UNIVERSITY OF CALIFORNIA, SAN DIEGO

On Improving Communication in Emergency Response at Network and Organizational Levels

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Electrical Engineering (Communication Theory and Systems)

by

Raheleh B. Dilmaghani

Committee in charge:

Professor Ramesh R. Rao, Chair
Professor Theodore Chan
Professor Rene Cruz
Professor Sujit Dey
Professor Ingolf Krueger

2010

The dissertation of Raheleh B. Dilmaghani is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2010

DEDICATION

In loving memory of my father,
this dissertation is dedicated to my mother and my sister.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude and deepest appreciation to my advisor, Professor Ramesh Rao, for his invaluable guidance and tremendous support for my research over the past six years. His continuous support, insights, countless discussions, reviews, and long weekend meetings were indispensable in making this research and dissertation possible. Thank you for always finding time to meet with me, for reviewing all of my work, and for believing in me and hearing me out when I needed it most.

I would like to thank Professor Sujit Dey, Professor Ingolf Krueger, Professor Rene Cruz, and Dr. Theodore Chan, Professor of Clinical Medicine, Medical Director of UCSD Emergency Department, for their support and comments, and for serving on my committee.

I am grateful to Professor Sujit Dey for serving as my academic advisor during my first quarter at UCSD and for his guidance and discussions. I would like to thank Professor Robert Lugannani and Professor Ian Galton as I had the opportunity of learning from them as their teaching assistant during my first quarter at UCSD. I would like to also thank Professor Paul Yu for his support. I also thank Professor Laurence Milstein and Professor William S. Hodgkiss for their inspirational teaching.

I would like to thank Dr. Sassan Sheedvash for his invaluable support and comments reviewing this dissertation. I truly appreciate his time and insights for countless discussions, practices, and reviews during my graduate school.

It is my pleasure to acknowledge researchers at Calit2, in particular my co-authors, Dr. Babak Jafarian, and Dr. B.S. Manoj for discussions and developing research ideas. I would also like to thank Dr. James Dunford, Professor of Clinical Medicine and Surgery, UCSD Department of Emergency Medicine and City of San Diego Medical Director (EMS), and Dr. Colleen Buono, Assistant Professor of Clinical Medicine and medical director for the San Diego County Metropolitan Medical Strike Team (MMST), for their generous help in facilitating my participation in the drills and interviews as part of WIISARD project.

My special thanks go to my mother and my sister for their support and encouragement. I owe everything I have achieved in my life to my family who provided me

with unconditional love and guidance. I also like to thank my aunt for her kindness and love and all my friends in Iran for their friendship.

It is my pleasure to acknowledge the kind support and friendship of Calit2 staff: Maureen Curran, Cristian Horta, Allyson Hearst, Alex Hubenko Baker, Alice Dignazio, Karen Stecher, Lovelle Cacho, Vanessa Pool, Samantha Romanin, Yuki Marsden, Lynda Tran, and former researchers and staff, Rajesh Mishra, Dr. Kameswari Chebrolu, and Helena Bristow.

I would like to acknowledge support of the colleagues who I have had the opportunity to work with during my internships at Hewlett Packard, Cisco, and Google who encouraged me to finish this work, in particular, Dr. James Mercer, Roxanna Ganji, Elisa Castillo Gambon, Henry Yu, and Dr. Tony Radi.

It is a pleasure to thank those who contributed during my higher education, in particular Professor Masoud Soltani, Professor Shahrokh Farhangi, Professor Hossein Mohseni, and Dr. Navid Lashkarian in Tehran University, and Professor Mahmoud R. Azimi-Sadjadi, Professor Anura Jayasumana, Professor Carl Wilmsen, and Dr. Indrajit Ray at Colorado State University. I would also like to thank Mahmoud Ahmadipour and Mohammad Ali Moallemi for providing great work opportunity and experience during my four years in Moshanir.

I would like to acknowledge ECE staff and graduate advisors, Karol Previte, M'Lissa Michelson, Gennie Miranda, Bernadette Villaluz, and UCSD Office of Graduate Studies (OGS) and former Assistant Dean for Student Affairs Office of Graduate Studies and Research, Tim Johnston. It is my pleasure to acknowledge all my friends and also my colleagues at UCSD for their friendship and support, in particular, Amir H. Djahanshahi, Mohammad Farazian, Zeinab Taghavi, Seyhan Karakulak, Salih Ergut, and Patrick Amihood.

Lastly, I offer my regards and gratitude to all friends who supported me in any respect during the completion of this work.

This research was supported mainly by National Science Foundation (NSF) as part of RESCUE project (Responding to Crises and Unexpected Events), NSF award number 0331690.

Chapter 1, in part, is a reprint of the material of the following paper: R.B. Dil-

maghani, and R.R. Rao, “A Wireless Mesh Test bed with Applications for Emergency Response” 4th IEEE International conference on Wireless, Mobile Computing, Networking and Communication (WiMob), 2008. The dissertation author was the primary researcher and author of this paper.

Chapter 2, in part, is a reprint of the material of the following papers: R.B. Dilmaghani, and R.R. Rao,“Hybrid Communication Infrastructure and Social Implications for Disaster Management,” Hawaii International Conference on System Sciences (HICSS), January 2007 *, and* R.B. Dilmaghani, R.R. Rao, “A Wireless Mesh Test bed with Applications for Emergency Response” 4th IEEE International conference on Wireless, Mobile Computing, Networking and Communication (WiMob), 2008 *, and* R.B. Dilmaghani, and R.R. Rao “Ph.D. Student Observes Medical Command Center at Rock ‘n’ Roll Marathon,”, 2009. The dissertation author was the primary researcher and author of these papers.

Chapter 3, in part, is a reprint of the material of the following papers: R.B. Dilmaghani, B.S. Manoj, B. Jafarian, and R.R. Rao, “Performance evaluation of RescueMesh: a metro-scale hybrid wireless network,” Proceedings of WiMesh Workshop, IEEE Workshop on Wireless Mesh Networks, held in conjunction with SECON, Santa Clara, 2005 *, and* R.B. Dilmaghani, and R.R. Rao,“Hybrid Communication Infrastructure and Social Implications for Disaster Management,” Hawaii International Conference on System Sciences (HICSS), January 2007 *, and* R.B. Dilmaghani, and R.R. Rao, “Hybrid Wireless Mesh Network with Application to Emergency Scenarios,” Journal of Software Engineer, Volume 3, number 2, Feb. 2008. The dissertation author was the primary researcher and author of these papers.

Chapter 4, in part, is a reprint of the material of the following paper: R.B. Dilmaghani, and R.R. Rao,“A Systematic Approach to Improve Communication for Emergency Response,”Hawaii International Conference on System Sciences (HICSS), January 2009. The dissertation author was the primary researcher and author of this paper.

Chapter 6, in part, is a reprint of the material of the following paper: R.B. Dilmaghani, and R.R. Rao,“A Systematic Approach to Improve Communication for Emergency Response,”Hawaii International Conference on System Sciences (HICSS),

January 2009. The dissertation author was the primary researcher and author of this paper.

VITA

1996	Bachelor of Science, Electrical Engineering, Tehran University, Tehran, Iran
1996-2001	Lead Engineer, Moshanir Power Engineering Consulting Company, Iran
2001-2003	Teaching Assistant, Electrical and Computer Engineering,Colorado State University, Fort Collins, Colorado
2002	Internship, Hewlett Packard, Roseville, California
2003	Master of Science, Electrical and Computer Engineering, Colorado State University, Fort Collins, Colorado
2004	Teaching Assistant, Electrical and Computer Engineering, University of California, San Diego
2007	Internship, Cisco Systems, San Jose, California
2007	Internship, Google, London, United Kingdom
2005-2010	Research Assistant, Electrical and Computer Engineering, University of California, San Diego
2010	Doctor of Philosophy, Electrical and Computer Engineering, University of California, San Diego

PUBLICATIONS

R.B. Dilmaghani, R.R. Rao, "A Systematic Approach to Improve Communication for Emergency Response" *, Hawaii International Conference on System Sciences (HICSS)*, 2009.

R.B. Dilmaghani, R.R. Rao, "A Wireless Mesh Test bed with Applications for Emergency Response" *, 4th IEEE International conference on Wireless, Mobile Computing, Networking and Communication (WiMob)*, 2008.

R.B. Dilmaghani, R.R. Rao, "A Wireless Mesh Infrastructure Deployment with Application for Emergency Scenarios" *, International conference on Information Systems for Crisis Response and Management (ISCRAM) Conference*, 2008.

R.B. Dilmaghani, R.R. Rao, "Hybrid Wireless Mesh Network with Application to Emergency Scenarios" *, Journal of Software Engineer, Volume 3, number 2, Feb. 2008*.

R.B. Dilmaghani, R.R. Rao, "Future Wireless Communication Infrastructure with Application to Emergency Scenarios" *, IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WOWMOM)*, 2007.

R.B. Dilmaghani, R.R. Rao,"Hybrid Communication Infrastructure and Social Implications for Disaster Management" *, Hawaii International Conference on System Sciences (HICSS)*, 2007.

R.B. Dilmaghani, R.R. Rao, "On Designing Communication Networks for Emergency Situations" *, International Symposium on Technology and Society (ISTAS), IEEE*, 2006.

R.B. Dilmaghani, B.S. Manoj, R.R. Rao, "Emergency Communication Systems and Privacy" *, International conference on Information Systems for Crisis Response and Management (ISCRAM) Conference*, 2006.

R.B. Dilmaghani, B.S. Manoj, B. Jafarian, R.R. Rao, "Performance Evaluation of RescueMesh: A Metro-Scale Hybrid Wireless Network " *, WiMesh workshop, IEEE workshop on wireless Mesh networks*, 2005.

ABSTRACT OF THE DISSERTATION

On Improving Communication in Emergency Response at Network and Organizational Levels

by

Raheleh B. Dilmaghani
Doctor of Philosophy in Electrical Engineering (Communication Theory and Systems)
University of California San Diego, 2010
Professor Ramesh R. Rao, Chair

The scale and frequency of large-scale disasters and the wide range of severities and the myriad ways a nation has been affected by, reveals the importance of a reliable communication system in emergency response scenarios. Communication may fail for a broken network component, infrastructure failure, or unreachability. A variety of communication technologies have been deployed at crisis sites but the problems of interoperability, unreachability, unclear communication plan and resource allocation still exist.

Communication in emergency response applications has unique demands for the minimum or no *a priori* knowledge, unpredictability, and short or no advance warning. Through participation in several real-life scenario exercises, analysis of network data, examination of after-incident reports, and interviews with first responders, this dissertation investigates the complex communication problem at both network and organizational levels. This dissertation argues that just the deployment of a robust and reliable communication infrastructure does not solve communication problem in emergency response without investigating and improving communication protocols at organizational level. This work justifies that the communication problem strongly correlates with the dynamics of communication protocols at organizational level among first responders in addition to a good choice of communication technology.

This effort presents event-driven models based on real-life scenarios to exam-

ine structural and behavioral properties of communication protocols at organizational level and to develop supervisory control solutions to assist decision makers with information exchange. This study illuminates the ways in which a supervisor can improve performance by reducing complexity of event-driven models and predicting deadlock in advance. This work incorporates communication technology and organizational communication protocols into traditional stereotypes of communication perspective.

Chapter 1

Introduction

1.1 Background

Consideration of both the scale and frequency of large-scale disasters such as World Trade Center, Hurricane Katrina and California wild fires, reveals the clear need for the continuous availability of a robust communication infrastructure. A reliable communication system helps to improve emergency planning, response and recovery by facilitating timely information exchange. Additionally, the myriad ways the nation has been affected by each of these large-scale disasters proves the importance of developing research in a variety of disciplines.

Communication and data sharing may fail for several reasons, such as a broken links, nodes' failure, loss of line of sight, physical infrastructure failure, power failure, lack of knowledge about radio channel frequency in use, or unsupervised organizational protocols. During an emergency response, access to a reliable communication infrastructure is required to transfer precise information in a timely manner.

1.2 Problem Statement

A failure in communication and timely delivery of precise information has impeded emergency response efforts to date. This made us think why does communication still fail in an emergency response scenarios despite all the advances in the technology?

How can we improve communication for an efficient emergency response? Questions like these initiated our research to investigate communication problems in real-life scenarios. This work has been motivated by the strong need of its application: the impact of communication in an emergency response on human's life and economy. A variety of communication technologies have been deployed at crisis sites but the problems of interoperability, unreachability, unclear communication plan and resource allocation still exist.

Emergency response scenarios have unique properties as dynamic heterogeneous environments and are chaotic in many cases, with small *a priori* knowledge, unpredictability or little advance warning prior to a disaster. We have concluded that these unique properties are the reason that not just any communication technology can solve the problem.

Investigation of network performance based on the measurements conducted over our test bed, clarified that reliable communication and precise information sharing goes beyond communication technology alone, and is very much driven and impacted by the first response organizations' collaboration policies and communication protocols. Therefore this effort to improve communication has tackled communication problems at both network and organizational levels. At the network level, we investigate network performance quantitatively based on the data captured over the network. At the organizational level, we develop an appropriate model such as an event-driven model that allows us to examine the structural and behavioral properties of the system. From an engineering perspective, we must identify unnecessary dependencies to reduce complexity. We also identify properties that help us predict performance. In a hierarchical architecture, not all the detailed information is available to a higher level supervisory controller under normal circumstances, however, when needed, the abstraction can unfold to reveal the details.

1.3 Literature Survey

Interoperability among different devices in a heterogeneous environment has always been a problem. The National Institute of Standards and Technology (NIST) has

a Distributed Test bed for First Responders (DTFR) which is a vehicle that provides wireless connection for first responders at a disaster site. The Advanced Network Technologies Communication for public safety is a division within NIST whose contribution to DTFR is a wireless ad hoc network using 802.11 off-the-shelf devices to provide users with network connectivity. The users can work with several hand held communication devices equipped with an IEEE 802.11b WLAN card and uploaded with NIST software. These communication devices are easy to configure as they automatically form a network as long as they are within radio frequency range of each other. This includes the transfer of voice communications, video streaming, data, short messaging, and access to Internet over wired and wireless networks with the legacy standards [1]. Many existing communication infrastructures became unavailable after Hurricane Katrina hit New Orleans including one of the 800 MHz communication towers. Therefore first responders did not have access to a reliable communication infrastructure to coordinate emergency response activities, transfer data and exchange information locally and to the off site command centers [2]. FEMA provided the affected area with MERS (Mobile Emergency Response Support) to provide a communication network to enable first responders to exchange information via video teleconferencing to the off-site command center. MERS includes trained personnel and a vehicle which provides mobile communication support for voice and data [3].

According to area Sniper investigations report in Washington DC [4], Montgomery County Police Department (MCPD) communicated to the county through available devices including radios working with different frequencies during the emergency response. Each organization (for the city or county) had a dedicated channel to talk to their own dispatcher at all times. Federal agents and MSP (Maryland State Police) officers who were initially assigned to the task force were using their cell phone and portable radios on Very High Frequency (VHF) channels which was not interoperable with task force members. MCPD UHF system was overloaded as the result of incoming traffic to respond to the citizens. A new 800 MHz voice radio system which was fully tested before, but not installed at the time, was placed into service earlier than planned to serve as a patch to provide interoperability with the UHF legacy system. MCPD distributed a few 800 MHz portable radios to be used by the other agencies to provide

interoperability. Audio cross-connect switches were deployed in some counties and the City of Alexandria for sniper investigations. This switch provides interoperability between radio systems working on different frequencies in a heterogeneous environment. Therefore different organizations were able to communicate across systems by switching to the operating frequency.

Based on an article in the New York Times on the 9/11 disaster, police helicopters were asking for evacuation of all people in the area of the second tower before the second airplane hits that tower, but the fire department did not hear those warnings. It is clear that more lives could have been saved if those warnings had been heard. The Wireless Emergency Response Team (WERT) was established after September 11, 2001 to provide wireless connectivity for first responders to facilitate emergency response and rescue efforts [5].

In another disaster scenario, the response to the Virginia Tech Massacre on campus in April 2007, different first response organizations used radios working on different frequencies such as UHF, VHF, 800 MHz, etc. which caused interoperability problems according to this incident report [6]. The Virginia Tech telecommunication network infrastructure is a fiber optic core IP-based data network which provides a reliable communication network for users (wireless or wired) throughout the campus. During the response to the Virginia Tech Massacre on campus in April 2007 which was comprised of two shootings, the load on the network was increased but the network was able to perform reasonably well with no significant breakdown under the high traffic circumstances. During the crisis, the network staff was responsible to support the infrastructure to ensure reliability for data transfer and voice. According to the report communication and exchange of information relying on the internal phone system, instant messaging, e-mail and cellular short message service worked fine during the response. However voice service provided by external vendors reached saturation levels at times making it unreliable communication tool for exchange of information. The network engineers quickly reconfigured a National LambdaRail [7] network connection on-the-fly to increase the network capacity by 1 Gbps to avoid potential bottleneck or congestion in the network and avoid delay or loss of information. According to this report the campus telephone system was able to receive and transfer calls without failure although some

delay was experienced on the lines connecting the campus and the city of Blacksburg to the outside network. There have been several radio coverage problems experienced indoor for different first response organizations according to the report. The signal was attenuated or there was no coverage at some indoor locations. Mobile broadband Internet access was not used by Virginia Tech Police and Virginia Tech rescue Squad during the April 16th incident. Different first response organizations used radios working on different frequencies such as UHF, VHF, 800 MHz, etc. which caused interoperability problems among them.

In 2001, the Miami-Dade Fire Rescue (MDFR) in Florida began to utilize the initial Voice over IP (VoIP) which takes advantage of already existing infrastructure for data to transfer voice at a lower cost. This has enabled fire fighters to talk to each station they need to which is very convenient as many of them are transferred to a different station and need to call other stations [8].

Push-to-talk systems have been used for emergency response for years. The problem with them is the lack of interoperability and incompatibility with the other existing communication technologies.

Text messaging or instant messaging are also beginning to be used for early warnings or exchanging information when a disaster happens. For example during the earthquake in Los Angeles in July 2008, the data network for sending text was not as congested as voice. There was no outage but there was an increase in the volume of calls [9]. Text messaging is a good alternative to share critical information as short messages does not consume as much bandwidth resources and there is a better chance of successful transmission compared to the voice network which is often busy and unreachable due to sudden network load increase after crisis.

We believe that the use of an electronic post-it board is a useful method to share information regardless of technologies used by individual organizations. Information can be accessed via a web-browser on a computer or cellular-phone. There are several software tools available, one being WebEOC which is a web-based video conferencing tool which is not available to all first responders for the high cost of the software license. Therefore creating and developing electronic pages such as wiki pages accessible to public will be very beneficial.

1.4 Contributions

To address communication problem in emergency response, we began our research by deploying a wireless mesh network, a promising candidate for improving network performance. The wireless mesh network provides a robust and reliable low-cost infrastructure which is easy to reconfigure. Our efforts included establishing a wireless mesh network consisting of several wireless access nodes: commercial outdoor units and wireless mesh access nodes developed in-house by researchers, and handling interoperability among them. The interoperability issues of the heterogeneous environment was resolved through the use of multiple interface cards. We performed multiple measurements over the test bed to capture network data and monitored the routing table updates to verify that it concludes paths with better signal strength. The network proved to be reliable, as a new route was seamlessly established when objects blocked the line of sight or when the nodes were moved around in the environment. We present the hybrid Wireless Mesh architecture, RescueMesh which was deployed as a test bed on one floor in a campus building, Atkinson Hall, with nodes placed in different offices, labs, and conference rooms.

With the gained confidence in the network performance and interoperability, the infrastructure was deployed as a test bed for emergency response drills conducted on campus and in the city of San Diego. We captured data over the test bed to analyze network performance. The results revealed a large amount of overhead in network data caused by sending a large number of small control packets [10]. Investigating network traffic patterns further revealed that the problem with communication and information sharing goes beyond communication technology and is impacted by the organizational collaboration policies and protocols. Wireless Mesh infrastructure, although proven to be an effective technology, did not solve all the information flow and data sharing problems observed at emergency response scenarios. We discovered that communication problem strongly correlates with the dynamics of communication protocols at the organizational level among first responders in addition to choosing a reliable communication technology.

For this effort, we approached communication problem in a radically different way. In addition to the infrastructure technology, which turned out to be insufficient,

we developed event-driven models to examine structural and behavioral properties of communication protocols at the organizational level and to develop supervisory control solutions to assist decision makers with information exchange. Systematic methodologies are presented that allow reuse of these methods for similar applications to identify deadlocks and predict performance. The models developed in this work are based on real-life scenarios from emergency response exercises to assess real communication protocols and are used to reach a concrete conclusion.

To bridge the gap between network level and human at the organizational level, we extended our research to consider organizational communication protocols. We contributed to the crisis response field by reducing unnecessary complexities in interactions to achieve efficient performance in one, or more of the following ways: minimizing dependencies, cancellation of an scheduled event and providing an alternative option in case of unreachability of a peer or supervisor, enforcing *liveness* in case of a deadlock, resource allocation and ensuring the receipt of updated messages by all, to prevent the distribution of multiple copies of data and to ensure reliability. The analysis of a large system is simplified by abstracting details of subsystems in a distributed hierarchical architecture. The supervisory solution cannot determine the next state but it can limit the possible set of events that occur by enforcing constraints to ensure deadlock free performance.

1.5 Dissertation Organization

In chapter 2 we present methods of data collection to assess the basis for this research. We provide a description of several drills that we participated at, the process of patients' admission at UCSD emergency medical department, findings from interviewing first responders, examining existing after-incident reports, and the lessons learned. This is followed by a discussion on technical, social, and organizational challenges of deploying a new technology.

In chapter 3 we present the wireless mesh infrastructure which has been deployed as the test bed at several drills, the data that was captured and an analysis on sources of bottlenecks and network performance. In chapter 4 we discuss real-life

emergency response scenarios from a network perspective that bridges network level to organizational level. In chapter 5 we present background information and properties on Petri nets that we have chosen to model and analyze communication protocols at organizational level. In chapter 6 we identify and discuss structural properties that allow complexity reduction and deadlock of large Petri net models.

In chapter 7 we present behavioral performance analysis of the models developed in previous chapters. We present a framework to schedule the execution order or events based on known costs and analyze communication complexity of incident command architecture. In chapter 8 we present systematic methodologies to ensure deadlock free performance by applying supervisory control solutions. Finally, in chapter 9 we conclude the contributions of this work and propose future research directions.

Certain organizational structures have been considered as the basis of the communication. Incident Command System (ICS) or Standardized Emergency Management System (SEMS) is followed in California and National Incident Management System (NIMS) is followed nationwide. In appendices A and B we provide a brief description of ICS and NIMS respectively. In appendix C we present the list of questions that we have asked first responders during the interviews to obtain the organizational knowledge presented in this work. Finally, in appendix D we present a description of the last exercise to assess communication and the methodologies developed in this work.

This chapter, in part, is a reprint of the material of the following paper: R.B. Dilmaghani, and R.R. Rao, "A Wireless Mesh Test bed with Applications for Emergency Response" 4th IEEE International conference on Wireless, Mobile Computing, Networking and Communication (WiMob), 2008.

Chapter 2

Case Studies

2.1 Introduction

We have uniquely collected data at both network level and organizational level based our participation at several real-life case studies conducted on campus, city, and county levels, learning patients' admission process onto UCSD emergency medical department, investigating several after-incident reports and conducting interviews with first responders. We present a description of data collection methods in section 2.2. In section 2.3 through section 2.5 we present the case studies that we have participated at, studied and analyzed. As part of these case studies, we have participated in several exercises run by the city of San Diego, San Diego County MMST (Metropolitan Medical Strike Team) as part of NSF-funded RESCUE project (Responding to Crises and Unexpected Events). San Diego MMST is a team of local responders who work together to develop and coordinate medical response for real-life disaster scenarios. MMST is drawn from different jurisdictions within the county and city to help with the overall emergency response efficiency and resolve territorial issues. In section 2.6 we present a description of our observations at the communication and command control in 12th annual Rock 'n' Roll Marathon in San Diego in May 2009. In section 2.7 we present a brief description of our observation and learning of patients' admission process at UCSD emergency medical department. In section 2.8 we present observations and results learned from evaluating after-incident reports. In section 2.9 we present a set of technical and social challenges that we have observed as well as a few feedbacks from

interactions with social scientists from Hazard workshop [11]. In section 2.10 we summarize the problems that have been investigated to improve communication, and finally, in section 2.11 we highlight the concluding remarks of this chapter.

2.2 Data Collection

We used the following four methods to collect data and information we needed to assess communication problem and improve performance at both network and organizational level.

- Participated at several drills simulating different emergency response scenarios conducted on campus, city, and county levels:
  Network data was captured both at packet level and organizational level. The communication and interactions among responders were investigated to examine the flow of information and to identify flaws to improve the quality of response.

- Learned patients admission process onto UCSD emergency medical department:
  We observed and studied the process of patients' admission at UCSD emergency medical department to model flow of information in one organization. This provided me with a typical system information to allow development of an event-driven system and develop systematic methods to analyze its structural properties such as deadlocks or conflicts in sharing resources, etc.

- Conducted and participated at several interviews with first responders:
  We interviewed several first responders to assess organizational level communication before and after the drill for the *Silver Bullet* drill. Reaching out to first responders to obtain the answers was not an easy task as a few people never responded. We developed a set of questions that is presented in the questionnaire in appendix C. Section 2.5.2 presents valuable information that we gathered based on responses from first responders.

- Investigated several after-incident reports:
  Finally to extend the data set to include a larger number of events that we were

not involved, we reviewed existing after-incident reports to identify the communication problems.

2.3 Mardi Gras San Diego

2.3.1 Description

For Mardi Gras 2006, we deployed a wireless mesh network for a large scale, public event with over 20,000 attendees in downtown San Diego. The goal was to provide connectivity locally and to the Internet. This event needed to be safe and in-control, therefore law enforcement needed to be aware of surroundings and behavioral changes so that they could respond effectively and quickly when necessary. Data from sensors was streamed to San Diego Police department command post and UCSD technology operation center. A report of this event can be found at [12].

2.3.2 Lessons learned

It was a unique opportunity to verify interoperability and observe unexpected signal interference. The interoperability was achieved between commercially available outdoor units and CalMesh boxes developed in house to provide a larger coverage and more reliable network [13]. Signal interference was measured during testing before the event as cell phones passed by the mesh nodes. This was important as based on theory they operate in different range of frequencies and they should not interfere while in practice, the measurements conducted by other researchers in the group confirmed interference [14].

2.4 Operation College Freedom

2.4.1 Description

This drill took place in the home of the UCSD Division of Calit2 on campus, Atkinson Hall, in August 2006 as part of San Diego MMST exercise which involved a

terrorist attack and chemical spill. The purpose of the drill was to practice and test cooperation and responsiveness among San Diego emergency and law-enforcement agencies. It was also an opportunity to deploy several technologies including wireless mesh network and to collect data. The wireless mesh access nodes were distributed across the exercise site such that they provide connectivity and no node was more than two hops away from the gateway. When a truck unpredictably blocked the line of sight between two nodes, the network was partitioned. The design and deployment of the wireless mesh network was such that this failure was transparent to other nodes and they were not impacted. The node connecting to the one which failed was re-routed through another wireless access node. The affected node was moved to another location where a line of sight existed for reconnection.

Several other devices and responders were communicating data over this network including medical devices communicating victim's vital signs to medical team. The network results presented in chapter 3 is based on the measurements conducted over the wireless mesh test bed in this event.

2.4.2 Lessons Learned

On the network level, packet traffic was captured and the throughput, retransmission and communication pattern between nodes were analyzed to identify sources of bottleneck. We found out that the communication between a wireless mesh access node and the gateway experienced larger amount of network data . Network data analysis which will be presented in chapter 3 revealed the source of bottleneck. Sending a large number of small control packets caused a large amount of overhead. There was also a large number of retransmission due to the loss between the same pair of nodes [15].

On the organizational level, there were other problems experienced during this drill, one being the loss of coordination in entering the building. The organizational hierarchy was not communicated well to all groups: there were instructions for agency A to not enter the building until they received direct command from organization B. For various reasons, organization B was not aware that they had to grant permission to group A and as a result there were conflicts when group A entered the building without B's permission. This is critical in safety and saving lives and at occasion, it has led to

observers being arrested which slows down the response.

After noticing such communication problems which do not resolve by deploying a suitable physical network infrastructure, we realized the importance of communication protocols among organizations. Therefore in the following set of exercises, we moved our focus to the study of organizational level communication to develop supervisory solutions to predict and prevent potential potential deadlocks and to improve the overall response.

2.5 Silver Bullet Drill

2.5.1 Description

This was a unique large scale, multi-incident, and multi-location drill that we participated in February 2008 in San Diego county CA which simulated a bomb explosion inside the Coors amphitheater followed by gas spill. The two locations represent two different cities in the exercise (Chula Vista and National City). Different organizations participated to practice and to improve their speed of response, collaboration, and resource allocation within a unified command. San Diego city, county and several other neighboring cities sent Police, Sheriff, fire department, *Emergency Medical Service (EMS)* paramedics from multiple agencies such as HAZMAT, Special Forces (bomb squads and SWAT), FBI, and the Metropolitan Medical Response Service. There was also a public safety officer in each drill. An *Area Command (AC)* was activated to oversee the management of multiple incidents that was each managed by a separate *incident command (IC)* center B.4. There were about 100 victims and patients consisting of volunteers who received instructions on their mock injuries and what they needed to do in advance. The practice started with an explosion followed by an emergency phone call. A few local police officers who were present at the site for general safety during the amphitheater event, attended the contaminated hot zone to help with very primary safety goals. This was mainly to secure the victims from possible second incident which was the gas spill. The police officers were contaminated after entering the hot zone and hence they could not leave the hot zone. However, they were in touch with dispatchers to update them with the information at the site. It took approximately over an hour until

HAZMAT and Bomb Squad arrived at the site, and took even longer until they entered the hot zone. Several factors contributed to this delay: normal city traffic and distance to get to the site from their base locations. Additionally, as squads entered the site, they could not hurry into the contaminated zone. They needed to set up the Decon facility first to make contaminated people pass through before leaving the site, and put on the special sealed-out uniforms. HAZMAT was also provided with the same protective uniforms but they were instructed to follow Special Forces to enter. After they entered the site and cleared the area from chemicals and explosives, they declared the area safe and that was when the medical team entered the hot zone to attend the victims and injured people. Patients were equipped with wireless devices to transfer their vital signs to the remote center [16]. The main activities were to reach the overall objective to develop *an effective communication, information sharing and information management* system. Activities that were included in this exercise were:

- Arrival of different teams
- All responders attending the incident got debriefed on the shared communication plan
- Determine and announce organizational chart as to whom the team members should report)
- Declare the area safe
- Send victims through DECON process as needed
- Initiate triage and treatment efforts
- Resource status/request/release allocation and monitoring
- Conclude the exercise

There was a communication plan for the drill which was relayed to all teams at debriefing. Debriefing occurs at the beginning of an incident response as different team become present at the site and a communication plan is to be announced. The IC is the basic structure followed in any incident response in California A. In Silver

Bullet drill, it was scripted that the first responder to the Coors Amphitheater would request MMST as part of their assessment of the incident. MMST is generally assigned to supervise the progress of response towards meeting the overall objective and ensure that organizational territorial obstacles are not slowing down the effort. Two-way radios and cellular phones are among common devices used for communication.

For this drill, an *area command (AC)* was established to oversee the management of two incidents and to contribute to the efficiency of overall response. The AC chief was responsible for setting overall incident-related priorities and allocating critical resources accordingly. The AC is formed when multiple incidents are handled by separate *incident command systems (ICS)* or when the incident is spread geographically or extended over longer periods of time and needs special coordination for resource allocation. Each incident has its own IC center located near the hot zone of each crisis site. The main functions of each emergency center include coordinating and communicating resource allocation status, information collection and dissemination. The IC organization also needs to establish communications with the activated local *Emergency Operation Center (EOC)*. Major functional disciplines like fire departments or law enforcement usually comprise the *department operation center (DOC)*. A more detailed description of ICS, AC, EOC, DOC and MMST is presented in appendices A and B.

Figure 2.1 shows ICS organizational chart from [17] followed in the drills. This graph illustrates how different jurisdictions, as decoupled systems, collaborate and cooperate to meet the overall emergency response objectives. In this scenario, the area commander supervises the overall emergency response progress by communicating with MMST task force leaders of each site to meet the response objective. Incident management should collect information and share it on a need-to-know basis for efficient communication to conserve on resources and avoid confusion.

2.5.2 Lessons Learned

We interviewed several first responders including MMST task force leader, the area commander, EMS supervisor for Coors amphitheater and medical group supervisor, before and after the incident. There was several instances in which *operations* was not

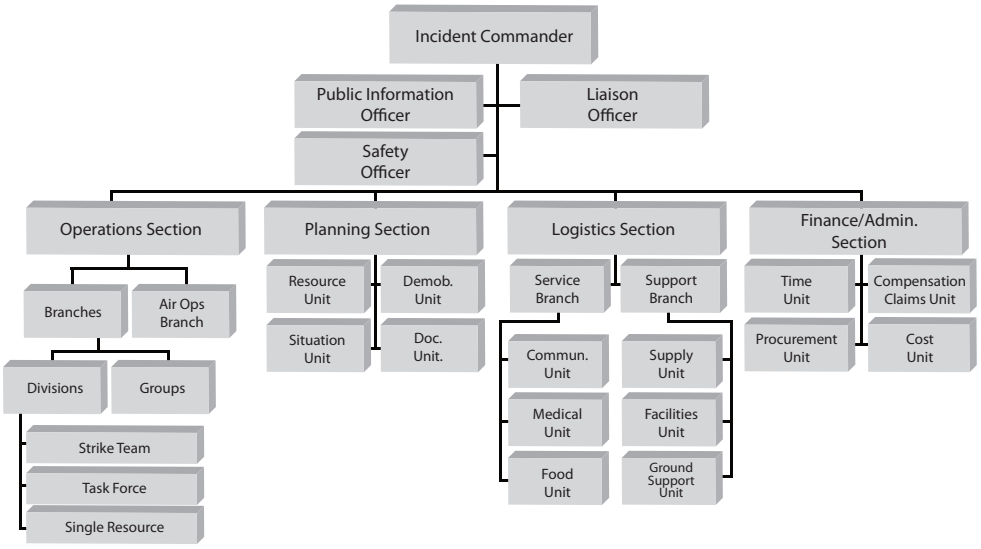

Figure 2.1: Incident Command System from Wikipedia

reachable. This will be addressed in section 7.3. We learned about the organizational structure of the drill and expected problems and communication unreachability issues that occurred. A summary of key roles:

The role of the medical group supervisor was to oversee the medical operation within MMST division at Coors. The MMST became a "single resource" to the overall incident commander, and fit into ICS structure under the incident's "Operations Section". Considering the scale of the incident, the MMST task force leader (TFL) arrived at the site after all first responders were debriefed by the incident commander. The MMST TFL provided the IC with an overview of what MMST offers, and integrated into ICS organizational structure by taking over the supervision of all specialized resources including HAZMAT, SWAT, EOD (Explosive Ordnance Disposal), and MMST Medical Team who were engaged in the response within the designated "Hot Zone". This simplified the reachability and communication problem for the incident commander as they could treat all those diverse resources as a single resource referred to as MMST. The MMST point of contact to obtain information from hot zone is through the "MMST Operations", not the incident Operations. The MMST Operations do not need to continuously monitor communications, only when it does not detract from what an

MMST supervisor (TFL) is engaged in.

The following problems have been identified in this exercise:

- There were a few people who preferred face-to-face communication as opposed to communication over physical network. This may lead to longer delays to reach the person, forgetting a message, missing a designated receiver, unnecessary repeats, and being interrupted by people or radios. Nonetheless, face-to-face communication is advantageous as the information is not abbreviated and there is no ambiguity about the content.

- There were delays that were inevitable but there were cases where waiting for a command or resource caused a deadlock.

- Problem of unclear communication when organizational communication protocols were not relayed clearly to all parties involved which impeded life saving activities. For example, there were instructions for agency A to not enter a building until they received direct command from organization B. For various reasons, organization B was not aware that they had to grant permission to group A and conflicts arose when group A entered the building. In some cases, this has led to people being arrested by mistake and the response and triage activities have been slowed down.

- Unreachability of a few first response supervisors unexpectedly.

- Lack of a public bulletin board to share information among all responders. There was a web-based software (webEoC) available to a very limited number of responders who shared a downloaded map of the site taken by a helicopter.

- Loss of information flow as a new task force leader was assigned to the second site and the connection with the first incident site got lost. The information did not flow properly to the second site.

- Deadlock caused by occasional independent actions or interaction procedures.

2.6 San Diego Rock ‘n’ Roll Marathon

2.6.1 Description

Cool, overcast weather drastically reduced the number of casualties at the 12th annual Rock ‘n’ Roll Marathon in San Diego, held on May 31, 2009. With over 18,000 participants and 13,000 finishers, it still made an excellent testbed for studying the flow of medical incidents and response.

In previous years at the Rock ‘n’ Roll Marathon, there were large numbers of casualties including runners suffering from heat stroke, seizures, low sodium (water intoxication) and altered levels of unconsciousness. However, this year - thanks to the cool weather - there were only 14 transfers to area hospitals and no one was labeled as “acute”. Over the 12 year history of the marathon there have been numerous improvements in the coordination of the Medical Command Center. Currently, the San Diego Fire Communications Center (FCC) deploys a satellite team of emergency medical dispatchers to a tent located near the end of the finish line on the grounds of the San Diego Marine Corps Recruit Depot (MCRD). Four FCC dispatch personnel (whose job is normally to screen and deploy fire and EMS resources throughout the city of San Diego), establish an entirely separate communications command center dedicated to race-related needs.

There are over $15,000$ runners each year and the potential to overwhelm the day-to-day resources of the city EMS system is real and potentially very dangerous. Therefore, San Diego Fire-Rescue utilizes this stand-alone command center so that they do not interfere with normal city call traffic. Dispatchers employ laptop computers running the ‘live’ FCC *Computer-Aided Dispatch (CAD)*. All City EMS and fire activity is displayed in real-time on a large screen inside the tent. In addition, there is a specific color allocated to ambulances and fire engines dedicated to the race. The race event coordinator is Elite Racing, a San Diego-based organization that operates numerous Rock ‘n’ Roll marathons around the country. A marathon medical director is hired to supervise Elite personnel but the responsibility for the overall organization and management of the system rests with the San Diego Fire-Rescue-Department, and many fire personnel are collocated at the medical command center, including deputy and battalion chiefs.

An efficient mechanism has been created to respond to runners' medical needs on the race course. At each of 26 medical aid stations (one for each mile along the race track), a medical doctor is equipped with a cellphone. There are also several spotters distributed along the course. Any of these individuals may phone an Elite call taker co-located in the Medical Command Center to report a runner 'down' and to provide basic medical information. These individual spotters are told to identify the runner's bib number along with nearby geographic landmarks to aid crews in locating the incident site among the large crowds. This Elite call-taker sits immediately behind FCC dispatch personnel who operates the CAD. They receive the information and pass it to a fire dispatcher, on a piece of paper, who triages the call's acuity and assigns the most appropriate resource.

Both basic and advanced life support ambulances are deployed on the course: *Basic life support (BLS)* ambulances are staffed by EMS's who can provide first aid and essential airway support, while *advanced life support (ALS)*, paramedic-staffed, units can provide definitive airway care and medications. On race day, the San Diego EMS system deploys additional BLS and ALS resources to support the needs of the city given the potential for large numbers of casualties if the weather is hot. Depending on the perceived acuity of the call, FCC dispatchers (using Medical Priority Dispatch System algorithms) assign one of four resources:

Level 1 (simultaneous fire engine with a paramedic) plus an ALS ambulance 'lights-and-sirens'

Level 2: ALS ambulance, 'lights-and-sirens'

Level 3: ALS ambulance, no 'lights-and-sirens'

Level 4: BLS ambulance, no 'lights-and-sirens'

Fire dispatchers assign the resources using the CAD and communicate basic information to the responding crews. Callers are told not to hang up until told to do so by dispatchers to ensure the receipt of complete information. Elite call-takers may direct questions to the marathon medical director as needed.

This communication protocol is efficient for FCC dispatchers as it frees them from becoming involved in information collection. The physical proximity between Elite call-takers and fire dispatchers added to the efficiency. Old-fashioned pen and paper works well and quickly. Elite dispatchers do not have a computer in front of them,

just a basic land line phone. Before runners leave the marathon finish line area, they are informed that this is their last chance to receive any medical attention. Upon their consent to leave the area, they are responsible for their own well-being. This year, only 14 of approximately 18,000 runners were transported to area hospitals. This report was published in [18].

2.6.2 Enhanced 9-1-1 Call Handling in San Diego

Outside the race, typical 9-1-1 calls continue to be managed in the normal way. Land line 9-1-1 calls in the city of San Diego are directed to its Public Safety Answering Point (PSAP) at the San Diego Police Department (SDPD) call-center located downtown at police headquarters. Police call-takers initially sort the incoming calls for police, fire and ambulance need. Should fire or EMS resources be needed, the call is automatically forwarded the city's Secondary Safety Answering Point (SSAP) at the Fire Communications Center near Montgomery Field airport. There are two types of 9-1-1 systems, regular 9-1-1 and "enhanced" 9-1-1. The city is equipped with enhanced 9-1-1, which automatically identifies the requesting party's phone number and address. This is referred to ANI-ALI, for "Automatic Number Identification-Automatic Location Information." Emergency dispatchers are required to ask for the caller's location and the nature of problem to verify the address of the actual emergency. Unfortunately, the SDPD and FCC CADs are not linked to each other for easy data sharing; this remains a goal of an upgraded communications system for the city.

Currently in the San Diego region, 9-1-1 calls from cellphones are forwarded to a different PSAP, located at the California Highway Patrol (CHP). Precise statistics are not available on how long it takes to receive and process CHP calls, although delays are common. One reason is the large numbers of cellphone calls which are redundant, reporting the same incident. No call can be ignored; therefore multiple calls increase the waiting time in the "queue" for other callers. Therefore, it is crucial to use a land line phone, when possible, to call 9-1-1 in a life threatening situation. As cellphone towers and global position systems are enhanced, it is expected that individual cellphone calls will be able to be routed directly to the most appropriate PSAP. Many thanks go to Dr. James Dunford, Professor of Clinical Medicine and Surgery, UCSD Department of

Emergency Medicine and City of San Diego Medical Director (EMS), for his insights, conversations and review of this article, and to Maureen C. Curran for editing it for website publication.

2.6.3 Lessons Learned

The physical proximity of great medical expertise and higher authority in the same control and command tent was a plus as the presence of the city of San Diego EMS medical director helped with resource allocation and life-and-death-related decision making. However distribution of local controllers at the site, helped with removing the bottleneck where everyone may need to talk to one supervisor or in the case of a broken or busy communication line. Secure access to the data among different responders including fire departments and hospitals is possible to achieve over a wireless network such as the mesh network explained in chapter 3. This is an ad hoc network of small, lightweight, and easily reconfigurable nodes that quickly self-configure to form a reliable wireless mesh network locally. The local network can connect to Internet utilizing satellites, wireless, or other available technologies to provide connectivity with decision makers off-site.

2.7 UCSD Emergency Medical Department

We interviewed nurses and doctors at UCSD medical center's emergency department at Hillcrest at several occasions to learn the process of patients' admission, treatment procedure and resource allocation from a network perspective. The process starts as the phone operator for emergency department receives a phone call and if a patient is admitted to the hospital as the result of this call, he or she arrives in one of three ways: walk-in, on ambulance, or by a taxi. Patients first wait in a queue to be seen by a triage nurse to get the vital signs checked and depending on the severity of the case, join different queues (more urgent versus medium or none-emergency). At this stage, the triage nurse routes the patient to the appropriate queue. Depending on the patients' priority and resources available, they get scheduled to receive the treatment which may include one or all of the following: lab resources, nurses, beds and doctors.

The process is rather efficient because of a well-taught intra-organization communication protocol. If the hospital is already short of resources, new patients are not admitted to the system so the arrival rate is controlled. In that case, they are re-routed to another medical center. They have also developed a web-application that has color-coded flags so the patients' status is updated on a web-based application on a computer which is accessible to all nurses, labs, and doctors at the same time. A Petri net modeling the general patients' admission and treatment has been developed and illustrated in figure 6.17.

2.7.1 Lessons Learned

The observations at UCSD emergency medical department and several other occasions from interviewing individual organizations, have confirmed that communication is well-conducted and followed within each organization. Over time and by experience they have developed mechanisms to communicate effectively and allocate resources and response more efficiently within their own organization. However, problems arise as the communication and interactions occur *among different organizations*. Several factors contribute to communication failure, including lack of overall system knowledge, limited access to information, and delay. Additionally, inter-organizational communication protocols can create potential bottlenecks and slow down the process unnecessarily. Territorial obstacles can also slow down the interactions as well, as some organizations have restricted policies to share information and communicate and collaborate with other organizations.

2.8 After-incident Reports

One of the after-incident reports we investigated and analyzed, concerned a fire rescue and response incident where two fire personnel including the fire captain and the fire engineer were killed [19]. A fire started in a residence and an alarm was sent to the alarm company. The alarm company called the residence to confirm the fire and they asked if the residents wanted them to contact the emergency number on their behalf. The alarm company agent called the emergency department on a non-emergency number.

They forgot to indicate that they had already confirmed the fire and that the residents asked them to call emergency on their behalf. The call on a non-emergency line was placed on hold as there was another call on the emergency line. Another dispatcher attended the call on the hold later, but the alarm company agent failed to provide precise and complete information again. There was more delay as the line was busy when the dispatcher tried to call the resident number. Eventually the fire department sent a truck. After about 15 minutes delay (or more), the resident called to report a fire but the first fire truck was already on its way. Another fire truck got released at this point.

During this incident, the communication was not was not clearly understood by the designated group as it was meant to. Radio transmission was directed to the engine $E74$ while second arriving unit was $E73$. The source of failure was identified as broken communication and lack of personnel accountability. In this incident, a breakdown in the transfer of command led to a response without proper coordination which initiated an independent action. According to the report, part of the message was not copied precisely and therefore it initiated an independent action (not coordinated).

In another report, we learned about a shooting incident at a school in California in an interview. HAZMAT was not supposed to enter the building until they had permission from Special Forces for protection. However, HAZMAT ignored the hierarchy and entered the building as they had the same protective uniforms and in fact they were able to save lives.

2.8.1 Lessons Learned

Lack of clear communication, lack of personnel accountability, broken line of communication, and fail to transfer information precisely have impeded emergency response.

2.9 Technical and Social Challenges of Communication

There are several technical, social, and organizational challenges that responders face during an emergency response. Below we present a set of challenges that we noticed or learned about during our involvement at several drills, as well as results from

interviewing team leaders at first response organizations, before or after the response and investigating after-incident reports.

- Inter-organizational communication:
  Technical interoperability in a heterogeneous environment is crucial for communication and coordination among different agencies where different technologies are used [20]. Interoperability is not only limited to technology, but also it concerns the protocols that each individual organization follows internally and across organizations, which become more of an issue at an international level or large scale scenarios. The horizontal communication with peers or the hierarchical vertical communication across different levels in the organizational chart is the source of communication failure as a person is unreachable or when communication protocol has not been clearly communicated. Intra-organizational communication between a team leader and team members often works well.

- Resource allocation:
  The term "resource" refers to network resources such as bandwidth, human resources and apparatus. The information on the status of resource requests, allocation and release must be available and updated in real-time to enable accurate decisions and efficient response.

- Choice of backhaul:
  To establish communication between disaster sites and the outside world, the infrastructure needs to get connected to Internet through a backhaul link. The choice of backhaul depends on available technologies and the amount of bandwidth or data transfer required.

- Network reconfiguration:
  In an emergency response, there is often a strong need to move nodes and reconfigure the topology. If the signal strength through a relaying node falls below a certain threshold, the network establishes a new connection seamlessly through a different path providing better signal strength. Additionally, in a multi-hop ad hoc network, it is preferred to be connected through the paths that are not more than

two hops away from the gateway. This is to keep the signal strength above a satisfactory threshold and ensure acceptable throughput. This concept can be extended to the human communication as new team members emerge and collaboration and communication has to be adjusted on-the-fly.

- Network reliability and message integrity:
  It is highly important to repeat or send information or messages intact throughout the network. Failure to transmit a message precisely has been the cause of broken communication in past.

- Traffic management and quality of service:
  Network management can allow the network to provide different services for different priority tasks depending on resource availabilities. Different traffic types can be scheduled with respect to the constraints, quality of service requirements and availability of the resources.

- Limited support for encryption or mobility:
  An emergency response depending on the type of incident may require different levels of encryption or information privacy. The communication infrastructure also needs to support mobility for the users when it is needed. The network should support mobility as users or access nodes move around the site.

- Minimized dependencies when applicable:
  Emergency response is an interdependent process by nature. It is desired to reduce dependencies in technical specifications by planning a back up infrastructure like power supply [21] and reaching at functional cohesion with minimum relational dependencies as will be explained in section 4.3. This eliminates unnecessary dependencies to protect the system performance and achieve the overall objective in case of an individual's failure. We do not want to turn one organization into a bottleneck and a single point of failure in case of unpredictable disconnects in communication. In sequential relationship, section 4.3, the following event is dependent on the completion of the preceding event.

- Scale and nature of disaster:

Decision making and resource allocation in a disaster response varies from one incident to another depending on the scale and the nature of the disaster. The degree of urbanization or the geographic spread may require different actions for a specific response. When the incident(s) is geographically dispersed, it requires a lot of coordination [11].

- Training and Financial obstacles:
  There is a natural resistance to try out new technology while an old system still works. This might be attributed to the cost of replacing or upgrading the existing technology, and the cost of training people to learn how to work with new technology [22]. They need to be convinced of benefits of new technology to be willing to spend money on the deployment and training. Lack of knowledge in operating new technology is another issue [23]. New technology needs to be installed and fully tested before final field deployment. Decision makers must select systems that fulfill public safety operational requirements while also considering all possible natural hazards. Some levels of redundancy is recommended as back-up to handle emergency situations more efficiently and quickly. However, there is always a trade-off between cost and redundancy. This trade-off point varies among different scenarios depending on the application, the probability and risk of failure, and the degree and scale of possible outcomes.

- Cultural barriers:
  Different organizations may resist deploying new technologies for different traditional and cultural reasons [22]. Each organization has pride in their departments and there is a tendency to be self-sufficient rather than to cooperate with others. This slows down the effectiveness of a response as it creates delay in sharing resources and coordination which are vital to an efficient response. To address this, a higher authority (like the MMST Task Force Leader) is responsible to set the overall objective during an incident response and ensure that all organizations work together to achieve that objective for an efficient response.

- Precise personnel accountability:
  Another practical issue in an emergency response scenario is obtaining account-

ability of personnel attending the response. Responders arrive at different times and before the mobile check-in staff can reach to them, the special protection force might have the sealed-out uniform on which does not allow scanning of the responder's magnetic badge. When a precise list of attendees is not available, communication is seriously affected. It is difficult to carry out a well-planned and well-organized response if there is not precise knowledge of the list of personnel attending the crisis site. We believe that an electronic check-in and check-out system would assist with the problem so that every one entering the site or leaving the system has to check in or out prior to attendance and departure. This solution is similar to the accountability check-in and check-out process in use by cruise lines. We suggest that a unique magnetic ID card get activated and distributed in an emergency; this ID would be linked to personal badges which can be retrieved upon departure and demagnetized to be re-used. The only difference is that on cruise ships, the card is not retrieved upon final departure but it will not be of any use as the date on it expires. With demagnetizing the same cards and redistributing them, the cost is a one-time expense which makes it a good candidate to consider.

2.10 Scope of Work

Considering the large scope of all possible communication problems with application to emergency response, a summary of the problems that are presented with a solution within the content of this work is:

- Accountability: solution presented in section 2.9
- Physical proximity and face-to-face communication presented in section 2.6.3
- Network reconfiguration: wireless mesh network, chapter 3
- Minimize dependencies: sections 4.3 and 6.3.3
- Scalability: in wireless mesh network infrastructure, chapter 3 and the structure of the model, chapter 6

- Breaking down the complexity of a large system: hierarchical supervisory solutions, chapter 8
- Predict and prevent deadlocks, chapter 6
- Impact of a supervisor or local controller with higher authority and knowledge, chapter 8

2.11 Concluding Remarks

We deployed many technologies including a communication infrastructure to provide connectivity at the disaster site, both locally and to the Internet. The test bed described in chapter 3 has been used to collect a large amount of data for post-analysis to improve network performance, enhance routing algorithms and several other issues by several researchers. During the *Operation College freedom* exercise, we captured network traffic and measured the amount of network throughput, application data, number of retransmissions and several other statistics. The feedback obtained from the analysis was relayed to the researchers in charge of developing next generation routing algorithms for wireless mesh access nodes.

During *Silver Bullet Drill* and *Rock 'n' Roll* exercises, we observed communication protocols at organizational level. To obtain further information, we also participated and conducted several interviews with first responders. Additionally, to clear other questions raised after the exercise, a questionnaire was submitted to the first responders who were willing and available to help. Real-life scenarios reflecting communication problems from networking perspective are based on the activities and observations gathered in this chapter and are presented in chapter 4.

The author would like to acknowledge Richard Leap for providing information in response to questionnaire submitted to him after the exercise. The author would like to also acknowledge Art Botterell, Community Warning System Manager for the office of the Sheriff of Contra Costa County, California, for providing us with a copy of the report.

This chapter, in part, is a reprint of the material of the following papers: R.B.

Dilmaghani, and R.R. Rao,“Hybrid Communication Infrastructure and Social Implications for Disaster Management,” Hawaii International Conference on System Sciences (HICSS), January 2007 *, and* R.B. Dilmaghani, and R.R. Rao, “Hybrid Wireless Mesh Network with Application to Emergency Scenarios,” Journal of Software Engineer, Volume 3, number 2, Feb. 2008 *, and* R.B. Dilmaghani, and R.R. Rao “Ph.D. Student Observes Medical Command Center at Rock ‘n’ Roll Marathon,”, 2009.

Chapter 3

Communication Infrastructure and Performance Evaluation

3.1 Introduction

Wireless Mesh Networks (WMNs) have become very popular over past couple of years for the use of unlicensed spectrum and low cost of IEEE 802.11b/a/g-based off-the-shelf devices. Wireless mesh networks have been used as a solution to extend wireless coverage in many cities [24] [25] [26] [27] [28] [29]. In section 3.2 we explain why a wireless mesh infrastructure is an appropriate candidate for the application. In section 3.3 we present the wireless mesh test bed which has been deployed in several emergency response exercises. In section section 3.4 we present data analysis, performance evaluation, and lessons learned from deploying the wireless mesh test bed. Finally, in section 3.5 we highlight the concluding remarks.

3.2 Choice of Wireless Mesh Network

A communication infrastructure within the context of emergency applications should be reliable, robust, easily configurable, quickly deployable at low cost and interoperable in a heterogeneous environment with minimum interdependencies. When different organizations use different radios operating on different frequencies, interop-

erability is the main common problem. There has been a better use of existing technologies to provide interoperability by establishing patches throughout the network [6].

Wireless mesh infrastructure is a promising candidate for emergency response applications as well as many other applications mainly for their quickly deployable and easily re-configurable features. In a mesh architecture, nodes form a local network when there is a line of sight and only gateway(s) needs to be connected through an available backhaul technology to the Internet. This is advantageous as fewer nodes need to be configured. The infrastructure consists of three-tiers considering the connectivity of the local mesh network to Internet [30]: the first tier consists of clients using different devices and technologies such as *Personal Digital Assistant (PDA)*, laptop and iTAGs [16]. ITAGS are small handhold wireless devices that communicate and transfer patients' medical status, treatment record, and vital information to medical personnel at the site of the disaster. The second tier consists of wireless mesh nodes and the third tier is the backhaul link(s) provides connectivity to Internet. Figure 3.1 shows the general architecture of wireless mesh network deployed at several drills.

The mesh architecture is resilient to the failure of nodes or links as there are alternate paths to take if a link or an access point fails. In case of a failure, a node re-establishes communication through other nodes when a neighboring node fails. This characteristic improves reliability, as unavailability or failure of sub-components of the system does not affect the overall performance of the system and the service will be continuously available. This architecture is robust in the sense that it is able to operate in a heterogeneous environment with a variety of technologies. Additional wireless access nodes can join or leave the network transparently without causing a service interruption by finding the closest node with best signal strength. Finally, at a disaster site, when nodes need to be moved for evacuation or better connectivity, reconfiguration is trivial since the wireless access nodes will automatically re-configure the network. These wireless access nodes allow users to communicate with each other when there is no wired configuration. Each access node has two interface cards to provide interoperability and redundancy through the use of different channels for communication. Users using different technology such as CDMA or 802.11 can connect and talk to each other through these nodes. VoIP can be deployed over wireless mesh network to provide an oppor-

tunity to call other responders regardless of their location. This allows responders to utilize their human and network resources efficiently with a smaller connection waiting time suitable for short-term deployments such as crisis cite or at rural areas.

3.3 Wireless Mesh Test bed

The measurements taken over the test bed during the drills have allowed us to capture real network traffic to identify sources of bottleneck. The wireless mesh nodes in our test bed use Soekris net4521 board running on modified Linux distribution with Linux 2618 kernel. The operating range of each access node is about $300m$ outdoors. It provides connectivity throughout the disaster site and to the outside world through Internet. *Evolution Data Only (EVDO)* was deployed as backhaul to connect the crisis site to Internet to share information with remote command center. CalMesh access nodes with multiple interface cards, provided the possibility of different users getting connected uniformly to exchange data regardless of their individual technology as shown in figure 3.2 [13]. There are sensors installed on a remote controlled toy truck which can be sent to the hot zone where people cannot yet enter for safety or security reasons [31]. It provides broadband coverage and situational awareness via its video and audio outputs. This device is equipped with a *Global Positioning System (GPS)* module that allows the device to record location-based measurements. The integrated environmental sensors allow the mobile truck to collect data on gas sensor and send measurements via *Short Message Service(SMS)* to a designated cell phone. Here are the steps taken to analyze communication at the network level:

- Deploy a wireless mesh infrastructure to provide reliable communication locally and to the outside world
- Analyze network performance and identify sources of bottleneck
- Establish interoperability in a heterogeneous environment where various devices

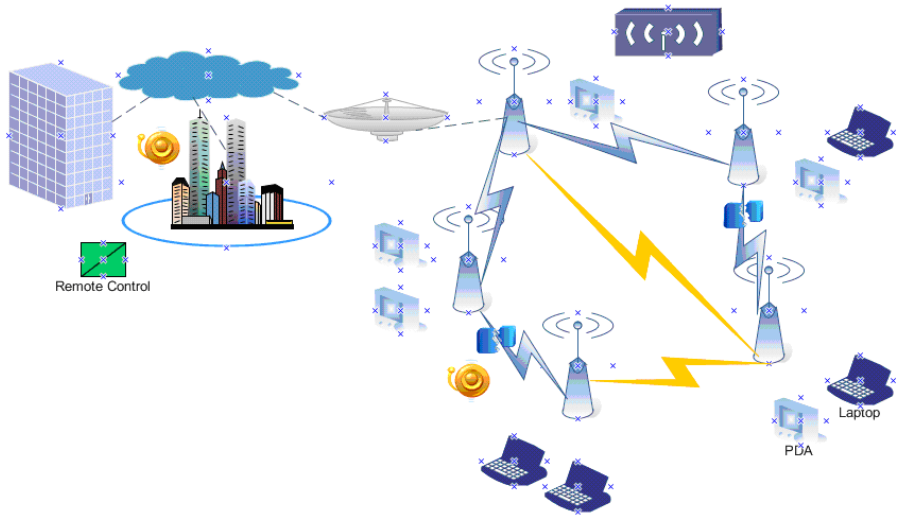

Figure 3.1: A Wireless Mesh Infrastructure

Figure 3.2: CalMesh Wireless Access Node

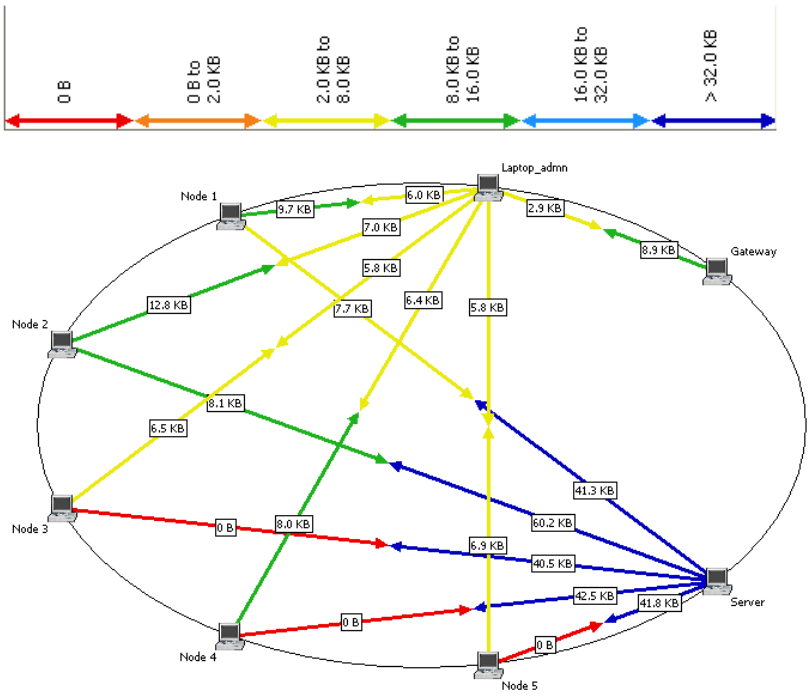

Figure 3.3: Network Data (Bytes) (Directional)

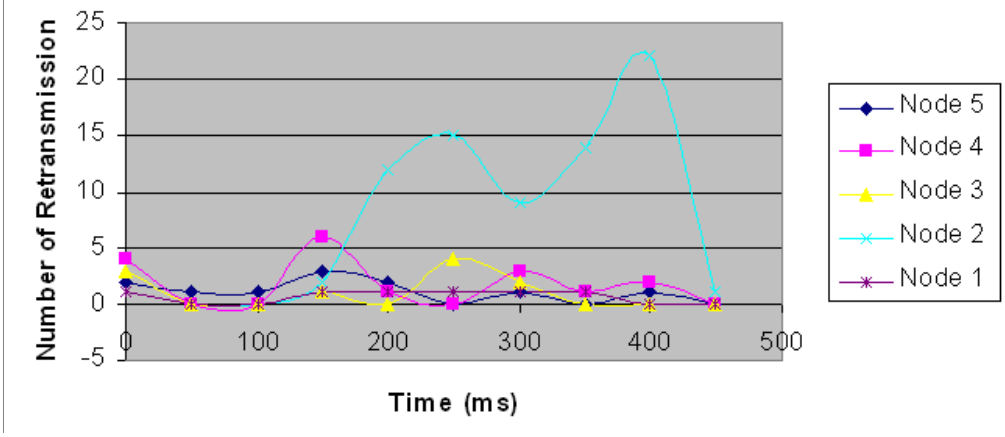

Figure 3.4: Retransmission from server to the mesh access nodes

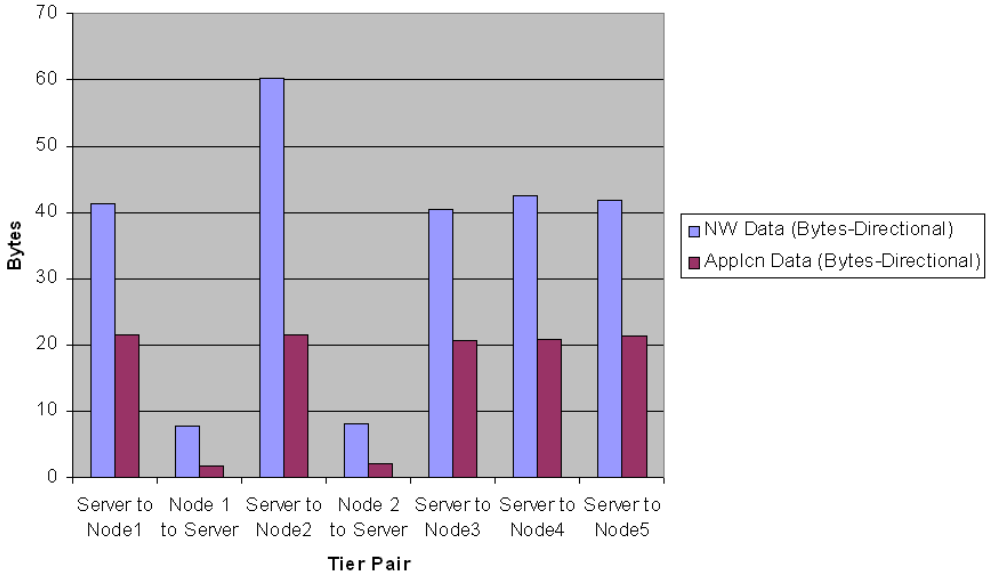

Figure 3.5: Application and Network Data from server to the mesh nodes (Bytes-Directional)

utilize different communication technologies

3.4 Data Analysis

Real measurements obtained directly over the test bed are studied for analysis and network performance evaluation. Figure 3.3 illustrates the amount of network Data in bytes in a directional graph. The measurements show the aspects of the network that need to be improved and identifies sources of bottleneck in the network. Below we present sample results obtained from the drills presented in section 2.4. For this set of data, the total response time is 456 seconds. 137.6 KB of application data and 328.7KB of network data were transferred. Table 3.1 shows the amount of application data and network throughput between the server and mesh nodes. The difference between these two values shows the amount of protocol overhead in the network. Figure 3.5 shows the total amount of data and network data between server and wireless mesh access nodes. The large number of small packets traveling over the network caused by the protocol chattiness is the main source of bottleneck in this set of data [15] [10]. TCP window-

Table 3.1: Data (bytes-directional) from Server to Mesh Access Nodes

Server to nodes	Network data	Application data
Node 1	41.3	21.6
Node 2	60.2	21.7
Node 3	40.5	20.7
Node 4	42.5	20.9
Node 5	41.8	21.4

ing and Nagle's algorithm are not the main source of bottleneck. Processing delay at the

Table 3.2: Total Number of Retransmission from Server to Mesh Access Nodes

Time (sec)	Node 1	Node 2	Node 3	Node 4	Node 5
0	1	1	3	4	2
50	0	0	0	0	1
100	0	0	0	0	1
150	1	2	1	6	3
200	1	12	0	1	2
250	1	15	4	0	0
300	1	9	2	3	1
350	1	14	0	1	0
400	0	22	0	2	1
450	0	1	0	0	0

server is a source of delay which accounts for approximately 68.7% of total delay [7]. Sending large number of small packets cause potential bottleneck in the network which can be optimized by sending fewer larger packets instead. This will reduce the number of request and respond set of messages. Application throughput is approximately the same between server and all mesh nodes as they run similar applications. Network throughput includes all application data and network protocol overhead which is similar for all mesh nodes except node 2. Node 2 has a large number of retransmission which is due to existence of error-prone links in this scenario, where network congestion was not an issue, and as a result this node's throughput varies drastically from other nodes. Table 3.2 shows the total number of retransmission for each node and figure 3.4 illustrates the same statistics. Node 2 has a large number of *application turns* indicating the number of times the direction of communication changes between source and destina-

tion. Satellite link, EVDO, wireless and wired have been the main choice of backhaul links. In *Operation College Freedom* drill, the local wireless mesh network was connected through wired link to Internet while in downtown San Diego, satellite provided the connection and EVDO at Silver Bullet drill. At Silver Bullet Drill, the infrastructure deployed at the site consisted of two independent networks at each site. Considering the small geographic spread of each incident site, it was sufficient to deploy a small network. Each site had a network consisting of one gateway and three access nodes where gateways were connected to Internet via EVDO technology. The default choice for CalMesh boxes is EVDO where it is not possible to connect through Ethernet or WiFi for faster speed or higher bandwidth. We have observed that the choice of EVDO decrease the network throughput as it is sharing the bandwidth with cellular phones and other devices deploying EVDO technology.

3.5 Concluding Remarks

We experienced interference among different devices employing various communication technologies which does not quite agree with theory as they operate in different frequency. Additional practical observation was the effect of network getting partitioned due to physical obstructions. This was experienced when a truck blocked the line of sight between two access nodes. To overcome this problem, the nodes were moved around to re-establish the connection where there was a good line of sight. During Silver Bullet Drill, there was a need to uniquely label GPS equipped devices as they show up on different monitoring system belonging to different organizations which made it challenging to identify and distinguish the function of each device.

In this chapter we also identified that large amount of control data was the main source of bottleneck.

We realized that a better communication technology deployment does not improve emergency response alone. This is not just about reducing the response time, although the time and a quick response is of crucial importance for the serious life threatening situations and economy impacts. It is impossible to reach at one efficient response solution with a reliable network without revising collaboration and commu-

nication protocols among organizations. We have chosen high-level Petri nets as an appropriate event-driven mathematical modeling tool to examine inter-organizational protocols and develop a supervisory control model at crisis site.

This chapter, in part, is a reprint of the material of the following papers: R.B. Dilmaghani, B.S. Manoj, B. Jafarian, and R.R. Rao, "Performance evaluation of RescueMesh: a metro-scale hybrid wireless network," Proceedings of WiMesh Workshop, IEEE Workshop on Wireless Mesh Networks, held in conjunction with SECON, Santa Clara, 2005 , R.B. Dilmaghani, and R.R. Rao,"Hybrid Communication Infrastructure and Social Implications for Disaster Management," Hawaii International Conference on System Sciences (HICSS), January 2007 *, and* R.B. Dilmaghani, and R.R. Rao, "Hybrid Wireless Mesh Network with Application to Emergency Scenarios," Journal of Software Engineer, Volume 3, number 2, Feb. 2008.

Chapter 4

Perspectives on Modeling Communication in Emergency Response

4.1 Introduction

Given the most effective and robust communication technology, do we have the right communication protocol at the organizational level? Does a more suitable communication infrastructure at network level solve the communication problem?

Such open questions identified the need to widen our perspective on communication, collaboration and information sharing concepts. There was no doubt that a wide range of organizational problems surrounding the emergency response scenarios impeded the collaboration at different stages. To the best of our knowledge, currently there is not a well-defined systematic methodology to model, predict and analyze communication at organizational level. The lack of a systematic methodology motivated our research and its unique contribution to provide a systematic foundation to assess communication problem from both network and organizational perspectives.

Inter-organizational interactions require sharing information of various types according to the organizational communication protocols. Different organizations participate in crisis response with a variety of responsibilities and special skills. There is not

a single solution to fit all scenarios, instead methodological solutions can be mapped to common problems based on real scenarios with the advantage of being reused, and in expediting the process of decision making capturing past experiences. In section 4.2 we present our perspectives on several real-life scenarios in emergency response communication. The current issues have been identified and solutions have been proposed which inspired the models developed and presented in this work. In section 4.3 we present the type of inter-organizational dependency that suits emergency response applications and finally, in section 4.4 we highlight the concluding remarks.

4.2 Perspectives on Emergency Response Scenarios

Scenario 4.1 *Description: Silver Bullet drill was an example of multiple incident scenarios. The information was supposed to follow the organizational chart which did not occur when* operations *was not reachable for the second incident. A few authorities preferred face-to-face communication which caused delay and failure in precise delivery of all detail information. In Operations College freedom drill on campus, network partitioning was source of failure which occurred when a truck blocked line of sight. This in turn led to communication failure and interruption in flow of information. Communication protocol at organizational level was not communicated clearly to the team leaders and therefore there was confusion and delay in entering the building.*

Issues identified: The inter-organizational communication works under normal condition when everything goes as planned. However in practice this has not been the case for several practical reasons. One scenario is raised when the primary incident evolves to a secondary incident unexpectedly which requires developing an on-the-fly communication plan and resource allocation. Another scenario is when the communication plan exists but it is not well-communicated or not uniformly executed among organizations.

Proposed infrastructure solutions: A wireless mesh infrastructure provides reliability and rapid reconfiguration for the physical network requirements when network is partitioned.

Scenario 4.2 *Description: When in an emergency response to an incident, there is a request for more resources, like Silver Bullet drill where a second incident happens asking for more resources, the request travels through the network to reach decision making authorities. There are limited resources available through logistics. An authority can reserve some resources to accommodate new requests, ask for more resources, or change priorities in allocating resources.*

Issues identified: The new incoming resource request receives service depending on its urgency.

Proposed supervisory solutions: For non-preemptive requests, a supervisor can wait to finish the execution of the current task before starting to serve the new request. A timer is set and upon expiration, to prevent infinite wait leading to deadlocks and the supervisor can enforce an alternative path to the new incoming request. For preemptive requests, a supervisor sends a token to cancel an already scheduled path immediately to serve the preemptive request first and resumes the one interrupted upon completion of the preemptive task.

Scenario 4.3 *Description: Ambiguity on the termination time of a process that is followed by another action, causes problems. In College Operations Freedom the entry team was supposed to declare the hot zone safe for the rest of the responders to enter the hot zone.*

Issues identified: Unclear termination of a process, when another task is dependent on precise termination of this process, can cause premature entry, delay the entry of responders or endanger responders' lives.

Proposed supervisory solutions: Correct and revise the structure of communication protocol to prevent premature entry or unclear termination time. This has been illustrated in the model of figure 6.12.

Scenario 4.4 *Description- An emergency response slows down when the data traveling over the network is delayed or lost.*

Issues identified: Waiting for an event dependent on the data that never arrives, causes deadlock.

Proposed supervisory solutions: To avoid deadlock, a supervisor can cancel a path that leads to the deadlock. A supervisor with the higher-level authorities and knowledge cancels a token and enforces liveness. Another process consumes the token so that the other transition cannot fire. This is illustrated in figure 6.8.

Scenario 4.5 *Description: Responders arrive at different times and before the mobile check-in staff gets to them, the special protection force might have their sealed-out uniform on which does not allow scanning of the magnetic badge. In this scenario, precise information on first responders attending the site will not be available and hence the communication is seriously affected. It is difficult to carry-on a well-planned and well-organized response if there is not precise knowledge of personnel attending the crisis site! This has been observed during Silver Bullet drill and fire incident report.*

Issues identified: Communication plan fails when there is no precise accountability of the personnel attending the response.

Proposed supervisory solution: A distributed architecture with several check-in locations throughout the cold zone can help. We propose an electronic check-in check-out system will assist with the problem so that every one entering the site or leaving it, has to go through the check in or out system prior to entrance or departure as explained in item 2.9.

Scenario 4.6 *Description: In Silver Bullet drill and at the fire incident report, the message was not completely transferred. Important updated information obtained by dispatchers were not relayed to responding unit in a timely manner. Disregarding organizational chart led to independent action and loss of coordination in the fire incident. At times, it was necessary to repeat the information to remove ambiguity caused by information being cut-off or lost, or organizational chart was discarded due to reachability problem or broken communication.*

Issues identified: Failure to transfer message precisely or causing bottleneck by turning a key person as a single point of communication. This is like the well-known send-receive packet loss over the network. Packets may get lost and therefore a packet sequence number and acknowledgment are used so that the sender knows whether the receiver has received the packet or not. If the acknowledgment is not received before the

time expires, the sender resends the packets. There are well-known protocols to resolve this problem in networking [32].

Proposed supervisory solution: Choose a strategically well-thought geographic location for the key person, operations or logistics, to be reachable at all times. Additionally team leaders are encouraged to stand together in the same location to facilitate horizontal communication (inter-organizational) as information will travel faster and team leaders will be able to reach each other. A supervisor must ensure that there is only one update at a time and that is received by all destinations. The supervisory Petri net solution is presented in figure 4.1 where a supervisory token circulates among members.

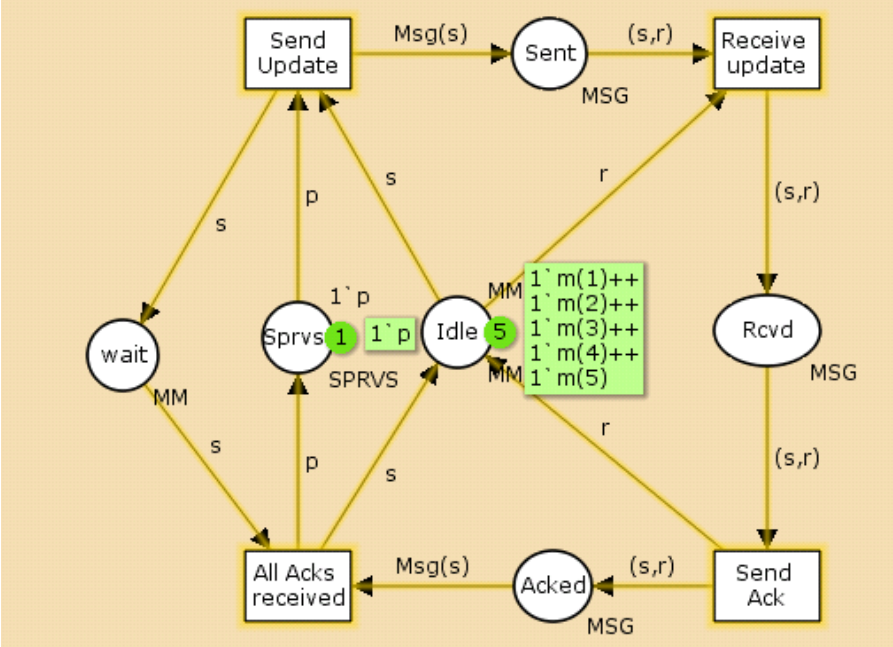

Figure 4.1: Petri net model corresponding to Scenario 4.6

Scenario 4.7 *Description: Emergency response by nature is an interdependent process. It is desired to reduce dependencies in technical specifications such as existence of a back up power source or battery for short-term deployments in case power infrastructure goes down. Similarly, minimizing dependencies in an inter-organizational chart is desired.*

Issues identified: Clearly it is preferred to eliminate unnecessary dependencies and interactions to protect the system performance and achieve the overall objective in case of individual's failure. However, at the same time, we do not want to turn one organization to a bottleneck and a single point of failure in case of unpredictable disconnection in communication. This can be interpreted as the functional cohesion corresponding with loose coupling as discussed in section 4.3.

Proposed supervisory solution: Minimize dependencies and try to achieve functional cohesion. Supervisory solutions can enforce liveness with forcing an alternate path or move the dependencies to the farthest possible point of communication as explained in section 6.3.3. If there is a need to synchronize tasks, for a soft synchronization, a supervisor allows each organization to proceed individually and at a later point get the acknowledgment from the slower organization to prevent unnecessary delay. However if this is a key task and the output depends on processing the information (a firm synchronization request), the supervisor enforces a new path if the person or resource is not reachable.

4.3 Inter-Organizational Dependencies

In any application where several entities communicate with each other, dependencies are created. It is desired to minimize dependencies so that in case of a component's failure, the rest of the system is impacted minimally. To investigate different types of interactions we use concept of cohesion in object oriented programming to fit this particular application [33]. It is important to decide on how the components are related within a system.

The relationship is presentable in one or more of the following ways [34]:

- Functional (best): when all organizations contribute to a single well-defined task
- Sequential: when all organizations work on an ordered list of data, i.e. the output of one is the input to another
- Communication: when all organizations work on the same data but not in any particular order

- Procedural: when all organizations follow a certain sequence of execution
- Temporal: when all organizations have a particular time during the execution
- Logical: when all organizations have different activities that logically categorized as doing the same activity
- Coincidental (worst): if none of the above holds true; organizations have a random relationship

We adapt functional cohesion as the best option representing dependencies among different groups in a system contributing to a single well-defined task (they work on one problem-related task) or one objective [35]. Figure 4.2 presents the flowchart based on the concept of cohesion represented in [33]. In system design and supervisory control

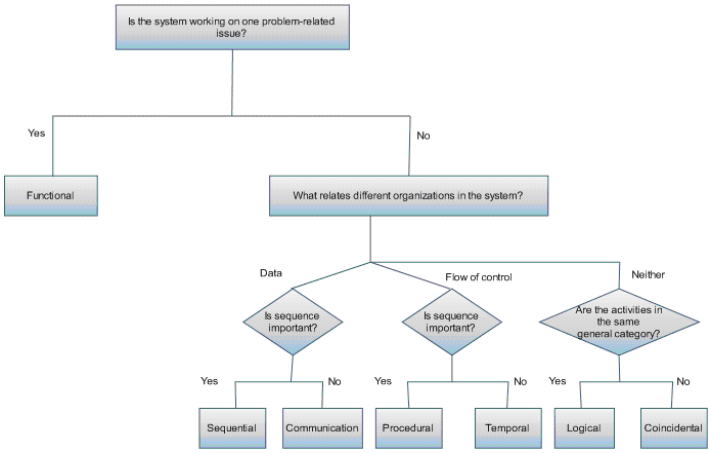

Figure 4.2: Dependencies in an Organizational Chart

solutions to emergency response application, minimum dependencies is the objective. This best correlates with loose coupling where minimum dependencies among different organizations allows them to work individually to achieve the overall objective.

Supervisory functions include tasks such as granting services or requests, interrupting or canceling an scheduled task, and introducing a new path will be discussed further in chapter 6 and chapter 8.

4.4 Concluding Remarks

During our study, we have observed and learned on several occasions that most of the times, vertical communication (intra-organizational) follows smoothly according to an organization chart. Team leaders are able to successfully communicate to their team members. This is attributed to the smaller scale of cooperation, within one organization, and the experience of people working with each other. There usually exists a well-established communication plan on what communication devices in what frequencies are in use within an organization. However problems arise when we consider horizontal communication among different organizations. This problem can be attributed to different factors such as lack of interoperability in a heterogeneous environment or unadaptable collaboration.

It is important to notice that the term "resource" has a wide concept in this application as it refers to humans, network resources such as bandwidth or storage capacity, and apparatus like fire trucks, radios, etc. There are limited resources at each level and communication planner, incident commander, and operations needs to be aware of the status of resources, limitations, and conflicts at all times in order to allocate resources efficiently and based on priorities.

In this chapter different scenario describing emergency response communications have been discussed. The networking issues have been identified and supervisory solutions were proposed which will be discussed in further details in the upcoming chapters.

The rest of this work presents an analysis and organizational communication enhancement techniques by:

- Developing event-driven models to investigate the structural properties of inter-organizational communication protocols
- Investigating behavioral properties to predict performance
- Proposing supervisory control solutions to improve inter-organizational communication

This chapter, in part, is a reprint of the material of the following papers: R.B.

Dilmaghani, and R.R. Rao,“Hybrid Communication Infrastructure and Social Implications for Disaster Management,” Hawaii International Conference on System Sciences (HICSS), January 2007 *, and* R.B. Dilmaghani, and R.R. Rao, “Hybrid Wireless Mesh Network with Application to Emergency Scenarios,” Journal of Software Engineer, Volume 3, number 2, Feb. 2008.

Chapter 5

Modeling Communication at Organizational Level by Petri nets

5.1 Introduction

In this chapter we present background information for understanding Petri nets in section 5.2. Petri nets will be used to model and analyze communication in large complex systems with an event-driven perspective. In section 5.3 we present definitions and properties of Petri nets. In section 5.4, we present structural invariants, followed by special classes of Petri nets in section 5.5. In section 5.6 we present different types of high-level Petri nets. In section 5.7 we present existing Petri net models from literature and finally, in section 5.8 we highlight the concluding remarks of this chapter.

5.2 A Review on Petri Nets

This section presents the required definitions that are relevant to the rest of this dissertation. For more details the reader is referred to [36] [37] [38] [39] [40] [41] [42] [43]. A Petri net consists of places, transitions, arcs and tokens. *Places* represent *conditions*, *transitions* represent *events*, and a *token* assigned to a place represents the execution of a task in a Petri Net. A *distribution of tokens* over the places of a Petri net is called a *marking*, and is denoted by μ. *Initial marking*, μ_0 represents the distribution

of tokens over places before the process begins. Transitions and places are connected through *arcs* to show the connection, $p \rightarrow t$ or $t \rightarrow p$. Transitions act on input tokens by a process known as *firing*. When a set of conditions are met, i.e. a transition becomes enabled and if it fires, it indicates the occurrence of an event. A transition becomes *enabled* when there are tokens on every incoming arcs (input places). When a *transition fires (executes)*, it *moves* the tokens from its input places and *puts* them on its output places. A *source place* is a place with no incoming arcs and a *sink place* is a place with no outgoing arcs. A *connected Petri net* is a Petri net where all transitions and places (excluding source and sink places) are on a path from the source place to the sink place where they contribute to the completion of the task. In a connected Petri net all transitions have at least one incoming arc and one outgoing arc. A Petri net is a 5-Tuple

$$(P, T, A, \mu_0, W) \tag{5.1}$$

where P is a finite set of places with m members, m is the total number of places,
T is a finite set of transitions with n members, n is the total number of transitions,
A is a finite set of all arcs from transitions to places, $t \rightarrow p$ and from places to transitions, $p \rightarrow t$
μ_0 is the initial marking
and W is a finite set of arc weights, natural numbers, denoting the number of arcs connecting a place to a transition or a transition to a place. Therefore multiple arcs between the same pair of transition and place is represented by an integer number greater than one and equal to the number of arcs.

A flow matrix is an $m \times n$ matrix where m is the number of places and n is the number of transitions. Flow matrix indicates the flow relation between transitions and places. Places form the rows and transitions form the columns of the flow matrix. The elements of the flow matrix are defined as following:

$$f_{i,j} = \begin{cases} -1 & \text{if } P_i \rightarrow T_j \text{and } T_j \nrightarrow P_i \\ 0 & \text{if } (P_i \rightarrow T_j \text{and } T_j \rightarrow P_i) \text{or } (P_i \nrightarrow T_j \text{and } T_j \nrightarrow P_i) \\ 1 & \text{if } P_i \nrightarrow T_j \text{and } T_j \rightarrow P_i \end{cases} \tag{5.2}$$

Definition 5.1 *A pair of a place and a transition is called a self-loop if that place is*

both an input and output to the transition [37]. A pure *Petri net is a self-loop free Petri net.*

Remark 5.1 *A flow matrix is uniquely representing a Petri net model if the Petri net is self-loop free (pure) and all arc weights equal to zero or one.*

Based on definition, there is more than one Petri net corresponding to the element zero of the flow matrix, one Petri net with at least one self-loop and the other one with no connection between that pair of place and transition. There is no indication of number of arcs in the corresponding element of the flow matrix and therefore Petri nets with different number of arcs than one between a pair of transition and a place are represented with the same flow matrix. Hence to have a flow matrix uniquely representing a Petri net, the Petri net has to be self-loop free and have all arc weights smaller or equal to one.

Definition 5.2 *The set of input transitions to a place P_i is denoted by ${}^\bullet P_{P_i}$ and the set of output transitions of a place is represented by $P_{P_i}^\bullet$ [36].*

Traps and siphons are special sets of places with properties that help us identify the existence or absence of deadlocks in a Petri net depending on the topology of the transitions connecting to those places.

Definition 5.3 *A set of places P is a siphon iff ${}^\bullet P \subseteq P^\bullet$ [36]. P_{ms} is a minimal siphon if there is no other siphon $P^{'}$, such that $P^{'} \subset P_{ms}$.*

For example if $P_{S1} = \{p_1, p_2, p_3\}$ and $P_{S2} = \{p_1, p_2\}$ are the only siphons in a system, P_{S2} is a minimal siphon since there is no other set of siphons that is a subset of P_{S2}, hence $P_{ms} = P_{S2}$. Clearly P_{S1} is not a minimal siphon because $P_{S2} \subset P_{S1}$.

Definition 5.4 *A set of places P is a trap iff $P^\bullet \subseteq^\bullet P$ [36]. P_{mt} is a minimal trap if there is no other trap $P^{'}$, such that $P^{'} \subset P_{mt}$.*

Remark 5.2 *Once a set of places in a trap become marked, they remain marked for all the future reachable markings. Similarly once a set of places in a siphon become empty, they remain empty for all future reachable markings [36].*

Since the siphon remains unmarked if it looses all its tokens, therefore the corresponding transitions cannot fire and are dead. Computational techniques to find a generating family of siphons and traps can be found in [44].

Remark 5.3 *'The existence of at least one empty siphon in the model indicates deadlock in the system ' [36].*

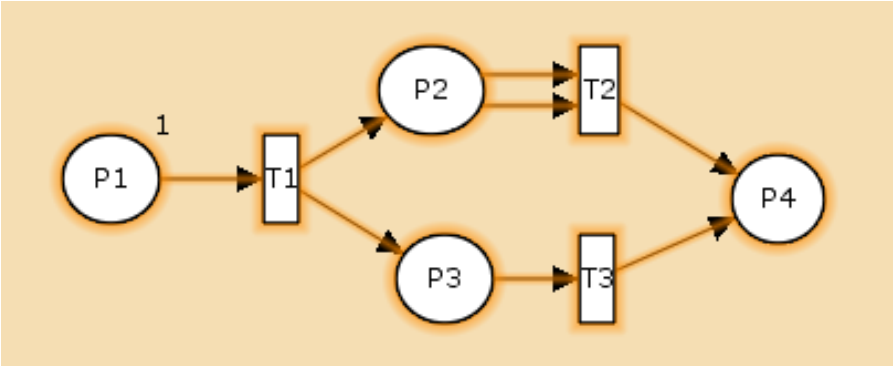

Figure 5.1: Illustration of a simple Petri net

Example 5.1 *For the Petri net model given in figure 5.1, the elements of Petri net definitions have been illustrated below.*

$P = \{P_1, P_2, P_3, P_4\}$

$T = \{T_1, T_2, T_3\}$

$\mu_0 = (1\,0\,0\,0)$

$A = \{P_1 \rightarrow T_1,\ T_1 \rightarrow P_2,\ T_1 \rightarrow P_3,\ P_3 \rightarrow T_3,\ T_3 \rightarrow P_4,\ P_2 \rightarrow T_2,\ T_2 \rightarrow P_4\}$

$W = \{1,\ 1,\ 1,\ 1,\ 1,\ 2,\ 1\}$

${}^{\bullet}P_{P_2} = \{T_1\}$

$P_{P_2}^{\bullet} = \{T_2\}$

$$F = \begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

μ_0 indicates the number of tokens at each place, corresponding to Ps, at initial marking. Note that the elements of weights in W correspond to the elements of arcs in A. $P = \{P_1, P_2, P_3, P_4\}$ is both a siphon and trap because ${}^\bullet P \equiv P^\bullet$. This leads to a property that enables us to specify whether the model is deadlock free or not, see section 5.4.

All reachable markings μ_i are obtained from initial marking, μ_0, and matrix multiplication of flow matrix by the firing vector [38]:

$$\mu_{i+1} = \mu_i + F.\psi_i \tag{5.3}$$

where ψ_i is the ith firing vector, a $n \times 1$ column vector with *one* in place of the transition that fires and *zero* in all other rows. μ_i is the ith marking, a $m \times 1$ column vector where element jth indicates the number of tokens in place j for $j = 1, 2, ..., m$. F is the flow matrix, $m \times n$. Therefore

$$\begin{aligned}
\mu_1 &= \mu_0 + F.\psi_0 \\
\mu_2 &= \mu_1 + F.\psi_1 \\
&= \mu_0 + F.\psi_0 + F.\psi_1 \\
&\cdot \\
&\cdot \\
&\cdot \\
\mu_N &= \mu_0 + F.(\psi_0 + \psi_1 + ... + \psi_{N-1}) \\
&= \mu_0 + F.\Psi
\end{aligned} \tag{5.4}$$

where Ψ represents the the sum of N firing vectors.

5.2.1 Reachability Tree

Reachability tree is used to illustrate the reachability of all possible markings and identify potential deadlocks in the model. Reachability tree develops from the initial marking, μ_0, based on the firing of enabled transitions according to equation 5.3 and 5.4. Any tree is specified with its root, nodes (vertices), and edges which are represented by μ_0, μ, and Ψ_i respectively. When a transition fires, one of the following three cases occurs:

- a new edge creates a new marking (node)
- a new edge connects to an already existing marking
- An already existing edge connects to an already existing marking

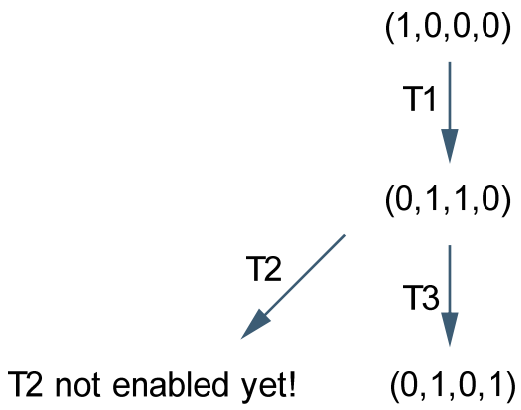

Figure 5.2: Reachability tree for $\mu_0 = (1, 0, 0, 0)$

The reachability tree for the Petri net in figure 5.1 for the initial marking $\mu_0 = (1, 0, 0, 0)$ is developed in figure 5.2 based on equation 5.4 and for the initial marking $\mu_0 = (2, 0, 0, 0)$ in figure 5.3. This is a behavioral property which depends on the initial marking. The existence of a potential deadlock indicates an obstacle in the proper flow of information which can be monitored and used as a feedback alert to the supervisory controller to take an alternate route as necessary.

5.3 Petri Net Properties and Definitions

In this section we present a review of Petri net properties from literature that are used in this dissertation. Reader is referred to [38] and [45] for further details. A use-case to the properties in the field of communication with application to emergency response follows.

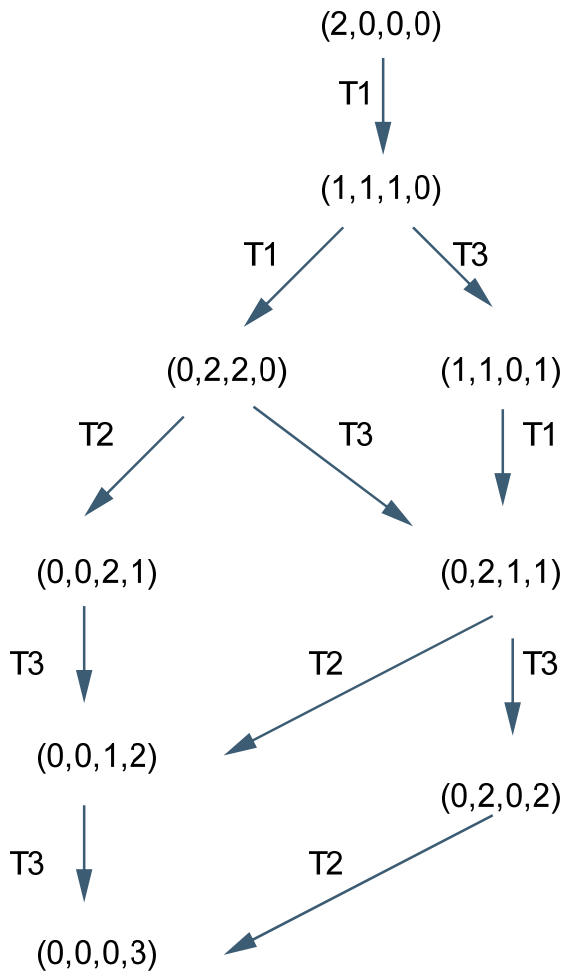

Figure 5.3: Reachability tree for $\mu_0 = (2, 0, 0, 0)$

Definition 5.5 *Liveness: A Petri nets can have different degrees of liveness, $L_1 - L_4$. Liveness property is defined for transitions. The Petri net is considered L_k-live if and only if all of its transitions are L_k-live [45] where a transition is:*
- Dead, if and only if it can never fire and hence it does not appear in any firing sequence
- L_1 live if and only if it may fire, i.e. it is in some firing sequence
- L_2 live if and only if can fire often, for every positive integer k, the transition appears at least k times in some firing sequence
- L_3 live if and only if can fire infinitely often
- L_4 live if and only if it may always fire, i.e. it is L_1 live in every reachable marking

Note that these are increasingly stringent requirements [45]: L_{i+1}-liveness implies L_i-liveness, for $i = 1, 2, 3$.

Predicting and preventing deadlocks in the analysis of work flow allows us to prevent bottlenecks, ensure proper flow of information, and completion of task. The liveness definition of a Petri net refers to the transitions in a Petri net. A Petri net model has a deadlock if there is at least one transition that can never be fired in any firing sequence. Figure 5.1 has one dead transition, T_2. This transition does not become enabled under given initial marking. This was verified in figure 5.2. In modeling communication protocol among different organizations, it is crucial that information reaches all designated destinations which requires that no transition is dead.

Definition 5.6 *Conservation: A Petri net is conservative if total number of tokens in the model is constant [38].*

Conservation is a structural property and is independent of initial marking of Petri net. It is an important property as it indicates that tokens that are representing resources are not created or destroyed in the model. The Petri net of figure 5.1 is not conservative while the Petri net of figure 5.7 is conservative.

Definition 5.7 *Strictly conservative: the number of inputs to each transition is equal to the number of outputs [38].*

Association in Emergency Response: no victims, patients or physical resources disappear or are created during an event. Loss of a victim or resource is modeled in a place

allocated for such purposes.

Definition of conservation with respect to a weighting vector is presented in [38].

Definition 5.8 *Boundedness: A Petri net is n-bounded where n is a natural number, if and only if for all reachable states the number of tokens in each place does not exceed n [38].*

The Petri net model of figure 5.1 is bounded while Petri net models of figure 6.5 are not bounded. The reachability tree of a bounded Petri net is finite [13], i.e. the number of nodes (markings) is finite. For an infinite reachability tree, the tree may expand with a pattern such that succeeding markings is similar to the preceding one where the only difference is the incremented number of tokens in a place in the marking. In such cases, the place is not bounded and the number of tokens is incrementing and represented by a parameter called ω. Hence instead of dealing with infinite reachability tree, we can build up a coverability tree which is a finite representation of the same system utilizing ω to replace the incrementing number of tokens in one place. ω represents infinite number of tokens in a place and therefore a tree that contains ω is not bounded, figure 6.4.
Association in Emergency Response: Places that represent resources are bounded to the number of resources available. Places that represent number of victims or patients are also bounded with an exception that if the scale of the incident grows or if the incident expands to a second incident (multiple incident scenarios), there is need to ask for more resources and the bound on the number of tokens on places representing such resources need to be updated.

Definition 5.9 *Safeness: A place in a Petri net is safe if the number of tokens in that place never exceeds one i.e. it is bounded by one [38].*

Association in Emergency Response: safety slows down the process but provides reliability. Individual networks may be safe but when put together, they may not be safe. Decontamination process is an example where victims arrive in different queues and when put together, the place will not be safe anymore.

Remark 5.4 *A Safe or binary Petri net is a Petri net where for all reachable markings from initial marking μ_0, each place is safe.*

A Petri net is well-formed if it is live and bounded for all reachable markings [39].

Definition 5.10 *Persistency: A Petri Net is persistent if for any two enabled transitions, the firing of one transition does not disable the other transition [40].*

This implies that a choice can be made in the order of execution of the two enabled transitions.

Remark 5.5 *A decision(choice)-free Petri net is a Petri net in which there exists only one enabled transition at any given time and there is no decision or choice to be made on which transition to fire [40].*

5.4 Structural Invariants

The structural invariants are dependent on the topology of the Petri net and independent of the initial marking of the Petri net. Structural invariants are important in analyzing the structural behavior of a given Petri net to extract certain properties such as existence or absence of deadlocks.
Place invariants are a set of places where the weighted sum of tokens is constant for all reachable markings, equation 5.5. Place invariants are represented by an n-dimensional integer vector P_{inv}, where n is the number of places in the Petri nets. Non-zero entries of P_{inv}, correspond to the places that belong to the set of place invariants [36].

$$P_{inv}^{T}\ \mu = P_{inv}^{T}\ \mu_0 \tag{5.5}$$

The place invariants of a Petri net are calculated by finding integer vector to the equation 5.6:

$$P_{inv}^{T}\ F\ =\ 0 \tag{5.6}$$

where F is the $n \times m$ incidence matrix (or flow matrix) of the Petri net model. Transition invariant is the dual of place invariant. A transition invariant is computed by finding integer vector to the equation below:

$$F\,T_{inv}\ =\ 0 \tag{5.7}$$

Place invariants are important in designing supervisory control and analyzing the behavior of a given Petri net for existence or absence of deadlocks. The following definitions and properties are used to investigate the existence or absence of deadlocks.

Remark 5.6 *A place invariant vector with nonnegative elements indicate a set of places that is both a trap and siphon [36].*

Definition 5.11 *A controlled siphon is a siphon that remains marked for all reachable markings. A controlled siphon is either trap-controlled or invariant controlled. In the former case the trap is initially marked and therefore prevents the siphon from ever loosing all of its tokens. An invariant-controlled siphon's marking is guaranteed by the place invariant property which ensures a fixed number of tokens for the set and therefore not only it never looses all its token but even the total number of tokens remains unchanged [36].*

Property 5.1 *A Petri net is deadlock free if every siphon in the net is a controlled siphon, i.e. for all reachable markings from initial marking, the siphon stays non-empty (always marked) [36].*

Example 5.2 develops Petri net definitions and properties explained above.

Example 5.2 *For the Petri net model given in figure 5.4, the following definitions have been developed:*
a. flow matrix
b. reachability tree
c. set of input transitions and output transitions to a place
d. set of place invariants
e. set of siphons and traps and whether there is a minimal siphon

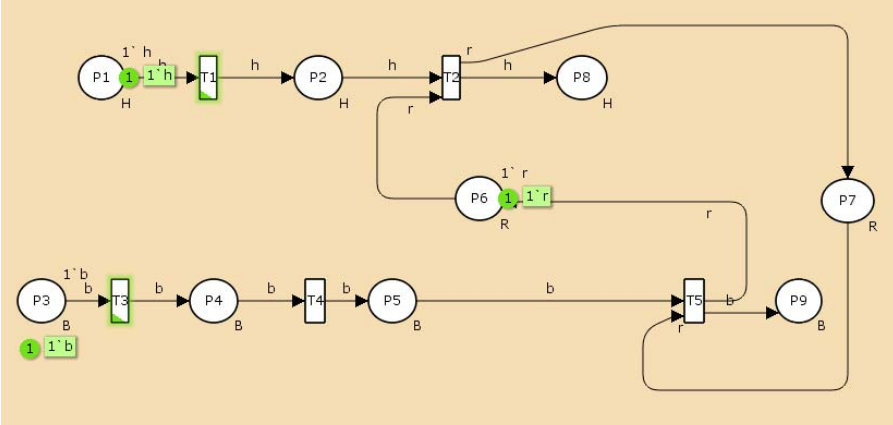

Figure 5.4: A Petri net model

a. The flow (incidence) matrix is formed as follows:

$$F_p = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.8}$$

This Petri net is safe and conservative.

b. The reachability tree is shown in figure 5.5. This is a different representation to summerize the markings at each state. The name of places which are marked by one token appear on the reachability tree. For example, instead of root node represented by $\mu_0 = (1, 0, 1, 0, 0, 1, 0, 0, 0)$, the root node is represented by$P_1 P_3 P_6$.

c. set of input transitions and output transitions to place P_5:

$$\begin{aligned} {}^{\bullet}P_{P_5} &= \{T_4\} \\ P_{P_5}^{\bullet} &= \{T_5\} \end{aligned} \tag{5.9}$$

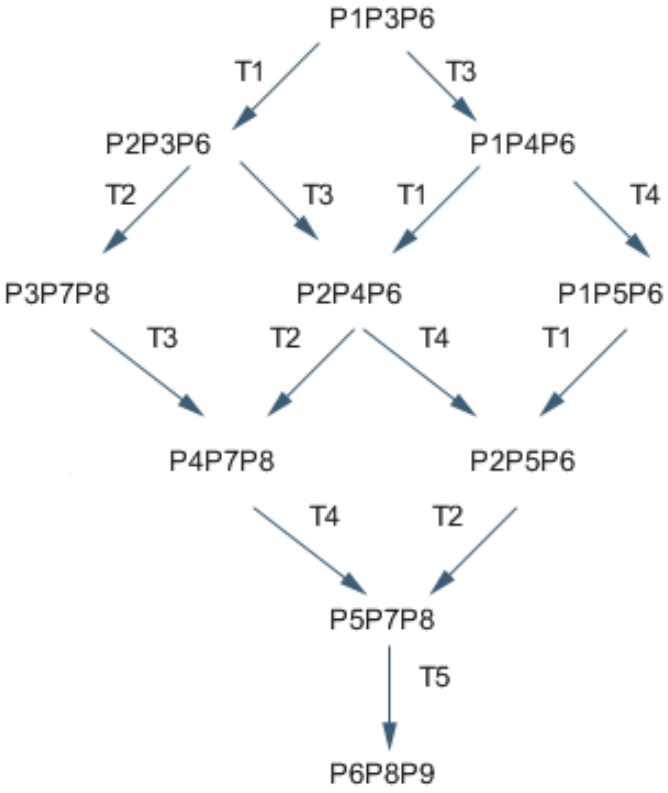

Figure 5.5: Reachability tree of the Petri net model of figure 5.4

A complete list of input transitions and output transitions to all place is presented in table 5.1.

Table 5.1: Set of input and output transitions for places in figure 5.4

Place	Set of Input Transitions	Set of Output Transitions
P_1	None	T_1
P_2	T_1	T_2
P_3	None	T_3
P_4	T_3	T_4
P_5	T_4	T_5
P_6	T_5	T_2
P_7	T_2	T_5
P_8	T_2	None
P_9	T_5	None

d. The set of place invariants:

$$\begin{aligned} &\{P_6, P_7\} \\ &\{P_1, P_2, P_8\} \\ &\{P_3, P_4, P_5, P_9\} \\ &\{P_3, P_4, P_5, P_6, P_8\} \end{aligned} \tag{5.10}$$

the place invariant vectors are found as:

$$\begin{aligned} X_1^T &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \\ X_2^T &= \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\ X_3^T &= \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \\ X_4^T &= \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \end{bmatrix} \end{aligned}$$

The place invariant sets are both siphons and traps. The complete list of siphons and traps and minimal siphons and minimal traps is presented in table 5.2. This Petri net is

deadlock free as it verified by the reachability tree developed in 5.5.
e. The set of place invariants presented in 5.11 are both siphons and traps.

$$\{P_6, P_7\}$$
$$\{P_1, P_2, P_8\}$$
$$\{P_3, P_4, P_5, P_9\}$$
$$\{P_3, P_4, P_5, P_6, P_8\} \tag{5.11}$$

The list of siphons and traps is not limited to table 5.2 but rather than providing an exhaustive list, we turn our attention to the properties and conditions that help us develop properties which help us to determine the existence or non-existence of deadlocks before extracting a comprehensive list of siphons and traps. Based on property 5.1 and definition 5.11, a Petri net model is deadlock free if all siphons are controlled and that implies either invariant siphons or trap-controlled siphons. In 5.11, we presented the list of place invariants which are both siphons and traps. The property of place invariant guarantees a fixed number of tokens and since they are traps, when they become marked, they remain marked. The strict or flexible conditions to guarantee a trap-controlled siphon are extracted as follows:

a. If P_1 is a place on the listed set of siphons, in order for the siphon to be trap-controlled, P_2 has to be on the set as well.
b. If P_3 is a place on the listed set of siphons, in order for the siphon to be trap-controlled, P_4 has to be on the set as well.
c. If P_4 is a place on the listed set of siphons, in order for the siphon to be trap-controlled, P_5 has to be on the set as well.
d. If P_2 is a place on the listed set of siphons, in order for the siphon to be trap-controlled, P_7 or P_8 has to be on the set as well.
e. If P_5 is a place on the listed set of siphons, in order for the siphon to be trap-controlled, P_6 or P_9 has to be on the set as well.
f. If P_6 is a place on the listed set of siphons, in order for the siphon to be trap-controlled, P_7 or P_8 has to be on the set as well.

g. If P_7 is a place on the listed set of siphons, in order for the siphon to be trap-controlled, P_6 or P_9 has to be on the set as well. The list of other siphon sets which are not minimal but trap-controlled are presented in table 5.4. PI in the last column indicates place invariants.

5.5 Special Classes of Petri Nets

Definition 5.12 *: A state machine (SM) is a Petri net in which all transitions have one input and one output. The significant nodes in a state machine are the places [36].*

A single token can enable multiple transitions leading to conflicts. A Petri net is conflict free if each place has at most one output transition.

Definition 5.13 *: A marked graph (MG) is a Petri net in which all places have a single input and a single output transition. The significant nodes in a marked graphs are transitions which might have multiple inputs or outputs and allow synchronizations [36].*

A Petri net is conflict free if each place has at most one output transition. A marked graph by definition is conflict free.

Definition 5.14 *: A free choice (FC) net is a Petri net such that for every arc from a place P to a transition T, i.e. $P \rightarrow T$,*
1. T is the only output transition of P (no conflict), or 2. P is the only input place of T (no synchronization) [36].

This implies that if place P_i is the input to a set of transitions, P_i is the only input place to all transitions belonging to the set. This also indicates that if any output transition of a place p is enabled, then all output transitions of p are enabled. All state machines and marked graphs belong to the class of free choice nets [36]. The example shown in figure 5.1 is a *free choice net*.

Remark 5.7 *The definition implies that if any output transition of a place p is enabled, then all output transitions of p are enabled. Both state machines and marked graphs are FC nets [36].*

Definition 5.15 *: An extended free choice (EFC) net is a Petri net such that for every arc from a place P to a transition T, $P \rightarrow T$,there exists an arc from all input places of t to all output transitions of p [36].*

This is a relaxed definition of FC to extend the class of nets maintaining the basic property. This implies that if another transition has another input place except P_i, this other place is input to other transitions of the set as well. From definition 5.2: if $p_1^\bullet \bigcap p_2^\bullet \neq 0$, then $p_1^\bullet = p_2^\bullet$
An EFC is live iff every siphon in the net are trap-controlled [36].

Definition 5.16 *: A Petri net is a work flow net if the following properties hold true [39]:*

- *There is only one source place such that ${}^\bullet P = \oslash$.*
- *There is only one sink place such that $P^\bullet = \oslash$*
- *Every node, transition or place, should be located on a path from source to sink or output place which holds true for a connected Petri net (connectedness property). Hence in a work flow net there are no dangling tasks or conditions and every task contributes to the completion of process.*

The example shown in figure 5.1 is a *work flow net*.

Corollary 5.1 *A Petri net that is a work flow is connected.*

Definition 5.17 *: A work flow net is sound if the following requirements are satisfied [39]:*

- *a. For any reachable marking, it is possible to terminate the process through a firing sequence such that there is only one token left at output place, P_o.*
- *b. When the process terminates (i.e. a token appears in P_o), there is no token left in any other places.*
- *c. There is no dead transition, i.e., starting with a token in the input place P_i, a task will be completed by following the appropriate route through the work flow net.*

Soundness property refers to the dynamics of a work flow net and is the minimal property any work flow process definition should satisfy. Note that soundness implies the absence of livelocks and deadlocks. In a livelock, the states of the processes involved in the livelock constantly change with regard to one another but none progresses [46]. Livelock is a special case of resource starvation [47]. The example shown in figure 5.1 is not *sound*.

In an emergency response application, we cannot enforce all the conditions of soundness property despite its useful guarantee on the state of being deadlock and livelock free. The reason is the limiting definition of item b. When a place represents a resource, the corresponding place gets its token back upon the completion of the partial task. Therefore when the system execution is completed, the sink is not the only place with tokens. Resource places keep their tokens as well as we assume resources do not disappear and hence the definition can be relaxed depending on the type of tokens for generality.

Definition 5.18 *: An asymmetric choice (AC) net is a Petri net such that for all pairs of places,* p_1 *and* p_2 *if* $p_1^\bullet \bigcap p_2^\bullet \neq 0$, *then* $p_1^\bullet \subseteq p_2^\bullet$ *or* $p_2^\bullet \subseteq p_1^\bullet$.

An AC is live iff every siphon is a controlled siphon [36]. The Venn diagram of figure 5.6 is taken from [36] which illustrates the relation of different Petri net classes.

5.6 High-level Petri Nets

High-level Petri nets have been developed to extend the modeling capability of Petri nets in more complex systems. Color Petri nets and Timed Petri nets are examples of high-level Petri net which we utilize in modeling emergency communication. In the rest of this chapter we briefly present background information necessary to understand the application of high-level Petri nets and readers are encouraged to refer to the given references for more details.

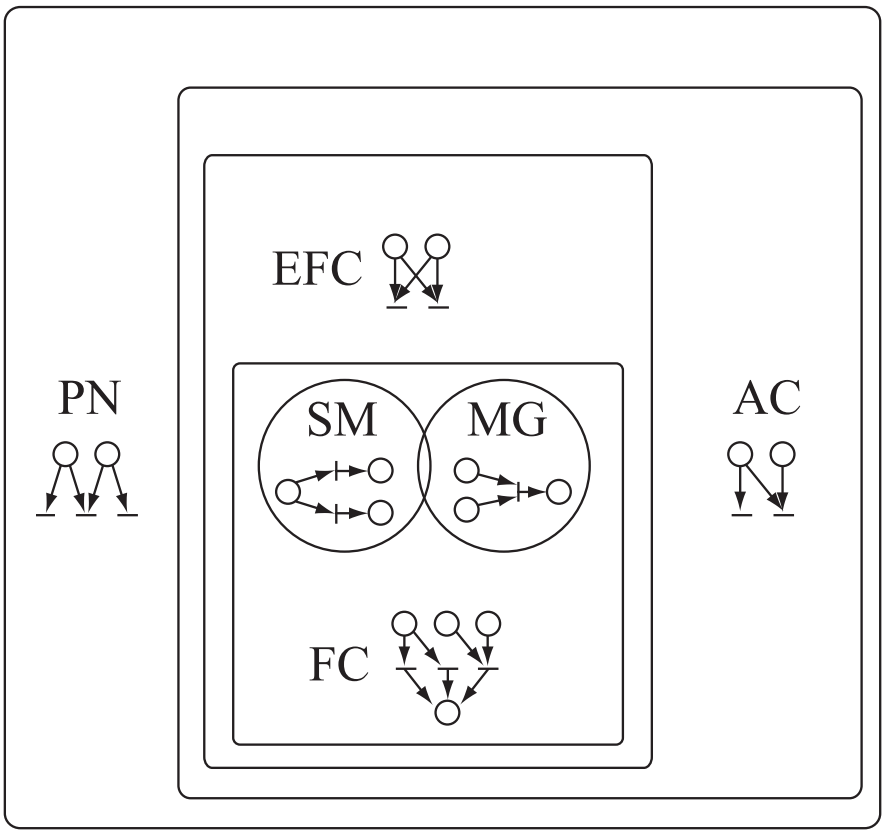

Figure 5.6: Relation of Petri net classes in a Venn diagram by Moody and Antsaklis

5.6.1 Color Petri Nets

Colored Petri nets allow users to define different types of data to label tokens. This is similar to concept of data type in programming languages to distinct between different types of tokens [41]. Figure 5.7 illustrates a color Petri net with variable

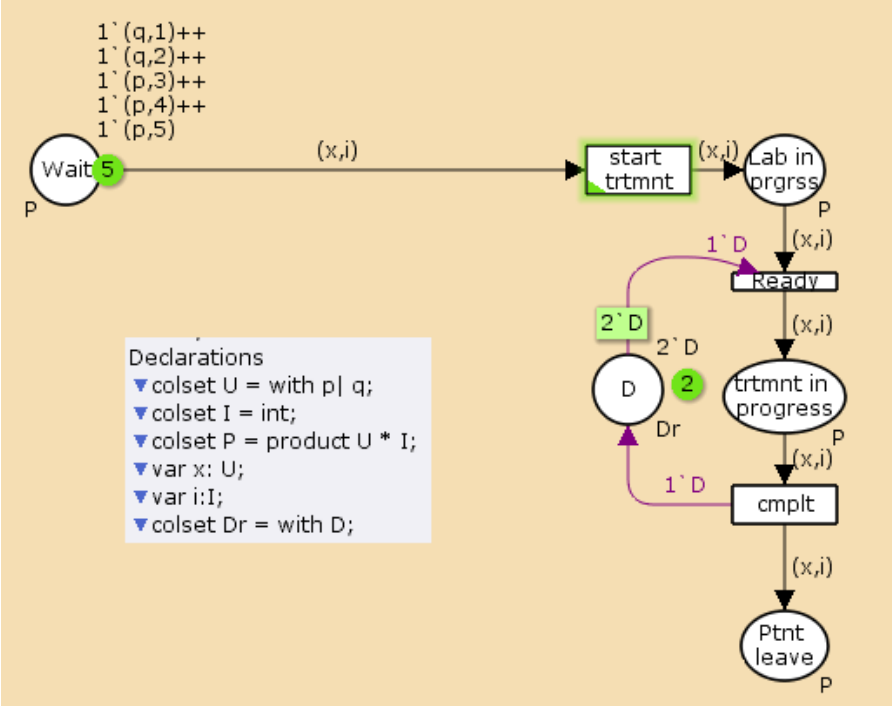

Figure 5.7: Illustration of a color Petri net

declarations to define and distinct different types of tokens. Different token types are defined by different colors. Token type is defined for places and indicated on arcs. Upon availability of the right type of tokens, a transition becomes enabled. All the Petri net models developed in this work are created using CPN tool [42].

5.6.2 Timed Petri Nets

Timed Petri nets are an extension of Petri nets or color Petri nets with a time value or time stamp associated to one or more of the components in the system including transitions, places and tokens. Time variable can be discrete or continuous, deterministic

or stochastic. Deterministic timed transition fire when the system time reaches the value. Similarly the tokens become available as the time stamp associated with them become valid. The execution of a timed Petri net is similar to handling time in programming languages to simulate discrete event systems. The system remains at a given time as long as there is a process ready for execution. The system time advances the clock to the next simulation time when there is no more task scheduled for the current time.

In stochastic timed Petri net models, typically a random delay with negative exponential probability distribution function (pdf) is associated to each timed transition [43]. Negative exponential distribution is the only continuous pdf with memoryless property which makes the analysis of discrete-state stochastic processes in continuous time possible by drastically reducing the amount of dynamic memory [48]. The Petri nets that are obtained by such approach are called Stochastic Petri Nets (SPN) which were independently proposed by [49] and [50] and similar ideas in [51] [48]. The only parameter of exponential pdf is the rate which is equal to the inverse of its mean.

> whether the timer is reset or not, the pdf of the remaining time to fire always follows the negative distribution function with the same mean due to the memoryless property of such pdf.

Stochastic timed Petri nets have been utilized in chapter 7. Figure 5.8 shows a simple timed Petri net model with a discrete clock. The @ indicates the time stamp. Tokens of type P are timed while the tokens of type r are not which represent resources. This implies that token type r are available at any time. Token P proceeds through the system as the system clock advances to the given time stamp on transitions. In this example all transitions are immediate.

5.7 Related Work

Figure 5.9 shows a Petri net that models the processes for the department of energy by [52]. This model holds the following properties:

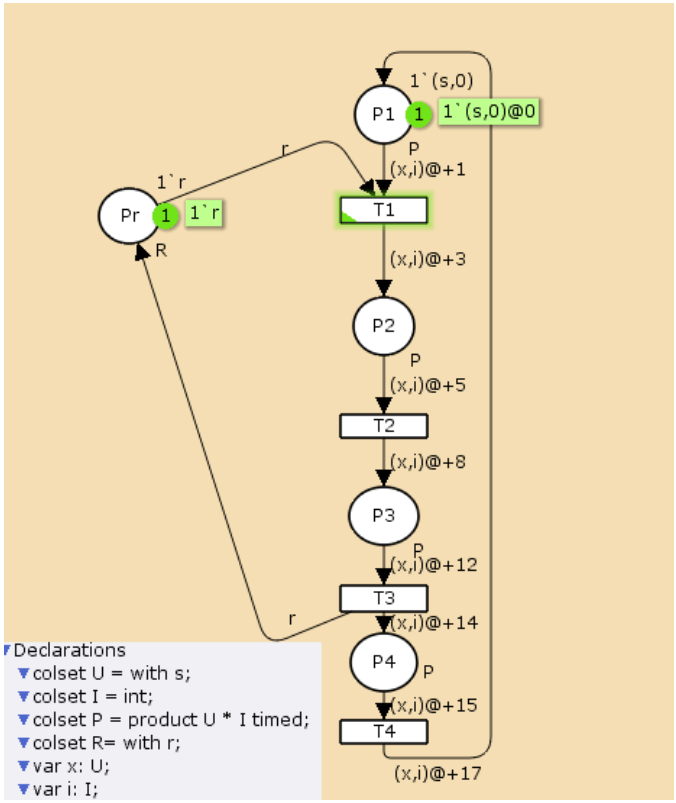

Figure 5.8: Illustration of a timed Petri net

- Boundedness: All places are 1-bounded and therefore the Petri net is safe.
- Liveness and Deadlock: There is no deadlocks; The Petri net is live.

Since there is no dead transition, all the valid markings including final marking are reachable from the initial marking. However, the drawback with this model is the existence of dangling tasks (transitions) without input or output conditions (places). This leads to unclear fire time and therefore it cannot contribute to a successful completion of the task. Figure 5.10 and figure 5.11 show the Petri net models to illustrate the process

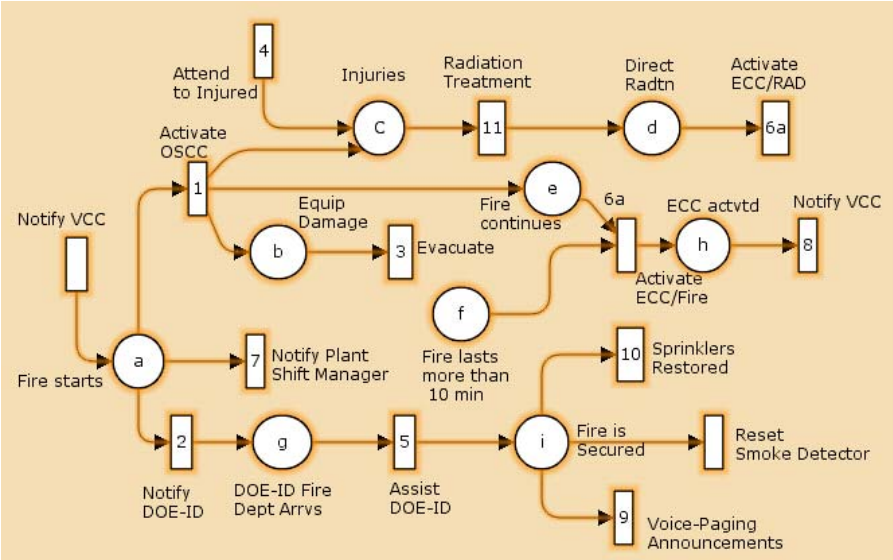

Figure 5.9: A Petri Net model developed for Department of Energy (DoE)

of patients treatment at a hospital without and with priority [53].

The first one assumes that all patients are in the same priority class and they are served in the order of arrival to the medical center. The second model has been developed for patients with priorities. Patients join the queue for registration, then they get into another queue to be assigned a bed; and next is being seen by a doctor. Beds, doctors and nurses are considered as resources in these models. Despite our effort to

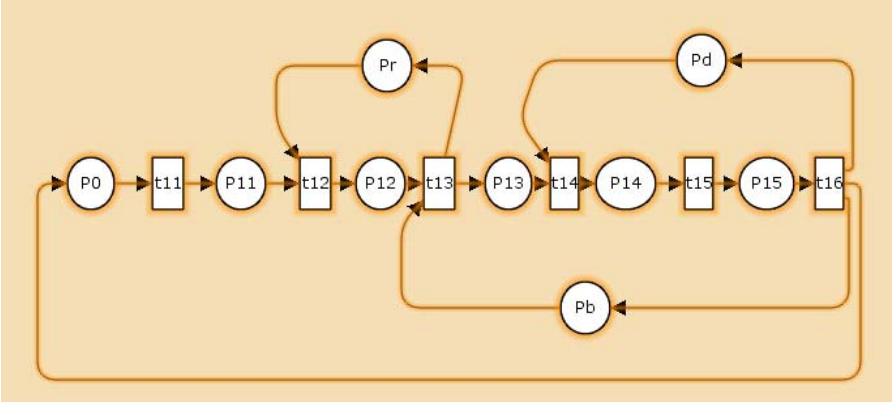

Figure 5.10: A Petri net model for patients with no priority

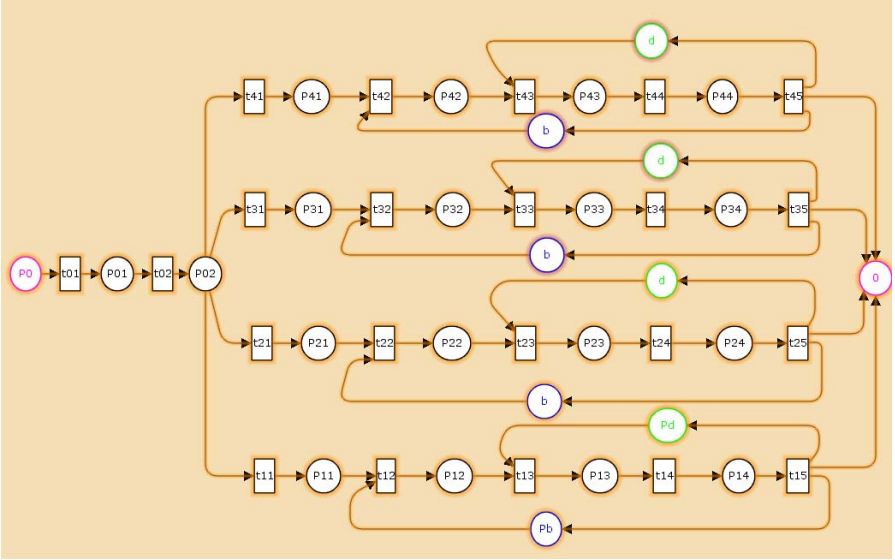

Figure 5.11: A Petri net model for patients with priority

contact the authors, we have not been able to find a complete draft of paper to examine the model in details.

5.8 Concluding Remarks

Petri nets enable us to decompose a system analysis to its structural and behavioral properties. Petri nets are investigated based on two different sets of properties: structural properties and behavioral properties. Structural properties refer to properties like liveness, boundedness, safeness, persistency, and structural invariants which are independent of initial marking of Petri net. Behavioral properties are based on the initial marking of Petri net and how it develops to other reachable markings. This is mainly carried out by investigating the behavioral properties mathematically including the incidence matrix and reachability tree. Structural analysis and performance prediction of Petri nets are presented in chapters 7 and 8.

Table 5.2: Set of Siphons and Traps for figure 5.4

Set of Places	Siphon	Trap	Minimal
P_1, P_2	Yes	No	Yes
P_1, P_2, P_3	Yes	No	No
P_1, P_2, P_3, P_4	Yes	No	No
P_1, P_2, P_3, P_4, P_5	Yes	No	No
$P_1, P_2, P_3, P_4, P_5, P_6$	Yes	No	No
$P_1, P_2, P_3, P_4, P_5, P_6, P_7$	Yes	Yes	No
$P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8$	Yes	Yes	No
$P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9$	Yes	Yes	No
P_1, P_3, P_4	Yes	No	No
P_1, P_3, P_4, P_5	Yes	No	No
P_1, P_3, P_4, P_5, P_6	Yes	No	No
$P_1, P_3, P_4, P_5, P_6, P_7$	Yes	Yes	No
$P_1, P_3, P_4, P_5, P_6, P_7, P_8$	Yes	Yes	No
$P_1, P_3, P_4, P_5, P_6, P_7, P_8, P_9$	Yes	Yes	No
P_1, P_6, P_7	Yes	No	No
P_1, P_6, P_7, P_8	Yes	No	No
P_1, P_6, P_7, P_8, P_9	Yes	No	No
$P_2, P_3, P_4, P_5, P_6, P_7$	No	Yes	No
$P_2, P_3, P_4, P_5, P_6, P_7, P_8$	No	Yes	No
$P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9$	No	Yes	No
P_2, P_4, P_5, P_6, P_7	No	Yes	No
$P_2, P_4, P_5, P_6, P_7, P_8$	No	Yes	No
$P_2, P_4, P_5, P_6, P_7, P_8, P_9$	No	Yes	No
P_2, P_5, P_6, P_7	No	Yes	No
P_2, P_5, P_6, P_7, P_8	No	Yes	No
$P_2, P_5, P_6, P_7, P_8, P_9$	No	Yes	No
P_2, P_6, P_7	No	Yes	No
P_2, P_6, P_7, P_8	No	Yes	No
P_2, P_6, P_7, P_8, P_9	No	Yes	No
P_2, P_7, P_8, P_9	No	Yes	No
P_2, P_8	No	Yes	Yes
P_2, P_8, P_9	No	Yes	No

Table 5.3: Set of Siphons and Traps for figure 5.4, Cont'd

Set of Places	Siphon	Trap	Minimal
P_3, P_4	Yes	No	Yes
P_3, P_4, P_5	Yes	No	No
P_3, P_4, P_5, P_6	Yes	No	No
P_3, P_4, P_5, P_6, P_7	Yes	Yes	No
$P_3, P_4, P_5, P_6, P_7, P_8$	Yes	Yes	No
$P_3, P_4, P_5, P_6, P_7, P_8, P_9$	Yes	Yes	No
P_3, P_6, P_7	Yes	No	No
P_3, P_6, P_7, P_8	Yes	No	No
P_3, P_6, P_7, P_8, P_9	Yes	No	No
P_4, P_5, P_6, P_7	No	Yes	No
P_4, P_5, P_6, P_7, P_8	No	Yes	No
$P_4, P_5, P_6, P_7, P_8, P_9$	No	Yes	No
P_5, P_6, P_7	No	Yes	No
P_5, P_6, P_7, P_8	No	Yes	No
P_5, P_6, P_7, P_8, P_9	No	Yes	No
P_5, P_7, P_8, P_9	No	Yes	No
P_5, P_8, P_9	No	Yes	No
P_5, P_9	No	Yes	Yes
P_6, P_7	Yes	Yes	Yes
P_6, P_7, P_8	Yes	Yes	No
P_6, P_7, P_8, P_9	Yes	Yes	No
P_6, P_8	No	Yes	Yes
P_6, P_8, P_9	No	Yes	No
P_7, P_8, P_9	No	Yes	No
P_7, P_9	No	Yes	Yes

Table 5.4: Set of Siphons and Traps for figure 5.4, Cont'd

Set of Places	Siphon	Trap	Trap-controlled
P_1, P_2, P_8	No	Yes	Yes (PI)
$P_1, P_2, P_4, P_5, P_7, P_9$	No	Yes	Yes
$P_1, P_2, P_4, P_5, P_6, P_7, P_9$	No	Yes	Yes
$P_1, P_2, P_4, P_5, P_7, P_8, P_9$	No	Yes	Yes
$P_1, P_2, P_3, P_4, P_5, P_7, P_9$	No	Yes	Yes
$P_1, P_2, P_3, P_4, P_5, P_6, P_8, P_9$	No	Yes	Yes
$P_1, P_2, P_3, P_4, P_5, P_7, P_8, P_9$	No	Yes	Yes
P_2, P_6, P_8	No	Yes	Yes
P_3, P_4, P_5, P_6, P_8	No	Yes	Yes(PI)
P_3, P_4, P_5, P_9	No	Yes	Yes(PI)
P_3, P_4, P_5, P_7, P_9	No	Yes	Yes
P_4, P_5, P_6, P_8	Yes	Yes	Yes
P_4, P_5, P_6, P_7, P_8	Yes	Yes	Yes
P_4, P_5, P_7, P_9	Yes	Yes	Yes
P_4, P_5, P_7, P_8, P_9	Yes	Yes	Yes
P_5, P_7, P_9	No	Yes	Yes

Table 5.5: Legends for no priority

Place	Description
P_0	Patients source
P_{11}	Arrival of patients
P_{12}	Registration
P_{13}	Bed (treatment area) assignment
P_{14}	Doctor assignment
P_{15}	Treatment
P_r	Registration nurse available
P_b	Bed (treatment area) available
P_d	Doctor available

Table 5.6: Legends for with priority

Place	Description
P_0	Patients source
P_{01}	Arrival of patients
P_{02}	Classification of patients urgency level
P_{i1}	Patients of urgency i, for i=1,2,3 and 4
P_{i2}	Bed assignment for patients of urgency i, for i=1,2,3 and 4
P_{i3}	Physician assignment for patients of urgency i, for i=1,2,3 and 4
P_{i4}	Treatment for patients of urgency i, for i=1,2,3 and 4
P_b	Bed (treatment area) available
P_d	Doctor available

Chapter 6

Structural Properties and Complexity Reduction Transformations

6.1 Introduction

In section 6.2 we present properties of Petri nets that lead to a unique Petri net representation of a system for a given reachability tree. We present reducing Complexity of Petri nets of certain structures in section 6.3 to simplify the complexity of analysis of larger Petri nets. Certain properties have been examined to determine whether they have been preserved within the transformation or not. In section 6.4 we investigate the set of conditions that lead to deadlock free control design. In section 6.5 we present conditions for liveness of interacting Petri net models with application to emergency response following ICS structure. This includes topologies and classes of Petri nets that can be simplified for deadlock analysis in a large Petri net model. Finally, in section 6.6 we highlight the concluding remarks of this chapter.

6.2 Unique Petri Net Representation of A System for a Given Reachability Tree

A system can be uniquely represented either by its Petri net or by its reachability tree if the system holds certain properties discussed in theorem 6.1.

Theorem 6.1 *For a given reachability tree, there exists only one corresponding Petri net model if the Petri net holds the properties of connectedness, absence of dead transition, boundedness and persistency. This implies that the two representations are equivalent only if all the above properties hold true.*

Proof- The proof is achieved by contradiction. We assume there is more than one Petri net corresponding to a given reachability tree holding the properties mentioned above and then see how this contradicts the assumptions.

For the same reachability tree, based on definition of the reachability tree in section 5.2.1 and equations 5.4, μ_0, Ψ and μ are the same. Hence there has to be at least one place which does not appear in the marking (nodes of the tree) or a transition which does not appear on the edges of the reachability tree. The former indicates the existence of at least one place that never has a token and never receives one while the latter indicates the existence of a transition that never fires. In either case, one or more of the following properties does not hold true: connectedness, deadlock free, or persistency. A place that never receives a token is either located after a dead transition or it is not connected to the rest of the model. Similarly a transition that never fires is an enabled transition that becomes disabled for the lack of persistency property or it is not connected to the rest of the model (never becomes enabled). Two Petri nets with different number of arcs between a place and a transition have the same reachability tree with parameter ω, definition 5.8, abstracting out the specific number of connections. The two Petri nets are not bounded by the same integer number.

Therefore with the properties of connectedness, deadlock free, persistency and boundedness, there is a unique Petri net model corresponding to any given reachability tree.

This allows us to uniquely analyze such Petri nets based on the reachability tree in stochastic modeling of the systems to identify the probability and the cost of reaching one state within certain number of steps as described in section 7.2. The following two examples illustrate the lemma.

Example 6.1 *For the reachability graph given in figure 6.1, there is more than one Petri net representing the same reachability tree when the property of connectedness does not*

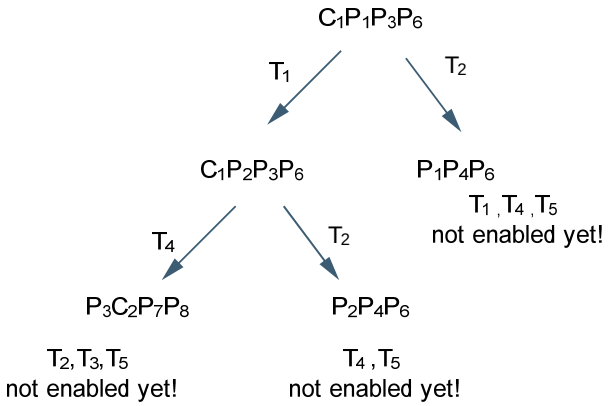

Figure 6.1: Reachability Tree of example 6.1

hold true.

The Petri net model in Figure 6.2 shows one possible Petri net corresponding to the given reachability tree in 6.1. Figure 6.3 represents another Petri net corresponding to the same reachability tree where the places that never receive a token and dead transitions have been removed (hence not appear on the reachability tree). In general if the

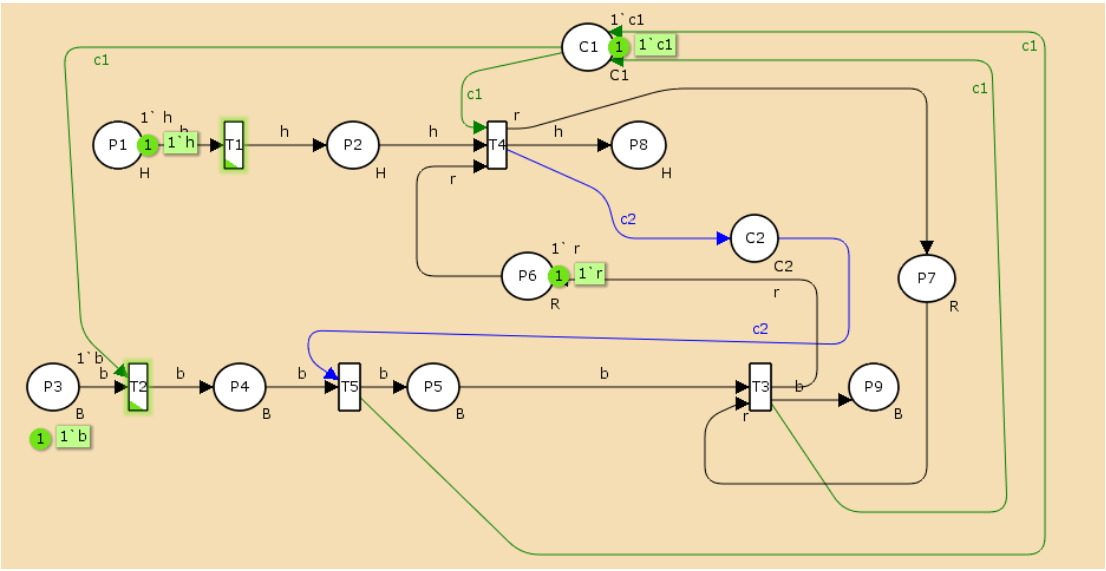

Figure 6.2: An acyclic Petri Net model with reachability tree given in figure 6.1

network is not connected, there would be multiple models associating with the same reachability tree. If the arrival rate of input tokens is higher than the processing time

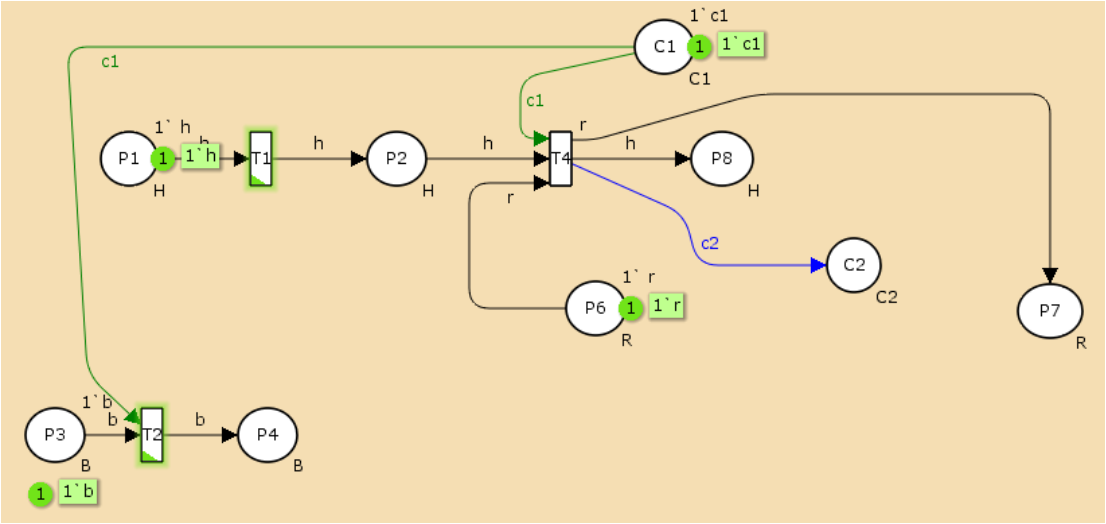

Figure 6.3: A non-connected Petri net model with reachability tree given in figure 6.1

of intermediate places in figure 6.2, the number of tokens increases and the model will not be bounded. The number of tokens in absorbing places P_8 and P_9 is not limited to a number and therefore it can be represented by parameter ω. This also leads to non-unique equivalent representations and will be illustrated within the next example.

Example 6.2 *For the reachability tree presented in figure 6.4, there is more than one Petri net corresponding the same reachability tree as the property of boundedness does not hold true.*

In this example tokens get accumulated in places P_5 and p_5. The w in the reachability tree represents this accumulation. In the top Petri net model in figure 6.5, w refers to the two arcs connecting T_4 to P_5 which implies an increment by two each time the transition T_4 fires. In the second Petri net model of this figure, every time transition t_4 fires, only one token is transmitted to the place p_5. This illustrates how the appearance of w buries the information and therefore for a unique representation, the reachability tree should be finite and tokens get recycled back to the model rather than being accumulated or symbolically can be interpreted as such. Now, if we add another arc from transitions

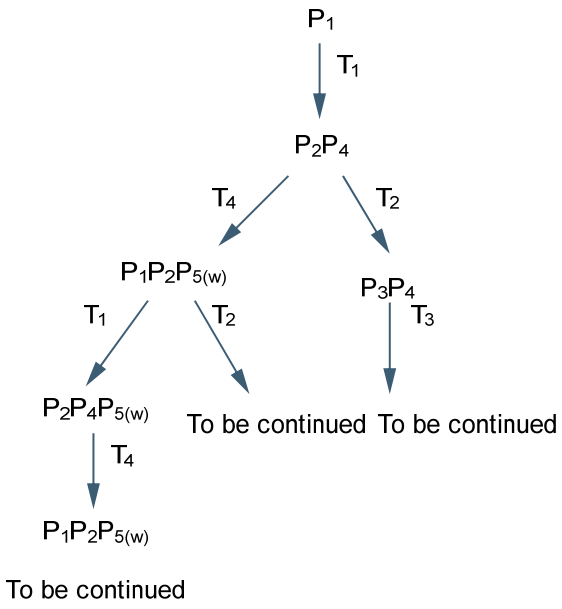

Figure 6.4: Reachability Tree of example 6.2

T_4 to P_5, the same ω appears. The networks are clearly different but they develop the same reachability tree. If the Petri net contains no self-loops, i.e. no transition has inputs and output to and from the same place, the enabling condition can be expressed based on the incidence matrix. However if a self-loop exists, this topology can not be uniquely expressed in the reachability tree.

6.3 Reducing Complexity of Communication among Different Organizations

For a large Petri net model consisting of individual Petri nets, the goal is to find methods to simplify the model. We present key properties that need to hold true to allow the replacement of the larger model by a simpler model with respect to the properties and the application. We present methods that allow us to extract a set of transitions $\{T_i, T_{i+1}, ..., T_j\}$ and places $\{P_k, P_{k+1}, ..., P_l\}$ out of a larger set of transitions $\{T_1, T_2, ..., T_m\}$ and places $\{P_1, P_2, ..., P_n\}$, $1 \leq i \leq j \leq m$ and $1 \leq k \leq l \leq n$, such

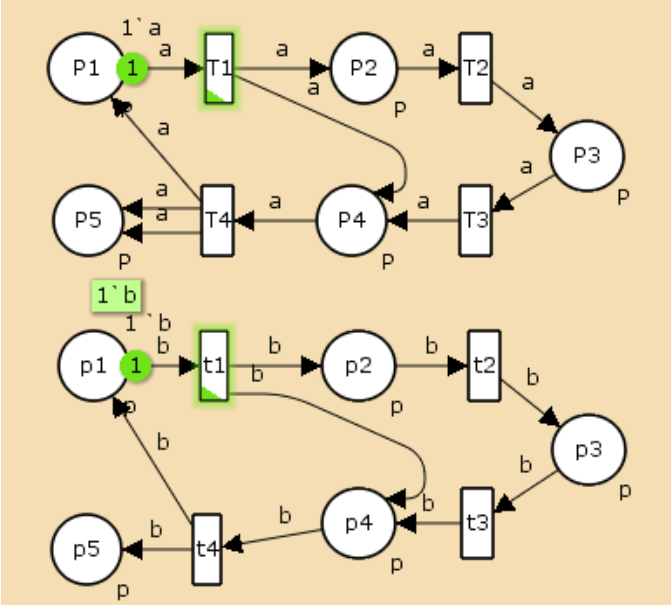

Figure 6.5: Two Petri Nets corresponding to the reachability tree of figure 6.4

that for the same initial marking and valid firing sequences in each model, the final marking is the same in both Petri net models.

6.3.1 General Complexity Reduction Transformations

A limiting factor in modeling and analyzing larger systems is the size of the Petri net. It is very helpful to reduce the size of a Petri net model which may or may not preserve the main properties of the Petri net [37] [54]. In this section, we present transformations that reduce the complexity of certain prototypes and investigate whether the properties of liveness, boundedness and safeness are preserved or not. The prototypes presented for the logical *AND*, logical *OR* and directed cascade topology can be extended to a larger number of places and transitions with the same structure. Existing linear algebra techniques are used to reduce the complexity. Below we present fundamental definitions and remarks that are used in this work.

Definition 6.1 *Notation $t(i, T_j)$ indicates the i^{th} firing of transition T_j. For the applications in this chapter, the i^{th} firing of a transition is dropped without loss of generality and hence the notation for firing time of the transition is reduced to t_{T_j}.*

Definition 6.2 *A* transition directed cascade *is a sequence of transitions and places, starting with a transition T_i and ending with transition T_n, $T_iP_jT_kP_lT_n$, such that each place or transition in the cascade is the output to the transition or place prior to it and the input to the following transition or place.*

In the above definition P_j is an output place to transition T_i and an input to transition T_k, and similarly transition T_k is an output transition to place P_j and an input transition to place P_l.

Definition 6.3 *A* place directed cascade *is a sequence of places and transitions, starting with a place P_i and ending with a place P_n, $P_iT_jP_kT_lP_n$, such that each transition or place is the output to the place or transition preceding it and the input to the place or transition following it.*

In the above definition P_k is an output place to transition T_j and an input place to transition T_l, and similarly transition T_j is an output transition to place P_i and an input transition to place P_k.

Definition 6.4 *When an equivalent transition, T_{eq}, replaces a set of consecutive transitions in a transition or place directed cascade, the firing time of T_{eq} is equal to the sum of firing time of all transitions in the series.*

In color Petri net (CPN) tool used in this work [42], the time constraint on each transition represents the time units that elapses after the previous event. This means that if T_{j+1} follows T_j in a directed cascade, the firing time of transition T_{j+1} includes the firing time of the preceding transition, T_j, plus the additional time for this event, i.e. $t_{T_{j+1}} = t_{T_j} + \delta$ where δ represents the time that elapses after T_j. Hence if T_{eq} replaces a directed cascade consisting of only two transitions T_j and T_{j+1}, The firing time of T_{eq} is represented by: $t_{T_{eq}} = t_{T_{j+1}}$.

Definition 6.5 *A* place cascade loop *is a directed place cascade where the last place in the sequence is the same as the first place.*

Remark 6.1 *When the first transition in a directed cascade transition or directed cascade place fires, the token traverses through the cascade and all the following transitions become enabled and fire as well.*

This definition concerns structural property of Petri net not behavioral. This is true when all line of communications, connecting arcs, between conditions (places) and events (transitions) in a cascade topology are valid and working. A failure in a condition or event stops the process and needs to raise a flag for supervisor.

Definition 6.6 *Two Petri nets are equivalent if for the same initial marking, i.e. same input, the final marking is the same for any valid firing sequence in both models.*

Definition 6.7 *A self-loop transition is a transition which has only one incoming and one outgoing arc from the same place in the Petri net model.*

Remark 6.2 *A self-loop transition may indicate the existence of an infinite loop if no supervisory or monitoring mechanism employed.*

Definition 6.8 *A self-loop place is a place which has only one incoming and one outgoing arc from the same transition in the Petri net model.*

Remark 6.3 *A self-loop place has to be* initially marked *to prevent the corresponding transition from being dead (live at level zero). A self-loop place can be eliminated safely.*

Figure 6.6a illustrates a self-loop transition which can lead to an infinite loop if the token never moves out of the loop. The replacement of a self-loop place is illustrated in figure 6.6b. The place in the self-loop can always be removed as it represents a condition that is always met and hence redundant.

Definition 6.9 *The structural representation of a token traversing through a set of places and transitions is represented by $P_i T_j P_l$ which indicates transition T_j is the output transition of place P_i and place P_l is the output place of transition T_j.*

Theorem 6.2 *Prototype I represents a transformation that reduces a Petri net by decomposing the possible paths that a token takes into directed cascades while preserving the logical* OR *merge.*

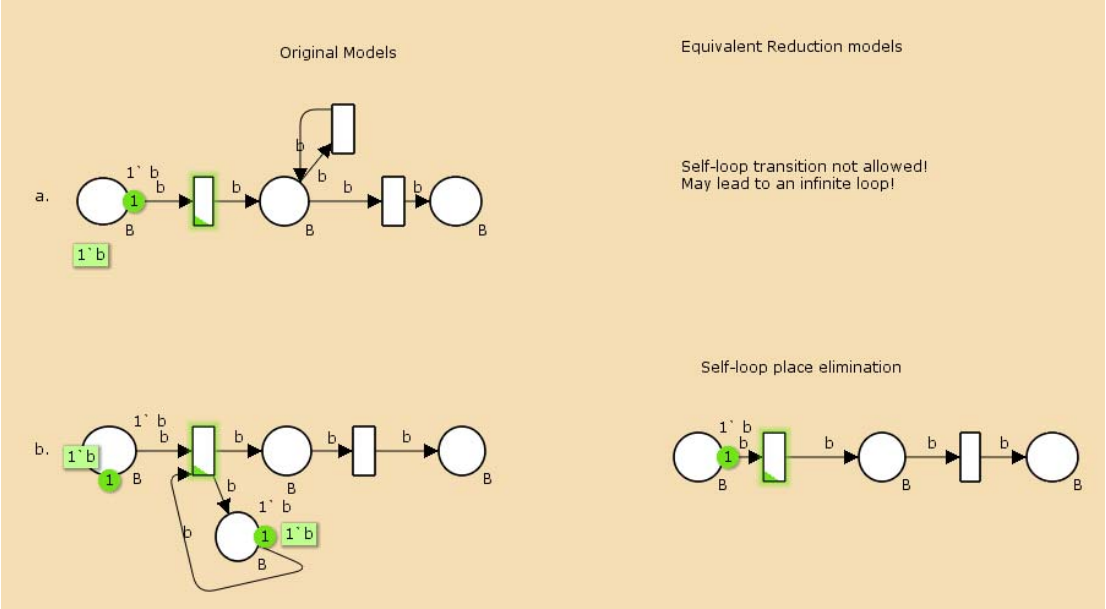

Figure 6.6: Transformation that allows self-loop place reduction preserving liveness and boundedness

Proof- A token k that appears in place P_6, could have taken any of the following paths:

$$\begin{aligned} &P_1T_1P_4T_4P_5T_5 \quad \text{OR} \\ &P_2T_2P_4T_4P_5T_5 \quad \text{OR} \\ &P_3T_3P_5T_5 \end{aligned} \tag{6.1}$$

Reduction transformation for prototype I in figure 6.7 replaces the three transition directed cascades with three equivalent transitions and merges inputs entering system at different points of time via logical *OR*. The firing time of the equivalent transitions is equal to the firing time of transition T_5.

$$\begin{aligned} T_{eq1} &= T_1P_4T_4P_5T_5 \\ T_{eq2} &= T_2P_4T_4P_5T_5 \\ T_{eq3} &= T_3P_5T_5 \end{aligned} \tag{6.2}$$

which results in the reduced model on the right hand side of figure 6.7. A token k that appears in place P_{6*}, could have taken any of the following paths:

$$P_1 T_{eq1} \quad \text{OR}$$
$$P_2 T_{eq2} \quad \text{OR}$$
$$P_3 T_{eq3} \quad \text{OR} \tag{6.3}$$

Replacing equations 6.2 for T_{eqi} in 6.3 results in the same paths as 6.1 of the original

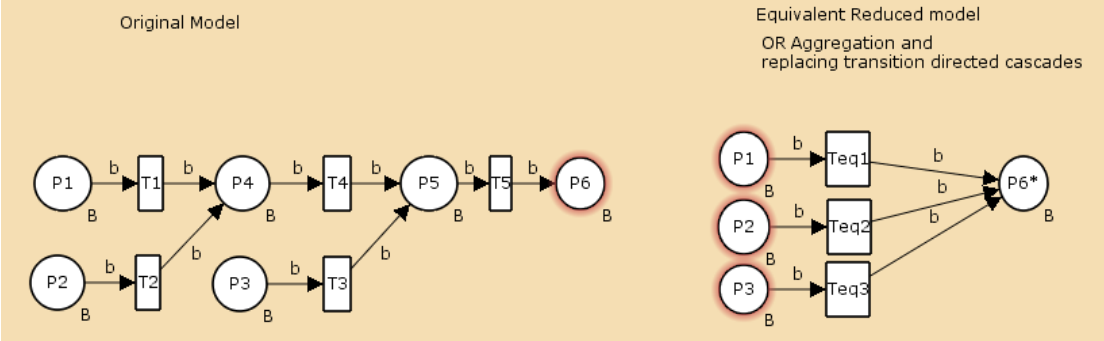

Figure 6.7: Reduction Transformation for Petri net Prototype I

model. Therefore based on definition 6.6, the two models are equivalent. The proof to show that the two models are equivalent for a logical *OR* split where all the arcs have reversed direction is similar.

An emergency response application of prototype I is the entry to the hot zone which follows a logical *OR*. Depending on the strategy, specialized forces and law enforcement groups can enter the site only or they can additionally take medical and HAZMAT in as well. This is represented by the required number of input places as demonstrated above. This prototype is scalable to any larger number of inputs preserving the merge *OR* structure. Then each organization in the entry team follows their intra-organizational protocol to complete the process before declaring the hot zone safe. If a person with higher authorities who is able to make changes to the response plan in real time, enforces another condition before an event takes place, the reduction transformation of this prototype does not work. The only simplification possible is replacement of a smaller transition directed cascade by an equivalent transition. Such a scenario is illustrated in figure 6.8a. Furthermore if this additional input creates a conflict, the reduction

is not applicable as illustrated in figure 6.8b. In both examples, the limiting structure is the introduction of synchronization (figure 6.8.a) or the addition of a conflict (figure 6.8.b). A token k that appears in place P_6 of figure a, could have taken any of the paths presented in equation 6.4

$$\begin{aligned}
&(P_1T_1P_4 \wedge P_7)\,T_4P_5T_5 \quad \text{OR} \\
&(P_2T_2P_4 \wedge P_7)\,T_4P_5T_5 \quad \text{OR} \\
&P_3T_3P_5T_5
\end{aligned} \tag{6.4}$$

and a token k that appears in place P_{6*} of figure b, could have taken any of the paths presented in equation 6.5

$$\begin{aligned}
&(P_1T_1P_4 \wedge P_7)\,T_4P_5T_5 \quad \text{OR} \\
&(P_2T_2P_4 \wedge P_7)\,T_4P_5T_5 \quad \text{OR} \\
&P_3T_3P_5T_5 \quad \text{OR} \\
&P_7T_7P_5T_5
\end{aligned} \tag{6.5}$$

For the Petri net of figure 6.8.b, the structure is not scalable as the conflicting situation impacts the merge *OR* such that a new path can cancel the other and the logical *OR* does not hold true. If T_4 and T_7 are both enabled, firing of one disables the other one and hence decomposition of different paths is not possible as the conflict brings dependency and impacts the protocol.

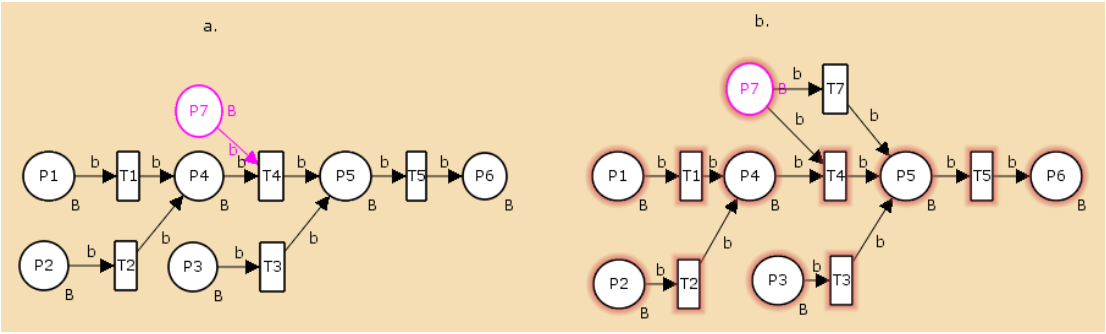

Figure 6.8: Alternative Structures that do not allow reduction of Prototype I

Theorem 6.3 *Prototype II represents a transformation that reduces a Petri net by aggregating the logical* AND *and replacement of a directed transition cascade with an equivalent transition.*

Prototype II in figure 6.9 illustrates replacement of parallel places between two Transitions (an *AND* combination of places) in a transition directed cascade. A token k that appears in place P_{n+1}, is obtained by the logical *AND* combinations of all places between two transitions:

$$P_1 T_1 \left(P_2 \wedge P_3 \wedge ... \wedge P_n\right) T_2$$

The logical *AND* of several places can be represented by one place:

$$P_{im} = P_2 \wedge P_3 \wedge ... \wedge P_n$$

and hence the path the token takes is summarized as:

$$P_1 T_1 P_{im} T_2 \tag{6.6}$$

The directed transition cascade in the reduced model is represented by $T_{eq} = T_1 P_{im} T_2$.

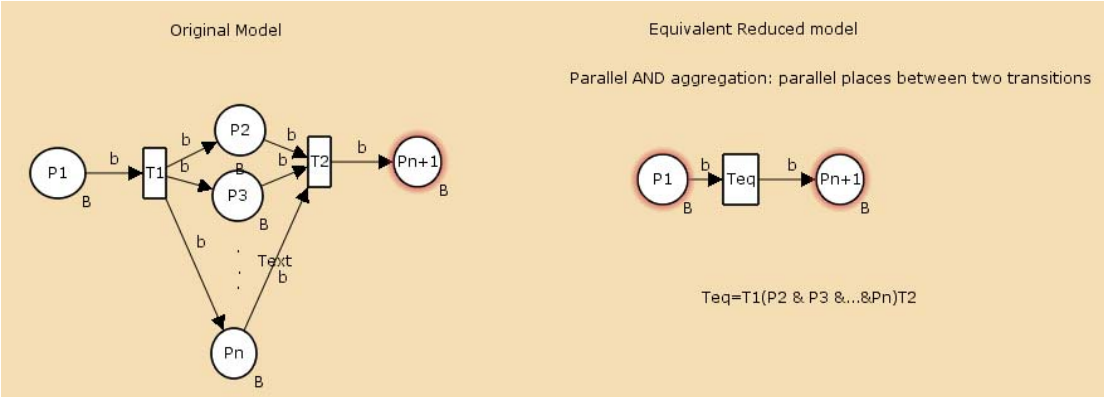

Figure 6.9: Reduction Transformation for Petri net Prototype II

Therefore a token k that appears in place P_{n+1} in the reduced model, has taken the path $P_1 T_{eq}$ which is equivalent to the equation 6.6, and hence the two models are equivalent.

Theorem 6.4 proves that property of boundedness is not necessarily preserved in reduction in a logical *OR* structure because replacing a directed transition cascade with

one equivalent transition changes the firing sequence. However the boundedness property remains unchanged for the logical *AND* aggregation as it models synchronization.

Theorem 6.4 *The property of boundedness is not preserved in the logical* OR *reduction transformation of prototype I and it is preserved in the logical* AND *reduction transformation of prototype II.*

Proof- For prototype I, proof is achieved by a counterexample. If $t_{T_4} < t_{T_5} < t_{T_3}$, place P_6 is a safe place, i.e. bounded to $n = 1$. However in the reduced model, the firing time of all equivalent transitions are the same and equal to t_{T_5} and therefore for an initial marking of $(1, 1, 1)$, place P_{6*} gets three tokens and it is not bounded to the same number, i.e. $n = 1$. For prototype II, the number of arcs from transition T_2 to place P_{n+1} in the original model is the same as the number of arcs from transition T_{eq} to place P_{n+1} in the reduced model and therefore the boundedness property is preserved. Places P_2 through P_n get the same number of tokens at the same time each time transition T_1 fires and all these enable transition T_2. When transition T_2 fires, it places one token in place P_{n+1} in both original and reduced models.

Theorem 6.5 *The property of liveness is preserved in the reduction transformations presented in theorem prototypes I and II.*

Proof- For prototype I, each equivalent transitions, T_{eq1} and T_{eq2} and T_{eq3}, replaces a transition directed cascade without changing the number of incoming arcs to the transition. Hence the equivalent transitions inherit the *liveness* property from the original model and the reduced model preserves the liveness property. Based on definition 6.4, the time at which T_{eqi} fires is equal to the firing time of transition T_5.

$$t_{T_{eq1}} = t_{T_{eq2}} = t_{T_{eq3}} = t_{T_5}$$

For prototype II, the transitions remain unchanged after reduction transformation and therefore the *liveness* property is preserved. The time at which a token enables T_{eq}, is equal to the time that the slowest place among place P_2 through place P_n takes to complete the processing of token in the original model.

The prototypes presented so far were conflict free. State machines are also examples of Petri nets that allow conflicts and this is when a place can enable more than one transition but firing of one disables the other transition.

Based on definition 5.14, a free choice Petri net, either allows synchronization or conflicts (not both at the same time). This is illustrated in figure 6.10. Marked graph and state machine classes of Petri nets both are free choice nets.

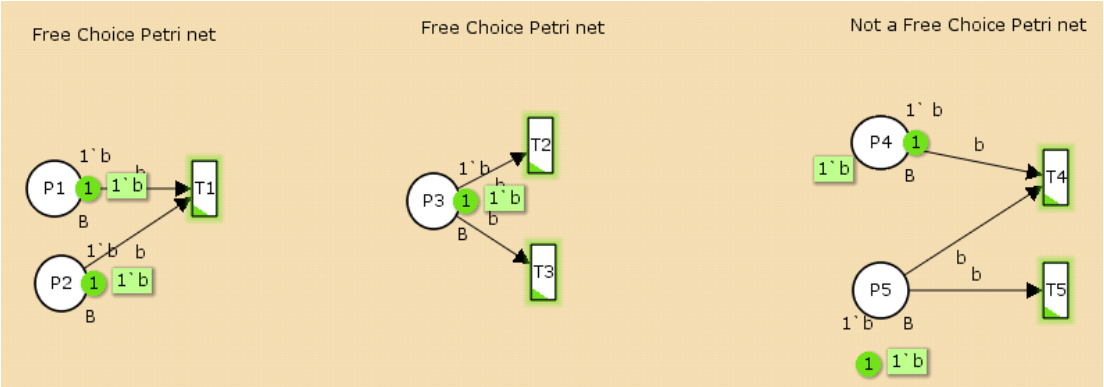

Figure 6.10: A Free Choice Petri net Illustration

Remark 6.4 *The complexity of free choice Petri nets is reducible.*

Remark 6.5 *At special cases, presented in section 6.3.2, conflict free property is required to preserve safeness or the exact execution order.*

6.3.2 Special Requirements of Complexity Reduction for Petri nets with Conflicts

A place that is input to two transitions creates a conflict. For example place P_5 of figure 6.11 shows a Petri net where an initial marking of one token for places P_1 and P_2, creates a conflict (or decision) on which transition to fire, i.e. T_5 or T_6. Both transitions are enabled but firing of one disables the other enabled transition. If there is more than one token at the common input place, a priority mechanism can be introduced to ensure fairness such that the transitions fire alternatively or any other design required priority.

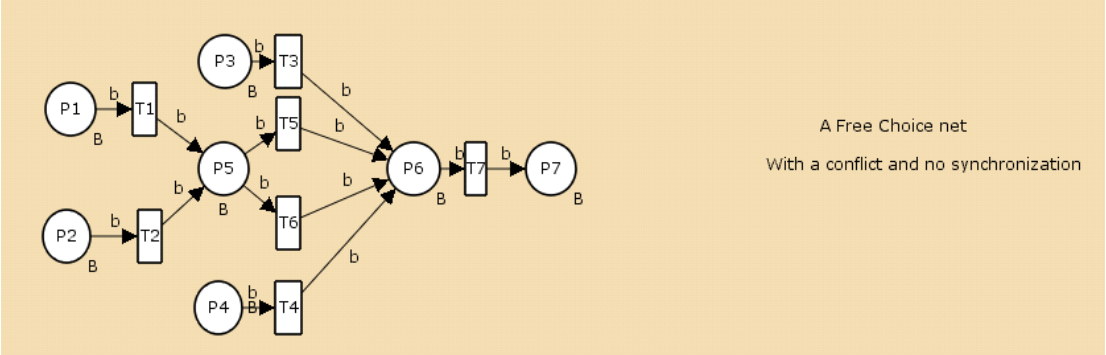

Figure 6.11: A Petri net with a Conflict

In the logical *OR* merge of figure 6.12, there are two input transitions, T_2 and T_3, for one place, P_5. If the sink place represents the termination of a process, there will be more than one termination time as a result of the logical *AND* merge which places tokens in both preceding places. The first transition that fires, sends a token to the sink place and when the other enabled transition fires, it sends another token to the sink place and creates a second termination time. This creates ambiguity on the initiation of another task which is dependent on precise termination of this process. A real scenario where lack of a clear termination causes problems is when the entry team (tactical law enforcement and special forces) are to declare the hot zone safe for the rest of the responders to enter the site. Communication has failed in such scenarios where an unclear signal has either delayed the entry or has put the other first responders' team in danger. If instead the model is reduced to the one shown on the right hand side of figure 6.12 where the output of place P_3 connects to transition T_2, the task completes only when both teams have completed their processes and hence the problem of unclear termination is removed. This model illustrates *a reduction that corrects the structure* of the Petri net model where an *AND* split is followed by an *AND* merge.

Attention needs to be paid in reduction models that are not decision free. In replacement of a directed transition or place cascade by an equivalent transition or place, the order of real execution of tasks or boundedness property or safeness in special case might be changed [43]. Petri net models representing such scenarios are discussed next. Example 6.3 illustrates a case where the reduction does not preserve the firing sequence.

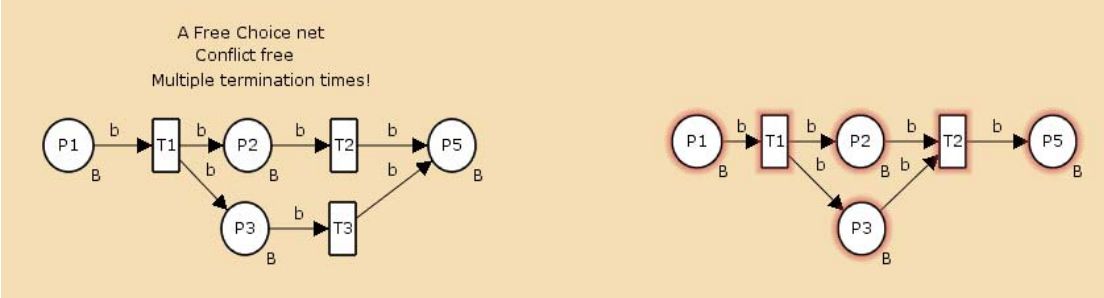

Figure 6.12: A Petri net with Multiple Termination Time

Example 6.4 presents a scenario where the safeness property does not hold true after the reduction.

Example 6.3

Figure 6.13 shows the original model on the left and the reduced model on the right. Figure 6.13.b is the reduced model replacing a transition directed cascade of $T_3P_4T_4$ by T_{4b} which does not preserve the original firing sequence. The execution time of T_{4b} is larger than T_{1b}, and hence transition T_{1b} fires first which hides the original firing time of the first transition in each path which indicates the same start time for both choices. In figure 6.13.a, T_1 and T_3 become enabled at the same time. If there is only one token in place P_1, a decision needs to be made while both transitions can fire at the same time. However the fusion of transition directed cascade in figure b into one transition, T_{4b}, does not preserve the firing sequence and the equal choice in start time of each process.

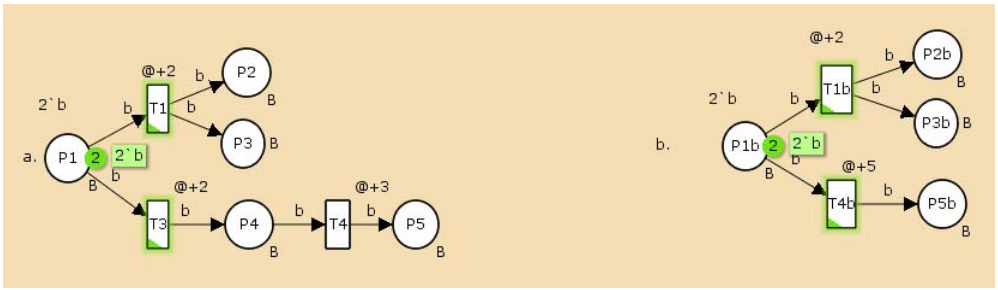

Figure 6.13: The reduced Petri Net model that does not preserve the firing sequence

Example 6.4 *Figure 6.14 illustrates that the reduction technique does not preserve the safeness property for place* P_5.

In figure 6.14.a, the execution time of transitions hold the following relationship:

$$\begin{aligned} t_{T_3} &\leq t_{T_2} \\ t_{T_2} &\geq t_{T_4} \\ t_{T_3} &\leq t_{T_4} \end{aligned} \tag{6.7}$$

and as a result place P_5 is safe. T_4 consumes the token from P_5 generated by T_3 before transition T_2 adds another token into place P_5. However in figure b, T_{5eq}'s firing time is greater than both T_{2b}'s and T_{3b}'s firing time and therefore they both place a token in place P_5. Hence this place is not safe and the property is not preserved in reduction when the original Petri net model is not decision free. A Petri net model has to be decision

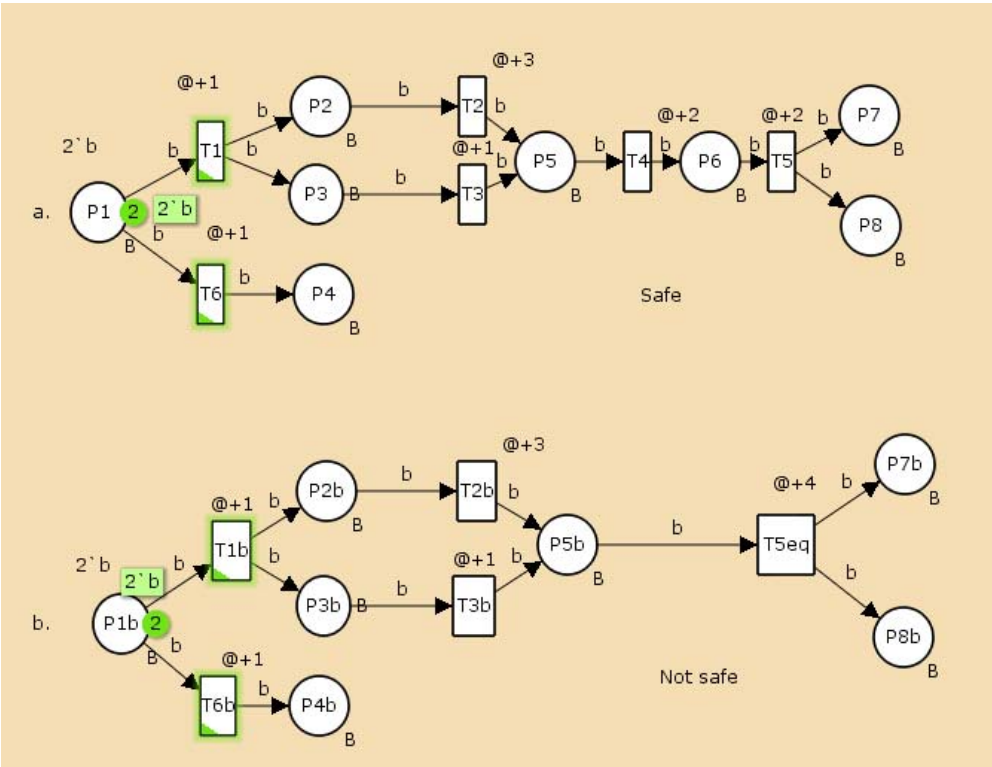

Figure 6.14: The equivalent Petri Net model that does not preserve safeness

free to ensure equivalence in preserving firing sequence, safeness or boundedness prop-

erty. Reduction techniques are helpful to reduce the complexity of the analysis of large Petri nets and are applicable to decision free models directly while for non-decision free models, properties mentioned above need to be checked first before replacement.

All detailed information of the lower levels produce a large amount of information to process which may not be available to a high-level supervisory or not practical to be transferred and processed. A higher-level supervisory controller does not need to know all the details if it can rely on lower level controllers to verify selective criteria and pass on necessary information only. Abstracting detailed information provides efficiency in faster analysis and real-time decision making.

Example 6.5 *The Petri net model in figure 6.15, represents two organizations communicating with each other. The reduced Petri net model is discussed below.*

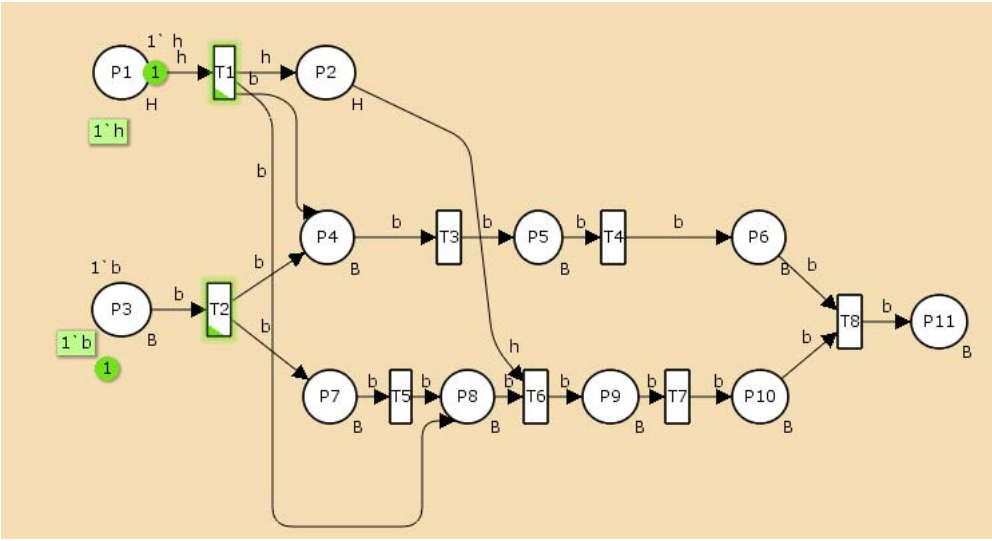

Figure 6.15: A large Petri Net model

A sample firing sequence for figure 6.15 is equal to $\psi = \{T_1T_6T_3T_7T_4T_8T_2T_5T_3T_4\}$. For any given firing sequence like ψ, the final marking achieved from the initial marking of one token at places P_1 and P_3 is equal to one token at places P_6, P_8 and P_{11}. The final marking for figure 6.16 is the same, one token at places P_{4eq}, P_{8eq}, and P_{11}, and hence two models are equivalent. The two arcs in the original model from T_1 and P_2 to P_8 and

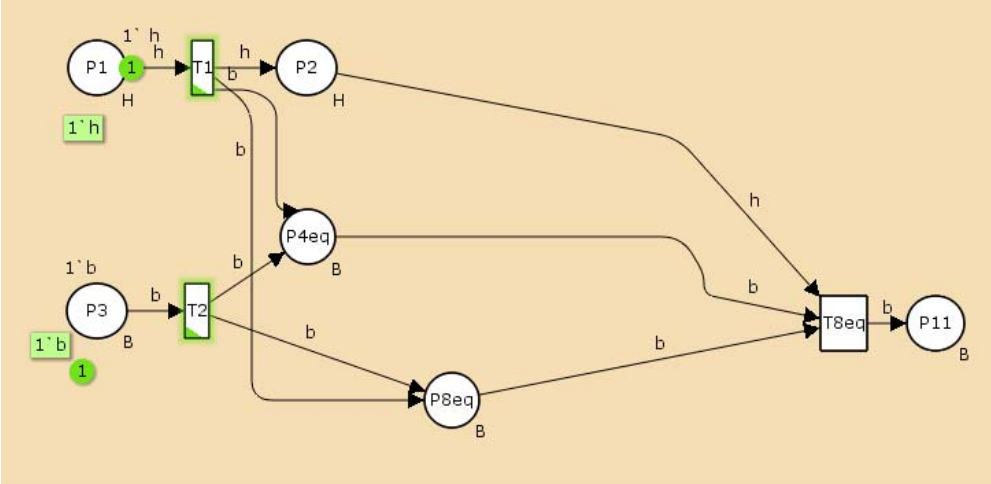

Figure 6.16: The equivalent Petri Net model that preserves boundedness and liveness

T_6 cause redundancy and therefore we can get rid of the outgoing arc from place P_2 in the reduced model as a token from T_1 always flows to P_2. Transition T_{8eq} inherits its liveness property from T_8. The difference between the arcs from P_{4eq} and P_{8eq} in figure 6.16 and P_6 and P_{10} of figure 6.15 is a directed cascade which does not change the liveness property. When transition T_1 in the original model fires, it puts one token on all its outgoing arcs including the one to place P_2. Instead of taking this token to transition T_6, the consumption of the token is delayed and it is directly taken to transition T_{8eq}. Hence, the incoming arcs to this transition preserve the liveness property.

The number of incoming arcs to P_{4eq} and P_{8eq} remain unchanged and therefore the boundedness property is preserved. Firing time of transition, $t_{T_{8eq}}$, is set to the *maximum* of firing time of T_4 and T_7. The reduction transformation presented here demonstrates communication complexity reduction technique which is presented in next section.

6.3.3 Communication Complexity Reduction Recommendations

Remark 6.6 *If two organizations are entangled in a particular way such that it creates redundancy, the model can be simplified with respect to the source of redundancy.*

Remark 6.7 *If interactions between two organizations cause redundant dependencies at several points of contact, the model can be simplified by reducing the dependency to*

one point only.

Remark 6.8 *We propose to delay dependencies to the last possible point of contact to allow each individual organization to progress independently for as long as possible. However, if this conflicts with system requirements, back propagation tracing algorithm backs up to the previous point of connection and continue till a valid solution is achieved.*

6.4 Deadlock Free Control Design

Theorem 6.6 *A Petri net model which is both a marked graph and a work flow net is deadlock free.*

Proof- The following conditions guarantee the Petri net is deadlock free:

a. There is only one sink place in a given Petri net model: the property of work flow net guarantees the existence of one sink place in the model and therefore the task continues for completion without improper termination of the process by a token ending at another sink place.

b. There is no conflict: a conflict in the model means at least two transitions become enabled and firing of one, disables the other one. This happens when there is a choice or decision to be made in the model which is when a place has two outgoing arcs to two transitions. A marked graph by definition is conflict free and all places have only one incoming and one outgoing arc.

c. There is no dangling tasks or conditions in the model: this means the execution of a task starts from the source place and ends at sink place and all routes contribute to the completion of task. This is guaranteed by work flow net property.

Corollary 6.1 *A Petri net that is both a marked graph and a work flow net does not have an empty siphon and is self-loop free.*

Proof- In a marked graph, all places have only one input and one output with the exception of source and sink place. The only way that a self loop can exist is as the sink place and that conflicts with the definition of a sink place and the existence of only one

sink place in a work flow net. If a Petri net is both a marked graph and a work flow net, according to theorem 6.6, the Petri net model is deadlock free. Based on remark 5.3, since there is no deadlock, there is no empty siphon.

Remark 6.9 *If the work flow net is sound, the Petri net is deadlock free by definition 5.17.*

Remark 6.10 *The non-existence of an empty siphon does not guarantee deadlock free because there may exist a dead transition leading to a place which does not belong to any set of siphons.*

Remark 6.11 *A set of place invariant is both a trap and siphon but the siphon never gets empty.*

For a set of Place invariants, weighted sum of tokens is constant. When a siphon gets empty, it remains empty. In order to be a trap, it has to have at least one live transition which means there has to be at least one token. Hence, such a set never gets empty to hold the property of place invariant (fixed number of tokens). This implies that a set of place invariant never has an empty siphon. Still the structural property of siphon holds true, i.e. ${}^\bullet P \subseteq P^\bullet$. It is worth noticing that a set of place invariants is not even initially a siphon. The catch with the property of place invariant set is that not only the set never gets empty, but also the total number of tokens stay constant. This is important as empty siphons indicate deadlocks.

Note that this is a one way conclusion, i.e. empty siphons indicate deadlocks but a judgement about the existence of deadlock in a Petri net cannot be given if the siphons are non-empty.

Remark 6.12 *A trap that never becomes marked indicate the existence of at least one dead transitions.*

Based on definition, a trap remains marked once it becomes marked. If a trap never becomes marked, that confirms the existence of at least one dead transition which never fires and hence no token flows to the set of places that form the trap.

Remark 6.13 *If a set of places is both a trap and siphon, the set is not necessarily a place invariant set.*

For a set of places that are both siphon and trap, there is no guarantee on the fixed number of tokens in the set, and therefore, no conclusion can be derived on the place invariance property.

The process of patients' admission and treatment at UCSD emergency medical department was explained in section 2.7 and the Petri net model developed to predict and analyze the performance is illustrated in figure 6.17 [35]. There are several instances of synchronization in this model when the patient and a resource must be present to proceed. This Petri net model has been developed based on the proposed structural prototypes to reduce dependencies when communication to other organizations is needed. The structural properties and theorems have been directly applied to expedite deadlock analysis and to enhance system performance

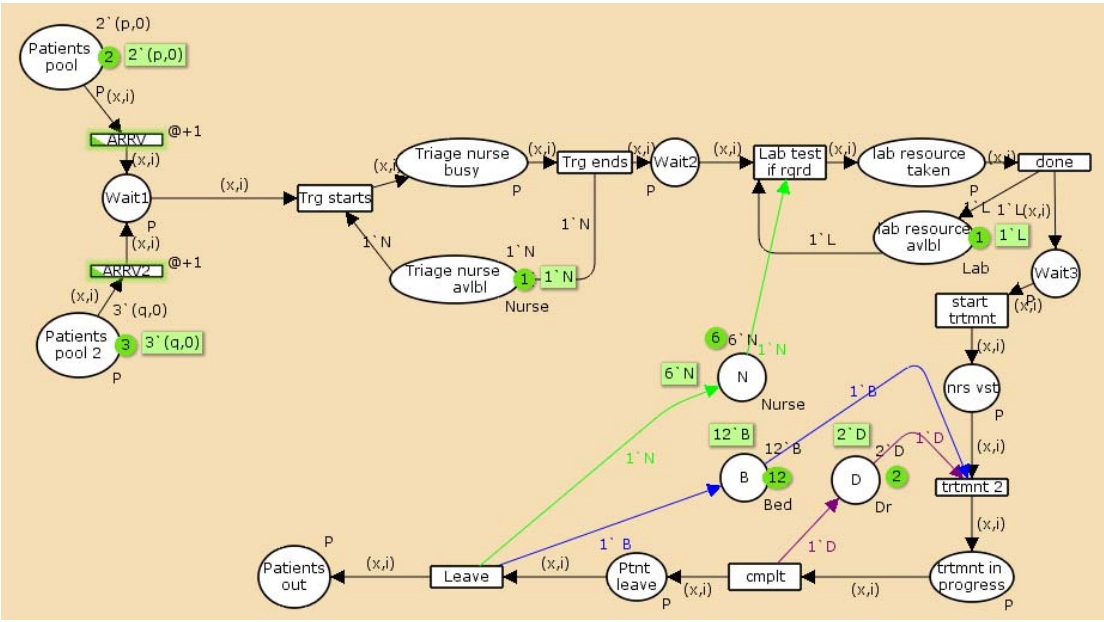

Figure 6.17: UCSD Emergency Medical Department

This Petri net holds the following properties:

- This Petri net model is a *Marked graph* Petri net because all places have only one incoming and one outgoing arc.

- This model is *conflict free* which by definition is the property of a marked graph, deducted from definition 5.13.

- This model is a *work flow Petri net* because the two input places with two different types of tokens can be represented in one place with both token types. There is only one sink place and each transition or place is on a path from input to the output.

- This model is *deadlock free* according to the theorem 6.6.

The structure of Petri net of figure 6.17 prototypes the system modeling of one organization. This modeling style is suitable to analyze hierarchical organizational structure where communication among different systems are the main focus. Each system is abstracted with its inputs and output where details of interim tasks can be hidden without loss of generality.

6.5 Conditions for Liveness of Special Petri nets with Application to ICS or MMST Communication Protocol

Interactions among different organizations is modeled into Petri nets with different topologies. Depending on the topology of the model following ICS organizational chart, analysis of the liveness property of the larger model can be simplified if each Petri net model holds specific properties and follows certain topologies. This is a methodology that reduces the complexity of the system by divide and conquer method and makes it appropriate for similar communication protocols. Below we present necessary conditions for liveness of specific topologies agreeing to ICS communication protocol. We present supporting theorems that help to deduct a conclusion on liveness property of large Petri net models holding certain topologies.

- Cascade- no feedback: in a cascade topology the output place (sink) of each system connects to the input place (source) of the next system through a connecting transition. The output place of the last system does not connect to the source place

of the first model and therefore this structure has one fewer connecting arc than the ring model. This is illustrated in figure 6.18. Each organization can have its own Petri net model and the communication among organizations is provided by connecting the appropriate input and output ports. There are two additional places, $P1to2$ and $P2to3$ to provide interaction between any two organizations.

Theorem 6.7 *If the Petri net model of each individual organization is a marked graph and work flow net and they interact in a cascade topology, the result keeps both properties of marked graph and work flow net, and therefore it is deadlock free.*

Proof- The proof relies on the topology of the new model. Since each system is connected to the other one via a transition, the larger model still has one input place, one output place, each place has at most one input and one output transition and transition or place contributes to the completion of task. Therefore the larger model is still a work flow net and marked graph. Hence the model is deadlock free according to theorem 6.6.

A real-life application to such structure is a sequential process: in an ICS structure, HAZMAT has to wait for special forces and tactical law enforcement teams to declare the area safe before entering the hot zone. The output provided by them is fed as input to the HAZMAT system. Another example is passing through decontamination process (DECON): as each victim leaves the decontamination process, a signal feeds the next victim into the process. As the victim leaves the system, next victim enters the system. Note that this is not a cyclic process. Both of these scenarios are examples of cascades topology and where the model of each organization is deadlock free and the interacting model remains deadlock free.

- Ring: one more transition connecting the output place of the last organization to the input place of the first one, turns a cascade to a ring. Since the source does not exist any more, the resulting model is not a work flow net but keeps the property of marked graph.

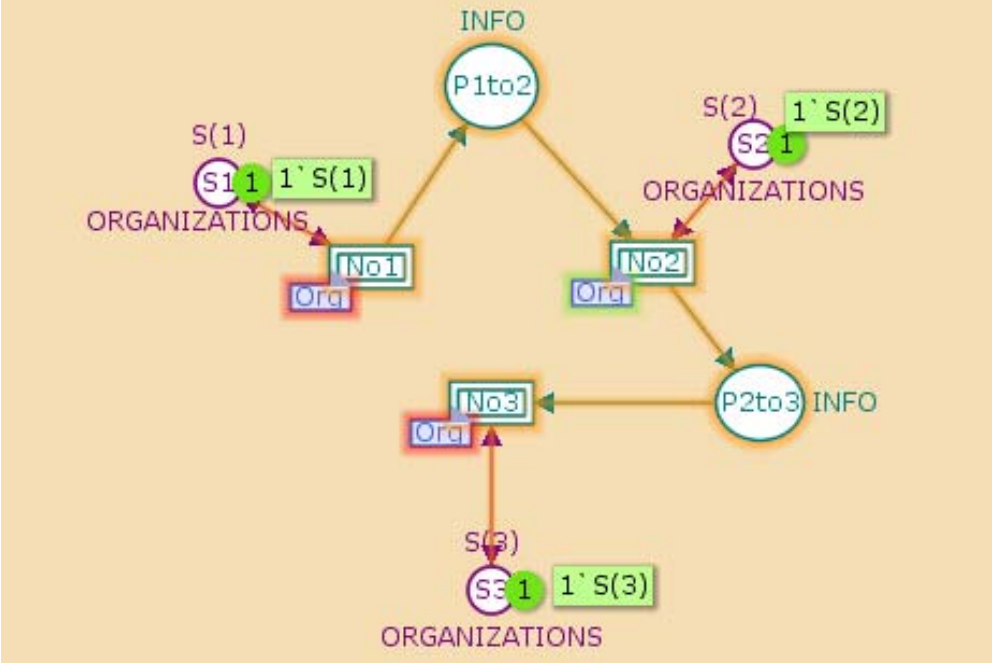

Figure 6.18: A Petri Net model with Cascade topology

Remark 6.14 *The only addition in the ring model compared to the cascade is a transition plus an incoming arc to this transition and an outgoing arc from it. Since the rest of the net is live, the transition with one incoming and one outgoing arc is live.*

This is a cyclic process. It is like the process of updating information and data in a ring topology with order. The input of each organization comes from the output of the previous organization and this is repeated among all participants keeping the order. This is similar to cascade topology with the difference of the last one feeding the first organization. The data can be updated and shared among them in a particular order as changes are required by each organization before arriving at next destination.

In an emergency response application, police calls for special forces and they arrive at the site. They enter the hot zone, clear the area, transfer victims, declare the area safe and allow HAZMAT to come in. HAZMAT performs their duties and the output is fed back to Police.

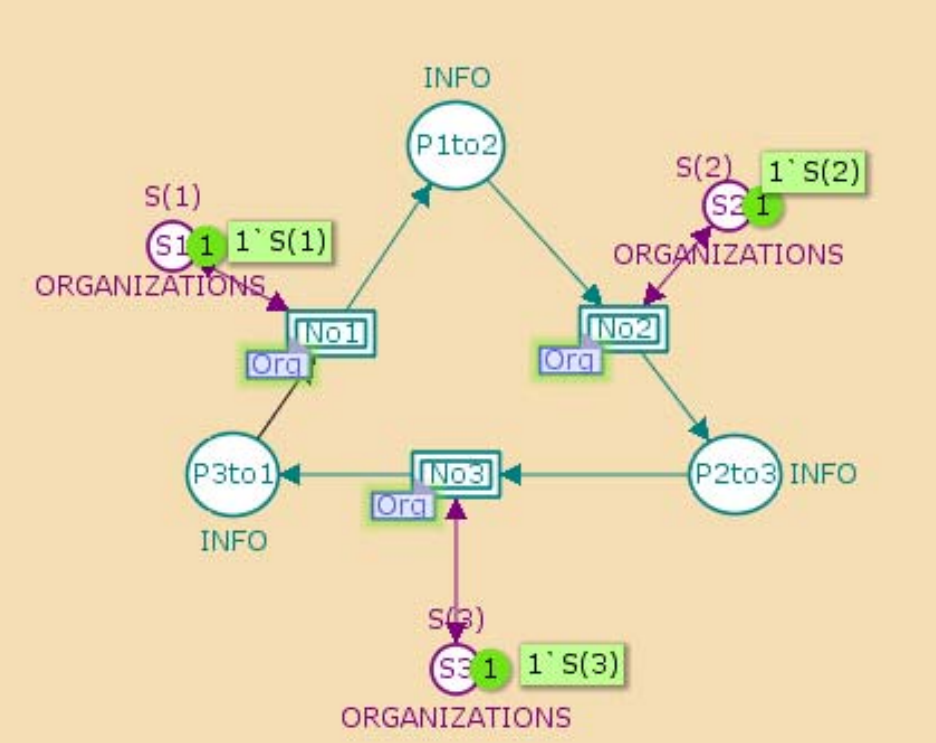

Figure 6.19: A Petri Net model with Ring topology

- Star: an application to this case is the 1 to n communication among a supervisor and several team leaders or among team leader and the team. Another example is subscription to the same set of data when there are several members in a system accessing data. The data is updated and shared among them in a push-pull method while a central supervisor has to maintain the latest copy at all times. The latter corresponds to scenario 4.6 presented in chapter 4. The star topology of figure 6.20 consists of a place and a transition to establish communication. The Petri net model for an individual organization is shown in figure 6.21 and the result of completion of tasks is presented in figure 6.22.

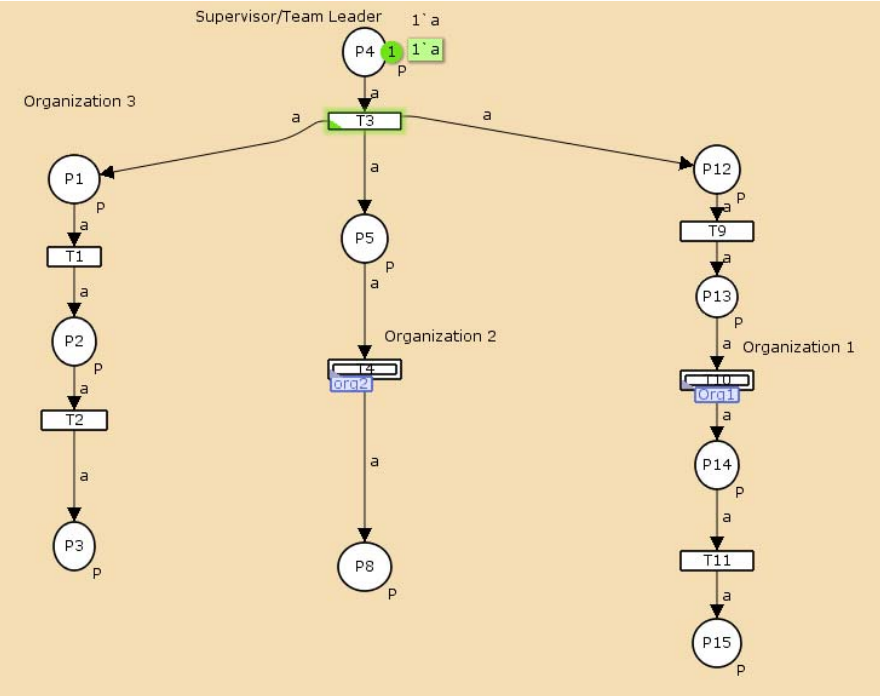

Figure 6.20: A Petri Net model with Star topology

Transition T_3 is connected to each team member or lower-hierarchy team leader's place. The star topology inherits its liveness from the liveness of transition T_3. If this transition is live and each individual model is a marked graph and work flow net, then each line of communication forms a cascade and according to theorem

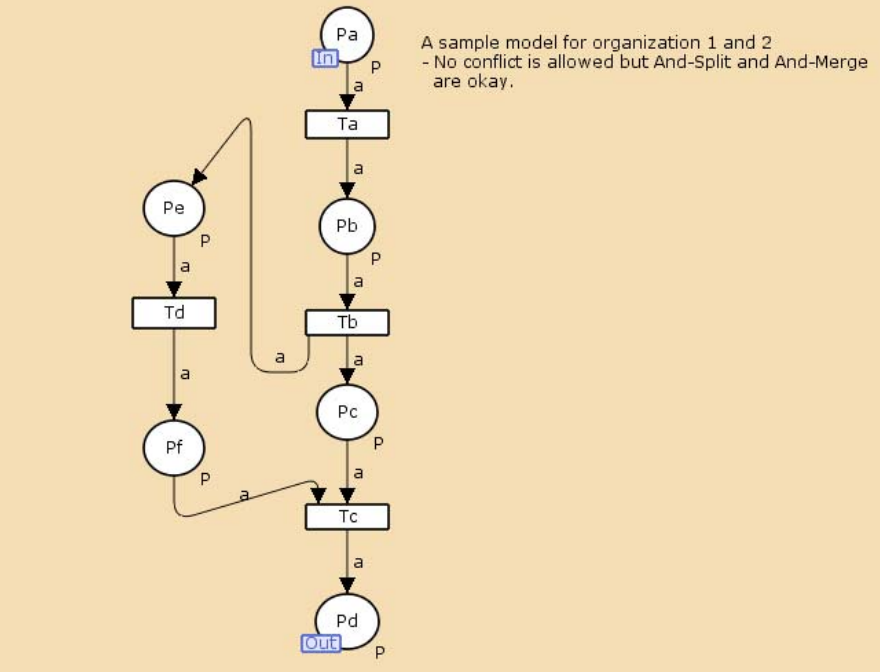

Figure 6.21: The Petri Net model of individual organization for figure 6.20

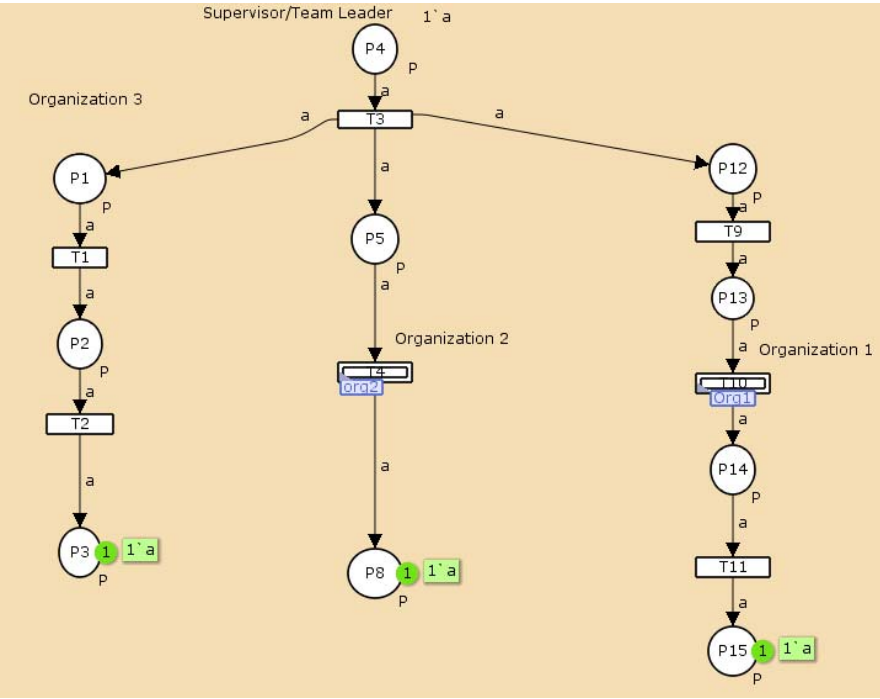

Figure 6.22: The Petri Net model of figure 6.20 upon completion

6.7 the system is deadlock free .

- Cascade or star- with feedback or forward interactions:
    1. from a place to a transition
    2. from a transition to a place

For such a general structure like presented in figure 6.23, the model needs to be examined even if each individual system is a marked graph and work flow net.

An arc from a place further down in the cascade, to a transition in an earlier process creates deadlock! The place is located after a transition which is not enabled yet! Liveness property can be preserved: if there is no dead transition in each Petri net model, and if the interacting arc is valid, the new interacting model preserves the liveness property. Hence valid connections are in the forward direction: an arc from a place in the preceding model to a transition following further down in the or from an arc from a transition which fires prior to a preceding place.

An interactive scenario is presented in figure 6.23 and the results of completion of process is shown in figure 6.24.

This figure presents a problem with information flow that needs to be revised and monitored by controllers. The connecting arc from place P_{14} to transition T_{10} creates a deadlock and hence needs to be removed. Otherwise there will never be a task executed by organization one. In some cases, organization one needs to transfer the job to organization two. In such scenarios, a controller place can remove its token from the input of transition T_{10} and the job is transferred to organization two.

6.6 Concluding Remarks

In this chapter we used Petri nets to represent an event-driven model of a large complex system. This allows us to analyze structural properties and reachability of states. We developed structural prototypes and theorems that help us to analyze the structural behavior of systems such as deadlock, synchronization, and conflict. We

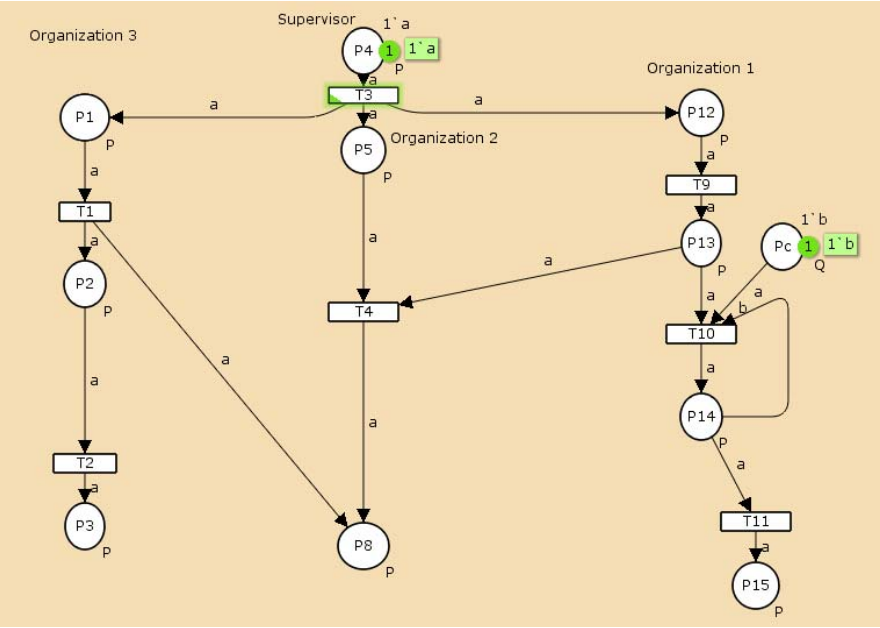

Figure 6.23: A Petri Net model with Star topology and additional interactions

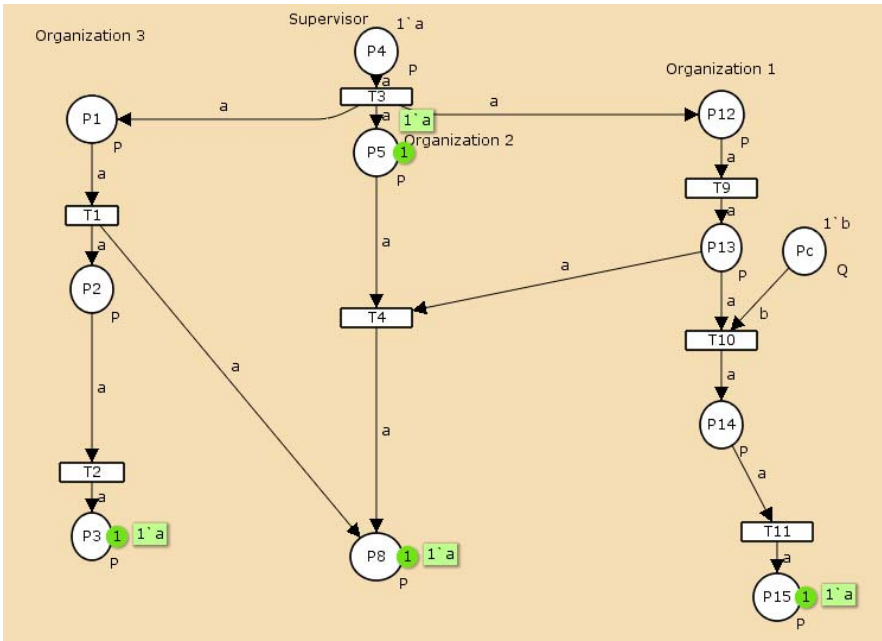

Figure 6.24: The Petri Net model of figure 6.23 upon completion

presented transformations to reduce the structural complexity of large Petri nets and investigated whether the properties like deadlock free, boundedness and safeness are preserved within or not. We contributed to the field of emergency response by making several recommendations on communication protocol and system design and techniques to reduce entanglement and dependencies among several organizations. Performance analysis based on behavioral properties of Petri net will be presented in next chapter.

This chapter, in part, is a reprint of the material of the following paper: R.B. Dilmaghani, and R.R. Rao,“A Systematic Approach to Improve Communication for Emergency Response,”Hawaii International Conference on System Sciences (HICSS), January 2009. This helps the supervisor to cancel an already scheduled path depending on the costs

Chapter 7

Behavioral Properties and Communication Complexity Analysis

7.1 Introduction

So far we have considered Petri nets with deterministic time delays for transitions, places and the arcs. In deterministic timed Petri nets, transitions fire when their associated time delay has been reached and the incoming tokens become available. The reachability tree of a Petri net can be transformed into a state transition diagram where the states are the markings of the Petri net and connecting arcs represent costs. In this chapter, we investigate the behavioral performance of the model rather than the events deriving the states.

Not all real systems have deterministic time delays. When uncertainty exists with execution time of events, a probability distribution function is associated to the firing of the transition which turns the timed Petri net into Stochastic timed Petri nets (STPN). In a stochastic timed Petri net the probability of the next transition to be fired is determined by its probability distribution function. The firing time of transitions are typically modeled by mutually independent random variables with negative exponential distribution. The memoryless property of the distribution makes the marking of the stochastic Petri net a continuous Markov chain [55]. In stochastic timed Petri nets there is no actual conflict since the probability of next transition to be fired is determined by

its probability distribution function.

In section 7.2 we present a framework that allows to control events at detailed level based on the state transition diagram derived from the reachability tree. In section 7.3 we present communication complexity analysis among several systems sharing one resource or entity and finally, in section 7.4 we highlight the concluding remarks.

7.2 Framework for Behavioral Performance Analysis

A supervisor can avoid communication problems if early symptoms such as exceeding the number of tokens in a place or number of firings of a transition are detected. The number of states in a state transition diagram is equal to the number of reachable markings in the reachability tree [56]. The model reaches its final state through one of several possible paths.

A supervisor can facilitate communication and improve performance by knowing real path costs in advance. If those costs are not known, they are assumed to be equal. Since communication may break unexpectedly or an entity may suddenly become unreachable, liveness is achieved if a supervisor can enforce a new path by placing a token in a place or by preventing an event from happening by introducing a condition to prevent a path from entering deadlock. To illustrate the subject, we revisit the Petri net of example 5.2, where two organizations communicate with each other.

Example 7.1

This is applicable to a medical scenario where a nurse or a doctor visits two patients, and takes turn in providing care to them. It is also applicable to the scenarios where victims are brought to decontamination (DECON) by two different organizations such as Bomb Squads and Police. Priority for organization one has been implemented such that T_2 always fires before T_5. The state transition diagram is illustrated in figure 7.2 based on the reachability tree presented in figure 5.5. When organization 1 releases the resource, the resource (token) becomes available for organization two to execute T_5. Similarly as T_5 fires, the resource token becomes available to organization one. This

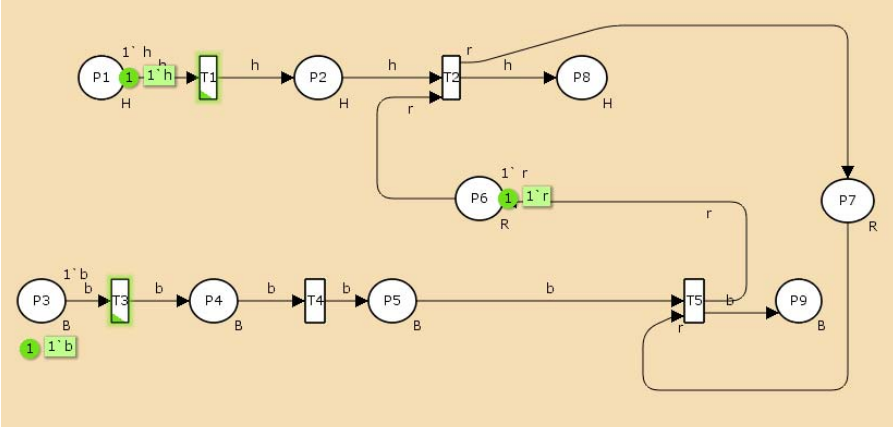

Figure 7.1: A Petri net model

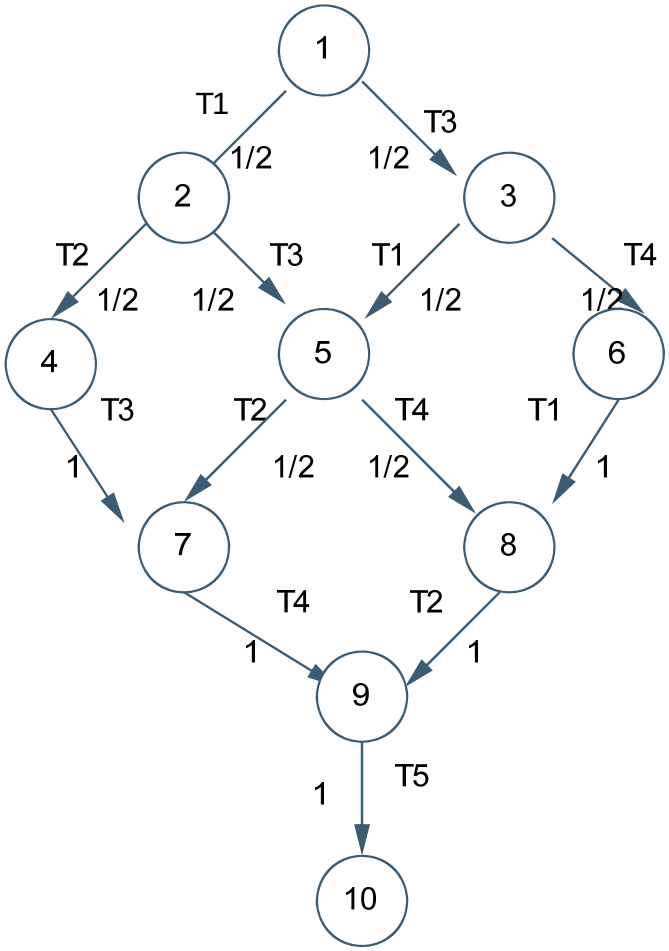

Figure 7.2: State transition diagram for the reachability tree of figure 7.1

system has one input place and one output place at each one of the parallel branches. Starting from input place the system always ends at the output place and therefore there is no dangling task or event. The transition matrix is derived from the state transition diagram of figure 7.2:

$$T_p = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The system is sound and hence it is deadlock and live lock free based on definition 5.17. The final state is reachable in five hops. The state transition diagram for this Petri net is presented in figure 7.2. This system always reaches state 7 or 8 in three hops under normal circumstances and from any of these states, the system will be absorbed in final state, i.e. state 10. The last two hops to reach final state are equal to one. If the path costs are not known in advance, they are assumed equally likely and therefore the final

cost to reach final state through any of those paths is as following:

$$\{1 \rightarrow 2 \rightarrow 4 \rightarrow 7\} = \left(\frac{1}{2} \times \frac{1}{2} \times 1\right) = \frac{1}{4}$$
$$\{1 \rightarrow 2 \rightarrow 5 \rightarrow 7\} = \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) = \frac{1}{8}$$
$$\{1 \rightarrow 2 \rightarrow 5 \rightarrow 8\} = \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) = \frac{1}{8}$$
$$\{1 \rightarrow 3 \rightarrow 5 \rightarrow 7\} = \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) = \frac{1}{8}$$
$$\{1 \rightarrow 3 \rightarrow 5 \rightarrow 8\} = \left(\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}\right) = \frac{1}{8}$$
$$\{1 \rightarrow 3 \rightarrow 6 \rightarrow 8\} = \left(\frac{1}{2} \times \frac{1}{2} \times 1\right) = \frac{1}{4}$$

The corresponding element of transition matrix to the power of *five*, $p_{1,10}$ of matrix T_p^5 is equal to 1 which verifies that the final marking is reachable in *five hops*.

$$T_p^5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

However when transition T_2 does not fire, organization two cannot proceed and this is when a supervisor can enforce liveness by preventing system from entering deadlock. The solution is illustrated in figure 7.3 where a supervisor enforces a new path by placing a token in place p_s. The supervisory control place does not have a token under

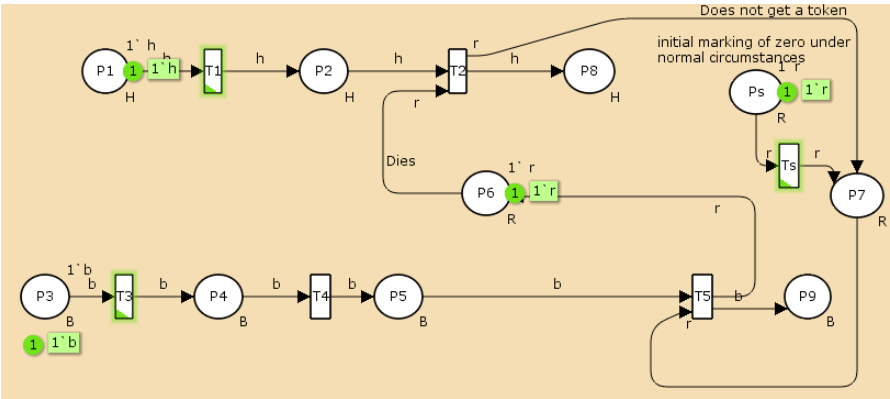

Figure 7.3: The Petri net model of example 7.1 with supervisory solution

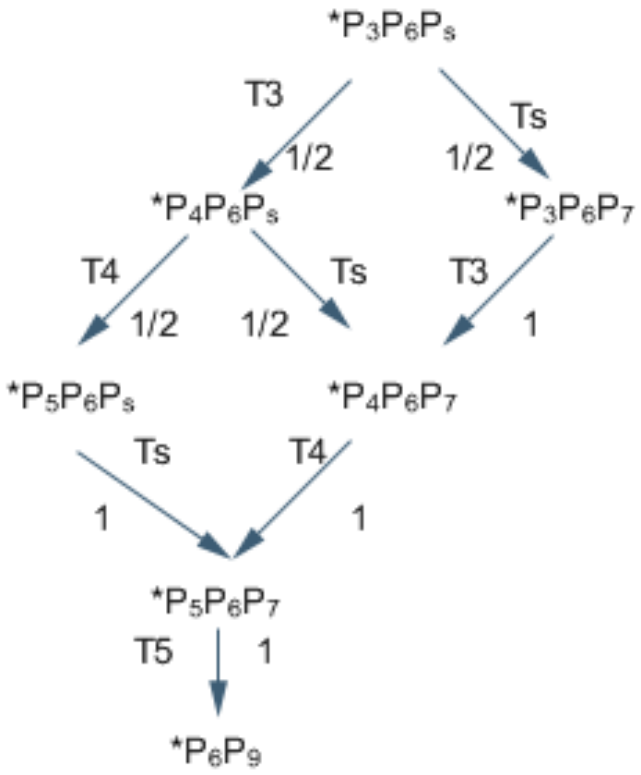

Figure 7.4: The reachability tree of the Petri net with supervisory solution

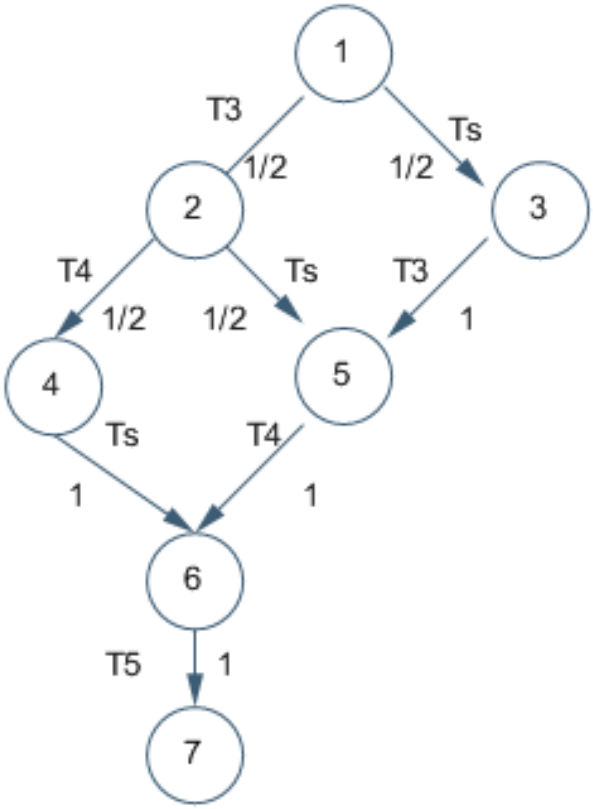

Figure 7.5: State transition diagram of the Petri net with supervisory solution

normal circumstances but when event T_2 does not fire, the supervisor places one control token there so that organization two can proceed. The reachability tree of the Petri net with supervisory solution is shown in figure 7.4 where $*$ represents place p_1 or place p_2. The state transition diagram is presented in figure 7.5. The system will be absorbed in final state in four hops:

$$\{1 \rightarrow 2 \rightarrow 4 \rightarrow 6 \rightarrow 7\} = \left(\frac{1}{2} \times \frac{1}{2} \times 1 \times 1\right) = \frac{1}{4}$$
$$\{1 \rightarrow 2 \rightarrow 5 \rightarrow 6 \rightarrow 7\} = \left(\frac{1}{2} \times \frac{1}{2} \times 1 \times 1\right) = \frac{1}{4}$$
$$\{1 \rightarrow 3 \rightarrow 5 \rightarrow 6 \rightarrow 7\} = \left(\frac{1}{2} \times 1 \times 1 \times 1\right) = \frac{1}{2}$$

To conclude the results, if supervisory event fires first, the process will be only dependent on organization two internal delays. If supervisor provides a faster response by setting the firing time of transition T_s smaller than the events in organization two, the final state is achieved in such order with a higher probability.

7.3 Communication Complexity Analysis

In an ICS organizational chart, there are several scenarios that can be mapped to a one-to-n communication such as communication among team leader and the team members or among team leaders (like HAZMAT, special forces and medical group) and *Operations*. In all these scenarios, different organizations communicate with each other and with the one entity. Below we develop transition matrix and find complexity of a one-to-n communication.

Example 7.2 *Consider a discrete event system consisting of two organizations with their own internal communication protocols and a single resource to communicate one at-a-time. Both organizations share the same clock and execute through their internal communication several times before sending a request to talk to the shared entity or resource. The single shared entity or resource can also communicate back to any of the organizations. The one shared entity or resource can be thought as the* operations.

The system behavior is modeled by a five-sate transition diagram with the following five possible states:

- Both organizations are running internal communication within themselves.
- Organization A is talking to operation and organization B is running internal communication protocol.
- Organization B is talking to operation and organization A is running internal communication protocol.
- Organization A is talking to operation and organization B is waiting to communicate to operation.
- Organization B is talking to operation and organization A is waiting to communicate to operation.

The system is memoryless as the next state only depends on the immediate previous state. Transition matrix is found as follows:

$$T_p = \begin{bmatrix} \overline{p_A}\,\overline{p_B} & \overline{p_A}\,p_B & p_A\,\overline{p_B} & 0 & p_A\,p_B \\ q_B\,\overline{p_A} & \overline{p_A}\,\overline{q_B} & p_A\,q_B & 0 & p_A\,\overline{q_B} \\ q_A\,\overline{p_B} & q_A\,p_B & \overline{q_A}\,\overline{p_B} & p_B\,\overline{q_A} & 0 \\ 0 & q_A & 0 & \overline{q_A} & 0 \\ 0 & 0 & q_B & 0 & \overline{q_B} \end{bmatrix}$$

Where p_i represents communicating to operations and q_i represents releasing the communication connection to operations. State transition diagram of the system described above is illustrated in figure 7.6.

Example 7.3 *A discrete event system consists of three organizations with their own internal communication protocols with one communication at a time possible to operation sharing one resource.*

The system takes one of the following ten states:

- All three organizations are running internal communication within themselves.
- Organization A is talking to operation and organization B and C are running internal communication protocol.

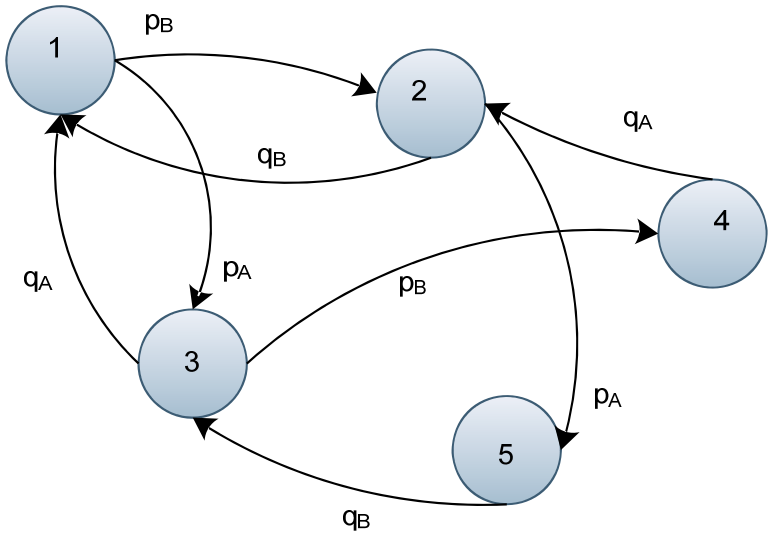

Figure 7.6: State Transition Diagram of two organizations with one Shared Resource

- Organization B is talking to operation and organization A and C are running internal communication protocol.
- Organization C is talking to operation and organization A and B are running internal communication protocol.
- Organization A is talking to operation, organization B is waiting to communicate to operation and organization C is running internal communication protocol.
- Organization A is talking to operation, organization C is waiting communicate to operation and organization B is running internal communication protocol.
- Organization B is talking to operation, organization A is waiting to communicate to operation and organization C is running internal communication protocol.
- Organization B is talking to operation, organization C is waiting to communicate to operation and organization A is running internal communication protocol.

- Organization C is talking to operation, organization A is waiting to communicate to operation and organization B is running internal communication protocol.
- Organization C is talking to operation, organization B is waiting to communicate to operation and organization A is running internal communication protocol.

The system behavior can be modeled by a ten-state transition diagram. Transition matrix can be described as following:

$$P = \begin{bmatrix}
\overline{p_A}\,\overline{p_B}\,\overline{p_C} & \overline{p_A}p_B\overline{p_C} & p_A\overline{p_B}\,\overline{p_C} & \overline{p_A}\,\overline{p_B}p_C & p_Ap_B\overline{p_C} & \overline{p_A}p_Bp_C & p_Ap_B\overline{p_C} & p_A\overline{p_B}p_C & \overline{p_A}p_Bp_C & p_A\overline{p_B}p_C \\
\overline{p_A}q_B\overline{p_C} & \overline{p_A}\,\overline{q_B}\,\overline{p_C} & p_Aq_B\overline{p_C} & \overline{p_A}q_Bp_C & p_A\overline{q_B}\,\overline{p_C} & \overline{p_A}\,\overline{q_B}p_C & 0 & 0 & 0 & 0 \\
q_A\overline{p_B}\,\overline{p_C} & q_Ap_B\overline{p_C} & \overline{q_A}\,\overline{p_B}\,\overline{p_C} & q_A\overline{p_B}p_C & 0 & 0 & \overline{q_A}p_B\overline{p_C} & \overline{q_A}\,\overline{p_B}p_C & 0 & 0 \\
\overline{p_A}\,\overline{p_B}q_C & \overline{p_A}p_Bq_C & p_A\overline{p_B}q_C & \overline{p_A}\,\overline{p_B}\,\overline{q_C} & 0 & 0 & 0 & 0 & \overline{p_A}p_B\overline{q_C} & p_A\overline{p_B}\,\overline{q_C} \\
0 & 0 & q_B & 0 & \overline{q_B} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & q_B & 0 & \overline{q_B} & 0 & 0 & 0 & 0 \\
0 & q_A & 0 & 0 & 0 & 0 & \overline{q_A} & 0 & 0 & 0 \\
0 & 0 & 0 & q_A & 0 & 0 & 0 & \overline{q_A} & 0 & 0 \\
0 & q_C & 0 & 0 & 0 & 0 & 0 & 0 & \overline{q_C} & 0 \\
0 & 0 & q_C & 0 & 0 & 0 & 0 & 0 & 0 & \overline{q_C}
\end{bmatrix}$$

State transition diagram of the system described above is illustrated in figure 7.7.

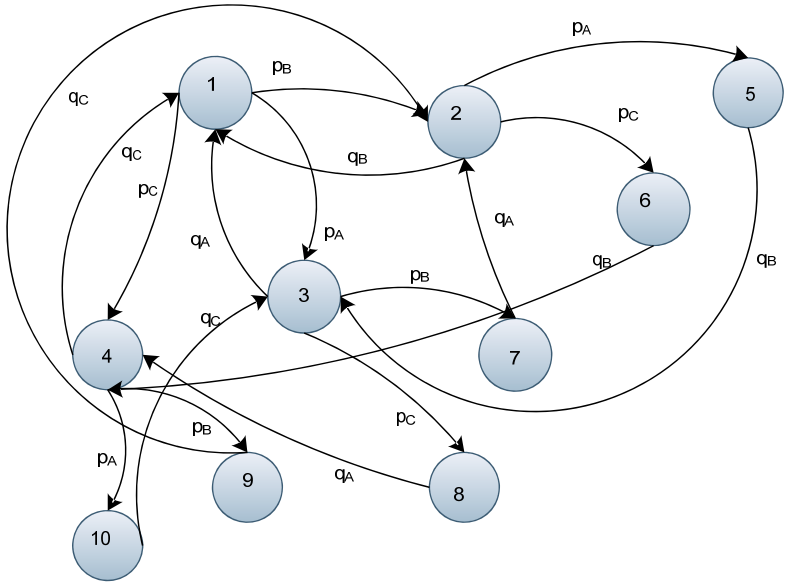

Figure 7.7: State Transition Diagram of three organizations with one Shared Resource

Theorem 7.1 *In a discrete event system consisting of n organizations sharing one resource, total number of possible states is equal to $n^2 + 1$ and the corresponding total*

number of communication is equal to $2 \times n^2$ *if only one organization is allowed to access the shared resource at a time.*

Proof- There is one shared entity or resource which can communicate to each organization in either direction, i.e. being accessed or released. For the next step there is one less communication for each organization. Therefore there is a total of $2n + 2n \times (n-1)$ communication which is equal to $2 \times n^2$. The total number of states is equal to $1 + n + n(n-1) = n^2 + 1$.

Total number of communication for 3 organizations is shown in figure 7.8 and for n organizations is shown in figure 7.9. The term $2\times$ originates from for bi-direction communication between each two states.

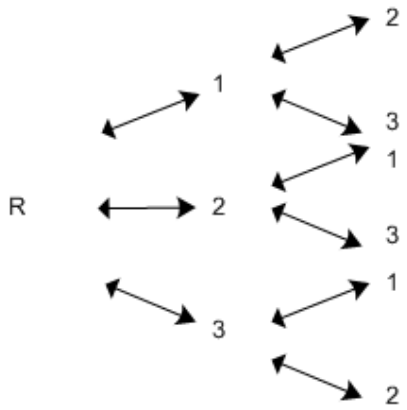

Figure 7.8: Total number of communication for 3 organizations

As proved by theorem 7.1, *the total number of communication increases exponentially* with number of organizations involved in sharing one resource or communicating to one entity. In emergency response application following ICS organizational chart, figure 2.1, span control of four to seven organizations or people can lead to a long delay, potential bottleneck or unreachability. This justifies the need to *elaborate communication between team leaders and the operations or local supervisor to develop horizontal interactions between peers*, scenario 4.6, figure 4.1. Rather than repetitive and redundant communication between each two entities and dealing with a long delay

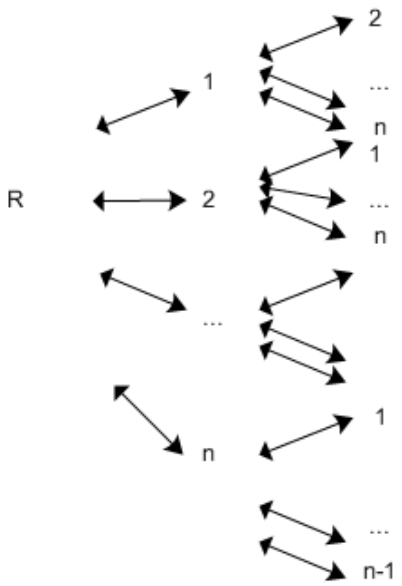

Figure 7.9: Total number of communication for n organizations

for an update, the communication protocol can be revised such that an updated message or data can be shared by the one entity who has it to ensure the existence of the last updated copy of the message at all times and avoid unnecessary delays.

7.4 Concluding Remarks

In this chapter we used transition matrix developed from reachability tree of Petri nets to predict and analyze system performance. This concerns behavioral properties of the system as it is developed from the initial marking. This chapter makes the following unique contribution: a framework is presented that allows a supervisor to schedule and control the execution of events based on the costs known from the state transition diagram. For a robust organizational structure, the performance does not vary by the change in the execution order of the same events.

Chapter 8

Supervisory Solutions to Improve Decision Making

8.1 Introduction

In a large-scale event different organizations with their individual Petri net models collaborate to meet the overall objectives. Detailed information on internal events within an organization may not be accessible from outside of that organization for territorial reasons. A supervisor with knowledge on the system can enforce constraints to prevent deadlocks and to improve performance.

In section 8.2, we present definitions and existing control solutions. In section 8.3, we present systematic methodologies to analyze properties and performance of hierarchical Petri nets, and to develop supervisory control solutions. In section 8.4 we present two hierarchical Petri net models where systematic methodologies presented in previous section have been applied to demonstrate properties of the models and to develop supervisory control solutions. In section 8.5 we present the set of functions by which a supervisor can improve decision making and finally, in section 8.6 we highlight the concluding remarks.

8.2 Control Solutions

Details of certain events within an organization may not be available to a higher-level supervisory controller.

Definition 8.1 *A transition is uncontrollable if it is an internal event within an organization such that its firing cannot be controlled by an external agent or supervisor [36].*

Definition 8.2 *A transition is unobservable if the firing of that transition is not directly visible to observer but the output of the system, that is dependent on that event, is visible [36].*

The theory of supervisory control of Petri nets has been introduced and developed by earlier work by Ramadge and Wonham in [57] and [58]. The constraint in equation 8.1 enforces a condition that restricts the set of reachable markings of a place [59] [60]:

$$l^T \,.\mu_p \leq b \tag{8.1}$$

l is the constraint vector of $n \times 1$ with integer elements representing the coefficient of places that appear in the constraint and 0 for the places that do not appear in the constraint, μ_p is the marking vector of the system, $n \times 1$, n is the number of places, and b is an integer representing the constraint. All constraints of this type can be grouped in matrix format:

$$L \,.\mu_p \leq B \tag{8.2}$$

where L is the constraint matrix of $n_c \times n$ with integer elements representing the coefficient of the places appearing in the constraint, and B is an integer vector representing the constraints, $n_c \times 1$, and n_c is the number of constraints of type 8.1. With a slack variable μ_c, we can re-write equation 8.2 as:

$$L.\mu_p + \mu_c = B \tag{8.3}$$

where μ_c is an integer vector of size $n_c \times 1$ representing the initial marking of control places where n_c is the number of control places. The set of place invariants are utilized mathematically to determine the set of constraints where all transitions are controllable and observable. Vector x is the desired place invariant:

$$x^T = \left[\; l_1 \; l_2 \; ... \; l_n \; 1 \; \right]$$

and the set of place invariants are:

$$X^T = \left[\begin{array}{cc} L & I \end{array} \right]$$

The flow matrix of the controlled system includes the original flow matrix of the model, F_p, and the flow matrix of the controller, F_c.

$$F = \left[\begin{array}{c} F_p \\ F_c \end{array} \right].$$

$\mu \in Z^{n_c+n}$ and initial marking μ_0

$$\mu = \left[\begin{array}{c} \mu_p \\ \mu_c \end{array} \right] \qquad \mu_0 = \left[\begin{array}{c} \mu_{p0} \\ \mu_{c0} \end{array} \right] \tag{8.4}$$

$$X^T.F = \left[\begin{array}{cc} L & I \end{array} \right] . \left[\begin{array}{c} F_p \\ F_c \end{array} \right] = 0 \;\Leftrightarrow\; F_c = -L \,.\, F_p \tag{8.5}$$

In a system with uncontrollable or unobservable transitions, the set of constraints need to be transformed to an admissible sets such that firing of those transitions can be prevented or monitored [60]. The following definition gives the sufficient and necessary condition to verify the admissibility of the constraint.

Definition 8.3 *For a model with initial marking μ_{p0}, an admissible constraint satisfies $L\,\mu_{p0} \leq b$, and $L\,\mu_p \leq b$ for all reachable markings μ_p developed from initial marking [36].*

Definition 8.4 *For a model with uncontrollable transitions represented by flow matrix of F_{uc}, and a constraint $l_T\,\mu_p \leq b$, if $l^T\,F_{UC} \leq 0$, then the constraint is admissible [36].*

Transformation to an admissible set of constraints for a Petri net model with uncontrollable and unobservable transitions can be found in [36].

We propose to apply reduction transformations according to section 6.3 prior to developing supervisory solutions to reduce the complexity of the flow matrix and initial marking of the controller.

8.3 Systematic Methodologies

In this section we present two systematic methodologies that provide necessary and sufficient conditions to ensure deadlock free performance. In systematic methodology 8.1 we examine structural properties of the Petri net model including conditions of liveness. In systematic methodology 8.2 we investigate the behavioral properties of the model and develop supervisory solutions. The model with supervisory control solution should be checked again for the conditions of liveness.

Methodology 8.1

- check for boundedness or safeness property
- check for conservation property
- develop the flow matrix
- find the set of place invariants
- verify whether the system is deadlock free or not

Methodology 8.2

- develop reachability tree
- calculate transition matrix
- reduce complexity of the model if possible
- find an admissible constraint or verify whether a given constraint is admissible; transform if not admissible
- develop supervisory control solution of section 8.2 to find F_C
- check the closed loop supervisory control model to determine whether it is deadlock free or not
- if there is a deadlock, revise the constraint and reapply supervisory control solution

The type of constraint used in this chapter covers a broad range of conditions in emergency response application concerning limits on total number of tokens in certain places, or simultaneous enabling of multiple events. These methodologies will be demonstrated in example 8.1 and example 8.2.

8.4 Hierarchical Petri nets with Supervisory Control Solutions

The Petri net model in the example 8.1 demonstrates a case where supervisor synchronizes two processes and controls an event in one of the organizations. We are interested in the points of interactions among organizations.

Example 8.1

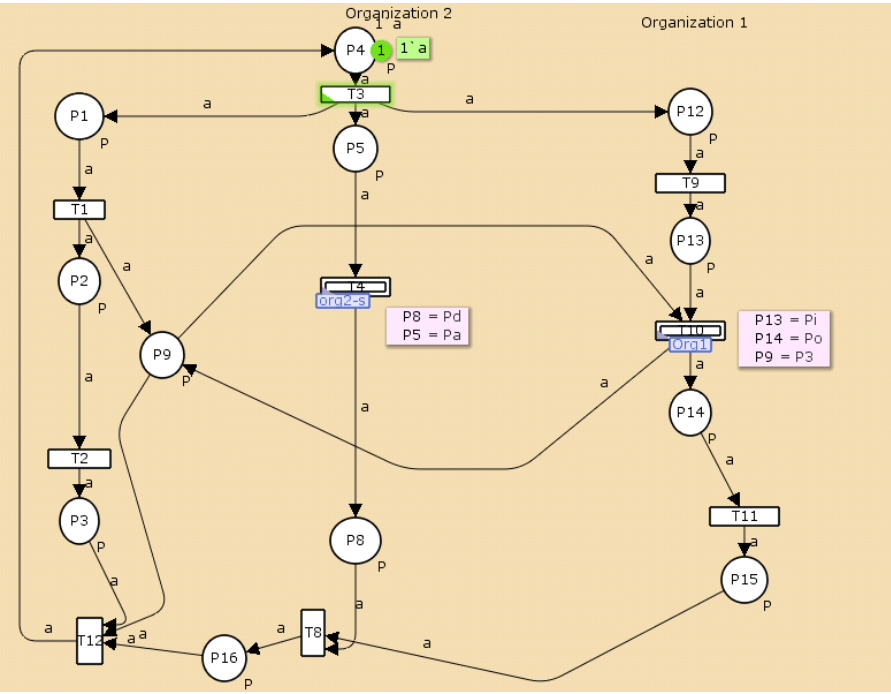

Figure 8.1: A Petri net model illustrating interactions among two Organizations and Supervisor

Figure 8.1 represents a hierarchical Petri net model where T_{10} and T_4 are substitution transitions to represent organization one and two respectively. Each organization can

have its individual Petri net model of any scale and complexity which does not affect the interactions on the higher-level supervisory model. The Petri net model of organization one is shown in figure 8.2. Organization two is revisiting the Petri net model of patients' treatment illustrated in figure 8.3. In CPN tool [42], interaction between two Petri nets is modeled by *ports and sockets*. Point of communication on higher level is referred as sockets and on lower level as ports. Ports are labeled as input, output or I/O to indicate direction of communication to the corresponding socket on the higher-level.

The number of tokens at all places except place P_9 is bounded by 1 as there is one incoming arc connected to all these places and therefore those places are safe. Place P_9 has two incoming arcs but the incoming arc from transition T_{10} always returns the consumed token on the outgoing arc from this place to that transition which is used to enable transition T_{12} and therefore does not cause a deadlock. Place P_9 has maximum of one token at any given time and hence it is safe. The total number of tokens in the model is constant and therefore this Petri net model is conservative. When T_3 fires, a token that enabled this transition, produces three tokens and when T_{12} fires, three token that enabled this transition, create only one token. The flow matrix of top page in figure 8.1 is shown in equation 8.6.

$$F_p = \left[\begin{array}{ccccccccc} -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{array}\right] \tag{8.6}$$

The flow matrix of this Petri net model will not be uniquely representing a Petri net because the model is not persistent according to remark 5.1 and therefore, developing the

state diagram of this model will be applicable to other Petri net models corresponding to the same reachability tree. There are four sets of place invariants:

$$\{P_1, P_2, P_3, P_4\}$$
$$\{P_4, P_5, P_8, P_{16}\}$$
$$\{P_4, P_{12}, P_{13}, P_{14}, P_{15}, P_{16}\}$$
$$\{P_1, P_4, P_9\} \tag{8.7}$$

The set of place invariants are both siphons and traps. These siphons are never empty, but this is not sufficient to guarantee that the model is deadlock free. To determine whether the model is deadlock free or not, the conditions of theorem 6.6 must be verified. This model is a work flow net where the sink place is identical to the source place. All places except place P_9 have one incoming and one outgoing arcs. The loop between place P_9 and transition T_{10} creates a redundant dependency which is not causing any deadlock and therefore the model is deadlock free. Table 8.1 shows a firing

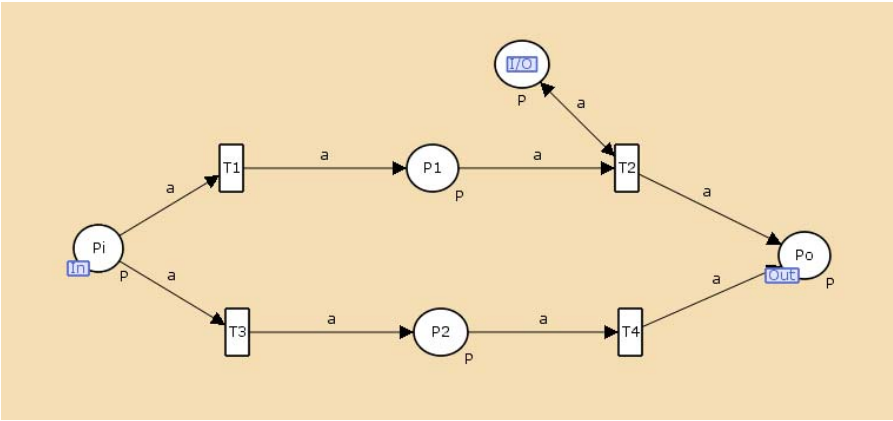

Figure 8.2: The Petri net model for Organization 1

sequence for one realization (execution) of this example which is completed in seventy nine steps. Each cycle that starts with firing of T_3 and ends with T_{12} consists of thirteen hops. Reachability tree is illustrated in figure 8.4. The Transition state diagram based on the reachability tree is presented in figure 8.5. The transition matrix T_p for the Petri

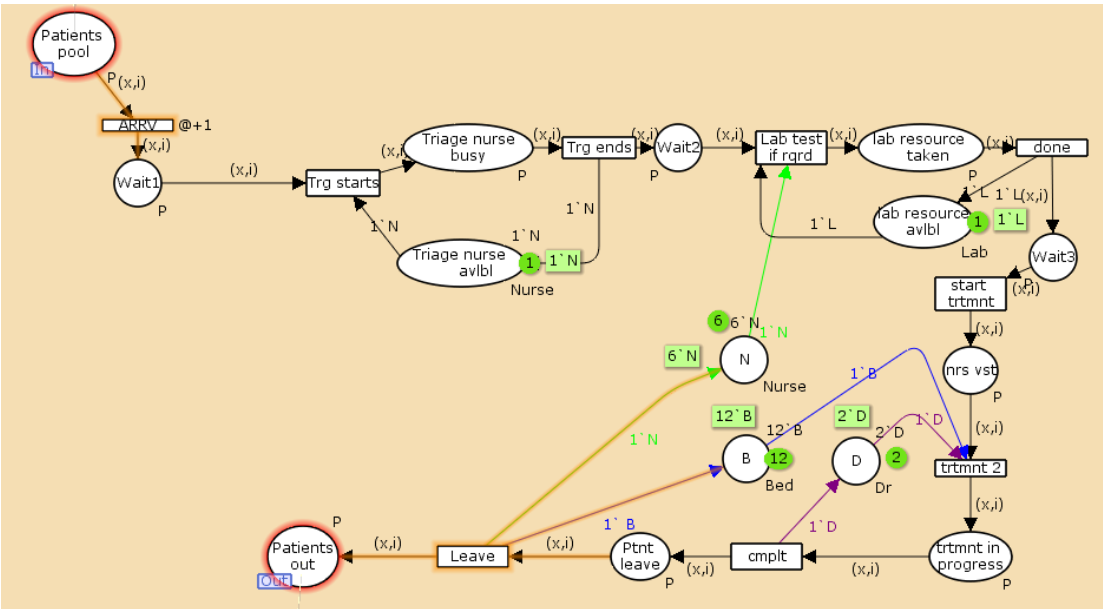

Figure 8.3: The Petri net model for Organization 2

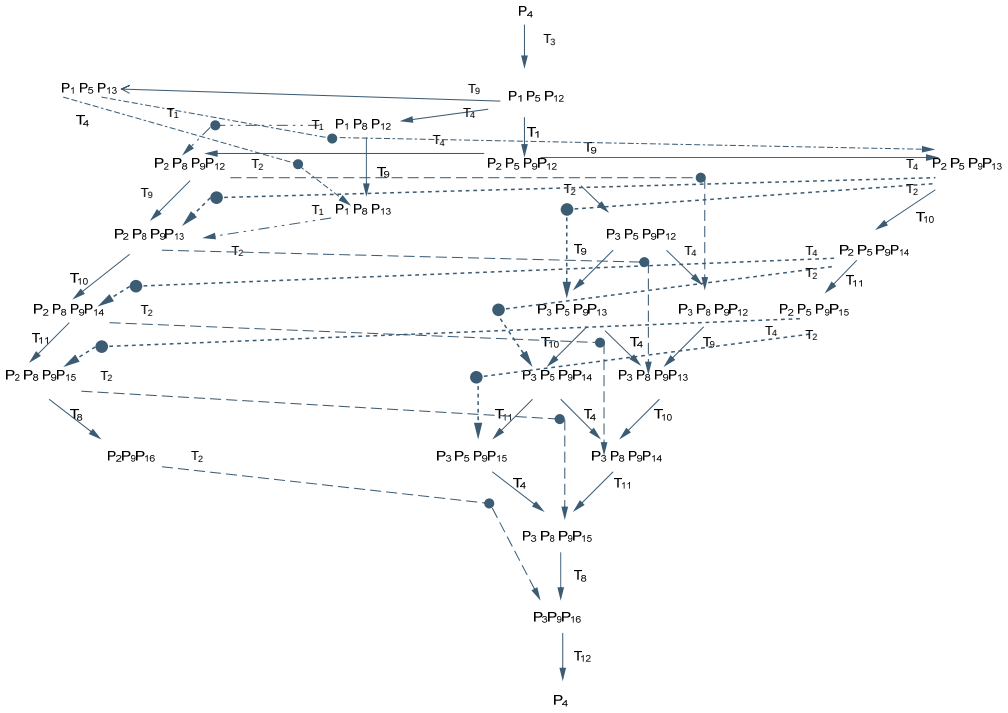

Figure 8.4: The Reachability Tree for Petri net of figure 8.1

net of figure 8.1 is presented below:

$$\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

To reduce the complexity of supervisory controller in hierarchical Petri nets, we apply reduction transformation to simplify the Petri net model. The loop between place P_9 and transition T_{10} can be removed without loss of generality. Two parallel paths will remain between transition T_1 and T_{12} which can be reduced to one according to section 6.3 as shown in figure 8.6. In this figure, there are three parallel paths, between an *AND* split

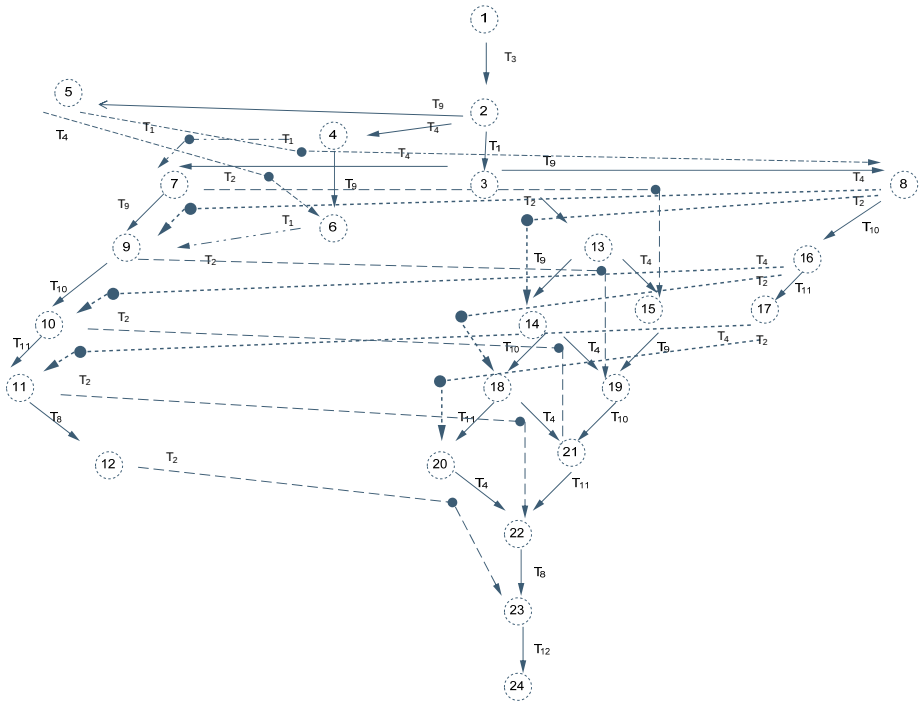

Figure 8.5: Transition State Diagram for Petri net of figure 8.1

and an *AND* merge which can be reduced further according prototype II in chapter 6 which leads to the simple Petri net model of figure 8.7.

Reachability tree is illustrated in figure 8.8 and the transition matrix T_p is reduced to:

$$T_p = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{8.8}$$

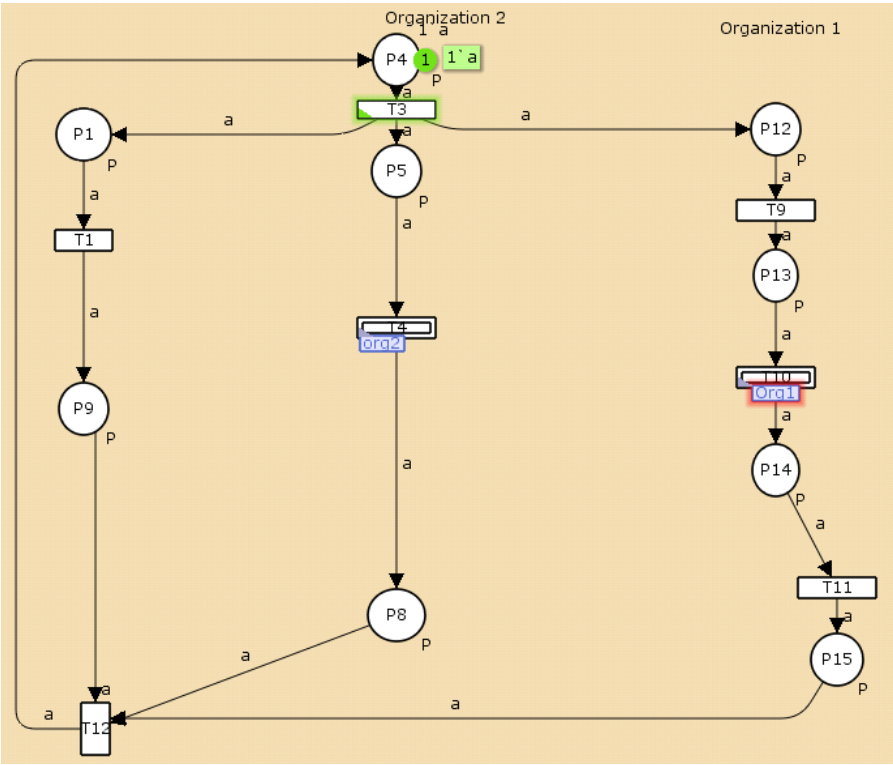

Figure 8.6: The reduced Petri net model

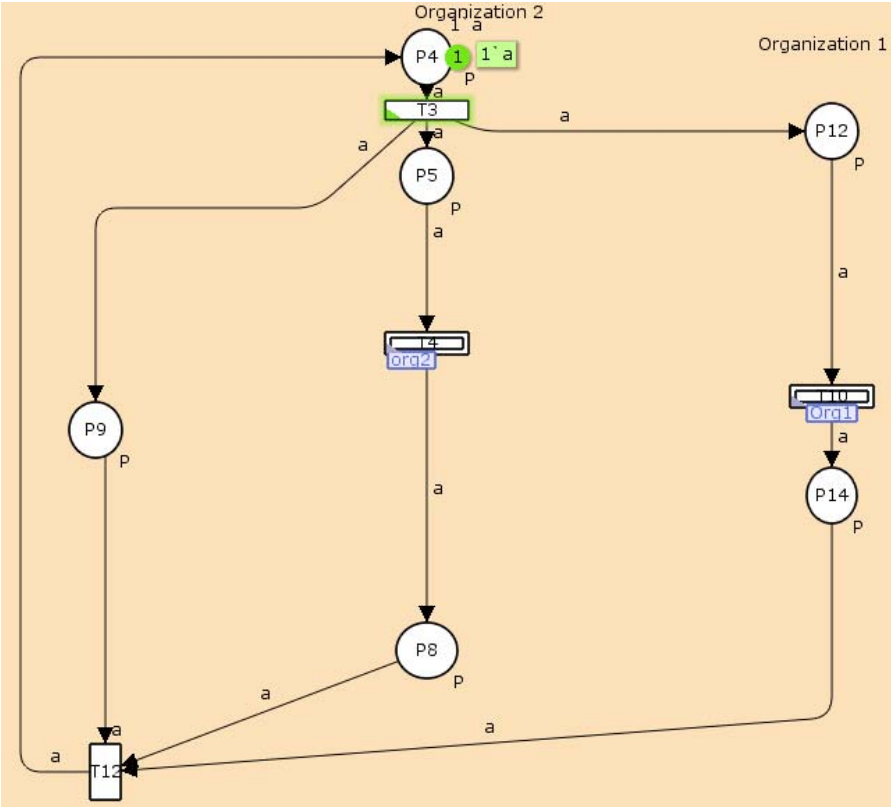

Figure 8.7: The reduced Petri net model of figure 8.1

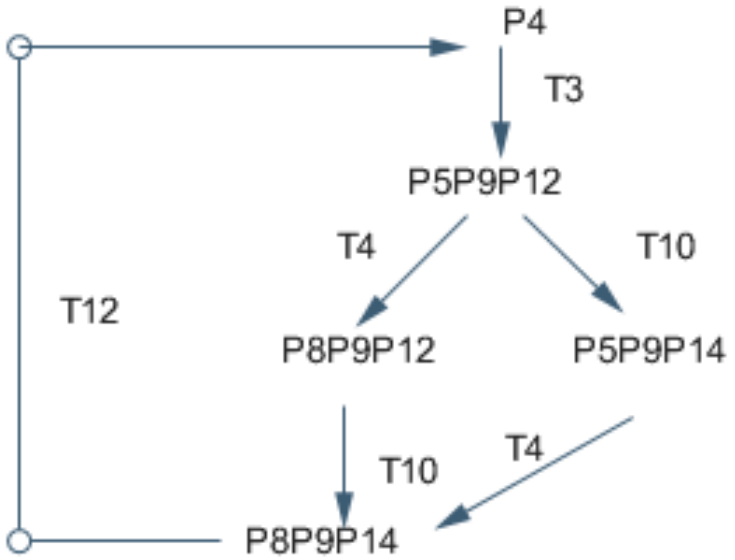

Figure 8.8: Reachability tree corresponding to the Petri net of figure 8.7

The flow matrix of The Petri net of figure 8.7 is reduced to the following:

$$F_p = \begin{bmatrix} -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \tag{8.9}$$

With flow matrix presented in equation 8.9, and a constraint $p_9 \leq 1$, we calculate flow matrix of the controller from equation 8.5:

$$l^T = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$F_c = \begin{bmatrix} -1 & 0 & 0 & 1 \end{bmatrix}$$

and initial marking of the control place is obtained as $\mu_{c0} = 1 - l\mu_{p0} = 1$. The Petri net model with supervisory controller for the given constraint is illustrated in figure 8.9. In this example the Petri net with supervisory solution ensures deadlock free performance according to the reachability tree of figure 8.10.

Example 8.2

This example revisits the Petri net of example 5.2, figure 8.11. In designing a supervisory Petri net, we need to identify any uncontrollable or unobservable transition in the system and verify whether the constraints are admissible or not. The constraint is such that P_2 and P_5 never have more than one token:

$$P_2 \; + \; P_5 \; \leq \; 1 \tag{8.10}$$

There is no unobservable transition in this model and transitions T_1, T_3 and T_4 are uncontrollable. The firing of uncontrollable transitions cannot be prevented. This can be justified for the inaccessibility of the internal transitions of an organization in a hierarchical system by an another entity. In this Petri net model, dependencies are placed

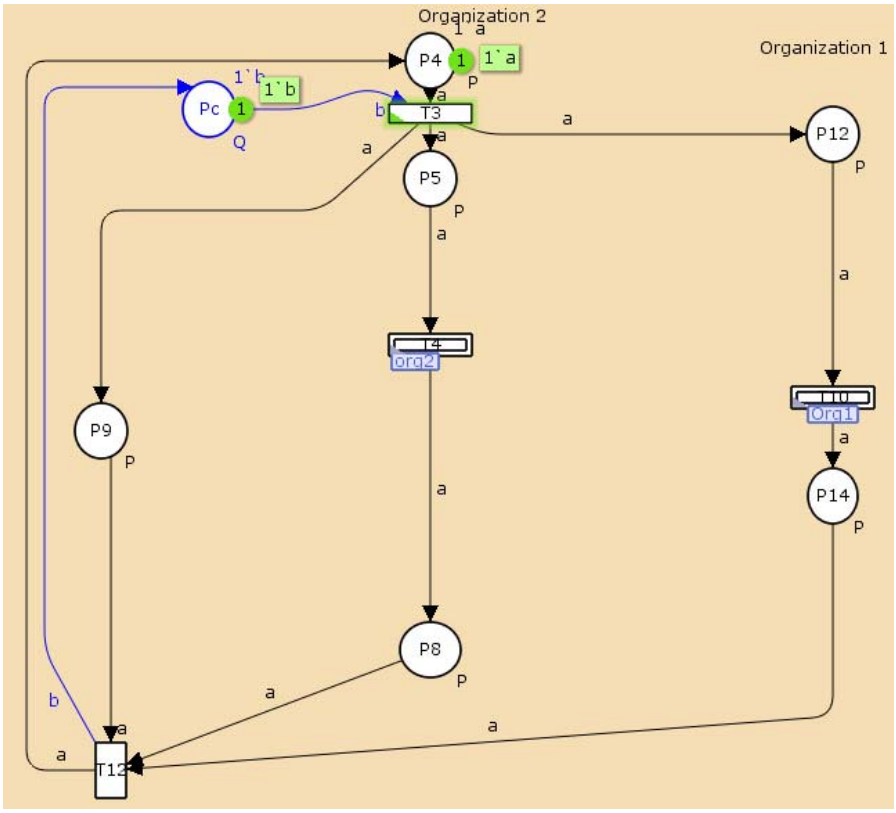

Figure 8.9: The supervisory Petri net model of figure 8.7

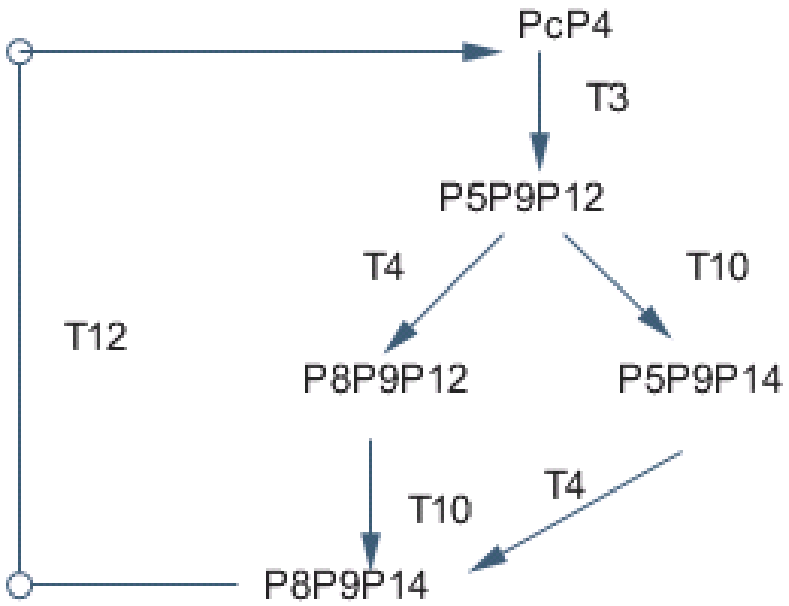

Figure 8.10: Reachability tree corresponding to the Petri net of figure 8.9

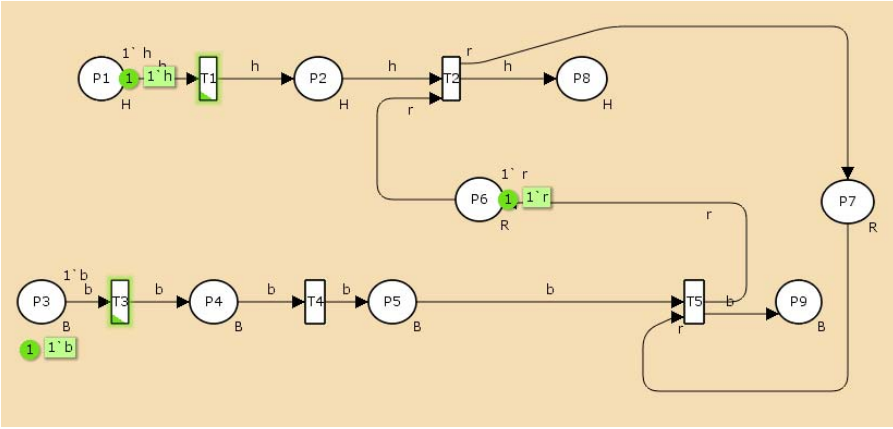

Figure 8.11: A Petri net model

in the farthest possible event so that each organization can proceed individually to that event without waiting for communication from another organization. F_{UC} represents the part of flow matrix that corresponds to the uncontrollable transitions:

$$F_{UC} = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Constraint 8.10 is admissible according to the condition presented in section 8.2, hence:

$$\mu_{P_0} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$L \times \mu_{P_0} = 0$$

$$L = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$b = 1$$

With a slack variable for the controller place, the initial marking and the flow matrix of the controller are calculated applying equations 8.3 and 8.5:

$$F_c = -LF_p = \begin{bmatrix} -1 & 1 & 0 & -1 & 1 \end{bmatrix}$$

$$\mu_{c_0} = 1 - L\mu_{p_0} = 1$$

The supervisory Petri net model for the given system is presented in figure 8.12. The topology of the arcs connecting to control place is obtained from the flow matrix, F_c. However, adding this controller leads the system to a deadlock: if T_4 uses the token in

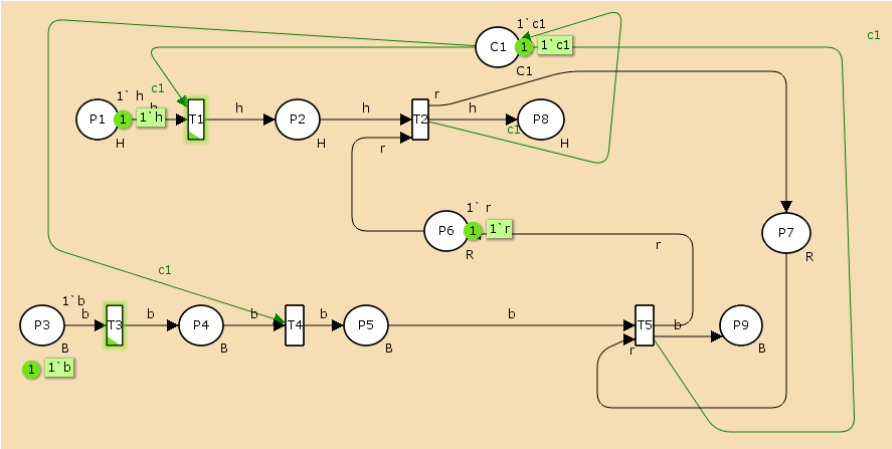

Figure 8.12: The Supervisory controller Petri net model

C_1 before T_1, the deadlock occurs! T_5 cannot fire because P_7 does not have a token; P_6 has the resource and organization one cannot proceed either because there is no token at controller place to enable T_1 and therefore T_2 cannot fire either! While supervisory controller can enhance the performance in many scenarios by preventing deadlocks, it can also cause deadlock. To solve the problem, the model can be reduced first or the supervisory control solution can be re-applied to the closed loop control system which adds the complexity.

We propose to reduce the complexity of Petri net by replacing transition directed cascades by equivalent transitions as presented in chapter 6 before applying the supervisory control solutions. The Petri net model of figure 8.11 can be reduced to the Petri net model of figure 8.13 where no further simplification is possible. This model is deadlock free. If for any reason transition T_{2eq} never fires, organization two faces infinite delay. To prevent this, a supervisor can place a token in a control place connecting through a transition to place P_7 which enables transition T_{5eq}. The Petri net model with supervisory solution is presented in figure 8.14.

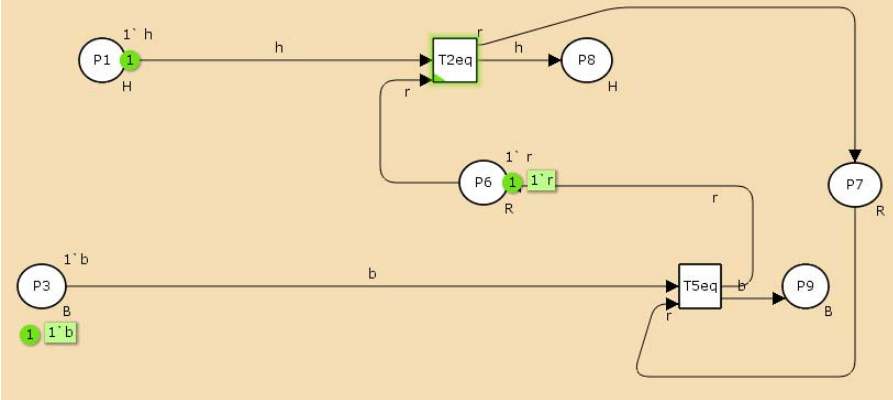

Figure 8.13: Reduced Petri net model

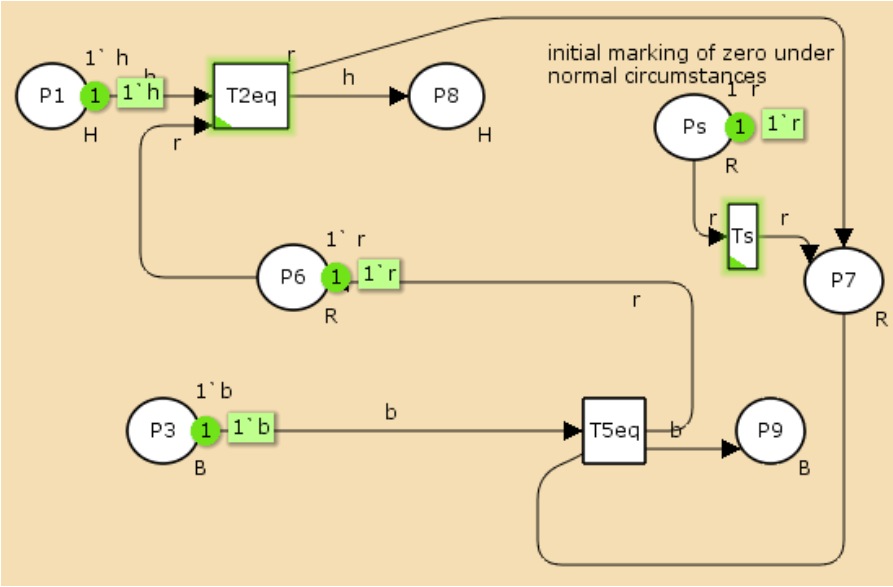

Figure 8.14: The Supervisory controller Petri net model for figure 8.13

8.5 Supervisory Control Functions

In this section we summarize supervisory functions with application to emergency response which have been presented in this work. A supervisory control solution can contribute to a more efficient response in any of the following ways:

- Prevent multiple incident notifications or prevent multiple updates at the same time (by enforcing clear termination time)
- Ensure safeness, liveness or fairness as applicable
- Minimize dependencies for a more efficient organizational relationship
- Control execution order based by calculating path costs based on transition diagram if real costs are known
- Ensure precise message delivery to all designated recipients by a supervisory token that ensures one copy of data at any given time
- Efficient resource allocation by monitoring resources, keeping track of number of tokens in places
- Distributed control architecture to prevent single point of failure by establishing local controllers, deputies

To achieve:

- Efficiency in resource utilization by knowing the resources and the limits; making recommendations; having a big picture
- Smoother inter-organizational communication protocols and making educated real-time decisions
- Situational awareness when dealing with conflicts

8.6 Concluding Remarks

In this chapter we developed two methodologies to analyze structural and behavioral properties of a hierarchical Petri net system consisting of several smaller Petri nets interacting with each other. These methodologies provide techniques to predict performance and design supervisory solutions to facilitate information sharing and decision making. We apply reduction transformations to reduce complexity and simplify analysis. Removing unnecessary dependencies and moving entanglements to the farthest possible point, reduces the complexity of analysis. It is possible that a supervisory control solution causes deadlocks and creates more complexity and dependencies in which case, the communication protocol and interactions need to be revised and the model has to be examined again for deadlock free state. However a supervisor with system knowledge and higher authorities can improve performance by revising flow of information or adding control places to prevent long delays or deadlocks.

Table 8.1: Firing Sequence of figure 8.1

Sequence	Network	Sequence	Network	Sequence	Network
1	T3 @ (1:Top)	28	T1 @ (1:Top)	55	Tb @ (1:org2)
2	T9 @ (1:Top)	29	T2 @ (1:Top)	56	T9 @ (1:Top)
3	Ta @ (1:org2)	30	T9 @ (1:Top)	57	Td @ (1:org2)
4	T3 @ (1:Org1)	31	T1 @ (1:Org1)	58	T1 @ (1:Org1)
5	T4 @ (1:Org1)	32	Ta @ (1:org2)	59	Tc @ (1:org2)
6	T11 @ (1:Top)	33	T2 @ (1:Org1)	60	T1 @ (1:Top)
7	T1 @ (1:Top)	34	T11 @ (1:Top)	61	T2 @ (1:Top)
8	Tb @ (1:org2)	35	Tb @ (1:org2)	62	T2 @ (1:Org1)
9	Td @ (1:org2)	36	Td @ (1:org2)	63	T11 @ (1:Top)
10	Tc @ (1:org2)	37	Tc @ (1:org2)	64	T8 @ (1:Top)
11	T2 @ (1:Top)	38	T8 @ (1:Top)	65	T12 @ (1:Top)
12	T8 @ (1:Top)	39	T12 @ (1:Top)	66	T3 @ (1:Top)
13	T12 @ (1:Top)	40	T3 @ (1:Top)	67	Ta @ (1:org2)
14	T3 @ (1:Top)	41	Ta @ (1:org2)	68	Tb @ (1:org2)
15	T1 @ (1:Top)	42	T9 @ (1:Top)	69	T1 @ (1:Top)
16	T9 @ (1:Top)	43	T1 @ (1:Org1)	70	T9 @ (1:Top)
17	T2 @ (1:Top)	44	Tb @ (1:org2)	71	Td @ (1:org2)
18	Ta @ (1:org2)	45	Td @ (1:org2)	72	Tc @ (1:org2)
19	T1 @ (1:Org1)	46	Tc @ (1:org2)	73	T2 @ (1:Top)
20	Tb @ (1:org2)	47	T1 @ (1:Top)	74	T3 @ (1:Org1)
21	Td @ (1:org2)	48	T2 @ (1:Org1)	75	T4 @ (1:Org1)
22	Tc @ (1:org2)	49	T2 @ (1:Top)	76	T11 @ (1:Top)
23	T2 @ (1:Org1)	50	T11 @ (1:Top)	77	T8 @ (1:Top)
24	T11 @ (1:Top)	51	T8 @ (1:Top)	78	T12 @ (1:Top)
25	T8 @ (1:Top)	52	T12 @ (1:Top)	79	T3 @ (1:Top)
26	T12 @ (1:Top)	53	T3 @ (1:Top)		
27	T3 @ (1:Top)	54	Ta @ (1:org2)		

Chapter 9

Conclusion

Communication in emergency response applications has unique requirements and critical importance for the myriad ways a nation will be affected by in case of a failure. This work developed performance prediction and supervisory solutions to improve performance. Next section presents a summary of our contributions to enhance communication with application to emergency response.

9.1 Summary of Contributions

This work incorporates communication technology and organizational communication protocols into traditional stereotypes of communication perspective. Traditional solutions have mostly focused on the communication infrastructure. This dissertation revealed that deployment of a robust and reliable communication infrastructure does not solve communication problem in emergency response without investigation of communication protocols at organizational level.

Through participation in several real-life scenario exercises, network data was captured to monitor and analyze network statistics, network performance, and to identify sources of bottleneck. It turned out that in our deployment, a large number of small control packets was causing overhead. Through examination of after-incident reports, and interviews with first responders, this dissertation developed its focus to consider complex communication problem at organizational level. Event-driven models based on real-life scenarios were developed to examine structural and behavioral properties of

communication protocols at organizational level. We presented a framework by which a supervisor can control the execution order of events given the costs. A supervisor can contribute to the crisis response field by reducing unnecessary complexities of interactions to achieve efficient performance in one or more of the following ways: minimizing dependencies, cancellation of an scheduled event and providing an alternative option in case of unreachability of a peer or supervisor, enforcing liveness in case of a deadlock, resource allocation and ensuring the receipt of updated message by all, to prevent the distribution of multiple copies of data and to ensure reliability.

This work presented systematic methodologies in analyzing organizational communication protocols and to enforce constraints to limit set of possible states to prevent deadlocks. These methodologies assist decision makers to improve quality of response. These methodologies fully benefit from real-life scenario and can be directly applied to currently existing structures in emergency applications.

We believe that developing event-driven models for internal organizations and analyzing the hierarchical structure of them when they gather to respond to a disaster once, is worth the time and cost for the benefits of monitoring critical statistics and preventing potential deadlocks. The hierarchical design allows the model to be robust to lower-level changes in each individual organization.

9.2 Future Directions

This work has the potential to extend in the following areas:

9.2.1 Automation

In this work we presented an event-driven modeling technique to predict and analyze system performance. The modeling technique is developed from the building blocks of its process, i.e. conditions, events and cases. In this modeling technique, Petri nets, events were represent by transitions, conditions were represent by places, and cases were represent by tokens. A case is completed as the designated token moves throughout the system from the input place to the output place. The objective is to ensure that the right action is taken place at the right time by the right person.

The same objective applies to work flow management. Petri net modeling and work flow representation can be transformed to each other. There are several work flow management tools which allow real-time monitoring of case progression. For example, PNMsoft [61] is a tool that allows real-time monitoring of process. Whenever more than one user is accessing same data or participating in the execution of the same step, an additional branch is automatically created. The work flow is highlighted in colors to indicate whether the tasks completed on-schedule or with delay.

The systematic methodologies developed in section 8.3 have the potential to be implemented on Programmable Logic Controller (PLC) to automate the process of decision making.

9.2.2 Security

Another possible extension to our work is for scenarios where the message needs to be transmitted via a secure channel or be encrypted due to a security threat. A higher authority supervisor may need to cancel an already scheduled procedure and send information securely.

Depending on the application, the information may need to be sent to one or n organizations securely. Considering that in an emergency response most communication take place over an ad hoc wireless network, sending over secure channel may not be an option at earlier stages. Instead the message can be encrypted. There are two main mechanisms to encrypt a message, *public key system (asymmetric) and Private key mechanism (symmetric key encryption). RSA or El-Gamal* are examples of public key mechanism and Date Encryption Standard (DES) is the well-known private key mechanism.

Each mechanism has its own advantages and disadvantages but since accessing a secure channel to exchange the secret key for private key mechanism is not realistic, public key mechanisms are more suitable candidates. Public key of sender is sent over any communication channel and the sender will be able to decrypt the message with its own private key. If there are different types of information shared among different entities, key distribution and key management can cause a problem as there is a large number of keys for each end user to handle. Digital signatures are another mechanism to

ensure authenticity and integrity of data. Practical data modification for privacy reason can be achieved to remove sensitive part of data by *redaction*. This allows information on names or association of individuals to be removed.

We believe that the systematic methodologies that we presented in this work form a good basis for further developments in both automation and security areas.

Appendix A

Incident Command System (ICS)

The purpose of this appendix is to provide background information on ICS structure [62]. The ICS defines the operating characteristics, interactive management components, and structure of incident management and emergency response organizations engaged throughout the life cycle of an incident. ICS is a field management system with the following primary functions: incident commander, operations, planning and intelligence, logistics and finance/administration. Figure 2.1 illustrates organizational chart for ICS. A brief definition of each function of ICS structure is as follows:

- Incident commander

  The incident commander is responsible for all incident activities. It defines the objective and priorities for overall emergency response.

- Operations

  This division directs the tactical actions and all resources to meet incident objectives to carry out the response and achieve the incident objective. The operations chief is normally from the agency with the greatest involvement in terms of resources assigned or area of concern. Strike team and task force belong to this section.

- Planning and intelligence

  This group is in charge of collection, evaluation, display of incident information, maintaining resource status and preparing the incident action plan and incident-related documentation.

- Logistics

  Logistics provides adequate services and resources to support all incident needs such as obtaining and maintaining essential personnel, facilities, equipment and supplies. Communication unit, medical unit, food unit, supply unit, facilities and ground support unit belong to this section.

- Finance/Administration

  This division monitors costs related to incident, provides accounting, time recording and cost analysis.

According to ICS structure a person in charge of an organization is responsible until the authority is delegated to another person. There is a Safety officer, a liaison officer and an information officer. Information officer is the only point of contact for media and liaison officer is the point of contact for larger incidents for other agencies to coordinate their agency's involvement. Note that ICS/SEMS(California Standardized Emergency Management System) is followed in California while NIMS (National Incident Management System) is followed across the nation.

A.1 Resources

The selection of the right resources for the task is critical to accomplish the objective properly and cost-efficiently. The person in charge of *planning section* maintains the status of all resources at all times.

A.1.1 Resource Types and Kinds

Resource *kind* describes what the resource is. Same kind of resources might be used by different agencies. Resource *type* describes performance capability for the kind of resource. Resource type helps with planning, resource allocation, and capacity monitoring. Knowing the capacity of each kind helps planners decide the type and quantity of resource. A clear indication of what is needed helps the planner to monitor below-capacity or over-capacity to make adjustments in allocating those resources.

A.1.2 Resource Management

Resources can be managed in one of three ways. Below is a brief description and the advantages of each method:

- As single resources: most common way, its type reflects its capability
- As task forces (TF): Any combination and number of single resources (within span of control limits) gathered for a particular need. Task force refers to any combination of the same kind but different types of resources. They have a designated task force leader who operates under common communication plan. This method is very flexible in the arrangements with the span of control being the only limit.
- As strike teams (ST), squads, platoons: Same requirements as task force. Strike team refers to any combination of the same kind and same type of resources.

Advantages of task force and strike teams are to facilitate resource planning, reduce unnecessary communication, efficient resource control and the abaility to scale for larger incidents within the span of control limit.

A.2 Incident Management

Management by the objective represents the approach that is communicated throughout the entire ICS organization. This approach includes establishing the objective, and developing plans and protocols and comprehensive resource management to achieve the objective. Incidents are managed under a single, collaborative approach including common organizational structure, single incident command post, unified planning process and unified resource management. The incident manager is responsible to establish incident objective and to choose an appropriate strategy. There is a concept of *unity and chain of command* which ensures everyone has a supervisor. *Unity of command* means that every individual has a designated supervisor to whom they report at the scene of the incident. *Chain of command* refers to the orderly line of authority within the ranks of the incident management organization. These principles clarify reporting relationships and eliminate the confusion caused by multiple, conflicting directives.

Advantages of Using Unified Command is that all efforts is directed to meet the incident objective. Information flow and interactions is improved among all different first response organizations. Possible problems with jurisdiction territorial proud that may create obstacles in communication among them is resolved as all the individual participants are recognized. Incident managers at all levels must be able to control the actions of all personnel under their supervision.

Appendix B

National Incident Management System (NIMS)

The purpose of this appendix is to provide the reader with a brief background information on Federal NIMS structure [63]. The NIMS defines standardized mechanisms and establishes requirements for processes to describe, inventory, mobilize, dispatch, track, and recover resources over the life cycle of an incident. NIMS standard incident command structures are based on three key organizational systems:

- Incident Command System(ICS)
  ICS is a management system designed to enable effective and efficient domestic incident management by integrating a combination of facilities, equipment, personnel, procedures, and communications operating within a common organizational structure.

- Multi-agency Coordination System
  This defines the operating characteristics, interactive management components, and organizational structure of supporting incident management entities engaged at the Federal, State, local, tribal, and regional levels through mutual-aid agreements and other assistance arrangements; and

- Public Information System
  Public Information System refers to processes, procedures, and systems for com-

municating timely and accurate information to the public during crisis or emergency situations.

B.1 Resource Management

Resource management includes processes for categorizing, ordering, dispatching, tracking, and recovering resources. Resources include personnel, facilities, equipments and supply items. Resource management involves coordinating and overseeing the application of tools, processes, and systems that provide incident managers with timely and appropriate resources during an incident. Generally *Emergency Operations Center (EOC)*f is responsible for resource management coordination. Multi-agency coordination entities when established may also prioritize and coordinate resource allocation during incidents. Resource management involves four primary tasks:

- Establish systems to describe, request and track resources
- Activate these systems prior to and during an incident
- Dispatch resources prior to and during an incident
- Deactivate or recall resources during or after incidents

Maintaining an accurate and up-to-date status of resource utilization is a critical component of domestic incident management.

B.2 Communications and Incident Management

Effective communications, information management, and information and intelligence sharing are critical aspects of domestic incident management. Establishing and maintaining a common objective and ensuring accessibility and interoperability are principal goals of communications and information management. NIMS identifies the requirement for a standardized framework for communications, information management (collection, analysis, and dissemination), and information-sharing at all levels of incident management. These elements are briefly described as follows:

- Incident Management Communications:
  Incident management organizations must ensure that effective, interoperable communications processes, procedures, and systems exist to support a wide variety of incident management activities across agencies and jurisdictions.

- Information Management:
  Information management procedures ensure that information, including communications and data, flows efficiently through a commonly accepted architecture supporting numerous agencies and jurisdictions responsible for managing domestic incidents. Effective information management enhances the response and decision making.

A good decision making system must be driven based on precise and updated information. The decision making process must adapt automatically to the type of information accessible to the decision maker. The value of information becomes greater as it becomes scarce and more difficult to access. A small update may lead to a different set of actions which may change the outcome from failure to success.

B.3 Interoperability

Interoperability provides the necessary framework to communicate operational decisions at an incident site, as well as off-the-site across different jurisdictions. Prior to an incident, entities responsible for taking appropriate pre-incident actions use communication and information management processes to inform and guide various critical activities. These actions include mobilization or pre-deployment of resources, as well as strategic planning by prepared organizations, multi-agency coordination entities, agency executives, jurisdictional authorities, and EOC personnel. During an incident, incident management personnel communicate to coordinate and execute operational decisions.

B.4 Area Command

An *Area Command* is only activated upon the need, depending on the complexity of the incident and the incident management span of control. An area Command is

established either to oversee the management of multiple incidents that are each being handled by a separate ICS organization or to oversee the management of a very large incident that involves multiple ICS organizations, such as incidents that are not site specific, geographically dispersed, or evolve over longer periods of time like a bioterrorism event. Area Command is also used when there are a number of incidents in the same area and of the same type, such as two or more hazardous material (HAZMAT) or oil spills and fires. These represent incidents that may compete for the same resources. When incidents do not have similar resource demands, they are usually handled separately and are coordinated through an EOC. An agency administrator or other public official with jurisdictional responsibility for the incident usually makes the decision to establish an area Command. If the incidents under the authority of the Area Command are multi jurisdictional, then a *Unified Command (UC)* should be established. This allows each jurisdiction to have representation in the command structure. Area Command should not be confused with the functions performed by an EOC. An Area Command oversees management of the incident(s), while an EOC coordinates support functions and provides resources support. An Area Command is responsible to set overall incident-related priorities, allocate critical resources according to the priorities and ensure proper incident management. It should also identify critical resource needs and report them to EOCs or multi-agency coordination entities.

B.5 Multi-agency Coordination

A multi-agency coordination system is a combination of equipments, personnel, procedures and communications integrated into a common system with responsibility for coordinating and supporting domestic incident management activities. The primary functions of multi-agency coordination system is to support incident management policies and priorities, facilitate logistics support and resource tracking, inform resource allocation decisions using incident management priorities, coordinate incident related information, and coordinate inter-agency and inter-governmental issues regarding incident management policies, priorities, and strategies.

Multi-agency coordination systems may contain EOCs. The principal functions

and responsibilities of multi-agency coordination entities regardless of form or structure, typically include the following:

- To ensure that each agency involved in incident management activities is providing appropriate situational awareness and resource status information
- To establish priorities between incidents or area Commands in concert with the IC or UC(s) involved
- To acquire and allocate resources required by incident management personnel in concert with the priorities established by the IC or UC
- To anticipate and identify future resource requirements
- To coordinate and resolve policy issues arising from the incident(s); and
- To provide strategic coordination as required

B.6 Emergency Operations Center (EOC)

EOC represents the physical location at which the coordination of information and resources to support incident management activities normally takes place. EOCs may be permanent organizations and facilities or may be established to meet temporary, short-term needs. The physical size, staffing, and equipping of an EOC will depend on the size of the jurisdiction, resources available, and anticipated incident management workload. EOCs may be organized and staffed in a variety of ways. Regardless of the specific organizational structure used, EOCs should include the following core functions: coordination, communications, resource dispatch and tracking, and information collection, analysis, and dissemination. EOCs may also support multi-agency coordination and joint information activities. Upon activation of a local EOC, communications and coordination must be established between the IC or UC and the EOC, when they are not collocated. ICS field organizations must also establish communications with the activated local EOC, either directly or through their parent organizations. Additionally, EOCs at all levels of government and across functional agencies must be capable of

communicating appropriately with other EOCs during incidents, including those maintained by private organizations. Communications between EOCs must be reliable and contain built-in redundancies.

The *Incident Command Post (ICP)* located at or in the immediate vicinity of an incident site, although primarily focused on the tactical on-scene response, may perform an EOC-like function in smaller-scale incidents or during the initial phase of the response to larger, more complex events. EOCs are organized by major functional discipline like fire department, law enforcement, or emergency medical services, by jurisdiction, i.e. city, county, or region, or more likely by a combination of them. *Department Operations Centers (DOCs)* normally focus on internal agency incident management and response and are physically represented in a higher level EOC. ICPs should also be linked to DOCs and EOCs to ensure effective and efficient incident management.

For complex incidents, EOCs may be staffed by personnel representing multiple jurisdictions and functional disciplines and a wide variety of resources. For example, a local EOC established in response to a bioterrorism incident would likely include a mix of law enforcement, emergency management, public health, and medical personnel including representatives of health care facilities, pre-hospital emergency medical services, patient transportation systems, pharmaceutical repositories, and laboratories.

Appendix C

Questionnaire

The questionnaire was conducted to better understand communication protocols and information flow among first response organizations in an emergency response. This has been also used as a set of data to extract real scenarios in this work. Interviewees have been asked a set of questions from the following questionnaire depending on time and their expertise:

- What role do you serve in the MMST drill?
- Describe what you know is happening at the drill. What is the communication protocols and organizational chart for this drill?
- Who is in charge of planning communication for the whole drill?
- What information must be quickly communicated? What information will one need from other organizations?
- Is there an information flow map available from any drills in past? will there be one for this drill? Any lessons learned?
- Any complexity issues when a choice needs to be made
- Any scenario where decision makers postpone a decision for a reason?
- What are possible problems that may raise if communication is delayed?

- Is there a well-communicated plan on what channels are in use by each organization?
- How is horizontal communication different from the vertical communication?
- Who gets to Operations? where is MMST TFL compared to ICS and ICS Operations? is there always a deputy TFL?
- Any mobile command center? power issues? any disconnection or traffic abbreviations using radios?
- Where will the key responders be at and who do they follow?
- Who establishes zones?
- Does HAZMAT get the Hot Zone always?
- Does the operational control shift from operations to MMST?
- A call for MMST goes to SWAT, Sheriff, who else? Who is higher than MMST TFL? Is MMST on a different channel?
- What is MMST point of contact to obtain info from hot zone? Through operations? Continuous monitoring?
- Do cities and counties have their own EOC? Is EOC off the site, the one referred as dispatcher maybe? Do they make decisions? Were they part of the drill? Who decided to call MMST? Any rearrangement as MMST joins?
- What resources each organization provide? who makes decision on resource split?
- When do they call for an Area Command?
- What is a Department operation (Doc)?
- What communication tools will be used? radio, cell phone, email, WebEoc?
- Any comments on competition and territorial problem among organizations that might slow down the efficient response?

Appendix D

Golden Guardian Drill On May 18th 2010

Golden Guardian 2010 exercise was the sixth full-scale exercise since 2004 which took place on May 18th at California State University, San Marcos [64]. This exercise was developed to assess selected state emergency functions and regional capabilities in a collaborative process. There was a shooting on campus at one of the dorms followed by a bomb that went off at 10:01 AM at soccer field. We observed communication and coordination among several capabilities that were assessed in this exercise. UCSD San Diego Health system participated in this exercise along with area hospitals, clinics the state, federal, regional, and local responders [65]. In this exercise the expected casualties included trauma, burn, possible radiation exposures and contamination from *dirty bomb* scenario, a conventional bomb containing radioactive material. Contamination required the use of radiation survey meters and activation of the Health System's Decontamination Team.

There were approximately 100 victims in the soccer field and a few at the dorms. San Marcos fire department got notified first but since the shooter was believed to be on campus, law enforcement got involved. The incident command was established at Campus Police and fire department supported them.

Sheriff department is the primary law enforcement for many cities and counties around San Diego which includes but not limited to Encinitas, Del Mar beaches and Poway.

MMST is called in for multiple hazard incidents including explosion and radiation. For this drill, they were called in after explosion to support the response. MMST task force leader (TFL) offers their specialized resources to *Incident Commander*. MMST consists of tactical law group, special forces, Explosive Ordnance Disposal (EOD:SWAT and bomb squad), HAZMAT, Tactical Medical, Paramedics, and decontamination tent (DECON). Every emergency response organization with regional responsibility has its own department operation (DOC) such as medical DOC and Sheriff DOC.

MMST TFL helps managing expectationsions. MMST TFL brings situational awareness, tries to anticipate, and reminds others to check on items based on experience. MMST TFL has also a deputy who can talk to the incident commander and make sure that all communication is clear. Incident commander can defer decisions to MMST TFL. MMST TFL provides safety comfort to the entry team. This is because they have specialists in hazards of different nature. MMST TFL takes control of hot zone and after victims leave DECON, sends them back to triage for paramedics and possible hospital transportations which is under incident command. To record all face-to-face communication, there is a scribe who takes notes of all communications. In this exercise, the authorities to make real-time changes to previously assigned communication protocols or tasks are: MMST director and exercise controller. MMST director is outside the hot zone and is concerned with safety of MMST team. Exercise controller had about twenty controllers in this drill who control traffic, victims and each one had an evaluator. MMST director supports what MMST needs and exercise controller acts under incident command operation.

We learned that there was a decision point where a choice could have made a difference for this drill. It was interesting to see what fire department captain would do when radiation pager of responders at the soccer field went off. They were ordered to evacuate the hot zone and the ones affected needed to pass through DECON for safety. In past when individuals did not understand capabilities due to *unclear communication*, the entry would be delayed but that problem is fixed now.

In this drill both incidents were managed under one incident command because the location was the same. If multiple incidents are spread in different locations or cities,

an area commander is usually formed to oversee multiple incident commanders. In *Silver Bullet drill* as explained in section 2.5, an area command was formed to oversee the two incidents and reported to EOC (Emergency Operation Center). An area commander prioritizes and coordinates the operations of multiple incident commanders and determines which one gets access to what resources. As usually there is not enough resources of all types available to all. There can be more than one area commander depending on the scale of disaster and in that case they all communicate to EOC and can request more resources.

The author would like to acknowledge Donna Johnson, Director of MMST, for her support to provide partial information presented in this appendix.

Bibliography

[1] "Advanced Network Technologies Division Communication and Networking Technologies for Public Safety," *National Institute of Standards and Technology (NIST)*.

[2] "The Federal Response to Hurricane Katrina, Lessons Learned," *http://www.whitehouse.gov/reports/katrina-lessons-learned.pdf*, 2006.

[3] "Advanced Network Technologies Division Communication and Networking Technologies for Public Safety," *http://www.fema.gov/media/archives/2007/061207.shtm*, 2008.

[4] "Washington DC, Area Sniper Investigation Communications After-Action Report," *SAFECOM*, 2003.

[5] "Wireless Emergency Response Team (WERT), Final Report for the September 11, 2001 NEW York City World Trade Center Terrorist Attack," *http://www.wert-help.org/index.php*, 2001.

[6] C. P. W. P. T. W. G. Virginia Tech, "Information and Communications Infrastructure," *Copyright Virginia Tech, http://www.vtnews.vt.edu/documents/2007-08-22 communications infrastructure.pdf*, 2007.

[7] "National LambdaRail," *http://www.nlr.net/*.

[8] "Disaster Preparedness and Service Using Communication Technology," *Computer World Honors Case Study, http://www.cwhonors.org/laureates/government/20055426.pdf*, 2005.

[9] "Los Angeles earthquake chokes phone calls, not Twitter," *http://alpha.news.yahoo.com/s/cnet/8301103531000191294, Yahoo News*, 2008.

[10] R. B. Dilmaghani and R. R. Rao, "Hybrid Wireless Mesh Network with Application to Emergency Scenarios," *Journal of Software Engineer*, vol. 3, pp. 52–60, February 2008.

[11] http://www.colorado.edu/hazards/workshop/, "Natural Hazard Center."

[12] "Calit2 Researchers Deploy Communications Network at San Diego Mardi Gras Festivities," *http://www.calit2.net/newsroom/article.php?id=810*, 2006.

[13] "CalMesh," *http://calmesh.calit2.net/*.

[14] "Meet Don Kimball," *http://www.calit2.net/newsroom/article.php?id=237*, 2006.

[15] R. B. Dilmaghani and R. R. Rao, "Hybrid Communication Infrastructure and Social Implications for Disaster Management," *40th HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES*, 2007.

[16] "UCSD Tests Intelligent Triage, Other Technologies in San Diego Disaster Drill," *http://www.calit2.net/newsroom/article.php?id=745*.

[17] http://en.wikipedia.org/wiki/File:ICS Structure.PNG), "ICS Structure."

[18] "Ph.D. Student Observes Medical Command Center at Rock 'n' Roll Marathon," *http://www.calit2.net/newsroom/article.php?id=1547*, 2009.

[19] "Fire Incident Report," *Art Botterell, Community Warning System Manager for the Office of the Sheriff of Contra Costa County, California, provided us with a copy of the report.*

[20] R. B. Dilmaghani and R. R. Rao, "A Wireless Mesh Test bed with Applications for Emergency Response," *4th IEEE International conference on Wireless, Mobile Computing, Networking and Communication (WiMob)*, 2008.

[21] R. Zimmerman, "Social Implications of Infrastructure Network Interactions," in *JOURNAL OF URBAN TECHNOLOGY*, vol. 8, 2001, pp. 97–120.

[22] S. J. Tierney K., "Cost and Culture: Barriers to the Adoption of Technology in Emergency Management," in *Proceedings of the 3rd International ISCRAM Conference*, ...

[23] H. Zimmerman, "Availability of Technologies versus Capabilities of Users," in *Proceedings of the 3rd International ISCRAM Conference*, 2006.

[24] "MIT Roofnet," *http://pdos.csail.mit.edu/roofnet/doku.php*.

[25] "Bay Area Wireless Users Group," *http://www.bawug.org*.

[26] "Champaigne-Urbana Community Wireless Network," *http://www.cuwireless.net*.

[27] "WiFi Mesh News," *http://wi-fi-mesh-news.newslib.com/feed.xml*.

[28] "Seattle wireless," *ttp://www.seattlewireless.net*.

[29] "wireless leiden," *http://www.wirelessleiden.nl*.

[30] R. Dilmaghani, B. Manoj, B. Jafarian, and R. Rao, "Performance evaluation of RescueMesh: a metro-scale hybrid wireless network," in *Proceedings of WiMesh Workshop, IEEE Workshop on Wireless Mesh Networks, held in conjunction with SECON, Santa Clara*, 2005.

[31] "Gizmo: Toy Truck Delivers," *http://www.calit2.net/newsroom/article.php?id=989*.

[32] J. Kurose and K. Ross, *Computer Networking: A Top-Down Approach*. Addison-Wesley Educational Publishers Inc., 2000.

[33] M. Page-Jones, *The practical guide to structured systems design*. 2nd edition.

[34] http://en.wikipedia.org/wiki/Cohesion-(computer science), "Cohesion-Wikipedia."

[35] R. B. Dilmaghani and R. R. Rao, "A Systematic Approach to Improve Communication for Emergency Response," *42nd HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES*, 2009.

[36] J. Moody and P. Antsaklis, *Supervisory control of discrete event systems using Petri nets*. Kluwer Academic Pub, 1998.

[37] T. Murata, "Petri nets: Properties, Analysis and Applications," *http://www.cs.unc.edu/montek/teaching/spring-04/murata-petrinets.pdf*.

[38] J. Peterson, *Petri net theory and the modeling of systems*. Prentice Hall PTR Upper Saddle River, NJ, USA, 1981.

[39] W. Van der Aalst and K. van Hee, *Workflow management: models, methods, and systems*. MIT press, 2002.

[40] C. Cassandras, *Discrete event systems*. Springer, 1993.

[41] K. Jensen, *Coloured Petri nets: basic concepts, analysis methods, and practical use*. Springer, 1992.

[42] http://wiki.daimi.au.dk/cpntools/cpntools.wiki, "Color Petri Net tools."

[43] J. Wang, *Timed Petri nets: Theory and application*. Kluwer Academic Publishers, 1998.

[44] J. Ezpeleta, J. M. Couvreur, and M. Silva, "A new technique for finding a generating family of siphons, traps and st-components. Application to colored Petri Nets," *Lecture Notes in Computer Science; Advances in Petri Nets 1993*, vol. 674, pp. 126–147, 1993.

[45] http://en.wikipedia.org/wiki/Petrinet, "Petri net-Wikipedia."

[46] J. Mogul and K. Ramakrishnan, "Eliminating receive livelock in an interrupt-driven kernel."

[47] J. Anderson, Y. Kim, and T. Herman, "Shared-memory mutual exclusion: Major research trends since 1986," *Distributed Computing*, vol. 16, no. 2, pp. 75–110, 2003.

[48] M. Marsan, G. Balbo, G. Conte, S. Donatelli, and G. Franceschinis, "Modelling with generalized stochastic Petri nets," *ACM SIGMETRICS Performance Evaluation Review*, vol. 26, no. 2, 1998.

[49] F. G. and N. S., "Feedback control of Petri nets based on place invariants," *technique et Science Informatiques*, vol. 32, pp. 143–160, 1985.

[50] M. Molloy, "On the integration of delay and throughput measures in distributed processing models," 1981.

[51] F. Symons, "Modelling and analysis of communication protocols using numerical Petri Nets." *Dissertation Abstracts International Part B: Science and Engineering[DISS. ABST. INT. PT. B- SCI. & ENG.],*, vol. 49, no. 1, 1988.

[52] P. Bammidi and K. Moore, "Emergency management systems: a systems approach," in *1994 IEEE International Conference on Systems, Man, and Cybernetics, 1994.'Humans, Information and Technology'.*, vol. 2, 1994.

[53] H. Xiong, M. Zhou, and C. Manikopoulos, "Modeling and Performance Analysis of Medical Services Systems Using Petri Nets," in *IEEE INTERNATIONAL CONFERENCE ON SYSTEMS MAN AND CYBERNETICS*, vol. 3. INSTITUTE OF ELECTRICAL ENGINEERS INC (IEEE), 1994, pp. 2339–2339.

[54] R. David and H. Alla, *Discrete, Continuous, and Hybrid Petri Nets*. Springer Berlin Heidelberg, 2005.

[55] F. G. and N. S., "Evaluation based upon stochastic Petri nets of the maximum throughput of a full duplex protocol," *Application and theory of Petri nets: selected papers from the first and the second European workshop*, vol. 52, pp. 280–288, 1982.

[56] R. H. A. van der Aalst1 W. M. P., van Hee1 K. M., "Analysis of discrete-time stochastic Petri nets," *Statistica Neerlandica pp.*, vol. 54, pp. 237–255, 2000.

[57] W. Wonham and P. Ramadge, "On the supremal controllable sublanguage of a given language," in *Decision and Control, 1984. The 23rd IEEE Conference on*, vol. 23, 1984.

[58] P. Ramadge and W. Wonham, "The control of discrete event systems," *Proceedings of the IEEE*, vol. 77, no. 1, pp. 81–98, 1989.

[59] K. Yamalidou, J. Moody, M. Lemmon, and P. Antsaklis, "Feedback control of Petri nets based on place invariants," *Technical Report of the ISIS Group, ISIS-94-002.2*, 1994.

[60] J. Moody, L. M. Yamalidou, K., and P. Antsaklis, "Feedback control of Petri nets based on place invariants," in *Decision and Control, 1994., Proceedings of the 33rd IEEE Conference on*, vol. 3, 1994, pp. 3104–3109.

[61] http://www.pnmsoft.com/workflow tutorial.aspx, "Workflow Tutorial."

[62] "Incident Command System (ICS) Orientation," *http://www.spokares.org/IncidentCommand.html*.

[63] "National Incident Management System (NIMS)," *http://www.nimsonline.com/*.

[64] "Golden Guardian 2010 (California Emergency Management Agency Training and Exercises)," *http://www.scribd.com/doc/25059978/Golden-Guardian-2010-California-Emergency-Management-Agency-Training-Exercises*.

[65] "Golden Guardian Drill," *http://healthspan.ucsd.edu/2010/04/Pages/mc-ann-drill.aspx*.