

Arquitectura para el manejo de congestión en una red de datos corporativa con participación del usuario, basado en inteligencia computacional

Architecture for managing congestion in a corporate data network with user involvement, based on computational intelligence

Luis Fernando Niño*

Erick Ardila**

Joaquín F Sánchez***

Fecha de recepción: 11 de marzo de 2013

Fecha de aprobación: 30 de abril de 2013

Resumen

En este artículo se hace la presentación de un modelo de detección de congestión el cual se implementó en la red LAN de Universidad Nacional sede Bogotá. La idea de este modelo es que con ayuda de los usuarios se haga una detección de los problemas de congestión en la red, de modo que se hace también el diseño de un sistema multiagente que es la base para el modelo de detección. En el proceso se han definido cinco variables de interés las cuales definen el comportamiento de la red LAN y son presentadas en los resultados obtenidos del proceso de detección.

Palabras clave

Redes de datos, inteligencia artificial, agentes móviles.

* Ing. Electrónico. Coordinador área de redes U.N. Universidad Nacional Bogotá, Colombia. Correo electrónico: lfninov@unal.edu.co

** Ing. Electrónico y de Telecomunicaciones. Esp. Telecomunicaciones móviles Universidad Nacional Bogotá, Colombia. Correo electrónico: eardilat@unal.edu.co

*** PHD. Matemáticas de la Universidad de Memphis. Universidad Nacional Bogotá, CICOM 2013 Cartagena, Colombia, jofsanchezci@unal.edu.co

Abstract

This article is presenting a congestion detection model which was implemented in the LAN of the National University in Bogotá. The idea of this model is that with the aid of users make a detection of congestion in the network, so that it is also the design of a multi-agent system that is the basis for detection model. In the process defined interest five variables which define the behavior of the LAN and are presented in the results of the detection process.

Keywords

Data networks, artificial intelligence, mobile agents.

1. Introducción

Para los administradores de una red corporativa de gran tamaño se hace necesario el contar con herramientas de diagnóstico que ayuden a determinar los problemas físicos o lógicos que se presentan en la infraestructura [4, 5].

Dentro de los problemas lógicos que se presentan está el manejo inadecuado de la congestión de tráfico y conlleva a que la red funcione cerca de la máxima utilización que puede alcanzar. Dicha congestión aparece debido al uso compartido de los recursos de los componentes de una red, como los enlaces, conmutadores (switches) y enrutadores (routers). Si esta situación no se controla, se llegará a la saturación de dichos recursos. Esta situación deriva en un aumento en los tiempos de respuesta y se propagará teniendo como efecto un deterioro global en el desempeño de la red [4, 6].

En la actualidad existen en el mercado varios productos que ayudan a los administradores a identificar los problemas de red. Estos productos son generalmente un grupo de herramientas que mediante un afinamiento y adaptación de parámetros en la herramienta, permite monitorear los elementos componentes de una red e indicar a los administra-

dores si se presentan problemas por una alta utilización de los recursos disponibles de un equipo [2, 3]. La actividad de configuración de estas herramientas no se realiza con la periodicidad necesaria, pues cada vez que la red cambia hay que volver a configurarla y la herramienta termina no siendo utilizada adecuadamente o sin utilizar en el peor de los casos. En este trabajo de investigación se pretende ayudar en este sentido, de forma que con la ayuda de los usuarios, al utilizar la red de datos, se pueda hacer un diagnóstico de la infraestructura que ellos utilizan, y los administradores no tengan que estar modificando y configurando las herramientas de gestión de redes en sus empresas.

Los problemas de congestión repercuten en la percepción que los usuarios tienen al respecto de la lentitud de la red. Dada que esta percepción es subjetiva y depende del usuario, este trabajo busca desarrollar un modelo que les permita a los administradores de red determinar, sin prejuicios de subjetividad, una línea base para el comportamiento normal de desempeño de la red corporativa.

En este trabajo se propone un modelo que permite identificar los problemas de congestión, utilizando conceptos de sistemas multiagentes (SMA) [1, 7, 8], en una red LAN. De la misma forma, la utilización de un sistema

inteligente facilita el diagnóstico de problemas por parte de los administradores de red. El modelo se desarrolló utilizando tres tipos diferentes de agentes, el primero funciona como un sensor que recopila información de comportamiento una de las subredes, un segundo que con la información recolectada determina si hay congestión en esa subred y un tercero que indica, en una forma visual a los administradores de red, la condición de congestión en todas las subredes. También se estableció la línea base de comportamiento de las diferentes subredes en la red de área local (LAN) de la sede Bogotá de la Universidad Nacional de Colombia (UNC), mediante la recolección de datos para las variables de comportamiento y la aplicación de análisis estadístico a dichos datos.

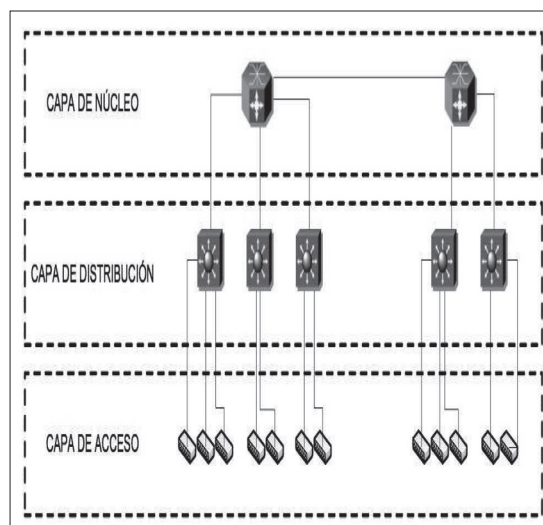
En el laboratorio se simularon condiciones de congestión, esto se realizó generando tráfico de datos mayor al que el canal de comunicaciones podía manejar y por lo tanto los recursos del equipo de red del laboratorio se saturaban, dando como resultado congestión en la subred de pruebas. Los resultados obtenidos muestran que el modelo planteado identifica adecuadamente la congestión, en un ambiente de laboratorio que hace parte de la red de área local de la UN.

En la sección 2 se hace la descripción del ambiente donde se desarrolló la investigación, para dar paso a la sección 3 en la que se hace la propuesta del modelo de detección de congestión. En la sección 4 se presentan los resultados de las pruebas hechas sobre una de las subredes en las que se hicieron las pruebas. Por último, en la sección 5 se dan las conclusiones del trabajo.

2. Ambiente de desarrollo de la investigación

Tomando en cuenta que uno de los objetivos de esta investigación es la implementación una herramienta prototipo de la arquitectura propuesta para la agrupación y clasificación de datos de medición para determinar los posibles problemas de congestión de red

Figura 1. Esquema de la red LAN de la sede Bogotá, Universidad Nacional



Fuente: elaboración propia.

se propone utilizar la red de datos de área local (LAN) de la sede Bogotá de la Universidad Nacional de Colombia (UNC).

Esta red LAN está compuesta por aproximadamente 600 equipos de red y unos 9000 computadores, la cual brinda servicios de acceso a Internet, correo electrónico, sistemas de información en red, servicios de videoconferencia, servicios con otras sedes de la UNC, entre otros. En la figura 1 se muestra el esquema de la red LAN de la sede Bogotá de la Universidad Nacional.

La topología de esta red LAN tiene tres capas, una es la capa que se denomina capa de núcleo, capa de distribución y la otra capa se denomina de acceso. La capa de núcleo está conformada por conmutadores robustos, por los cuales transita tráfico a altas velocidades; las conexiones de los conmutadores de núcleo trabajan a velocidades de GigaEthernet (1000Mbps) y DecaGiga (10000Mbps). La capa de distribución está compuesta por conmutadores con características de concentradores de fibra óptica o que concentran varias conexiones en edificios. Las velocidades de conexión de la capa

de distribución son FastEthernet (100Mbps), y GigaEthernet (1000Mbps). La capa de acceso está conformada por diferentes conmutadores y concentradores (Hub) que utilizan tecnología FastEthernet (100Megabit/s) o Ethernet (10Megabit/s) y son los que proporcionan el acceso a la red de trabajo, a los computadores y usuarios finales.

2.1 Determinación de las variables de medición de congestión

Utilizando el hecho que la congestión se presenta cuando hay una demanda mayor a los recursos disponibles en alguno de los equipos que intervienen en una transacción de red, se hizo un análisis de cada uno de los recursos y los diferentes métodos de medición de estos. Se deduce que una de las posibles variables a considerar para la medición de congestión es el tiempo de respuesta (TTL), por ser de fácil obtención desde un equipo cliente. La herramienta más simple para la medición de tiempos de respuesta en una red es mediante el uso del protocolo ICMP y de este protocolo la herramienta de ping es la más utilizada. Pero hay dos inconvenientes en cuanto a la utilización de esta herramienta:

- Dentro de una red local (LAN) los tiempos de respuesta son muy bajos, por lo general son de 1ms o menos.
- El tamaño de los paquetes que se utilizan para la herramienta ping es pequeño, de 32 bytes de tamaño predeterminado, comparado contra el tamaño máximo (MTU) del datagrama IP que es de 1500 bytes. Dado el tamaño reducido de los paquetes estos no son fácilmente afectados por la congestión.

Por lo anteriormente expuesto, la utilización de la herramienta ping es limitada en el caso de medición de congestión. Pues por sus bajos tiempos de respuesta y tamaño reducido de los paquetes, no sería posible determinar una variación significativa, se habría de realizar la medición de varios ping sucesivos, pero esto podría en un momento dado in-

crementar más la congestión y volver lentas las transacciones de los clientes. Otra posible solución para la medición de tiempos de respuesta y que pueda ser afectada por congestión es la utilización de la transferencia de un archivo, y tomar la diferencia de tiempos desde el inicio de la transmisión hasta la finalización. Para la transmisión de un archivo que siempre tiene el mismo tamaño, el tiempo de respuesta se puede asumir constante bajo condiciones de no congestión en la red. Pero de la misma forma que en la otra solución, esto incrementaría la congestión e incrementaría el tiempo de respuesta en las transacciones de red de los clientes.

Una variable más para la medición de congestión es la cantidad de paquetes perdidos. Aunque esta medición requiere de una mayor cantidad de muestras con la herramienta ping para realizar una medición adecuada. Una variable que puede ser de fácil obtención en los equipos de red es el porcentaje de utilización del ancho de banda disponible en los canales de datos. Si la utilización está por encima del 80 % de su capacidad total, puede afectar el tiempo de respuesta en las transacciones de los clientes. Por último, la cantidad de clientes en una red también afecta la lentitud en las transacciones de red.

De acuerdo con el análisis se escogieron cinco variables, estas son: tiempo de respuesta (TR), porcentaje de utilización entrante y saliente de la subred (% UI y % UO), porcentaje de paquetes perdidos (% PP) y cantidad de equipos conectados en la subred (CE). Y en cuanto a las herramientas de medición se utilizar el ICMP/ping para tomar las mediciones de TR y % PP, y se utilizar SNMP para las mediciones de % UI, % UO y CE.

2.2 Servicios por utilizar para la recolección de variables

Dado que se piensa hacer uso del cliente para la recolección de las variables mientras se hace uso normal de la red, se han de buscar los servicios que son más utilizados por los clientes en la red. A continuación se enu-

meran algunos posibles servicios a evaluar para utilizar en la toma de datos.

Servicio de PROXY para navegar. Este servicio es utilizado en la red de la UNC para hacer una autenticación del usuario y llevar una bitácora de las transacciones realizadas durante la navegación. Dado que este servicio ejecuta un script en el navegador del usuario para determinar el comportamiento del servicio, es un servicio ideal para la recolección de datos para el experimento. Este es el servicio más utilizado y solo por los clientes internos de la red local en estudio. Una de las dificultades encontradas con este servicio es que el script tiene unas utilidades muy limitadas y es muy difícil modificarlo para realizar las mediciones de tiempo de respuesta o paquetes perdidos, por lo que no es recomendable utilizarlo.

Página web principal de la universidad. Al ser la página principal de la universidad es consultada frecuentemente, pero no solo por los clientes internos a la red local, sino que también es necesario filtrar estos accesos. Ya que es una página web hay varias formas para realizar las mediciones con un script, con lenguajes como java, php, o activex. Debido a la variedad de navegadores y los diferentes niveles de seguridad en estos, es más complicado ejecutar un script en el equipo del cliente que en el equipo servidor que brinda el servicio. Con diferentes pruebas realizadas se evidenció que ejecutar un script en el lado del cliente es más difícil por la seguridad incorporada en los navegadores, mientras que en el lado del servidor, aunque implica más carga al servidor, siempre se ejecuta independiente del tipo navegador o el nivel de seguridad.

Servicio de correo electrónico. Este es otro de los servicios más utilizados y también garantiza que los usuarios sean exclusivamente de la universidad, aunque no todos están dentro de la red local de la universidad. Hay dos formas de acceder al servicio de correo de la universidad, utilizando un cliente (p.ej. outlook) o por una página web. Dado que los clientes de correo son software propietario

es casi imposible modificarlos para tomar las mediciones del experimento. Sin embargo, la página web tiene las mismas ventajas que la página principal de la universidad y se pueden ejecutar scripts en el lado del servidor.

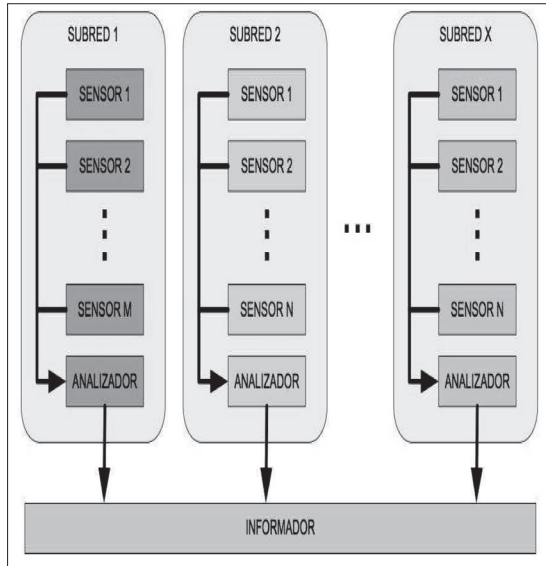
Servicio de mensajería instantánea (chat). También es uno de los servicios mayormente utilizados por los clientes. Pero la desventaja es que son programas propietarios y no son fácilmente modificables para este experimento.

3. Propuesta del modelo de detección de congestión

Como sea mencionado anteriormente la red de la sede Bogotá de la Universidad está compuesta por diferentes subredes. Este tipo de separación en subredes no es solo una característica exclusiva de la red de la UNC; en general, es una mejor práctica, porque facilita la administración de la red separando segmentos lógicos de una red de área local (por ejemplo por los departamentos de una empresa) que no deberían intercambiar datos usando la red local. Adicionalmente brinda una independencia de tráfico entre las subredes. Dado que el modelo de detección de congestión ha de funcionar para todas estas subredes, se hace necesario implementar en el modelo, uno o varios recolectores (sensor) de información para cada subred. De la misma forma debe existir un analizador de esta información por cada subred. Por último, un componente informa el estado general de la red LAN, que recopila la información de todas las subredes. En la figura 2 se muestra el esquema global de detección de congestión.

En la figura 2 se muestra el modelo propuesto. Para cada subred hay uno o varios sensores, estos sensores son invocados por los usuarios de cada uno de las redes al utilizar el servicio de una página web en la red. Al ser invocado este sensor toma el valor de las variables TR, % UI, % UO, % PP y CE en ese instante, y envía estos valores al analizador de esa subred. El analizador tiene los valores de los umbrales para cada una de estas varia-

Figura 2. Esquema global de detección



Fuente: elaboración propia.

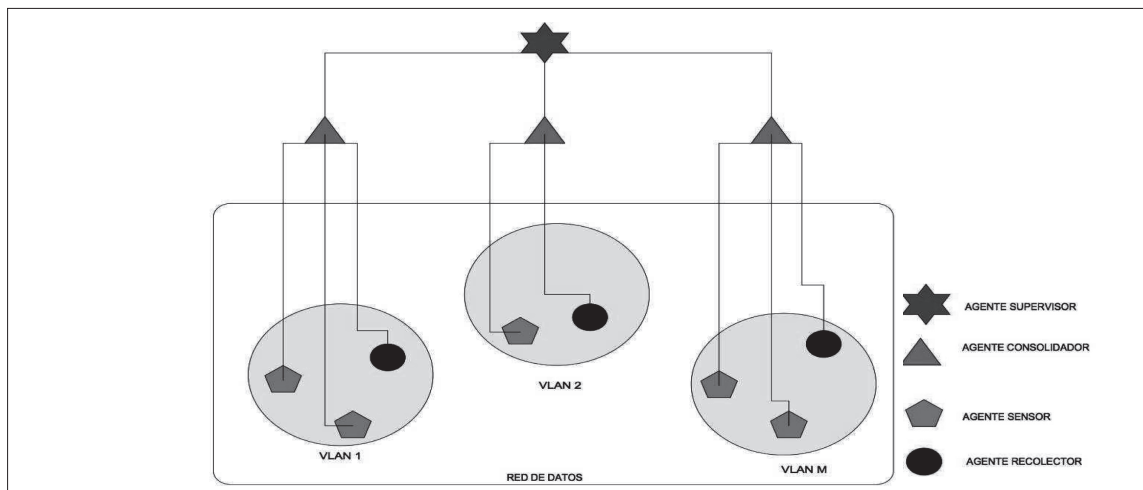
bles y al recibir una nueva información de un sensor, toma la decisión al respecto de si esa subred está o no congestionada. Por último, el analizador envía la información del estado de la subred al informador, este se encarga de mostrar la información del estado de todas las subredes a los administradores de red.

3.1 Diseño del Sistema Multiagente

De las ventajas de un sistema multiagente y para esta propuesta de modelo, se hace evidente que la velocidad y escalabilidad son factores importantes para el correcto funcionamiento del modelo [9, 8]. La velocidad permitirá que el modelo reaccione oportunamente a los cambios en congestión presentados en la red, dado que los tiempos de respuesta en una red LAN son de 1 milisegundo o menos, es necesario que el modelo tome decisiones e informe en forma oportuna. La escalabilidad también es una característica clave, pues las redes LAN pueden variar en tamaño entre una compañía y otra y por lo tanto la cantidad de usuarios y equipos en la red también. El modelo debe ser capaz de ajustarse fácilmente a los diferentes tamaños de redes de área local. El sistema multiagente implementó utilizando la herramienta JADE (Java Agent Development Framework). Del esquema general del modelo propuesto se haría la implementación utilizando agentes del tipo colaborativos y reactivos. El sistema multiagente propuesto es del tipo de interacción simple entre agentes.

En la figura 3 se muestra el esquema de sistema multiagente propuesto para el modelo de detección de congestión.

Figura 3. Esquema del sistema multiagente para detección de congestión



Fuente: elaboración propia.

En la figura 3 se observa que hay cuatro tipos de agentes en el modelo particular para la detección de congestión. El primer tipo de agente es el supervisor, en el esquema global del modelo se denomina informador, se encarga de mostrar la información del estado de todas las subredes a los administradores de red, de este agente solo hay 1 en el sistema. El segundo tipo de agente es el consolidador, en el esquema global del modelo se denomina analizador, al recibir una información de las variables de red de un agente sensor o un recolector, toma la decisión al respecto de si esa subred está o no congestionada, de este agente hay uno por cada subred. El tercer tipo de agente es el sensor, en el esquema global del modelo se denomina sensor, cuando el usuario utiliza el servicio de red seleccionado, este agente toma el valor de las variables TR, % UI, % UO, % PP y CE en ese instante, y envía estos valores al agente consolidador de esa subred, puede haber varios de estos agentes por subred, dependiendo de la cantidad de usuarios que utilicen el servicio. El cuarto tipo de agente es el recolector, en el esquema global del modelo se denomina también como sensor, a diferencia del agente sensor éste no se invoca en el momento en que un usuario utiliza el servicio de red, este agente es invocado por el agente consolidador si no ha recibido información de algún sensor durante un periodo de tiempo determinado, este sensor toma el valor de las variables TR, % UI, % UO, % PP y CE en ese instante, y envía estos valores al agente consolidador de esa subred, de este agente puede haber uno por cada subred.

A continuación se muestran algunas de las funciones de cada uno de los agentes del sistema:

- **Sensor.** Determinar el cliente a que subred pertenece. Ejecutar la recolección de variables de congestión en forma pasiva. Enviar la información de las variables de la prueba al agente consolidador de esa subred.
- **Recolector.** Ejecutar la recolección de variables de congestión en forma activa.

Enviar la información de las variables de la prueba al agente consolidador de esa subred.

- **Consolidador.** Agrupar y clasificar los datos recibidos por los agentes recolectores y sensores, de los clientes en una subred. Guarda información de los clientes en cada subred. Aplicar el algoritmo de decisión para determinar congestión en la subred. Si un cliente lleva tiempo sin responder, solicita al agente recolector realice prueba activa. Enviar información de la subred al agente supervisor.

3.2 Recolección de datos

Para la determinación de la línea base de comportamiento de las diferentes variables de comportamiento de red, se realizó la siguiente recolección de datos. Se tomaron muestras de 11 subredes de la red en estudio. La selección de estas subredes se hizo tomando en cuenta los siguientes atributos:

- **Tamaño de la subred,** en cuanto a la cantidad de usuarios y equipos. En la sede Bogotá hay redes de 1024, 512 y 256 equipos, y se escogieron subredes de cada uno de los tamaños.
- **Dado que es necesarios que las muestras se tomen 24 horas al día, siete días a la semana,** entonces es necesario que al interior de la subred exista un equipo servidor que pueda responder a las solicitudes durante el periodo de recolección de variables.
- **La distancia física de una subred en la red de la sede Bogotá altera los tiempos de respuesta obtenidos,** y dado que el campus de la sede es bastante amplio en área física, se hace necesario utilizar subredes cuya distribución física sea por todo el campus de la sede Bogotá.

Se tomaron muestras en cada una de las subredes de las cinco variables de interés para determinar congestión: tiempo de respuesta, porcentaje de utilización entrante y saliente

de la subred, porcentaje de paquetes perdidos y cantidad de equipos conectados en la subred. Se tomaron muestras cada cinco minutos en cada una de las once subredes, durante un periodo de 99 días (desde el día 15 diciembre 2011 al 22 marzo 2012), las 24 horas, todos los días de la semana. Fueron 23585 muestras recolectadas.

La obtención de cada variable se hizo utilizando el mecanismo descrito a continuación:

- Tiempo de respuesta y porcentaje de paquetes perdidos: se utilizó la herramienta ICMP/ping, con una secuencia de veinte pruebas hacia los servidores en cada subred y se toma el promedio del tiempo de respuesta en milisegundos obtenido y la cantidad de paquetes pedidos.
- Porcentaje de utilización entrante y saliente de la sub-red: se utilizó la herramienta SNMP para su recolección. Esta muestra se toma en el equipo enrutador principal de la red en estudio, el cual gestiona cada uno de los canales de comunicación de las diferentes subredes. Para cada una de las subredes se consulta por SNMP los valores de *averageInput*, *averageOutput* y *ifSpeed*, estos representan el promedio de utilización de entrada, el promedio de utilización de salida en bps de los últimos 5 minutos y la capacidad del canal en bps de la subred, respectivamente. Con estos valores se hace aplicó la siguiente fórmula para calcular el porcentaje de utilización tanto entrante como saliente del canal para cada VLAN.

$$\%UI = \frac{\text{averageInput}}{\text{ifSpeed}} \times 100 \quad (1)$$

$$\%UI = \frac{\text{averageOutput}}{\text{ifSpeed}} \times 100 \quad (2)$$

- Cantidad de equipos conectados en la subred: se utilizó la herramienta SNMP para su recolección. Para cada una de las subredes se consulta por SNMP los valores de *dot1dTpFdbTable*, es la tabla donde está la información de todas las direc-

ciones unicast que ha detectado el equipo enrutador principal de la red en estudio. Para obtener la cantidad de equipos por cada subred se cuenta la cantidad de líneas de la tabla.

Se realizó un análisis estadístico a las muestras recolectadas. La primera prueba a realizada fue un análisis de correlación entre las diferentes variables, este análisis ayudó a determinar si el comportamiento de una de estas variables se ve o no afectada por el comportamiento de las demás. La siguiente prueba que se realizó fue un análisis de varianza de las variables entre subredes, con el fin de determinar la independencia de comportamiento entre las subredes estudiadas. Por último se realizó un análisis de estadística descriptiva básica de las variables, con el fin de obtener los valores del mínimo, primer cuartil, mediana y tercer cuartil para las muestras de cada una de las subredes. Con estos valores se determinaron los umbrales para las variables en el algoritmo para determinar congestión. De este análisis se obtuvieron los siguientes resultados:

- De las cinco variables seleccionadas (TR, % UI, % UO, % PP y CE) para determinar congestión, se presenta una alta correlación entre las variables de porcentaje de utilización entrante y saliente de la subred. Esto indica que para el modelo se podrían utilizar solo cuatro de estas variables propuestas, pero se toma la decisión de continuar trabajando con las cinco originalmente definidas.
- El comportamiento de cada subred es independiente de todas las demás, lo cual significa que problemas de congestión en una subred, no genera congestión en las otras.
- Por el comportamiento observado por los administradores de red de la sede Bogotá, se toma la decisión de hacer una separación en tres grupos de todas las muestras obtenidas para cada variable así: primer grupo, muestras obtenidas en días laborales (lunes a viernes) y en horario laboral (de 7 a.m 7 p.m.); segundo grupo,

muestras obtenidas en días laborales (lunes a viernes) y en horario no laboral (7 a.m 7 p.m.); y el tercer grupo, muestras obtenidas los fines de semana (sábado y domingo) durante todo el día.

- Después del análisis descriptivo, para cada una de las variables y por cada grupo de los mencionados en el punto anterior, se obtuvo los valores para el mínimo, el primer cuartil, la mediana y el tercer cuartil. Si el % PP es mayor que cero se supone congestión y por lo tanto no es necesario hacer análisis estadístico. Estos valores son utilizados por el algoritmo de toma de decisión de congestión, explicado en la próxima sección.

Es importante aclarar que por variaciones en las subredes, como por ejemplo la cantidad de equipos, es necesario realizar este mismo procedimiento de toma de muestras y obtención de valores en forma periódica. La periodicidad será determinada por los administradores de red, y su experiencia al respecto de cada cuánto tiempo va variando las subredes en la red. Para el caso de la universidad, los administradores de red consideran que este procedimiento de recolección de variables ha de hacerse por lo menos cada semestre académico.

3.3 Algoritmo para determinar la congestión

A continuación se explica el algoritmo que se utilizará para la toma de decisión de si hay o no congestión a partir de las variables obtenidas en cada una de las subredes.

- De las muestras obtenidas mediante el mecanismo explicado se obtuvo para las variables de tiempo de respuesta un porcentaje de utilización entrante y saliente de la subred, y cantidad de equipos, su valor mínimo, primer cuartil, mediana y tercer cuartil.
- Para cada una de las variables se determinará un indicador de congestión, lo cual se representa utilizando la siguiente

escala de colores: verde indica un valor normal para esa variable, amarillo indica que la variable muestra un poco de congestión y rojo indica congestión. Ver la figura 4.

- Para la variable porcentaje de paquetes perdidos (% PP), si esta es mayor que cero entonces hay congestión y su color es rojo; si es cero, entonces es normal y su color es verde. Esta variable no toma el color amarillo.
- Para determinar si hay o no congestión en una subred se usan cinco variables, cada una toma un resultado de la tabla 1. Si dos o más variables están en rojo, entonces hay congestión en esa subred. Si hay una roja y dos amarillas, entonces

Tabla 1. De decisión para congestión en una subred

Rojos	Amarillos	Verdes	Resultado
5	0	0	Congestión
4	1	0	Congestión
4	0	1	Congestión
3	2	0	Congestión
3	1	1	Congestión
3	0	2	Congestión
2	3	0	Congestión
2	2	1	Congestión
2	1	2	Congestión
2	0	3	Congestión
1	4	0	Congestión
1	3	1	Congestión
1	2	2	Congestión
1	1	3	Poca Congestión
1	0	4	Poca Congestión
0	5	0	No se Puede
0	4	1	Poca Congestión
0	3	2	Poca Congestión
0	2	3	Poca Congestión
0	1	4	Normal
0	0	5	Normal

Fuente: elaboración propia.

Figura 4. Determinación del color para cada variable

RANGO	CONGESTIÓN
Mayor al tercer cuartil	CONGESTIÓN
Entre la mediana y el tercer cuartil	POCA CONGESTION
Entre el primer cuartil y la mediana	NORMAL
Entre el mínimo y el primer cuartil	

Fuente: elaboración propia.

hay congestión en esa subred. Si hay dos o más amarillas o hay una roja, hay un poco de congestión. Para el resto de las opciones se supondrá que la subred no presenta congestión. (ver tabla 1)

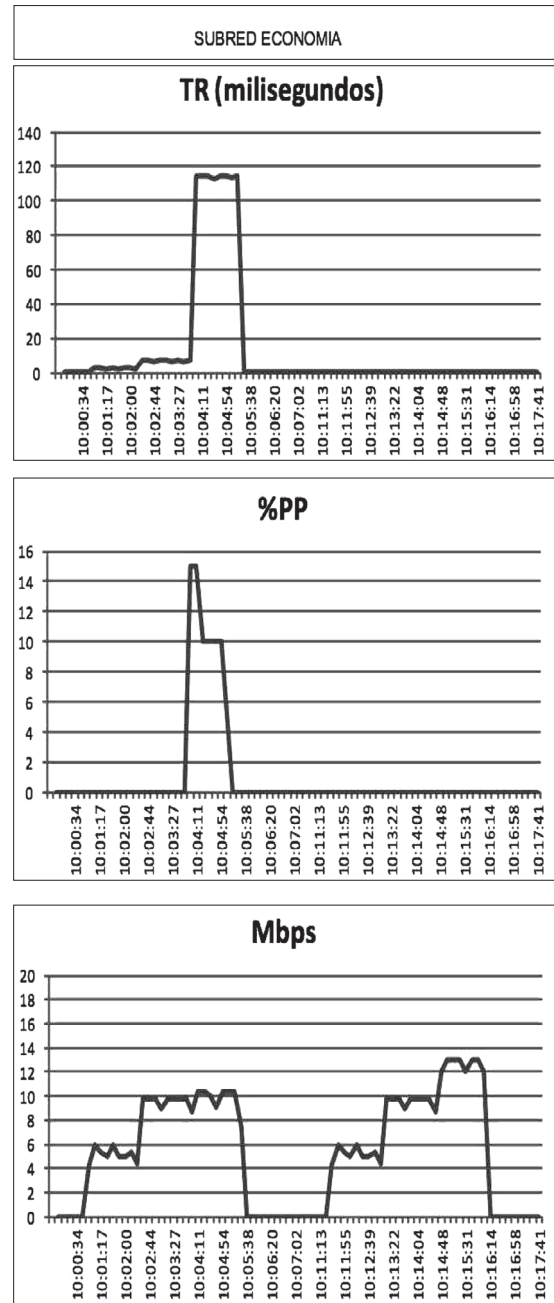
En la tabla anterior se muestran todas las combinaciones posibles de los colores para las cinco variables y el resultado del nivel de congestión para la subred.

4. Resultados de las pruebas de detección

En esta sección se presentarán los resultados obtenidos del laboratorio y se comprobará si el laboratorio simula, en un entorno controlado, la congestión en una red. Ya que a las muestras obtenidas se les hizo un análisis estadístico diferenciando tres grupos (horario laboral, no laboral y fin de semana), entonces el laboratorio se ejecutó para cada uno de esos horarios. A continuación se muestran los resultados obtenidos de las pruebas efectuadas el día 25 de octubre de 2012 a las 10 a.m. En la figura 5 se muestran las gráficas de las variables TR, % PP y consumo de canal, para la subred de economía.

En la figura 5 durante la primera parte del laboratorio, cuando el ancho de banda del canal de la subred de economía estaba reducido, se observa que los tiempos de respuesta subieron hasta casi los 120 ms, hubo un porcentaje de paquetes perdidos (% PP)

Figura 5. Resultados del laboratorio de la subred de economía



Fuente: elaboración propia.

no pasa del 15 % y el consumo del canal no pasó de los 10 Mbps. Dado que los valores normales para esta subred de los tiempos de respuesta son menores a los 0,5 ms, se tendría que la variable TR obtendría un co-

lor rojo por congestión. El mismo resultado obtendría la variable % PP, color rojo. Si se aplica el modelo se obtendrían dos variables en rojo y, por lo tanto, se indicaría congestión en la subred de economía. Para la segunda parte del laboratorio, los tiempos de respuesta y porcentaje de paquetes perdidos permanecieron en valores cercanos o iguales a cero, lo cual indica que la subred de economía presenta un comportamiento normal y no hay congestión.

4.1 Pruebas del modelo de detección

En esta sección se presentarán los resultados obtenidos de la prueba del modelo en el laboratorio y se comprobará, en un entorno controlado, si el modelo detecta adecuadamente congestión en las subredes del laboratorio. Durante las pruebas realizadas en el laboratorio este servidor adicionalmente iba guardando una bitácora con los eventos de recepción de información de la diferentes subredes y la toma de decisión al respecto de si existía o no congestión. La figura 6 presenta una muestra de una de la bitácora de eventos del modelo.

5. Conclusiones

El modelo hizo una exitosa identificación de congestión en las subredes de la red de la universidad, utilizando cinco variables definidas y que ayudan a modelar el comportamiento de las subredes. De las pruebas estadísticas aplicadas a las muestras de las variables se obtuvo que de las cinco variables es posible utilizar solo tres de estas y el modelo continuaría funcionando adecuadamente. Dado que tres de las variables tienen una alta correlación entre ellas, indica que el comportamiento de una de ellas afecta a las otras dos, y por lo tanto, para el modelo sería suficiente utilizar una sola de estas. Pero esta conclusión solo fue evidente luego de aplicar diferentes pruebas estadísticas a las muestras obtenidas.

El concepto de sistema multiagentes desarrollado permite asignar adecuadamente las tareas que son necesarias para la identificación de congestión en redes. También permite establecer un sistema distribuido, el cual facilita la identificación segmentada de comportamientos que indican congestión.

Figura 6. Muestra de los registros de eventos de las pruebas

fecha	hora	ip	VLAN	%PP	TR	%UI	%UO	CE	estado
4/01/13	16:51:16	168.176.26.57	Ingeniería	0	0,281	0,2733	0,1082	89	NORMAL
4/01/13	16:51:16	168.176.16.181	Economía	0	114.393	0,4396	0,3687	33	POCA-CONGESTION
4/01/13	16:51:19	168.176.85.124	Unisalud	0	0,347	0	0	101	NORMAL
4/01/13	16:51:21	168.176.54.143	Ibun	0	0,287	0,0029	0,0587	22	NORMAL
4/01/13	16:51:28	168.176.26.57	Ingeniería	0	0,267	0,2751	0,1063	90	NORMAL
4/01/13	16:51:28	168.176.16.181	Economía	10	115.257	0,51	0,3657	33	CONGESTION
4/01/13	16:51:30	168.176.85.124	Unisalud	0	0,352	0	0	101	NORMAL
4/01/13	16:51:32	168.176.54.143	Ibun	0	0,312	0,0029	0,0618	22	NORMAL
4/01/13	16:51:39	168.176.26.57	Ingeniería	0	0,277	0,2698	0,1027	90	NORMAL
4/01/13	16:51:40	168.176.16.181	Economía	10	113.702	0,5232	0,3536	33	CONGESTION
4/01/13	16:51:42	168.176.85.124	Unisalud	0	0,342	0	0	101	NORMAL
4/01/13	16:51:43	168.176.54.143	Ibun	0	0,305	0,0028	0,0588	22	NORMAL

Fuente: elaboración propia.

En el modelo se utiliza el principio que los usuarios finales al hacer uso de la red para sus actividades cotidianas, pueden ayudar a identificar problemas como fallas, congestión o lentitud. Pero dado que el obtener información de los usuarios finales no es una labor simple, se plantea en el modelo hacer pruebas en la red para identificar los diferentes problemas, pero sin que el usuario tenga que ejecutar alguna acción adicional a su normal comportamiento. Se genera un sistema distribuido natural, dado que los usuarios están dispersos en las diferentes subredes. Esto también simplifica la tarea de los administradores de redes, ya que no se hace necesario instalar programas sensores en cada subred, sino que la distribución natural de los usuarios permite crear un sistema distribuido robusto.

La precisión en la medición de congestión está fuertemente atada a la cantidad de usuarios en una subred y al tipo de utilización que hagan de la red los usuarios. Al respecto a la cantidad de usuarios, hay una relación directa entre el número de usuarios en la subred y que tanta información recibe de esta subred el modelo, por lo tanto, entre mayor cantidad de usuarios, mayor cantidad de variables recolectadas y mejor las decisiones de si hay o no congestión. Es posible que el modelo en una red con pocos o ningún usuario activo, no detecte adecuadamente un problema de congestión o lo detecte cuando ya sea muy tarde. Adicionalmente, si el servicio de red escogido en el modelo para disparar la recolección de datos, no es usualmente utilizado por los usuarios, por ejemplo que los usuarios prefieran utilizar correos electrónicos externos, entonces aunque haya muchos usuarios activos, para el modelo el comportamiento es similar a una red en donde no hay usuarios y por lo tanto habrá una pobre detección de la congestión.

6. Agradecimientos

Los autores agradecen al centro de cómputo de la Universidad Nacional sede Bogotá,

ya que brindo amplio apoyo en la construcción del ambiente controlado en sus instalaciones.

7. Referencias

- [1] C. H. Adolfo López Paredes. "Sistemas Multiagente en Ingeniería de Organización. Técnicas Computacionales de Simulación de Sistemas Complejos". II Conferencia de Ingeniería de Organización., Vigo, Espasa. 2002.
- [2] L. P. y. D. J. Baydal, E. "A family of mechanisms for congestion control in wormhole networks. Parallel and Distributed Systems", *IEEE Transactions*, vol. 16, num. 9: 772-784, 2005.
- [3] K. D. S. M. . S. P. Calyam, P. "Active and passive measurements on campus, regional and national network backbone paths". *Computer Communications and Networks*, vol. 1, num. 1: 537-542, 2005.
- [4] S. Floyd and K. Fall. "Promoting the use of end-to-end congestion control in the internet". *Networking, IEEE/ACM Transactions*, vol. 7, num. 4: 458-472, 1999.
- [5] R. Jain, "Congestion control in computer networks: Issues and trends". *IEEE Network*, vol. 4, num. 3: 24-30, 1990.
- [6] S. Keshav. Congestion Control in Computer Networks. PhD Thesis. s.l., UC Berkeley, 1991.
- [7] P. Maes, "Artificial life meets entertainment: Life like autonomous agents". *Communications of the ACM*, vol. 38, num. 11: 108-114, 1995.
- [8] C. A. Nikolai, "Tools of the trade: A survey of various agent based modeling platforms". *Journal of Artificial Societies and Social Simulation*, vol. 12, núm. 2: 1-5, 2009.
- [9] . N. P. Russell, S. *Inteligencia Artificial, un Enfoque Moderno*. Pearson Prentice Hall, Madrid, Espasa, 2004.