An introduction to Linear Network Coding

Alberto Ravagnani

within the EBSIS project, `http://ebsis.info.uaic.ro/`

In 2000, Ahlswede, Cai, Li and Yeung discovered that the information rate of a network communication might be improved employing coding at the intermediate nodes of the network. If a source of information **S** attempts to transmit messages $v_1,...,v_k$ to certain receivers $\mathbf{R}_1,...,\mathbf{R}_N$ via a network $\mathscr{N}$, then a network coding strategy would consist in injecting the messages in the network $\mathscr{N}$, and then make the intermediate nodes cooperate to spread information faster towards the receivers.

In [1] it was shown that the multicasting technique illustrated above allows in practice to increase the number of delivered messages per channel use. The authors gave evidence of this phenomenon via the celebrated "Butterfly network", which we will discuss in detail.

Two years later, Li, Yeung and Cai showed in [5] that the maximum multicast rate of a network communication is bounded by a certain graph-theoretic invariant associated to the underlying network $\mathscr{N}$. The invariant is well-known in mathematics as the **min-cut** between vertices of $\mathscr{N}$ (viewed as a graph). The upper bound of [5] naturally leads to the following question.

Question 1. Is the min-cut upper bound achievable? In affirmative case, which are the operations that the intermediate nodes should perform in order to achieve the maximum rate?

Question 1 has a very surprising positive answer. More precisely, in [5] and [3] it was shown that the min-cut bound can always be achieved letting the intermediate nodes of the network perform appropriate *linear* operations on the received inputs, provided that the messages are vectors over a sufficiently large field. The result is known as the "Max-Flow-Min-Cut" theorem, and it essentially gave birth to linear network coding as a research field. In the mini-course we will present a partial proof for the theorem that uses an original algebraic idea based on polynomials.

Notice that the "Max-Flow-Min-Cut" theorem shows the existence of linear node operations that achieve the maximum rate, but it does not provide an effective method to concretely find them. The second question that we will investigate in our introductory course is therefore the following.

Question 2. How can one concretely design node operations that achieve the maximum multicast rate over a given network?

In [2], Médard, Kötter, Karger, Effros, Shi and Leong provided a very interesting probabilistic answer to Question 2. More in detail, they showed that the maximum multicast rate over any network $\mathscr{N}$ is achieved with high probability letting the intermediate nodes of $\mathscr{N}$ perform *random* linear operations on the received inputs.

Motivated by the results of [2], Kötter and Kschischang proposed in [4] a mathematical description of network communications in which the transmitted messages are vector spaces rather than vectors. The approach of [4] also allows to give a concise algebraic description of errors and erasures in the context of linear network coding. The model presented in [4] is based on a new class of error-correcting codes called **subspace codes**. In the last part of our mini-course I will present the main ideas behind (subspace) error correction.

The fundamental results in linear network coding have been collected in 2012 in the first chapter of a book edited by M. Médard and A. Sprintson (see [6]). The chapter was written by F. Kschischang, and it will be the main reference for the mini-course. The schedule is as follows.

1. Motivating examples: the "Butterfly network". Routing versus coding.
2. Definition of network and mincut between two vertices. Examples.
3. The min-cut bound (idea of the proof).
4. The "Max-Flow-Min-Cut" theorem (with proof).
5. Subspace codes and subspace error correction.

References

[1] R. Ahlswede, N. Cai, S.-Y.R. Li, R.W. Yeung, *Network information flow*. IEEE Transactions on Information Theory 46 (2000), 4, pp. 1204 – 1216.

[2] T. Ho, M. Médard, R. Kötter, D. R. Karger, M. Effros, J. Shi, and B. Leong, *A random linear network coding approach to multicast*. IEEE Transactions on Information Theory, 52 (2006), pp. 4413 – 4430.

[3] R. Kötter, M. Médard, *An algebraic approach to network coding*. IEEE Transactions on Information Theory, 11 (2003), 5, pp. 782 – 795.

[4] R. Kötter, F. R. Kschischang, *Coding for Errors and Erasures in Random Network Coding*. IEEE Transactions on Information Theory, 54 (2008), 8, pp. 3579 – 3591.

[5] S.-Y.R. Li, R.W. Yeung, N. Cai, *Linear network coding*. IEEE Transactions on Information Theory, 49 (2003), 2, pp. 371 – 381.

[6] M. Médard, A. Sprintson (eds.), *Coding Theory: Fundamentals and Applications*. Elsevier 2012.