Learning condition-specific networks

Sushmita Roy[1], Terran Lane[2], and Margaret Werner-Washburne[3]

***Short Abstract* — Condition-specific networks are networks of genes, where edges correspond to functional interactions that exist in a condition-specific manner. Existing approaches for inferring condition-specific networks consider data from each condition separately, and therefore ignore shared information across conditions. We have developed an approach for jointly learning condition-specific networks for two conditions, thus exploiting the shared information across the conditions. Our approach is able to identify more shared edges than an approach that learns the networks independently. Importantly, by jointly learning the networks from both conditions we are able to identify many more biologically meaningful dependencies capturing both condition-specific and shared response information.**

***Keywords* — Condition-specific networks, gene expression.**

I. Introduction

CONDITION-SPECIFIC NETWORKS describe functional interactions among genes that are induced in cells under different environmental stresses. Condition-specific response is a complex phenomenon, which is likely to involve a combination of stress-specific and general stress response factors. A systematic understanding of condition-specific response mechanism of cells requires us to identify interactions involved in both shared and condition-specific response.

Existing approaches for analyzing condition-specific behavior of networks typically identify sub-networks specific to, or general across, different conditions *after* learning separate networks per condition [1]. Simultaneous learning of specific and generic sub-networks can better exploit the shared information across conditions, and therefore construct a more complete picture of cellular stress response mechanism.

We present a novel approach for learning condition-specific networks capturing functional interactions that are specific, as well as shared across different conditions. Similar to existing approaches we learn a network (represented as undirected probabilistic graphical models, [2]) per condition. However, instead of assessing the benefit of adding an edge in individual conditions, our approach identifies the benefit of adding an edge to any subset of conditions. While adding edges to multiple conditions, we pool the data from those conditions, thus explicitly incorporating the shared information across the conditions. As a consequence, we are able to identify functional interactions that are useful not only to one condition, but to multiple conditions simultaneously.

We applied our approach to microarray expression data from two yeast, *Saccharomyces cerevisiae*, cell types, *quiescent* and *non-quiescent*, isolated from glucose-starved, stationary phase cultures. We treat each cell-type as a condition. Compared to an approach that learns independent networks per condition, our networks have more edges representing biologically meaning dependencies, including edges representing shared dependencies across the two conditions.

II. Results

We first assessed the quality of the inferred networks by identifying Gene ontology biological process terms that were significantly overrepresented in our sub-networks. Networks inferred by our pooling approach had a higher fraction of sub-networks enriched in a biological process, as compared to an approach that learned networks independently. This suggests that by pooling data we are able to not only capture shared structures, but also infer networks with higher overall quality.

We also analyzed our sub-networks for enrichment of targets of known transcription factors. We identified several sub-networks that were specific to each population and included targets of transcription factors involved in stress response. We also found that sub-networks from both cell-types were enriched in targets of known global regulators (HAP4, HAP2) of respiratory genes. The presence of these sub-networks in both cell types makes sense, as both cell populations must be able to transition from fermentation to respiration under global glucose starvation stress.

III. Conclusion

We have developed an approach for learning condition-specific networks that can identify both shared and condition-specific functional interactions. Preliminary analysis on microarray data from two conditions shows that our approach learns more biologically plausible network structures, thus effectively exploiting shared information across conditions.

Acknowledgements: This work was funded by NIMH grant 1R01MH076282-01 and NSF grant MCB-0645854.

[1]Department of Computer Science, University of New Mexico.
[2]Department of Computer Science, University of New Mexico.
[3]Department of Biology, University of New Mexico.

References

[1] Oleg, R. et al., 2007. Similarities and differences of gene expression in yeast stress conditions. Bioinformatics 23(2):e184–e190.

[2] Friedman, N. (2004). Inferring cellular networks using probabilistic graphical models. Science 303:799–805.