A study of AGN inference with Tsallis Entropy

Cassio Henrique dos Santos Amador, Fabrício Martins Lopes

Universidade Tecnológica Federal do Paraná

Abstract

In the field of gene networks, network inference is an open problem. This inference is handicapped by the low number of samples and the great network complexity (number of genes and even more interactions between them). These networks can be modeled using Probabilistic Boolean Networks, where the gene relationships are modeled as Boolean operators, and these operators can depend on one, two or more genes. Besides, the network structure can be inferred if the correct criterion function is chosen, which can be the information entropy, for instance. Previous work has shown that Tsallis entropy as a criterion function is more accurate to infer network structures than traditional Shannon entropy, if one can find the best non-extensive parameter. This parameter is related to the interaction distance between elements, the complexity of this interaction, and the possible absence of probability configurations between the elements, in this case, the genes. The present work investigates artificial gene networks (AGN) to analyze the reasons for this entropy to be more efficient, and what is the effect of a network structure topology on the non-extensive parameter used for its inference. The use of artificial gene networks allow a controlled environment, and the analysis with a larger number of samples than experimental results could allow for. Different genes with various number of links in the network, from different scale-free networks were studied. We also analyzed the relationship between these different genes and the inference of the network as a whole, and the results are compared with inferences obtained from the Shanon entropy.