A Survey on Methods for Broadband Internet Access on Trains

Daniel T. Fokum and Victor S. Frost

ITTC-FY2009-TR-41420-09

August 2008

Project Sponsor:
Oak Ridge National Laboratory
Abstract

We present a survey of approaches for providing broadband Internet access to trains. We examine some of the barriers that hinder the use of broadband Internet on trains and then discuss some of the opportunities for broadband deployment to trains. This survey considers some of the basic concepts for providing broadband Internet access and then reviews associated network architectures. The review of network architectures shows that we can subdivide networks for providing broadband Internet access to trains into the train-based network, the access network—for connecting the train to the service provider(s)—and the aggregation network—for collecting user packets generated in the access network for transmission to the Internet. Furthermore, our review shows that the current trend is to provide Internet access to passengers on trains using IEEE 802.11; however, a clear method for how to connect trains to the global Internet has yet to emerge. A summary of implementation efforts in Europe and North America serves to highlight some of the schemes that have been used thus far to connect trains to the Internet. We conclude by discussing some of the models developed, from a technical perspective, to test the viability of deploying Internet access to trains.
## CONTENTS

I  Introduction  1

II  Difficulties and Opportunities  2
   II-A  Difficulties .................................................. 2
   II-B  Opportunities .................................................. 4

III  Reference Architecture  5
   III-A  Other Architectures .......................................... 7
   III-B  Handoff Issues ............................................... 8
   III-C  Addressing Issues ............................................ 10

IV  Initial Concepts  10
   IV-A  FAMOUS Architecture and Its Extensions .................... 12
   IV-A1  Aggregation Network Design ................................ 12

V  Taxonomy of Architectures by Access Network Technology  16
   V-A  Radio-over-Fiber ............................................... 17
   V-B  Leaky Coaxial Cable-based Architecture ................. 17
   V-C  IEEE 802.11-based Architectures ............................. 18
   V-D  Satellite-based Architectures ................................ 19
   V-E  WiMax-based Architectures .................................. 20
   V-F  Architectures Based on Emerging Standards ............... 21
   V-G  High-Altitude Platforms for Internet Access ............. 21
   V-H  Internet Access through Direct Connectivity with Customer Terminals .......... 22

VI  Testbed Results  22

VII  Implementation Efforts  25
   VII-A  Implementation in Europe .................................... 25
   VII-B  Implementation in North America .......................... 27

VIII  Business Models and a Feasibility Study  28

IX  Lessons Learned  30
X Conclusion
Acknowledgments
References

LIST OF FIGURES
1 Architecture for Internet Connectivity between Rail Cars ........................................... 5
2 Reference Architecture for Internet Access on Trains .................................................. 6
3 FAMOUS Architecture for Providing Broadband Internet Access to Fast Moving Users, based on [8], [44] ................................................................. 12
4 Reference Architecture for Internet Access on Trains using Radio-over-Fiber, based on [5] . 18

LIST OF TABLES
I Abbreviations ................................................................................................................. 3
II Classification of Access Technologies ............................................................................ 23
I. INTRODUCTION

WITH the explosion in growth of the Internet in the last 20 years, people have a much higher expectation of being able to get on the Internet independent of location. Until recently trains and airplanes have been two locations where passengers have not necessarily been able to achieve high-speed Internet connections. In the particular case of trains, providing Internet access to passengers on board trains makes good business sense: Internet access for passengers can provide a revenue stream for the train company while attracting more travelers. For example, a 2004 study in the United Kingdom found that 72% of business travelers were more likely to use trains than cars or airplanes if Wi-Fi access was available on trains. This study also found that 78% of these business travelers would use Wi-Fi access if it was made available on trains [1]. In the case of freight trains, Internet access can allow for real-time or near-real-time tracking of freight-related events on board the train, potentially resulting in a decrease in insurance charges to the freight carrier. In addition to these benefits, broadband Internet access on trains can also enhance the safety of the train by allowing an operations center to monitor, in real-time, train-related data, as in [2].

Internet access on board trains is already available today in parts of Europe. For example, beginning in July 2004, a British train operator, GNER\(^1\), began offering Internet access on some of its trains [3]. In 2005 another British company, Nomad Digital, claimed to have addressed the problem of providing high-speed Internet access to passengers on Southern Trains’ London to Brighton route using WiMax [4]. In what follows, we provide an overview of communications on board trains, beginning with some of the earliest papers discussing broadband Internet access for users on the move.

The main contribution of this paper is to provide a survey of research and implementations that aim to make Internet access available on trains. The conditions of a rail environment that make communications from trains difficult are highlighted. For reasons that shall become apparent later, we make distinctions between work done in Japan, Europe, and North America, due to the different characteristics of rail transportation in those locations. The rest of this paper is laid out as follows: Section II lists the issues hindering high-speed communications from trains. Section III introduces a reference architecture for Internet access on trains, and also provides some context on handoff and addressing issues related to trains. Section IV provides discussion on initial concepts that have guided the deployment of broadband

\(^1\)GNER subsequently lost its license to operate the East Coast Mainline, where the Wi-Fi-enabled trains were deployed. National Express replaced GNER on the East Coast Mainline, and they offer free wireless (Wi-Fi) Internet access on all trains on the East Coast line.
Internet access to trains. In Section V we provide a taxonomy of technologies used to connect trains to the Internet. Section VI discusses results from testbeds that have examined how to deploy broadband Internet to trains. In Section VII we present the efforts made, or those efforts underway, to carry out high-speed communications from trains. Section VII is further subdivided into examining implementation efforts underway in Europe and North America. Section VIII presents business models developed to determine the viability of providing broadband Internet access on trains. In Section IX we provide a summary of the lessons learned from deploying broadband Internet to trains. Finally, in Section X we provide concluding remarks.

II. DIFFICULTIES AND OPPORTUNITIES

A. Difficulties

Communications on board trains are complicated by several factors. Lannoo et al. [5] state that railcars have Faraday cage-like characteristics which can lead to high penetration losses for signals. Beeby [6] adds that other complicating factors include:

- A “high vibration environment” that may require mechanical isolation of communication devices.
- A “thermally challenging” environment, since heat may be a significant issue in certain parts of the train.
- A harsh electrical environment due to:
  - The proximity of high voltages, as in electrical trains.
  - High magnetic fields, as in magnetic levitation (Maglev) trains.
  - Trains are not designed to provide a “clean” electrical supply for computers.
- The need to have equipment with minimal maintenance schedules—this may result in equipment with near military-grade specifications.
- The presence of trackside features, such as railway signaling equipment.

Some other factors hindering communications on trains include:

- Railway companies constantly add or remove rail cars from trains. As a result, it is necessary for the communications network to discover these changes automatically [7].
- Poor coupler contacts on rail vehicles, which may introduce communications failures [7].
- Tunnels may limit visibility to wireless communication infrastructure.
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>WiMax</td>
<td>Worldwide Interoperability for Microwave Access</td>
</tr>
<tr>
<td>FLASH-OFDM</td>
<td>Fast Low-latency Access with Seamless Handoff-Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>GPRS</td>
<td>General Packet Radio Service</td>
</tr>
<tr>
<td>UMTS</td>
<td>Universal Mobile Telecommunications System</td>
</tr>
<tr>
<td>UMTS-FDD</td>
<td>Universal Mobile Telecommunications System - Frequency Division Duplex</td>
</tr>
<tr>
<td>HSDPA</td>
<td>High-Speed Downlink Packet Access</td>
</tr>
<tr>
<td>ADSL</td>
<td>Asymmetric Digital Subscriber Line</td>
</tr>
<tr>
<td>VLAN</td>
<td>Virtual Local Area Network</td>
</tr>
<tr>
<td>TAT</td>
<td>Train Access Terminal</td>
</tr>
<tr>
<td>MAR</td>
<td>Mobile Access Router</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>GVC</td>
<td>Ground-to-vehicle communications</td>
</tr>
<tr>
<td>OVC</td>
<td>On board vehicle communications</td>
</tr>
<tr>
<td>CL-IP</td>
<td>Convergence Layer IP</td>
</tr>
<tr>
<td>CL-MAC</td>
<td>Convergence Layer MAC</td>
</tr>
<tr>
<td>BS</td>
<td>Base station</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless local area network</td>
</tr>
<tr>
<td>PDF</td>
<td>Policy Decision Function</td>
</tr>
<tr>
<td>SCTP</td>
<td>Stream Control Transfer Protocol</td>
</tr>
<tr>
<td>MMP-SCTP</td>
<td>Mobile Multi-Path Stream Control Transport Protocol</td>
</tr>
<tr>
<td>TMS</td>
<td>Train Management System</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple Input-Multiple Output</td>
</tr>
<tr>
<td>AGW</td>
<td>Access Gateway</td>
</tr>
<tr>
<td>SGW</td>
<td>Service Gateway</td>
</tr>
<tr>
<td>GARP</td>
<td>Generic Attribute Registration Protocol</td>
</tr>
<tr>
<td>GVRP</td>
<td>GARP VLAN Registration Protocol</td>
</tr>
<tr>
<td>G2RP</td>
<td>GARP Reservation Parameters Registration Protocol</td>
</tr>
<tr>
<td>RSTP</td>
<td>Rapid Spanning Tree Protocol</td>
</tr>
<tr>
<td>MCFA</td>
<td>Motion-aware Capacity and Flow Assignment</td>
</tr>
<tr>
<td>RAU</td>
<td>Remote Antenna Units</td>
</tr>
<tr>
<td>LCX</td>
<td>Leaky Coaxial cable</td>
</tr>
<tr>
<td>DVB-S</td>
<td>Digital Video Broadcasting-Satellite</td>
</tr>
<tr>
<td>SWiFT</td>
<td>Seamless Wireless Internet for Fast Trains</td>
</tr>
</tbody>
</table>
• Frequent handoffs\(^2\) in the cellular network. These handoffs can result in packet loss and packet reordering [8].
• The train’s mobility complicates the provision of quality of service to different traffic flows [9].

In spite of these difficulties, there are several opportunities to provide Internet access on trains using a variety of technologies, including Wi-Fi, WiMax, satellite technologies, and radio-over-fiber. In Section II-B we discuss some of these opportunities.

### B. Opportunities

The growth in wireless communication technologies over the last two decades opens up several opportunities for supporting communication on board trains. For example, customers in a stationary train can have Internet access through the existing cellular infrastructure without many modifications, except for an antenna on the outside of the train. Issues only arise when the train begins to move, particularly at high speeds, and requires several handoffs in a short period of time. Beeby [10] argues that communications capabilities on mobile terminals is constantly improving, with some phones now having multiband and Wi-Fi capabilities. Currently it is standard to have Wi-Fi integrated on laptops, and eventually WiMax might also be commonly available. These factors, especially the latter, have the potential to drive Internet usage higher, particularly because as connectivity becomes more prevalent, usage increases [10]. Beeby [10] goes on to argue that there are significant opportunities available for Internet access on trains if access to the technology can be made: simple, ubiquitous (as in not requiring any special software or terminal), and useable (that is, acceptable throughput and delay with few service interruptions). In this respect, Fourth Generation (4G) communications technologies, such as WiMax, IEEE 802.16m [11], or LTE [12] may be good solutions for offering Internet access on trains. It has been reported that WiMax is being used in the UK to provide Internet access for Southern Trains [4], [13]–[15]. We expect further growth in broadband Internet access availability on trains as more train operators are convinced of the business viability of negotiating for wireless coverage along their tracks using WiMax or some other 4G technology.

Another application for broadband communication on trains is railway signaling. Aguado et al. [16] note that standards-based communications systems such as IEEE 802.16 (WiMax) [17] and IEEE 802.20 (Mobile Broadband Wireless Access) [18] can be used for railway signaling instead of the cable-based systems currently in use.

\(^2\)To see why this might be a problem, consider a train travelling at 60 m/s (216 km/h) through an environment with cell sizes on the order of 3 km, then we would have handoffs every 50 s.
III. REFERENCE ARCHITECTURE

In this section we present a reference architecture to guide our discussion of broadband Internet access on trains. We also provide some context on handoff and addressing issues that are common to all Internet deployments on trains.

Fig. 1 shows a logical architecture for the computer networks aboard the trains used to provide Internet access to passengers. This architecture, which incorporates aspects of the train communication management platform from [19], uses gateways in each train car to build a train-level network. Broadband Internet access on the train is provided through the Train Access Terminal (TAT). This terminal, which can support one or many technology types, connects to the access network using an antenna mounted on the outside of one train car. The incoming signal from the train access terminal is then fed to gateways and wireless access points in all the rail cars in the train. Within each rail car IEEE 802.11 [20] is commonly proposed to provide connectivity to passengers; however, passengers may also connect to a wired network in the railcar, if one is available. The benefits of using such an architecture include the following:

- The cellular network system is not put under strain attempting to make handoffs for many fast moving users simultaneously [21]. Furthermore, emerging Internet Engineering Task Force (IETF) protocols for network mobility [22] can be deployed to manage TAT handoffs.
- The train access terminal can combine different access technologies. The TAT can also implement
some “intelligence” to select the best means of communication between the train and the access network, as in [21].

Fig. 2 shows a train connected to the Internet using our reference architecture. Our architecture for Internet access on trains is layered and consists of the access network, aggregation network, and the service providers’ networks. The access network (shown here as composed of base stations) is close to the train tracks, and it provides the last hop communications for the train access terminal. The aggregation network lies between the access network and the service providers’ networks, and it forwards data from the access network to the global Internet. The access gateway in the architecture combines the data from a group of users into a tunnel and forwards that data to the service gateway. The service gateway serves as an interface between the aggregation network and service providers’ networks. Van Quickenborne et al. [23] argue that aggregated tunnels per train are ideal for this architecture since they are more manageable and efficient than a per user connection scheme. From the reference architecture diagram, we can also see that there are different technology options, including satellite technologies, for the access and aggregation.
networks. This observation is in agreement with Conti [14], who states that currently there is general agreement on how to provide Internet access to passengers aboard trains. A disagreement arises on the best method to connect moving trains to the Internet backbone, i.e., how to connect the antenna on the train access terminal to the access network. However, we expect that the widespread deployment of 4G technologies may lead to some consensus on the best way to do this.

It should be noted that Fig. 2 combines features of several proposed architectures, including the FAMOUS architecture [8] that we will see later in this paper. Some other features of this architecture include:

- The access network is a wireless network with base stations along the train tracks. The access network can use either GPRS [13] and [14], UMTS/HSDPA [24], WiMax [4], [13]–[15], Wi-Fi [25], FLASH-OFDM [26], satellite links [14], IEEE 802.20 [27], or radio-over-fiber [5].
- The aggregation network can use the following technologies for forwarding data: IEEE 802.11 [28], Ethernet [8], ADSL [14], or optical fiber [15].
- Virtual Local Area Networks (VLANs) are preferred in Ethernet-based aggregation networks to carry aggregated traffic flows from the access gateway to the service gateway [8].
- Satellite links\(^3\) can be used to provide Internet access to trains; however, they do not fit this architecture neatly, since the satellite ground station cannot be easily classified as either a service gateway or an access gateway. Consequently, the satellite links in Fig. 2 are shown straddling the different networks.
- The train access terminal may support different access technologies. In [29] Rodriguez et al. implement a prototype that combines several wireless access technologies to create a mobile hotspot. Experimental results show that such a device provides much better network throughput than if a single access technology was used.

A. Other Architectures

Kanafani et al. [30] propose an architecture for Internet access on trains that is based on open standard radio technologies, such as IEEE 802.11 and IEEE 802.16, Mobile IP [31], in-train network components, train to backhaul architecture components, a trackside communication system, a homeland security surveillance system, and command and control centers. In addition, this architecture has a subsystem

\(^3\)Lannoo et al. [5] state that satellite communications are not ideal for high-speed access to trains since satellite links have limited bandwidth and long round trip times (RTT).
that handles handoffs as the train moves from the coverage area of one trackside unit to the next [30].
The train to backhaul component here is analogous to the train access terminal in Fig. 1. The trackside
communication system is the access network, while the in-train network is the same as the network shown in Fig. 1.

Riihimaki et al. [32] introduce an architecture that divides train communications into backhaul con-
nections, Ground-to-vehicle communications (GVC) and on board vehicle communications (OVC). The
GVC is analogous to the access network in our reference architecture in Fig. 2, while the OVC network
consists of customer devices as well as other networking devices, such as a train server, placed in the
train. The OVC network is similar to the train-based network shown in Fig. 1. On board each train the
OVC and GVC are connected through a connection manager (CM), which is analogous to our train access
terminal in Fig. 1 [32].

B. Handoff Issues

In 2003 it was observed that popular Internet applications may not be available at high speeds due to
lack of bandwidth, poor quality of service, and frequent handoffs [33]. These problems could be partially
addressed by: increasing network bandwidth using smart antenna systems and MIMO technologies, as
well as improved handoff protocols that prevent connection loss when moving from one base station to
another. Van Leeuwen et al. [33] state that the technologies discussed above are not sufficient to support
broadband communications at high speeds; new modulation schemes and context-aware applications are
also needed to achieve high data rates in fast moving vehicles.

De Greve et al. [8] stated, in 2005, that high link speeds for end users could only be achieved in
 cellular networks by reducing the cell size to efficiently reuse spectrum. However, small cells also mean
more handoffs between cells. Furthermore, Mobile IP is not a good protocol for delivering high link
speeds to fast moving users since Mobile IP does not work well with frequent handoffs due to handoff
latency, handoff packet loss and control message load. As a result, [8] stated that higher link speeds could
be offered to fast moving users on a train by using small cells operating in the millimeter wave band.
In addition, these authors suggest using radio-over-fiber with moveable cells to reduce handoff times, an
idea that is an extension of Gavrilovich’s moving base stations model [34]. We will revisit this concept
in Section V-A.

In 2005 Jooris et al. [35] studied seamless handoff, roaming, Quality of Service (QoS), and connections
between heterogeneous wireless networks, such as the on board network and the trackside network. On
each train the Mobile Access Router (MAR)—for connecting the train to the outside Internet—will have
one interface for each type of technology, and it will constantly choose the best link from the train to the outside world. It should be observed that the mobile access router is analogous to the train access terminal (TAT) in Fig. 1. Aboard a train, handoffs can occur when a mobile device is either unplugged from the train’s wired network or when a mobile user moves from one Wi-Fi hotspot on the train to another. In each case the user’s session must be protected. Jooris et al. [35] propose carrying out this protection by creating a convergence layer that hides the Ethernet and WLAN interfaces and instead creates a single virtual interface that has a single IP (CL-IP) and MAC (CL-MAC) address assigned to it. Outgoing packets will be encapsulated with the CL-IP and CL-MAC, while devices connected to the train LAN will only see one device and one MAC address. In Jooris et al.’s implementation, [35] every wireless user device is associated with a unique software object, which they call the access point. This software object is installed on the nearest base station (BS) on the train, but it is moved from one WLAN base station to the next as the user moves. In this architecture, each base station is configured with two interfaces, but the BS operates on a fixed frequency. The first interface runs an access point for all WLAN stations—for example, wireless user devices—within range of the BS, whereas the second interface listens to neighboring base stations’ frequencies and measures the signal strengths of the broadcast messages. If the second interface detects a stronger signal from a station than the signal measured by the station’s current base station, then the station’s access point is changed to that of the measuring interface. The station is also informed that its access point has changed frequency. This handoff mechanism has been simulated successfully, and it should allow passengers to be mobile while using the networks on board trains.

In 2007 Pareit et al. [21] assumed that several different access technologies need to be combined to provide broadband Internet access on trains. As a result, they tackle the issue of handoffs as the train moves from the coverage area of one access technology to another. To prevent the access network from having to make several simultaneous handoffs, it is proposed that train passengers connect to the Internet via on board Wi-Fi access points that are connected to the local train network. The architecture proposed in [21] places a Policy Decision Function (PDF) on the gateway, i.e., the train access terminal, between the train’s network and the outside world. The PDF decides which interface should be used to provide the connection between the train and the access network. This decision is based on link quality, train location and speed, and possibly cost or load balancing. Mobility Management modules are the other key part of the architecture. They reside partly on the train and partly on the Central Management System. These modules take input from the PDF to make handoffs as smooth as possible.

Pareit et al. [21] evaluate the feasibility of using either Mobile IP or MMP-SCTP (Mobile Multi-Path
Stream Control Transport Protocol) [36] for a mobility management handoff protocol. Mobile IP allows nodes to change their point of attachment to the Internet without changing their IP address [21], while Stream Control Transport Protocol (SCTP) is a reliable transport protocol that resides above an unreliable connectionless packet service [21]. SCTP allows for the detection and retransmission of packets that might be lost during a handoff. In addition, SCTP endpoints allow for multihoming. In [21] it is shown that MMP-SCTP displays better performance than Mobile IP after a slow start for TCP performance without a handoff. Pareit et al. emulate the case where a train passenger gets Internet access using a satellite link and an HSDPA [37] link. Reference [21] shows that for a satellite link, Mobile IP exhibits better performance than MMP-SCTP (also after the slow start). When there is a handoff between satellite and HSDPA, we see that Mobile IP does not require any retransmissions, and all packets arrive in order. Very similar results were obtained when the same test was performed using MMP-SCTP [21]. Pareit et al. [21] conclude by noting that MMP-SCTP and Mobile IP are able to handle handoffs seamlessly when handoffs can be predicted. In spite of its overhead, MMP-SCTP can be a better choice for a mobility management protocol since it does automatic retransmissions.

C. Addressing Issues

Thus far we have presented a reference architecture for Internet access on trains as well as a discussion of handoff issues. However, we also need to account for the network topology and addresses on the train-based network. Network topology on board trains changes constantly [19], hence, there is a need to create a robust management infrastructure that will establish and maintain connectivity on the train while providing logical and IP addressing services [19]. Verstichel et al. [19] propose a connectivity management platform that uses the Train Communication Standard (TCN) [7], an adapted Dynamic Host Configuration Protocol (DHCP) [38] and Network Address Translation (NAT) [39] to link the devices in a coach-level network into one network across the entire train. Addresses are statically assigned in the coach-level network and Network Address Translation is used to enable communication throughout the train.

IV. INITIAL CONCEPTS

In this section we provide an overview of some of the initial concepts that have guided deployment of broadband Internet on trains. We also examine the FAMOUS architecture, which was developed to

4Note that [21] only studied handoffs between satellite and HSDPA; however, we expect similar results for other cellular-based systems.
provide Internet access to Fast MOving USers.

Due to their mobility, Internet access can be provided on trains only by use of wireless links. Correia and Prasad [40] present some of the technical challenges involved in providing wireless broadband services. The reader is referred to [40] for a more complete treatment of the important attributes of a wireless broadband system. References [34] and [41] address how to provide broadband communications to fast moving users. In 2001 Gavrilovich [34] argued that a large number of small cells operating at high frequencies was the most economical and practical infrastructure for providing wireless broadband access to many users. In Gavrilovich’s model these small cells were provided by moving base stations that travel along a track beside the roadway. The moving base stations were then linked to fixed base stations using wireless links. The fixed base stations were uniformly distributed along the roadway and were also interspersed with the mobile base stations. This combination of mobile and stationary base stations allowed the realization of broadband wireless communications while also yielding fewer handoffs due to the mobile base stations [34]. However, a moving base station may not be practical.

An architecture for providing communications and entertainment aboard a high-speed public transport system is proposed in [41]. This architecture is composed of the following components:

- A mobile subsystem that consists of a mobile subnetwork, access to an infrastructure network, and a mobility management component. This mobile subsystem is analogous to the access network in this paper; however, it does not include any of the wireless communication technologies incorporated at the access network in our architecture.

- A wireless transport subsystem that handles radio transmission between the mobile subsystem and the infrastructure network. This subsystem is analogous to the wireless communication technologies found in the access network of our reference architecture.

- A land subsystem consisting of an infrastructure network and a network management component. This would be analogous to the aggregation network and the service providers’ core networks.

In addition Lin and Chang [41] argue that the link between the passenger device and the base station can be provided by IEEE 802.11, Bluetooth, or one of the Third Generation (3G) wireless standards. As we have seen already, WiMax—which is one of the 4G wireless standards—has been chosen in one deployment [4], [13] to provide the link between the train and a terrestrial network, while Wi-Fi has been chosen to provide the link between the passenger terminal and the train network. Finally, [41] also notes that for a train, the cell planning problem reduces to a one-dimensional problem, which should greatly facilitate frequency planning.
In this subsection we examine work done on the FAMOUS architecture; an architecture designed to support broadband Internet access for Fast Moving Users. All of this work ([5], [8], [19], [21], [23], [28], [33], [35], [42]–[51]) was conducted by researchers in Belgium.

In the FAMOUS architecture, users do not connect directly to the base stations in the access network; instead, the entire train has a single connection to the access network. This connection is then shared amongst all the users on the train. The FAMOUS architecture assumes that seamless connectivity is not guaranteed for users in fast moving vehicles; instead they will hop from one access gateway (AGW) to the next [45]. Within the aggregation network, VLANs are used to group the different base stations in an access network that satisfy a given train’s traffic demands [8]. Another component of this architecture is the service gateway, where connections are made between service providers and the aggregation network. The FAMOUS architecture is summarized in Fig. 3.

1) Aggregation Network Design: In [8], [43], and subsequent papers ([23], [44]–[47], [49]), Ethernet is chosen for the aggregation network since it is simple, cost-effective and bandwidth flexible. In spite
of Ethernet’s advantages, it requires some modifications to support fast moving users. One of Ethernet’s issues is the rapid depletion of VLAN databases in Ethernet switches. Ethernet already has an extension, called GVRP\(^5\) [43], that can register VLANs automatically in a consistent and reliable manner; however, standard GVRP distributes VLAN IDs of all tunnels to all the switches in the network, thereby flooding the VLAN databases. This issue is resolved by developing a “scoped refresh” of GVRP, such that Ethernet switches determine whether or not they are part of a given tunnel. If they are, then the switch will issue deregistration messages on all its interfaces that do not have the VLAN registered, otherwise the switch would attempt to register the VLAN. G2RP, on the other hand, has been developed to support fast moving users by allowing for the separate distribution of traffic reservation parameters and VLANs to Ethernet switch hardware. When combined, GVRP and G2RP allow switched Ethernet to be used as a transport technology for an aggregation network.

De Greve et al. [44] present the Motion-aware Capacity and Flow Assignment (MCFA) algorithm to optimize the use of network resources, determine paths for dynamic tunnels in an aggregation network, and minimize the impact of packet loss and packet reordering when designing an aggregation network. The schemes discussed in [44] include:

- An ideal routing algorithm for minimal network cost, which does not take any additional constraints into account when solving the MCFA problem.
- A limited Hop Count Variations routing scheme, which guarantees maximum delay by limiting the variation in hop counts between two different paths.
- A shared routing algorithm, which requires the paths assigned to a given connection between the node and the aggregation network to share some nodes in common.
- An incremental routing method, which is an even stricter form of shared routing in which the different paths share even more nodes in common.

Of the schemes presented in [44], incremental routing exhibits excellent packet loss features but poor scalability, while Limited Hop Count Variations (LHCV) routing yields a network that has a slightly higher network cost than if ideal routing was used. However, LHCV routing shows better congestion performance. De Greve et al. [44] also present a heuristic, called Subpath Assignment (SpA), for mapping aggregation network routes onto a minimal set of spanning tree instances. When this heuristic is compared with other path aggregation schemes, it is seen that SpA can perform the path mapping in the shortest amount of time.

\(^5\)GARP VLAN Registration Protocol, where GARP is Generic Attribute Registration Protocol.
For switched Ethernet to be used in a carrier-grade network, it must provide a mechanism for fast recovery from link failures in the aggregation network. De Greve et al. [45] present an extension to Ethernet’s Rapid Spanning Tree Protocol (RSTP) that uses a fast detection mechanism for link and node failures. This mechanism, which is resilient to node or link failures, bypasses the RSTP failure detection process and monitors links by examining incoming and outgoing packets at a given switch. De Greve et al. [45] show that if reliability constraints are added to the MCFA optimization problem, then it is possible to have good recovery times in the aggregation network, even when there are dynamic VLANs present.

De Greve et al. [46] argue that aggregation networks are not optimally designed for broadband services from fast moving vehicles; therefore, they develop an integer linear program (ILP) to calculate the exact dimensioning and tunnel paths needed to satisfy traffic demands from a train to the global Internet. For large network cases, the ILP can take several days or weeks to solve; therefore, De Greve et al. [46] develop and apply a heuristic—which achieves low congestion and optimizes the use of network resources—to solve the problem, i.e., meeting the traffic demands of fast moving users in the FAMOUS architecture. In the ILP model, each train is assumed to generate a certain amount of traffic, where these traffic demands can be defined as one of the following:

- Exact, which would require optimization of network resources with knowledge of the exact access gateway (AGW) where two trains cross each other, and the exact instant when the crossing occurs.
- Static, which results from neglecting all time-related aspects of an exact demand. This is required if a network lacks a dynamic reservation mechanism; however, it results in overdimensioning of resources.
- Train delay insensitive (TDI), which results from ignoring the exact point when and where the trains cross each other. This implies the network is dimensioned to allow trains to cross at any AGW along their respective paths.

It is shown in [46] that using TDI demand results in a more complex optimization problem; however, if traffic demands are defined as train delay insensitive, the QoS guarantees of passengers can be fulfilled always. In addition, [46] concludes that for optimal network design, the links that need to be considered for connecting the service gateway to the access gateway are those closest to the rail line end terminuses.6

6In this problem assume that the different towns/stations in the rail network represent the vertices of a graph, while the rail lines represent the edges of the graph. Then, only links between the service gateway and the access gateways closest to the vertices need to be considered when using the heuristic approach. For more details please consult [46].
Van Quickenborne et al. combine the findings from [43], [46] in [47]. Reference [47] deals with designing an aggregation network that combines data from several users as they move from one access network to the next. The access network traffic is aggregated into tunnels in the aggregation network, and these tunnels have to move with the users from one access network to the next. In designing the aggregation network, Van Quickenborne et al. [47] rely on an objective function that minimizes the number of hops between the train and the service gateways. The objective function’s constraints include link capacity restrictions and ensure that only one path is needed from source to destination. Using this optimization model it can be shown that if each train requires two dynamic tunnels—one for basic demand and the other tunnel for transient spikes in traffic demand—then the solution to the optimization problem can be obtained quickly. On the other hand, this problem takes longer to solve if we seek to minimize the costs of the network interface cards and routing subject to the same constraints. Another result from this paper shows that dynamic tunnel configuration and activation reduce network cost, since the basic traffic demand is routed over a shorter path, while the transient spikes in traffic demand are routed over longer paths [47].

The FAMOUS architecture has also been used in [23] to show that a hierarchical wired Ethernet aggregation network in combination with Ethernet-based wireless access networks may be used for providing broadband Internet access to fast moving users. Reference [23] assumes the use of dynamic tunnels, as proposed in [47]. Here the dynamic tunnel management takes one of three forms:

- Management-based approach, which uses location information, e.g., from GPS, to set up tunnels to a train. When the train arrives at an access gateway, the train’s location information is sent to a management platform that sets up the train’s tunnels. When the train moves to another access gateway, the previous tunnel is torn down.
- Signaling-based approach, in which a train announces its presence at a given access gateway, resulting in tunnel setup for the train. After a timer expires, the tunnels are torn down.
- Hybrid approach, which incorporates portions of the schemes described above, i.e., a signaling-based approach in the tunnels nearer the train, and a management-based approach in the higher parts of the network.

Simulation results from [23] show that the signaling-based approach is hard to use in aggregation networks, since tunnel-setup times increase with tunnel length, i.e., number of hops in the aggregation network.

Reference [23] presents an example of an Ethernet-based wireless access network that has a single WiMax station per access network. Each base station is linked to the aggregation network via an Ethernet link.
network. As a result, the hybrid approach is recommended. This approach has the added benefit of reducing packet loss while providing accurate tunnel-setup triggers.

The FAMOUS architecture is also extended in [28] and [48] to support the case where several leaf nodes (trains) require connectivity with a limited set of service gateways through a wireless mesh network, i.e., the aggregation network is built using wireless mesh networks. De Greve et al. [28] say this is possible because wireless mesh networks are cheaper to deploy than their wired equivalents. In [28] and [48] the access gateways are replaced by wireless gateways. In addition, the underlying aggregation network technology is replaced with IEEE 802.11e instead of switched Ethernet [28]; in the future, we expect that such a wireless aggregation network can also be provided by emerging Ethernet-based gigabit radios. Wireless networks can sometimes be subject to reduced throughput due to interference from neighboring stations. Therefore, De Greve et al. [48] suggest wireless throughput may be improved in mesh networks by intelligent distribution of neighbor mesh nodes and minimizing link interference levels by assigning different channels to the different interfaces of the wireless gateways. These objectives can be achieved by using a distributed channel assignment module that tries to minimize interference levels on links by assigning different channels to various interfaces on the wireless gateways. Fast moving users can then be supported by using a wireless mesh node placement algorithm that minimizes the hop count of the service gateway-wireless gateway paths [48].

In [49] the FAMOUS architecture is used to provide high-bandwidth and low latency traffic to fast moving users. In this case, the MCFA optimization problem from [45] is used to determine optimal aggregation gateway location, the number and speeds of interface cards, and traffic tunnel setup. The routes computed by MCFA are then mapped onto VLANs and spanning tree instances for routing in the FAMOUS architecture. Results from a testbed show that low latency high bandwidth links can be provided to fast moving users, and that rapid recovery with spanning trees is feasible without a centralized system [49].

V. Taxonomy of Architectures by Access Network Technology

In this section we provide a description of some technologies that have been proposed as options to provide broadband Internet access to fast moving users. We examine radio-over-fiber, leaky-coaxial cable-based, IEEE 802.11-based, satellite-based, WiMax-based, and high-altitude platforms. We also examine architectures based on emerging standards.
A. Radio-over-Fiber

In 2005 and 2007 Lannoo et al. ([5], [50]) proposed extensions to Gavrilovich’s [34] moving base stations model. Lannoo et al. [5] argue, just as in [8], that frequent handoffs greatly reduce the bandwidth available to fast moving users. Consequently, they propose using radio-over-fiber, as suggested in [8], to feed base stations along the rail track. Unlike in Gavrilovich’s model there are no moving base stations; instead there is a fiber-fed distributed antenna network. These distributed antennas are located along the railroad tracks, and they are called remote antenna units (RAU) (These correspond to the base stations in Fig. 2.). The remote antenna units are supervised by one control station via an optical ring network. For communications from the access network to the train, data is modulated at the control station and sent optically to each remote antenna unit using wavelength division multiplexing, i.e., each RAU has a unique wavelength for communications. The remote antenna unit will convert the optical signal to radio waves and transmit to the train. For communications from the train to the access network, the data will typically be captured by the remote antenna unit closest to the train. In order to reduce handoff times for the train access terminal, Lannoo et al. propose using “moving cells,” i.e., a cell pattern that is constantly reconfigured at the same speed as the train so that the train access terminal communicates on the same frequency during a trip. For a more complete treatment of Lannoo’s moving cell concept, please consult [5]. Fig. 4 presents a reference architecture for the radio-over-fiber deployment.

B. Leaky Coaxial Cable-based Architecture

Ishizu et al. [52] observed in 2007 that leaky coaxial cable (LCX) has been used throughout Japan for radio communications on trains (Leaky coaxial cables are regular coaxial cables with slits cut in the jacket so that the radio waves can “leak out.”). Reference [52] argues that new data services, such as broadband Internet and video, can be provided to passengers by using a new frequency in the leaky coaxial cable. To support these data services, the system has to carry out seamless handoffs between leaky coaxial cable segments at high speeds. The authors of [52] propose an architecture for communications on “bullet trains” that consists of a base station with an Ethernet interface and mobile devices. Testbed results indicated that data rates of up to 768 kbps could be achieved using leaky coaxial cable. The proposed architecture for Internet access on trains uses an Ethernet extension called Mobile Ethernet, which allows switches to learn paths and suppress unnecessary broadcasts once paths are learned. Each LCX transceiver has a transmission range of 17.3 km and about 62 of these transceivers are needed to cover all of Japan’s “bullet train” network. Assuming full utilization of each of the LCX transceivers, a gateway node between the aggregation network will need to support a peak data rate of 48 Mbps
(768 kbps × 62). On board each train is a Mobile Bridge (MB) that has interfaces for different types of communication technologies, such as IEEE 802.11, IEEE 802.16, and an LCX mobile device. The LCX mobile device in the architecture addresses handoffs by detecting new LCX base stations and transmitting beacons in the aggregation network that update the forwarding tables for the mobile bridge in the aggregation network. Reference [52] concludes by noting that a testbed is being developed to test this architecture.

C. **IEEE 802.11-based Architectures**

In 2003 Bianchi *et al.* [53] thought that it may be expensive to wire a train for network access. In addition, [53] stated that rewiring may be needed every time the train is reconfigured. Therefore, they proposed using IEEE 802.11 to construct a wireless network between the train cars. In their basic architecture, the train is connected to the Internet through a “train server” using satellite links. The train server here is analogous to the train access terminal in Fig. 1. Aboard the train IEEE 802.11 is used to: link all the railcars on the train into a computer network, provide Internet access to passengers, and
connect the train to the Internet when the satellite links become too expensive. Bianchi et al. proposed two topologies, based on IEEE 802.11, for constructing the computer network aboard the train. In their first topology, the railcars are linked into a network using IEEE 802.11 access points with antennas on the outside of each railcar, i.e., in this case the gateways shown in Fig. 1 are IEEE 802.11 access points. In order to minimize interference between adjacent access points, Bianchi et al. state that directional antennas should be used in this deployment. Furthermore, channels should be chosen on each access point, such that neighboring access points do not interfere with each other. Additional gains in performance may be achieved by using IEEE 802.11a [20] for the wireless network between railcars, and IEEE 802.11b [20] within the rail car. These technology choices imply that the computer network on the outside of the train would not interfere with that inside the railcars. An alternative topology for the network aboard the train arranges the access points in each railcar such that each access point serves as a client station for the access point in the previous car, while also serving as an access point for all the stations within its car. In other words, given train cars 1 and 2: the access point in car 2 serves as a client (station) of the access point in car 1 while also serving as the host (access point) for all stations within car 2. Since an access point may not transmit and receive simultaneously, this topology requires that each access point possess two interface cards—one for transmitting and the other for receiving. Bianchi et al. conclude by noting that their proposed topologies need to be tested in a real-world deployment to assess the impact of interference.

D. Satellite-based Architectures

Trains may be connected to the Internet via a satellite link. One proposed architecture has been developed by ACCORDE, a company specializing in developing satellite RF equipment. Their architecture consists of communications, pointing, and distribution subsystems. The communications subsystem consists of an antenna, a satellite transmitter, and a modem (same as the train access terminal in Fig. 2). The pointing system performs satellite acquisition and tracking, while the distribution system uses optical fiber links to distribute the signal between the modem and each of the cars on the train. Within each car is an IEEE 802.11 or IEEE 802.16 access point. It should be noted that emerging technologies

---

8For example, if a given train station has IEEE 802.11 access points, the train can be connected to the Internet through those access points instead of through the train server.

9It should be noted that trains using satellite links for Internet access typically use other technologies as “gap fillers” for areas where satellite coverage is limited.
such as multi-beam lens antennas [55] may lead to improvements in the quality of the satellite signals received on trains.

E. WiMax-based Architectures

In 2008 Aguado et al. [56] presented a network architecture based on WiMax for use in railway environments. This architecture can also be used for railway signaling. Reference [56] states that WiMax can be considered for communications in railway contexts because it [WiMax] can provide mobility support at speeds up to 500 km/h. In addition, WiMax also provides QoS support.

One emerging trend in railways is the use of broadband wireless communications to control trains [56]. The communications between the train and the operations center can be used to enable safer railway operations. Aguado et al. [56] proposed an architecture for train signaling that consists of a train network, a wayside to train network, and the wireless backbone network. The train network is similar to the coach level network from Fig. 1. It contains a WiMax Type C node that also has Ethernet and IEEE 802.11 connections. The Ethernet connections are used to link all the carriages on the train into a network, while the IEEE 802.11 connections are used to provide connections to customers’ laptops. The wayside to train network is analogous to the access network in Fig. 2. It consists of WiMax nodes that are regularly spaced along the trackside. These nodes have two interfaces; one interface for connecting to the node C on the train, and the other for connecting to nodes in the backbone network. Finally, the wireless backbone network is analogous to the aggregation network in Fig. 2. It consists of WiMax type A nodes, where each node has three interfaces. Two interfaces are used for connecting to other nodes in the backbone network, while the other interface connects to nodes in the wayside to train network. Simulation results show that this network architecture can satisfy European Union requirements on end-to-end delay in railway communication networks.

Kumar et al. [57] introduced an architecture called SWiFT (Seamless Wireless Internet for Fast Trains) in 2008. This architecture consists of IEEE 802.11e [20] access points within train carriages for the on-train network, IEEE 802.16m base stations at the trackside for the access network, and an optical backbone (aggregation network) for linking the IEEE 802.16m base stations to the global Internet. Kumar et al. argue that this architecture is viable because customers can continue using their access terminals. Furthermore, Doppler effects for customers are reduced since the IEEE 802.11 access point is within the train carriage, while handoffs are simplified since the train appears as a single access terminal to the IEEE 802.16 network. The proposed architecture is used in conjunction with the proposed IEEE 802.21 standard [58] for smart handoffs. Simulation results show that by using triggers that predict Layer 2
handoffs, one can reduce the number of packets lost during a handoff [57].

**F. Architectures Based on Emerging Standards**

Most of the papers we have seen thus far use existing radio technologies, such as IEEE 802.16 [17] or cellular technologies. In 2004 Zou et al. [27] deviated from most of the previous work, and called for using IEEE 802.20 [18], which is technology under development, in the access network to provide broadband Internet access for trains. IEEE 802.20 is chosen because existing 3G technologies do not offer sufficiently high data rates to support many users on a high-speed train. IEEE 802.20, on the other hand, is being designed to support data delivery at high bit rates to vehicles travelling at up to 250 km/h, while using the wireless spectrum efficiently [18]. As in many of the other systems reviewed thus far, Zou et al. use an IEEE 802.11 WLAN on board the train to provide Internet access to passengers. In order to allow for smooth handoffs between base stations, they call for the train to make two IEEE 802.20 connections to base stations, i.e., the train access terminal in Fig. 1 will make connections to two separate base stations in the access network. However, the train would maintain a single IP address, using Mobile IP, throughout its journey. Furthermore, they argue that since the train’s schedule is known, handoff instances should be handled by a Predictive Pre-handover (PPH) algorithm that would precompute the routes needed after a handoff. The access node on the train would actively monitor the received signal strength from IEEE 802.20 stations, and it would trigger a handoff whenever the received signal strength from the new station exceeds that of its current base station [27].

**G. High-Altitude Platforms for Internet Access**

With the exception of the system proposing the radio-over-fiber methods for Internet access on trains, hitherto all the systems that we have studied examine communication protocols for providing Internet access. White and Zakharov [59], on the other hand, deal strictly with physical layer issues. They argue that high-altitude platforms, such as airplanes and airships at stratospheric altitudes, are a less costly yet feasible method of providing Internet access to trains. Digital Signal Processing (DSP) algorithms for tracking high-altitude platforms are presented in [59]. The algorithms’ purpose is to estimate the direction of arrival (DOA) for signals transmitted from a high-altitude platform (HAP) to a train. Some of the methods applied for DOA estimation include Spectral-based [59] and Polynomial-based [59] techniques. An Extended Kalman Filter (EKF) is used to track the train location, while beam forming is used on the satellite uplink. Finally, the paper shows that EKF can track slow variations in train velocity and account
for sudden HAP motion. Null steering (beam steering) is also shown to be advantageous in HAP-train data communications.

H. Internet Access through Direct Connectivity with Customer Terminals

It is also becoming more common for laptops to ship with 3G modems built-in. Furthermore, train passengers are able to get Internet access by connecting 3G-capable telephones to their laptops. Since each user connects directly to the access network in this case, there is no train access terminal. Sauter [24] describes tests conducted by connecting an HSDPA terminal directly to a communications network outside the train as he travelled from Paris to Frankfurt on a German Intercity Express high-speed train. Sauter [24] observed peak download speeds of about 1.5 Mbps with average download rates of 850 kbps. Sauter concludes by arguing that dedicated 3G coverage along railway tracks would give a “smooth user experience” without any onboard 3G/Wi-Fi train access terminals.

VI. Testbed Results

Thus far we have reviewed the reference architecture, initial concepts underpinning broadband Internet deployment on trains, and a taxonomy of access network technologies. In this section we review results from testbed implementations of Internet access on trains. The transition from the more theoretical to prototypes and deployment begins in [61] and [62]. In 2004 Sivchenko et al. [61] presented simulation results showing that Internet traffic performance on high-speed trains decreases as the number of users increases, which is an expected result. The performance of several existing radio technologies with respect to data rates experienced on fast moving trains is investigated in [62]. Gaspard and Zimmermann [62] evaluated the relationship between throughput as a function of Doppler shift (speed) in 2005. This investigation was carried out in two phases: in the first stage, a channel sounder was used to take channel measurements for different placements of a mobile receiver, while the mobile transmitter was moved along the track. In the next stage, different access network radio technologies were evaluated using a hardware emulation of the channel characteristics. The experiments evaluated how throughput would vary for a channel between a trackside transmitter and a receiver on board a train. Experimental results indicate that:

- TCP/IP throughput of a UMTS-FDD downlink does not vary much with receiver input power; however, it is relatively low, i.e., ∼0.06–0.35 Mbps.
- At 300 km/h, TCP/IP throughput of an IEEE 802.11b link between a trackside transmitter and a receiver on the train varies with receiver input power due to multipath channels. It should be noted
<table>
<thead>
<tr>
<th>Access Network Technology</th>
<th>Data rates</th>
<th>Handoff Frequency</th>
<th>Technology Maturity</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEEE 802.11</td>
<td>Up to 54 Mbps</td>
<td>High</td>
<td>Mature</td>
<td>Used in the access network of [60] as a gap-filler. Frequently used for the link between customer terminals and the access point in each car.</td>
</tr>
<tr>
<td>WiMax</td>
<td>Up to 42 Mbps</td>
<td>Can be relatively high</td>
<td>Mature. Other draft standards are being added to improve performance at high speeds</td>
<td>Used by Utah Transit Authority and Southern Trains</td>
</tr>
<tr>
<td>GPRS</td>
<td>Up to 171 kbps</td>
<td>High</td>
<td>Mature</td>
<td>Used by Southern Trains</td>
</tr>
<tr>
<td>UMTS (HSDPA)</td>
<td>Up to 2 Mbps</td>
<td>High</td>
<td>Mature</td>
<td>Deployed by Vodafone along InterCity Express track from Paris to Frankfurt [24]</td>
</tr>
<tr>
<td>FLASH-OFDM</td>
<td>Up to 1.5 Mbps</td>
<td>High</td>
<td>Mature</td>
<td>Unclear if used in any real deployments</td>
</tr>
<tr>
<td>Satellite</td>
<td>512 kbps (upload), 2 Mbps (download)</td>
<td>Low</td>
<td>Mature</td>
<td>Used by Thalys and SNCF trains</td>
</tr>
<tr>
<td>Leaky coaxial cable</td>
<td>Up to 768 kbps</td>
<td>High</td>
<td>Mature</td>
<td>Deployed in Japan</td>
</tr>
<tr>
<td>Radio-over-fiber</td>
<td>Not known, but will be high</td>
<td>High</td>
<td>Immature</td>
<td>Not yet deployed. Proposed in [5].</td>
</tr>
<tr>
<td>IEEE 802.20</td>
<td>Will be high</td>
<td>High</td>
<td>Standards being drafted</td>
<td>Not yet deployed. Proposed in [27]</td>
</tr>
</tbody>
</table>

that IEEE 802.11b provides high data rates under the measurement conditions. In addition, the authors state that one would need several access points along the track to have good coverage.

- The IEEE 802.16 system evaluated in [62] was not suited for high-speed trains since TCP/IP throughput decreased sharply with increasing speed. However, the authors note that the amendments to the IEEE 802.16e standard for mobility should enhance the performance of the IEEE 802.16 system.
Ceprani and Schena [63] presented implementation details on their Fast Internet for Fast Train Hosts (FIFTH) project in 2004. The FIFTH architecture consists of Mobile Train Terminal Prototype (MTTP) and FIFTH Access Network Infrastructure (FANI) modules. The MTTP is composed of a Satellite Access Terminal (SAT), which uses the Ku band to provide satellite access for the train, and the Train User-Local Area Network (TU-LAN), which constitutes the LAN on board the train. The satellite access terminal is analogous to the train access terminal in Fig. 1, while the Train User-Local Area Network is akin to the rest of the computer network shown in Fig. 1. The antenna for the SAT is adjusted by a Navigation and Tracking Unit during a trip to optimize reception conditions. The TU-LAN consists of a coach LAN (within a train car) and a train LAN (between cars on the train). The TU-LAN is implemented by using Ethernet connections between train cars, as well as Ethernet connections and IEEE 802.11 links for passengers to use. Unfortunately, additional details are not available on what bit rates were seen during the trial.

A lot of the work coming from North America is experimental, given the lack of widespread Internet access on board trains. One example of some experimental work comes from the University of Nebraska, where Hempel et al. [25] deployed a wireless testbed for IEEE 802.11 along a train track in 2006. In this testbed, IEEE 802.11 access points were placed along the tracks with line of sight paths to neighboring access points. This arrangement allowed for seamless IEEE 802.11 coverage along the tracks. IEEE 802.11a channels were used to provide backhaul links between the testbed access points, while IEEE 802.11b was used to provide wireless Internet connectivity to the train car used in the tests. Results from the testbed showed that IEEE 802.11b could support data rates of up to 11 Mbps; however, IEEE 802.11b was also subject to interference from passing trains. Additional test results showed that train velocity does not appear to have a significant effect on the throughput experienced by the node on board the train. The conclusion from this paper is that while it is feasible to deploy IEEE 802.11 along the train track, IEEE 802.11 has a limited coverage area; therefore, such a deployment would be expensive [25].

Most of the work we have reviewed in this paper has discussed providing Internet access to passengers on a train. However, a train operator might also like to collect operational data from its trains. Edwards et al. [2] discuss just such a scheme that allows for controlling and monitoring various sensors and supervision modules on a freight train. This scheme uses IEEE 802.11b for intratrain communications to allow for braking, coupling and uncoupling, etc. This scheme uses a Controller Area Network (CAN) bus to collect data from sensors on board the train. The data is then coupled with GPS information and reported to a web server via a CDMA-based transmitter. In this case, the train access terminal is a 1xRTT.
radio, whereas the links between the cars are IEEE 802.11b links, unlike the wired links shown in Fig. 1.

VII. IMPLEMENTATION EFFORTS

In the previous section we reviewed results from testbed implementations of Internet access to trains. In this section we look at how those ideas have been implemented in Europe and North America. As we mentioned in Section I, broadband Internet access is increasingly becoming available on trains in Europe. In Europe, the preponderant demand for Internet access is from passengers, while in North America, train traffic is dominated by freight [64]. As a result, efforts to carry out communications from trains have evolved in slightly different directions on these two continents due to market forces. We review the implementation efforts in Europe and North America separately, since conclusions drawn from one continent might not necessarily apply to the other. Furthermore, implementation efforts in Europe are much more advanced than those in North America. In 2007 zu Hörste [65] observed that railway operators are moving away from proprietary solutions to commercial off-the-shelf (COTS) solutions to reduce cost while improving bandwidth and reliability. In this section we will see implementations based on open standards such as WiMax (IEEE 802.16) and cellular technologies—a trend that appears to validate zu Hörste’s observation.

A. Implementation in Europe

One of the earliest accounts of Internet access on trains comes from the Railway Open System Interconnection Network (ROSIN) project. In 1999 Fabri et al. [66] presented a report on a web-based tool deployed to a train to allow maintenance staff to supervise railroad equipment using a GSM connection between the train and an operations center. Aboard the train the railcars were linked into a network using the Train Communication Network (TCN) standard [67]. Unfortunately, reference [66] does not provide any additional details on the bit rates seen during the trial or the network topology.

Conti [14] provides a contemporary (2005) view of the implementation of Internet access on trains in Europe. In his paper he argues that telecommunications operators have offered Internet access to passengers using GPRS [68] or 3G wireless cards; however, this is not sufficient for most users. Furthermore,

---

10The TCN specification consists of a train bus and a vehicle bus. The train bus can self-configure itself by connecting a new node (railcar) to the network and dynamically assigning it a new address. The vehicle bus is optimized to handle small packets originating from a large number of devices. The train bus and the vehicle bus are connected through a gateway, which allows for exchange of data between devices in the same railcar, or in two different train cars. The TCN can also be linked to the Internet by means of a radio link between the train and a ground station.
he states that there is now agreement that Internet access should be provided on board trains using IEEE 802.11 access points within the train; however, there is not much agreement on how to connect moving trains to the Internet backbone. In the United Kingdom GNER trains use a combination of satellite and cellular links to provide a backhaul link from the train. Therefore, the train access terminal in this instance supports both satellite and cellular technologies. The Internet connection is shared with all cars on the train using the train’s lighting circuit; this implies that the topology of the computer network on the train is not radically different from that shown in Fig. 1. Unfortunately, additional details are not available on how the wired portion of the network aboard the train operates. GNER’s system favors satellite access for the backhaul link, but when the train enters a tunnel, the system automatically switches over to GPRS (The technical details of how this switch is accomplished are not clear from [14]; however, it may be assumed that the GPRS signal is brought into the tunnel via a leaky cable, or some similar mechanism.). For redundancy purposes, the train connects to base stations from two different mobile carriers. In addition, up to six parallel cellular phone links are established for redundancy purposes when the train passes through a tunnel. It is worth noting that this same technology is also used by the Swedish train operator, SJ, to provide Internet access [14].

Conti [14] discusses Southern Trains’ efforts to provide Internet access on its trains along the 96 km London-Brighton route using WiMax [4]. It is interesting to note that this system does not use any of the enhancements found in IEEE 802.16e, which is designed for mobile access. Instead, this system uses a draft implementation of IEEE 802.16d [13]. T-Mobile and Nomad Digital collaborated on the implementation of this venture, however, it is now operated by T-Mobile. In addition to the pre-WiMax standard, GPRS and 3G technologies are also used for robustness with each train having three GPRS modems for redundancy [13]. As of 2005 there were 37 WiMax base stations deployed along the track, with plans to install up to 60 base stations [13]. The base stations operated in the 5.4–5.8 GHz band with a 1 W maximum power output [69]. Each base station is equipped with a 2 Mbps ADSL link to the Internet [13] even though each base station could achieve data rates of up to 32 Mbps for both the uplink and the downlink wireless channels [14]. In Southern Trains’ implementation the train access terminal consists of a server with support for WiMax and GPRS technologies. The architecture of the in-train network is akin to that shown in Fig. 1, with passengers connecting to the in-train network using an IEEE 802.11b link [13]. Finally, the access network in this case uses WiMax and GPRS [13], while the aggregation network uses ADSL [13].

As we have already seen, satellite links can also be used for Internet access on trains. By 2005 a European train company, Thalys, [14] was using a bidirectional satellite link, designed by 21Net, and
operating in the Ku-band to support link speeds of up to 2 Mbps, i.e., the train access terminal only supports satellite links. The downside of relying on satellite links is that operational costs are probably higher than for links that rely on either WiMax or 3G technologies [14]. By 2008 Thalys claimed that all of its trains provided broadband Internet access. Thalys has routes from Marseille to Paris, Paris to Brussels, Brussels to Amsterdam, and Brussels to Cologne.

Lomas [60] discusses a 2008 deployment of Internet access to three SNCF trains\textsuperscript{11} in France using a combination of satellite and Wi-Fi links. This system uses the Eutelsat satellite and then switches over to a Wi-Fi relay when the antenna on the train loses line of sight with the satellite, for example, at stations or in tunnels. The system provides speeds of 512 kbps and 2 Mbps for upload and download respectively. It should be noted that the SNCF system can only support up to 50 out of 375 passengers on the train. Hence, if more passengers want to get Internet access, the system will need to be upgraded.

Echensperger [26] discusses work done by T-Mobile in Germany to bring Internet access to Intercity Express trains. He discusses the Railnet effort, which aims to provide WLAN access on board trains while also providing a broadband radio connection between the train and the land side. The Railnet system uses a Central Train Unit to control traffic and store onboard content, several antennas to maintain the train to base station link, an IEEE 802.11 network to link the rail cars into a train level network, and IEEE 802.11 access points on board the train for passenger access. The onboard network for the Railnet effort is very similar to that shown in Fig. 1, except that there are no wired links between the railcars. Instead, this time we have IEEE 802.11 links between the cars. The train access terminal in this case supports T-Mobile’s access technology.\textsuperscript{12} Since T-Mobile (the service provider) owns its network, and also provides service on board the train, there is not much of a distinction between the access and aggregation networks in this case. It is worth noting that FLASH-OFDM has also been evaluated in the course of the Railnet effort, and its throughput has been found to be nearly independent of velocity [26]. By 2009 broadband Internet service was available on routes between Dortmund and Cologne, Cologne and Frankfurt, Frankfurt and Munich, and Frankfurt and Hamburg.

B. Implementation in North America

As previously mentioned rail transportation in North America and Europe have very different characteristics. Consequently, broadband Internet deployment to trains on those continents has evolved differently.

\textsuperscript{11}These trains travel at 320 km/h (88.8 m/s).

\textsuperscript{12}Unfortunately, technical details on the access technology are not available in [26].
In fact, it could even be argued that these deployments are in their infancy in North America. However, there are some efforts underway for North America. For example, Conti [14] points out that PointShot Wireless has worked on initial deployments with Canada’s VIA Rail and California’s Altamont Commuter Express and Capitol Corridor operators.

In 2008 Nomad Digital collaborated with the Utah Transit Authority (UTA) and Wasatch Electric to provide a wireless broadband connection on a 64 km long commuter line between Ogden and Salt Lake City. In this case the access network consists of WiMax radios from Redline Communications. On board the train, passengers get Internet access from a free Wi-Fi connection [15]. The onboard network for this rail deployment is very similar to that shown in Fig. 1, while in this instance the train access terminal supports WiMax. Unlike in any of the examples seen thus far, the aggregation network in this instance is composed of fiber optic links, some of which run trackside [15].

VIII. Business Models and a Feasibility Study

As we have seen in previous sections, broadband Internet access is increasingly being deployed to trains. However, for us to see more widespread deployments, train operators would have to be convinced of the business advantages of such a deployment. In this section we present different business models for paying for Internet service on trains.

One of the earliest business models developed studied deploying Internet access to intercity trains in California [30]. In developing this model, the authors say that the provision of Internet access on trains would likely lead to an increase in ridership on intercity trains. The train operators, on the other hand, could collect revenue from this service either by applying “per use or time charges, subscription fees,” or by negotiating an arrangement with a third party to pay for the service through advertising, or sponsorship, or an increase in ridership [30]. In the case of California trains, the authors present two business models for providing Internet access:

- Option 1 is a conservative model that uses satellite and cellular networks for backhaul, with an IEEE 802.11 access network on the train. This option has a low operational cost with low bandwidth and a high operational cost with high bandwidth, but it generally results in low revenue for the train operator. This option is aimed at capturing mobile Internet users on trains in a conservative manner.
- Option 2, uses WiMax for backhaul access with an onboard Wi-Fi network, but it has a high initial cost (due to the cost of deploying WiMax antennas) with low operational costs. Kanafani et al. [30] state that this model should result in high revenue for the train operator, and that it should help capture mobile Internet users as the market grows.
The next two business models were developed for use in Europe. Using data from Belgian railways, Lannoo et al. [51] present business models that investigate the possibilities and economic viability of providing Internet access on trains. Recall that these researchers are part of the same group that proposed the FAMOUS architecture. As in previous work, they argue that broadband Internet access on trains can be provided by using an in-train network, and a network between the train and the service provider for Internet access. For the backhaul network, trains can use cellular networking technologies such as GPRS/UMTS/HSDPA, or wireless networking technologies such as Wi-Fi, WiMax, FLASH-OFDM, or even a satellite networking standard, such as DVB-S/DVB-S2/DVB-RCS. These backhaul networks can be classified as either incumbent networks, for example, GPRS, UMTS, or HSDPA, or dedicated networks, such as, WiMax or FLASH-OFDM, or satellite networks. With incumbent networks the goal would be to provide Internet access on trains without making a major capital expenditure. The business model presented in [51] considers using incumbent networks until their capacity requirements are exceeded, then one can roll out a dedicated network. Satellite networks would only be used as gap fillers, i.e., in areas where the other networking standards do not provide adequate coverage, just as we saw in [14]. The analysis carried out in [51] assumes revenue schemes where either every passenger pays for Internet service, or only first class passengers get free Internet access, while all other passengers pay. Their analysis also includes the capital expenses required for deploying Internet service, as well as the operational costs required to maintain service. The model then presents results to show that train operators would realize a net profit if only first class passengers get free Internet access. Lannoo et al. [51] conclude by noting that using a combination of technologies is the best way to provide broadband Internet access to trains, and that in the particular case of Belgian railways it would be better to use a mix of WiMax and UMTS for Internet access [51].

More recently Riihimaki et al. [32] have studied Finnish railroads to determine the feasibility of deploying broadband Internet to trains. They argue that revenue from providing Internet service to train customers may come from the following sources:

- An increase in passenger volume, if a train operator offers free Internet access for passengers.
- An increase in the number of first class passengers, if first class passengers get free Internet access.
- Reduced personnel costs, if passengers who buy their tickets online get free on board Internet access.
- Direct revenue, if train tickets and data connections are sold separately.

From the standpoint of the train operator, Internet access on trains could allow for more efficient train operations, e.g., allowing real-time traffic control, or more efficient staff who can verify passenger tickets
in real-time.

Hitherto, we have focussed on Internet access to passengers, Riihimaki et al. state that train operators shipping freight could use a broadband Internet connection to allow their customers to perform accurate cargo monitoring. In the case of the Finnish railroads, it is argued that the cost of building a network for Internet access from trains can be spread out over a period of time if the network is built in two or more phases, for example, by using GPRS or FLASH-OFDM in the first phase, and then using mobile WiMax in the second phase. Furthermore, in the case of WiMax they show that the average revenue collected per user, and the cell range of the WiMax network are the most critical parameters influencing this technology’s viability for Internet access on trains [32]. For example, their analysis is based on an estimated WiMax cell size of 5 km. However, if this cell size is decreased by 10% then it becomes unprofitable to provide Internet access using WiMax [32].

Lundberg and Gunningberg [70] study the feasibility of using IEEE 802.11 networking equipment to provide Internet access for a train traveling at 200 km/h between Uppsala and Stockholm. Here they observe that commercial solutions for providing Internet access on trains are available, but note that they are either limited or expensive. Furthermore, they observe that if IEEE 802.11 technology is used, the technology choice will depend on the possible impact of fading and related problems, such as the Doppler effect due to the train’s motion [70].

Given that most of the train traffic in North America is freight traffic [64], possibly the best avenue for getting broadband Internet access on trains would be to forge some kind of partnership between the train companies and telecommunications companies. If the train operators can see a reduction in their insurance payments by allowing freight customers to gain visibility into their shipments or other gains in efficiency, then the long-term viability of broadband Internet on trains may be achieved in North America. In the case of the United States, Amtrak passengers can also benefit from a deployment of broadband Internet access to trains, and perhaps even more people can be lured to riding trains, resulting in lower greenhouse emissions.

IX. LESSONS LEARNED

In the last section we presented a review of models that can be used to test the viability of broadband Internet access on trains. In this section we present a summary of some of the lessons that have been learned from broadband Internet deployments on trains.

- Most of the deployments that we have seen in this paper provide a single access terminal per train, and the connection(s) made by the access terminal is shared by all passengers on the train. Such an
architecture prevents the mobile telephony provider from having to make handoffs for many users simultaneously.

- Rodriguez et al. [29] stated that an access terminal displays better performance if several technologies are combined. Except for the Thalys deployment, each of the systems discussed in Section VII combines several access technologies.
- Switched Ethernet may be used in a carrier-grade network to support fast moving users; however, extensions are needed to improve Ethernet’s recovery from link failures. Optimization techniques can also be applied to determine paths for dynamic tunnels in the aggregation network while minimizing the impact of packet loss and packet reordering.
- Fourth-generation communication technologies such as WiMax have already been used in broadband Internet deployments to trains.
- A modified version of IEEE 802.11 has been used as a “gap filler” technology to provide Internet access to trains in areas where the train might not have a line of sight path to a satellite.
- It has been shown [62] that IEEE 802.11 provides high data rates at high speeds. However, except for the gap-filler case, IEEE 802.11 is typically not used in the access network because it costs too much to install sufficient access points at the trackside.
- While Lannoo et al. [5] have argued that satellite links are not suited for broadband Internet access to trains, we have seen two cases from Europe (Thalys and SNCF) that rely primarily on satellite links.
- Passengers are beginning to use 3G-capable hardware to connect to the cellular infrastructure directly from within trains [24]. This is in spite of Lannoo et al.’s argument that a direct link between the passenger and a base station is “too liable” to the Faraday cage characteristics of the railcar [5].

X. Conclusion

The availability of broadband Internet access on trains should prove to be a revenue source for operators. Previous studies from the United Kingdom show that train companies can attract more passengers if Wi-Fi access is made available [1]. In this paper we have presented some of the initial approaches, current technologies, and future ideas, such as IEEE 802.20 and radio-over-fiber, related to Internet access on trains. We have also provided an account of implementation efforts for broadband Internet access on trains in Europe and North America. These efforts, particularly from Europe, show that broadband Internet access on trains is realizable. Furthermore, business models, developed to test the viability of Internet access on trains, show that broadband Internet access on trains is best realized by using a combination
of access technologies. However, efficient operation requires proper system design. North America does not share the same rail traffic characteristics as Europe [64], and so broadband Internet access on North American trains is not as readily available. In North America, broadband Internet access on trains may be used for collecting operational data from trains, as well as freight monitoring. Future work could be to develop a business model for broadband Internet access on North American trains that takes into account the fact that North American rail traffic is dominated by freight. A good business model might serve to accelerate the deployment of broadband Internet access in North America.

ACKNOWLEDGMENT

The authors would like to thank Yewande Lewis and Ann Francis for reading and commenting on previous versions of this paper.

REFERENCES


