EFFECT OF BURSTINESS

A. Ephremides

Since the early days of Information Theory it was realized that a major limitation of its use in Communication Systems is the inability to handle bursty traffic and queueing delay. On the other hand in the context of communication networks bursty data and delay in store-and-forwarding of packets are central and indispensable concepts. Moreover, while the data "primitives" in Information Theory are endless strings of symbols, the basic data units in networks are groups of symbols called packets. Usually the time duration of a packet and/or its content (in terms of number of symbols) are suppressed and considered irrelevant since rate is measured in "packets per slot".

These differences have given rise to three distinct measures of "rate" performance:

1) **Capacity**: This is the traditional Shannon theoretic maximum achievable numbers of symbols/s over a communication resource that can be transmitted reliably. In multi-user systems this is a multi-dimensional region. It is unknown except for the most rudimentary of multi-user systems. A crucial element in its definition is that all users have infinite reservoirs of data from which they pick to transmit. In other words, all users are "backlogged" and their "queues" are assumed "saturated" (which means that they are never empty).
2) **Maximum Throughput**: This is a traditional networking measure that is usually, again, a multi-dimensional region of "rates" expressed in packets/slot that are achievable over a given network. The users are considered backlogged here as well. Hence, this quantity measures the "maximal" set of packet flows that can be sustained in the network when all users are loaded with unlimited amounts of data traffic. This is usually easier to characterize in most networks. If the conversion of packets/slot to bits/sec can be defined (and it is not always easy or even possible to do that), the maximum throughput region usually coincides with the Capacity Region. It is certainly outer-bounded by it.
3) **Maximum Stable Throughput**: This is a nontraditional but recently increasingly popular measure of rates (in packets/slot) that is especially meaningful and relevant in wireless networks. It is defined only for users that are **not** backlogged who receive bursty traffic that is queued up at their buffers while awaiting transmission. The "arrivals" of the bursty traffic must have the same rate as the "departures", that is the successful deliveries to the destinations. Equivalently, this requires the "queues" to be **stable**. There are several definitions of stability but they all model, more or less, the need for the queue lengths to either stay finite or have a non-degenerate limiting probability distribution. The important property of this measure is that it need not coincide with the maximum throughput measure. In fact, in many cases it outer-bounds the latter. The reason for this is that in a network (especially, but not only, a wireless network) the queues of the users are "coupled" or, as we say, "interact". Thus, when, as it must with probability one, under the stability requirement, a queue empties, it ceases (albeit

temporarily) to complete for communication resources with the other queues. Therefore, the latter can achieve higher rates as long as that queue remains idle.

Based on these observations we want to point out a rather important and fundamental observation that may appear counter-intuitive. **When traffic is bursty, the capacity region of a multi-user system generally increases.** Even though the conversion of packets/slot to bits/sec has the same difficulties as before, it is clear that such a correspondence is possible and can be quantified. So, the expansion of the achievable packet delivery rate region under stability constraints opens up a vast and new research vista for the "traditional" Shannon Theory.

Examples of cases in which the maximum stable throughput region outer-bounds the maximum throughput region include the random access channel as well as cooperative multi-access systems, with, or without, a relay node, first described in [1]. The ramifications and consequences of this reality are only beginning to be understood and/or explored.

Bottom Line: The traditional view (i.e. when all users remain backlogged) Leads to lower data rates!

References

1. A. Sadek, A. Ephremides, R. Liu, "Cognitive Multiple Access via Cooperation: Protocol Design and Performance Analysis", **IEEE Transactions on Information Theory**, Vol.10, pp.3677-3696, October 2007.