

Final Presentation for BiGCaT Internship

GÜNGÖR BUDAK
2013 SUMMER

Outline

- Introduction
- Methods
- R Scripts, Results & Scores
- Discussion/Conclusion/Ending notes
- Acknowledgements

Introduction

HPN-DREAM Challenge

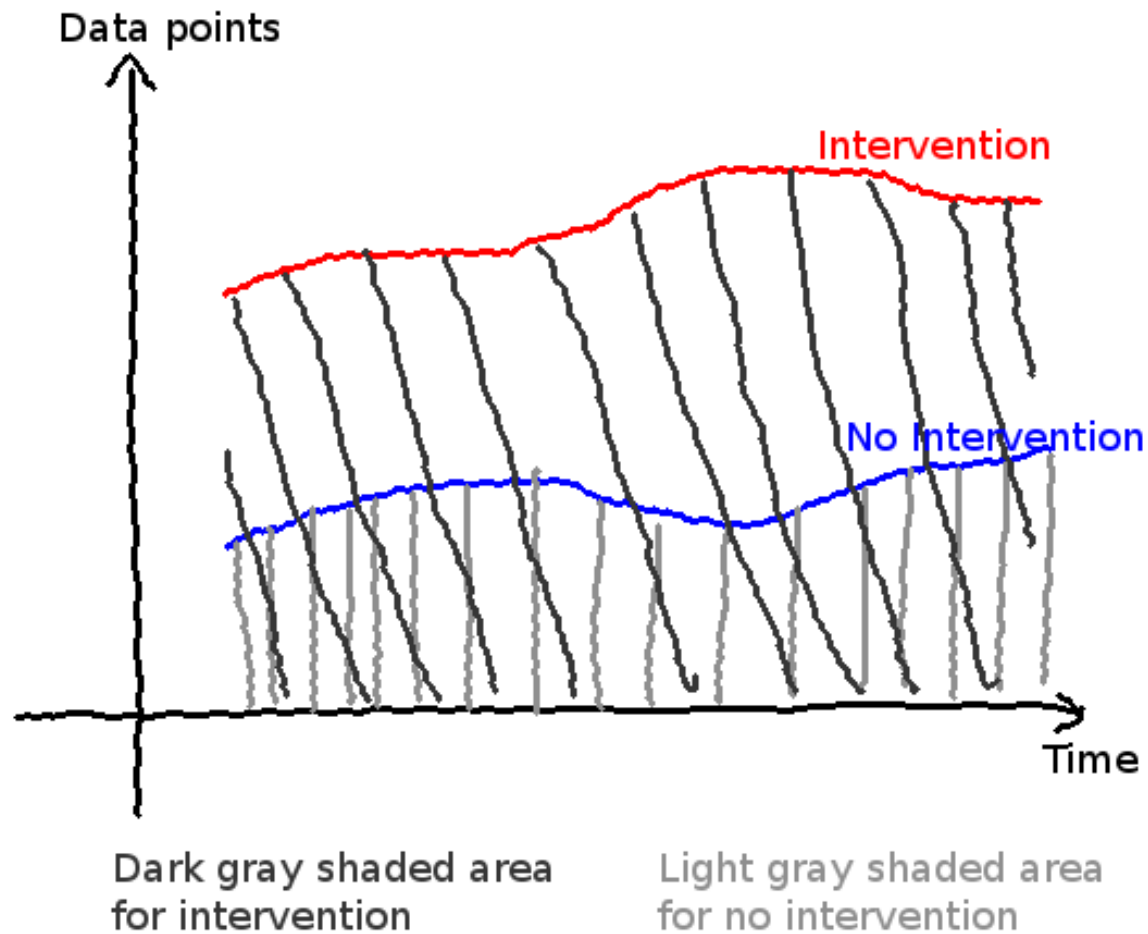
Network inference using (breast cancer)
experimental data & *in silico* data

- Finding out **directed** and **causal** edges with confidence **scores** having nodes from data

Methods

I applied the method given by the Challenge. I take time series data for intervention and no intervention cases, integrate over both series, and compare those results. If intervention is larger than no intervention, this means there is an inhibition and *vice versa*.

Methods



in silico Data Part

R Scripts for Network Inference using *in silico* Data

netInfPlotter_insilico.R – reads MIDAS file &
plots expression profiles

Parameters:

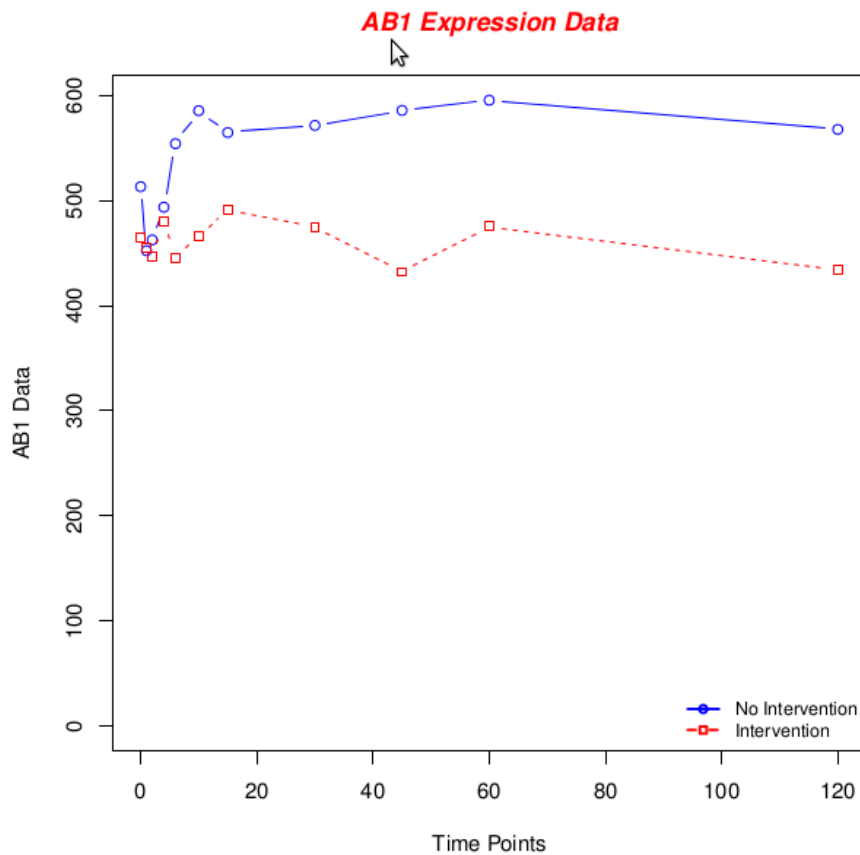
- MIDAS file
- Child antibody
- No intervention case
- Intervention case

R Scripts for Network Inference using *in silico* Data

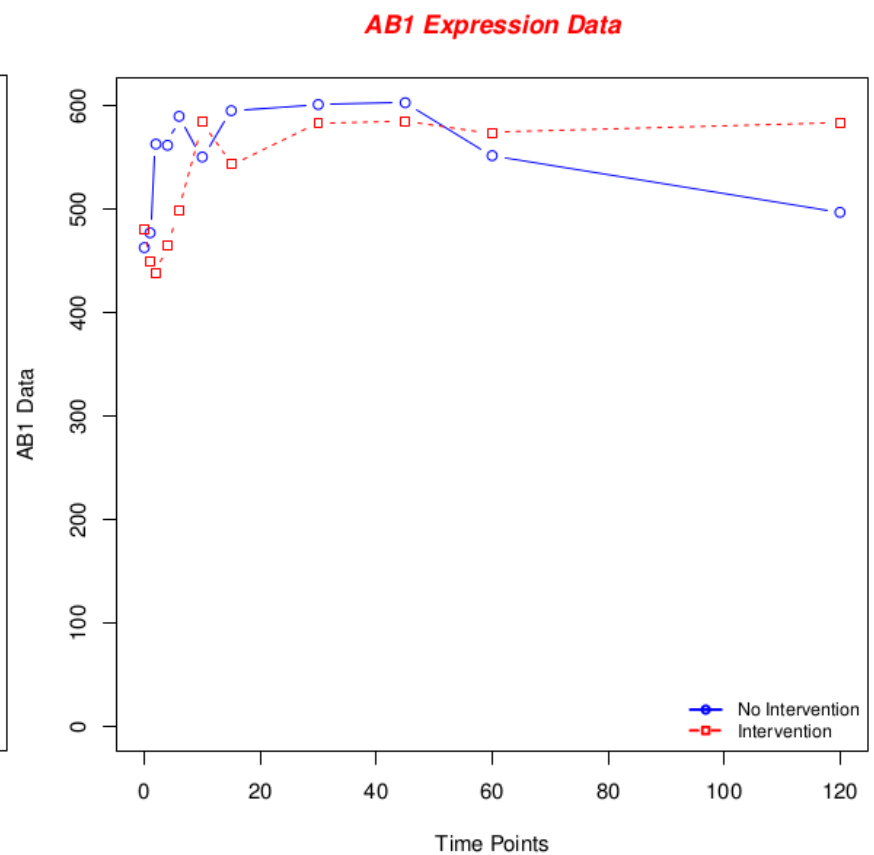
It reads **MIDAS** file and **gets treatments, time points & data points** for the child antibody given as parameter.

Next, it **plots the graph** using intervention information given as parameters.

R Scripts for Network Inference using *in silico* Data



AB1, **lo**LIG1 & INH1+**lo**LIG1



AB1, **hi**LIG1 & INH1+**hi**LIG1

R Scripts for Network Inference using *in silico* Data

netInf_insilico.R – reads MIDAS file & infers
network

Parameters:

MIDAS file

It has pre-defined relations of inhibitors, stimuli & targets. Using this definition, it infers the network.

R Scripts for Network Inference using *in silico* Data

Functions in **netInf_insilico.R**:

- **midasReader** – reads entire dataset into (and returns) a list
- **dataOptimizer** – optimizes data by organizing data according to intervention / no intervention cases and returns a list

R Scripts for Network Inference using *in silico* Data

Functions in `netInf_insilico.R`:

- **networkInferer** – gets optimized data, integrates over time series data for each case (using `simp.R` from `StreamMetabolism R` package), sends this to **edgeScrorer** and gets scores, then compares integration results and prints network and edge scores

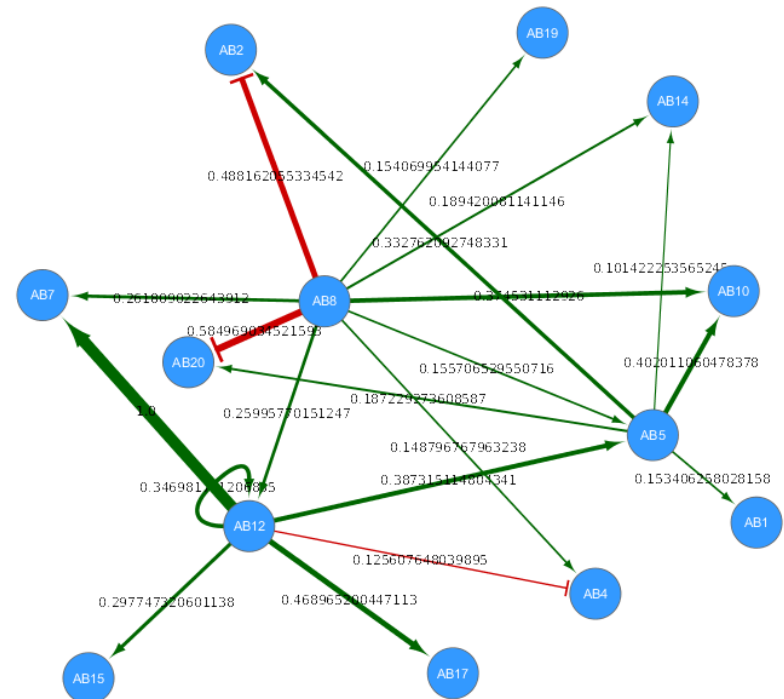
R Scripts for Network Inference using *in silico* Data

Functions in **netInf_insilico.R**:

- **edgeScorer** – gets re-optimized data, takes differences of intervention - no intervention pairs and divides each by the maximum difference (relative scoring)

R Scripts for Network Inference using *in silico* Data

NETWORK	EDGE SCORES
AB12 -1 AB4	AB12 (-1) AB4 = 0.125607648039895
AB12 1 AB5	AB12 (1) AB5 = 0.387315114804341
AB12 1 AB7	AB12 (1) AB7 = 1
AB12 1 AB12	AB12 (1) AB12 = 0.346981771206885
AB12 1 AB15	AB12 (1) AB15 = 0.297747320601138
AB12 1 AB17	AB12 (1) AB17 = 0.468965200447113
AB5 1 AB1	AB5 (1) AB1 = 0.153406258028158
AB5 1 AB2	AB5 (1) AB2 = 0.117095967751232
AB5 1 AB1	AB5 (1) AB1 = 0.107801371147316
AB5 1 AB2	AB5 (1) AB2 = 0.332762092748331
AB5 1 AB10	AB5 (1) AB10 = 0.402011060478378
AB5 1 AB14	AB5 (1) AB14 = 0.101422253565245
AB5 1 AB20	AB5 (1) AB20 = 0.187229273608587
AB8 1 AB4	AB8 (1) AB4 = 0.148796767963238
AB8 1 AB14	AB8 (1) AB14 = 0.114983286025325
AB8 -1 AB2	AB8 (-1) AB2 = 0.118139381346698
AB8 1 AB5	AB8 (1) AB5 = 0.155706529550716
AB8 1 AB7	AB8 (1) AB7 = 0.261809022643912
AB8 1 AB12	AB8 (1) AB12 = 0.25995770151247
AB8 1 AB14	AB8 (1) AB14 = 0.187371053492683
AB8 1 AB19	AB8 (1) AB19 = 0.154069954144077
AB8 -1 AB2	AB8 (-1) AB2 = 0.111852460947932
AB8 1 AB10	AB8 (1) AB10 = 0.110119528817645
AB8 -1 AB2	AB8 (-1) AB2 = 0.488162055334542
AB8 1 AB4	AB8 (1) AB4 = 0.123109029618681
AB8 1 AB10	AB8 (1) AB10 = 0.374531112926
AB8 1 AB14	AB8 (1) AB14 = 0.189420081141146
AB8 -1 AB20	AB8 (-1) AB20 = 0.584969034521593



Experimental Data Part

R Scripts for Network Inference using Experimental Data

netInfPlotter_experimental.R – reads MIDAS
file, optimizes data (this will be mentioned later)
& plots expression profiles

Parameters:

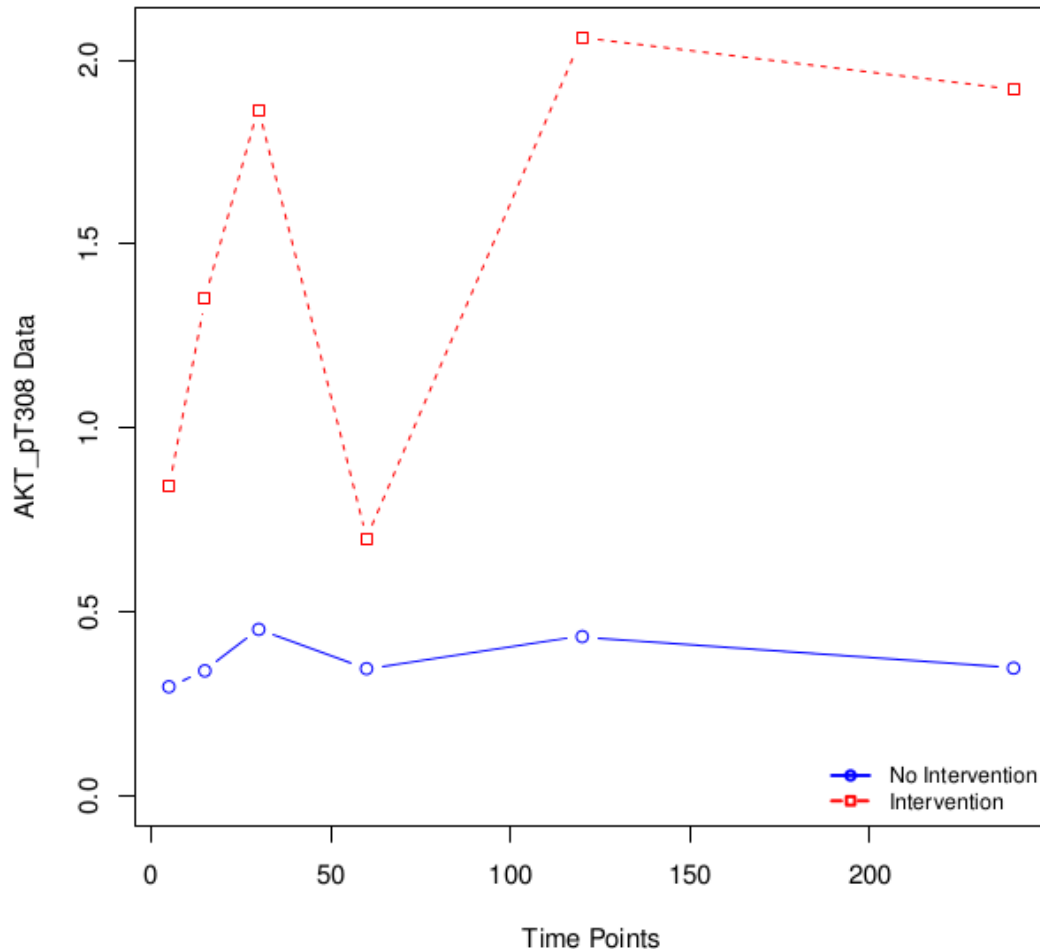
- Cell line
- Child antibody
- No intervention case
- Intervention case

R Scripts for Network Inference using Experimental Data

netInfPlotter_experimental.R – steps in the same manner as in *in silico* script. It has some functions for handling experimental data.

R Scripts for Network Inference using Experimental Data

AKT_pT308 Expression Data (Cell line: MCF7 & Stimulus: Serum)



Parameters:

- MCF7
- AKT_pT308
- Serum
- Serum__GSK690693

This result shows an **inhibition**

R Scripts for Network Inference using Experimental Data

netInf_experimental.R – reads MIDAS file into a list, **optimizes data** and **infers network**

Unlike *in silico* script, while being read, data is ordered. Duplicates are removed; missing data points are checked and NAs are put and missing data points in middle are estimated.

R Scripts for Network Inference using Experimental Data

Additional Functions in this script:

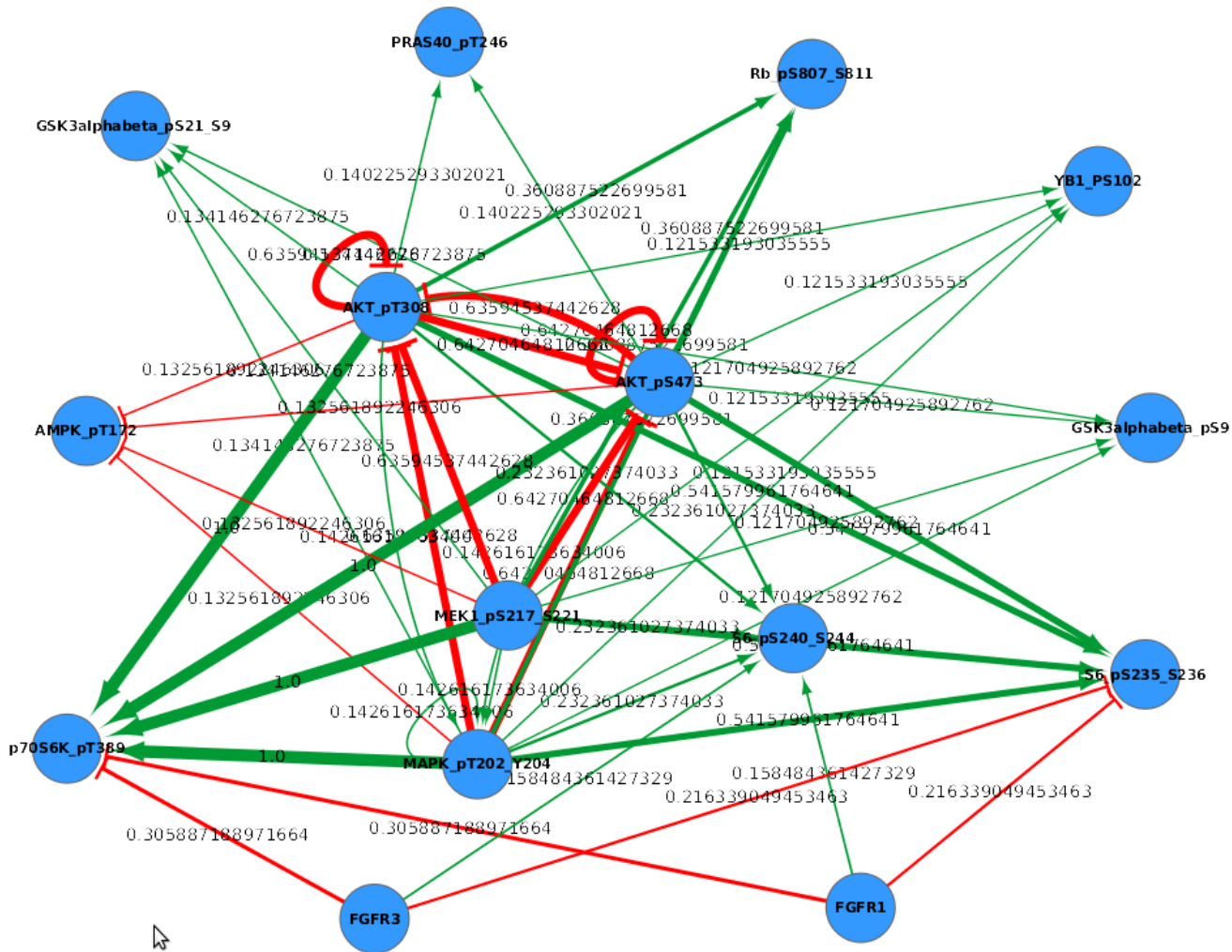
- **orderData** – gets data and orders it in a special way
- **duplicateHandler** – data points for the same condition are detected and by taking mean of them they are reduced to singles

R Scripts for Network Inference using Experimental Data

Additional Functions in this script:

- **fashionChecker** – checks if there is any missing data point and adds NAs to them
- **estimateDataPoints** – if missing data points are in the middle, they are estimated by taking mean of neighboring data points

R Scripts for Network Inference using Experimental Data



Parameters:

- MCF7
- HGF

Scores on Leaderboards

Experimental part

- Rank – **70**/83
- Z-score – -0.01

in silico part

- Rank – **49**/86
- Z-score – 1.3

Discussion/Conclusion/Ending notes

- I had to use better statistical methods for network inference
- I couldn't make use of existing tools due to nature of the data
- The challenge was a nice work to deal with to learn more on bioinformatics

Discussion/Conclusion/Ending notes

- It was nice to learn R and program in R
- It was a good experience for me to be a part of BiGCaT group and live in beautiful Maastricht

Scripts & Notes Online

Scripts in a **GitHub** repo



[gungorbudak](#) / [netinf-bigcat](#)

Notes on **Blogger**

<http://biyoenformatik.blogspot.com>

Acknowledgments

- Martina Kutmon
- Anweshha Bohler
- Susan Coort
- Chris Evelo