Cell Model of Multidimensional Networks

SYNOPSIS

Submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

S R K PRASAD TALASILA

Under the Supervision of

Prof. NEENA GOVEAS

and

Co-supervision of

Prof. BHARAT M DESHPANDE

BITS Pilani

Pilani|Dubai|Goa|Hyderabad

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE,

PILANI

2018

Synopsis

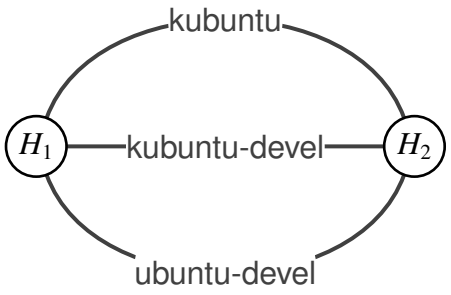
Multidimensional networks are a useful modelling paradigm for representing the multirelational interactions between entities. Such multirelational interactions frequently occur in the application domains of computer networks, social networks and multi-modal transit networks. Figure 1 shows examples of multirelational interactions in three different application domains. In prior literature, multidimensional networks are also referred to as *multilayer networks*. In some application domains like computer networks, the multilayer networks are used with strict sequence of layers and with additional restriction of inter-layer connections to adjacent layers.

1 Models for Multidimensional Networks

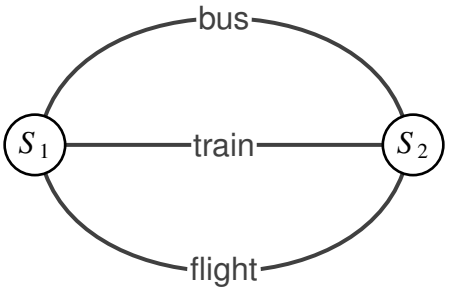
Network models help us create the right abstractions for study of interactions in networked systems. Suitable notation and semantics for nodes and connections is necessary for any valid network model. Researchers have extended the regular graph models to represent multidimensional networks. The notational emphasis of the existing models has the following highlights.

- A node may be present in all the dimensions. Thus effective set of nodes in a multidimensional network is a Cartesian product of nodes and network dimensions. Let this set be represented as V^d.
- Connections may be any subset of $V^d \times V^d$.
- Some network models attribute functions to connections in order to model weighted and dynamic graphs.

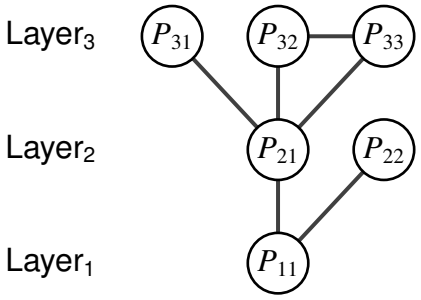
A common theme running through many of the existing models is the focus on the connections. *Network structure* is defined by the connections among



(A) Social network between IRC channel users depicting multiple relations. The text on the connections refers to the IRC channel on which both the users are present.



(B) Transit network between stations depicting multiple relations. The text on the connections refers to the vehicles using a particular mode of transport to service the transit stations.



(C) Protocol graph between protocols depicting use relations. A connection between two protocol nodes refers to a use relation from a higher layer protocol to a lower layer protocol.

FIGURE 1: Examples of multidimensional networks.

the nodes. The study of structural properties of networks such as the diameter of the network, betweenness centrality measures, clustering coefficients etc. is part of the *structural analysis* of a network. Existing literature on multidimensional networks focuses on structural representation and analysis. As a consequence, the existing network models of multidimensional networks do not adequately address the following aspects of network modelling.

- Nodes and connections are passive entities with no ability to compute or to respond.
- Real-life networked systems have cascade events happening in the system and messages flowing through the system. This requires process oriented models for networks. The existing network models do not handle the cascade events.

- Multiple relations between two nodes get represented as seemingly independent connections in different dimensions. The mutual dependence among different relations of any two nodes is not considered.
- There have been studies on the networks of networks for their structural properties and resilience. But the notion of hierarchy to replace a subnetwork with an equivalent node is not considered.
- There is a heavy emphasis on the structural view of the network with little attention paid to functional and behavioural views of network nodes and connections.
- The semantics of the interactions between the participants cannot be clearly specified.

2 Research Objectives

This thesis attempts to overcome the limitations of network models mentioned in the previous section. In order to systematically pursue the stated goal of developing a new network model, the research objectives have been defined. The research objectives of this thesis are to:

1. Create datasets for multidimensional networks from the application domains of transit networks and social networks.
2. Apply the existing network models to perform structural analysis on the multidimensional networks found in social networks and transit networks.
3. Develop a new network model for multidimensional networks. The new network model should be active, must have a process view and account for hierarchy in networks.
4. Demonstrate the proposed model to the research problems in the application domains of protocol analysis and multi-modal transit networks.

3 Thesis Outline

The first two chapters of the thesis introduces the research problem and provides a survey of the relevant literature. The rest of the thesis is broadly structured in the following sequential order.

3.1 Structural View of Network

The initial research contribution is in the structural analysis of multidimensional networks in the domains of social networks, data networks and multi-modal transit networks. The emphasis is on the study of multidimensional networks by performing data collection and analysis. The network diagnostic test (NDT) dataset is considered for the study of the properties of broadband network connections. A statistical analysis of the connection properties has been performed to deduce the Quality of Service (QoS) experience of the Indian broadband users.

The transit timetable data from the public transit service providers of India is analyzed. The transit timetables of Indian Railways are represented as multidimensional network. The multidimensional network representation of transit timetable builds on the previous work and also considers the shared transit vehicles. From the domain of social networks, the data on the Internet relay chat (IRC) community is used to create and analyze the multidimensional structure of the online chat communities.

To overcome the limitations of purely structural models mentioned in Section 1 above, the thesis proposes the Cell Model – a new network model for multidimensional networks.

3.2 The Cell Model

The Cell Model enriches the network participants with structural-functional-behavioral (SFB) views; such a combination of three views provides a chance for better modelling of network participants. In the Cell Model, nodes and connections together are referred to as *network participants*. The interactions between all the participants of a network form the structural view of a network. The nodes and connections are active, i.e., they can perform computation. The Cell Model uses either the functional or the behavioural views to express the computational models of network participants.

The functional view can represent the external effect of a network participant as seen by an external observer. The functional view of a participant does not consider the state of a participant. When considering the state, we utilize the behavioural model of a participant. The behavioural view represents internal

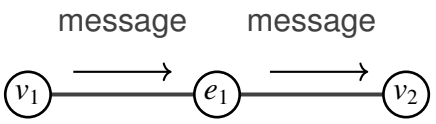


FIGURE 2: A typical scenario of message exchange. A message is sent from node v_1 to its connection e_1; e_1 transfers the message to node v_2.

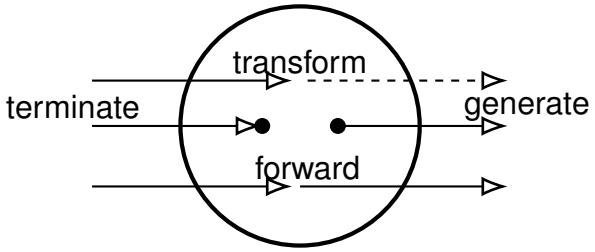


FIGURE 3: Operations on messages performed by a participant of a network. Each keyword denotes the kind of operation performed on a message. The participants are nodes and connections.

changes and external responses to stimuli. We use the state machines to provide a behavioural view and restrict the type of sate machines to finite state machines (FSM).

The Cell Model gives node-like prominence to relations between nodes which are modeled as connections. More importantly, multiple relations between any two nodes are modeled as one single connection, thus leading to a notational consistency. The Cell Model of a network is built using four axioms.

Message Exchange All network interactions happen via message exchanges. In the Cell Model of a network, a connection (ex: e_1 in Figure 2) is drawn from one participant of a network to another if they exchange messages. A message is sent from a node to another node via an edge. The message transfer between two nodes facilitated by a connection is illustrated in Figure 2.

Activeness All the participants (nodes and connections) are active. A participant can generate, terminate, transform or forward a message. The permitted operations are illustrated in Figure 3. Nodes can perform any of the four operations; Connections cannot perform the generate message operation, but can perform terminate, transform or forward operations. The restriction of generate operation to nodes has been put in place to limit causality of events in the network to only nodes. Only nodes can generate new messages which can trigger further events in the network. Connections only act on the messages sent to them.

There is an influence of environment on all the participants of the network. The environmental influence is modeled using independent variables. The computational models of network participants are dependent on the environment variables.

Equivalence Two behavioural models of a participant are equivalent if both have equivalent Finite State Machines (FSM), i.e., both the FSM's are observational equivalent.

Hierarchy A node can represent a sub-network. The representative node must satisfy the equivalence axiom with the sub-network being represented. The external connections of the sub-network remain the same and representative node provides observational equivalent behaviour to the sub-network it replaced.

In the context of message passing cell network, we utilize a very restricted form of hierarchy via FSM equivalence. The semantics of our hierarchical approach have been adopted from David Harel's StateCharts.

3.3 Applications

The Cell Model has been applied to two research problems: protocol analysis in computer networks and itinerary search in multi-modal public transit networks. To apply the Cell Model on the protocol analysis problem, the protocol analysis parse graph needs to be embedded in the Cell Model of network. The graph embedding problem is $\mathcal{NP}$-hard; a heuristic embedding scheme is chosen for demonstration purpose. The Cell Model-based solution is implemented in a software package named Darshini. Darshini has the following features.

- Performs analysis of the selected protocols.
- Provides for configurable tradeoffs between memory consumption and execution time.
- Stores the analysis results into a database.

The Cell Model has been applied to the itinerary search problem in the domain of multi-modal public transit networks. To understand the itinerary search problem better, a non-graph algorithm known as the connection scan algorithm (CSA) has been implemented and extended for direct connections and private transport services. Based on the lessons learnt, the itinerary search problem has been modeled in terms of the Cell Model. The Cell Model-based solution for the has been applied to itinerary search on the public transit networks of India.

In summary, the thesis describes two solutions to the itinerary search problem: tCSA algorithm and the Cell Model-based solution. Both the solutions search the timetables of the public transit networks of India and provide the itinerary search results in less than a second. Thus both the proposed solutions are suitable for creating interactive itinerary planning applications.

4 Research Contributions

This thesis makes the following research contributions.

1. A study of the network conditions experienced by the Indian broadband users. The dataset of the measurement lab's (M-Lab) network diagnostic test (NDT) is used for this study.
2. Representation of the transit timetable of the Indian Railways as multidimensional networks. The case of shared transit vehicles that create multiple entries in the transit timetable is considered in this study. The thesis contains a structural analysis of the resultant multidimensional network.
3. A social network representation and structural analysis for the Internet Relay Chat (IRC) channel logs of Ubuntu and Slackware IRC communities.
4. The Cell Model for the modelling of multidimensional networks. The Cell Model uses the axioms of activeness, message exchange, observational equivalence of FSMs and hierarchy to help model active multidimensional networks. The Cell Model provides flexibility to choose from structural, functional and behavioural views of network participants.
5. An implementation of a protocol analysis tool using the Cell Model. In the process of developing the implementation, the protocol analysis problem has been redefined as a graph embedding problem. The resulting application, named as Darshini, can perform a selective analysis of the protocols, provides for configurable tradeoffs between memory consumption and execution time, and persists the analysis results into a database.
6. Two implementations of multi-modal journey planner. The first implementation is an extension to Connection Scan Algorithm (CSA); the proposed extension adds optimization for direct connections between transit stations and for private transport facilities.

The second implementation uses the Cell Model for modelling of the multi-modal transit networks of India. Both the implementations provide an interactive journey planner that can produce the itinerary search results in less than a second.

5 Conclusions

Based the research contributions listed in the previous sections, the following conclusions can be drawn.

1. Active network models are much better at modelling the process-oriented phenomenon in networks.
2. The Cell Model provides a flexible approach to creating the network models. The computational models of nodes and connections can be as simple as those of regular graph models or as complex as finite state machines.
3. The Cell Model is applicable across multiple domains. This thesis showcases the applicability of the Cell Model to the domains of computer networks and transportation networks.
4. The Cell Model with its message passing paradigm is very close to the message passing method of concurrent computation.

6 Publications

Conference Publications

1. Prasad Talasila, Mihir Kakrambe, Anurag Rai, Sebastin Santy, Neena Goveas, Bharat M. Deshpande, BITS Darshini: A Modular, Concurrent Protocol Analyzer Workbench, 19th International Conference on Distributed Computing and Networking (ICDCN), Varanasi, 4 - 7 January, 2018. **(Core B Rank)** **(Indexed in https://www.scopus.com)**
2. Prasad Talasila, Shaik Asifullah, Neena Goveas and Bharat Deshpande, Transit Timetables as Multi-Layer Networks, 4th Workshop on Intelligent Transport Systems, 10th International Conference on Communication Systems & Networks (COMSNETS), Bengaluru, 3 - 7 January, 2018.**(Indexed in https://www.scopus.com and http://www.scimagojr.com/)**

3. Prasad Talasila, Aparajita Haldar, Suhas S. Pai, Neena Goveas, and Bharat M. Deshpande, Multimodal Transit Scheduler: An Actor-based Concurrent Approach, 20th IEEE International Conference on Intelligent Transportation Systems (ITSC), Yokohama, JAPAN, 16 - 19 October, 2017.**(Indexed in https://www.scopus.com and http://www.scimagojr.com/)**
4. TSRK Prasad, Dhruv Shekhawat, Sukanto Guha, Neena Goveas and Bharat Deshpande, Analysis of Impartial Quality Measurements on Indian Broadband Connections, 22nd National Conference on Communications (NCC), Guwahati, 4 - 7 March, 2016.**(Core B Rank) (Indexed in https://www.scopus.com and http://www.scimagojr.com/)**
5. TSRK Prasad, Kartik Sathyanarayanan, Sukriti Tiwari, Neena Goveas and Bharat Deshpande, t-CSA: A fast and flexible CSA Implementation, 2nd Workshop on Intelligent Transport Systems(COMSNETS), Bengaluru, 8th International Conference on Communication Systems & Networks, 5 - 9 January, 2016.**(Indexed in https://www.scopus.com and http://www.scimagojr.com/)**

Communicated

1. Prasad Talasila, Rohan Goel, Dhruv Shekhawat, Neena Goveas and Bharat Deshpande, Discovering Patterns in Activities of Online Chat Communities: A Case Study on Internet Relay Chat Channels, Springer Journal of World Wide Web (WWW).
2. Prasad Talasila, Neena Goveas and Bharat Deshpande, Cell Model for Multidimensional Networks, IEEE Global Communications Conference (GLOBECOM), 2018.