Sistema de detección de anomalías de red basado en monitorización y predicción de tráfico

A System for Detecting Network Anomalies based on Traffic Monitoring and Prediction

P. Barlet, H. Pujol, J. Barrantes, J. Solé y J. Domingo

Resumen

SMARTxAC es un sistema de monitorización y análisis de tráfico para enlaces de alta velocidad. Este artículo describe su módulo de detección de anomalías y propone el uso de algoritmos de predicción adaptativa de tráfico como una técnica efectiva para la detección de anomalías de red en tiempo real.

Palabras clave: Detección de anomalías, monitorización pasiva, predicción de tráfico

Summary

SMARTxAC is a passive monitoring and analysis system for high-speed links. This paper describes its anomaly detection module and proposes the usage of adaptive traffic prediction as an effective technique to detect network anomalies in real-time.

Keywords: Anomaly detection, Passive monitoring, Traffic prediction

SMARTxAC es un sistema de monitorización y análisis de tráfico para enlaces de alta velocidad

1.- Introducción

Tradicionalmente, los sistemas de monitorización de tráfico han sido de gran ayuda para las tareas clásicas de gestión y administración de la red. Además, con la reciente proliferación de usos irregulares de la red, debida en gran parte a la progresiva universalización de las redes conectadas a Internet y a la diversificación de su uso, estos sistemas se han convertido también en herramientas muy útiles para la detección de anomalías de red.

En la actualidad estos sistemas se han convertido también en herramientas muy útiles para la detección de anomalías

Consideramos como anomalías de red a aquellas situaciones en que el tráfico de una red difiere notablemente de su perfil habitual, teniendo en cuenta que el tráfico ya presenta intrínsecamente una serie de variaciones debidas principalmente a la diversidad de usuarios, servicios y aplicaciones, así como a los ciclos temporales diarios y semanales.

Las anomalías se deben normalmente a causas físicas (como puede ser la caída de un enlace); a situaciones concretas en las que se producen aumentos de tráfico temporales hacia determinados destinos (por ejemplo porque se haya publicado un contenido exclusivo en un servidor), o a ataques deliberados que tienen como objetivo una intrusión o una denegación de servicio en un sistema, ya que todos estos tipos de acciones, en general, producen cambios considerables en el comportamiento habitual de una red.

Existen dos enfoques distintos que tienen como finalidad la detección de situaciones anómalas. Por un lado, la detección de intrusiones de red, que normalmente consiste en la búsqueda de determinados patrones o cadenas de bits que aparecen en algunos ataques conocidos dentro de la información contenida en los paquetes capturados, como en el caso de SNORT y de otros sistemas de detección de intrusiones de red (NIDS). Normalmente, estos sistemas están diseñados para proteger equipos finales o redes pequeñas.

Por otro lado, mediante la localización de desviaciones apreciables respecto al perfil de trafico habitual de la red. Tradicionalmente, este proceso se ha realizado mediante la revisión manual de la

información ofrecida por los equipos de red, como SNMP o NetFlow, o de las estadísticas ofrecidas por sistemas de monitorización de tráfico, como MRTG, CoralReef o Ntop. No obstante, este mecanismo de por sí es poco efectivo y requiere una elevada intervención humana.

Basado en este segundo enfoque, el objetivo de este trabajo es la detección de manera automática de aquellas anomalías que puedan degradar el rendimiento global de redes de alta velocidad en las que participan miles de usuarios.

La siguiente sección describe los dos métodos utilizados en el sistema SMARTxAC para la detección de anomalías (mediante predicción de tráfico y mediante umbrales) y además presenta la clasificación por tipos de anomalías de red utilizada por el sistema. La tercera sección muestra los resultados de detección de anomalías obtenidos en un entorno real, y finalmente, la cuarta sección resume las características del sistema y algunas posibles ampliaciones.

Los dos métodos utilizados en el sistema SMARTxAC para la detección de anomalías son mediante predicción de tráfico y mediante umbrales

2.- Detección de anomalías de red en el sistema SMARTxAC

SMARTxAC es un sistema de monitorización continua y de análisis de tráfico en tiempo real para enlaces de alta velocidad (1 Gb/s o más) desarrollado en el marco de un convenio de colaboración entre el CESCA y la UPC. Desde julio de 2003, este sistema está siendo utilizado por el CESCA para la monitorización permanente de la *Anella Científica*. Una descripción detallada del sistema SMARTxAC puede encontrarse en [1].

Para facilitar la laboriosa tarea de revisión de las estadísticas generadas por el sistema en busca de situaciones irregulares se ha desarrollado un módulo de detección automática de anomalías para el sistema SMARTxAC basado en técnicas de predicción de tráfico.

Dado que SMARTxAC monitoriza un enlace troncal, interesa la detección de aquellas anomalías que impliquen una degradación notable del servicio más desde el punto de vista del funcionamiento de la red que desde el de la protección de los equipos finales, minimizando el número de falsos positivos e intentando apuntar cuáles pueden ser sus causas. Concretamente, estamos interesados en detectar aquellas anomalías que puedan provocar cambios en el patrón habitual de tráfico de alguna de las instituciones conectadas a la red monitorizada.

Dado que SMARTxAC monitoriza un enlace troncal, interesa la detección de aquellas anomalías que impliquen una degradación notable del servicio

Sin embargo, el tráfico de Internet comporta de manera inherente múltiples variaciones, tanto a diferentes escalas temporales como por su heterogeneidad, de manera que la caracterización del patrón habitual de tráfico de una institución es muy compleja. Así pues, parece viable detectar cambios inesperados en el perfil de tráfico mediante técnicas de predicción que permitan resaltar diferencias significativas entre el tráfico real medido y el predicho para un mismo intervalo de tiempo. La principal ventaja de estas técnicas radica en que no es necesario un conocimiento explícito a priori sobre el patrón de tráfico habitual de cada institución. Además, su implementación es suficientemente sencilla y ligera para ser aplicada en tiempo real.

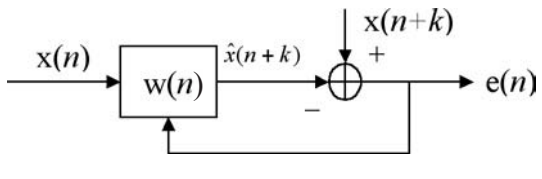
2.1.- Detección de anomalías basada en predicción de tráfico

El algoritmo de predicción utilizado en SMARTxAC está basado en un filtro lineal adaptativo, muy utilizado en aplicaciones de procesado de voz, la descripción detallada del cual puede encontrarse en [2]. Su principio de funcionamiento, que esquematiza la figura 1, es el siguiente: el valor predicho de una serie para un determinado intervalo de tiempo, $x(n+k)$ se calcula como una combinación lineal de las últimas p medidas de tráfico, dando más peso a las observaciones más recientes.

A cada iteración del algoritmo, a partir de la diferencia entre los valores predichos y los medidos realmente, $e(n)$, los p coeficientes del filtro se actualizan para minimizar el error cuadrático medio; de esta manera, se adapta al comportamiento del tráfico en cada momento.

El valor de μ debe encontrarse entre 0 y 2: los valores altos proporcionan una adaptación más rápida a los cambios, mientras que los valores cercanos a 0 conceden más relevancia a las observaciones más antiguas y en consecuencia proporcionan una convergencia más lenta. Tanto μ como los coeficientes del filtro influyen en gran medida en la capacidad de predicción del algoritmo, y su ajuste se ha efectuado para minimizar el número de falsos positivos. Esto puede implicar que las anomalías más discretas no se detecten, pero es muy probable que éstas no afecten de manera relevante al comportamiento de la red.

FIGURA 1: ESQUEMA DE BLOQUES DE UN PREDICTOR LINEAL ADAPTATIVO



$$\hat{x}(n+1) = \sum_{l=0}^{p-1} w(l)x(n-l)$$

$$w(n+1) = w(n) + \frac{\mu e(n-1)\, x(n-1)}{\|x(n-1)\|^2}$$

En SMARTxAC se puede definir un umbral mínimo y uno máximo para el tráfico

Los intervalos de captura adoptados son de 10 segundos, que aportan una buena resolución a los resultados sin comportar una carga computacional significativa al sistema. Por otro lado, $w(n)$ se compone de 6 coeficientes, de manera que cada valor predicho se basa en el tráfico capturado en el minuto anterior.

La detección de anomalías en SMARTxAC se realiza sobre tres unidades de medida del tráfico: bits, paquetes y flujos por segundo

2.2.- Detección de anomalías basada en umbrales

El sistema de detección de anomalías de SMARTxAC también incorpora técnicas de detección basadas en umbrales de tráfico, para mitigar las limitaciones de la predicción adaptativa en presencia de anomalías que implican cambios progresivos y continuos en el tráfico.

En SMARTxAC se puede definir un umbral mínimo y uno máximo para el tráfico. Si durante un determinado intervalo de tiempo el tráfico se encuentra por debajo del umbral mínimo, o por encima del máximo, se dispara una alarma. Así se pueden detectar tanto las situaciones en las que el tráfico aumenta excesivamente, como aquellas en las que es inusitadamente bajo, debido por ejemplo a la caída de un enlace, un error de configuración, etc.

2.3.- Clasificación de las anomalías de red

La detección de anomalías en SMARTxAC se realiza sobre tres unidades de medida del tráfico: bits, paquetes y flujos por segundo. El proceso de clasificación de anomalías está basado en el descrito en [3], donde se propone justamente una clasificación por tipos de anomalías en función de en cuál o cuáles de estas tres métricas se han detectado variaciones inesperadas en el perfil de tráfico. Los tres tipos de métrica son importantes para la detección de anomalías, ya que conjuntamente permiten visualizar algunas situaciones irregulares que por separado pueden no resultar evidentes.

En [3] las anomalías se clasifican en ocho categorías, que hemos adoptado en nuestro sistema: (1) *Alpha*: Tasa de transferencia inesperadamente alta entre dos puntos de la red; (2) *Ataques DoS y DDoS*: Ataques de denegación de servicio (distribuidos o no) contra una sola máquina; (3) *Flash crowd:* Aumento repentino del número de usuarios que acceden a un mismo servicio o a un recurso concreto simultáneamente; (4) *Escaneo*: Escaneo de un equipo o una red en busca de un puerto vulnerable; (5) *Gusano*: Código que se autopropaga a través de una red explotando la falta de seguridad; (6) *Punto a multipunto*: Distribución de contenidos desde un solo servidor hacia un gran número de destinos; (7) *Caída*: Disminución repentina en el tráfico intercambiado entre una pareja de equipos, y (8) *Ingress shift*: Cambio del tráfico de un usuario de un punto de entrada a la red a otro.

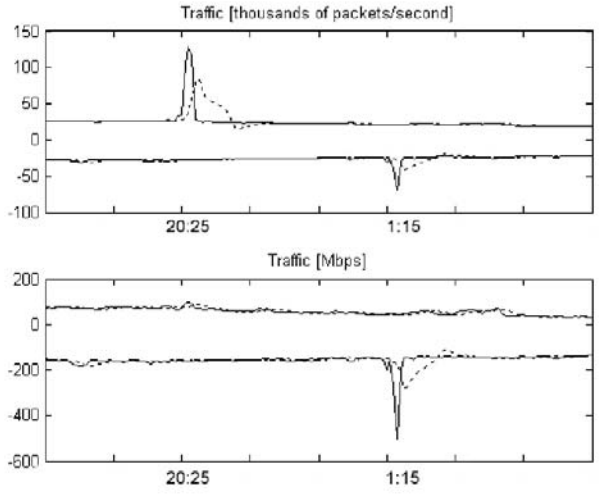
Además, cuando se detecta una anomalía, el sistema almacena todos los flujos relacionados con ésta durante un periodo de 5 minutos. Posteriormente, el sistema puede mostrar esta información o analizarla con más detalle para proporcionar, entre otros datos, el volumen de bytes, paquetes y flujos para cada puerto, protocolo y dirección IP.

Cuando se detecta una anomalía, el sistema almacena todos los flujos relacionados con ésta durante un periodo de 5 minutos

3.- Resultados

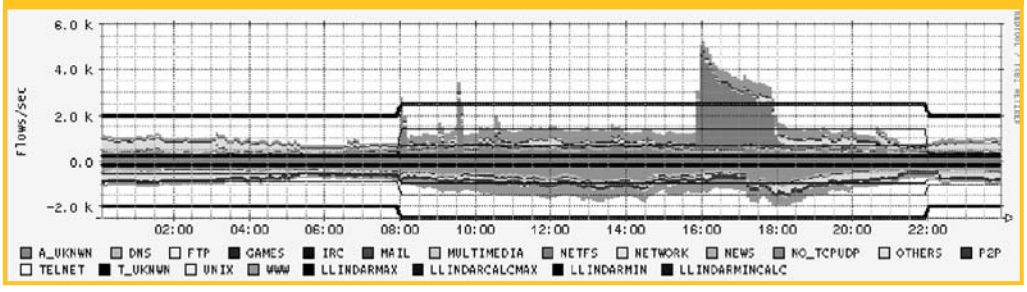
La Figura 2 muestra un ejemplo real del tráfico de una institución particular de la *Anella Científica*, en unidades de paquetes/segundo y bits/segundo respectivamente. Las líneas continuas representan el tráfico de entrada y salida, mientras que las líneas discontinuas muestran el tráfico predicho. Como puede observarse, una anomalía fue detectada a las 20:25 cuando el número de paquetes de entrada se multiplicó por 5 inesperadamente, aunque esto no afectara significativamente al número de bytes. Posteriormente, a la 1:15 se detectó otra anomalía debida no sólo al número de paquetes de salida, sino también al de bytes.

FIGURA 2: TRÁFICO REAL VS. PREDICHO PARA UNA INSTITUCIÓN DETERMINADA



La Figura 3 muestra el número de flujos/segundo por aplicación de una institución particular de la *Anella Científica*. Tres anomalías fueron detectadas a las 8:06, 9:36 y 15:50 respectivamente, cuando el umbral superior del tráfico de entrada fue superado. Por ejemplo, después de revisar la información almacenada por el sistema, se concluyó que la última fue debida a un *escaneo* de puertos TCP.

FIGURA 3: DETECTOR DE ANOMALÍAS BASADO EN UMBRALES DE TRÁFICO (FLUJOS/SEGUNDO)



4.- Conclusiones y trabajos futuros

Para combatir la proliferación de situaciones irregulares en Internet se requieren nuevas herramientas que permitan automatizar su detección. Con este objetivo se ha implementado un módulo de detección automática de anomalías para el sistema SMARTxAC, basado en técnicas de predicción adaptativa de tráfico. Además, el módulo se complementa con la definición de umbrales de tráfico, que permiten detectar las anomalías que presentan una progresión lenta. Los primeros resultados muestran una buena capacidad de detección, sin necesidad de definir explícitamente el perfil de tráfico habitual de la red.

Actualmente, el módulo está funcionando de manera independiente a modo de prueba, y en breve se realizará su integración final en el sistema SMARTxAC. Asimismo, otra de las ampliaciones del sistema será la clasificación automática del tipo de anomalía a partir del tipo de alarmas generadas por el módulo de detección.

Para combatir la proliferación de situaciones irregulares en Internet se requieren nuevas herramientas que permitan automatizar su detección

Este trabajo ha sido financiado parcialmente por el CESCA (convenio SMARTxAC) y por el Ministerio de Ciencia y Tecnología dentro del proyecto TIC2002-04531-C04-02.

Referencias

[1] Barlet-Ros, P.; Sole-Pareta, J; Domingo-Pascual, J. SMARTxAC: "Sistema de monitorización y análisis de tráfico para la Anella Científica". "Boletín de RedIRIS". (66-67): 27-30, 2004 (http://www.rediris.es/rediris/boletin/66-67/ponencia6.pdf.

[2] Adas, A. "Supporting Real Time VBR Video Using Dynamic Reservation Based on Linear Prediction". Presentada en: IEEE INFOCOM. 1996.

[3] Crovella, M.; Diot, C.; Lakhina, A. "Characterization of Network-Wide Anomalies in Traffic Flows". Presentada en: Internet Measurement Conference. 2004.

Actualmente, el módulo está funcionando de manera independiente a modo de prueba, y en breve se realizará su integración final en el sistema SMARTxAC

Pere Barlet-Ros, Helena Pujol
(pbarlet@ac.upc.edu), (hpujol@ac.upc.edu)
Javier Barrantes, Josep Solé-Pareta
(jbarranp@ac.upc.edu), (pareta@ac.upc.edu)
Jordi Domingo-Pascual
(jordi.domingo@ac.upc.edu)
Centre de Comunicacions Avançades de Banda Ampla (CCABA)
Dept. Arquitectura de Computadors
UPC