Parameters Estimation of Biological Regulatory Networks: An Alternate Approach Based on Computation of Network Centrality

By

Darrak Moin Quddusi

(Fall 2017-MS BI-2 00000205563)

A thesis submitted in partial fulfillment of

The requirement of the degree of

Master of Science Bioinformatics

Department of Computational Sciences

Research Center for Modeling and Simulation (RCMS)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

August 2019

Declaration

I, *Darrak Moin Quddusi* declare that this thesis titled "Parameters Estimation of Biological Regulatory Networks: An Alternate Approach Based on Computation of Network Centrality" and the work presented in it are my own and has been generated by me as a result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in the candidature for a Master of Science degree at NUST
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at NUST or any other institution, this has been clearly stated
3. Where I have consulted the published work of others, this is always clearly attributed. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work
4. I have acknowledged all main sources of help
5. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself

---

Darrak Moin Quddusi

Fall 2017-MS BI-2 00000205563

Copyright Notice

I'd like to dedicate this thesis to the two strongest pillars of my life;

my beloved parents

Acknowledgments

In the name of Allah, the Most Gracious and the Most Merciful, all praises to Allah for the strengths and His blessings in completing this thesis.

First and foremost, I would like to express my greatest gratitude to my supervisor Dr. Muhammad Tariq Saeed, who has motivated, guided and helped me a lot throughout this research work and believed in me. His invaluable help of constructive comments and suggestions throughout the research and thesis work have contributed to the success of this study.

I'd like to extend my thanks to NUST and I treasure the research environment provided by RCMS. I am thankful to my Guidance and Examination Committee (GEC) members; Dr. Ishrat Jabeen (RCMS) and Dr. Amjad Ali (ASAB) who provided me with constant assistance throughout my research phase.

I would like to express my appreciation to Principal RCMS, Head of Department of Computational Sciences, all my teachers, all the technicians, attendants, exam branch, student affairs and admin staff for their co-operations. Sincere thanks to all my friends and class fellows for their kindness and moral support during my study. Thanks for the friendship and memories. My Special thanks goes to my research group fellow Mr. Talha Junaid for the keen interest shown to complete this research work successfully.

I am extremely grateful to my parents (Mr. M. Moin Quddusi and Mrs. Kausar Moin Quddusi) for their love, prayers, care and sacrifices for educating and preparing me for

my future. I am very much thankful to my siblings Dr. Hajrah, Mrs. Maarij, Mr. Nasik and Mr. Basiq; brother-in-law Dr. James and Mr. A. Muqsit for their love, understanding, prayers and continuous support to complete this research.

I'd like to express my humble thanks to my mentor and the best teacher Dr. Jamil Ahmad for introducing me to a new world of Bioinformatics. I am grateful for his support, motivation and invaluable knowledge.

Last but not the least, my deepest gratitude goes to my very special teacher without whom my success is incomplete; Dr. Mehak Rafiq, for her endless moral support, guidance, prayers and encouragement.

To those who indirectly contributed in this project, your kindness means a lot to me. Thank you very much.

List of Figures

2.1 Time line of major contributions in the previous studies 15

3.1 Methodology work flow. Work flow starts by inputting a BRN, which is processed with the help of several libraries and in the end results are produced in the form of heat map which is then further analyzed. 17

3.2 Comparison of the model checking work flow with the proposed approach. **(a)** Work flow for model checking approach. **(b)** Generic work flow of the proposed approach. 18

3.3 Dummy tendency graph of biological regulation. **(a)** Step function of activation of B by A. **(b)** Step function of inhibition of B by A. 20

3.4 State graphs (6 out of 324) of pseudomonas aeruginosa constructed on the basis of all possible parameter combinations. 24

3.5 Pseudo code for parameters estimation of BRNs using betweenness centrality. 28

3.6 Operon model of pseudomonas aeruginosa and regulation of its mucus production. 31

3.7 Heat map of all the parameter values for state (2,1) in state graphs (along x-axis) where its betweenness centrality value is maximum 32

4.1 Pathway for Cerebral Malaria [2] **(i)** Pathway is initiated when PAMPS are recognized by TLR2/4. **(ii)** It leads to activation of BTK. **(iii)** This in turn phosphorylates MAL and activates it. **(iv)** Kinases and MyD88AP are recruited and are activated around MAL. **(v)** Followed by degradation and activation of I$\kappa\beta$ and NF$\kappa\beta$ respectively. **(vi)** INCY is produced after activation of cytokine genes. **(vii)** INCY activates its regulators. **(viii)** Meanwhile, INCY produces inflammation as well. **(ix)** Again NF$\kappa\beta$ gets activated. **(x)** NF$\kappa\beta$ induces SOCS-1 production via an alternate pathway. **(xi)** SOCS-1 inhibits MAL via polyubiquitination. **(xii)** SOCS-1 blocks expression of NF$\kappa\beta$. 36

4.2 This is the MAL associated BRN abstracted from TLR2/4 signal transduction pathway. The nodes represent the proteins involved in this pathway and the arrow heads direct the flow of pathway. Integers (-1, +1, +2) depict the threshold values of expressions which these proteins have to attain in order to have their influence on the targeting one. The positive and negative signs show the type of influence, i.e. activation and inhibition respectively [2] 37

4.3 Graphical User Interface for developed approach. 38

4.4 Heat map for estimated parameters of state 00000 sorted on the basis of low betweenness centrality. 40

4.5 Heat map for estimated parameters of state 00121 sorted on the basis of low betweenness centrality. 41

4.6 Heat map for estimated parameters of state 00000 sorted on the basis of high betweenness centrality. 44

4.7 Heat map for estimated parameters of state 00121 sorted on the basis of high betweenness centrality. 45

List of Tables

4.1 Occurrence of reported states in state graphs generated by parameter sets for 00000 sorted on the basis of low betweenness centrality. 42

4.2 Occurrence of reported states in state graphs generated by parameter sets for 00121 sorted on the basis of low betweenness centrality. 42

4.3 Information of cycles generated in state graphs due to parameter sets sorted on the basis of high betweenness centrality of state 00000 46

4.4 Information of cycles generated in state graphs due to parameter sets sorted on the basis of high betweenness centrality of state 00121 46

Contents

Acknowledgments iv

List of Figures vii

List of Tables viii

List of Abbreviations xii

Abstract 1

1 Introduction 3

1.1 Biological Regulatory Networks 4

1.2 Modeling Frameworks 4

1.3 Parameters Estimation of BRNs 5

1.4 Network Centrality and Biological Circuits 6

1.5 Problem Statement 7

1.6 Aims and Objectives 8

1.7 Formulation of Thesis 8

2 Literature Review 9

2.1 Qualitative Modeling Approach 9
2.2 Model Checking 10
2.2.1 Parameters Estimation of Biological Regulatory Networks Using Model Checking 11
2.2.2 Complexity of Model Checking 12
2.2.3 Existing Tools 12
2.3 Network Centrality 13
2.3.1 Betweenness Centrality in Selection of Trajectories 14
2.3.2 Complexity of Betweenness Centrality 14

3 Methodology **16**

3.1 Methodology Overview 16
3.2 Proposed Approach 18
3.3 Qualitative Modeling 19
3.3.1 Network Modeling 20
3.3.2 Semantics of Qualitative Modeling 21
3.4 Parameters Estimation of Biological Regulatory Networks 22
3.4.1 SIF Preparation 23
3.4.2 All Possible Parameters Combinations 23
3.4.3 State Graphs Generation 24
3.4.4 Betweenness Centrality Calculation 25
3.4.5 3.4.5. Parameters Profiling and Analysis 25
3.5 Implementation in Python 26
3.6 Software and Libraries Used 29

3.6.1 Software 29
3.6.2 Libraries 29
3.7 Example: Mucoidy in *Pseudomonas aeruginosa* 30

4 Results and Discussion **34**
4.1 Results 34
4.1.1 Case Study: MAL Associated BRN 34
4.1.2 Parameters Estimation of MAL Associated BRN 37
4.2 Discussion 47
4.2.1 Limitations 47

5 Conclusion and Future Work **48**
5.1 Conclusion 48
5.2 Future Work and Extension 49

Appendices **50**
A Supplementary Content 51

References **54**

List of Abbreviations

BRN	Biological Regulatory Network
BTK	Bruton's Tyrosine Kinase
CTL	Computation Tree Logic
CSV	Comma Separated Values
GPI	glycosylphosphatidylinositols
IDE	Integrated Development Environment
INF-γ	Interferon gamma
iNOS	induced Nitric Oxide Synthase
LTL	Linear Temporal Logic
MAL	MyD88-adapter-like
NF$\kappa\beta$	Nuclear Factor kappa-light chain enhancer of activated B cells
PAMPS	Pathogen Associated Molecular Patterns
SIF	Simple Interaction File
SOCS-1	Suppressor of Cytokine Signalling-1
TLR2/4	Toll Like Receptors 2 and 4
TNF-α	Tumor Necrosis Factor alpha

Abstract

Biological systems are complex, diverse and dynamic in nature and these features make them difficult to study. Thus, an easy way is to abstract them in the form of simple regulatory networks. Qualitative modeling approaches based on the work of René Thomas are used extensively in the domain of computational systems biology to explore the dynamics of biological regulatory networks. Modeling and analysis on the basis of qualitative modeling framework reveals several behaviors of biological systems in the form of state graphs. These behaviors are driven by certain sets of parameters which are unknown and are very crucial to understand the dynamics of biological systems. There are several approaches which are meant for parameters estimation however, one important problem in these approaches is the exponential number of model parameters. Model checking is one of these approaches based on qualitative modeling framework. It has exponential complexity which when added to complexity of parameters estimation, aggravates the situation in case of large networks; moreover, complex file management and CTL formulas required by model checking approach are difficult to write by people with no programming background, thus, in this work, a simple but scalable approach is proposed to address this challenge by extending the use of betweenness centrality with René Thomas logical formalism to the selection of suitable model parameters. It has linear complexity as compared to that of model checking and it is easy to use for everyone (with or without programming background). The developed approach is executed on reported biological regulatory networks for bench marking purpose. This work has been

validated by running the approach on a case study of Cerebral Malaria, and comparing its results with those already published in the literature.

Chapter 1

Introduction

Most of the research and efforts being made by scientists in the stream of life sciences converge at one major goal of identifying the factors responsible for dynamic behavior of the living systems, in order to develop solutions for health problems and to improve the ecosystem. Scientists from different fields are developing their techniques and using their own methods to fulfill this aim. Systems Biology having a multidisciplinary approach brings scientists from different domains like Physics, Engineering, Mathematics, Computer Sciences, and Chemistry etc. at the same page to help achieve the goal. It provides complete understanding of all biological systems as it focuses on gene expressions, molecular interactions between proteins, and external cues not only in one process but between several processes taking place inside a living organism. But this interconnection of several biological processes due to cross talks between proteins makes the biological systems complicated to study, thus making it difficult to identify the underlying factors responsible for different behaviors. Therefore, different approaches are used in Systems Biology which not only decrease the difficulty level of this study but also help to highlight important parameters (factors). Systems under study are abstracted in the form of networks; models are prepared, parameters are shortlisted and analyzed using different modeling frameworks.

1.1 Biological Regulatory Networks

There are different ways to study the biological systems but an easy way is their abstraction in the form of networks i.e. Biological Regulatory Networks (BRNs). A BRN comprises of macromolecules as its biological entities; mainly RNAs and proteins, whose interactions control the expression level of certain genes, provided in a genome. Systems Biology models biological systems, mathematically and computationally; using their respective regulatory networks. BRNs make it very easy and simple to understand the biological processes by providing a clear picture to Systems Biologists which unveils the different dynamics of the system under study.

1.2 Modeling Frameworks

In order to study BRNs, continuous, discrete and hybrid modeling frameworks have been used so far [16] .Continuous modeling is a quantitative modeling framework which refers to the use of ODEs and PDEs, but it requires precise expression values which are unknown most of the time [42]. This limitation was overcome by Boolean logic formalism where expression levels are represented with 0s and 1s depicting inhibition and activation of entities respectively [31, 32, 57]. It was further extended by Thomas to kinetic logic formalism, where expression levels are represented by discrete values instead of 0s and 1s. Both Boolean and Kinetic logic formalisms form the basis of qualitative modeling framework. Discrete formalism jumps in when enough information isn't available about parameters and conditions [16]. Hybrid formalism incorporates both continuous and discrete changes in a system [1]. As BRNs are meant to model the interactions, mostly between proteins and genes; a graph can help to represent the static part of such model. Whereas, vertices and edges of graph represent the biological entities and interactions respectively. Negative sign - shows inhibition and positive sign + shows activation. Changes in numerical values assigned to the corresponding entities

with respect to time; constitutes the dynamic part of model, which relies on temporal evolution of concentration of these values. Initially, this dynamic behavior was studied using differential equation approaches, but later on Thomas came up with kinetic logic approach in order to study the qualitative nature of the dynamic behavior of regulatory networks [58].

1.3 Parameters Estimation of BRNs

Models are abstractions of real systems, and the values which are responsible in determining how close a model is to a real system are known as parameters. Estimation of parameters is a challenging task because it is difficult to get the right ones from the huge set of parameter values. Likewise, parameters estimation of biological regulatory networks to study the dynamical changes in biological systems is as challenging as it is important. Parameters estimation using qualitative modeling framework is preferred over quantitative modeling as expression levels are represented using discrete values, which decreases the difficulty for parameters computation job keeping the parameter state space finite; whereas, in the case of quantitative modeling where expression levels have continuous values, the difficulty for parameters computation is high as parameter state space is infinite. Moreover, qualitative modeling framework unveils important properties of Biological systems i.e., bifurcation points, cyclic behavior in the form of feedback loops [58] and stable or steady states [56] etc. Parameters estimation is not only helpful in the analysis of systems level diseases like Parkinson's disease, Alzheimer's disease, and cancer etc., but also advantageous in identification of potential therapeutic drug targets [54].

1.4 Network Centrality and Biological Circuits

Network centrality aims at ranking network elements. It also helps to identify interesting facts about these elements and brings the key players in a certain pathway to notice. Centrality refers to a function C which is meant to assign a numerical value to each vertex v of the graph i.e. C(v). In order to rank the vertices, Koschutzki et al [36] chose a convention where a vertex v carries more importance than w, iff C(v)> C(w). Likewise, network centrality measures when applied on biological networks, extract important information. Jeong et al [27] has explained that a protein which is highly connected in a particular protein-protein interaction network is functionally very important most of the time. Removal of such proteins (i.e. vertices) leads to lethality in the whole network.

Network centrality analysis is done using different centrality measures depending upon the kind of information required. There are several centrality measures but some major and related measures constitute degree centrality, closeness centrality, harmonic centrality, betweenness centrality, and cross-clique centrality.

Degree centrality of a node is the total sum of its edges. If the edges are outgoing, then it is out-degree centrality and if the edges are ingoing then it is in-degree centrality [36]. Degree centrality of a node u is:

$$D(u) = deg(u) \tag{1.4.1}$$

Closeness centrality of a node is the average of the sum of shortest paths from that node to all the other nodes in a network. It is only applicable to strongly connected networks. This measure was defined by [8] as reciprocal of farness, i.e. closeness centrality of v is:

$$C(v) = \frac{1}{\Sigma_{w \in V} dis(v, w)} \tag{1.4.2}$$

where $dis(v, w)$ is the shortest distance between two nodes v and w.

Harmonic centrality just reverses the reciprocal and sum operations in closeness centrality. Moreover, it isn't just limited to strongly connected networks only. Harmonic centrality of a node v can be found as:

$$H(v) = \Sigma \frac{1}{dis(v, w)} \tag{1.4.3}$$

where v $\neq$ w; and $\frac{1}{dis(v,w)} = 0$ if v and w doesn't connect with each other [41].

Betweenness centrality is a measure of node within a network. It accounts for the number of times a certain node has served as a bridge between two nodes on their shortest paths. According to Freeman et al. [22] nodes exhibiting highest probability of occurrence on a shortest path chosen randomly between two nodes depict high betweenness. Betweenness of a node v is defined as:

$$B(v) = \Sigma_{xvy}(\frac{\sigma_{xy}(v)}{\sigma_{xy}}) \tag{1.4.4}$$

where x $\neq$ v $\neq$ y; and σ_{xy} is the total number of shortest paths between x and y; while $\sigma_{xy}(v)$ represents the number of paths which pass through v.

Cross-clique centrality of a given node determines its connectivity to different cliques in a complex graph. Clique is a sub-graph where each node is connected to all the other nodes in that sub-graph. A node having high cross-clique centrality eases the dissemination of information in the graph. For a node u, it is defined as X(u), which gives the number of cliques constituting u [19].

1.5 Problem Statement

The estimation of logical parameters for modeling biological networks is a computationally intensive task. The formal approaches that utilize Model Checking to compute parameters use exhaustive state space search that leads to State Space Explosion. More-

over, use of CTL makes it limited to be used by Computer scientists only. Therefore, a potential solution to cater this issue is by extending the concept of betweenness centrality with René Thomas logical formalism, which will not only reduce state space but will also decrease processing time and providing the users with an easy to use approach.

1.6 Aims and Objectives

This study aims at developing a new and easy approach by extending the concept of betweenness centrality with René Thomas logical formalism to estimate parameters for BRNs with a reduced computation cost. Following are the objectives of the study;

- Study existing approach for parameters estimation with respect to Qualitative Modeling
- Develop a more user friendly and easy to use approach
- Benchmark developed algorithm using data sets of existing regulatory networks
- Validate the developed approach

1.7 Formulation of Thesis

Thesis is divided in three major parts. First one is the literature review of the pioneering work in this domain which focuses on gradual evolution of modeling stream and how the concept of parameters estimation originated. The second one elaborates the methodology flow used for the development of this approach, and the last one focuses on the results and discussions regarding the future prospective of developed approach. In the end supplementary files and bibliography formally close the thesis.

Chapter 2

Literature Review

Biological system as a whole is difficult to study; therefore, an easy way is to abstract them in the form of networks, where nodes of the network represent the biological entities and edges are responsible for communication between nodes. As biological systems are non-linear systems thus, a good approach to study their dynamics is qualitative modeling framework.

2.1 Qualitative Modeling Approach

Kauffman [30] proposed Boolean logic for regulatory networks; according to this logic, activations are represented with 1 and inhibitions are represented with 0. In order to add more meanings to the analysis based on this approach Thomas et al. [61] proposed kinetic logic formalism which represents the expression levels with multiple discrete values; rather than just 0 and 1.

This makes qualitative study of regulatory networks more dynamic and versatile; to make it more detailed Thomas [59] introduced the concept of feedback circuits and highlighted their importance in generating sustained oscillatory and multiple states behavior. Feedback circuits have two types; degenerative or negative circuits and regenerative or positive circuits. The degenerative circuit is responsible for generating sustained

oscillatory behavior and the regenerative one generates multiple steady state behavior.

Feedback circuits alone cannot give a proper and better insight to the dynamics of systems, thus, Snoussi [56] introduced the idea of logical parameters also known as model parameters. This notion of feedback loops leads to the concept of singular logical states and loop-specific state [60]. It helped in providing further insights to the study of regulatory systems by formally demonstrating that only the states involved in positive feedback loops lead towards steady states or fix points [56].

2.2 Model Checking

Model checking is one of the formal verification techniques, which was designed with the aim of checking the desired specifications of hardware circuits or computer programs. It permits the modeler to test all the possible upshots of the system kept under study. This is done in an exhaustive manner. Thus, conclusions made on the predictions coming from model checking are definite. This approach is used in different domains; hence, it is considered as a standard method in industrial sector to ensure correctness in complex software, hardware and embedded systems, security protocols, digital circuits and what not. It evolved gradually with the contribution of several scholars making it more reliable day by day. Devloo [17] used constraint programming in order to detect all the steady states of large regulatory networks. It is a form of declarative programming where relation between entities exists in the form of constraints. This approach was used to identify all the stable states in a system, where stable states are derived as a solution to the system of stable state equations. In the same year Peres et al. [47] applied model checking in Biological Regulatory Networks to understand the system dynamics, which is explained in article 2.2.1. Dynamics and complexity of biological systems mark them similar to parallelism in software systems distinguished by non-deterministic behavior [21] . Due to this similar behavior shown by concurrent systems as well; model checking

is used for the analysis of huge state space of all possible outcomes of the biological model under study. For this purpose the required behavior and the details of the model M are given in the form of a formula ϕ and transition system respectively. Model checker exhaustively explores the state space to check the correctness of formula ϕ.

2.2.1 Parameters Estimation of Biological Regulatory Networks Using Model Checking

For the analysis of BRNs, model checking is used by numerous tools. In this approach, information of the system under study is expressed as a transition system which has all the possible states with their transitions. Behavior (to be verified) of the system is expressed in the form of a temporal logic formula. Temporal logic formula is comprised of temporal logic quantifiers which depict the system's behavior. Model checker takes the system model and temporal logic formula as input and then via brute force technique it starts exploring the complete state space of entered model exhaustively. If the property is satisfied, the model is validated and it gives the set of parameters and if it is not satisfied, then it generates a counter example and whole process is repeated.

According to René Thomas logical formalism, state graphs represent the dynamic behavior of biological systems. The factors (parameters) responsible for these behaviors are unknown and their estimation is the most important step in the qualitative modeling of BRNs [11]. Bernot used model checking for parameters estimation of Biological Regulatory Networks. It is differentiated on how the time is interpreted; whether linear (Linear Temporal Logic) [50] or branching (Computation Tree Logic) [15]. As Biological systems are dynamic and non-deterministic in nature; thus, CTL is suitable for BRNs because it is branching time logic. The already known experimental observations are converted to the formula ϕ, which is actually a CTL. Model checker evaluates several parameter combinations and ultimately selects only those parameters against which CTL observations are satisfied.

State graph of a Boolean network having n nodes has 2^n numbers of states; each of which have a maximum of n outgoing transitions. Asynchronous state graphs may or may not have $n2^n$ possible transitions, thus, the total number of state graphs (asynchronous) corresponding to Boolean networks with n number of nodes turn out to be 2^{n2^n}. If an interaction graph having n vertices along with a certain property p is input while estimating parameters via model checking, then the output comprises of all those asynchronous state graphs which satisfy property p. Therefore, complexity for parameters estimation is $O(2^{n2^n})$ [13].

2.2.2 Complexity of Model Checking

In biological systems, regulatory functions are performed when biological entities i.e. genes or proteins interact with each other. These interactions make a biological network simple or complex or highly complex. The complexity of model checking approach is $O(2^n)$ and this complexity increases when the number of interacting entities increases followed by an increase in computational cost required for the analysis of such biological networks. This cost is large even for small and simple networks as it requires model building for huge set of parameters, and evaluation of each model; to meet the required properties. Model checking looks for all of this in the complete state space of the mode. The length of the state space for simple to complex and from complex to highly complex network starts from hundreds and thousands to trillions and so on. Therefore, adding difficulties to the approach in estimating parameters.

2.2.3 Existing Tools

Several tools have been developed in this domain on the basis of model checking approach. Four major tools are GINSim [23], SMBioNet [33], GenoTech and Parallel SMBioNet [54]. Usually CTL formulas are written corresponding to the properties of system under study and models are generated with respect to those formulas and param-

eters are computed. GINSim is Gene Interaction Network Simulator which aims at BRN construction and supports its export in different file formats to establish compatibility with other tools like NuSMV, snoopy etc. SMBioNet uses NuSMV to estimate logical parameters of Biological models which satisfy known observations [29]. It is applicable to small BRNs only. GenoTech helps in BRN construction, specifying logical parameters and generation of state graphs and their analysis. Parallel SMBioNet is the extension and parallel implementation of SMBioNet, which helps in parameters estimation of bigger BRNs. Moreover, Paulevé et al. [46] developed a π-Calculus framework known as Process Hitting. It is a stochastic framework which highlights the set of dynamics of a BRN which are highly functional. Sheikh et al. [55] incorporated time delays in this framework to model large BRNs dynamically.

2.3 Network Centrality

Network centrality is a local measure in order to determine the position of a certain node in the network with respect to the other nodes. It is a quantitative measure that helps in estimating the role of the node in the network. The concept of network centrality comes from social network analysis when Bavelas applied this idea on human communication to explore the relationship between their behavior and patterns of communication in small groups. It was concluded that the behavior is affected by communication patterns in which the groups operate; and network centrality is pertinent to this [38]. Cohan and applied network centrality to study how such large and heterogeneous Indian nation is administered; and it came out that these are the network centers that knit all the attributes of Indian social life into a refined, intertwined and well-coordinated structure. This approach further got boom when it was applied on communication paths by Pitts [49] for urban development. From social networks to urban development and establishing inter-organizational relations idea of network centrality is used in every manner [9]. Apart from this, the concept of network centrality has been extended to study the

computer networks, electrical circuits, and the regulatory networks of biological systems etc. to extract desired information. Thus, different centrality measures exist depending upon the type of the role that is expected from the node [3].

Network centrality plays an eminent role in the study of regulatory networks of biological systems. On the basis of graph spectral properties, centrality measures help in distinguishing important proteins in Protein Protein Interaction (PPI) networks. Moreover, in order to select potential drug targets; proteins if ranked as per their centrality measure can be of great help [18].

2.3.1 Betweenness Centrality in Selection of Trajectories

According to the directed graph theory, a path is made when distinct nodes are connected by directed edges in the form of a chain. These paths in the study of biological networks are transformation pathways from one entity to a certain other entity in the network [5]. Redundancy of such pathways in biological networks, accounts for their robustness. Betweenness centrality helps in measuring the effect caused by node perturbation on the pathway redundancy. The length of the paths determines the response time of perturbations [45, 51].

2.3.2 Complexity of Betweenness Centrality

Different algorithms are used for betweenness centrality calculations and they usually have a time and space complexity of $O(n^3)$ and $O(n^2)$ respectively. The algorithm used in Python language for computation of betweenness centrality is Brandes algorithm. It is named after Ulrik Brandes who proposed it. It has a space complexity of $O(n+m)$ and time complexity for unweighted graphs is $O(nm)$ and for weighted graphs is $O(nm+n^2 logn)$ [12].

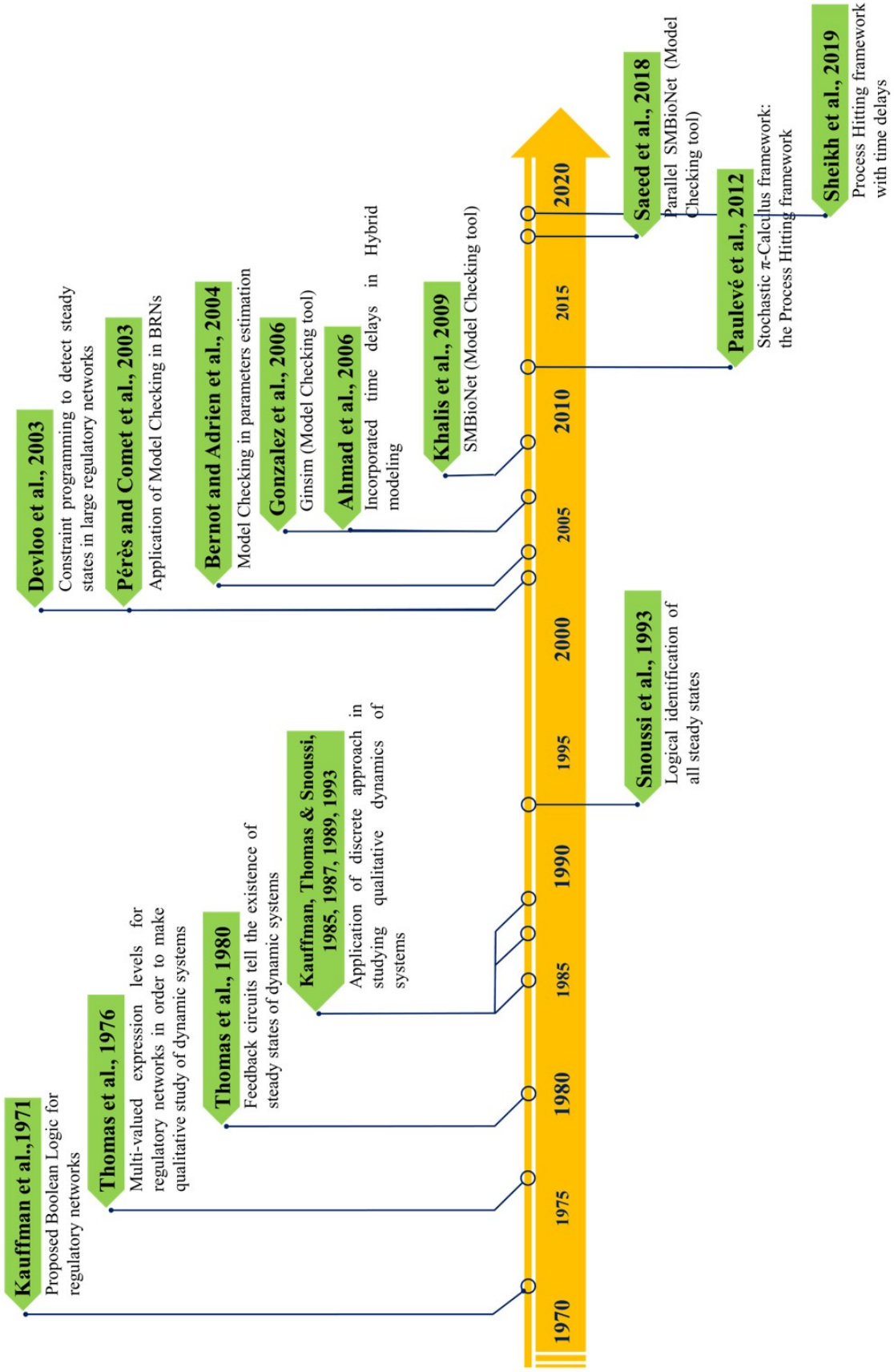


Figure 2.1: Time line of major contributions in the previous studies

Chapter 3

Methodology

3.1 Methodology Overview

This chapter focuses on the methodology used to develop the approach. Python language with its several libraries is used in the development process where BRN abstracted from the actual pathway is written in a SIF format. From that information parameters are computed and are used for the state graph generation. Betweenness centrality calculations are performed on the nodes of all state graphs. Centrality values are sorted and are used to trace back the state graphs and their parameters to generate results in the form of heat maps which are further analyzed. Figure 3.1 shows the methodology work flow of the approach. Every step is discussed one by one in detail.

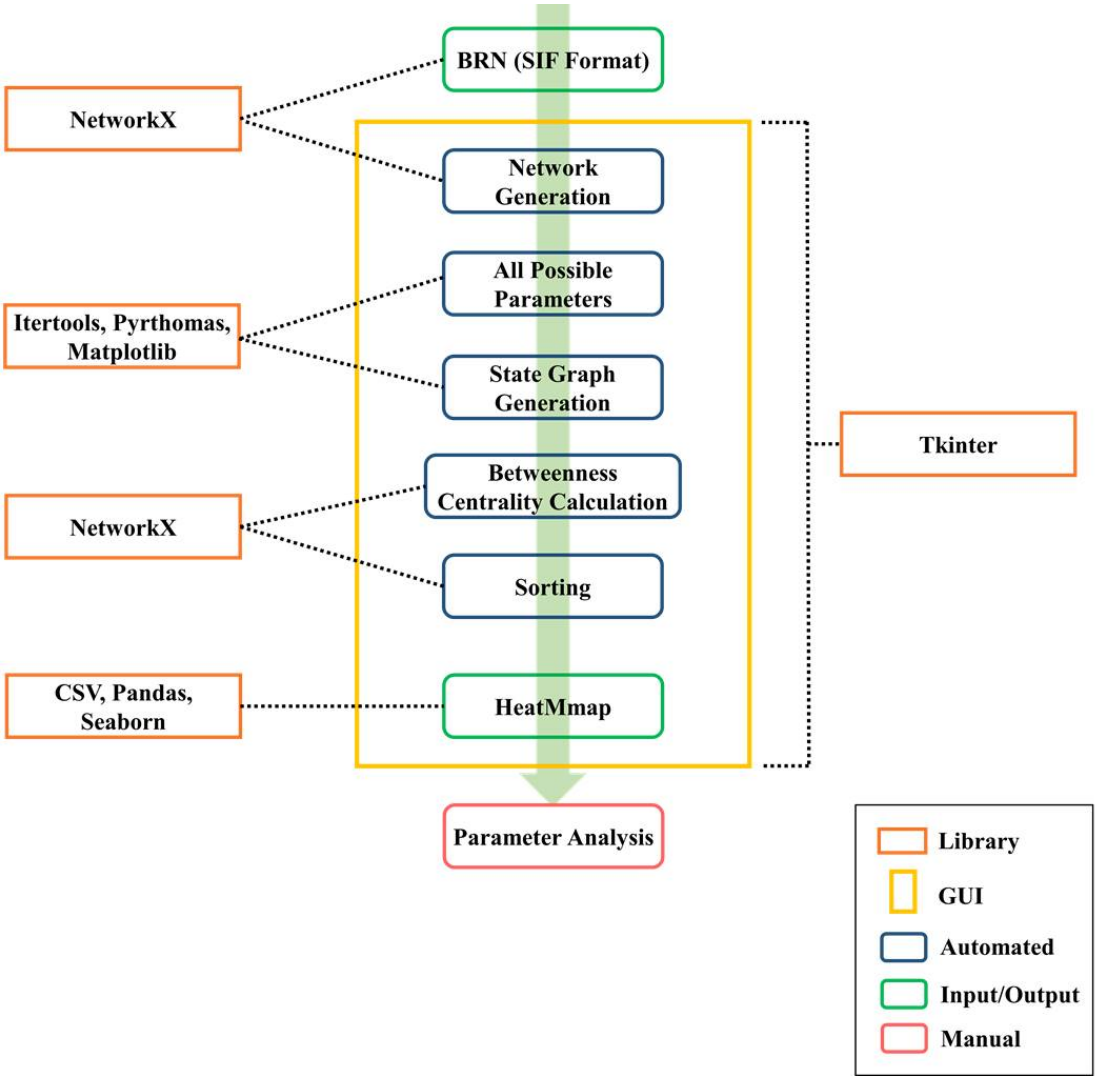


Figure 3.1: Methodology work flow. Work flow starts by inputting a BRN, which is processed with the help of several libraries and in the end results are produced in the form of heat map which is then further analyzed.

3.2 Proposed Approach

In this work, a new and easy to use approach is developed by extending the concept of betweenness centrality on qualitative modeling framework proposed by René Thomas to estimate parameters for BRNs with a reduced computation cost. It is free from all kinds of difficult file inputs and CTL formulas, which makes it easy to be used by everyone without having any programming background.

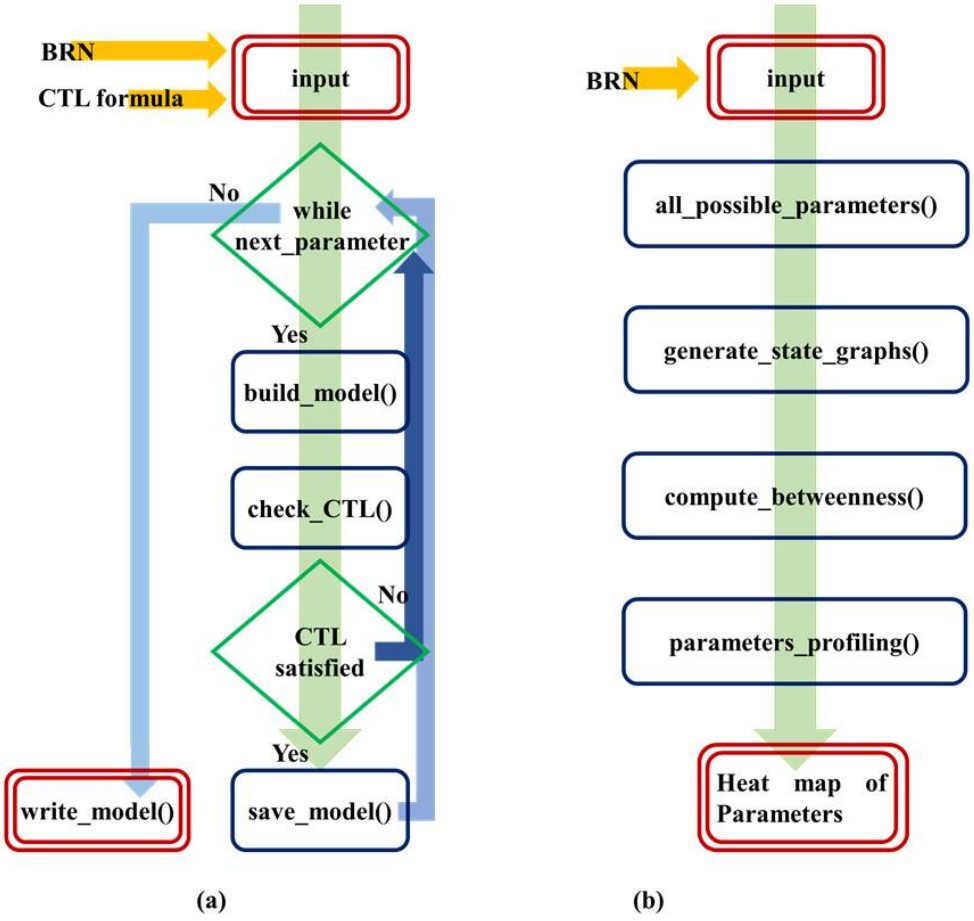


Figure 3.2: Comparison of the model checking work flow with the proposed approach. **(a)** Work flow for model checking approach. **(b)** Generic work flow of the proposed approach.

3.3 Qualitative Modeling

Qualitative modeling highlights the qualitative properties of a dynamic system by modeling its behaviors. These behaviors are driven by model parameters also known as logical parameters. In a biological system, biological entities are influencing each other; and most of the time it is impossible to measure this influence. Thus, finding the set of logical parameters which are suitable enough to describe a certain regulatory behavior occurring strictly in an intracellular process is pivotal in qualitative modeling of these biological systems. It helps in modeling the diseased systems in a way that these parameters point towards the responsible factors in causality of the disease; and once such factors are determined it is easy to predict potential drug targets via several computational models.

Qualitative modeling depicts the chemical concentrations of biological entities in BRN with discrete values. As already mentioned these biological entities influence each other; this influence can be positive or negative i.e. activation or inhibition respectively. This influence is described by change in their expression level by one discrete level. Positive influence (activation) means that the expression level of source entity causes an increase in the production or rate of activation of sink (target) entity. Contrary to this negative influence (inhibition) occurs when expression level of source entity causes the degradation of sink (target) entity.

These regulations can be represented using step functions; if entity A positively influences entity B, then A is known as the activator of B. A activates B only when its expression level reaches a certain threshold value. Whereas, if A negatively influences B; then A is known as its inhibitor. Expression level of B starts decreasing when expression level of A hits a certain threshold value. These changes in expression levels of both activator and inhibitor are shown in figure 3.3 [7].

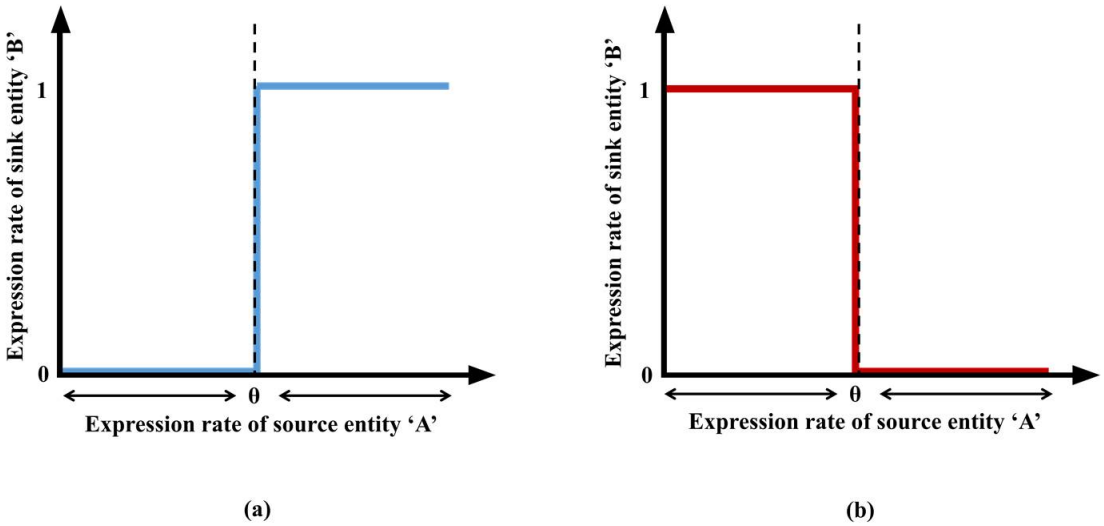


Figure 3.3: Dummy tendency graph of biological regulation. **(a)** Step function of activation of B by A. **(b)** Step function of inhibition of B by A.

3.3.1 Network Modeling

Graph Theory plays a pivotal role in modeling and analysis of biological regulatory systems as it represents them in the form of graphs (networks) to model their behaviors. For this purpose a directed graph is very useful and easy to understand data structure. A BRN is mapped onto a graph in a way that its biological entities make the vertices of graph and its interactions (positive and negative) make the edges of the graph, which are directed.

Graph based approaches are used to determine topological and structural parameters to discern important properties of biological networks. Structural properties of a graph appear very helpful in making biological predictions. Thus, a lot of operations are performed on graphs for this purpose; like degree and centrality measurement. These bring out important information about nodes a.k.a. biological entities. Betweenness centrality is the back of this study to characterize the nodes of different biological models in order to understand the dynamics of biological systems to which these models belong.

Given below are the formal definitions and properties of graphs taken from [28].

Definition 1 (Graph). *Graph G is an ordered pair i.e. G = (V, E) where, V represents the set of vertices i.e. biological entities and E represents the set of edges which marks the interactions between vertices.*

Definition 2 (Degree). *Degree is the total sum of edges a vertex v has. If the edges are originating from vertex v then it is known as out degree and if the edges are terminating towards vertex v then it is known as out degree.*

Definition 3 (Bipartite Graph). *Graph G = (V, E) is said to be bipartite iff:* $v = (A \cup B) \mid A \cap B = \phi$ *and* $E \subset (AXB) \cup (BXA)$.

3.3.2 Semantics of Qualitative Modeling

This section constitutes semantics of René Thomas qualitative modeling framework taken from [2, 10].

Definition 4 (Directed Graph). *A directed graph D is an ordered pair D = (N, I); where N represents finite set of all nodes or vertices and I represents set of ordered pair of nodes i.e.,* $I \subseteq N\ X\ N$. *Nodes are connected to each other via arcs, such that arc a = (x, y) connects x to y; where x is the head (source node) and is known as the tail (sink node).*

Definition 5 (Biological Regulatory Network). *A Biological Regulatory Network (BRN) is a labeled directed graph D = (N, I); where set of nodes N model the biological entities (genes, proteins etc.) and set of I subset N X N models interactions. Each interaction* (n_i, n_j) *is labeled by a pair* (τ, σ), *where* τ *depicts the threshold value x attains in order to regulate y; and* σ = {+, -} *represents the sign of interaction where + shows activation and - shows inhibition.*

Definition 6 (State). *A qualitative state of BRN is n-tuple* $S = \{s_{n_1} \ \ldots \ s_{n_j}\}$, $\forall s_{n_i} \in E_{n_i}$ *where* s_{n_i} *is the abstract expression level of* n_i.

The dynamics of BRN are dependant on model parameters (set of positive integers)

and resources. Different dynamics are shown by same BRN when model parameters are changed.

Definition 7 (Resources). *In D, availability or absence of an activator or inhibitor respectively for each biological entity is known as a resource; and resource set is Cartesian product of these activators and inhibitors.*

Definition 8 (Logical Parameters). *These are the set of parameters responsible for discrete evolution of biological entities in the BRN. Let D = (N,I,π) be the BRN and Para(D) the set of parameters of D,*

$\mathrm{Para(D)} = \{\mathrm{K}_{x,A,B} | \mathrm{A} \subseteq \mathrm{D}_{+}(\mathrm{x}),\ \mathrm{B} \subseteq \mathrm{D}_{-}(\mathrm{x})\}$

Valuation of Para(D) is a mapping κ *: Para(D)* $\implies$ *such that* $\forall\ x \in N$, $A \subseteq D_{+}(x)$, $B \subseteq D_{-}(x)$, $\kappa(K_{x,A,B}) \in S_x$.

Definition 9 (State Graph). *Let D = (N, I) denotes a BRN and* s_{n_i} *represents the expression level of* n_i *in a state* $s \in S$. *Then its state graph G = (S, T) is a directed graph, where S is the set of states and* $T \subseteq S\ X\ S$ *is the transition relation; such that* $s \rightarrow \acute{s} \in T$ *iff:* $\exists$ *unique* $p \in N$ *such that* $s_{n_p} \neq \acute{s}_{n_p}$ *and* $\acute{s}_{n_p} = s_{n_p}\ \Delta\ K_p\ (W_{n_p})$, *and* $\forall q \in V \setminus \{x\}\ \acute{s}_{n_q} = s_{n_q}$

3.4 Parameters Estimation of Biological Regulatory Networks

As per René Thomas Logical Formalism, dynamics of a biological system are studied by converting its regulatory network into a state graph on the basis of logical parameters. These logical parameters are unknown and a huge challenge for Qualitative Modeling is their estimation. Thus, a new approach has been introduced in this study which somehow eases the way for Biologists to estimate parameters without getting caught into CTL formula and complex file preparations. This approach constitutes of following

steps:

1. SIF Preparation
2. All Possible Parameters Combinations
3. State Graphs Generation
4. Betweenness Centrality Calculation
5. Parameters Profiling and Analysis

3.4.1 SIF Preparation

Different file formats exist for storing BRN information like GraphML, XML, and dot etc. but this study requires a SIF format for storing BRN information. SIF stands for Simple Interaction File; it is a file format which specifies nodes and their interactions. Unlike other file formats SIF file is easy to prepare in notepad. It is a tab separated file where each line corresponds to an interaction which takes three tab separated things i.e. source node, threshold value with + or - sign and sink node.

3.4.2 All Possible Parameters Combinations

In order to compute all the possible parameter combinations of the BRN, resources of each entity (node) are found first. This is done by using Definition 4. On the basis of threshold values i.e. expression level values, ranges (from minimum to maximum expression level value) are assigned to each resource of each entity; this leads to the computation of sets of all possible parameter combinations of BRN by assigning different values from the deputed ranges of each resource.

3.4.3 State Graphs Generation

State Graphs are generated corresponding to all sets of these possible parameter combinations of the entire BRN. State graphs represent different dynamics of the system under study. These parameter combinations are actually different factors that are playing their role in driving the system and state graphs generated on their basis depict the different behaviors that arise when these driving factors are changed. Out of all of these sets of possible parameter combinations, there are two types of sets one if followed keeps the system in healthy condition and the other if followed diverts the system from healthy condition to the pathogenic one.

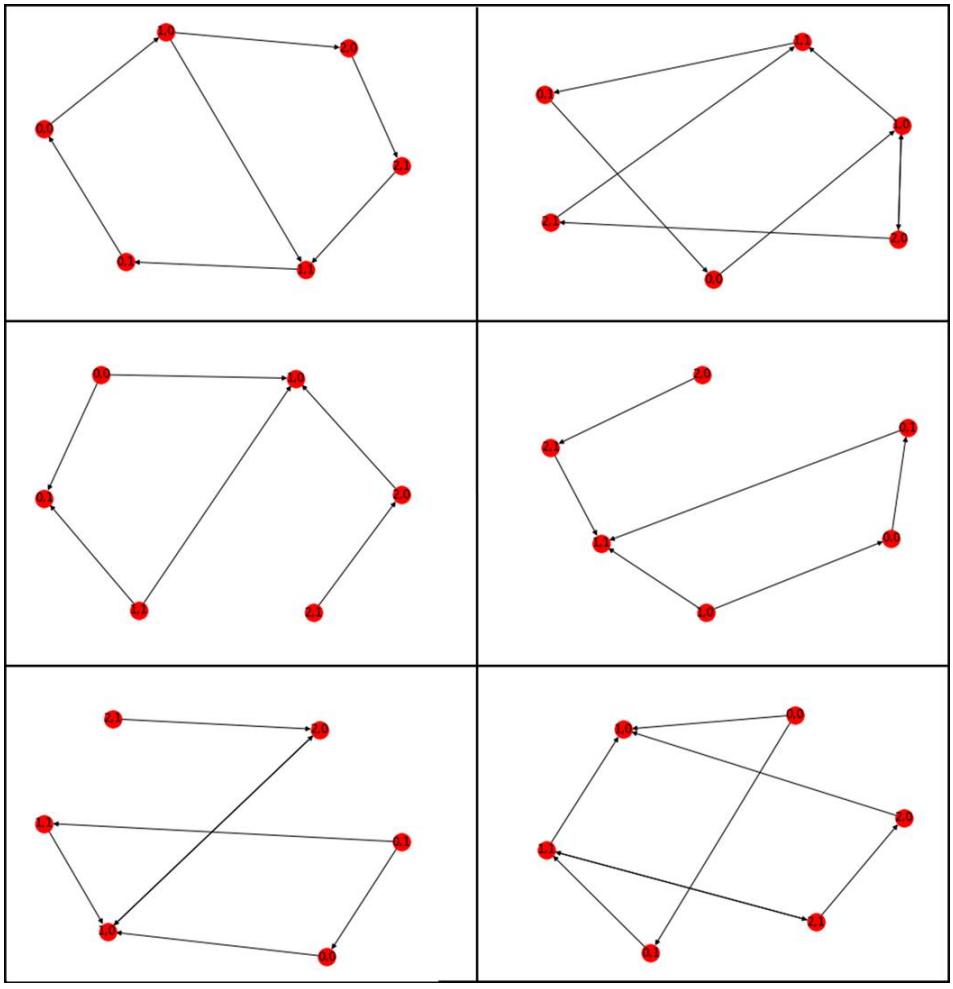


Figure 3.4: State graphs (6 out of 324) of pseudomonas aeruginosa constructed on the basis of all possible parameter combinations.

3.4.4 Betweenness Centrality Calculation

Definition 10 (Betweenness Centrality). *Let D = (N, I) denotes a state graph having a, b and c as three disparate states in D. Let P represents the set of all ordered pairs of all distinct states and Îija,b represents the total number of shortest paths from a to b; also, $\theta_{a,b}$ (c) be the total number of shortest paths from a to b passing through c. Then, the betweenness centrality B for state c is computed from the following:*

$$B(c) = \Sigma_{ab \in P}(\frac{\theta_{ab}(c)}{\theta_{ab}}) \tag{3.4.1}$$

The nub of this study is to get a bigger picture of all such factors and to categorize them on the basis of betweenness centrality. This is done by calculating the betweenness centrality of each node of each state graph. Each state can be analyzed individually, as the parameter sets of desired state in all the state graphs where its betweenness centrality value is highest or lowest are grouped together respectively. State graphs of states showing highest or lowest betweenness centrality are traced back to sets of parameter combinations which are making them, and then the parameters are displayed in the form a heat map.

3.4.5 3.4.5. Parameters Profiling and Analysis

Parameters profiling for the given qualitative BRN is done on the basis of normal and pathogenic states of corresponding state graph of the system, which user should know. It will yield the respective parameters sets, which if followed (for good states) keep the system in healthy state; and, the other if followed (for bad states), diverts the system to pathogenesis or diseased state. Analysis is done on the basis of the biological significance of state of interest.

3.5 Implementation in Python

Whole approach has been developed in Python version 3.7.0. Python was developed by Guide Van Rossum in 1991. It is a multi-purpose, high level, interpreted, procedural, functional and object oriented programming language. It is enriched with built in packages/libraries that are easily installed using "pip install" command. Thus, it is highly preferred by most of the organizations and institutions due to its various programming paradigms. Apart from this it also serves with automatic memory management [37].

In order to endue this approach with life, several python libraries are used which are explained in detail in section 3.6. While implementing the approach, only SIF format is considered as input for the retrieval of BRN information (Figure 3.5 line: 8). All the necessary information (i.e. source nodes, resource sets, threshold values) of each node is processed and a BRN is generated upon reading the respective file. Then all the possible parameter combinations of that particular BRN are computed on the basis of which all state graphs are generated (Figure 3.5 line: 9), this is achieved by using Pyrthomas library. A built in function supported by NetworkX is used for calculating betweenness centrality of each node (state) of each state graph (Figure 3.5 line: 11). All the centrality values are stored in a list of dictionaries, where each dictionary corresponds to each node for storing it's all betweenness centrality values in all state graphs in a way that the key refers to the node and values refer to its all centrality values (Figure 3.5 line: 11).

The time required by the system to calculate betweenness centrality varies from network to network. The denser a network is, more time it requires for calculation. After the betweenness centrality values are calculated and stored, the next step is input of state under study (Figure 3.5 line: 14). It can be any state of state graph. Following the input of state of interest is the retrieval of its centrality values which are stored in the list of dictionaries. Once the values are retrieved they are sorted in both ascending and descending manner (Figure 3.5 line: 17). Betweenness centrality of the state (node) in

the former manner points towards its deadlock behavior, because minimum centrality values mean that state is least visited and less connected and it depicts its stability. So if the entered state is a pathogenic state, then it means it is hard for the system to recover from that phase. If entered state is a healthy (good) state, then it means system will remain in a healthy condition or will show most ideal behavior. Whereas, betweenness centrality value in the later manner shows an oscillatory behavior, because a state having a high betweenness centrality value is highly connected and reachable state. So if the entered state is a pathogenic one, then it means it can come out of that phase or if the driving factors (parameter) are repeated it may fluctuate between healthy and diseased condition. Likewise, if the entered state is a healthy (good) one, it will persevere the healthy condition.

After sorting, a CSV file is made which contains three columns i.e. state graph reference number to track the parameter sets, resource sets of all the nodes in the BRN, and the parameter values which are making the respective state graphs (Figure 3.5 line: 34). This data is used to make the heat map for a clearer view and better understanding. Heat map is generated by using built in Pandas and Seaborn libraries (Figure 3.5 line: 35). Interface of developed approach has been created by using Tkinter. It incorporates all the above mentioned functionalities. Pseudo code of developed approach is given in figures 3.5.

PSEUDO CODE: Parameters estimation of BRNs using betweenness centrality

```
1:  procedure PARAMETERS_ESTIMATION(BRN b)
2:  import networkx as nx
3:  from pyrthomas.network_analyser import NetworkAnalyser
4:  import matplotlib.pyplot as plt
5:  import pandas as pd
6:  import seaborn as sns
7:  import csv
8:  G = nx.drawing.read(b.sif);
9:  state_graphs[ ] = NetworkAnalyser.get_possible_state_graphs(G);
10: for graph in state_graphs:
11:         bw_cen = nx.betweenness_centrality(graph)
12:         list_dic.append(bw_cen)
13: while True:
14:         node = input()
15:         for i in lst_dic:
16:                 bw_list.append(i[node])
17:         sorted_bw_list[ ] = sorted(bw_list)
18:         del sorted_bw_list[100:]
19:         with open jnp_csvfile
20:                 jnp_csvfile.write(sorted_bw_list)
21:         close jnp_csvfile()
22:         params[ ] = NetworkAnalyser.get_possible_parameters(G)
23:         with open jnp_csvfile:
24:                 sg_list[ ] = jnp_csvfile.read(sorted_bw_list[1])
25:         close jnp_csvfile()
26:         for s_ind, s_val in enumerate sg_list:
27:                 par_val_list.append(params[s_val])
28:         with open hmap_csvfile:
29:                 hmap_csvfile.write(sg_list, par_val_list, node)
30:         close hmap_csvfile()
31:         df = pd.read_csv(jnp_csvfile)
32:         sns.joinplot(df)
33:         helix = pd.read_csv(hmap_csvfile)
34:         table = helix.pivot('Resources', 'State Graphs', 'Parameters')
35:         sns.heatmap(table)
36:         d = input()
37:         if (d == '0'):
38:                 break
```

Figure 3.5: Pseudo code for parameters estimation of BRNs using betweenness centrality.

3.6 Software and Libraries Used

This approach has been developed in Python language using some of its very famous libraries like NetworkX, Math, Matplotlib, Itertools, Pyrthomas, CSV, Pandas, Seaborn, and Tkinter. The IDE used for programming is Spyder.

3.6.1 Software

3.6.1.1. Spyder

It is an open source IDE (Integrated Development Environment) for Python language. It was developed in 2009 by Pierre Raybaut. It is mainly used for scientific programming. It is integrated with a lot of Python libraries, and those not available can easily be downloaded using "pip install package name" command.

3.6.2 Libraries

3.6.2.1. NetworkX

NetworkX is meant for dealing with complex networks, it helps in their creation and studying the dynamics, structure and functions.

3.6.2.2. Math

This library makes access easy to all Mathematical methods, defined as per C standard. But it doesn't deal with complex numbers.

3.6.2.3. Matplotlib

Matplotlib is a 2D plotting library, which helps in visualizations. Visualizations can be of graphs, networks or plots.

3.6.2.4. Itertools

This library comprises of the collection of tools required to handle iterators i.e. for loop. It is used for iterating lists, dictionaries and other data structures.

3.6.2.5. Pyrthomas

This library corresponds to the Python implementation of René Thomas Logical Formalism. It helps in incorporating all the semantics of Thomas's Formalism.

3.6.2.6. CSV

CSV in Python is used for import and export of databases and spreadsheets. CSV stands for Comma Separated Values.

3.6.2.7. Pandas

Pandas is a BSD licensed library. Its name has been derived from âĂIJpanel dataâĂİ, which is a famous term of Econometrics. It helps in data manipulation and its analysis.

3.6.2.8. Seaborn

Seaborn is based on Matplotlib and it is a data visualization library which helps to draw attractive and informative statistical graphs.

3.6.2.9. Tkinter

It is a Python library for development of standard graphical user interface. It provides an interface to Tk GUI toolkit. It is fast and easy to use and holds a lot of functionalities.

3.7 Example: Mucoidy in *Pseudomonas aeruginosa*

Pseudomonas aeruginosa is a mucus producing bacterium. It is a non-fermenting, oxidase positive, gram negative aerobic rod which is ubiquist in nature, but in some lung diseases it turns to an opportunistic pathogen. It has been frequently found associated in nosocomial infections and its mucus production is the huge cause of mortality in patients suffering from cystic fibrosis. Figure 3.6 shows the regulatory network responsible for mucus production in pseudomonas aeruginosa. AlgU is its main regulator which not only activates its own production but also that of several other structural genes present in the operon. It activates mucB which in turn translates into an anti-AlgU protein.

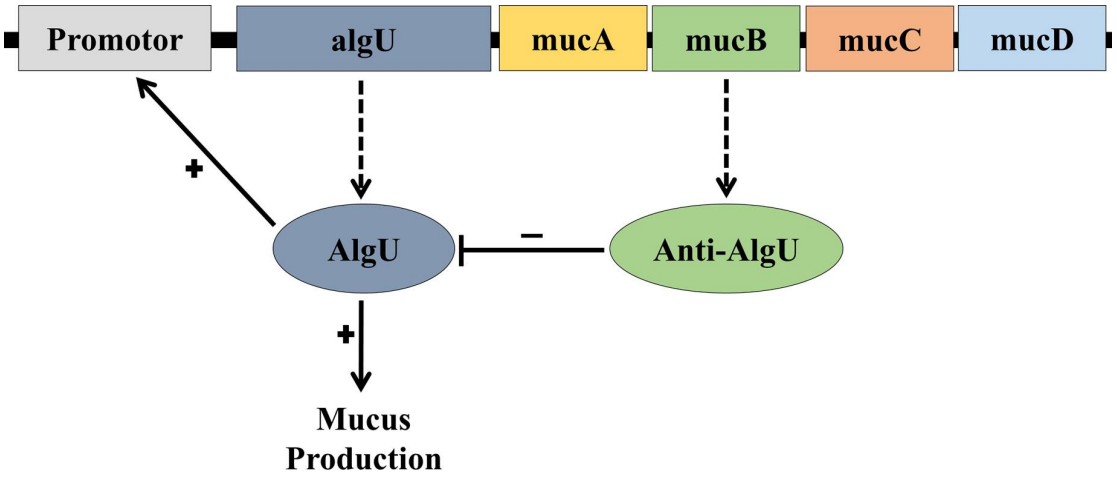


Figure 3.6: Operon model of pseudomonas aeruginosa and regulation of its mucus production.

This regulatory network comprises of two feedback circuits; the first one is a regenerative circuit responsible for excessive mucus production (where AlgU is activating itself) and the second one is a degenerative circuit for controlled mucus production (where AlgU is activating mucB and in turn expression of mucB is inhibiting it, resulting in decreased mucus production). In this case, if we consider the state graphs, (2,1) is a stable and pathogenic (mucoid) state. The first element represents the expression of AlgU (x) and the second one for anti-AlgU (y). This state marks itself as a part of regenerative circuit where AlgU when expresses itself up to threshold of 2, starts producing excessive mucus and being shown multistationary behavior its expression value will only be increased hence, producing more and more mucus. Contrary to this (0,0), (1,0), (1,1), (0,1), (0,0) are non-pathogenic (non-mucoid) states as they control mucus production. The expression of AlgU (x) is constantly kept under control by anti-AlgU (y) thus, maintaining a homeostasis or a stable oscillatory behavior [53].

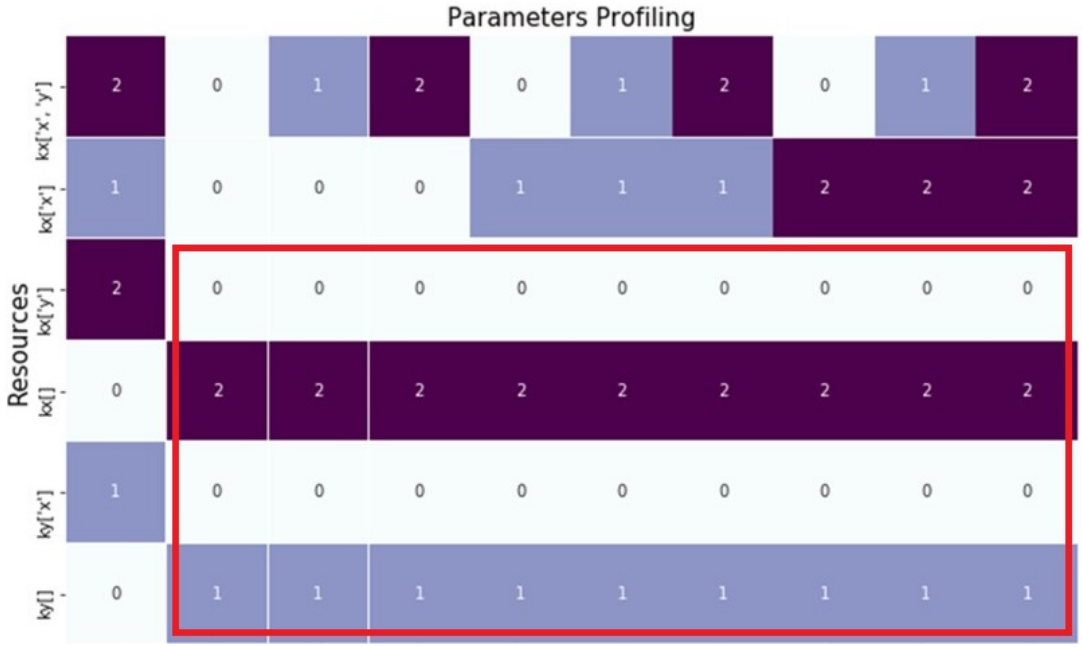


Figure 3.7: Heat map of all the parameter values for state (2,1) in state graphs (along x-axis) where its betweenness centrality value is maximum

Heat map shown in the figure 3.7 is a result of trace back of the state graphs where the centrality value is highest to get the parameter values responsible for these sate graphs. The continuous patterns (highlighted) of parameter values in the heat map are the factors responsible for system to converge at a particular state; i.e. (2,1) in this case. The parameter values in this highlighted chunk corresponds to resources of y and x respectively i.e. ky[ ], ky[x], kx[ ] and kx[y]. ky[ ] means that x is absent and thus y is not expressed, ky[x] means that x is activating y, kx[ ] means that x has no activators and it is being inhibited by y, whereas, kx[y] means x is not activating itself and y is absent thus, x is not being inhibited and maintaining its current expression. The pivotal of all of these is kx[ ], because here y (anti-AlgU) is inhibiting x (AlgU), yet x reaches an expression level of 2. These values are making the system to converge at (2, 1) which being a pathogenic state is leading the system to mucoidy. Whereas, rest of the resources have either no effect or mild effect in causing mucoidy; therefore, they are not picked for further analysis. This can help in devising a treatment plan from mucoid to normal behavior of bacterium either by decreasing concentrations of AlgU directly or

by injecting the system with anti-AlgU.

Chapter 4

Results and Discussion

4.1 Results

This section of the thesis provides with the results obtained by running the developed approach explained in the previous section on an already proposed case study to explain our results in correspondence to the literature and to evaluate the developed approach.

4.1.1 Case Study: MAL Associated BRN

The case study chosen to be run under this developed approach is that of cerebral malaria focusing on MAL (MyD88-adapter-like) protein, which has been found, associated with the onset of disease.

Cerebral Malaria is an aggravated form of Malaria. Usually in Malaria, a pathogen *plasmodium* infects red blood cells. It enters the human body from the bite of female Anopheles mosquito. It is present in its saliva [43]. As a response to infection, human body shows inflammation by producing proinflammatory cytokines like INF-γ (Interferon gamma) and TNF-α (Tumor Necrosis Factor alpha), which help in destroying the pathogen. But, when this pathogenesis becomes grave in its nature, persistent production of cytokines causes an increase in the rate of production of cellular messengers

like iNOS (induced Nitric Oxide Synthase). This produces nitric oxide in brain tissues. An increase in the production of nitric oxide along with ischemic hypoxia caused by plasmodium leads the body to a state of hyper inflammation, which in turn onsets the Cerebral Malaria also known as Diffuse Encephalopathy [6, 14, 26].

In Malaria, proinflammatory cytokines are produced by a signaling pathway which initiates when GPIs (glycosylphosphatidylinositols) are recognized by TLR2/4 (Toll Like Receptors 2 and 4) and it starts forming dimers. Innate immune system of human body uses TLRs to recognize the PAMPS (Pathogen Associated Molecular Patterns). As a result several kinases and primary response protein MyD88 (it differentiates myeloid) are recruited. This signaling cascade comes to its climax when NF$\kappa\beta$ is activated producing inflammatory cytokines [4, 20, 25, 34, 52]. It points towards the role of MAL protein in the pathogenesis followed by hyper inflammation. There are chances that MAL in its wild type can cause the host to develop Cerebral Malaria; whereas, in its mutated form, it can control the inflammation. Thus, it is considered as a potential therapeutic target while fighting against the Cerebral Malaria pathogenesis [39, 44, 64].

Focusing on MAL, there are some proteins associated with it like BTK, INCY, and SOCS-1. BTK is a kinase that falls in the Tec family of proteins. It positively regulates MAL by phosphorylating it [24, 48].

Contrary to this SOCS-1 negatively regulates phosphorylated MAL via polyubiquitination. It is induced by INCY, which apart from this causes inflammation as well. Moreover, it also blocks NF$\kappa\beta$ and thus, it is considered as the negative regulator of this signaling pathway [40, 63].

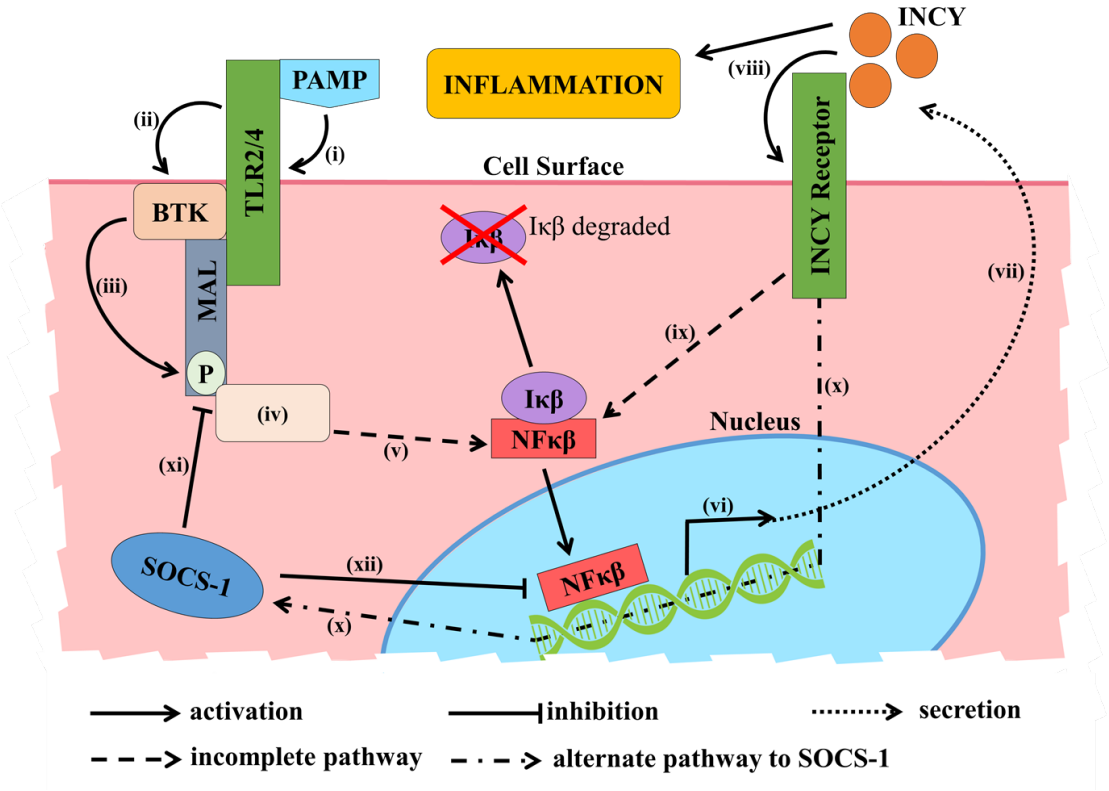


Figure 4.1: Pathway for Cerebral Malaria [2] **(i)** Pathway is initiated when PAMPS are recognized by TLR2/4. **(ii)** It leads to activation of BTK. **(iii)** This in turn phosphorylates MAL and activates it. **(iv)** Kinases and MyD88AP are recruited and are activated around MAL. **(v)** Followed by degradation and activation of I$\kappa\beta$ and NF$\kappa\beta$ respectively. **(vi)** INCY is produced after activation of cytokine genes. **(vii)** INCY activates its regulators. **(viii)** Meanwhile, INCY produces inflammation as well. **(ix)** Again NF$\kappa\beta$ gets activated. **(x)** NF$\kappa\beta$ induces SOCS-1 production via an alternate pathway. **(xi)** SOCS-1 inhibits MAL via polyubiquitination. **(xii)** SOCS-1 blocks expression of NF$\kappa\beta$.

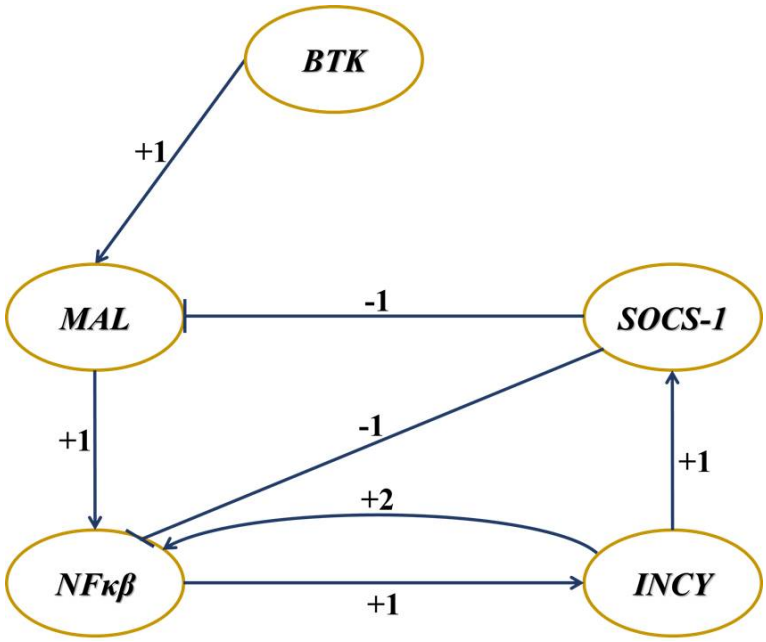


Figure 4.2: This is the MAL associated BRN abstracted from TLR2/4 signal transduction pathway. The nodes represent the proteins involved in this pathway and the arrow heads direct the flow of pathway. Integers (-1, +1, +2) depict the threshold values of expressions which these proteins have to attain in order to have their influence on the targeting one. The positive and negative signs show the type of influence, i.e. activation and inhibition respectively [2]

.

4.1.2 Parameters Estimation of MAL Associated BRN

Graphical User Interface for proposed approach has been developed using Tkinter library of Python. It simply takes the BRN in SIF format; shows the BRN on interface and generates the state graphs and calculates the betweenness centralities at the back end. The text box gets enabled to take input once all the calculations have been done. After that, the desired stable state can be input to get heat maps either for its highest betweenness centrality values or for lowest.

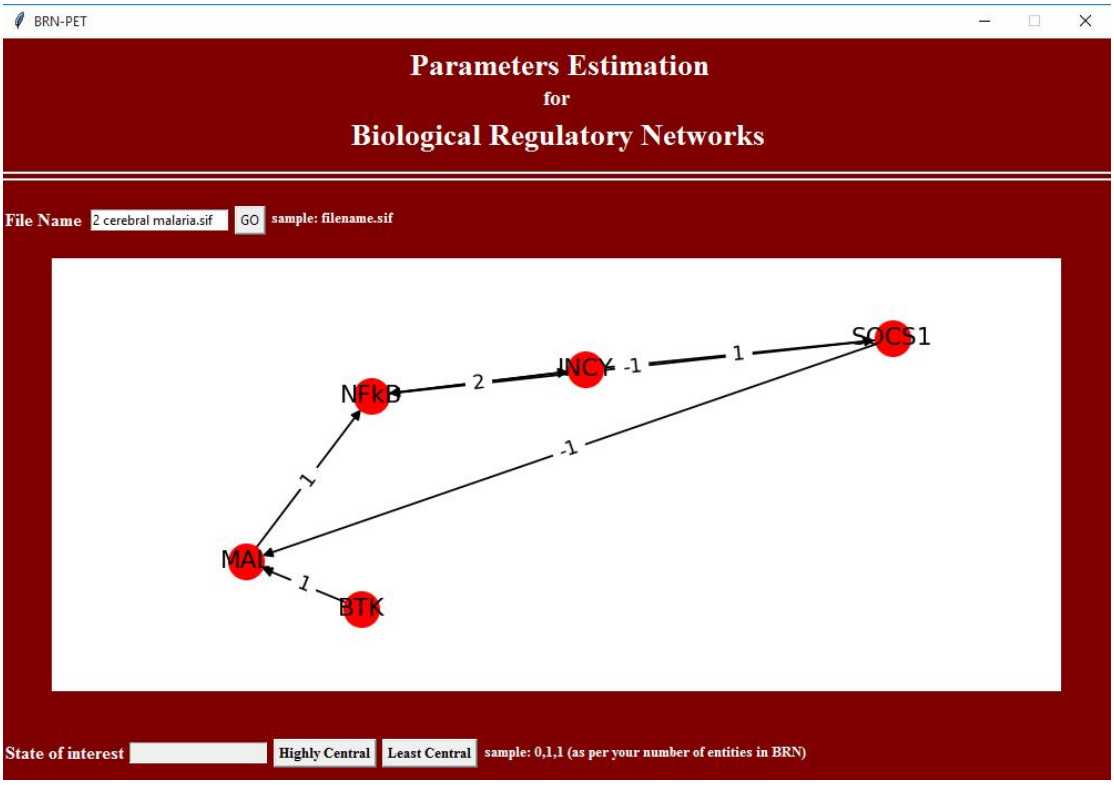


Figure 4.3: Graphical User Interface for developed approach.

In the case of MAL associated BRN, there are two stable states i.e. 00000 and 00121. The former one shows normal behavior where the disease has been cured and the later one shows pathogenic behavior leading to severe hyper inflammation. The order of states owing to the concentrations of expression levels of entities is like this: BTK, MAL, NF$\kappa\beta$, INCY and SOCS-1. For normal state the expression level values for all the entities are zero; whereas, for pathogenic state it is zero for BTK and MAL, 1 or NF$\kappa\beta$, 2 for INCY and 1 for SOCS-1. Both states exist simultaneously and depict two distinct behaviors of the system. System attains the first stable state (00000) after an inflammatory response has been generated against infection and has caused its elimination, whereas, it attains the second state (00121) when infection has been prolonged and cytokines (INCY) are kept on releasing, turning inflammation to hyper inflammation. System has this capacity to maintain its stability after perturbations iff the concentration of INCY remains 1 for any state in the system. If the concentration of INCY reaches level 2 which is the threshold concentration for pathogenic state, then most of the trajectories encountered

by system will only take it to pathogenic/diseased state i.e. 00121.

This syllogism is supported and backed up with experimental observations which were made in a bacterial sepsis case study, where the onset of disease was associated with the hyper inflammation caused by over production of INCY [35, 62].

According to [2] concentration of expression level for INCY is already elevated as an aftermath of prior inflammation due to infection, which increases the chance of system to enter pathogenic state. A reported and clear indication is state 00111, which diverts the system towards pathogenic state with one transition only. Analysis of the underlying system and validation of the developed approach has been done by studying and analyzing the stable states with respect to both their high and low betweenness centrality values.

4.1.2.1. Analysis on the Basis of Low Betweenness Centrality

Lowest betweenness centrality value of a particular node (state) in a network (state graph) means it is the least visited node by other neighboring nodes in the same network. There are total 294912 state graphs of MAL associated BRN as per René Thomas Logical Formalism. Betweenness centrality of each node of each state graph has been calculated and sorted in ascending order. Due to large state space and for the ease of manual analysis only the first 1000 centrality values are retrieved for both states (00000 and 00121). The first 1000 centrality values for both the states are zero. To narrow down further, analysis is performed on first 10 sets of parameters corresponding to first 10 state graphs.

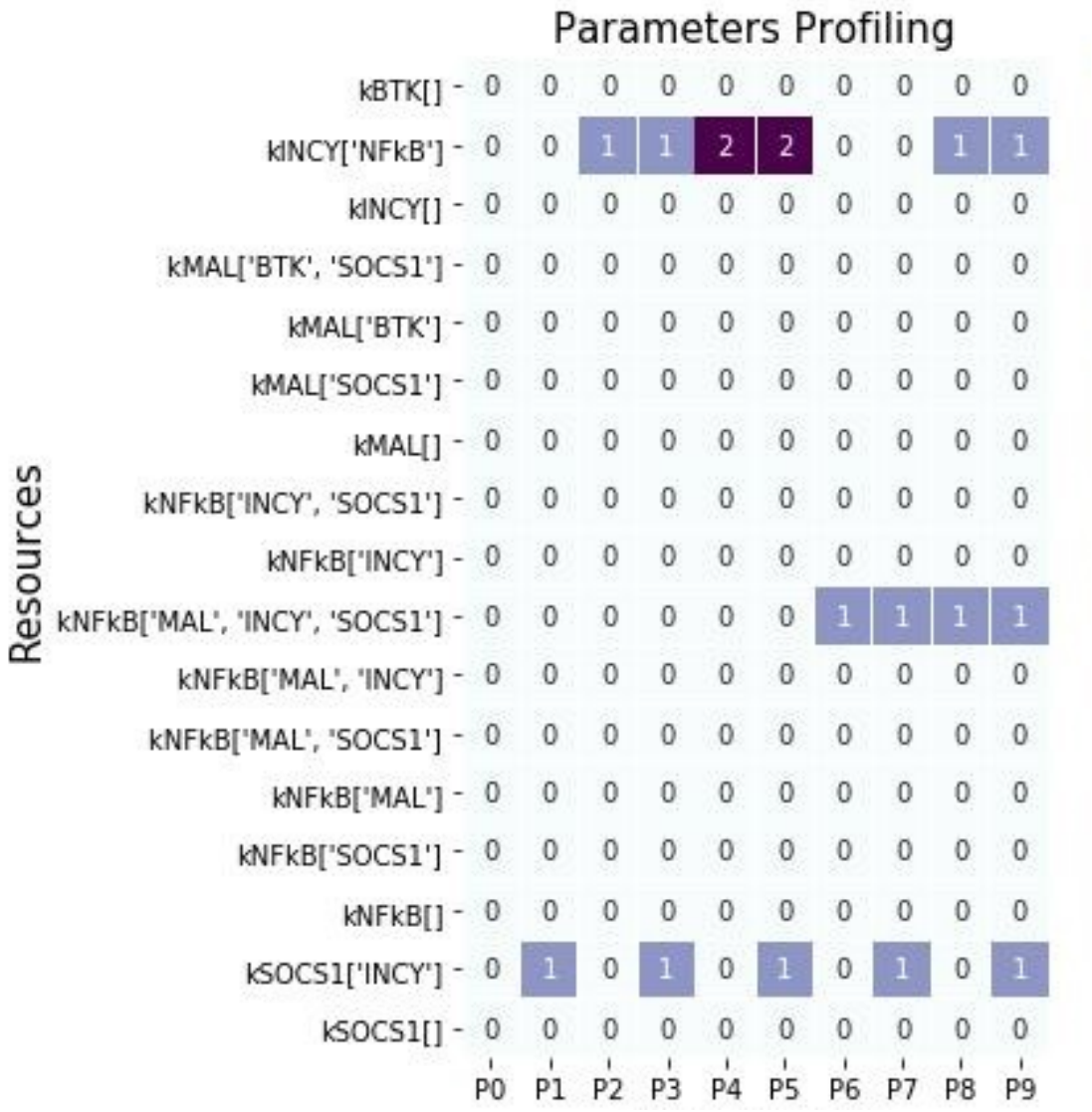


Figure 4.4: Heat map for estimated parameters of state 00000 sorted on the basis of low betweenness centrality.

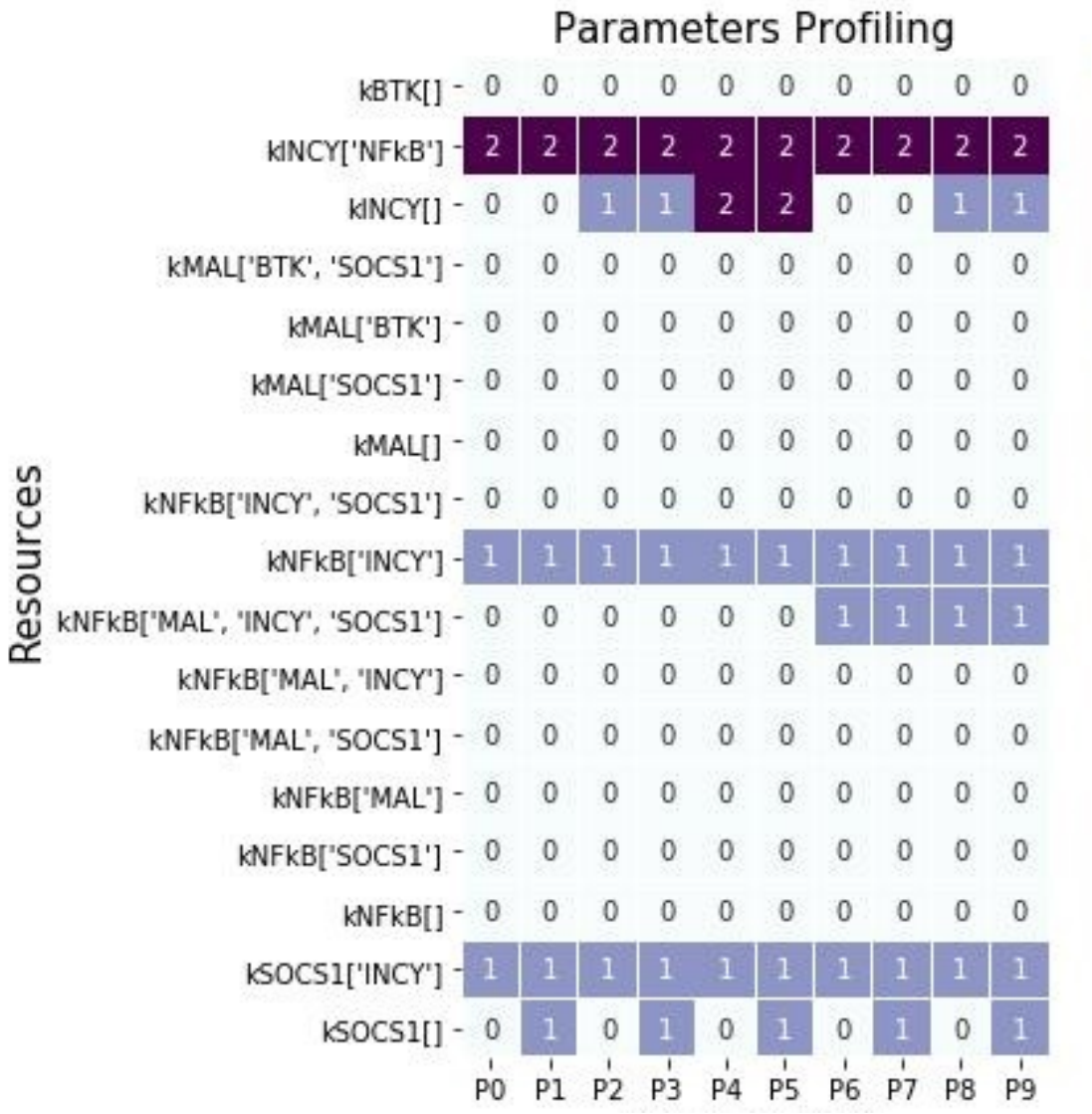


Figure 4.5: Heat map for estimated parameters of state 00121 sorted on the basis of low betweenness centrality.

Stable states (00000 and 00121) appeared as deadlock states in their corresponding state graphs. Both sets of state graphs for 00000 and 00121 respectively, are analyzed and the outcomes are matched with those reported in literature. The states reported in [2] are responsible for taking the system to normal state are 10000, 00001, 00011, 00111, 01111, 01110, 01100, 01000, 11000 and 11100; and those responsible for diverting the system to pathogenic state are 00111, 01111, 11111, 11110, 11100, 11120, 11121, 01120 and 01121. These states are exactly the same as the states isolated by analysis using the developed approach.

Table 4.1: Occurrence of reported states in state graphs generated by parameter sets for 00000 sorted on the basis of low betweenness centrality.

Parameter Sets	10000	00001	00011	00111	01111	01110	01100	01000	11000	11100
P0	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P5	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P6	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P7	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P8	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
P9	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 4.2: Occurrence of reported states in state graphs generated by parameter sets for 00121 sorted on the basis of low betweenness centrality.

Parameter Sets	00111	01111	11111	11110	11120	11121	01120	01121	11100
P0	✓	✓	✓	✓	✓	✓	✓	✓	✓
P1	✓	✓	✓	✓	✓	✓	✓	✓	✓
P2	✓	✓	✓	✓	✓	✓	✓	✓	✓
P3	✓	✓	✓	✓	✓	✓	✓	✓	✓
P4	✓	✓	✓	✓	✓	✓	✓	✓	✓
P5	✓	✓	✓	✓	✓	✓	✓	✓	✓
P6	✓	✓	✓	✓	✓	✓	✓	✓	✓
P7	✓	✓	✓	✓	✓	✓	✓	✓	✓
P8	✓	✓	✓	✓	✓	✓	✓	✓	✓
P9	✓	✓	✓	✓	✓	✓	✓	✓	✓

4.1.2.2. Analysis on the Basis of High Betweenness Centrality

High betweenness centrality value of a particular node (state) in a network (state graph) means it is the most visited node by the neighboring nodes in the same network. In state graphs where stable state has high betweenness centrality value, it is a part of certain cycles. Moreover, it means there exist trajectories which are responsible for taking the system towards that state, but such behavior won't last because it is a part of cycle; so this behavior comes and goes in a cyclic manner.

In case of a diseased state, parameter sets responsible for the generation of state graphs where the diseased state has high betweenness centrality values unveil the basic information regarding expression levels of entities required to avoid pathogenesis. Likewise, in case of normal stable state, parameter sets responsible for the generation of state graphs where the diseased state has high betweenness centrality values highlight the factors that can help in cure if kept under consideration. Parameter sets in both scenarios help in the identification of potential therapeutic targets.

During the analysis of both the states (00000 and 00121) with high betweenness centrality values, cyclic behavior is observed which is sure to occur. Like low betweenness centrality analysis, 1000 centrality values are retrieved and analysis is performed on first 10 sets of parameters corresponding to first 10 state graphs.

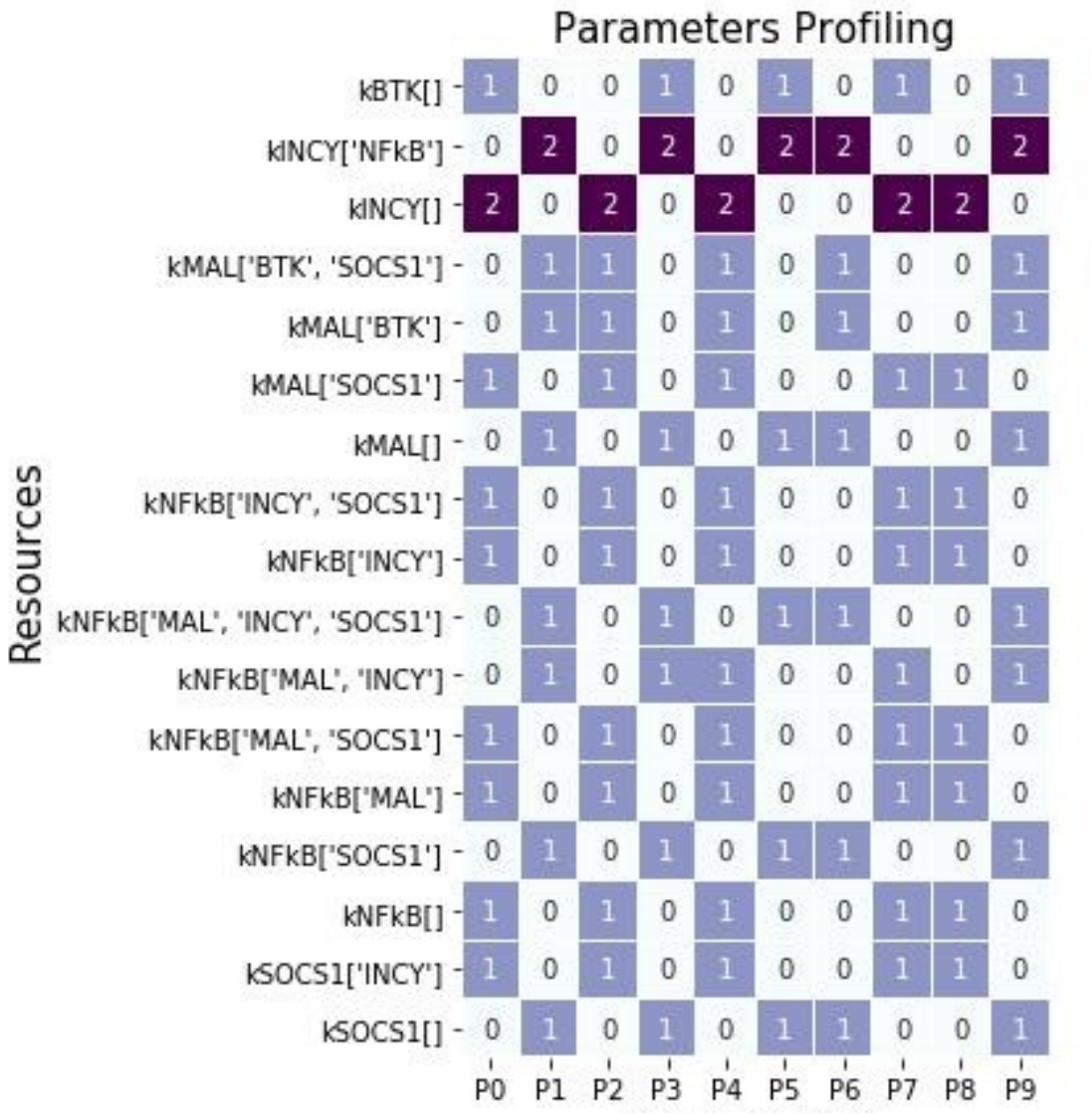


Figure 4.6: Heat map for estimated parameters of state 00000 sorted on the basis of high betweenness centrality.

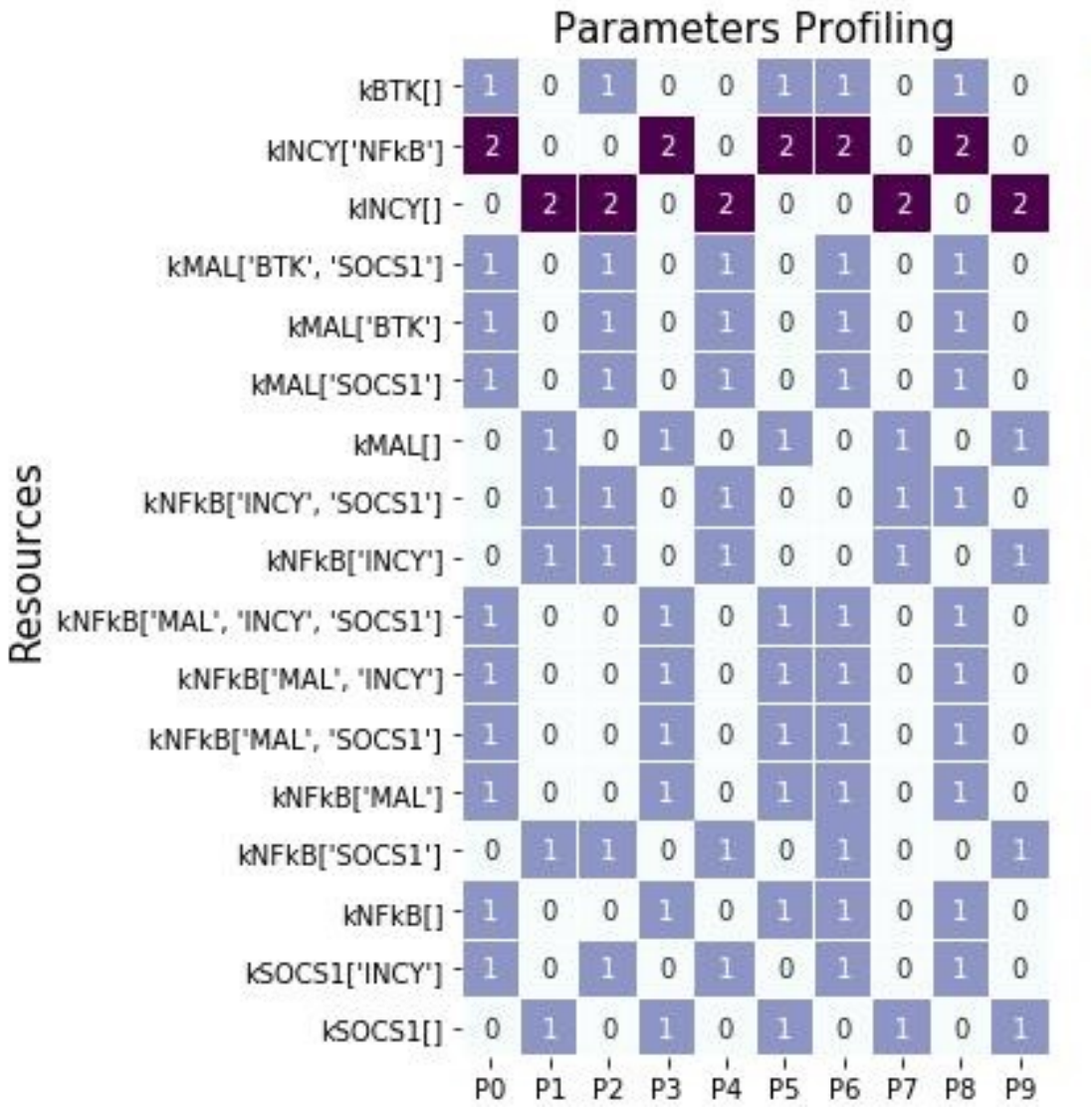


Figure 4.7: Heat map for estimated parameters of state 00121 sorted on the basis of high betweenness centrality.

Several cycles existed for both the states corresponding to each parameter set. While studying 00000, 00121 appeared in some of its cycles and likewise for 00121, 00000 appeared in a lot of its cycles. Details are mentioned in tables 4.3 and 4.4.

Table 4.3: Information of cycles generated in state graphs due to parameter sets sorted on the basis of high betweenness centrality of state 00000

Parameter Sets for 00000	Total No. of Cycles	Cycles for 00000	Cycles for 00121
P0	40	32	13
P1	40	32	13
P2	40	32	13
P3	40	32	13
P4	57	49	18
P5	57	49	18
P6	57	49	18
P7	57	49	18
P8	40	32	18
P9	40	32	13

Table 4.4: Information of cycles generated in state graphs due to parameter sets sorted on the basis of high betweenness centrality of state 00121

Parameter Sets for 00121	Total No. of Cycles	Cycles for 00000	Cycles for 00121
P0	106	10	83
P1	106	12	83
P2	106	94	85
P3	106	94	85
P4	106	94	85
P5	106	94	85
P6	152	72	117
P7	152	72	117
P8	45	9	36
P9	45	9	36

Similar states are responsible for the cyclic behavior of both the stable states. These states include 00100, 00101, 00111, 00110, 00120, 00020, 00010, 00011, 01011, 01010, 01000, 01111, 01110, 01120, 01021, 01020, 01101, 00021 and 01121.

For 00121, three states are identified which are responsible for diverting the system from normal behavior to diseased behavior. These include 00111 as reported in [2], 00021 and

01121. 00111 and 00021 have shown a huge responsibility in diverting the system towards pathogenesis. 01121 has also played role in it but it's not that significant. Thus, the parameter sets which are responsible for 00111 $\rightarrow$ 00121 and 00021 $\rightarrow$ 00121 carry high importance as they make these states as potential therapeutic targets in order to avoid hyper inflammation.

4.2 Discussion

Parameter sets sorted on the basis of high betweenness centrality are responsible for generating cycles. In case of pathogenic states, these parameters are responsible for diverging the system from normal behavior to diseased behavior. Likewise for normal states, they tell how system can be cured and reverted back to a normal path. Contrary to high betweenness centrality, parameter sets sorted on the basis of low betweenness centrality yield deadlocks. In short, parameter estimation on the basis of betweenness centrality will always yield valuable information and this has been validated in the previous section.

4.2.1 Limitations

The developed approach works well as per its proposed concept but nothing is perfect and limitations always exist. Limitation of this approach is the processor and RAM of computer system. As faster the system is, efficient the approach is. The case study of MAL associated BRN took approximately 22 minutes of processing time on a computer system of 3.6 GHz processor and 16 GB RAM whereas, it took approximately 80-90 minutes on a system having 2.4 GHz processor and 4 GB RAM. Moreover, the approach works efficiently on simple and partially dense networks but as the density and size of the network increases then it becomes problematic causing the system to freeze.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

As mentioned in the chapters 1 and 2, parameters estimation of BRNs is a challenging task which has been done using several techniques out of which model checking is one. This study focuses on the development of a new approach for parameter estimation of BRNs using concept of network centrality with René Thomas logical formalism. It is not as exhaustive as model checking because the complexity of parameters estimation using model checking is $O(2^{n2^n}+2^n)$ and that of using betweenness centrality is $O(2^{n2^n}+nm)$. Moreover, it is easy to use for Biologists as model checking requires properties to be written in CTL formulas, which is not an easy job specifically for people having no programming background. This approach has been developed using Python 3.7.0 by incorporating its several libraries. It works efficiently on small to medium size networks. The work is validated correctly by executing the approach on a case study of MAL associated BRN and then comparing its results with already published results. The estimated parameters are helpful in unveiling different kinds of information, especially for the study of mutagenesis and drug designing.

5.2 Future Work and Extension

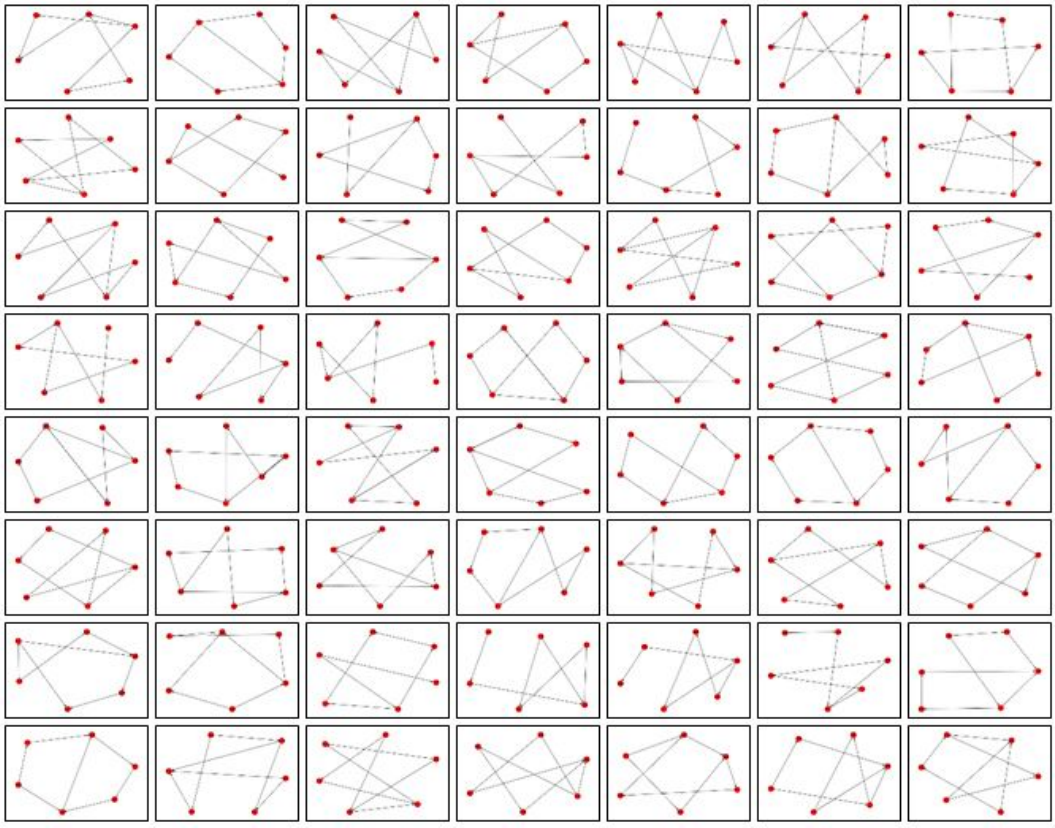
Parameters are very crucial to understand the dynamics of biological systems. Thus, their estimation is very important. Pioneer work of Thomas and other scientists in this domain has benefited biologists a lot. Excessive experiments are replaced by computational tools to decrease the resource and time cost required by the experiments, as these tools predict the behaviors and provide with specific sets particular to the problem to perform wet lab experiments. It is difficult to understand and use programming and computational tools based on it thus, there is a dire need to develop sophisticated softwares which are user friendly and easy to use by biologists. As mentioned in the section 4.2.1, that developed approach works well and efficiently with fast processor and bigger RAM, but not so efficient in the case of computer systems with slow processor and small RAM; therefore, it can be extended in the future to make it parallel in order to make it work efficiently on not so fast computer systems. Moreover, functionality of identification of stable states can be added to this and the manual post parameter estimation analysis can be automated in the future to lessen the difficulty of user.

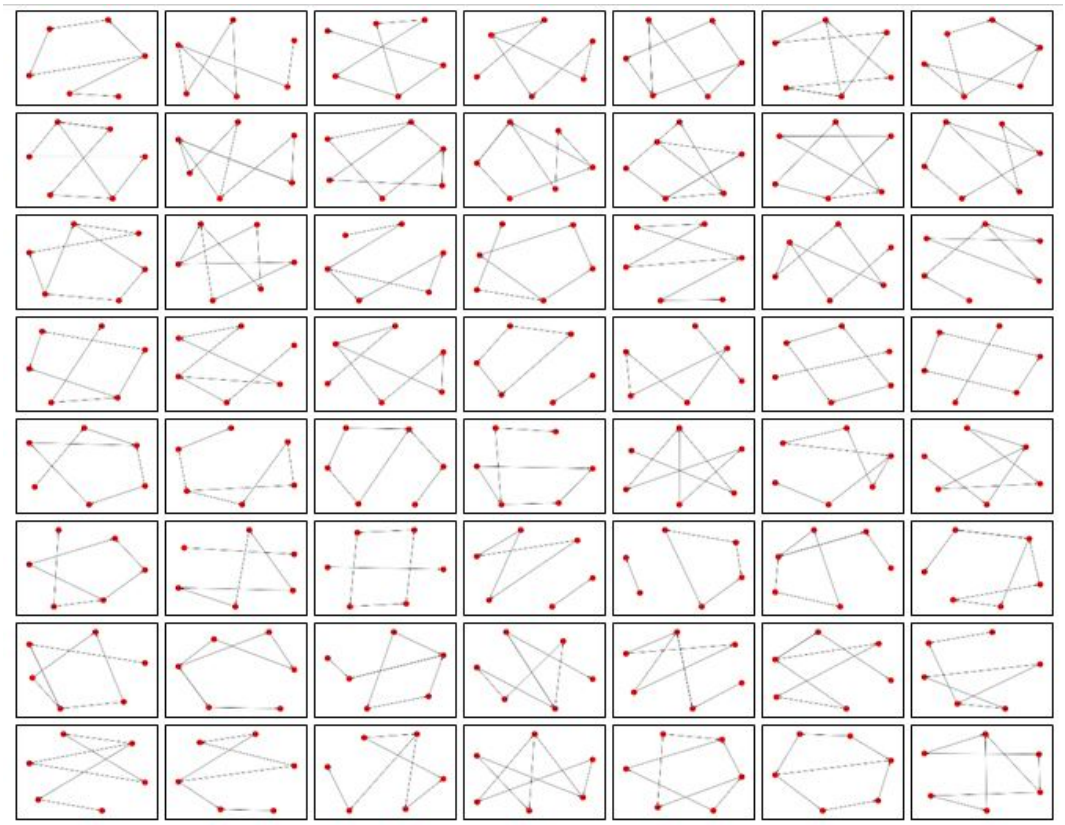
Appendices

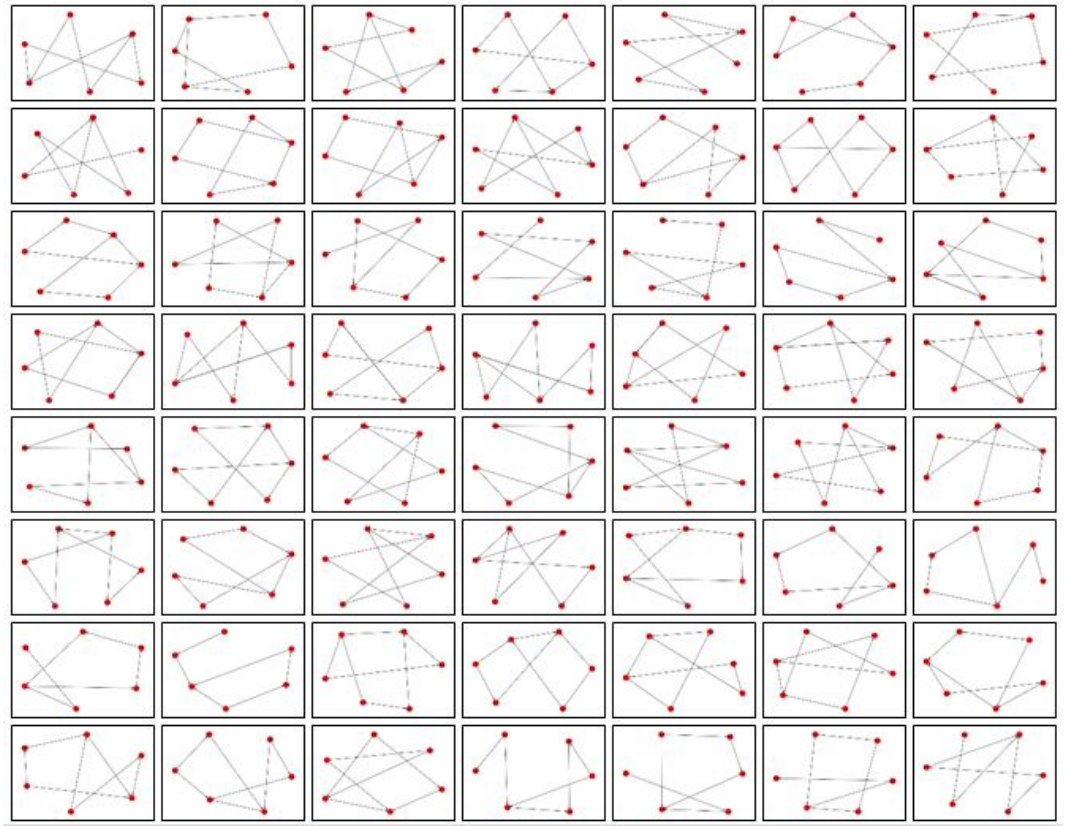
Appendix A

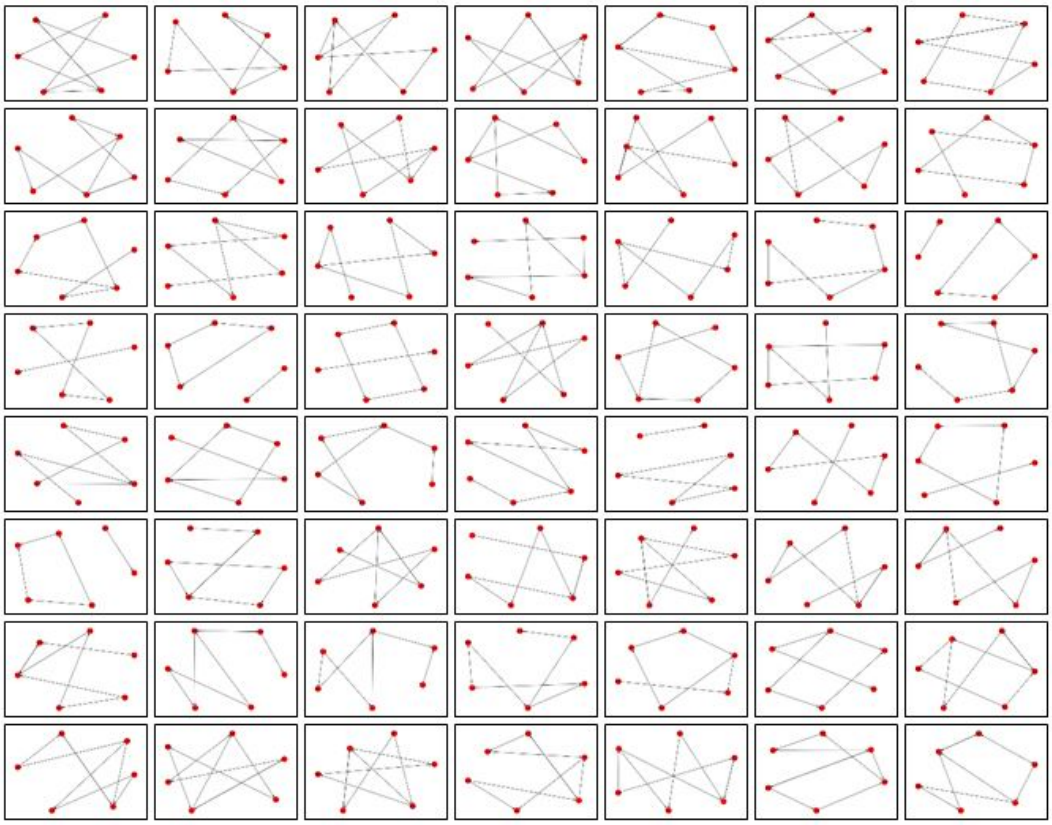
Supplementary Content

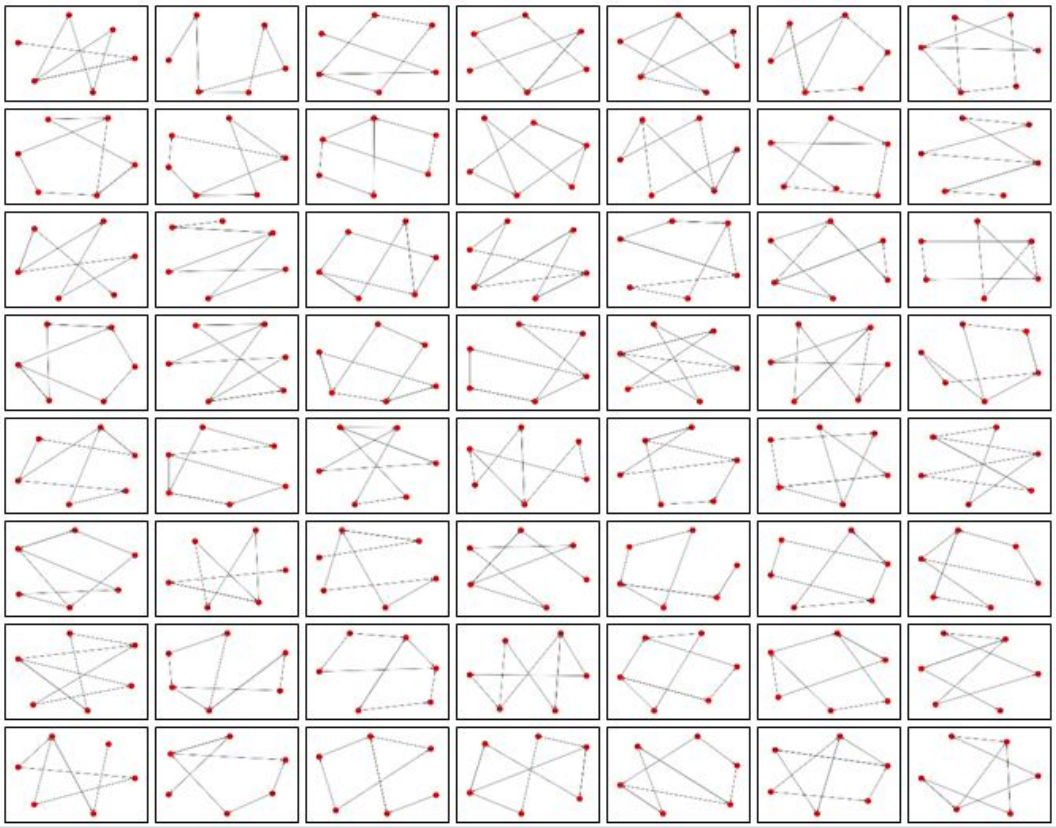
A.1 State Graphs of *Pseudomonas aeruginosa*

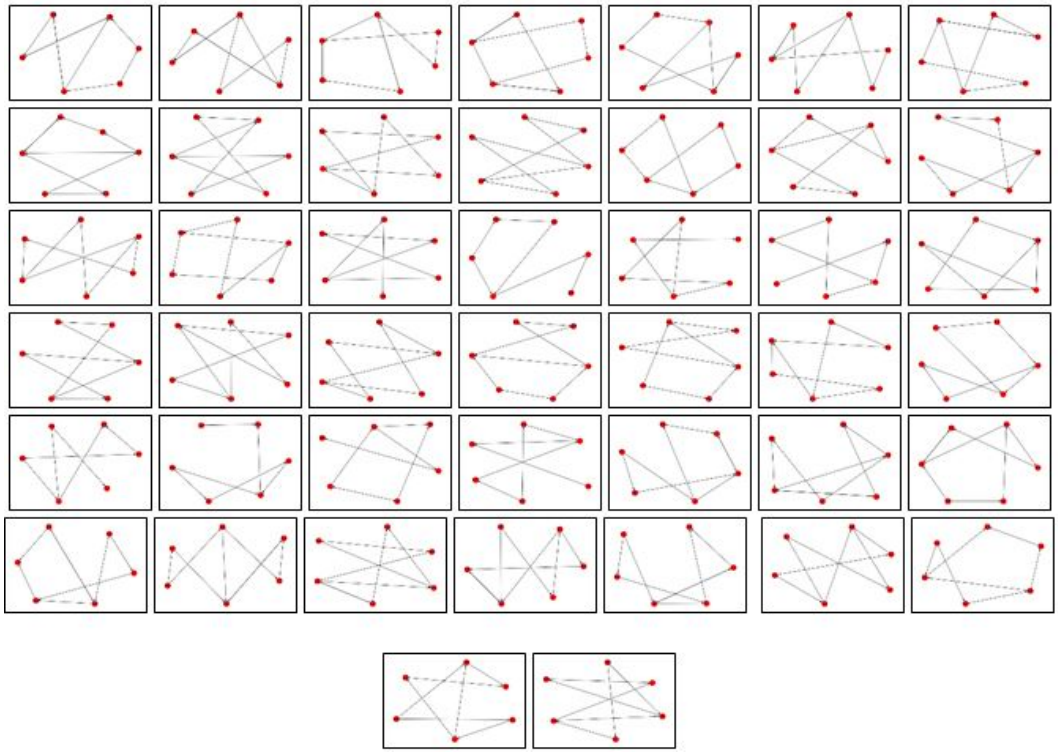












References

[1] Jamil Ahmad, Gilles Bernot, Jean-Paul Comet, Didier Lime, and Olivier Roux. Hybrid modelling and dynamical analysis of gene regulatory networks with delays. *ComPlexUs*, 3(4):231–251, 2006.

[2] Jamil Ahmad, Umar Niazi, Sajid Mansoor, Umair Siddique, and Jaclyn Bibby. Formal modeling and analysis of the mal-associated biological regulatory network: insight into cerebral malaria. *PloS one*, 7(3):e33532, 2012.

[3] Tero Aittokallio and Benno Schwikowski. Graph-based methods for analysing networks in cell biology. *Briefings in bioinformatics*, 7(3):243–255, 2006.

[4] Shizuo Akira and Kiyoshi Takeda. Toll-like receptor signalling. *Nature reviews immunology*, 4(7):499, 2004.

[5] Reka Albert. Scale-free networks in cell biology. *Journal of cell science*, 118(21):4947–4957, 2005.

[6] K Artavanis-Tsakonas, JE Tongren, and EM Riley. The war between the malaria parasite and the immune system: immunity, immunoregulation and immunopathology. *Clinical & Experimental Immunology*, 133(2):145–152, 2003.

[7] Babar Aslam, Jamil Ahmad, Amjad Ali, Rehan Zafar Paracha, Samar Hayat Khan Tareen, Umar Niazi, and Tariq Saeed. On the modelling and analysis of the regulatory network of dengue virus pathogenesis and clearance. *Computational Biology and Chemistry*, 53:277–291, 2014.

[8] Alex Bavelas. A mathematical model for group structures. *Human organization*, 7(3):16, 1948.

[9] Murray A Beauchamp. An improved index of centrality. *Behavioral science*, 10(2):161–163, 1965.

[10] Gilles Bernot, Franck Cassez, Jean-Paul Comet, Franck Delaplace, Céline Müller, and Olivier Roux. Semantics of biological regulatory networks. *Electronic Notes in Theoretical Computer Science*, 180(3):3–14, 2007.

[11] Gilles Bernot, Jean-Paul Comet, Adrien Richard, and Janine Guespin. Application of formal methods to biological regulatory networks: extending thomasâĂŹ asynchronous logical approach with temporal logic. *Journal of theoretical biology*, 229(3):339–347, 2004.

[12] Ulrik Brandes. A faster algorithm for betweenness centrality. *Journal of mathematical sociology*, 25(2):163–177, 2001.

[13] Miguel Carrillo, Pedro A Góngora, and David A Rosenblueth. An overview of existing modeling tools making use of model checking in the analysis of biochemical networks. *Frontiers in plant science*, 3:155, 2012.

[14] Ian A Clark and WB Cowden. Why is the pathology of falciparum worse than that of vivax malaria? *Parasitology Today*, 15(11):458–461, 1999.

[15] Edmund M Clarke and E Allen Emerson. Design and synthesis of synchronization skeletons using branching time temporal logic. In *Workshop on Logic of Programs*, pages 52–71. Springer, 1981.

[16] Hidde De Jong. Modeling and simulation of genetic regulatory systems: a literature review. *Journal of computational biology*, 9(1):67–103, 2002.

[17] Vincent Devloo, Pierre Hansen, and Martine Labbé. Identification of all steady

states in large networks by logical analysis. *Bulletin of mathematical biology*, 65(6):1025–1051, 2003.

[18] Ernesto Estrada. Virtual identification of essential proteins within the protein interaction network of yeast. *Proteomics*, 6(1):35–40, 2006.

[19] Martin G Everett and Stephen P Borgatti. Analyzing clique overlap. *Connections*, 21(1):49–61, 1998.

[20] Bart Ferwerda, Santos Alonso, Kathy Banahan, Matthew BB McCall, Evangelos J Giamarellos-Bourboulis, Bart P Ramakers, Maria Mouktaroudi, Pamela R Fain, Neskuts Izagirre, Din Syafruddin, et al. Functional and genetic evidence that the mal/tirap allele variant 180l has been selected by providing protection against septic shock. *Proceedings of the National Academy of Sciences*, 106(25):10272–10277, 2009.

[21] Jasmin Fisher and Thomas A Henzinger. Executable cell biology. *Nature biotechnology*, 25(11):1239, 2007.

[22] Linton C Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.

[23] A Gonzalez Gonzalez, Aurélien Naldi, Lucas Sanchez, Denis Thieffry, and Claudine Chaouiya. Ginsim: a software suite for the qualitative modelling, simulation and analysis of regulatory networks. *Biosystems*, 84(2):91–100, 2006.

[24] Pearl Gray, Aisling Dunne, Constantinos Brikos, Caroline A Jefferies, Sarah L Doyle, and Luke AJ O'Neill. Myd88 adapter-like (mal) is phosphorylated by bruton's tyrosine kinase during tlr2 and tlr4 signal transduction. *Journal of Biological Chemistry*, 281(15):10489–10495, 2006.

[25] Lutz Hamann, Oliver Kumpf, Ron P Schuring, Erkan Alpsoy, George Bedu-Addo, Ulrich Bienzle, Linda Oskam, Frank P Mockenhaupt, and Ralf R Schumann. Low

frequency of the tirap s180l polymorphism in africa, and its potential role in malaria, sepsis, and leprosy. *BMC medical genetics*, 10(1):65, 2009.

[26] Nicholas H Hunt, Jacob Golenser, Tailoi Chan-Ling, Sapan Parekh, Caroline Rae, Sarah Potter, Isabelle M Medana, Jenny Miu, and Helen J Ball. Immunopathogenesis of cerebral malaria. *International journal for parasitology*, 36(5):569–582, 2006.

[27] Hawoong Jeong, Sean P Mason, A-L Barabási, and Zoltan N Oltvai. Lethality and centrality in protein networks. *Nature*, 411(6833):41, 2001.

[28] Björn H Junker, Dirk Koschützki, and Falk Schreiber. Exploration of biological network centralities with centibin. *BMC bioinformatics*, 7(1):219, 2006.

[29] Michael Karin. Nf-κb as a critical link between inflammation and cancer. *Cold Spring Harbor perspectives in biology*, 1(5):a000141, 2009.

[30] Stuart Kauffman. Gene regulation networks: A theory for their global structure and behaviors. In *Current topics in developmental biology*, volume 6, pages 145–182. Elsevier, 1971.

[31] Stuart A Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of theoretical biology*, 22(3):437–467, 1969.

[32] Stuart A Kauffman. *The origins of order: Self-organization and selection in evolution*. OUP USA, 1993.

[33] Zohra Khalis, Jean-Paul Comet, Adrien Richard, and Gilles Bernot. The smbionet method for discovering models of gene regulatory networks. *Genes, genomes and genomics*, 3(1):15–22, 2009.

[34] Chiea C Khor, Stephen J Chapman, Fredrik O Vannberg, Aisling Dunne, Caroline Murphy, Edmund Y Ling, Angela J Frodsham, Andrew J Walley, Otto Kyrieleis,

Amir Khan, et al. A mal functional variant is associated with protection against invasive pneumococcal disease, bacteremia, malaria and tuberculosis. *Nature genetics*, 39(4):523, 2007.

[35] P Kłuciński and G Martirosian. Role of cytokines and pathogen associated molecular pattern receptors in sepsis. *Przeglad epidemiologiczny*, 59(3):695–701, 2005.

[36] Dirk Koschützki and Falk Schreiber. Centrality analysis methods for biological networks and their application to gene regulatory networks. *Gene regulation and systems biology*, 2:GRSB–S702, 2008.

[37] Dave Kuhlman. A python book: Beginning python. *Advanced Python, and Python Exercises*, 2013.

[38] Harold J Leavitt. Some effects of certain communication patterns on group performance. *The Journal of Abnormal and Social Psychology*, 46(1):38, 1951.

[39] Maria Loiarro, Vito Ruggiero, and Claudio Sette. Targeting tlr/il-1r signalling in human diseases. *Mediators of inflammation*, 2010, 2010.

[40] Ashley Mansell, Rosealee Smith, Sarah L Doyle, Pearl Gray, Jennifer E Fenner, Peter J Crack, Sandra E Nicholson, Douglas J Hilton, Luke AJ O'Neill, and Paul J Hertzog. Suppressor of cytokine signaling 1 negatively regulates toll-like receptor signaling by mediating mal degradation. *Nature immunology*, 7(2):148, 2006.

[41] Massimo Marchiori and Vito Latora. Harmony in the small-world. *Physica A: Statistical Mechanics and its Applications*, 285(3-4):539–546, 2000.

[42] Thomas Mestl, Erik Plahte, and Stig W Omholt. A mathematical framework for describing and analysing gene regulatory networks. *Journal of theoretical Biology*, 176(2):291–300, 1995.

[43] Louis H Miller, Dror I Baruch, Kevin Marsh, and Ogobara K Doumbo. The pathogenic basis of malaria. *Nature*, 415(6872):673, 2002.

[44] Umar HK Niazi, Jaclyn Bibby, and Michael J Sutcliffe. In-silico characterization of the effects of phosphorylated tyrosines 86 and 106 on structure and binding of mal: insight into hyperinflammatory response to infection by the human malaria parasites. *Journal of Receptors and Signal Transduction*, 31(1):53–65, 2011.

[45] Maria Concetta Palumbo, Alfredo Colosimo, Alessandro Giuliani, and Lorenzo Farina. Functional essentiality from topology features in metabolic networks: a case study in yeast. *FEBS letters*, 579(21):4642–4646, 2005.

[46] Loïc Paulevé, Morgan Magnin, and Olivier Roux. Refining dynamics of gene regulatory networks in a stochastic π-calculus framework. In *Transactions on computational systems biology xiii*, pages 171–191. Springer, 2011.

[47] Sabine Peres and Comet Jean-Paul. Contribution of computational tree logic to biological regulatory networks: example from pseudomonas aeruginosa. In *International Conference on Computational Methods in Systems Biology*, pages 47–56. Springer, 2003.

[48] Wenji Piao, Chang Song, Haiyan Chen, Larry M Wahl, Katherine A Fitzgerald, Luke A O'Neill, and Andrei E Medvedev. Tyrosine phosphorylation of myd88 adapter-like (mal) is critical for signal transduction and blocked in endotoxin tolerance. *Journal of Biological Chemistry*, 283(6):3109–3119, 2008.

[49] Forrest R Pitts. A graph theoretic approach to historical geography. *The professional geographer*, 17(5):15–20, 1965.

[50] Amir Pnueli. The temporal logic of programs. In *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)*, pages 46–57. IEEE, 1977.

[51] N Pržulj, Dennis A Wigle, and Igor Jurisica. Functional topology in a network of protein interactions. *Bioinformatics*, 20(3):340–348, 2004.

[52] Rajendranath Ramasawmy, Edecio Cunha-Neto, Kellen C Fae, Susan CP Borba, Priscila C Teixeira, Susanne CP Ferreira, Anna C Goldberg, Barbara Ianni, Charles Mady, and Jorge Kalil. Heterozygosity for the s180l variant of mal/tirap, a gene expressing an adaptor protein in the toll-like receptor pathway, is associated with lower risk of developing chronic chagas cardiomyopathy. *The Journal of infectious diseases*, 199(12):1838–1845, 2009.

[53] Usman Rauf, Umair Siddique, Jamil Ahmad, and Umar Niazi. Formal modeling and analysis of biological regulatory networks using spin. In *2011 IEEE International Conference on Bioinformatics and Biomedicine*, pages 304–308. IEEE, 2011.

[54] Muhammad Tariq Saeed, Jamil Ahmad, Jan Baumbach, Josch Pauling, Aamir Shafi, Rehan Zafar Paracha, Asad Hayat, and Amjad Ali. Parameter estimation of qualitative biological regulatory networks on high performance computing hardware. *BMC systems biology*, 12(1):146, 2018.

[55] Iftikhar Ali Sheikh, Jamil Ahmad, Morgan Magnin, and Olivier Roux. Incorporating time delays in process hitting framework for dynamical modelling of large biological regulatory networks. *Frontiers in physiology*, 10:90, 2019.

[56] El Houssine Snoussi and Rene Thomas. Logical identification of all steady states: the concept of feedback loop characteristic states. *Bulletin of Mathematical Biology*, 55(5):973–991, 1993.

[57] Roland Somogyi, Stefanie Fuhrman, Manor Askenazi, and Andy Wuensche. The gene expression matrix: towards the extraction of genetic network architectures. *Nonlinear Analysis*, 30(3):1815–1824, 1997.

[58] René Thomas. Logical analysis of systems comprising feedback loops. *Journal of Theoretical Biology*, 73(4):631–656, 1978.

[59] René Thomas. On the relation between the logical structure of systems and their

ability to generate multiple steady states or sustained oscillations. In *Numerical methods in the study of critical phenomena*, pages 180–193. Springer, 1981.

[60] René Thomas and Richard d'Ari. *Biological feedback*. CRC press, 1990.

[61] René Thomas, Anne-Marie GATHOYE, and Lucie Lambert. A complex control circuit: Regulation of immunity in temperate bacteriophages. *European Journal of Biochemistry*, 71(1):211–227, 1976.

[62] Hironori Tsujimoto, Satoshi Ono, Philip A Efron, Philip O Scumpia, Lyle L Moldawer, and Hidetaka Mochizuki. Role of toll-like receptors in the development of sepsis. *Shock*, 29(3):315–321, 2008.

[63] Akihiko Yoshimura, Tetsuji Naka, and Masato Kubo. Socs proteins, cytokine signalling and immune regulation. *Nature Reviews Immunology*, 7(6):454, 2007.

[64] Jiankun Zhu and Chandra Mohan. Toll-like receptor signaling pathwaysâĂŤtherapeutic opportunities. *Mediators of inflammation*, 2010, 2010.