens on a rolling basis, allowing us to gather more extensive data about the actual uses to which PAWS was put.

4. DATA ANALYSIS

Certain ISP behaviours affected the measurement data we could collect. Specifically, it appears that Sky blocks ICMP traffic to the first hop within their network for all the Sky ADSL based access links we examined, and so we could not easily determine latency related measurements for those links. We also had issues conducting capacity measurements for one particular Sky ADSL link: we suspect the ISP was throttling the access link, perhaps due to a data cap. We only collected throughput data via 15 sharers: as the deployment progressed we had to cease collecting throughput measurement data from new sharers to prevent them falsely ascribing performance problems they experienced from our measurements to citizen behaviour and thus risking them withdrawing from the experiment. Finally, we could not collect meaningful data from two gateways that were left unplugged most of the time (Routers #2 and #7 in Figure 2).

Nevertheless, the data we collected was rich and we now report results of its analysis. We begin with the Bismark data and its measurement of the sharers' home broadband capacity (§4.1) before looking at citizen usage (§4.2).

4.1 Analysis of PAWS Network

We analyse measurements from the sharer's access points to understand (*i*) gateway availability, as this will impact PAWS service availability; (*ii*) how much capacity is available to be shared and so (*iii*) whether there are broadband subscription packages that would be unsuitable for sharers; and (*iv*) whether effective sharing would require addition of more complex technology such as active queue management or other quality-of-service mechanism.

4.1.1 Availability

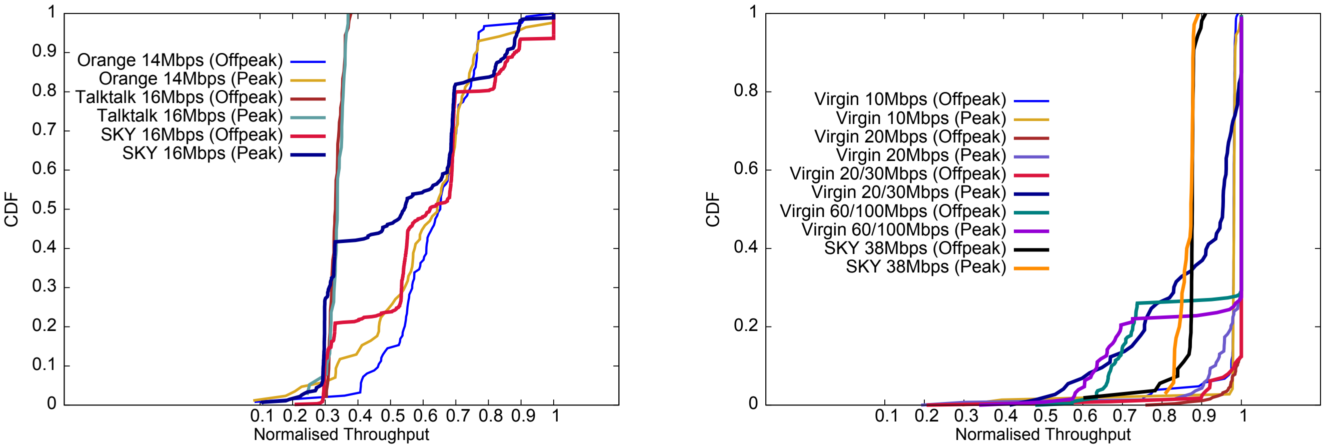
A system like PAWS requires that gateways be available for use particularly given that the access patterns we observed suggest a tendency for specific citizen–sharer pairs to form, rather than citizens making use of the service from any nearby sharer while roaming. Figure 2 presents the availability of the PAWS gateways over this initial deployment period. In this data we can see the reality of human behaviour imposing itself on PAWS. Disregarding one system-wide outage caused by an electricity supply failure in the area (around Oct. 21), no PAWS gateway was continuously available for the entire deployment period with some being turned off apparently at random and others being turned off with a somewhat regular period. When we followed up with sharers we discovered that the main reasons were: the limited number of wired Ethernet ports on home routers (often only one or two are provided) resulting in them unplugging the PAWS gateway from their home router so they could use the Ethernet port for something else; and either personal financial straits or general thriftiness resulting in them turning off *all* unnecessary electrical equipment overnight and when not home to save money.

4.1.2 Throughput

We next consider the access link capacity measurements, both download (from the ISP to home) and upload (from home to ISP). We define *peak* and *off-peak* times as 4pm–midnight and midnight–4pm respectively.

To compare the quality of ISP service between sharers subscribed on different packages, we normalise these values against the sharer's subscription. As part of the PAWS installation process each sharer was interviewed to understand their perception of the service they have from their ISP. Unfortunately, most did not remember either their subscription package or how much they were paying for it – some did not even remember who their ISP was! We thus assigned each sharer to the nearest package our capacity data indicated. For example, Virgin currently offers three packages with download limits at 30Mb/s, 60Mb/s and 120Mb/s, and has previously also offered packages with download limits at 10Mb/s, 20Mb/s and 100Mb/s. We thus matched the observed download throughput from Virgin customers to these values to determine their subscriber package. For those sharers who appear to have changed package during the measurement period, we label their data with all the packages they appear to have had, i.e., *Virgin 20/30* indicates a Virgin user who upgraded from 20Mb/s to 30Mb/s during the experiment. We report upload performance directly as we could not reliably match observed upload rates to subscriber package upload values, particularly for ADSL packages.

Figure 3 presents the normalised download throughput for the different subscriber packages, distinguishing between off-peak and peak hours. On the left, Figure 3a presents data for ADSL links. These exhibit substantial variation, achieving 80% of the advertised speed less than 20% of the time. In particular the link provided by TalkTalk, an ADSL-based provider, does quite badly, barely reaching 35% of advertised download capacity throughout, and the Sky links exhibit clear stepping behaviour. Latency and loss measurements (not reported in this paper due to lack of space) do not corroborate these effects being caused by significant and varying contention in the upstream network. We believe that, in the first case, low link speed is due to ADSL performance being dependent on distance from the DSLAM: subsequent data showed that this low speed was extremely stable and not accompanied by high loss, or high or variable latency as one might expect if low



(a) ADSL providers (Orange, TalkTalk, Sky).

(b) Fibre providers (Sky, Virgin).

Figure 3: Measured peak *vs* off-peak download capacity across subscription packages normalised by estimated subscription value. Where the contract label indicates two speeds, measurements indicated an upgrade to the contract during the lifetime of the project.

ISP	min.	μ	σ	max.
Virgin 10	0.33	0.68	0.10	1.00
Virgin 20	1.01	1.12	0.03	1.18
Virgin 20/30	1.05	1.23	0.26	2.01
Virgin 60	2.45	2.94	0.10	3.00
Virgin 60/100	2.50	5.15	3.21	9.95
Sky 38	1.05	1.97	0.10	2.33
Sky 16	0.47	0.91	0.10	1.31
TalkTalk 16	0.72	0.87	0.02	0.90
Orange 14	0.03	0.60	0.27	1.31

Table 2: Per-subscription package upload capacity (Mb/s).

performance was due to heavy multiplexing; and in the second, the stepping behaviour is being caused by step-changes in the multiplexing behaviour of the ISP's network.

On the right, Figure 3b presents the same measurements from the fibre networks. The first observation is that the throughput observations meet the estimated contract values notably better: all fibre-based links achieved 80% of their advertised download throughput at least 75% of the time. Virgin appears to do best, though the Virgin 20/30Mb/s package sees a significant reduction in download throughput during peak hours, only achieving the advertised download speeds 18% of the time compared to 85% of the time during offpeak hours.

Table 2 presents the upload capacities per contract which tended to vary more significantly. All ADSL based subscribers experienced upload capacity of less than 1Mb/s. The Orange 14Mb/s ADSL download contract exhibits significant variation in upload speeds, although the TalkTalk contract does not seem to. All fibre based contracts had upload capacities of 1Mb/s and greater, with Virgin 20Mb/s contracts achieving 1Mb/s upload, 30Mb/s contracts achieving 2Mb/s upload, 60Mb/s contracts achieving 3Mb/s, and the 100Mb/s contract achieving the highest upload speed, 10Mb/s. All fibre upload speeds were consistent, with no notable variation during the measurement period.

These results give us an indication that fibre customers are good potential sharers i.e. sharing a small portion of the network capacity by throttling the download speeds at 2Mb/s should not be an issue as all measurements exceeded 2Mb/s 100% of the time (even during peak hours) indicating there was sufficient capacity of atleast upto 2Mb/s which could be potentially shared. For e.g. Figure 4 presents the average loss rate and E2E RTT (log scale) of sharer of the primary citizen (§4.2), a Virgin 20/30 fibre customer. Although citizen activity became significant between Oct. 07 and Nov. 15 we can see no corresponding effect on either measured E2E RTT or loss rate for sharer. This suggests that, at least in this case, throttling download and upload rates to 2Mb/s and 512kb/s was successful in preventing citizen usage notably impacting the sharer. For ADSL based links where speeds reaching only 35%-50% of advertised speeds are common, sharing 2Mb/s may have an impact on the network. Similarly the rate at which we throttle the upload speeds may also be an issue as most of the upload speeds are limited to less than 1Mb/s (with the exception of the higher speed fibre links) hence sharing 50% of the upload capacity may not be an ideal solution especially with the current trend of uploading more user generated content. Although from the wardrive we were able to infer that fibre infrastructure accounted for atleast 25%, ADSL broadband is still considered less costly compared to fibre and more prevalent. Hence its mandatory for wider availability of PAWS, we need to take into account ADSL based sharers. This coupled with the need for better sharing of the upload capacity indicate that it is mandatory for us to enable better active queue management and/or QoS at the access points to ensure that traffic from PAWS citizens do not have any significant impact on the sharer.

4.2 Analysis of Citizen Usage

Having presented an analysis of PAWS' network performance, we turn to the use made of PAWS by citizens. We correlated packet data from *tcpdump* with RADIUS accounting records accounting use of PAWS back to citizen accounts. During the deployment a total of 36GB traffic (15GB upload, 21GB download) was generated by 18 citizens, with one citizen (the primary citizen) standing out as responsible for 28GB. From the HTTP headers, we observed that this citizen used both a Windows PC and an iPhone with PAWS. Six other citizens used Windows PCs connecting via the single PAWS gateway we deployed in a public space (a local community centre). Everyone else used a mixture of Android and iOS (iPad and iPhone) mobile devices.

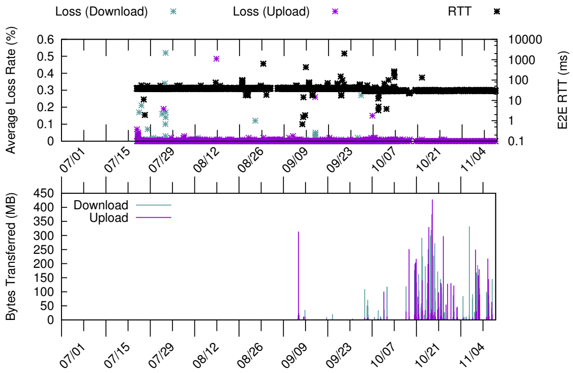


Figure 4: Primary citizen's usage (below) *vs* the average loss rate and E2E RTT (log scale) of their sharer (above).

We determine citizen usage by extracting all DNS request/response data alongside the destinations of outbound HTTP/HTTPS traffic from the *tcpdump* logs. By matching the IP addresses of the HTTP/HTTPS traffic to the DNS data we map traffic on a particular flow to service. If the *tcpdump* logs did not contain relevant DNS data, we performed a DNS reverse lookup while processing the data. We then categorise traffic to service based on analysis of domain names, as indicated in Table 3. This process accounted for 90% of all bytes and 85% of all flows.

Per-category proportions by bytes are summarised in Table 4. The most significant point to observe from all the usage data (Figures 5a and 5b) is that *advertising* dominates everything: of all traffic, around 60% (by bytes or sessions) is classified as advertising[2]. The dominance of advertising contrasts strongly with traditional examinations of backbone network usage, where streaming video tends to dominate, though it is less unusual in the context of mobile network usage [16]. We further examine use of PAWS splitting the traffic of the primary user from the rest: the primary user was responsible for over 75% (by bytes) of the total use of PAWS, perhaps unsurprising given they could access PAWS using desktop PCs from home, in contrast to the others who largely could only access PAWS occasionally, generally from the shared public space of the community centre. This allows the primary citizens to use PAWS much more freely, for more purposes, and at all hours of the day and night. We break down use by both hour-of-day and day-of-week: we find different patterns of use occur at different times of day but no noticeable distinctions in use between days of the week, including weekday *vs* weekend and hence this is not reported.

Category	**Sample keywords/netblocks matched**
Ads	doubleclick, 2mdn, advert, analytics
BBC iPlayer	iplayer, bbci
Bittorrent	torrent
CDN	edgecast, akamai, cdn
Games	game, mochi, nextgenhabbo, playfatal
Information	google, 173.194/16, facebook, fbcdn, yahoo, wiki, edu, .ac.uk, bbc., itv, channel4, telegraph, sun
Jobs	job, vacancy, reed, career, work
Porn	xxx, raunch, porn, strip [others]
Shop/Bank	tesco, asda, lloyds, natwest, halifax, ebay, amazon, gumtree
Skype	skype
Updates	avast, mcafee, microsoft, apple, norton
Youtube	youtube
Other	[unmatched names, IP addresses]

Table 3: Observed categories of use

Category	**% bytes transferred**
Ads	60%
Youtube	8%
Games	5%
Anticipated "legitimate" uses (Information, Facebook, Shopping, Banking)	4%
Porn	3%
Update	2%
CDN/Other	18%

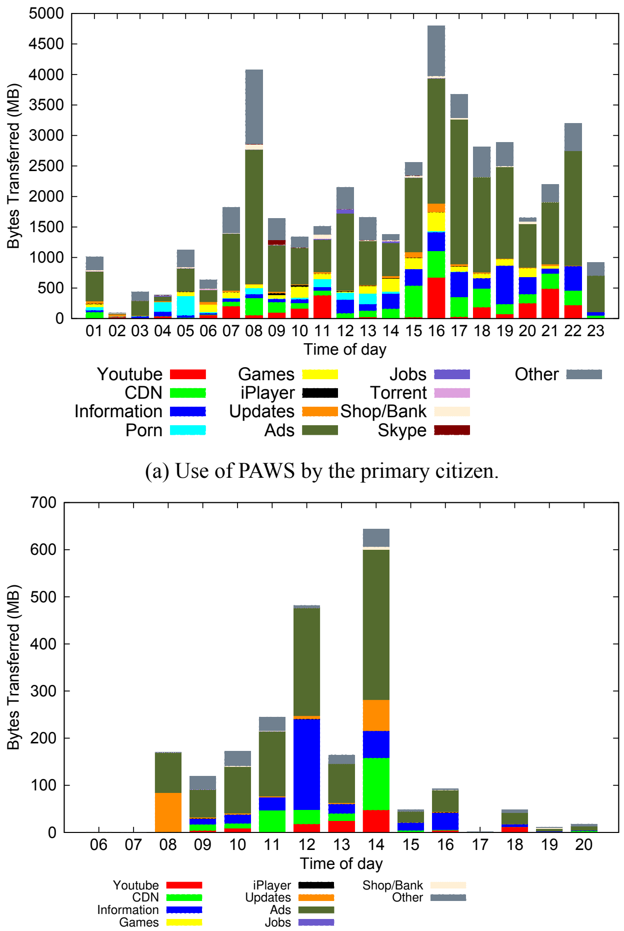
Table 4: Proportions of use.

In detail, considering first the traffic consumption of the primary citizen (Figure 5a), we see that they make quite extensive use of their sharer's network, primarily for entertainment (including porn), focused in the afternoon and evening but significant during the peak hours (8 am and in the evenings). The citizen also accesses more socially acceptable services: information sources such as Wikipedia, the BBC, other television news channels, newspapers, and Google are a noticeable contributor from mid-morning to early evening. Shopping and banking online, where the increased efficiency and availability of better value deals are often cited as key benefits that the digitally disconnected miss out on where also accessed.

Finally, we observe that the primary citizen accessed PAWS through two sharers, gateways #17 and #12 in Figure 2, both in range of the citizen's property. Sharer #1 was used almost 70% of the time to transport around 20GB of data, and Sharer #2 for the remaining 30% (8GB). Even given the relatively heavy and continuous use of Sharer #1 by the primary citizen, the sharer reported no problems or perceived performance impact and was happy to leave their PAWS gateway running for 89% of the time.

Turning now to the other citizens' use of PAWS (Figure 5b), we see a quite different pattern. Most noticeably, the total amount of use is dramatically less, just 23% of the total, and the hours of use are much more restricted, from 08.00–20.00 rather than 06.00–01.00. The relatively large burst of update traffic between 08.00–09.00 is likely due to the public access Windows PCs being turned on at that time when the community centre opens. At the same time, the uses to which PAWS is put are much more concerned with access to information through Google and Yahoo, viewing videos on YouTube and social networking on Facebook. Pornography does not feature at all – perhaps due to the public nature of the space.

[2] Perhaps giving the lie to the oft-repeated claim that "The Internet is for porn" (Avenue Q, The Musical). It actually appears to be for advertising in this context.



(a) Use of PAWS by the primary citizen.

(b) Use of PAWS by other citizens.

Figure 5: Use of PAWS

5. CHALLENGES

The purpose of this study was to sensitise us to the problems associated with providing a free-to-use Internet service aimed at connecting the currently disadvantaged. We make the following observations about the challenges we experienced in carrying out this research, all of which have implications for larger-scale deployment.

First, as is often the case, recruitment was significantly harder than we expected. Although things improved as we streamlined our recruitment and deployment process, we were surprised how many people were *not* interested in receiving free Internet access, even though they would be financially compensated for any inconvenience due to the research activities surrounding that provision (e.g., time taken by follow-up calls and interviews). This situation did improve somewhat once we deployed PAWS nodes in shared public spaces (specifically, the node in the local community centre), suggesting there is still a place for initiatives such as the provision of free Internet access via for e.g. libraries. It also seems possible that in these social contexts, where lifestyles are relatively chaotic, self-signup is a more appropriate means of recruitment than approaching people door-to-door.

Second, a particular difficulty in carrying out research like this is the geographical constraints placed on recruitment: citizens must be within range of sharers. The well-known vagaries of WiFi range and performance instability, coupled with the dynamics of the population in a community such as Aspley meant that this was relatively difficult to achieve. Citizens that the map suggested would be able to see sharers turned out to be unable to; citizens that expressed initial interest decided to drop out due, in many cases, to them obtaining their own broadband; and one sharer (sharer #2 for the primary citizen) had their electricity cut off for approximately 2 weeks part-way through the deployment leading to them cancelling their broadband contract in order to afford their electricity bills. The sharer then started using PAWS instead. This particular incident demonstrated that a system like PAWS could be an extremely useful alternative.

Finally, quotes such as “*"I’ll definitely say no. We’re all struggling." "Even if you got a hundred pounds for it?" "It wouldn’t matter if you pay me a thousand pounds. At the end of the day, if you can’t afford something, you shouldn’t have it."* ” demonstrate some of the social issues with attempting to engage in research like this. Others, e.g., “*"No – doesn’t sound safe [worried about security]"* ”, “*"I don’t understand any of it. It’s all foreign to me."* ”, and “*"Do you have the Internet?" "No" "Would you like to?" "No, that’s alright. I’m not interested." "May I ask why not?" "Cos it’s a load of fucking bullocks."* ” show that the benefits of Internet access are still not universally appreciated. They also suggest that the compensation offered did not significantly sway participation: the money offered was not sufficient to cause those who were uninterested to decide to take part, at least in those cases where responses were given.

Our deployment also shed light on four technical challenges that would need addressing for a larger-scale deployment, as well as providing evidence that one challenge we anticipated may not actually exist in practise.

The first two technical challenges were rather mundane. We chose to secure access to PAWS by relying on a VPN for authentication of PAWS to citizens and privacy of citizen data. However, based on the feedback we received from the self-signup users who created an account but did not use the system, first, the system PPTP VPN implementation on earlier Android devices was buggy and did not actually work, requiring us to add support for L2TP which has a more complex setup phase. Second, however, this is not really a satisfactory solution without providing customised client software as VPN setup was rather complex on all devices encountered, relying on individuals correctly maintaining and using multiple password credentials, something that even experts have trouble doing. Subsequent experimentation with per-device client software for OpenVPN indicates that auto-generating device configuration files on participant signup would have been a more effective strategy.

The third was an unanticipated interaction with the way that WiFi stacks on many current smartphones appear to behave, at least in some circumstances. We found that several sharers experienced difficulties using their own WiFi networks after becoming sharers, and it transpired that some of their devices were preferring to connect to the open PAWS SSID rather than their own WPA2 secured home network. This did not happen in all cases so we believe it had more to do with the specific configuration of those devices, but this effect did cause us to stop registering sharers as citizens as a result.

The fourth was anticipated but requires a larger scale trial deployment involving other stakeholders (e.g., government, ISPs) to satisfactorily address: citizens accessed several websites that would be arguably inappropriate to support over a national service such as PAWS. Given the dynamic nature of the Internet, this cannot be reasonably addressed using blacklists and so a clearer picture of precisely the services to be whitelisted needs to be built, although the data gathered about citizens' use of PAWS can certainly inform this, e.g., the need to provide access to software update sites along with those providing government services, social security, public information, etc. Such a solution would though raise serious ethical questions around the provision of a tier of Internet that is not only bandwidth limited, but also has its content limited according to the whims of a central body.

Finally, a challenge that we anticipated finding but which the data we gathered suggests is not significant is the need to apply Less-Than-Best-Effort (LBE) protocols in this context. Broadband users on fibre contracts do seem to have sufficient spare capacity that PAWS did not intrude on their own experience (certainly none of our measurements indicated such problems, nor did our sharers report any). However, given the poor quality of service that we observed ADSL users can experience, there may be a case for LBE protocols to be applied there. In such cases, effectively implementing this feature may require mapping IP-layer QoS down to the link-layer (e.g., by VLAN tagging to the DSLAM).

6. RELATED WORK

The Bismark [14] and Netalysr [17, 18] projects have performed extensive studies of the performance and other characteristics of home networks. Although we made use of MLab and the Bismark infrastructure in gathering these data, we are not trying to provide such large-scale statistically representative data. Rather, within PAWS we examined use and performance of a free-to-use broadband network over a reasonable period of time, in depth, and with the ability to engage with the individuals and their use of the network.

The problem of providing community broadband is well-studied from a technical point-of-view. Academic work has considered the construction of opportunistic networks using available bandwidth specifically using available home broadband [19]. In such cases a key concern has been to address problems such as reciprocity and incentives to share through technical means. For e.g., [20] demonstrates a decentralised approach based on indirect reciprocity using previously introduced algorithms. Members form a club and provide free WiFi access through their home broadband and in exchange are able to access controlled wifi sharing in populated cities. Users use signed digital receipts which are used to recognise contributing users in a reciprocity algorithm developed by the authors. Unlike PAWS, they do not have a central authority for registering or authentication which could lead to accountability issues. Another basic difference between PAWS and work including both the above and Fon, is that they only facilitate existing home broadband users while PAWS could enable access for all.

The problem of resource management and fair sharing in such contexts has been addressed using probabilistic load balancing techniques to give sharers priority over citizens [21]. A key part of the value of the in-the-wild approach used in PAWS is that it demonstrates that such problems may not actually occur in practice, obviating the need for solutions.

There have also been several previous community driven initiatives to provide community broadband [22] as well as academic initiatives that have served as research testbeds [23, 24, 25]. Among the better known academic initiatives that also took place in the UK are the Tegola project [26] which uses long distance wireless mesh networks to connect few communities in Scottish Highlands and the Wray project which also uses long distance wireless mesh to connect the Wray community [27].

Bristol Wireless [28] aimed to start a community wireless network by sharing a portion of a community centre's bandwidth with a number of local residents by using wireless technology. They identified and experienced a number of similar problems to PAWS: setting up a wireless network in an actual community was much easier said than done due to the need to recruit and obtain permission to enter homes to do installation and equipment configuration, finding people within wireless coverage range, dealing with practical issues such as rental properties where installation options are more limited etc.

These research testbeds use a combination of both long range wireless and mesh networking to serve rural regions. PAWS was particularly designed with urban digital exclusion in mind and is the only available testbed that utilises the notion of benevolence through sharing existing home broadband connections with fellow citizens to provide free Internet access. Thus PAWS provides an unique opportunity to understand not just the technical challenges but also as a research testbed to understand social and behavioural challenges. Although PAWS is similar to Fon, Fon is economically driven where homeowners share a part of their bandwidth with other users mainly other Fon members or those who purchase Fon credits.

7. CONCLUSIONS

This experimental deployment demonstrated the potential utility and viability of a Public Access WiFi Service. Even in the relatively complex form it was presented, users signed up and used it for legitimate, socially beneficial purposes. The primary citizen was clearly sufficiently happy to use the service extensively, and all citizens used it for a wide range of purposes: this suggests that the overall performance of the system was acceptable once past the interactional overheads of signing up and connecting to the VPN. We also observed that several users, not just the primary users, made use of PAWS for the legitimate (foreground) uses it was intended, such as banking and retail. They also brought to light the value of other (background) uses that we did not originally envisage, e.g., software updates.

We believe we can now partially answer the questions we posed: (*i*) many existing domestic broadband deployments do have sufficient spare capacity that sharing some of it would not be a problem for the owner; (*ii*) such a capability will be used in a range of ways, some more acceptable than others, suggesting there may be a need to manage the uses to which the shared bandwidth can be put; (*iii*) the service was useful to at least some of the target demographic; and (*iv*) given that the compensation we offered failed to sway people who refused to participate, and some sharers did state a willingness to share bandwidth simply for the common good without requiring any further incentive. It is important to emphasise that it was easier to recruit the sharers than the citizens. Further research is required to better understand the incentive structure for both the sharers as well as the network operators to enable a service such as PAWS. One plausible incentive structure would be for enabling third party stakeholders such as grassroot user communities or local government who may have a socio-environmental objective rather than purely economical to manage the PAWS service (where sharers could get a small council tax rebate for sharing their connection) while for the network operator they get paid (again) for leasing out the unused capacity (which has already been paid for) [29].

Finally, the ISPs clearly have a role to play in this. Traditionally they have argued against giving free access based on the claim that traffic on their backbone networks costs them money. However, networks are (typically) provisioned for peak use, and citizens using PAWS would not significantly increase peak use (they will not increase the number of access links into ISP networks, and there would be little spare capacity for them to use at peak times anyway). Thus they would increase off-peak but not peak use; any slight increase in peak use could be further mitigated through deployment of less-than-best-effort protocols from the access points into ISP networks, enabling ISPs to further degrade the free service if they found it necessary. We might anticipate regulation would have a part to play in controlling such degradation. There is evidence that some ISPs (e.g., AT&T[3]) are already applying such selective traffic grooming practices for commercial gain; PAWS would encourage them to do so for social benefit.

Acknowledgements.
This work was supported by the EPSRC grants EP/K012703/1, EP/G065802/1 and the EU CONFINE project (EU FP7 288535). We thank Srikanth Sundaresan, Nick Feamster, Heidi Howard and Achilles Petras for support and feedback.

8. REFERENCES

[1] Internet world statistics. http://www.internetworldstats.com/stats.htm.

[2] Cabinet Office. Digital by Default proposed for government services. http://www.cabinetoffice.gov.uk/news/digital-default-proposed-government-services, 2010. Accessed Dec. 10th, 2013.

[3] OFCOM. Communications Market Report: UK – Smartphone Data Tables (Adults). http://stakeholders.ofcom.org.uk/market-data-research/market-data/communications-market-reports/cmr11/, 2011.

[4] J.B. Horrigan. What are the consequences of being disconnected in a broadband-connected world? *Daedalus*, 140(4):17–31, 2011.

[5] Broadband Delivery UK, Broadband Delivery Programme: Delivery Model, 2011. UK Gov't Dept. Culture, Media and Sport: London.

[6] A. Sathiaseelan and J. Crowcroft. Lcd-net: Lowest cost denominator networking. *SIGCOMM Comput. Commun. Rev.*, 43(2):52–57, April 2013.

[7] Fon. http://www.fon.com.

[8] A. Sathiaseelan, J. Crowcroft, M. Goulden, C. Greiffenhagen, R. Mortier, G. Fairhurst, and D. McAuley. Public access wifi service (PAWS). In *Digital Economy All Hands Meeting*, 2012.

[9] Y. Rogers. Interaction design gone wild: Striving for wild theory. *interactions*, 18(4):58–62, July 2011.

[10] I. Deutscher. *What We Say/What We Do: Sentiments & Acts*. Scott, Foresman., Glenview, ILL, December 1973.

[11] A. Crabtree, P. Tolmie, and M. Rouncefield. "How Many Bloody Examples Do You Want?" Fieldwork and Generalisation. In *Proc. European Conference on Computer Supported Cooperative Work (ECSCW)*, pages 1–20. Springer London, September 21–25 2013.

[12] Project BISmark. http://projectbismark.net.

[13] M-lab. http://www.measurementlab.net.

[14] S. Sundaresan, W. de Donato, N. Feamster, R. Teixeira, S. Crawford, and A. Pescapè. Broadband internet performance: A view from the gateway. In *Proc. ACM SIGCOMM 2011*, pages 134–145. ACM, 2011.

[15] A. Botta, A. Dainotti, and A. Pescapè. A tool for the generation of realistic network workload for emerging networking scenarios. *Computer Networks*, 56(15):3531–3547, 2012.

[16] N. Vallina-Rodriguez, J. Shah, A. Finamore, Y. Grunenberger, K. Papagiannaki, H. Haddadi, and J. Crowcroft. Breaking for commercials: Characterizing mobile advertising. In *Proc. ACM Internet Measurement Conference (IMC)*, pages 343–356. ACM, 2012.

[17] C. Kreibich, N. Weaver, B. Nechaev, and V. Paxson. Netalyzr: Illuminating the edge network. In *Proc. ACM Internet Measurement Conference (IMC)*, Melbourne, Australia, 2010.

[18] C. Kreibich, N. Weaver, G. Maier, B. Nechaev, and V. Paxson. Experiences from netalyzr with engaging users in end-system measurement. In *Proc. ACM SIGCOMM Workshop on Measurements Up the Stack (W-MUST)*, Toronto, Canada, 2011.

[19] E. C. Efstathiou, Pantelis A. Frangoudis, and G. C. Polyzos. Controlled wi-fi sharing in cities: A decentralized approach relying on indirect reciprocity. *IEEE Trans. Mob. Comput.*, pages 1147–1160, 2010.

[20] E.C. Efstathiou and G.C. Polyzos. Trustworthy accounting for wireless LAN sharing communities. In *Proc. EuroPKI*, pages 260–273, 2004.

[21] I. Psaras and L. Mamatas. On demand connectivity sharing: Queuing management and load balancing for user-provided networks. *Comput. Netw.*, 55(2):399–414, February 2011.

[22] M. Mandviwalla, A. Jain, J. Fesenmaier, J. Smith, P. Weinberg, and G. Meyers. Municipal broadband wireless networks. *Commun. ACM*, 51(2):72–80, February 2008.

[23] J. Bicket, D. Aguayo, S. Biswas, and R. Morris. Architecture and evaluation of an unplanned 802.11b mesh network. In *Pro. International Conference on Mobile Computing and Networking (MOBICOM)*, pages 31–42. ACM, 2005.

[24] J. Camp, J. Robinson, Ch. Steger, and E. Knightly. Measurement driven deployment of a two-tier urban mesh access network. In *Proc. International Conference on Mobile Systems, Applications and Services (MOBISYS)*, pages 96–109. ACM, 2006.

[25] B Raman and K. Chebrolu. Experiences in using WiFi for rural internet in india. *Comm. Mag.*, 45(1):104–110, January 2007.

[26] G. Bernardi, P. Buneman, and M.K. Marina. Tegola tiered mesh network testbed in rural Scotland. In *Proc. ACM Workshop on Wireless Networks and Systems for Developing Regions (WINS-DR)*, pages 9–16. ACM, 2008.

[27] J. Ishmael, S. Bury, D. Pezaros, and N. Race. Deploying rural community wireless mesh networks. *IEEE Internet Computing*, 12(4):22–29, 2008.

[28] Bristol wireless. http://www.bristolwireless.net.

[29] A. Sathiaseelan, C. Rotsos, C.S. Sriram, D. Trossen, P. Papadimitriou, and J. Crowcroft. Virtual public networks. In *2nd IEEE European Workshop on Software Defined Networking (EWSDN)*, 2013.

[3] http://wired.com/threatlevel/2014/01/att-sponsored-data/