Measuring Trust in Wi-Fi Hotspots

Tim Kindberg
Hewlett-Packard Laboratories
Filton Rd, Stoke Gifford, Bristol BS34 8QZ, UK
timothy@hpl.hp.com

Eamonn O'Neill, Chris Bevan, Vassilis Kostakos, Danaë Stanton Fraser, Tim Jay
University of Bath, Bath BA2 7AY, UK
{E.O.Neill, C.R.Bevan, V.Kostakos,
D.StantonFraser, T.Jay}@bath.ac.uk

ABSTRACT

Pervasive systems provide services that are situated within specific contexts. An everyday example of this is Wi-Fi hotspots. Factors such as branding and presentation are known to affect whether users are prepared to invest trust in services, but little is known about trust in situated services. This paper describes an experiment to measure de facto trust in Wi-Fi hotspots in public places, as opposed to examining trust behaviour in a simulated lab setting. We investigated two hypotheses about the effect of location-specific images in the hotspot's pages on trust behaviours, compared to images of non-specific locations. We found a significant result which confirms that decisions to access an unfamiliar Wi-Fi hotspot can be affected by location-relevant images.

Author Keywords
Trust, Wi-Fi, Phishing, Security, Privacy, Pervasive Computing, Field Study Methodology.

ACM Classification Keywords
H.1.2 [User/Machine Systems]: Software Psychology. D.4.6 Security and protection

INTRODUCTION

As mobile devices with built-in wireless connectivity continue to proliferate, so do services that offer wirelessly transmitted content. In particular, researchers have been investigating *situated* services, which are embedded within particular locations [17]. A widespread present-day example of these are Wi-Fi hotspots, which are restricted to use only in the region of the cafés and other places that provide them. Situated services raise issues of privacy and security, and pervasive computing researchers have been investigating ways of reducing the potential for abuse using methods such as 'phishing', where electronic communications from trusted vendors are mimicked for malevolent purposes.

Trust is an important factor when considering privacy and security, since, on the one hand, users may mistakenly trust a malevolent but apparently trustworthy service and thereby open themselves to attack; and, on the other hand, they may distrust a bona fide service and thereby miss out on its benefits.

The present study investigated trust investment behaviour by deploying spoofed Wi-Fi service provision ('hotspots') in cafés, and gathered data about how users responded to them. A difficulty with investigating trust in general is that trust behaviour will be influenced by the setting in a traditional laboratory-based experiment [16]. In addition, we are interested specifically in situated services within urban environments. Thus, our first goal was to develop an experimental methodology by which we could invoke measurable trust behaviour from participants 'in the wild' who were unaware that they were participating in a controlled experiment. While difficult to achieve, the objective of this methodology was to gather data about *de facto* trust behaviour in public places, rather than trust behaviour that was possibly influenced by a lab setting.

Our second goal was to test two hypotheses about the presence within the Wi-Fi hotspot's introductory web pages of highly salient photographs that represented or did not represent the user's current location. Specifically, it was first hypothesised that an image representing the location would increase the likelihood of the user trusting the website enough to supply personal information in the form of his or her mobile phone number, when compared to the same website displaying an image of a location that did not represent the user's current location. By including an image of the location as a salient evidential cue – a *locative cue* – of the Wi-Fi service, we hypothesized that uncertainty about the source of the service would be reduced through 'anchoring' the service to the venue where it was deployed.

In addition, we investigated the converse of that hypothesis, that an image of a location that was specifically *unlike* the current location – an *anti-locative cue* – would increase uncertainty and thus decrease the likelihood of the user trusting the hotspot. In fact we found evidence supporting that second, anti-locative hypothesis, but not the first. This paper's contribution is an account of our novel experimental methodology for investigating trust in situated services, and an analysis of our first findings with respect to locative images.

CHI 2008, April 5-10, 2008, Florence, Italy.

BACKGROUND AND TRUST HYPOTHESES

This section first describes the phenomenon of Wi-Fi phishing, and then describes related work on trust behaviours in order to motivate our hypotheses about the effects of locative and anti-locative cues. Finally, it describes related work on methodologies for the measurement of trust needed to test those hypotheses.

Wi-Fi 'Phishing'

'Phishing' is the practice of attempting to acquire personal or sensitive information fraudulently through electronic communications [1]. Phishing communications are often engineered to appear trustworthy by closely resembling real communications from trusted online vendors both in terms of their branding and their visual appearance. Common examples of phishing attacks involve mimicking the on-line communications of the banking/financial services, and such attacks have been shown by recent experimental research to be worryingly effective [10, 11, 27].

With the growth of Wi-Fi service provision in urban areas, the potential for fraudulent abuse by 'phishers' is increased. A Wi-Fi phishing attack is relatively straightforward to mount via a laptop carried discreetly by the attacker. Like any other hotspot, the spoofed hotspot appears in the list of available wireless networks, and as a set of web pages when the user first tries to connect to it. An attacker who can convince an unwitting user to connect to the hotspot can, like any phishing website, capture data from the user and supply bogus responses carrying malware. But the threat from a Wi-Fi phisher is worse in that she can control all the users' network services, and thus act as a 'man in the middle', more easily than on the wired internet. A man in the middle can capture all data sent in the clear to or from the user and, if the user does not check certificates, can even capture data encrypted using Transport Layer Security (TLS) and Secure Socket Layer (SSL). The source of a wireless service is neither easily ascertained, nor verifiable. Thus, there remains a degree of trust required on the part of a user as to whether a particular Wi-Fi service is indeed what it purports to be.

Initial Situational Trust

Initial situational trust refers to ad hoc trust investment decisions based on limited knowledge and/or limited prior experience with the trustee, such as that found when encountering an unfamiliar Wi-Fi hotspot where some form of personal information is required of the user before the service can be used. This is in contrast with the 'basic trust' that refers to the normative degree of trust we place in everyday realities such as gravity [15]. The concept of situational trust can be described as 'a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intention or behaviour of another' [25]. 'Another' may refer to another human (interpersonal trust), and/or to a system (institutional or system trust) [2, 20].

We can regard a given decision to invest trust as being, at least in part, the calculated outcome of an assessment of the level of perceived risk involved and the degree of uncertainty present (as to the potential outcome of the interaction). The tolerance levels of these factors are weighed against the benefits that would be lost should trust not be invested. Thus, while trust is typically irrational (the potential maximum losses are more than the potential maximum gains [9]) a decision to trust emerges when a subjective threshold is reached beyond which trust effectively becomes a rational choice [14, 22]. Any attempt by designers to engineer trustworthiness must therefore seek to reduce uncertainty, and the perceived risk involved from the perspective of the trustor [24].

Traditional models of trust development, particularly those from economics, assume that novel situations where a degree of trust is required are approached initially with suspicion. If such models are accurate, initial interactions between a user and a new, unknown system do not provide conditions conducive to cooperation based on high degrees of trust [21]. Instead, trust-based ventures of the type involved in an interaction with an unknown Wi-Fi network, where cooperation is required (such as the entry of credit card details for payment), are sometimes only achievable in the first instance based on known external factors such as generic controls applied to the interaction.

The security of the connection between the user and host is one such form of control, and web browsers indicate the security status of the connection to the user in a variety of ways. However, as Dhamjia et al. [10] found when examining 'why phishing works', a significant proportion of people (23%) do not look at either the address or status bar, or pay particular attention to browser-based security indicators (see also Sheng et al, [27]). When assessing a website on initial encounter, 23% of Dhamjia et al's participants determined legitimacy by the content of the site only, and a further 36% by content and domain name alone. Well-engineered spoof websites successfully fooled 90% of the people that they tested.

Within the content of a website, evidence of trustworthiness is sometimes provided by graphical logos and accompanying verification links to indicate the presence of third party institutional safeguards such as Verisign or 'Verified by VISA'. Additionally, the investment of an established brand (and by extension the established reputation) of a service provider can provide a powerful cue to trustworthiness insofar as it is assumed that the service provider wishes to protect their positive reputation.

However, there remains a significant problem for pervasive systems designers keen to maximise user acceptance of their services without the benefit of an established brand history or control structure support. If such external structures are absent, assessments of trustworthiness can only be made by the subjective assessment of the trustee's incentive to renege upon the initial trust investment, by identifying and evaluating any available cues as to their intentions.

Fogg et al [13] have highlighted the importance of perceived credibility of web sites, and have produced a set of guidelines as to how the perception of credibility can be improved. Their guidelines include such factors as 'real-world feel', 'ease of use', transparency of information and meticulous attention to detail in respect of any typographical and functional errors within the content of the interface.

The notion of 'real-world feel' is particularly relevant to situated services. The physical provenance of a service can be indicated in a variety of ways, such as including the provider's postal address or telephone number. Further studies by Fogg [13], Zheng et al. [30] and Steinbruck et al. [28] found that by presenting photographs of authors accompanying on-line articles, and of staff accompanying banking websites, user perceptions of trustworthiness can be increased by implying that the information source is both credible and attached to or sanctioned by the company who owns the website.

The use of facial cues as an indicator of underlying intentions and credibility is a common method used by humans during interpersonal trustworthiness assessment [3,29]. As Riegelsberger et al. [23] note, significant efforts have been made to evaluate the effectiveness of emulating aspects of face-to-face interaction as a trust-building cue within interface design. However results are thus far inconclusive [23].

A question arises, however: if the considered and careful use of face imagery does indeed increase user perceptions of trustworthiness by reducing the uncertainty about where the accompanying information has originated, could the use of proximate location-based imagery as a salient evidential cue increase the real-world feel and reduce the uncertainty about where a situated service has originated, and so increase the tendency to trust that service? Conversely, might dissonant location-based imagery increase the uncertainty, and thus decrease the tendency to trust the service?

We use the term *locative cue* in the context of a situated service. This is an image or other form of media that is embedded within the service's content, and which represents the location where the (source of the) service is situated. In order to understand whether trust behaviours are related to locative properties, it is also important to examine what is meant by the absence of a locative cue. The idea of an image not representing a given location can be broken down into one of the following mutually exclusive categories:

1. *anti-locative*: the image represents a place that is specifically *unlike* the given location; or
2. *a-locative*: the image represents a location that could be in any of many places (e.g. a generic picture of an English cottage), including the given location.

Other types of cue may also exist as to the trustworthiness of a given service, such as a known brand, a known or assumed reputation, and prior experience with the service. But, focusing on locative and anti-locative properties, and their likely effect on uncertainty, we formulated *locative and non-locative hypotheses* as follows:

Locative hypothesis: The presence of a locative cue increases trust in a Wi-Fi hotspot compared to the presence of an a-locative counterpart.

Anti-locative hypothesis: The presence of an anti-locative cue decreases trust in a Wi-Fi hotspot compared to the presence of an a-locative counterpart.

In each case, the comparison is with an a-locative cue, since that is presumed to have a neutral effect on uncertainty: an a-locative image is not inconsistent with the user's location, but it does not specifically represent that location.

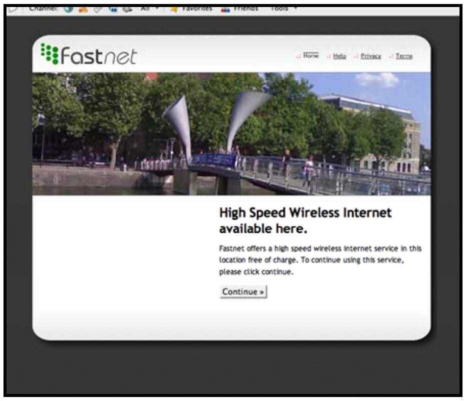
Methodological Issues: Working with Trust

"Risk, or meaningful personal investment, is a prerequisite of trust. The need for trust only arises in risky situations, and the trustor must be cognizant of the risks involved" Deutsch, [9].

Testing our hypotheses requires a measure of trust, but trust is a complex and multi-faceted concept that poses a significant challenge to researchers keen to isolate the phenomenon in a controlled experimental setting. As Deriaz [8] notes, notions of trust and risk are dissociable. This presents trust researchers with a methodological problem: how to create true risk in a controlled experimental environment. Experimental trust research, particularly in economics and the social sciences, has relied heavily upon the use of laboratory-based experimental trust games using variants of the 'Prisoner's Dilemma' protocol [4, 12, 7] such as the 'investment game' [5, 6] and 'trust game' [19].

Many such studies are based on monetary gambling, where a subject decides whether to trust that the other player will increase the trustor's pot through honouring an expected return, in the knowledge that the trustee is free to renege on that initial trust investment. Several studies have sought to involve the presence of risk by using real money and thus real potential loss. However, while at a conceptual level these games fulfill the criteria for trust (presence of risk and uncertainty, and a higher payoff for a successful trust investment over a non-investment), at a psychological level the lack of real world context may serve to reduce the perceived levels of risk involved to a degree where trusting behaviour in the lab is not indicative of trusting behaviour in reality.

Previous lab-based research into trust in situated payment mechanisms [18] investigated what trust-related statements emerged unprompted from interviews with participants, after demonstrations of the technology. Other lab-based research into the efficacy of phishing on the web examined trust behaviours by asking participants to assess whether or not they trusted a particular website [10]. But neither study measured *actual* investment of trust. By implying that some websites tested are not what they appear to be [10], a framing effect of suspicion is present before the participant even sees the first web page.



High Speed Wireless Internet available here.

Fastnet offers a high speed wireless internet service in this location free of charge. To continue using this service, please click continue.

Figure 1: Fastnet Splash Screen displaying the $L_{Bristol}$ Image Condition.

As Malhotra [19] notes, trust (a psychological state) and trusting acts (behaviours) are two different things: an apparent willingness to invest trust does not necessarily translate to an actual trusting action. As the investiture of trust can only be measured post hoc, we sought to develop a method for measuring trust through field experimentation where the experimenter is absent and the participant is unaware of his or her involvement. This kind of 'unattended' experimentation is designed to provide accurate data on de facto trust behaviours. On the other hand, the data we can gather is limited in that, with no experimenter present, we have no access to participants' reasoning. However, in investigating trust in situated digital services, we feel that the contamination effects of authority in lab-based trust research are a sufficient problem to warrant accepting this self-imposed limit on the data we can retrieve in the field.

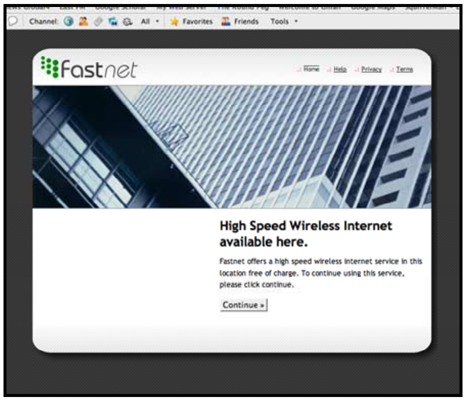
METHODOLOGY

This section describes the experimental design and then discusses the implementation and ethical issues that it raises.

Design

The experiment was based around a spoof Wi-Fi hotspot that we developed for public use. The hotspot was configured to appear in the list of available wireless networks as 'Fastnet'. When users attempt to connect to it, it provides a set of branded web pages purporting to be from Fastnet's wireless internet service. We installed the Fastnet hotspot in two public café locations in major U.K cities: Bristol and London. The venues selected were broadly comparable: each serves food and beverages with similar layouts; each has an existing Wi-Fi internet access point and established Wi-Fi user bases, with similar facilities for, and levels of, laptop use. The Bristol location is also licensed and is within a digital media centre, whereas the London location is close to a University. Consequently, on average, the Bristol location has a somewhat broader range of clientele and the London location a somewhat younger clientele.

In each city, we exposed users attempting to connect to the hotspot to a degree of apparent risk. We provided a pretext for entering their mobile phone numbers in order to connect to the 'service'. We chose mobile phone numbers because, on the one hand, they are personal information that carries a real potential risk of abuse and therefore requires a real trust investment, but, on the other hand, we were able to secure them to prevent any possibility of actual abuse. As explained below, we were able to record whether a user provided us with their bona fide number, and this user choice was the dependent variable. Not providing a number or providing a number other than the participant's own was counted as 'non-trusting' behaviour.



High Speed Wireless Internet available here.

Fastnet offers a high speed wireless internet service in this location free of charge. To continue using this service, please click continue.

Figure 2: Fastnet Splash Screen displaying the $NL_{Bristol}$ Image Condition.

To test our locative and anti-locative hypotheses, a between-subjects design was utilised. The independent variables were 'location' with two conditions (Bristol, London) and 'image' with two conditions ($L_{Bristol}$, $NL_{Bristol}$). Any user trying to connect to the hotspot was randomly but consistently assigned one of the two images. With the exception of this masthead image, the appearance and functionality of the website was identical.

Regardless of whether the hotspot was in Bristol or London, about half the users were presented with a masthead image $L_{Bristol}$ in the web pages (Figure 1). Image $L_{Bristol}$, which is of a scene immediately outside the Bristol venue, was chosen to be a locative cue for the Bristol location, and an anti-locative cue for the London location. The other half of the participants were presented with pages in which the locative cue was replaced by an a-locative image $NL_{Bristol}$, of the same size and in the same position on the pages

(Figure 2). That image is of a generic urban building which, we shall argue, is a-locative in both places.

The photographs used as masthead images were selected by means of a ranking exercise undertaken with members of the public at both venues. Participants (Bristol: n = $21_{[12 \text{ male, } 9 \text{ female}]}$) , London: n = $20_{[11 \text{ male, } 9 \text{ female}]}$) were asked to rank seven photographs in terms of the statement 'most clearly represents where I am now'. Three of the images were of the immediate area adjacent to the venue in Bristol, three were non-specific photos of urban-style architecture, and one was a 'wildcard' photo of a New York street scene.

The overall highest ranked image in Bristol was used as $L_{Bristol}$ and the second lowest ranked image in Bristol (disregarding the wildcard image) formed the $NL_{Bristol}$ image condition. We expected $L_{Bristol}$ to be anti-locative in London, since the bridge in this Bristol image is unique, and certainly there is nowhere like it anywhere near the London location. Results from London were consistent with this, since $L_{Bristol}$ was second *least* representative of that locale, ahead of only another photograph of the area around the Bristol venue, also with distinctive sculptural features and showing more water. Even the wildcard image of a New York street was ranked above those two in London.

Our $NL_{Bristol}$ image was second *most* representative in London, behind only another generic picture of a tall building, with different architecture. Despite its rankings near the opposite ends of the Bristol and London scales, image $NL_{Bristol}$ is a-locative in both places. It was deliberately chosen as a generic picture of a tall building; there are similar buildings in Bristol, London and most other major cities. The a-locativity of $NL_{Bristol}$ is consistent with the measured user responses in the ranking exercises used for image-selection, which, by their nature, are relative rather than absolute, and are to be taken in the context of what the images depict. Indeed, the foregoing characterisations of both images $L_{Bristol}$ and $NL_{Bristol}$ are consistent with the responses in the image-selection trials. One would expect users to rank an a-locative image above an anti-locative image and below a locative image. Given these characterisations, according to the locative hypothesis, the Bristol hotspot featuring $L_{Bristol}$ would elicit greater trust than that featuring $NL_{Bristol}$; according to the anti-locative hypothesis, the London hotspot featuring $L_{Bristol}$ would elicit less trust than that featuring $NL_{Bristol}$.

The participants in the study were 361 members of the public ($n_{[Bristol]} = 247$, $n_{[London]} = 114$), identifiable only by the unique MAC address of the devices they used to connect to Fastnet. Automatic MAC filtering performed immediately upon connection prevented a participant being phished more than once. Age and gender distributions were unknown.

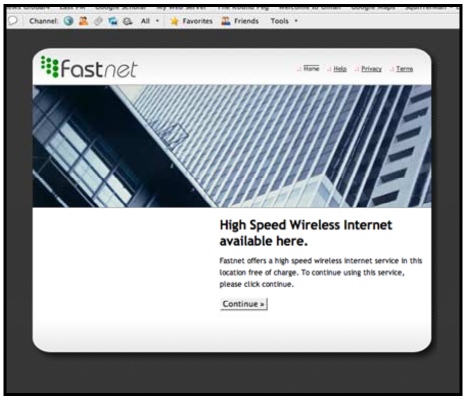
The Fastnet website

Discussing situational trust in terms of novel technology, Rutter [26] notes that when people approach a new experience, they tend to try to apply rules that have governed similar experiences in a similar domain. To this end, the site itself was deliberately minimalist in tone, content and colour, while adhering to design conventions of sites of a similar kind. To ensure maximum salience, around 50% of the active space of the website was devoted to the experimental image, and other imagery usage was minimal. Dhamjia et al. [10] found that 36% of participants tested on a variety of spoofed 'phishing' websites utilised the domain name as well as the content of the site when making judgements of site legitimacy. The use of an IP address as the URL was considered by their participants to be highly suspicious. To avoid this suspicion, our service presented the domain name 'www.fast-net.org' rather than an IP address.

The Fastnet website contained eight pages: four formed the 'login' process, three offered help and information on the use of the site. Instances of multiple logins from the same device, or attempts to submit a previously used mobile phone number were redirected automatically to a 'blocked' page. 'Blocked' participants were barred from hyperlink/direct URL access to any of the other pages. The path through the login process was forced: 'home' → 'login' → 'password' → 'thankyou'. Attempts to jump steps (e.g. through direct URL entry) were automatically redirected to the 'home' page.

Step 1: Splash Screen – 'Welcome to Fastnet'

Upon connection to the Fastnet server, the MAC address of the connecting device was recorded and used to assign the participant to one of the two image conditions. In the event of a repeat visit, the same device would always be presented with its originally assigned image.



High Speed Wireless Internet available here.

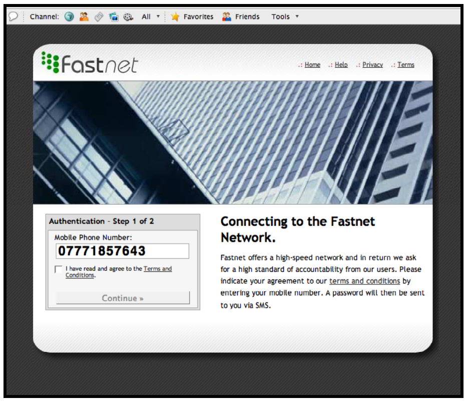
Fastnet offers a high speed wireless internet service in this location free of charge. To continue using this service, please click continue.

Figure 3: Fastnet Access: Step 1: Splash Screen.

When a web browser was opened, participants were automatically presented with an introductory splash screen that introduced our 'free wireless internet gateway service' offer and details of how it could be accessed (Figure 3).

Any attempt by the user to bypass Fastnet (e.g. through direct URL input) resulted in a redirect to the splash screen.

Step 2: Login – 'Please Supply Your Mobile Phone Number'
Participants who chose to continue were then asked for their mobile phone number in order to access the network (Figure 4). An explanation that the service offered 'high speed' access in return for a degree of accountability on the part of its users was provided as the reason for this process.



Authentication - Step 1 of 2

Mobile Phone Number: 07771857643

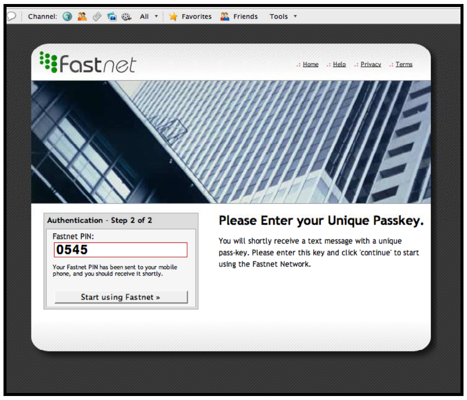
□ I have read and agree to the Terms and Conditions.

Connecting to the Fastnet Network.

Fastnet offers a high-speed network and in return we ask for a high standard of accountability from our users. Please indicate your agreement to our terms and conditions by entering your mobile number. A password will then be sent to you via SMS.

Figure 4: Fastnet Access: Step 2: Login.

Step 3: Authenticate – 'Please Enter Your Unique Passkey'
Upon submission of a valid mobile phone number, a unique personal identification number (PIN) was sent to the participant's mobile phone using the Short Messaging Service 'SMS', via a phone attached to the server.



Authentication - Step 2 of 2

Fastnet PIN: 0545

Your Fastnet PIN has been sent to your mobile phone, and you should receive it shortly.

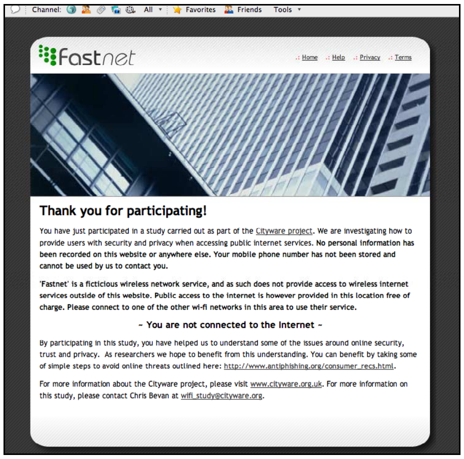
Please Enter your Unique Passkey.

You will shortly receive a text message with a unique pass-key. Please enter this key and click 'continue' to start using the Fastnet Network.

Figure 5: Fastnet Step 3: Authenticate.

The website then informed them that they would shortly receive the PIN, and that they would need to enter this PIN into the website in order to complete their authentication and start using the service (Figure 5).

Step 4: Debrief
When the correct PIN had been entered and submitted, the experiment ended and the participant was informed about the experiment (Figure 6). Since the server was not connected to the internet, we were not able to provide an internet connection at that point.



Thank you for participating!

You have just participated in a study carried out as part of the Cityware project. We are investigating how to provide users with security and privacy when accessing public internet services. **No personal information has been recorded on this website or anywhere else. Your mobile phone number has not been stored and cannot be used by us to contact you.**

'Fastnet' is a ficticious wireless network service, and as such does not provide access to wireless internet services outside of this website. Public access to the internet is however provided in this location free of charge. Please connect to one of the other wi-fi networks in this area to use their service.

~ You are not connected to the Internet ~

By participating in this study, you have helped us to understand some of the issues around online security, trust and privacy. As researchers we hope to benefit from this understanding. You can benefit by taking some of simple steps to avoid online threats outlined here: http://www.antiphishing.org/consumer_recs.html.

For more information about the Cityware project, please visit www.cityware.org.uk. For more information on this study, please contact Chris Bevan at wifi_study@cityware.org.

Figure 6: Fastnet Step 4: Debrief.

Implementation Issues
The chosen methodology entailed significant practical difficulties. As an experimenter was not present, our participants had to engage in the experiment without instructions. This posed a problem in that many mobile devices will automatically connect to a previously used network if it is detected again. As both our experimental venues have existing Wi-Fi networks to which devices might automatically connect, we were obliged to accept that engagement in our experiment would be limited by the participant's inclination to discover and connect to our service. Consequently, substantive participation required months rather than days or weeks. This timescale had an additional effect upon our implementation. Our system had to be robust enough to withstand a significant period of sustained activation. Error capture, recovery and access control were all critical factors in our design. In order to provide continual service over several months, our web servers were deployed on notebook computers running a UNIX-based operating system for maximum stability, especially where power sources are potentially variable. (The system can generally survive accidental power outage, and recovery can be performed by non-technical persons without difficulty.) Additional USB external fans were attached to both machines to provide additional cooling. Each web server transmitted a report of its status via SMS to the research team on a daily basis. In addition to providing us with an early warning of a system problem, it also supplied a daily report of usage and the number of 'phished' participants.

To secure our participants' mobile phone numbers, we sent them to the server from their mobile devices via an encrypted 'HTTPS' connection. To protect the participants' privacy further, we stored the secure hashes of the numbers on the server, rather than the numbers themselves. The stored hashes were used to prevent a malicious user from entering someone else's number repeatedly. The server itself was protected from external attack by being connected only to the local wireless network and not to the internet.

The unattended nature of our study also demanded a system that could cope with a myriad of connecting devices and operating systems. The site itself was designed to comply with W3C xHTML web standards, and was subject to an exhaustive testing schedule with Safari, Internet Explorer and Firefox for Macintosh, and Internet Explorer, Firefox, and Opera for PCs, in addition to the proprietary browsers of several types of mobile phone and PDA.

Ethical Issues

Ethical approval for the study was applied for and gained. The British Psychological Society (BPS) guidelines on ethical experimental design are clear that the interests of the participant must take precedence over the interests of the experimenter. From an ethical standpoint, the use of deception in unattended studies must therefore be handled with a much higher degree of care than conventional experimentation, insofar as prior consent cannot be obtained and neither can the standard practice of post-experimental debrief be observed. We addressed this issue with a clear textual debrief at the end of the experiment (when the participant was phished), including details of the experiment, the security of the mobile phone number supplied during the experiment, and the contact details of the experimenter and project staff.

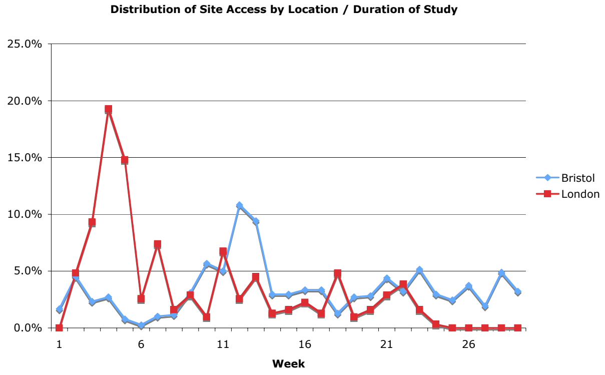


Figure 7: Distribution of Fastnet Site Accesses Recorded over the Duration of the Study.

RESULTS

The results presented in this section were generated through a combination of the server's raw web log and additional logs generated by the Fastnet system itself, collected over a period of 29 weeks between October 2006 and July 2007.

Patterns of Site Access and Instances of phishing

The distributions of unique MAC connections and phishing events recorded on Fastnet for each of our two locations are presented in Figures 7 and Figure 8 respectively.

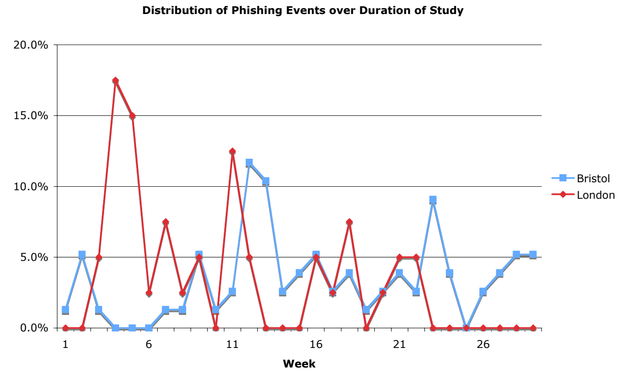


Figure 8: Distribution of Phishing Events Recorded over the Duration of the Study.

The general phishing success rate (independent of location) was approximately 32%. Around half (53%) of the participants who failed to be 'phished' exited the site without progression past the splash screen. 29% exited at 'login', 12% at 'password' and the remaining 6% left from one of the three 'help and information' pages. 80% of participants made only one visit to Fastnet. The spike in phishing rates noted in London between weeks four and five coincided with the period leading up to Christmas 2006 when it is assumed that the venue (being adjacent to the city shopping district) encountered much higher than usual customer traffic.

Location	Image	Total Participants	Phished	Not Phished
Bristol	NL[Bristol]	122	36 (29.5%)	86 (70.5%)
	L[Bristol]	125	41 (32.8%)	84 (67.2%)
	Total	**247**	**77 (31.2%)**	**170 (68.8%)**
London	NL[Bristol]	59	26 (44.1%)	33 (55.9%)
	L[Bristol]	55	13 (23.6%)	42 (76.4%)
	Total	**114**	**39 (34.2%)**	**75 (65.8%)**
Total		**361**	**116 (32.1%)**	**245 (67.9%)**

Table 1: Phishing Success Rates by Location and Image.

The Effect of Location and Locative Cue

A participant receiving the debriefing page indicated a successful phishing event. Raw hit counts for the debrief page were calculated for all participants by location/image condition and are presented in Table 1. As any given participant could be phished only once, any repeat hits (incurred by a page refresh) made by the same participant on the debrief page were removed from subsequent analysis.

Comparative phishing rates across the location and image conditions are presented as a plot in Figure 9. A three-way loglinear analysis was conducted, in order to determine the effect of location and image type on phishing rates. Results showed no significant main effect of either location

($x^2(1)=0.407$, $p=0.523$ n.s) or of image type ($x^2(1)=2.639$, $p=0.104$ n.s). However, there was a significant interaction between location and image type ($x^2(1)=4.886$, $p=0.027$), whereby phishing rates were lower for image type $L_{Bristol}$ and higher for image type $NL_{Bristol}$ in the London location only.

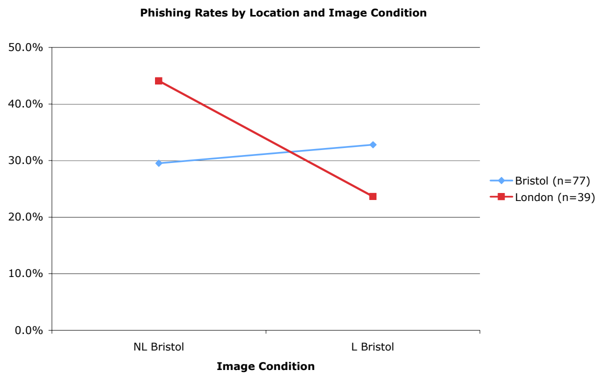


Figure 9: Proportional Phishing Rates by Location and Image.

To further investigate the interaction between location and image and their effect upon phishing success, additional chi-square tests were performed on the image and phished variables separately for the locations Bristol and London.

For London, there was a significant association between the image shown and the likelihood of subsequent phishing, $x^2(1)=4.41$, $p=0.036$ (using Yates' continuity correction); this was not true of Bristol, $x^2(1)=0.177$, $p=0.674$ n.s (using Yates' continuity correction). Odds ratios indicated that participants in London were 2.55 times more likely to be phished when presented with the image $NL_{Bristol}$ than if presented with image $L_{Bristol}$.

The analysis revealed a difference in the incidence of phishing for Bristol and London: while participants in Bristol were equally susceptible to being phished regardless of the image displayed, participants in London were much less susceptible to being phished when presented with image $L_{Bristol}$ (the image selected as a locative cue in Bristol) than with $NL_{Bristol}$ (the image selected as being a-locative in both locations).

DISCUSSION

The first major finding of this study is the high rate at which participants were phished, attesting to the vulnerability of Wi-Fi hotspot users in public places. In both London and Bristol, irrespective of the image used, about a third of people exposed to our spoof Wi-Fi hotspot trusted it with their mobile phone number, so even more might trust a hotspot that required nothing of them. A real attacker could not only have abused their phone number, but could also have observed all data sent to or from their machines, and installed malware on them.

The second finding is the evidence for location as a trust-relevant attribute in images presented to Wi-Fi hotspot users. The result of the experiment – that in London image $L_{Bristol}$ led to significantly less trust than $NL_{Bristol}$ – provides evidence to support the anti-locative hypothesis, but does not support the locative hypothesis. This outcome is illustrated in Figure 10. While somewhat inconclusive, this evidence is compelling enough to warrant further investigation. The questions our study has raised include: Are the chosen locations indeed equivalent for the purposes of the experiment, and how should such equivalence be determined? Are the distinctions we have made between locative, anti-locative and a-locative images valid according to users' perceptions – and are they categorical or do images fall in a continuum of 'locativity'? Last but not least, was the significant difference in trust investment in London definitely due to the anti-locative properties of the images, or might some other attribute have created this effect?

Implications for Wi-Fi provisioning

The results of this experiment have implications for the design of situated services such as Wi-Fi. Designers need (a) to protect consumers from *mistakenly trusting* spoofed services, and (b) to avoid *distrust* as a barrier to use of legitimate services.

Taking mistaken trust first, the fact that so many users entered their mobile phone number into our spoofed hotspot suggests that Wi-Fi providers should consider protecting their users. If users are prepared to follow them through, then protection mechanisms exist. For example, Wi-Fi providers could issue users with a challenge and response in the form of a slip of paper containing a random string to type into the hotspot's website, and the expected response. But users would have to follow the instructions assiduously, and it is not clear that they would do so because of the inconvenience. An attacker could, for example, provide a 'service granted' page immediately after entry of the challenge, relying on the user not noticing that they did not receive the expected response.

Turning to distrust, the evidence from this study suggests that, to avoid putting off some users, an a-locative image may be best. Wi-Fi providers are unlikely to use an anti-locative image deliberately, but some users might take an intended locative image to be anti-locative, if they are not sufficiently familiar with their surroundings.

Implications for experimental methodology

Although the foregoing findings are tentative and require further investigation, their existence validates the most significant contribution of our research: the experimental methodology that led us to them. The main features of our methodology are, first, that users are exposed to what, as far as they know, are real risks (in this case, the abuse of their telephone numbers); and secondly, that experimenters, who otherwise might influence the outcome, are absent. The

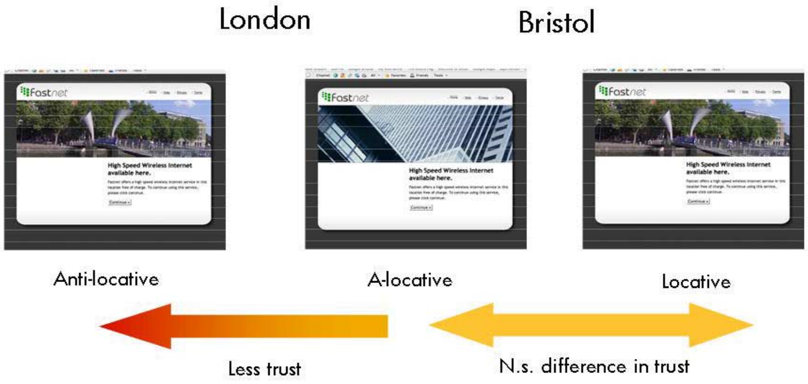


Figure 10: Support for 'anti-locative' hypothesis.

advantage of this set-up is that the experiment produces a measurement of trust as expressed through actual trust behaviours rather than merely through assertions of trust: we measured *de facto* susceptibility to an attack versus *de facto* avoidance of it. This methodology is applicable beyond Wi-Fi provisioning. Another example of a situated service is one invoked by reading a 2D barcode with a camera phone in a particular situation [17]. How is the user's readiness to trust such a service related to factors such as the locativity of the content printed around the tag, the initial content delivered from the tag, and *in situ* factors such as whether the tag is fixed to its position? Our methodology will enable us to investigate the effects of locativity by comparing barcode instances that differ only by those selected features.

On the other hand, a weakness of our method derives from the same factors underlying its strength: since we cannot engage with the participants, we cannot control or distinguish certain factors that might improve the accuracy of our measurement. In particular, some participants may have cut short interaction with the spoof site, not because they distrusted it but because they became distracted or didn't like one of its risk-neutral features. Other, more qualitative measures were also unavailable, including an analysis of which specific aspects of locativity or anti-locativity affected trust decisions. It was not possible to obtain feedback via the web site because non-phished participants could not be alerted to the true nature of the exercise and because, for security reasons, phished participants were not connected to the internet during engagement with the spoof site. Perhaps unsurprisingly, although we provided those who were phished with other ways of contacting us, we received no communications. In future experiments we will make it possible for participants to respond conveniently, and conduct discreet *in situ* interviews of participants and non-participants.

Finally, this type of methodology increases the onus on researchers to protect their participants. Stringent measures were taken to ensure that participants' phone numbers would not be abused in any way. Moreover, we were conscious of the lack of any immediate way of responding to the participants, should they have had concerns about their experiences.

CONCLUSION

This paper has described an experiment to measure the trust that users do or do not place in situated services, specifically Wi-Fi hotspots. The results tend to support the anti-locative hypothesis: that those exposed to an anti-locative cue are less likely to trust the service than those exposed to an a-locative cue. Both this result and the methodology we used to obtain it, in which we strove to measure de facto trust in the field rather than asserted trust in the lab, are novel.

In future work, we will be further investigating the effects of locativity and anti-locativity, and strengthening the method as outlined above and in the following ways. First, we will select candidate images in the light of a more developed analysis of 'locative cues', investigating further the distinctions between locative, anti-locative and a-locative content, and choosing images that are comparable and testable in better-understood ways. Moreover, we will investigate the notion of the salience and range of location-based cues. Secondly, we will extend the experiment to more locations, in order further to test the locative and anti-locative hypotheses.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers, Ava Fatah gen. Schiek, the Watershed Media Centre, Bristol, and Gordon's café, London. This research is funded by the UK Engineering and Physical Sciences Research Council grant EP/C547683/1 (Cityware: urban design and pervasive systems).

REFERENCES

1. 'Phishing'. Sourced from Wikipedia. http://en.wikipedia.org/wiki/phishing, accessed 05/03/07.

2. Abdul-Rahman, A., Hailes, S. Supporting Trust in Virtual Communities. Proc. of the 33rd Hawaii Conference on System Sciences, 2000.

3. Adolphs, R. Trust in the Brain. Nature Neuroscience Vol. 5(3), pp. 192-193, 2002.

4. Axelrod, R. The Evolution of Cooperation. Harper Collins, 1984.

5. Berg, J., Dickhaut, J., McCabe, K. Trust, Reciprocity and Social History. Games and Economic Behaviour. 10, pp. 122-142. 1995.

6. Cox, J.C. How to Identify Trust and Reciprocity. Games and Economic Behavior. 46, pp. 260-281. 2004.

7. Dasgupta, P. Trust as a Commodity. In: Gambetta, D. (Ed.), Trust. Making and Breaking Cooperative Relations. Basil Blackwell, Oxford, pp. 49–71, 1988.

8. Deriaz, M. What is Trust? Position Paper. Available at Http://scholar.google.com/url?sa=U&q=http://cui.unige.ch/ASG/publications/TR2006/5whatisTrust.pdf.

9. Deutsch, M. The Resolution of Conflict: Constructive and Destructive Processes. Yale University Press, 1973.

10. Dhamija, R., Tygar, J.D., Hearst, M. Why Phishing Works. Proc. of the SIGCHI Conference on Human Factors in Computing Systems, Montreal, Canada, 2006.

11. Downs, J.S., Holbrook, M.B., Cranor, L.F. Decision Strategies and Susceptibility to Phishing. Proc. of the Second Symposium on Usable Privacy and Security, Pittsburgh, Pennsylvania. 2006.

12. Flood, M.M. Some Experimental Games. Research Memorandum RM-789. RAND Corporation, Santa Monica, CA, 1952.

13. Fogg, B.J., Marshall, J., Kameda, T., Solomon, J., Rangnekar, A., Boyd, J. & Brown, B. Web Credibility Research: A Method for Online Experiments and Early Study Results. CHI2001: Extended Abstracts, 1-6 April, The Hague, The Netherlands, pp. 295-296, 2001.

14. Gambetta, D.(ed). Trust: Making and Breaking Co-Operative Relations. Blackwell, 1988.

15. Giddens, A. The Constitution of Society. Cambridge Press, 1984.

16. Jakobsson, M., Ratkiewicz, J. Designing Ethical Phishing Experiments: A Study of (ROT13) rOnl Query Features. In Proceedings of the 15th international Conference on World Wide Web, Edinburgh, Scotland, May 23-26, 2006.

17. Kindberg, T., and Barton, J. A Web-Based Nomadic Computing System. In Computer Networks, Elsevier, vol. 35, no. 4, March 2001, pp. 443-456.

18. Kindberg, T., Sellen, A., and Geelhoed, E. Security and Trust in Mobile Interactions: A Study of Users' Perceptions and Reasoning. in UbiComp 2004.

19. Malhotra, D. Trust and Reciprocity Decisions: The Differing Perspectives of Trustors and Trusted Parties. Organisational Behaviour and Human Decision Processes. 94, pp. 61-73. 2004.

20. McKnight, D.H., Chervany, N.L. Trust and Distrust Definitions: One Bite at a Time. Trust in Cyber-Societies: Integrating the Human and Artificial Perspectives. Springer Berlin, 2000.

21. McKnight, D.H., Cummings, L.L., Chervany, N.L. Initial Trust Formations in New Organisational Relationships. Academy of Management Review. Vol. 23(3), pp. 473-490, 1998.

22. Nooteboom, B. Trust: Forms, Foundations, Functions, Failures and Figures. Edward Elgar, 2002.

23. Riegelsberger, J. and Sasse, M. A. Face it - Photos Don't Make a Web Site Trustworthy. In CHI '02 Extended Abstracts on Human Factors in Computing Systems ACM Press, 2002.

24. Riegelsberger, J., Sasse, M.A., McCarthy, J.D. The Mechanics of Trust: A Framework for Research and Design. Int. Journal of Human-Computer Studies, 62, pp 381-422, 2005.

25. Rousseau, D.M., Sitkin, S.B., Burt, R.S., Camerer, C. Not so Difficult After All: A Cross-Discipline View of Trust. Academy of Management Review, 1998.

26. Rutter, J. From the Sociology of Trust Towards a Sociology of E-Trust. Int. Journal of New Product Development and Innovation Management. Pp. 371-385, 2001.

27. Sheng, S., Magnien, B., Ponnurangam, K., Acquisti, A., Cranor, L.F., Hong, J., Nunge, E. Anti Phishing Phil: The Design and Evaluation of a Game that Teaches People not to fall for Phish. In Proceedings of the 3rd Symposium on Usable Privacy and Security, Pittsburgh, Pennsylvania, July 18-20, 2007.

28. Steinbrück, U., Schaumburg, H., Duda, S., and Krüger, T. A Picture Says More Than A Thousand Words - Photographs As Trust Builders In E-Commerce Websites. 2002 CHI2002 Conference Proceedings, pp748-749, 2002.

29. Winston, J.S., Strange, B.A., O'Doherty, J., Dolan, R.J. Automatic and Intentional Brain Responses During Evaluation of Trustworthiness of Faces. Nature Neuroscience, Vol. 5(3), pp. 277-283, 2002.

30. Zheng, J., Veinott, E., Bos, N., Olson, J. S, and Olson, G. M. Trust without Touch: Jumpstarting long-distance trust with initial social activities. CHI2002 Conference proceedings, pp. 141-146, 2002.