Modems to Mars

Gregory J Pottie
Electrical Engineering Department
University of California, Los Angeles
Pottie@icsl.ucla.edu

Abstract

The practice of modem engineering has generated many interesting theoretical problems, which now apply to a vast range of systems. The title of this paper has dual meaning: the engineering path leading from voice-band modems to communications in distributed sensor networks for Martian exploration, and preliminary plans for the establishment of an interplanetary Internet.

1. Introduction

Modems are one of the basic building blocks of the Internet, and voice-band modem technology has shown the way to a wide variety of popular access devices including cable TV modems and digital subscriber lines. Recent advances in integrated circuit fabrication technology now permit the construction of compact and low-cost devices that include signal processing, sensing, and wireless communication. This will enable the extension of the Internet to the physical world. Simultaneously, new extensions of the Internet are being planned for our solar system, with an entirely new vision of how communication infrastructure will support science missions and human exploration. This paper presents an overview of the technological path from voice-band modems to ubiquitous communication networks.

2. Some Really Cool Things about Voice-Band Modems

A bias must be acknowledged in that my graduate digital communications course is centered on the study of V.34 modems, with excursions into other technologies. The main point I make about such modems is that the channel resource was extremely limited and thus the baud rates low, which meant that many signal processing cycles can go into each bit transmitted and received. The result was that very advanced techniques were used that only later were applied to higher speed systems. It is a consumer application in which Shannon's limits are approached, over a medium never intended to support data communications, in a product that has become a low-priced commodity. Along the way, new concepts in line probing, adaptive equalization, constellation shaping, coded modulation, combined coded and shaped modulation and equalization, echo cancellation and synchronization were developed. A backwards compatible family of devices was produced such that each succeeding generation could actually in a bug-free manner interoperate with preceding generations. A standards process took place that actually spurred rather than smothered technological change. These are all significant accomplishments that merit academic study on their own.

3. V.34 to GigaBit Ethernet

The very fast or very cheap communication systems usually first have an analog solution, which is quite fortunate for digital designers. Since it is difficult to get anything analog to work that involves loops of degree larger than 2, the channels are inadvertently engineered in such a fashion that adaptive least squares techniques are nearly certain to converge. The obvious success in adaptive techniques in raising voice band modem speeds from 1200 bits/s to 28.8 kb/s over "analog" lines has paved the way for many of these techniques to be applied to digital subscriber lines, cable modems, and high speed ethernet. However, the demands of higher speed were met not by simply having more general purpose DSPs, but by design of ASICs capable of carrying out the most common operations with a fraction of the chip area and thus cost and energy. Here the most impressive technological change from the design point of view is the improvement in computer aided design tools for ASICs, which enable experienced designers to very rapidly turn out new chips using standard cell libraries. New cells need to be created only for the most critical functions. These designers need to know both communications theory (what algorithms are suitable) and be versant in the tools. A recent example of the success of this approach is the very rapid growth of Broadcom, with a variety of high speed chips with extensive adaptive features being produced on short design cycles.

4. Sensor Networks

Wireless sensor networks are an example of systems in which energy resources are very highly constrained [1-6]. In such systems, each node includes sensors, signal processing, and a radio, with the objective of detecting/identifying events with high probability. Prior architectures would typically send raw data to a central site for processing. However, given that 3,000,000 operations can be executed for the same energy as it takes to transport 1 kbit 100 m (for typical parameters) [6], it makes much more sense to process and make decisions in the nodes insofar as possible. Shannon's limit and Maxwell's Law together dictate that there will always be a minimum energy cost to transmit the bits, while continuing advances in IC technology indicate we are far from the minimum energy costs for signal processing. Further, by coming to decisions rather than transporting the raw data, we ease bandwidth limitations in the network, promoting scalability. The resulting bit rates on any given link will be low, and many messages will be very short. Here advanced communications techniques are needed not for increasing the bandwidth efficiency, but for conserving transmit energy, and avoiding costly retransmissions. Early prototype systems are described in [1,2], with [5,6] describing later systems, and [6] providing a general overview of the principles of wireless sensor network design.

The high cost of communications dictates distributed rather than centralized algorithms for a wide range of sensor functions, from networking to data fusion. The large numbers of nodes further imply that the network should self-organize for these tasks, rather than relying on human intervention to set-up or adjust to failures of nodes. A set of algorithms to accomplish this are presented for example in [4]. The severe energy constraint is in part mitigated by the loose latency requirements for many of the applications. Indeed, progress on this problem is made only when considering the complete system from physical layer through to application. Challenges include questions of network hierarchy (flat or multi-tiered), interaction of signal processing, storage, and communication, heterogeneity in the hardware and environment, and interaction with remote networks such as the Internet with their associated resources. The first commercial products that deal with these questions are beginning to appear, but we have not even constructed the theoretical scaffolding for dealing with issues of fundamental limits. Like voice-band modems, this seems to be an area where very interesting theory will be spurred by practical need.

Figure 1 illustrates how sensing, networking, signal processing, and database management are all intertwined with wireless sensor networks. Consider a large network whose purpose is to detect and identify particular physical events. Individual nodes are powered off some combination of batteries and solar cells so that energy is at a premium. As noted above, this compels processing as much of the data at source as possible so as to limit communications. Further, there will be a signal processing hierarchy within the nodes themselves. Energy-detection circuits use little power, and can be constantly vigilant. However, to meet detection probability requirements they will produce an unacceptable rate of false alarms. Thus sampled data can be queued, and in the event that the energy detection threshold is exceeded some higher level of signal processing can be invoked, e.g., frequency analysis. By queuing data and proceeding through a sequence of operations of increased sophistication (lower false alarm rates, higher energy consumption), energy is expended only to the extent required while assuring performance targets are met. Yet a single node cannot on its own provide accurate position information for targets, and may have insufficient SNR for producing a reliable decision about an event. Thus, collections of nodes may form ad hoc networks to perform data fusion (exchange of likelihood function values) or coherent beamforming (raw data) [3]. This step is taken only when lower-energy procedures fail to produce the desired results. Finally, information can be multi-hopped back towards an end user, possibly being queued and aggregated along the way, for example to produce activity summary reports. This possibility of aggregation expands network scalability.

The end user may actually desire more information, and has the option to query the network to provide further details about particular events. Thus queued data may be recovered and transmitted. For example, for data on which the network has come to a murky decision a central site may perform highly sophisticated adaptive processing, the end result of which will be new detection table parameters being sent to all the nodes. For the highest detection/identification reliability, the end user must retain the ability to drill down to the lowest level in the network. At the same time, to avoid information overload, bandwidth contention, and massive energy usage the nodes must only rarely be in the situation where the transport of raw data to the end user is required.

Notice that in this example there is no clear division between signal processing, database issues, and networking. It is rather a unified process, in which design decisions at one level have a large impact on other levels.

5. Robot Ecologies

Sensor nodes scattered upon the ground can end up in disadvantageous locations with respect to antenna elevation or sensor field of view. The thought naturally occurs that it would be nice if either the whole node or an appendage could move to the more favorable location. The advantage of extending an arm is that we do not have to the energy or complexity cost of moving the whole unit. On the other hand, having some mobile elements may allow holes in the network to be filled, or afford the possibility of physical exchange of information and energy. In this scenario, the fixed robotic elements are the analog of plants--accumulating energy--and the mobile elements are the analog of animals--using high energy supplies. Such a robot ecology [7] provides the possibility of sustained autonomous operation. The plants provide the infrastructure that makes possible the long-term deployment of mobile objects, and in addition are useful in themselves in collecting information. A very wide variety of plants can be conceived: balloon-launchers (roots tap down to water, and generate hydrogen by electrolysis), burrs, tumbleweeds, lily-pads, seed launchers, creeper vines (say propelled by mobots), etc. Plants could provide navigational assistance, in effect trading information against energy consumption. The number and complexity of the interactions that are possible makes clear the need for a more formal theory of cooperation. At the moment, cooperation is pursued on an ad hoc basis, and typically with either a very rigid hierarchy (master/slave) or actions among homogeneous peers. Extensions to large numbers of heterogeneous devices is very challenging. Never the less, a broad range of applications are possible with even relatively simple robotic ecologies.

Figure 2 illustrates a fanciful ecology for performing a geological survey of Mars, with mother plants equipped with creeper vines and sensor seed launchers, mobile diggers and trucks, wind-propelled sensor tumbleweeds and balloons, and air-drop of fresh supplies. It is of interest that the rover in the celebrated Mars Pathfinder mission moved not all that much further than spring-launched sensor seeds could be sent. Energy constraints are even more important in planetary exploration applications than for Earth-based systems due to the huge expense of transporting heavy power-generation and storage units. Thus, unconventional systems which opportunistically use local resources (to live off the land) may be an interesting avenue for future research.

6. The Martian Internet

Plans are afoot for a sustained series of robotic missions to Mars. In contrast to the practice of the past, it is intended that missions will contribute to a common infrastructure with the goal of supporting robot and eventually human colonies on the red planet. Part of this task will be construction of an interplanetary network (IPN) [8]. Eventually, a network of communications/position location satellites will be created for Mars, in support of a local internet similar to the terrestrial internet. Management of traffic between Earth and Mars however will be challenging due to the large latency. Standard Internet protocols will not work; the signaling will be more akin to sending email than carrying on an interactive session. It is anticipated that the traffic flow will be highly asymmetric (from Mars to Earth, since the information consumers are terrestrial).

Presently, communications time is allocated according to committee decisions. This will clearly not be possible when there are many scientific missions in progress, as well as requests from the public at large. Market auction algorithms show some promise in reducing complicated allocation criteria to a single parameter that can determine access priority. Such mechanisms may also be used to provide incentives for missions to create the necessary infrastructure (e.g., awarding Mars phone cards which can then be sold for real money).

7. Grand Information Theory Challenges

Distributed sensor and robotic networks bring large new challenges to information theory. In sensor networks, the basic problem is to identify classes of events to a required level of fidelity (missed/false detection) using the minimum resources (energy, bandwidth) within latency constraints. Neglecting latency, this can be cast as a network rate-distortion problem with the resources in the role of rate and fidelity in the role of distortion. Resource usage includes that needed for determining which set of sensor nodes will be involved in the decision, including inhibition of others from participating. There are also capacity problems, in that groups of nodes may cooperatively signal to overcome gaps in the network. How to form the transmitting and receiving groups, what cost is associated with maintenance of a particular level of synchronism, and how to optimally code are all unknown. Further, we may have to deal with heterogeneous devices.

The challenges in robotics are even greater, in that we must deal with both of the above problems in the context of mobility, with interactions between fixed and moving assets. Vast latencies are possible in attempting to control the operations of robot colonies on distant planets, so that adjustments can only be made to parameters controlling behaviors, rather than individual actions. A very interesting interaction between theory and practice is before us.

References

1. K. Bult, A. Burstein, D. Chang, M. Dong, M. Fielding, E. Kruglick, J. Ho, F. Lin, T.-H. Lin, W. J. Kaiser, H. Marcy, R. Mukai, P. Nelson, F. Newberg, K. S. J. Pister, G. Pottie, H. Sanchez, O. M. Stafsudd, K. B. Tan, C. M. Ward, S. Xue, J. Yao, “Low Power Systems for Wireless Microsensors”, *Proceedings of International Symposium on Low Power Electronics and Design,* pp. 17-21, 1996.
2. G. Asada, M. Dong, T. S. Lin, F. Newberg, G. Pottie, H. O. Marcy, and W. J. Kaiser, "Wireless Integrated Network Sensors: Low Power Systems on a Chip" , *Proceedings of the 24th IEEE European Solid-State Circuits Conference*, 1998.
3. K. Yao, R.E. Hudson, C.W. Reed, D. Chen, F. Lorenzelli, “Blind Beamforming on a Randomly Distributed Sensor Array System,” *IEEE J. Select. Areas in Comm.*, vol. 16, No. 8, Oct. 1998, pp. 1555-1567.

4. K. Sohrabi, J. Gao, V. Ailawadhi, G. Pottie, “A Self-Organizing Sensor Network,” *Proc. 37th Allerton Conf. On Comm., Control, and Computing*, Monticello IL, Sept. 1999.
5. J.R. Agre, L.P. Clare, G.J. Pottie, N.P. Romanov, “Development Platform for Self-Organizing Wireless Sensor Networks” *Aerosense'99*, Orlando FL, April 1999.
6. G.J. Pottie and W.J. Kaiser, "Wireless Integrated Network Sensors," Commun. Of the ACM, vol. 43, No. 5, May 2000, pp. 51-58.
7. G.J. Pottie and R. Brooks, "Towards a Robotic Ecology," DARPA ISAT Study, 1999.
8. V. Cerf et al., "Interplanetary Internet Seminar," http://www.ipnsig.org/reports/CalTechSeminar1.html, Oct. 1999.