White Paper

BATS:

An Enabling Communication Technology for IoT and Beyond

by Shenghao Yang and Raymond W. Yeung

About the Authors

The authors are the inventors of BATS, a disruptive network coding technology that makes wireless multi-hop communication a reality. Network coding is a cross-disciplinary field of research co-founded by Raymond Yeung at the turn of the last century that has induced a paradigm shift in network communications. Shenghao Yang and Raymond Yeung are faculty members at The Chinese University of Hong Kong.[1] They are also co-founders of n-hop technologies that pioneers the applications of BATS.

The Wireless Challenge

Internet of Things (IoT) is a network that will connect billions of physical devices such as vehicles, home appliances, sensors, actuators, etc., and enable them to exchange data among themselves. Such a network can be realized by fog computing [9][10] that builds a continuum of interface between the cloud and the things. Due to the size of the network, when two devices communicate, most likely they need to go through a large number of other devices, referred to as relays. Many of these IoT devices are connected to a wireless instead of a wired network, and some of them can be mobile devices. Here, wireless may refer to technologies for short-range wireless (e.g., Bluetooth, Wi-Fi, ZigBee), medium-range wireless (e.g., LTE), long-range wireless (e.g., VSAT for satellite communication), or underwater communication (e.g., acoustic, optical). In some cases, data exchange between IoT devices can be achieved through power-line communication.

Packet Loss Accumulation

A main challenge for all wireless (and power-line) communications is data packet loss. Depending on the specific technology and the application scenarios, this may be due to noise, interference, path loss, multi-path fading, Doppler spread, etc. In the Internet, the transport and network layers are dominated by TCP and IP, respectively. TCP uses retransmission and rate control to resolve the end-to-end packet loss, and IP forwards received packets at intermediate nodes. When the source node transmits data packets to the destination node via a number of relay nodes, each packet must be successfully received on each hop before it can reach the destination node. As such, the end-to-end packet loss is simply an accumulation of packet loss on the individual hops.

For example, if one packet is transmitted on each hop per unit time and the packet loss on each hop is 0.2 (assume without retransmission), which is quite common in certain applications (e.g., low-power IoT, WiFi under high interference, terabit satellite), then after 10 hops, only $(1-0.2)^{10} \approx 0.1$, or 10%, of the packets are left. This means that the end-to-end throughput cannot exceed 0.1 packet per unit time. This theoretical upper bound can be attained by applying suitable end-to-end ***forward error correction*** (FEC), e.g., fountain code [1]. If instead TCP/IP is applied, the throughput would be even lower due to protocol overhead.

Hardly Goes Beyond a Few Hops

As such, we very rarely see wireless networks that have more than a few hops, not to say tens or even hundreds of hops. In other words, if a device communicates with another device

[1] Shenghao Yang and Raymond Yeung are with the Shenzhen and Hong Kong campuses, respectively.

through a number of other devices and the throughput is to be maintained at any reasonable level, only very few of those devices along the path can be wirelessly connected. In fact, in today's Internet, almost all network connections involve at most two wireless hops, namely the first hop and the last hop. This shortcoming of wireless communication will very soon become a bottleneck for the development of IoT because it prevents the massive deployment of wireless IoT devices.

Network Coding

In the past few decades of network communication, data packets are transmitted through the network by ***store-and-forward***, also known as ***routing***. The rationale supporting this telecommunication paradigm is the folklore that information behaves like a commodity. This, however, was refuted by network coding theory [2], which shows that in order to achieve the network capacity, coding needs to be employed at the intermediate nodes of the network.

Fig. 1a: Store-and-forward.

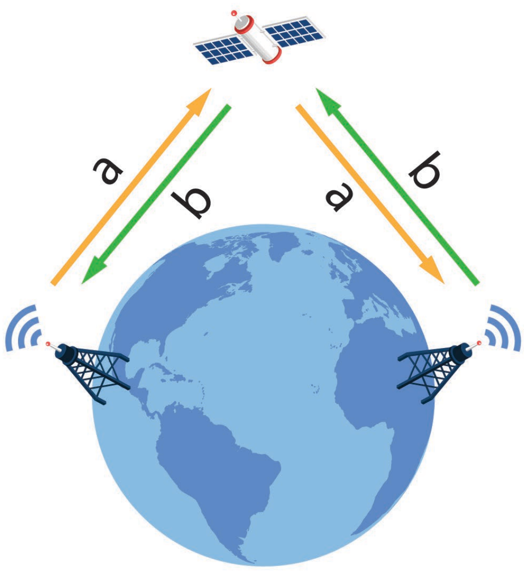
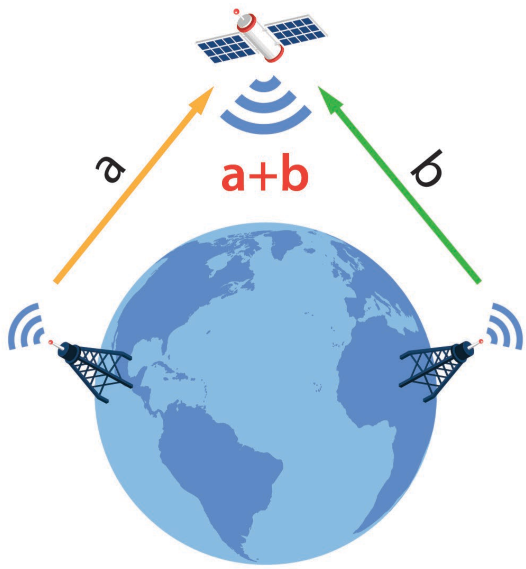


Fig. 1b: Network coding.



Information ≠ Commodity

The fundamental concept of network coding can be explained by means of a very simple example. In Fig. 1a, two transmission towers need to exchange two bits, a and b, through a satellite. With store-and-forward, a is first sent to the satellite by the tower on the left, and then broadcast by the satellite to the other tower. The same for b. This takes a total of 4 time slots to accomplish. In Fig. 1b, the two towers first send their own bits to the satellite in 2 time slots. Then the satellite performs a computation on the two bits and generates a new bit $a + b$ ("+" denotes *binary addition* or XOR) and broadcasts the bit to both towers. This recoding operation inside the network is referred to as ***network coding***. The tower on the left, upon receiving $a + b$, can decode the other bit b by adding its own bit a and the received bit $a + b$:

$$a + (a + b) = (a + a) + b = 0 + b = b.$$

Similarly, the tower on the right can decode the other bit a. Thus in only 3 time slots (as opposed to 4 time slots with store-and-forward), both towers can receive the other bit. So, although the two bits a and b are not "explicitly transported" to the other tower as in Fig. 1a, the information does go through. This very simple example clearly shows the inadequacy of the classical commodity view of network communication. Rather, network coding needs to be employed in order to achieve optimality.

The above example is the simplest case of ***linear network coding***, meaning that the mathematical operations involved are linear operations over a finite field. Linear network codes can be encoded, recoded, and decoded much more efficiently compared with nonlinear network codes. In general, optimal linear network codes exists for any given network topology [3][4]. This paves the way to application of network coding in real systems.

A Paradigm Shift in Network Communications

Network coding induces a paradigm shift in network communications. In the past, a network solution merely consists of a number of point-to-point solutions (e.g., channel coding, encryption) that are "glued together" by routers as if information is a commodity. With the advent of network coding, many of these network solutions need to be revisited. Since its inception, network coding has been applied to many different domains including channel coding, wireless communication, distributed data storage, cryptography, and even quantum computing. Over 10,000 research papers have been written on network coding and its applications.

Network coding theory unveils the non-commodity nature of information. Unlike traditional network protocols that put great emphasis on link-by-link reliability, network coding only focuses on end-to-end reliability. This principle of network coding has spawned a new class of codes to be discussed in the next section that provides a practical solution to the longstanding problem of packet loss accumulation in multi-hop networks, sometimes called the "multi-hop curse".

Breaking the Multi-hop Curse

As discussed, in a multi-hop network, if each hop has a certain packet loss, then the fraction of packets that can go through drops exponentially fast with the number of hops. This is inevitable with the commodity view, because each packet must be successfully received on each hop before it can reach the destination node.

Can network coding help? The answer is yes. By applying a network coding technique called random linear network coding (RLNC) [5], ***the throughput of a multi-hop network can be maintained regardless of the number of hops*** [6][7]. In RLNC, at a relay node, instead of forwarding the received packets, different random linear combinations of the received packets are transmitted to the next node. However, the complexity of RLNC is too high to be implemented on most devices, in particular many IoT devices that have low computing power and a small memory, and may even be battery-powered.

Introducing BATS

To meet this implementation challenge, a class of efficient network codes called BATched Sparse codes (BATS) was invented by the authors [8][11]. In BATS, the source node encodes the source file into ***batches*** of coded packets. These packets are transmitted through the network, where some of them may be lost along the way. At a relay node, ***recoding*** is applied to packets received at that node that belong to the same batch. This network coding operation compensates for the packet loss at the upstream of the network. At the destination node, the source file can be decoded as soon as a sufficient number of coded packets have been received.

The high efficiency of BATS enables it to be implemented on most platforms, including IoT devices. In particular, BATS satisfies all the requirements for a practical solution:

1. high throughput
2. low latency
3. low coding complexity
4. low storage requirement at the relay nodes.

BATS is the only scheme that can satisfy all these four requirements. Due to the pipelining nature of BATS, the recoding delay incurred at a relay node is minimal. This makes BATS applicable to real-time applications such as video streaming. For the technical details of BATS, we refer the reader to the monograph [12].

Sustains Hundreds of Hops

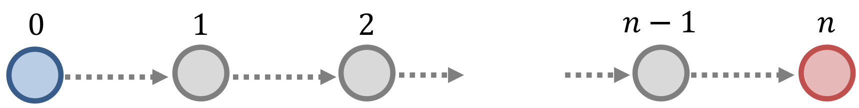


Fig. 2: A line network with n hops.

Consider the line network in Fig. 2 where the packet loss on each hop is 0.2 and the total number of hops is n. After 1 hop, the fraction of packets that remain is 0.8. This is also the theoretical upper bound on the throughput that can be achieved by any scheme on an n-hop network. When the number of hops increases, the throughput is expected to drop. If the relay nodes simply store and forward the packets (e.g., TCP/IP, fountain code), the throughput would drop exponentially fast with respect to the number of hops.

Fig. 3 compares the throughputs of BATS and TCP/IP for up to 50 hops. We see that after 50 hops the throughput is about 0.7, i.e., a 12% drop from 0.8, whereas the throughput of any store-and-forward based protocol such as TCP/IP is already close to 0. Fig. 4 shows the throughput of BATS for up to 1,000 hops. We see that even after 1,000 hops the throughput still maintains at about 0.66, i.e., only a 17.5% drop from 0.8. Thus BATS essentially converts a multi-hop network with packet loss into a single-hop network.

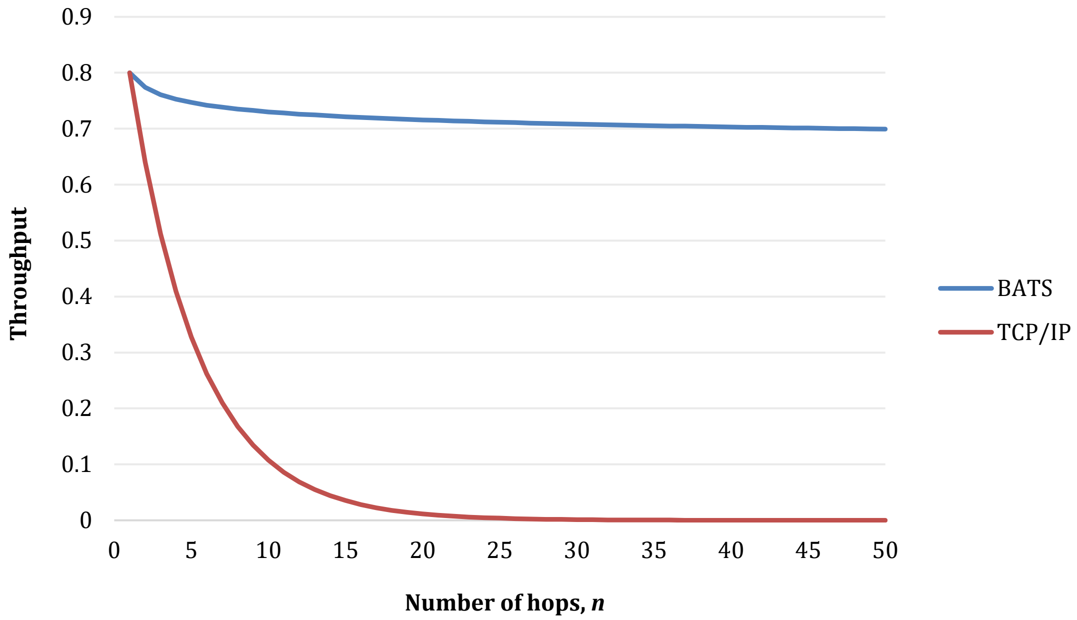


Fig. 3: Performance of BATS up to 50 hops.

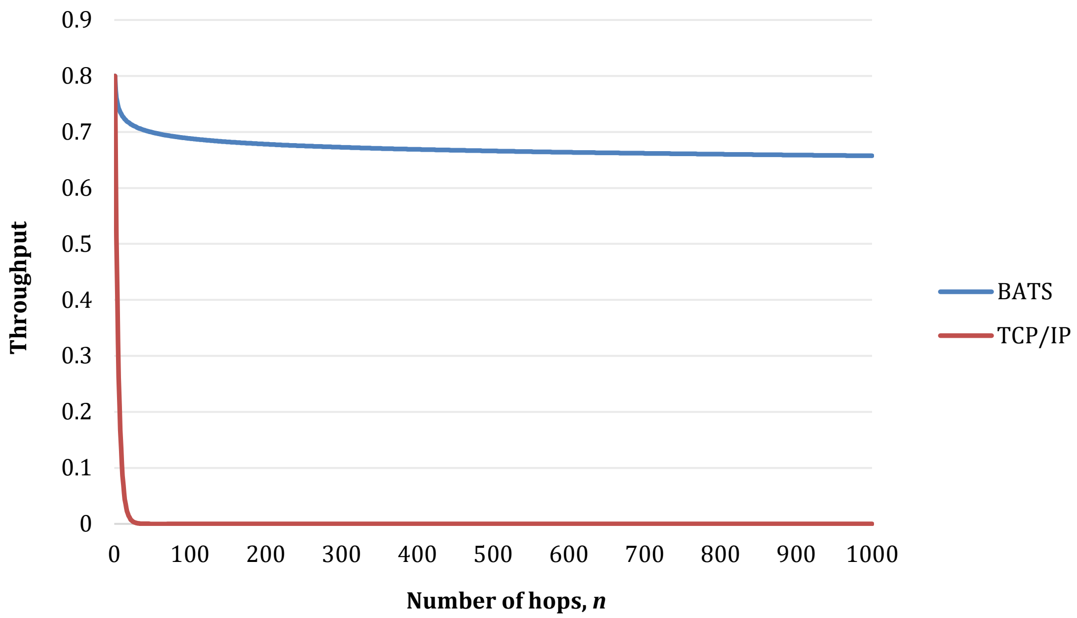


Fig. 4: Performance of BATS up to 1000 hops.

An additional benefit of BATS is enhanced security, because the coding employed by BATS is random. This means that the transmission by BATS is a randomization of the message data, so that the adversary cannot possibly decode the message by incepting part of the transmission.

Performance Proof

BATS is a disruptive technology developed by n-hop technologies that results from 20 years of research on network coding, all the way from a fundamental mathematical theory to a practical scheme that can be implemented on most IoT devices, for example, Raspberry Pi. The following video demonstrates the superior performance of BATS for video streaming over an 11-hop network:

n-hop.com/BATS.mp4

Use Case: Smart Lampposts

BATS can dramatically improve the performance of any multi-hop network with packet loss. Such networks are ubiquitous in 5G access networks, V2X, wireless mesh networks, low-orbit satellite systems, underwater communication, powerline communication, to name a few. In this section, we focus on our first deployment of the BATS technology in a smart cities application.

Smart lampposts have been recognized as the key infrastructure of smart cities worldwide. They support a wide range of smart cities applications including autonomous driving, intelligent transportation, real-time surveillance, and high-speed WiFi coverage on a city scale. In short, smart lampposts are indispensable for smart cites. Navigant Research estimates that the annual smart lamppost revenue will grow to nearly $8.3 billion globally by 2027.

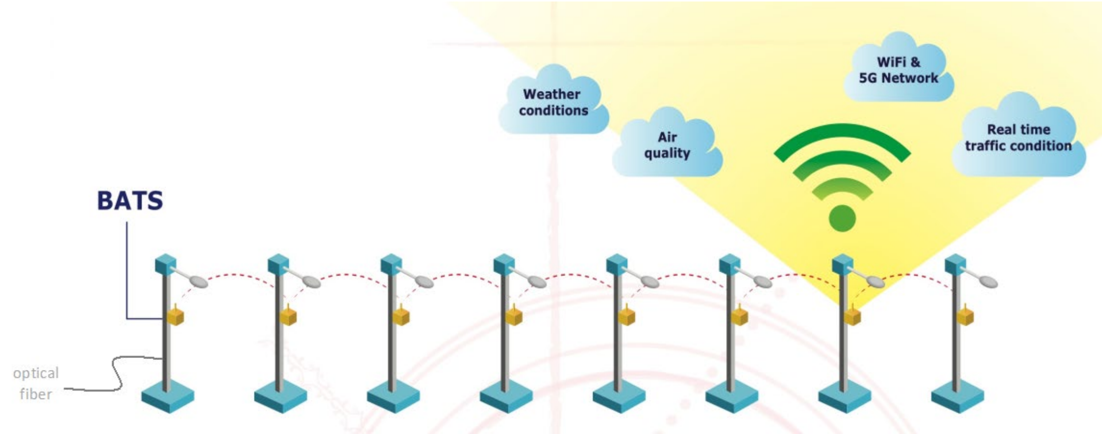


Fig. 5: A smart lamppost system powered by BATS.

In summer 2019, BATS developed by n-hop technologies was successfully deployed in the Hong Kong pilot smart lamppost system – one of the most advanced such systems in the world. Smart lampposts cannot be smart if they are not connected to the Internet. BATS can provide smart lampposts with wireless Internet connection without the need to lay optical fibre to every smart lamppost. Laying new optical fiber not only is expensive but also has a very long setup time. In some cities it can take a few years to obtain all the necessary licenses for the construction. Also, the construction process is very disturbing to the traffic and pedestrians.

An application scenario is illustrated in Fig. 5, in which the optical fibre connection at the smart lamppost on the left is extended wirelessly to the other smart lampposts in a multi-hop fashion by means of BATS. The hardware on each smart lamppost, called a ***BATS box*** (see Fig. 6), is co-developed with Advantech, the leading industrial PC vendor.

Fig. 6: A BATS box (for illustration only).

Instead of BATS, an alternative solution is to provide wireless Internet connection to each smart lamppost by means of 4G/5G. Compared with BATS which is based on WiFi, the operating expense (OPEX) of the 4G/5G solution is very high because a fee is needed for the cellular service for each smart lamppost. For a city with several ten thousand smart lampposts, the service fee can easily be a few hundred thousand US dollars per month as data are uploaded 24 hours a day. For 4G, the bandwidth provided by the network can drop drastically during rush hours.

Thus BATS is a proven technology that provides a city-friendly and economical solution for massive deployment of smart lampposts, which in turn will expedite the advent of smart cities and improve the quality of life around the world.

Conclusion

It is predicted that billions of IoT devices will be connected to the Internet by multi-hop wireless links, where TCP/IP, based on store-and-forward at the intermediate nodes, will meet difficulty. BATS is a disruptive network coding technology that solves the packet loss accumulation problem in wireless networks, making such networks of tens or even hundreds of hops possible. This enables the massive deployment of wireless IoT devices in the fog computing environment.

Finally, in addition to IoT, BATS can also be applied in satellite, deep-space, powerline, and underwater communication networks. These applications of BATS are expected to fundamentally change the landscapes of the related industries.

Licensing

Enquiry about licensing the BATS technology is most welcome. Please send email to *info@n-hop.com.*

References

[1] J. W. Byers, M. Luby, M. Mitzenmacher, and A. Rege, “A digital fountain approach to reliable distribution of bulk data,” in *Proceedings of ACM SIGCOMM’98*, pp. 56–67, New York, NY, 1998.

[2] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, “Network Information Flow,” *IEEE Transaction on Information Theor*y, IT-46: pp. 1204-1216, 2000.

[3] S.-Y. R. Li, R. W. Yeung and N. Cai, “Linear network coding,” *IEEE Transaction on Information Theor*y, IT-49: 371–381, 2003.

[4] R. Koetter and M. M ́edard, “An algebraic approach to network coding,” *IEEE/ACM Transactions on Networking*, 11: 782–795, 2003.

[5] T. Ho, M. Medard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, “A Random Linear Network Coding Approach to Multicast,” *IEEE Transaction on Information Theor*y, vol. 52, no. 10, pp. 4413-4430, 2006.

[6] A. F. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, “Capacity of wireless erasure networks,” *IEEE Transactions on Information Theory*, IT-52: pp. 789–804, 2006.

[7] Y. Wu, “A trellis connectivity analysis of random linear network coding with buffering,” 2006 IEEE International Symposium on Information Theory, Seattle, July 2006.

[8] S. Yang and R. W. Yeung, “Coding for a network coded fountain,” 2011 IEEE International Symposium on Information Theory, Saint Petersburg, Russia, July 31–August 5, 2011.

[9] F. Bonomi. Connected vehicles, the internet of things, and fog computing. VANET 2011, 2011.

[10] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, “Fog Computing and Its Role in the Internet of Things,” in *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*, pp. 13–16, 2012.

[11] S. Yang and R. W. Yeung, “Batched sparse codes,” *IEEE Transaction on Information Theor*y, vol. 60, no. 9, pp. 5322- 5346, 2014.

[12] S. Yang and R. W. Yeung, *BATS Codes: Theory and Practice*, Morgan & Claypool Publishers, 2017.