Network archeology: phase transition in the recoverability of network history

Young, Jean-Gabriel (1,2) (Présentateur), L. Hébert-Dufresne (1,3,4}, E. Laurence (1,2), Charles Murphy (1,2), G. St-Onge (1,2), P. Desrosiers (1,2,5)

(1) Département de physique, de génie physique et d'optique, Université Laval, Québec QC, Canada
(2) Centre interdisciplinaire de modélisation mathématique de l'Université Laval, Québec QC, Canada
(3) Department of Computer Science, University of Vermont, Burlington VT, USA
(4) Vermont Complex Systems Center, University of Vermont, Burlington VT, USA
(5) Centre de recherche CERVO, Québec QC, Canada

Networks are simplified charts of the structure of complex systems, where the only information that matters is how the basic elements (nodes) of these systems interact together (via edges). The foundational insight of network science is that much can be learned from this structure alone.
For example, it has been observed that many real complex systems---including the North---share universal structural characteristics, and that the emergence of these characteristics can be explained by simple growth rules, i.e., rules for how networks evolve in time from their origin.

It is usually understood that these simple growth processes only provide a simplified picture of real networks rather than a detailed description. They do not---nor are they expected to---give a perfect account of reality. But, as we show in this contribution, viewing growth processes as plausible description of the actual growth of real networks can be a fruitful endeavour. In particular, this point of view leads to a natural statistical inference task: That of reconstructing the past states of a growing network, from its current structure alone.

We introduce an importance sampling algorithm that allows us to estimate the distribution of past states, conditioned on the observed structure, for many such growth processes. This provides a quasi-automatic method for "network archaeology" across disciplines: Specify a model and the method can return the history of any network, given enough sampling time.

Importantly, we show that there are insurmountable limitations to our ability to reconstruct the past of networked system, due to the simplicity of the network representation. But despite these limitations, we show that it is possible to recover a significant quantity of information from incomplete data, and therefore that network archaeology is a worthwhile pursuit---as long as imperfect knowledge is tolerable.