

Architectures and Technologies for the Future Mobile Internet*

Dipankar RAYCHAUDHURI^{†a)}, *Nonmember*

SUMMARY This position paper outlines the author's view on architectural directions and key technology enablers for the future mobile Internet. It is pointed out that mobile and wireless services will dominate Internet usage in the near future, and it is therefore important to design next-generation network protocols with features suitable for efficiently serving emerging wireless scenarios and applications. Several key requirements for mobile/wireless scenarios are identified - these include new capabilities such as dynamic spectrum coordination, cross-layer support, disconnection tolerant routing, content addressing, and location awareness. Specific examples of enabling technologies which address some of these requirements are given from ongoing research projects at WINLAB. Topics covered briefly include wireless network virtualization, the cache-and-forward (CNF) protocol, geographic (GEO) protocol stack, cognitive radio protocols, and open networking testbeds.

key words: *future mobile Internet, virtualization, cache-and-forward, geographic routing, cognitive radio, network testbeds*

1. Introduction

Over the next 10–15 years, it is anticipated that significant qualitative changes to the Internet will be driven by the rapid proliferation of mobile and wireless computing devices. Wireless devices on the Internet will include laptop computers, personal digital assistants, cell phones (over 3 billion in use as of 2009 and growing!), portable media players, etc. along with embedded sensors used to sense and control physical world objects and events. As mobile computing devices and wireless sensors are deployed in large numbers, the Internet will increasingly serve as the interface between people moving around and the physical world that surrounds them. Emerging capabilities for opportunistic collaboration with other people nearby or for interacting with physical world objects and machines via the Internet will result in new applications that will influence the way people live and work. The potential impact of the future wireless Internet is very significant because the network combines the power of cloud computation, search engines and databases in the background with the immediacy of information from mobile users and sensors in the foreground. The data flows and interactions between mobile users, sensors and their computing support infrastructure are clearly very different from that of today's popular applications such

as email, instant messaging or the world-wide web.

As a result, one of the broad architectural challenges facing the network research community is that of evolving or redesigning the Internet architecture to incorporate emerging wireless technologies - efficiently, and at scale [1]. The Internet's current TCP/IP protocol architecture was designed for static hosts and routers connected by wired links. Protocol extensions such as mobile IP have been useful for first-generation cellular mobile services involving single-hop radio links from mobile devices to base stations or access points [2]. However, incremental solutions based on IP are inadequate for dealing with the requirements of fast-growing wireless usage scenarios such as multi-hop mesh [3], peer-to-peer [4], disruption-tolerant networks (DTN) [5], sensor systems [6] and vehicular applications [7]. These emerging wireless scenarios motivate us to consider "clean-slate" network architectures and protocols capable of meeting the needs of these and other emerging wireless scenarios.

2. Future Network Requirements

Considering the wide range of future wireless network usage scenarios (4G cellular/mobile, WLAN, mesh, P2P, DTN, sensor networks, vehicular networks, pervasive/ubiquitous systems), it is important to extract a set of common requirements general enough to meet these needs as well as those of future applications which cannot easily be predicted today. We suggest an approach for decomposing these requirements into two major categories, the first reflecting the intrinsic properties of the radio medium, and the second reflecting the needs of future mobility and pervasive services. It is important to note that these requirements should apply to future access networks and the Internet protocol stack as a whole in view of the increasingly predominant role of wireless end-user devices. The current approach of designing specialized networking solutions for cellular systems, ad hoc nets, sensor applications and so on leads to undesirable fragmentation (and hence poor scalability, lack of interoperability, inefficiencies in application development, etc.) among different parts of the network, and needs to be replaced by a unified end-to-end protocol architecture which supports emerging requirements of both wired and wireless networks.

To elaborate further, basic transport services of future Internet protocols should reflect intrinsic radio properties such as spectrum use, mobility, varying link quality, hetero-

Manuscript received September 5, 2009.

Manuscript revised November 7, 2009.

[†]The author is with WINLAB, Rutgers University, 671 Route 1, North Brunswick, NJ 09802, U.S.A.

*Research supported by US National Science Foundation grants # ANI-0335244, CNS-0626740, 0626959, 0725053.

a) E-mail: ray@winlab.rutgers.edu

DOI: 10.1587/transcom.E93.B.436

geneous PHY, diversity/MIMO, multi-hop, multicast, etc. and the capabilities of emerging radio technologies such as LTE, next-generation WLAN, Bluetooth, Zigbee, vehicular standards such as 802.11p, and of course, cognitive software-defined radio (SDR). In addition, Internet protocol service capabilities should be designed to serve emerging uses of wireless technology, not only for conventional mobile communications, but also for content delivery, cloud computing, sensing, machine-to-machine (M2M) control, and various other pervasive system applications. Examples of specific protocol features that may be useful are:

- Dynamic spectrum coordination capability
- Fast topology discovery and ad hoc routing
- Dynamic mobility for end-users and routers
- Cross-layer protocol stack for adaptive networks
- Incentive mechanisms for cooperation
- Routing protocols for intermittent disconnection
- Transport protocols for time-varying link quality
- Efficient multicasting and multi-path routing
- Geographic routing and multicast capabilities
- Content- and context-aware routing for sensors
- In-network storage for content caching
- Low delay modes for pervasive applications
- Enhanced security & privacy for radio medium

3. Protocol Architecture Considerations

Design of a single, unified protocol framework for the emerging wireless/mobile scenario is a challenging technical problem. The most straightforward approach would be to identify key protocol attributes needed for wireless access and mobility and then design these into a single protocol framework that works across multiple services and usage scenarios. Clearly, this is a complex design problem due to the large number of requirements such as those identified above, and it is questionable whether a single protocol architecture can be achieved through existing R&D and standardization processes. Because a universal wireless protocol solution is difficult to find, most researchers in the field tend to limit their work to specific usage scenarios (sensor, MANET, mesh, etc.) and then optimize specialized solutions for the requisite networking protocols. This has led to an explosion in the number of distinct protocols for wireless access networks, making it unlikely that these point solutions will eventually converge to a single unified solution. In our own work with clean-slate protocols in ongoing future Internet projects, we encountered dramatically different requirements — as an example, there is not much in common between reliable short message multicast in automotive applications and routing of media files to mobile handsets.

Rather than attempting to combine qualitatively different requirements into a single protocol, we have started to think about the use of network virtualization technology as a foundation capable of supporting a set of otherwise incompatible protocols which have been individually optimized for very different transport services. This

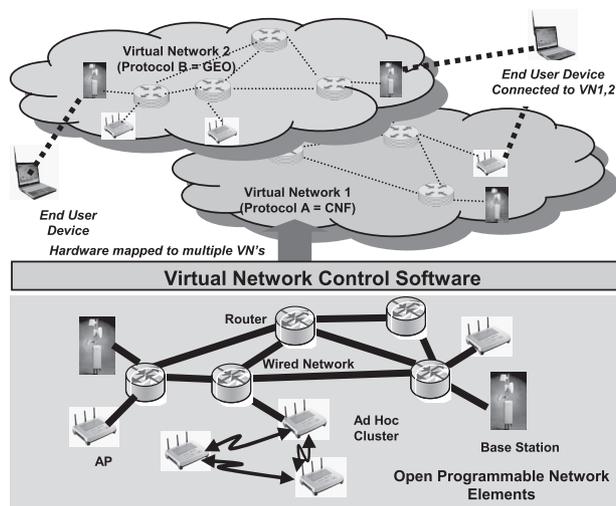


Fig. 1 Multi-protocol architecture using network virtualization.

so-called “pluralistic network architecture” [8], which first gained attention in context of experimental networks such as PlanetLab [9] and GENI [10]), uses network virtualization and programmability to accommodate multiple independent “slices” within the same physical infrastructure. In the proposed converged multi-protocol mobile network architecture (see Fig. 1), each virtual network “slice” corresponds to a distinct protocol optimized for a specific control or transport service. End-user devices attaching to the network will use one or more of the available services by opting in to the applicable virtual network slices. It is noted here that such a pluralistic architecture also offers the important features of legacy support and graceful migration because legacy protocols can be supported on one of the slices, while future services can be added later by appropriately provisioning a new slice.

4. Key Technologies

In this section, we briefly discuss some of the key technologies being considered at WINLAB as part of our research on the future mobile Internet. It is noted that the topics discussed next are only representative examples of enabling technologies still at the research stage, and are not intended to be a comprehensive list.

4.1 Wireless Virtualization

The pluralistic architecture outlined in Sect. 3 depends on the wireless network virtualization as the technology foundation. While there has been a lot of attention given to wired network virtualization [11], techniques for wireless virtualization are at an earlier stage. Note that the problem is more complicated than for wired networks because the capacity of a radio channel varies with signal strength at the mobile clients and interference from other radio devices, and these parameters change dynamically due to mobility.

As a first step, we are working on virtualization of

infrastructure mode access points and base stations such as WiFi and WiMAX. Initial results for WiFi APs [12], [13] show that such devices can be effectively partitioned, but traffic from each virtual network needs to be managed carefully (using mechanisms such as traffic shaping and/or admission control to maintain isolation and avoid unfairness between slices. In an ongoing study, we are evaluating virtualization of a WiMAX base station intended to serve as an open experimental platform in GENI. Such a WiMAX base station can be virtualized using the experimental setup shown in Fig. 2(a) in which an external virtual network controller with traffic shaping is used in conjunction with a standard 802.16e WiMAX radio. Our implementation uses cross-layer load and channel adaptive traffic shaping to maintain fairness between VNs when the channel approaches saturation. An example result with two virtual networks on WiMAX is shown in Fig. 2(b). The figure shows how the bit-rate of slice 1 varies with time due to signal fluctuations, and the corresponding bit-rate of slice 2 is also affected severely when traffic shaping is not used. The bit-rate trace with VNTS traffic shaping shows that this method is effective in maintaining bit-rate isolation between multiple virtual network slices. Ongoing work is aimed at studying interactions between the external virtualization controller and internal base station scheduler. Multi-node wireless network virtualization will also be considered next.

4.2 Cache & Forward (CNF) Protocol

Delay tolerant network (DTN) protocols [5] which were designed to handle intermittent link disconnections offer an interesting alternative to traditional TCP/IP for wireless scenarios. At WINLAB, we have been investigating a DTN-style clean-slate protocol for mobile content delivery called cache-and-forward (CNF) [14], under the NSF future Internet research program (FIND). CNF, shown schematically in Fig. 3, is based on the concept of routers with large storage, strict hop-by-hop transport of media files, content caching, and post-offices to support delivery to mobiles. This approach makes it possible to efficiently deliver content to mobiles which may occasionally be disconnected or suffer from poor link quality. Moreover, in-network storage and content routing makes it possible to implement media caching as a regular protocol feature, helping to further improve network throughput and delay.

The CNF protocol’s data plane consists of three layers: CNF transport protocol (CNF TP), CNF network protocol (CNF NP), and CNF link protocol (CNF LP). CNF LP is a reliable link layer protocol that transmits the entire content file on a hop-by-hop basis with storage at intermediate nodes. The CNF NP is an *integrated routing and storage protocol* which forwards content files based on long-term and short-term path metrics as well as the availability of storage along the path. The goal of the CNF routing algorithm is to maximize network throughput by holding files intended for a disconnected or poorly connected destination, while favoring those which currently have good quality

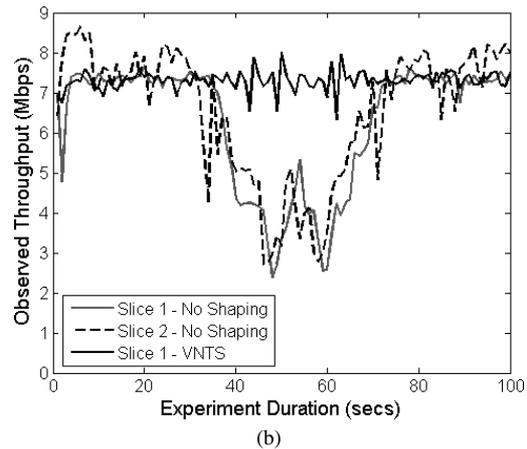
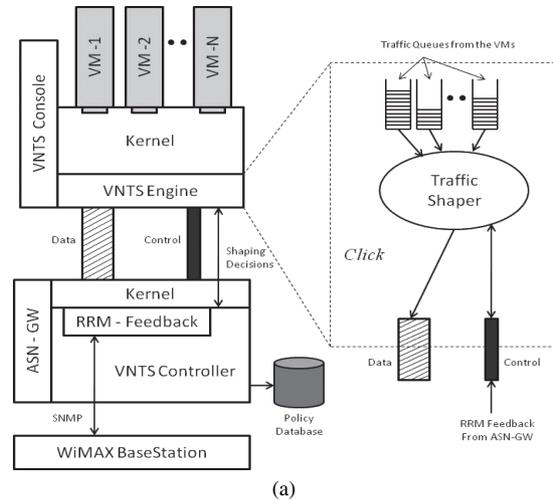


Fig. 2 (a) Architecture of virtualized GENI WiMAX BS, (b) VN performance with and without shaping.

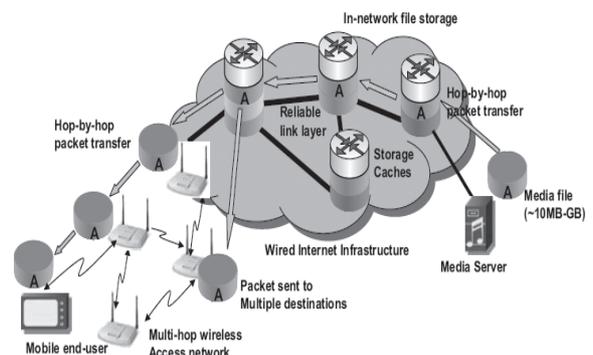


Fig. 3 CNF architecture concept.

paths. The CNF NP includes a simple form of congestion control based on buffer occupancy at the next-hop node; the complexity of flow control is avoided because of the strict store-and-forward architecture being employed.

In the ongoing FIND project, we have conducted extensive NS-2 simulations of CNF and compared the results with TCP/IP [15]. Results obtained so far show that the protocol offers significant (~2-5x) network throughput gain

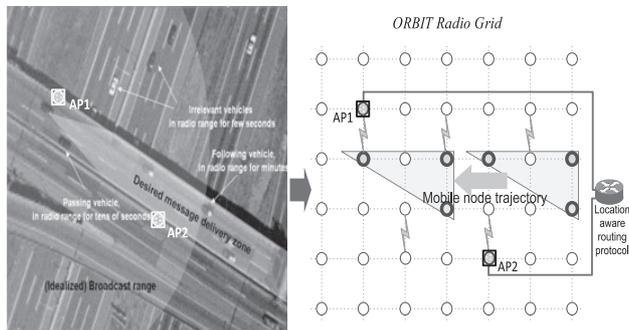


Fig. 4 Georouting & multicast scenario (figure courtesy of Prof. M. Gruteser).

(relative to TCP/IP) for representative mobile service scenarios. Additional CNF results with caching show that system capacity and file transfer delay can be further improved via in-network caching and content routing procedures. A proof-of-concept CNF prototype is also being developed on the ORBIT testbed.

4.3 Geographic Protocols

Network protocols with location awareness are becoming increasingly important for emerging pervasive computing applications in which the Internet is used to support physical world interactions. For example, in vehicular applications, safety messages may need to be delivered to cars approaching a road hazard. The group of vehicles that need to receive warning messages depends on their geographic position and their speed, as illustrated in Fig. 4. This is an example of a “geocast” service where the destination “address” is specified by GPS coordinates.

One of the clean slate FIND projects at WINLAB (led by Prof. Marco Gruteser) is aimed at the development of a “geometric stack (GEO)” with location-awareness at all protocol layers. The GEO stack has an interface to a localization plane which supports discovery of the position of neighbors and wireless access points in the area. It is expected that in most cases mobile nodes will obtain their own position through GPS receivers. In indoor environments, GPS may be augmented by WiFi localization methods which have been studied extensively during the past few years [16]. The GEO protocol stack consists of GeoMAC, Geocasting, and Geocaching protocol components that use location information to optimize message forwarding.

The GeoMAC protocol [17] uses location at the MAC layer for low-delay message forwarding in a vehicular ad hoc network. GeoMAC exploits spatial diversity, by allowing other nearby nodes to opportunistically forward and retransmit messages. The geo-backoff mechanism uses geographic distance to the destination as a heuristic to select the forwarder most likely to succeed. In trace-driven ns2 simulations using packet error measurements from a freeway environment, GeoMAC has been shown to provide bounded delays with up to 25% packet delivery rate gains, when compared with conventional AODV and GPSR routing proto-

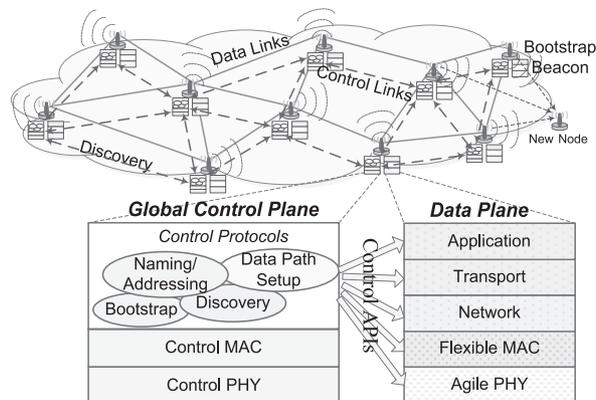


Fig. 5 CogNet protocol architecture.

cols.

4.4 Cognitive Radio Protocols

Cognitive software defined radios are an emerging wireless technology which are expected have a significant effect on the overall network architecture. One of the important features of cognitive radio is dynamic spectrum access (DSA), which involves identification of and opportunistic use of idle spectrum. Another key feature of cognitive radios is the ability to dynamically adapt physical layer, medium access control and routing parameters in response to observed spectrum usage, channel propagation and interference conditions. Cognitive radios can also cooperate with each other to improve communications quality, for example using cooperative PHY techniques such as network MIMO or network coding methods.

A key design issue for cognitive radio networks is that of defining a protocol architecture which supports DSA, spectrum etiquette, cross-layer adaptation and cooperative modes mentioned above. The *CogNet* protocol architecture under joint development by WINLAB, CMU and KU [18], [19] is based on the concept of a “global control plane (GCP)” which uses a separate physical channel for control messaging. The *CogNet* control plane shown in the figure below incorporates several key components including a radio bootstrapping function, a discovery process to provide global awareness of radios, a setup protocol for establishing multi-hop paths for collaborative data forwarding, and a naming/addressing scheme for allocating network addresses and name-to-address translation.

The data plane shown in Fig. 5 has an agile physical layer which can sense spectrum opportunities and rapidly move to newly available bands. In addition, the discovery module of the GCP can be used to implement a distributed spectrum coordination channel enabling dynamic spectrum sharing between radios with similar functionality. The flexible MAC layer supports switching between different media access mechanisms to achieve the best performance under different network topology and traffic conditions, e.g., in a sparse network, CSMA-based MAC may be appropriate,



Fig. 6 ORBIT radio grid at WINLAB tech center.

while in a dense network, it is preferable to use a TDMA-like MAC for scheduling to avoid excessive channel contention. The GCP provides a generic framework to exchange control information to implement these and other network adaptation functions. A prototype of the CogNet protocol architecture has been implemented using GNU radios on the ORBIT radio grid testbed, and preliminary results validate the use of this protocol stack for dynamic spectrum coordination [20] and adaptive MAC scenarios [21]. Future work is aimed at implementation of more general DSA algorithms and adaptive wireless networks with cognitive machine learning algorithms.

4.5 Open, Programmable Networks and Experimental Testbeds

Open, programmable networks are an important enabler for the future Internet because of their ability to support flexible experimentation and to evolve functionality as new network architectures are deployed on a trial basis. The NSF supported GENI initiative is an ongoing effort to build a national scale open programmable network using a combination of open switching, routing and wireless technologies. The main features of open networking devices used in such testbeds are: (a) an open API which provides access to link-layer technology parameters; (b) downloadable programmability of protocols used at the network layer; and (c) observability of key performance measures such as throughput and packet loss. At WINLAB, we have worked on open WiFi, SDR (GNU) and WiMAX devices with API's which support control and observation of key radio parameters such as frequency, transmit power, modulation/coding mode, MAC settings, etc.

The 400-node ORBIT radio grid at Rutgers (shown in Fig. 6) is a large-scale open wireless networking testbed which supports flexible and reproducible experimentation on future protocols [22]. The testbed currently incorporates a variety of open API radios including 802.11 a, b, g, n, Bluetooth, Zigbee, GNU/USRP and GNU/USRP2. The testbed also includes a high-speed wired control and management network used to download experiment code and collect mea-

surements. Backend servers can be used to emulate wired network routers, and gateways to wide-area testbeds such as GENI and PlanetLab are also available for end-to-end protocol studies. The radio grid is also supplemented by a number of outdoor and vehicular nodes (both WiFi and WiMAX) deployed on or around the Rutgers campus, to be used for real-world validation of results or for application trials. The ORBIT testbed described above has been used by over ~ 450 groups worldwide to conduct a wide range of future Internet experiments including protocols such as CNF, GEO and CogNet described earlier. Future work on ORBIT is aimed at further improving software defined radio capabilities, addition of WiMAX radios, integration with the GENI control framework, support for multiple simultaneous experiments (virtualization), and features for location-aware mobility scenarios.

5. Concluding Remarks

Future Internet research worldwide is reaching critical mass and will certainly have a major impact on future network architectures and protocols. As we point out in this position paper, it is important to recognize the central role of wireless and mobility in almost every future Internet usage scenario, and thus design in the necessary capabilities as "first class" protocol features rather than as extensions and overlays. This goal will require increased collaboration between different wireless networking areas and the wired/Internet protocol communities to avoid the risk of incompatible solutions. We believe that the development of large-scale open testbeds with protocol virtualization is an important step towards experimentation and eventual adoption of new networking architectures and technologies. Although clean-slate protocol ideas such as hop-by-hop transport, in-network storage, content routing, geographic routing and global control plane are still at an early stage, it is our expectation that many of these concepts will influence future Internet and wireless access network standards.

Acknowledgements

The author thanks his collaborators Ivan Seskar, Marco Gruteser, Max Ott, Sanjoy Paul, Yanyong Zhang, Shweta Jain, and PhD students Gautam Bhanage, Xiangpeng Jing, Lijun Dong and K.C. Huang for valuable discussions and contributions to the future Internet research topics discussed in this paper.

References

- [1] D. Raychaudhuri and M. Gerla, "New architectures and disruptive technologies for the future internet: The wireless, mobile and sensor network perspective," Report of NSF Wireless Mobile Planning Group (WMPG) Workshop, Aug. 2005, www.winlab.rutgers.edu/WMPG
- [2] C.E. Perkins, "Mobile networking for mobile IP," Proc. IEEE Internet Computing, Jan. 1998.
- [3] I. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: A

- survey,” *Computer Networks and ISDN Systems*, March 2005.
- [4] E.K. Lua, J. Crowcroft, M. Pias, R. Sharma, and S. Lim, “A survey and comparison of peer-to-peer networking schemes,” *IEEE Communications Surveys and Tutorials*, 2005.
- [5] K. Fall, “A delay tolerant network architecture for challenged inter-nets,” *Proc. ACM SigComm*, 2003.
- [6] D. Culler, D. Estrin, and M. Srivastava, “Guest editors’ introduction: Overview of sensor networks,” *IEEE Computer Magazine*, Aug. 2004.
- [7] A. Zanella, E. Fasolo, C. Padova, F. Chiasserini, M. Meo, M. Torino, M. Franceschinis, and A. Spirito, “Inter-vehicular communication networks,” 2nd Internal NEWCOM Workshop, May 2006.
- [8] J. Crowcroft, S. Hand, R. Mortier, T. Roscoe, and A. Warfield, “Plutarch: An argument for network pluralism,” *Proc. ACM SIGCOMM Workshop on Future Directions in Network Architecture*, Aug. 2003.
- [9] Planetlab testbed, “An open platform for deploying and accessing planetary scale services,” <http://www.planet-lab.org>
- [10] Global environment for network innovations (GENI), <http://www.geni.net>
- [11] A. Bavier, N. Feamster, M. Huang, L. Peterson, and J. Rexford, “In VINI Veritas: Realistic and controlled network experimentation,” *Proc. ACM SIGCOMM*, Aug. 2006.
- [12] R. Mahindra, G. Bhanage, G. Hadjichristofi, I. Seskar, and D. Raychaudhuri, “Space versus time separation for wireless virtualization on an indoor grid,” *Proc. IEEE NGI*, April 2008.
- [13] S. Singhal, G. Hadjichristofi, I. Seskar, and D. Raychaudhuri, “Evaluation of UML based wireless network virtualization,” *Proc. IEEE NGI*, April 2008.
- [14] S. Paul, R. Yates, D. Raychaudhuri, and J. Kurose, “The cache-and-forward network architecture for efficient mobile content delivery services in the future Internet,” *Proc. ITU NGN Conference*, 2008.
- [15] H. Liu, Y. Zhang, and D. Raychaudhuri, “Performance evaluation of the cache-and-forward (CNF) network for mobile content delivery services,” *Proc. ICC*, 2009.
- [16] Y. Cheng, Y. Chawathe, A. LaMarca, and J. Krumm, “Accuracy characterization for metropolitan-scale Wi-Fi localization,” *Proc. ACM Mobisys*, 2005.
- [17] S. Kaul, M. Gruteser, R. Onishi, and R. Vuyyuru, “GeoMAC: Geo-backoff based cooperative MAC for V2V networks,” *Proc. IEEE Conf on Vehicular Electronics & Safety*, 2008.
- [18] D. Raychaudhuri, N.B. Mandayam, J.B. Evans, B.J. Ewy, S. Seshan, and P. Steenkiste, “CogNet — An architectural foundation for experimental cognitive radio networks within the future Internet,” *Proc. ACM MobiArch*, Dec. 2006.
- [19] X. Jing and D. Raychaudhuri, “Global control plane architecture for cognitive radio networks,” *Proc. IEEE CogNets 2007 Workshop*, June 2007.
- [20] X. Jing, S. Anandaraman, M. Ergin, I. Seskar, and D. Raychaudhuri, “Distributed coordination schemes for multi-radio Co-existence in dense spectrum environments: An experimental study on the ORBIT testbed,” *Proc. DySpan*, 2008.
- [21] K.C. Huang, X. Jing, and D. Raychaudhuri, “MAC protocol adaptation in cognitive radio networks: An experimental study,” *Proc. ICCCN*, 2009.
- [22] D. Raychaudhuri, I. Seskar, M. Ott, S. Ganu, K. Ramachandran, H. Kremo, R. Siracusa, H. Liu, and M. Singh, “Overview of the ORBIT radio grid testbed for evaluation of next-generation wireless network protocols,” *Proc. IEEE WCNC*, 2005. [also see www.orbit-lab.org]
-