ARTICLE IN PRESS

Computer Communications xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Context classifier for position-based user association control in vehicular hotspots

Pedro M. Santos*,[a], Leonid Kholkine[a], André Cardote[b], Ana Aguiar[a,c]

[a] Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias s/n, Porto 4200-465, Portugal
[b] VENIAM, Rua dos Heróis e Mártires de Angola 59, 4th floor, Porto 4000-285, Portugal
[c] Instituto de Telecomunicações, Rua Dr. Roberto Frias s/n, Porto 4200-465, Portugal

ARTICLE INFO

Keywords:
Association control
Mobile device
Vehicular hotspot
Contextual classification

ABSTRACT

Unintentional associations of mobile devices to on-board WiFi access points (APs) can affect the outdoor Internet experience of mobile device users, as the on-going cellular connection is broken and a short-lived WiFi connection is initiated. This disruption of the user experience can be avoided if the on-board AP learns whether the user device is inside or outside the bus and decides to accept its connection request or not. In this article, we present a classifier-based mechanism for on-board APs that accepts or denies user device associations based on a classification of the relative position of the device. An analysis of the problem in terms of connection duration and RSSI is presented to motivate the selected approach. We then describe a classifier to identify the user relative position trained on features extracted from contextual information. The classifier was trained with a large dataset of real-world WiFi-usage and mobility patterns of a public bus fleet from Porto, Portugal. The training procedure indicated bus speed as the most relevant feature, and that the RSSI measured at the on-board AP does not contribute. Finally, we propose a mechanism that grants or denies connection access to users based on the classifier output. We discuss how to integrate this mechanism in the AP network stack and evaluate its performance in real-world tests. Our solution can avoid 40% of the associations from users outside of the bus.

1. Introduction

Hotspots with WiFi access points (AP) and a connection to the cloud (via 3G or DSRC) are becoming common in public transportation fleets to provide Internet service to passengers. For ease of connection over multiple occasions, the on-board APs advertise the same Service Set Identifier (SSID) in all vehicles. In addition, many users of mobile devices tend to leave multiple wireless interfaces active when they are on the streets, specially WiFi and cellular. In case a user has already connected to the WiFi network of a bus, the network will be memorized by the user device and connected to every time the device discovers it, even if sometimes the user is not on the bus. In such cases, the cellular connection will be broken and a new WiFi connection will be established with the on-board AP, that may last for a very short period of time if the AP becomes out of range. Consequently, the latter connection is bound to cause a bad experience to the user as seamless Internet access is disrupted. Additionally, the association process also drains AP resources into a connection that will be inconsequential, thus potentially deteriorating the overall quality-of-service of the AP. As such, it is an *undesired* connection or association. Users to whom such accidental connections occur are typically also customers of the bus service (as they need to have used it once to have the network registered on their device), and higher customer satisfaction can be achieved if customers do not have their outdoor Internet experience disturbed by occasional buses passing by.

The undesired connection can be avoided if the on-board AP can learn whether the user is inside or outside the bus, and use that information to decide on whether to accept its connection request or not. There are several strategies to identify the user's relative position (if the user is inside or outside), but we seek a solution that is fully contained in the on-board AP. An alternative approach would be to request user input or action (e.g., logging in to a captive portal, installing a dedicated app). An AP-side solution has several benefits over this solution: (i) No human intervention is needed; (ii) the service provider needs not to rely on costumer awareness to achieve high adoption ratios of the solution; and (iii) the service provider can control the quality of the solution by rolling out new versions of the classifier and mechanism. Other non-AP-only options could pass by installing support hardware (e.g., Bluetooth beacons) or integrate with the ticket validation system, which would involve some deployment overhead and/or leave the

* Corresponding author.
E-mail addresses: pmsantos@fe.up.pt, pedro.miguel.santos@fe.up.pt (P.M. Santos), leonid.kholkine@fe.up.pt (L. Kholkine), acardote@veniam.pt (A. Cardote), anaa@fe.up.pt (A. Aguiar).

https://doi.org/10.1016/j.comcom.2018.03.004
Received 30 April 2017; Received in revised form 3 March 2018; Accepted 7 March 2018

service provider dependent on third parties.

In this paper, we propose a classifier-based scheme to support the association decision at the on-board AP using available contextual information. We first describe the problem with field experiments designed to characterize these connections and that motivate the need of a classifier-based approach. The classifier was trained using a real-world dataset of mobility and WiFi connection traces collected at buses equipped with on-board APs and GPS. The automated training showed that the most relevant contextual information is the instantaneous speed and speed average over last 10 s. Finally, we also propose a design for a decision algorithm that incorporates the developed classifier and discuss some options of its implementation in software and integration in the network stack. An experimental evaluation of the solution performance with respect to blocking outside users is presented: The classifier manages to block about 40% of external users at the instant of the first contact. In practice, this means that 40% of outside users are prevented of engaging in an undesired connection, thus not having their outdoor Internet experience affected; and that the AP can avoid 40% of inconsequential connection requests, thus freeing resources for valid connections.

Our contributions are:

- Characterization of undesired connections and the conditions in which they are likely to happen;
- Development of a classifier for predicting the relative position of the on-board AP and a user device using real-world datasets;
- Design and implementation of an association decision mechanism in an open network stack, and real-world experimental performance evaluation.

The remainder of this article is as follows. In Section 2, we outline the existing state of the art. A characterization of undesired connections is discussed in Section 3. The development of a classifier to detect the relative position of the user is explained in Section 4. The design of a mechanism to identify relative position and decide on connection acceptance or refusal is presented in Section 5. The experimental evaluation of the mechanism performance is shown in Section 6. Final remarks and future work are laid out in Section 7.

2. Related work

We review the existing literature on decision schemes for association and handover, and solutions for relative position identification. Recent works on association in vehicular scenarios focus on service to vehicular users by infrastructural access points. The work of Xie et al. [1] presents an algorithm to improve long-term service duration by infrastructural WLANs. The authors of [2] propose a load balancing-aware scheme to decide association of user devices to heterogeneous cellular base stations. In [3], an algorithm framework that provides analytical performance guarantees in scenarios of multi-tier multi-cell environments is presented. Handover decisions schemes extend association schemes by considering additional criteria such as quality-of-service and/or logistics of sustaining on-going sessions. Given its close relation to association and wide body of literature, we review also the relevant works in this field. A taxonomy of handover decision schemes can be found in [4], with the proposed categories being RSS-based, QoS-based, Decision Function, Network Intelligence and Contextual. Our approach identifies with the later class; and within these, it sits among the decision mechanism that harness mobility prediction. In [5] the authors propose a decision scheme for handovers between WiFi and WiMax networks that takes into account the user's speed. For high user speeds, the authors defend that the handovers from WiMax networks to WiFi should not be easily triggered due to the WiFi base station's smaller coverage. The authors of [6] propose a handover decision scheme between WLAN and WWAN based on user micro-mobility prediction. In [7], a network selection mechanism for LTE user devices is presented: it seeks the best network (LTE or WiFi) to support application QoS requirements, using external user mobility prediction services. Works addressing the scenario of association/handover to in-vehicle networks typically focus on QoS issues. In [8], a mobility-aware call admission control (CAC) algorithm is proposed: when a hotspot-equipped bus stops to let passengers in, WLAN guard channels are reserved to support handover sessions from users coming in. In [9] a hybrid interworking scheme is proposed to support seamless vertical handover of IP sessions for vehicular passengers. We found no proposals of association or handover decision schemes for our target scenario.

Solutions for learning the relative position or distance between nodes fall into two classes: those that assume active participation from both nodes, or those in which a single node infers it. In the first case, nodes typically share a common communication technology. WifiHonk [10] is a vehicle-to-pedestrian (V2P) WiFi-based mechanism to avoid collisions, in which vehicular users advertise their positions via beacons. The authors of [11] describe an implementation of a DSRC stack in a smart phone-grade WiFi chip. Both solutions source GPS to obtain the nodes' position estimates; if the actual distance between nodes is close to the GPS receiver's position error, the computed relative distance can suffer from a substantial error [12]. Solutions that try to identify another node's position (or distance to it) passively are mostly based on RSSI. RSSI-based methods for localization include lateration methods, machine learning classification, probabilistic approaches and statistical supervised learning techniques [13]. For static users, the work of Krumm and Horvitz [14] tries to infer user motion and location from WiFi received signal strengths. The authors of [15] use the RSSI of V2V messages to predict vehicle collisions. The work of Parker and Valaee [16] proposes a collaborative RSSI-based localization solution. The use of infrastructural nodes is proposed in [17], in which the authors use RSSI and angle-of-arrival from infrastructural APs to improve their vehicle's position estimate. However, the studies of Parameswaran et al. [18] and Heurtefeux and Valois [19] show that, even if the target node is static, measured RSSI is not consistent enough through multiple measurements sessions to support reliable ranging/localization.

The topic of decision schemes for in-vehicle network associations (or handovers from urban WLAN/WWANs) has not been explored in detail, to the best of our knowledge. RSS-based localization solutions are a natural approach to explore, but existing solutions require specific software and/or hardware. Our solution abstracts from this shortcoming by being designed to operate on the AP side and sourcing contextual information available to the vehicle. In addition, our proposed scheme protects the QoS of the users that rest *outside* of the vehicle (an aspect that we have not seen explicitly addressed in literature), and harnesses a real-world dataset with a large number of users to develop a generic and universal solution.

3. Undesired connections to on-board access points

We present an introductory analysis to the problem of undesired connections between user devices and on-board access points (APs). Our definition of an undesired connection is as follows: a connection established between a stationary or slow-moving user device (smart phone) that stands by the road side and an on-board AP that passes by or is stopped for a short period near the user.

We developed and conducted two experiments to characterize connections between a user device and on-board APs. The experiments address the following cases: (i) The user device is by the road-side and the vehicle passes by; and (ii) the user device enters the vehicle. The characterization is made over time in terms of: (i) RSSI throughout connection; and (ii) duration of connection and connection stages (particularly in the first experiment). In performing these experiments, we expect to identify metrics and/or process behaviours that may indicate whether a user is outside or not with some degree of certainty. We detail next the methodology and results of both sets of experiments.