Report 1: Data Description and Analysis

Andres Chaves (706801)
achaves@student.unimelb.edu.au
Melbourne School of Information
The University of Melbourne

March 17, 2015

Abstract

The purpose of this document is to describe and analyse the available data from which the experiments will be executed.

Throughout this document we will understand what is the data, how the data is generated, received, stored and how much is available. We will see also some examples of relations between the data.

Introduction

The objective of my Research Project is to explore how Machine Learning technologies can be used for aid the Network Management process, specifically the Fault Management process.

One of the key Information Technology components in Fault Management is the Fault Management System (FMS). A FMS is a system that receives all the alarms from the different Network Elements (NEs), performs several data transformations, and at the end stores enriched alarms and allows their visualisation:

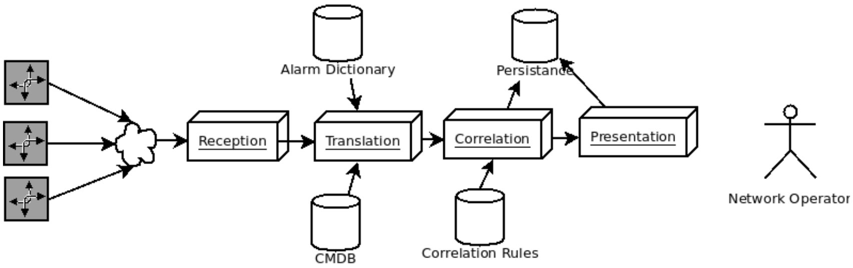


Figure 1: *A schematic of a generic Fault Management System*

The processing stages might include the following steps:

- Translation: Consist on receive the alarm in a given standard, de-codify it, translate it and enrich it into useful data. This is a required step.
- Correlation: Consist on establish rules for alarm suppression and correlation to present only key information to the network operator.

The proposed hypothesis of my Research is that given an alarm of interest an unsupervised learning algorithm can be run in order to find correlation between the alarm of interest and other alarms and therefore suppress the related alarms.

In order to conduct the experiments and analysis required for evaluating the hypothesis, it is required to have a data set from a FMS. For the scope of this project I will use the data from one FMS built in a company I worked on in the past.

We will see an introduction to the Simple Network Management Protocol(SNMP) and then some details of the specific FMS along with some insight on the network that is managed by this system. Finally, we will analyse the data by proposing five possible correlation rules.

Chapter 1

Simple Network Management Protocol

SNMP is a Layer 7 Network Protocol, specified in 1990 by the IETF, created to standardise a way to manage Network Elements. In this management protocol there are five message types and four of these are based on Request-Response interaction between a Network Management System (NMS) and a NE: GetRequest-PDU, GetNextRequest-PDU, GetResponse-PDU, SetRequest-PDU. With these four message types a NMS can read either a value or a set of values an also can write a value to override parts of the configuration of the NE. The full details of the protocol can be reviewed in RFC 1157[1].

The fifth message type, named Trap-PDU, allows a Network Element to inform the NMS about significant events or alarms in an unsolicited and asynchronous way[2]. This is the key mechanism of a typical Fault Management System.

The alarms can be either generic ones, specified by the protocol, or specific enterprise ones. Each alarm type has a specific Object Identifier (OID) to allow the system the identification of the alarm. Alarms can also have zero or more variables bindings according to the alarm type.

The set of alarm types that a given equipment can send is described in the Management Information Base (MIB) which is a collection of ASN.1 files that specify all the possible alarm types, measurements and configuration values that can be managed in the network element.

The following figure summarises the SNMP Trap mechanism:

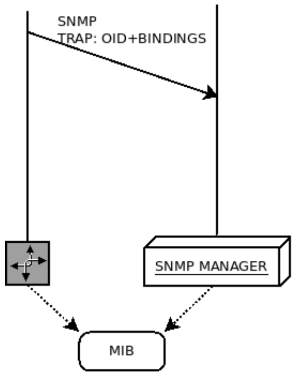
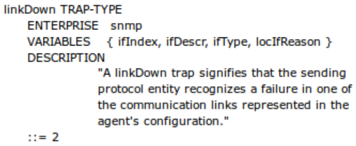


Figure 1.1: *A representation of the SNMP Trap message*

As an example of MIB specification we can see a definition of the LinkDown alarm specified by Cisco equipment vendor in CISCOGENERAL-MIB[3]:



```
linkDown TRAP-TYPE
    ENTERPRISE   snmp
    VARIABLES    { ifIndex, ifDescr, ifType, locIfReason }
    DESCRIPTION
                "A linkDown trap signifies that the sending
                protocol entity recognizes a failure in one of
                the communication links represented in the
                agent's configuration."
    ::= 2
```

Figure 1.2: *Cisco Trap Specification*

This LinkDown alarm specified in the example has the OID .1.3.6.1.2.1.11.0.3 and four variable bindings: ifIndex, ifDesccr, ifType and locIfReason.

Chapter 2

Fault Management System

Based on the overall components of a generic FMS in Figure 1 and using the SNMP protocol, we built in the company a specific FMS using several Open Source software. A schematic of the FMS, with all the open source components, can be seen in this figure:

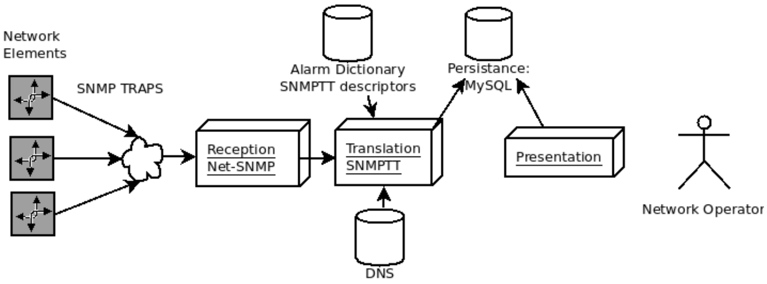


Figure 2.1: *Our built Fault Network Management System*

The FMS receives alarms using the SNMP protocol. The Open Source components used are described as follows:

- Net-SNMP: Net-SNMP is an open source Linux and Unix package that implements the SNMP protocol. The role of this component will be the reception of alarms by using the snmptrapd daemon[4].
- SNMPTT: SNMPTT or SNMP-Trap Translator is an open source component that takes the SNMP trap (alarm) received by Net-SNMP and by using an alarm dictionary (to interpret the alarm name and its attributes) and a DNS (to gather the NE name) translates it into a more useful event. SNMPTT can also enrich the alarm from a database by executing a shell script.[5]
- MYSQL: MySQL is an open source relational database. For this FMS, it was used for the storage of the Alarms[6].

Chapter 3

Network Description

In order to understand the scenarios and alarms that arrive to the FMS described in the last chapter, it is important to have an overview of the network managed by the FMS. The purpose of this chapter is to outline this network.

The network is designed hierarchically based on the Cisco hierarchical (three-layer) internetworking model[7]:

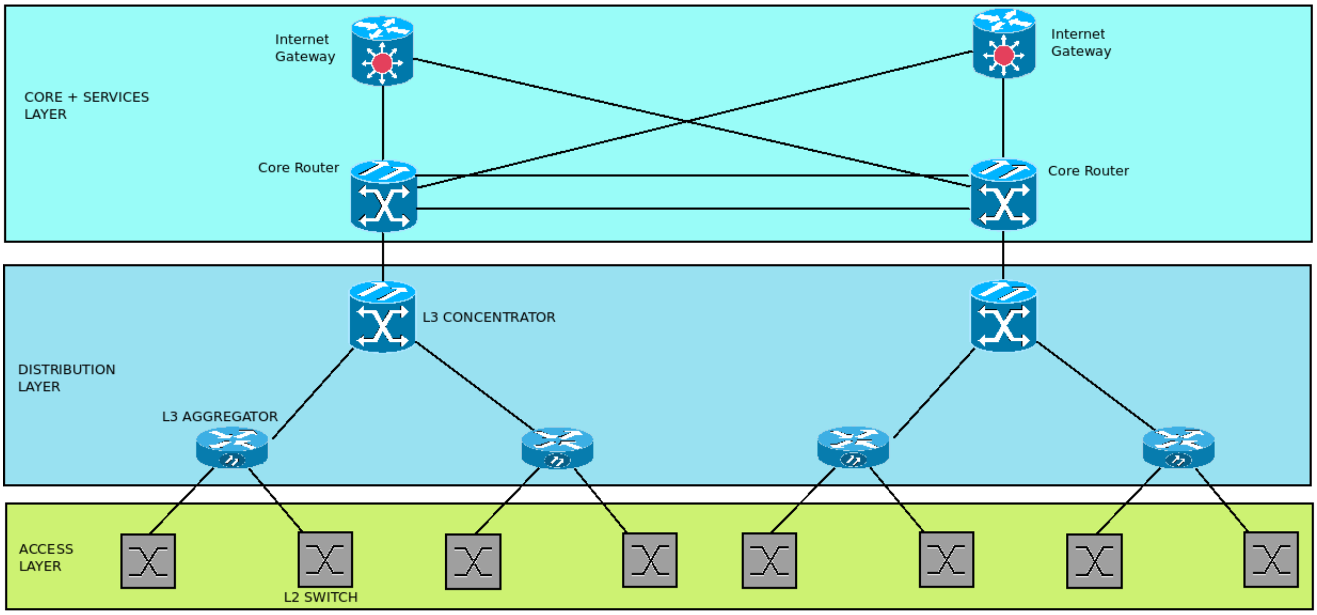


Figure 3.1: *Network Design*

It is important to mention that there are two different equipment vendors distributed geographically.

3.0.1 Access Layer

Correspond to the Layer 2 switching layer. The last miles and client services are connected to this layer. The most significant equipment of this layer is the Switch.

This is the bigger layer in terms of number of network elements, as in every town/city there should be at least one equipment. The switch is installed in an Access Node which can be thought as a small location/rack. The node is supported by a battery bank connected to a Rectifier element that converts AC to DC.

The switch has both 10G and 1G ports. The 10G optical ports are use to connect the location to a higher hierarchical equipment. The 1G ports are used to provide all the services in the town. Normally last mile equipments, such as Wifi or PointToMultiPoint (PMP) radios, are connected to these ports.

The following figure illustrates the key components on an access node:

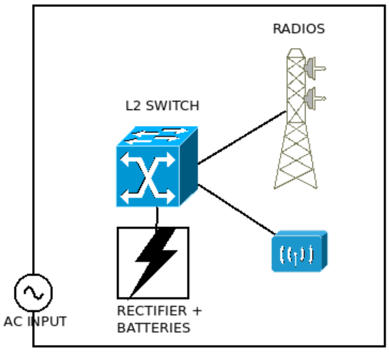


Figure 3.2: *Schematic of an access node*

3.0.2 Distribution Layer

Correspond to the Layer 3 Distribution layer. In this layer we have Routers that aggregate the traffic of several nodes and convert the Layer 2 domain into Layer 3. There are two types of node in this layer: the Aggregator node which correspond to a medium sized town that connects between 10 and 50 access nodes, and the Concentrator node which normally contains a bigger router and is installed in the main city of the state and process all the traffic in the region.

3.0.3 Core Layer+Internet Gateway

This layer is composed by big MPLS routers in charge of packet switching between all the concentrators and the different Internet Gateways or Data center services (broadband, tv, telephony, etc). These equipments are normally installed in datacenters.

3.0.4 Host Naming Convention

It is important to describe the host naming convention used for the network elements because it provides useful information. A host name is a 19 character string which follows the convention NEROLE_BRAND_MODEL_STATE_TOWN_C (i.e: AC_ER_SP10_ST_VCH_1):

- Network Element Role: Two letters that correspond to the type of node where the network element is located. The type of node can be AC (Access), AG (Aggregator), CN (Concentrator), CR (CORE), IG (Internet Gateway) and RC (Rectifier).
- Brand: Two letters that correspond to the vendor of the equipment. Some examples are: ER, AL, DT.
- Model: Four letters that codify the NE's model. Some examples are: SP10, SA30, DP29.
- State: Two letters that represent the state where the network element is physically located.
- Town: Three letters encoding the town where the network element is physically located.

Chapter 4

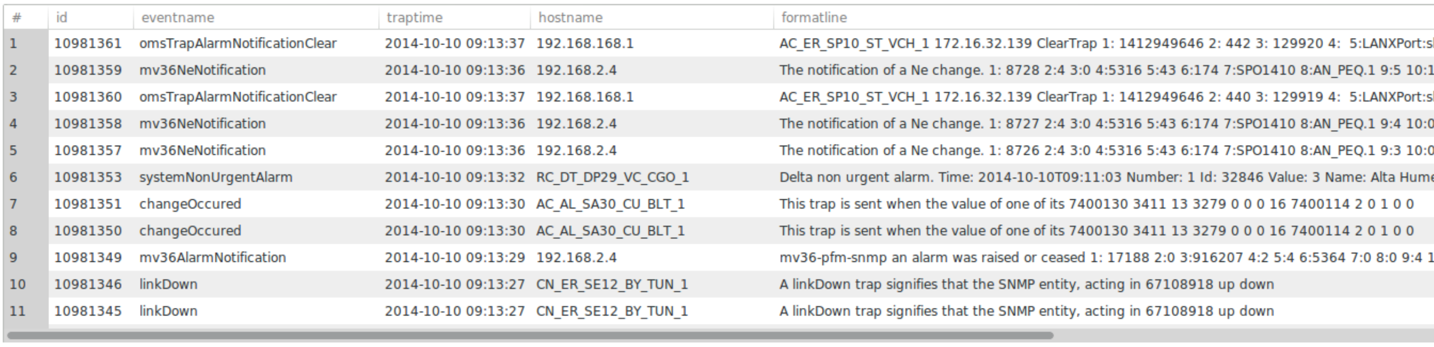
Data Specification

The FMS described in Chapter 2 stores the alarms in a table on a Relational Database System. Each month of alarms is stored in a separated table. For the purpose of this analysis we used the table for the month of October, 2014.

4.1 Data Attributes

The relevant attributes of the alarms stored are:

- Id: Unique numeric identifier of each alarm.
- Eventname: Correspond to the translated name of the alarm's OID as informed by the SNMPTT dictionary.
- Traptime: Timestamp of arrival of the trap.
- Hostname: Ip or host name of the entity that sends the alarm. A DNS is queried in order to translate the IP Address, if the translation is successful the name is stored otherwise the ip is inserted.
- Formatline: Semi-structured formatted line of the alarm. Normally it includes all the alarm's trap bindings.



#	id	eventname	traptime	hostname	formatline
1	10981361	omsTrapAlarmNotificationClear	2014-10-10 09:13:37	192.168.168.1	AC_ER_SP10_ST_VCH_1 172.16.32.139 ClearTrap 1: 1412949646 2: 442 3: 129920 4: 5:LANXPort:s
2	10981359	mv36NeNotification	2014-10-10 09:13:36	192.168.2.4	The notification of a Ne change. 1: 8728 2:4 3:0 4:5316 5:43 6:174 7:SPO1410 8:AN_PEQ.1 9:5 10:1
3	10981360	omsTrapAlarmNotificationClear	2014-10-10 09:13:37	192.168.168.1	AC_ER_SP10_ST_VCH_1 172.16.32.139 ClearTrap 1: 1412949646 2: 440 3: 129919 4: 5:LANXPort:s
4	10981358	mv36NeNotification	2014-10-10 09:13:36	192.168.2.4	The notification of a Ne change. 1: 8727 2:4 3:0 4:5316 5:43 6:174 7:SPO1410 8:AN_PEQ.1 9:4 10:0
5	10981357	mv36NeNotification	2014-10-10 09:13:36	192.168.2.4	The notification of a Ne change. 1: 8726 2:4 3:0 4:5316 5:43 6:174 7:SPO1410 8:AN_PEQ.1 9:3 10:0
6	10981353	systemNonUrgentAlarm	2014-10-10 09:13:32	RC_DT_DP29_VC_CGO_1	Delta non urgent alarm. Time: 2014-10-10T09:11:03 Number: 1 Id: 32846 Value: 3 Name: Alta Hume
7	10981351	changeOccured	2014-10-10 09:13:30	AC_AL_SA30_CU_BLT_1	This trap is sent when the value of one of its 7400130 3411 13 3279 0 0 0 16 7400114 2 0 1 0 0
8	10981350	changeOccured	2014-10-10 09:13:30	AC_AL_SA30_CU_BLT_1	This trap is sent when the value of one of its 7400130 3411 13 3279 0 0 0 16 7400114 2 0 1 0 0
9	10981349	mv36AlarmNotification	2014-10-10 09:13:29	192.168.2.4	mv36-pfm-snmp an alarm was raised or ceased 1: 17188 2:0 3:916207 4:2 5:4 6:5364 7:0 8:0 9:4 1
10	10981346	linkDown	2014-10-10 09:13:27	CN_ER_SE12_BY_TUN_1	A linkDown trap signifies that the SNMP entity, acting in 67108918 up down
11	10981345	linkDown	2014-10-10 09:13:27	CN_ER_SE12_BY_TUN_1	A linkDown trap signifies that the SNMP entity, acting in 67108918 up down

Figure 4.1: *An example of the stored data*

4.2 Data Overview

For the particular month of October 2014 we have the following statistics:

Total Alarms	2350714
Different Alarm Types	351
Different Host Names	3272
Average of Alarms Received per Minute	53
Peak of Alarms Received in a Minute	714

4.3 Data Analysis

In order to have a better insight of the data and to understand better what should be the work required by the learning algorithm, we will describe 5 different scenarios of correlation between two alarms.

For each correlation scenario we calculated the following: given in a time window of 1 minute, what is the probability of occurrence of the target alarm a_t, the probability of the second alarm a_i and the conditional probability of the alarm a_i given the occurrence of the alarm a_t.

4.3.1 Scenario 1: AC Outage + Link Down, Vendor 1

As stated in Chapter 3, the access node is composed by a Layer 2 equipment, a rectifier, a battery set, and possibly with several radios for the last miles. The radios are connected to the Ethernet ports of the switch.

The power source of the radios can be DC or AC, if it is DC and there is an AC Outage the radio will be backed up by the battery bank. Conversely, if the power source is AC and there is an AC Outage the radio will turn off and the switch equipment will register a LinkDown trap.

Therefore, the proposed scenario is the following: For vendor 1, given an AC Outage Alarm in a given town, it is expected to receive a LinkDown trap from the given town within the specified time window.

4.3.2 Scenario 2: AC Recovery + Link Up, Vendor 1

As stated in the last scenario, if the radio was connected using an AC power source, it is expected that when the AC power is restored, the radio will power on and the switch will send a LinkUp trap.

Specifically, the proposed scenario is the following: For vendor 1, given an AC Recovery Alarm in a given town, it is expected to receive a LinkUp trap from the given town within the specified time window.

4.3.3 Scenario 3: AC Outage + Link Down, Vendor 2

This is the same Scenario 1, applied to the second vendor

4.3.4 Scenario 4: AC Recovery + Link Up, Fabricator 2

This is the same Scenario 2, applied to the second vendor

4.3.5 Scenario 5: Fuse Breaker Alarm + Link Down, Vendor 1

The rectifier in the node is a smart equipment that can detect when a breaker opens in the energy distribution chain inside the node.

The last scenario analysed is the following: Given a breaker open alarm from the rectifier, it is expected to receive a LinkDown within the specified time window.

4.3.6 Summary

The following table summarises the frequencies:

Scenario	$P_t(a_t)$	$P_t(a_i)$	$P_t(a_i\|a_t)$
AC Outage + Link Down Vendor 1	0.2214	0.3886	0.11
AC Recovery + Link Up Vendor 1	0.1990	0.1228	0.117
AC Outage + Link Down Vendor 2	0.0421	0.6704	0.407
AC Outage + Link Up Vendor 2	0.0443	0.6645	0.366
Fuse Alarm + Link Down Vendor 2	0.0272	0.6704	0.027

Bibliography

[1] IETF RFC 1157, http://www.ietf.org/rfc/rfc1157.txt?number=1157.

[2] Cisco Traps Documentation, http://www.cisco.com/c/en/us/support/docs/ip/simple-network-management-protocol-snmp/7244-snmp-trap.html.

[3] Cisco General Trap MIB, http://tools.cisco.com/Support/SNMP/do/BrowseMIB.do?local=en&step=2&submitClicked=true&mibName=CISCO-GENERAL-TRAPS

[4] Net-SNMP website, http://www.net-snmp.org/.

[5] SNMPTT website, http://snmptt.sourceforge.net/docs/snmptt.shtml.

[6] MySQL website, http://dev.mysql.com/doc/refman/4.1/en/what-is-mysql.html.

[7] Cisco Networking Academy Connecting Networks Companion Guide: Hierarchical Network Design, http://www.ciscopress.com/articles/article.asp?p=2202410