Chapter 3 Analysis of IGMP and PIM-DM

After discussing IP Multicast, computer simulation, and the analysis of UMJaNetSim in the previous chapter, this chapter concentrates on the analysis of the IGMP and the PIM-DM.

IGMP is needed for IP Multicast. Hence, it is included in UMJaNetSim. The first section in this chapter is the analysis of IGMP, in which an introduction on the IGMP message and IGMP table are given. This is followed by an analysis of the IGMP timers. Then, protocol analysis is presented. In this section, an analysis of the IGMP function in a host and a router focuses on the IGMP message handling, timers handling, members joining and leaving, and so on.

In the next section, an analysis of the PIM-DM algorithm and functionality are detailed. Here, an introduction on the PIM-DM control message, the PIM-DM timers, and the PIM-DM tables are presented. Then, the protocol analysis of PIM-DM control message and timers handling, multicast packet forwarding, group joining, pruning, and so on are discussed. The final section of this chapter is the summary.

3.1 Analysis of IGMP

In this thesis, IGMP version 2 is used. Hence, [13] is used as the reference to develop *the IGMP in the UMJaNetSim.*

3.1.1 IGMP Message

IGMP messages are encapsulated in an IP packet. Figure 3.1 shows the format of an IGMP message. Three types of IGMP message are defined. There are Membership Query, Membership Report, and Leave Group. Due to the assumption that all routers implements IGMP version 2, only Version 2 Membership Report is used in this study.

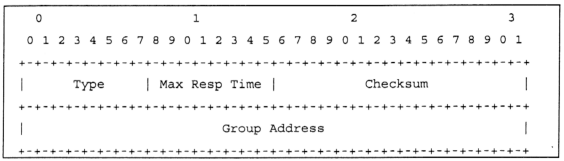

Figure 3.1 IGMP message format

a) Membership Query

Query message is used to determine the existence of a multicast group. There are two types of Membership Query message.

- General query

  It is used to learn and identify the groups that have members on the attached network. [13]

- Group-Specific Query

  It is used to learn if a particular group has any members on the attached network. [13]

b) Membership Report

It is used to reply the query message in order to let the router know the interest of a host in joining a particular multicast group.

c) Leave Group

Leave message is used when a member decides to leave a group.

Refer to [13] for other details of IGMP messages.

3.1.2 IGMP Group Table

According to [3], IGMP employs a group table to store information on group membership. An entry in the group table means that a multicast group has at least one process or one host as its member.

The IGMP group table for host and router is slightly different. Group table in a host maintains the following information.

- State

  It is used to indicate the status of an entry. Three states are available. There are FREE, DELAYING and IDLE. FREE indicates that the entry has expired and can be replaced by a new entry while DELAYING indicates that the entry is in reporting state and a timer is running. IDLE means that there is no timer running for that entry.
- Multicast group address
- Interface number

  It is used to identify the interface that receives the multicast packets
- Reference count

  It is used to state the number of processes that are still interested in joining a group.
- Report flag

  It indicates that whether it is the last party to send a Membership Report message

No timer is implemented in the IGMP group table at a host. The entry in IGMP group table is turned into FREE when there is no member joining that multicast group. On the other hand, routers keep the following information.

- State

  It is used to indicate the status of an entry. Two states are available. There are MEMBER and NON-MEMBER. MEMBER indicates that the entry is active and has at least one member for the multicast group in that entry. NON-MEMBER means that the entry has expired and can be replaced by a new entry.

- Multicast group address
- Interface number

  It is used to identify the interface where a member exists.

Each entry is assigned a Group Membership Interval Timer in the group table in a router. It is used to indicate the lifespan of the entry. The timer is reset when a report is received for the multicast group stated in the entry. The state of an entry is turned into NON-MEMBER when the timer has expired.

3.1.3 IGMP Timers and Variables

Several variables are used to increase efficiency in the implementation of IGMP. Table 3.1 shows the variables. Timer intervals are also defined in [13]. Table 3.2 shows the timer interval in IGMP. In general, timers in IGMP can be divided into 2 categories. The first category is for hosts and the other category is for routers. Table 3.3 and Table 3.4 list the timers used in IGMP for host and router respectively.

Table 3.1 IGMP Variable and Count

Variable/Count	Description	Value
Robustness Variable	The Robustness Variable allows tuning for the expected packet loss on a subnet. If a subnet is expected to be lossy, the value may be increased	2
Startup Query Count	The number of queries sent out on startup, separated by the Startup Query Interval Timer	Robustness Variable
Last Member Query Count	The number of Group-Specific Membership Query message sent before router declare that there are no existence of member for a group	Robustness Variable

Table 3.2 List of IGMP Timer Intervals

Timer Interval	Description	Default Value (second)
Query Interval	Interval between General Membership Query message sent out by router	125
Query Response Interval	Value inserted in Max Response Time in General Membership Query message	100
Group Membership Interval	The amount of time that must pass before router decides that there are no members for a group in a network	{(Robustness Variable) x (Query Interval Timer)}+(Query Response Interval Timer)
Other Querier Present Interval	The duration of time that pass before a multicast router decides that there is no other router which should be querier	{(Robustness Variable) x (Query Interval Timer)}+3 x (Query Response Interval Timer) ÷ 2
Startup Query Interval	Interval between General Membership Query sent by router at startup	(Query Interval Timer) ÷ 4
Last Member Query Interval	The amount inserted in Max Response in Group-Specific Membership Query message and also the time Interval between Group-Specific Membership Query messages.	10
Unsolicited Report Interval	Time between repetitions of initial report membership for a group in a host.	10

Table 3.3 List of IGMP Timers for Host

Timer		Description	Value (second)
Host Timer	Query Response Timer	Timer for Membership Report responding	A random value between zero to Max Response Time in General Membership Query message

Table 3.4 List of IGMP Timers for Router

Timer		Description	Value (second)
Router Timer	Query Timer	Timer for sending General Membership Query	Query Interval
		At startup, Timer of General Membership Query sent by router. This is done for [Startup Query Count] of times	Startup Query Interval
	Group Membership Timer	Timer for router to decide that there are no members for a group in a network	Group Membership Interval Timer
	Other Querier Present Interval Timer	Timer for non-querier	Other Querier Present Interval Timer
	Last Member Query Timer	Timer to declare that there is no other members exist for a multicast group after receiving Leave message	For querier: (Last Member Query Interval) X (Last Member Query Count) For non-querier: (Max Response Time in the General-Specific Membership Query message) X (Last Member Query Count)
	Group-Specific Membership Query Retransmit Timers	Timer to retransmit Group-Specific Membership Query by [Last Member Query Count] of times	Last Member Query Interval

3.1.4 Protocol Description

The description of IGMP is divided into two main parts for analysis. There are IGMP in hosts and IGMP in routers. Each part is further divided into three subsections as stated below.

- Handling IGMP Message

- Handling IGMP Event
  - Member Join
  - Member Leave
- Handling IGMP Timers

3.1.4.1 IGMP in Host

- **Receiving Membership Query Message**

  When General Membership Query message is received, the host will go through each active entry in the group table and set a response timer, namely Query Response timer, for each entry. The response time is randomly chosen between zero and maximum response time specified in the Membership Query Message. [13] If a Group-specific Membership Query message is received, a response timer is only set for the particular multicast group specified in the message. If membership for that particular group does not exist, the query is ignored.

- **Receiving Membership Report Message**

  When a host receives a Membership Report message sent by another host in the same network, it checks whether it has membership for the reported multicast group. If there is, the report process is cancelled and the response timer is stopped. Otherwise, the host ignores the Membership Report message.

- **Receiving Other Message**

  If other than the above-mentioned message is received, the message is ignored and does not need to be processed.

- **Join Event**

  When a process desires to join a multicast group, the host determines whether the entry for that multicast group already exists. If the multicast group already exists, nothing needs to be done except increasing the Reference Count in that existing entry. Otherwise, a report is sent to inform the router about the joining event.

- **Leave Event**

  When a process in a host decides to leave a multicast group, a host checks whether it is the last process in the entry from the group table. If it is, it checks

if it is the last party to send Membership Report message for that multicast group. A Leave message is sent only if the host is the last party to send a Membership Report.

- **Query Response Timer**

  When Query Response timer expires, Membership Report message for that particular group is sent to router.

3.1.4.2 IGMP in Router

- **Receiving Membership Query Message**

  When a router receives a General Membership Query message, it has to determine its status. A router is set to non-querier if the message comes from a router with a lower IP address. Otherwise, it has to reset the Other Querier Present Interval Timer. On the other hand, when the router (non-qurier) receives a Group-specific Membership message, a Last Member Query Timer is started.

- **Receiving Membership Report Message**

  Upon receiving a Membership Report message from a host, a router will update the corresponding entry in a group table and reset the Group Membership Interval Timer for that entry.

- **Receiving Leave Message**

  If a Leave message is received by a querier router, it will response by sending a Group-Specific Membership Query message of the leaving group to determine whether there is still a member interested in joining that multicast group. Last Member Query Timer and Retransmit Timer is started.

- **Notification Add**

  When a router receives a join event from a host, it will notify the multicast routing protocol. The multicast routing protocol in this study is PIM-DM.

- **Notification Remove**

  On the other hand, when a member no longer exists for a certain group, router will also notify multicast routing protocol.

- **Query Timer**

  When the timer expires, a General Membership Query message is sent.

- **Group Membership Timer**
  When the timer expires, router declares there is no member for that particular entry and the state of the entry is turned into NON-MEMBER.
- **Other Querier Present Timer**
  When the timer expires, a router will shift into querier state and start as a querier router.
- **Last Member Query Timer**
  When the timer expires, a router declares that there is no member for that particular entry and the state of the entry is turned into NON-MEMBER.
- **Group-Specific Membership Query Retransmit Timer**
  When the timer expires, a Group-Specific Query message is sent.

3.2 Analysis of PIM-DM

This study uses [18] as the main reference in developing PIM-DM in UMJaNetSim. Since PIM-DM employs the same packet formats used in PIM-SM, [26, 50] have also been referenced during development of PIM-DM.

3.2.1 PIM-DM Control Message

In general, PIM-DM uses five types of control message. These control messages are Hello Message, Prune Message, Assert Message, Graft Message, and Graft-Ack Message. All PIM-DM control messages have a common header before the respective control message. Figure 3.2 shows the common header for PIM-DM control message. PIM-DM control messages are encapsulated in IP packets before it is sent and its protocol number in the IP packet is 103. Basically, PIM-DM message are either unicast (such as Graft Message and Graft-Ack Message) or multicast to all PIM routers group (such as Hello Message, Prune Message, and Assert Message). [18]

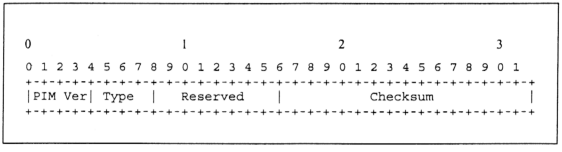

Figure 3.2 Common Header for PIM-DM Control Message.

According to [18], the number in PIM Version field in the control message is 2. This is because PIM-DM version 2 is used in this study. For the Type field, 8 types of control messages may be filled in. These include some control messages used in PIM-SM. Table 3.5 lists out the Type field number and its corresponding message type. Currently, Reserved field has not been defined for any other function. It is set to zero and ignored upon reception. The checksum is the 16-bit complement one of the complement one sums of the entire PIM-DM control message.

Table 3.5 Control Message Types for Both PIM-SM and PIM-DM

Type Number	Message Type
0	Hello
1	Register (PIM-SM only)
2	Register-Stop (PIM-SM only)
3	Join/Prune
4	Boostrap (PIM-SM only)
5	Assert
6	Graft
7	Graft-Ack
8	Candidate-RP-Advertisement (PIM-SM only)

The following sections describe the PIM-DM control message. Refer to [50] for detailed information and format of PIM-DM control messages.

3.2.1.1 PIM-Hello Message

PIM-Hello messages are sent periodically on each interface. A router learns about the neighboring PIM routers on each interface through this message. Routers record the received Hello information from each PIM-DM neighbor. [50] Hello messages must be sent on all active interfaces, including physical point-to-point links, and it is multicast to ALL-PIM-ROUTERS group address (refer to Appendix D).

A Generation Identifier (GenID) is included in a Hello message. GenID is a randomly generated 32-bits number. It must be unique for each router, and at every restart. This identifier is used to differentiate whether neighboring PIM-DM router has been restarted. This could help to reduce the join latency. [18] This feature is going to be discussed in detail in the later section.

3.2.1.2 Prune Message

A Prune message is sent by routers towards upstream sources to prune the source tree when members leave multicast groups. In other words, Prune message is sent to upstream PIM neighbor router when no member exists for a particular multicast group. A Prune message is only sent in one hop.

3.2.1.3 Assert Message

The Assert message is used to resolve forwarder conflicts between routers on a link. It is sent when a multicast data packet is received on an interface that the router would normally forward that packet. Assert messages may also be sent in response to an Assert message from another router. [50]

3.2.1.4 Graft Message

This message is sent by a downstream router to a neighboring upstream router to reinstate a previously pruned branch of a source tree. [18] This means when a router receives a join event from a downstream router or node for a multicast group that has been pruned before, the router has to send a Graft message to the upstream router

towards all sources of that group so that the upstream router would forward the multicast group packet to the router.

The format of the message is same as a Prune message, except that the value in the type field is 6. The source address should be included in the join section of the message. The hold time field is unused, and is ignored when a graft is received.

3.2.1.5 Graft-Ack Message

A Graft-Ack message is sent in response to a received Graft message. The format is same as Prune message, except that the value of the message type field is 7. The Encoded-Unicast-Upstream Neighbor Address field is unused and needs not be checked when this message is received. The hold time field in the packet is ignored when the packet is received. [18]

3.2.2 Timers

The timers maintained in PIM-DM are countdown type timers. A countdown type timer is a timer that is set to a value which will be decreased to zero. When timer reaches zero, the timer expires and an action is triggered.

In general, from [18, 26], timers for PIM-DM can be categorized into 3 types. There are per interface timers, per entry timers, and per entry per outgoing interface timer. Table 3.6 lists all types timer of in PIM-DM according with a brief description and default value for each timer.

Table 3.6 Timers in PIM-DM

Timer		Description	Default Value (second)
Per Interface	Hello Timer	Timer between Hello message sent out by router	30
Per Entry	Data Timeout Timer	Timer that indicates lifespan for an (S, G) entry.	210

	Assert Timer	Timer that indicates lifespan for information about upstream assert winner	180
	Neighbor Timer	Timer that Indicate lifespan for a neighbor entry.	3.5 x Hello Timer
Per Entry Per Outgoing Interface	Prune Timer	Timer for a pruned interface to return into forward state	210

In addition to the above timers, [50] also defines some other time parameters that could increase the efficiency of timers handling. Some of the parameters that are implemented in PIM-DM are specified in Table 3.7.

Table 3.7 Important Parameters in PIM-DM

Value Name	Description	Default Value (second)
LAN Delay Default	Expected propagation delay over the local link.	0.5
Triggered Hello Delay	Randomized interval for initial Hello message on boot up or triggered Hello message to a rebooting neighbor	5
Assert Override Interval	Short interval before an assert times out where the assert winner resends an assert message	3

3.2.3 PIM-DM Table

In PIM-DM, three types of table are available. There are PIM-DM Multicast-Forwarding Table, PIM-DM Neighbor Information Table, and PIM-DM Membership Table. Routers refer to these tables in forwarding the multicast packets.

3.2.3.1 PIM-DM Multicast Forwarding Table

PIM-DM has a multicast-forwarding table, which stores the forwarding information for each multicast group. This table stores a list of (S, G) entry. Each (S, G) entry consists the following.

- Source address
- Multicast group address
- Incoming interface
- A list of outgoing interfaces

Initially, when a (S, G) entry is created, the entry is in forwarding state. This means the list of outgoing interfaces in the entry is not empty. A forwarding state (S, G) entry will forward the appropriate received multicast packets. If there is no outgoing interface in the list, the (S, G) entry is turned into negative cache state.

Each (S, G) entry is assigned a Data Timeout Timer. This timer is reset whenever a multicast packet is forwarded based on this (S, G) while it is deleted when Data Timeout has expired.

3.2.3.2 PIM-DM Neighbor Information Table

This table is used to store the information of PIM-DM neighbor router. The information is obtained from the exchanged Hello message. Among the information recorded are listed below.

- Interface on which the Hello message arrived
- IP address of the PIM-DM neighbor
- Generation ID of the PIM-DM neighbor
- Hello hold time

A Neighbor timer is assigned to each entry of the table. The timer is reset once a new entry update is done after receiving a Hello message. It is deleted when the timer expires.

3.2.3.3 PIM-DM Membership Table

This table stores the information of multicast groups that are joined by hosts in the networks that attach to the router. This information is notified by IGMP. The information included in the table is listed below.

- Multicast group address
- Outgoing interface

An entry is added into the table when an IGMP Notification Add is received, and is removed when an IGMP Notification Remove is received.

3.2.4 Protocol Analysis

In PIM-DM, routers decide the multicast packet forwarding by referring to the multicast forwarding table. A source does not have to give any prior notifications to the network before it sends a multicast packet to a multicast group. If a receiving router does not have any forwarding entry for a (S, G) multicast packet, a forwarding entry is created when receiving the multicast packet. Otherwise, the received multicast packet is forwarded based on the existing forwarding entry.

If there is no member downstream, the router triggers a Prune message towards the source. If a router has a negative cache entry for a (S, G) multicast packet, a Graft message may be sent when a new member joins that multicast group.

Among routers in a network, they need to communicate with each other so that they can keep track of their neighbor. This is done by sending out Hello messages periodically.

Basically, an analysis of PIM-DM can be divided into several parts. There are stated below.

- Handling PIM-DM Control Message
  - Sending and Receiving Hello Message
  - Sending and Receiving Prune Message
  - Sending and Receiving Graft and Graft-Ack Message

- Handling IGMP Notification
- Handling PIM-DM Timers
- Receiving and Forwarding Multicast Packets

The following sections describe in detail the items stated above.

3.2.4.1 Sending and Receiving Hello Message

When a router is up, it sets a Hello timer for each interface to a random value between 0 and Triggered Hello Delay [50]. This is to prevent synchronization of Hello message sending in case multiple routers are up at the same time. Routers send a Hello message through the interface when its Hello timer expires. After sending a Hello message, the Hello timer is reset. The router waits for next Hello timer alert.

When receiving Hello message from a neighbor, the router records the information about the neighbor. Routers perform checking on a neighbor list. If there is an existing neighbor entry, it updates the information of the neighbor. Otherwise, a new neighbor entry is created to keep the information.

The router resets the Neighbor Timer to the Hello hold time value specified in the message if the hold time option is used in a received Hello message. Otherwise, Neighbor timer is reset to default value. The neighbor state is deleted when the neighbor timeout expires.

On existing neighbor entry, an addition comparing process on the Generation ID for the existing neighbor entry and the newer received Generation ID is done. This is to ensure that both Generation ID have the same value. If it is not, this means the neighbor to the router has been restarted. A restarted router will lose all multicast forwarding information. For the case where the upstream router of the restarted router is in negative cache state, new joining members have to wait for longer join latency. In other words, if the upstream router is in forward state before the router goes down, the multicast packet still could reach the router after restarting. In the worst case where a new member joins

a multicast group after the router restarting, the new member has to wait for the expiration of Prune timer for that particular multicast group in the upstream router before it could receive any multicast packets.

To reduce this join latency, some actions have to be taken to recover the prune state for the PIM-DM neighbor after the detection of not conformance of Generation ID received.

According to [18], if a new GenID is received on a pruned outgoing interface of a (S, G) entry, that interface should be turned into forward state. If the (S, G) entry transitions into forward state from pruned state, a graft must be sent to the upstream RPF neighbor. On the other hand, if the new GenID is received from a RPF neighbor router, for every negative cache (S, G) entry which involve that RPF as an incoming interface, a (S, G) prune should be immediately sent to that RPF neighbor router.

3.2.4.2 Sending and Receiving Prune Message

Pruning is an action to turn down an interface so that it does not forward packets through that interface. Upon receiving a Prune message for a particular multicast group and source pair from an interface of a downstream neighbor, a router will prune the interface and starts the Prune Timer. At the same time, the router determines if there is an outgoing interface in the list of outgoing interface for that particular (S, G) entry. If there is no outgoing interface, the (S, G) entry is turned into negative cache state and the Data Timeout for that entry is set to the maximum value of the remaining Prune timer among Prune timers for all outgoing interfaces. If there is a PIM-DM neighbor upstream, a Prune message has to be sent to the upstream neighbor router.

3.2.4.3 Sending and Receiving Graft and Graft-Ack Message

A Graft message is sent to turn a pruned interface into forwarding state so that multicast packets can be received. Upon receiving a Graft message from an interface, a router has to determine the interface state of the grafted multicast group. If it is in pruned state, it must be turned into forwarding state. If this action triggers the entry from negative

cache state to forwarding state, another Graft message has to be sent towards the source of the multicast group. This process is carried on until all the pruned interfaces toward a source of a multicast group are turned into forwarding state.

Graft-Ack message is sent for every successful action in processing Graft message. A successful action means the grafted interface has been turned into forwarding state.

A timer is started for the first sending of a Graft message. If the timer expires before receiving a Graft-Ack from its upstream neighbor, the Graft message is retransmitted. Sending of Graft message is terminated only upon receiving a Graft-Ack message.

The above procedures are performed for each entry for a multicast group.

3.2.4.4 IGMP Notification

When receiving a notification from IGMP, PIM-DM will add or remove the multicast group entry in the PIM-DM Membership table. This is followed by performing some checking on the PIM-DM forwarding table. If IGMP Notification Add is received, Graft message may be sent out to ensure that the newly joined host could receive data packets from the interested multicast group. If IGMP Notification Remove is received, a pruning action is carried out to prune the interface.

A host joins a multicast group by sending an IGMP Join message to a router. IGMP would inform the PIM-DM multicast routing protocol about the new joining member by triggering an IGMP Notification Add. There may be two situations in the handling of the new joining member.

- **Joining an Existing Group**

  Since a multicast group may have multiple senders, a multicast group could have more than one entry in the multicast forwarding table. If entries exist inside the multicast forwarding table for a multicast group, the router checks through all entries for that multicast group. A checking on whether the outgoing interface towards the new member is in prune state is done. If it is in prune state, then it has

to be turned into forward state. If this action triggers the entry from negative cache state to forward state, a Graft message has to be sent out towards the source of the multicast group. This process is carried on until all the pruned interfaces toward a source of a multicast group are turned into forwarding state. (Refer to 3.2.4.3)

- **Joining a New Group**

A new joining group can be defined as a multicast group that have not been joined by downstream networks, or no single entries available in the multicast routing table. These include the previously deleted entries. If a router finds that the new member is joining a new multicast group, it records the joining group and waits for the arriving of multicast packet for the group. An (S, G) entry is created when a multicast packet is received. Then, the multicast packet is forwarded.

3.2.4.5 Handling PIM-DM Timers

Here, actions to handle PIM-DM timers are discussed. Timer handling is focused on the response to the expiration of timers. Table 3.8 shows the expiration of timers and the appropriate response event.

If (S, G) entry is in negative cache state, Prune timer expiration will cause the (S, G) entry to change state to forwarding state.

Table 3.8 Timer Expiration and Response

Expired Timer	Response
Hello Timer	Send Hello message
Data Timeout Timer	Delete expired (S, G) entry
Assert Timer	Restores RPF information obtained by the RPF lookup
Neighbor Timer	Delete neighbor entry
Prune Timer	Turn interface into forward state

3.2.4.6 Receiving and Forwarding Multicast Packet

A router is waiting for multicast packet. Upon receiving a multicast packet, the router determines whether there is an existing (S, G) entry for that particular multicast group.

If it exists, the router just forwards the multicast packet based on the (S, G) entry. Otherwise, a new (S, G) entry for that particular source and multicast group is created. Initially, the router assumes that all outgoing interfaces are keen to receive the multicast packet.

Meanwhile, the router determines whether the received multicast packet is from RPF interface. If it is, this multicast packet is forwarded through all interfaces except the incoming interface of multicast packet. At the same time, the router checks whether there are members in the subnet that are attached to the router. Multicast packets are forwarded to the subnet where there is a member. Otherwise, the interface is pruned. If the multicast packet is not coming from an RPF interface, the router discards the multicast packet silently and prunes the incoming interface.

Once a multicast packet is forwarded, Data Timeout timer for that particular (S, G) entry is reset.

3.3 Summary

The above sections cover an analysis of IGMP and PIM-DM. The first part includes IGMP messages, IGMP timer, IGMP group table, and IGMP protocol description. IGMP is implemented to handle the joining and leaving of a multicast group between host and router. An analysis of IGMP provides an overview of the IGMP module design in UMJaNetSim. On the analysis of PIM-DM, PIM-DM control messages, PIM-DM timer, PIM-DM entries table, and protocol description of PIM-DM are discussed. PIM-DM is implemented at routers to establish multicast trees for multicast packets forwarding to appropriate destination. An analysis of PIM-DM provides an overview of the PIM-DM module design in UMJaNetSim.

In next chapter, system design for the study is discussed. This includes the design for IGMP, PIM-DM, and multicast simulation application and so on.