2003년도 대한전기학회 하계학술대회 논문집 2003.7.21-23

침입 탐지를 위한 FCM 기반의 네트웍 트래픽 데이터 클러스터링

곽미라, 조동섭
이화여자대학교 과학기술대학원 컴퓨터학과

Clustering Network Traffic Data Based on FCM for Intrusion Detection

Mira Kwak, Dong-sub Cho
Dept. of Computer Science and Engineering Ewha Womans University

Abstract - 여러 종류의 트래픽을 포함하는 네트웍 트래픽 데이터에서 각 종의 트래픽을 분류할 수 있는 능력은 네트웍 침입 탐지를 가능하게 하는 기본이다. 본 연구에서는 서비스 거부 공격과 사전 조사 행위 트래픽을 다른 트래픽으로부터 구분해 낼 수 있는 특징을 파악하고, 그것이 효과적인지 퍼지 c-means 기법으로 사용하여 실험하였다.

1. 서 론

네트웍 침입의 탐지를 위해서는 대량의 네트웍 트래픽 데이터로부터 공격 트래픽을 찾아낼 수 있어야 한다. 이것이 가능하도록 하기 위해서는 트래픽을 유형별로 구분지을 수 있는 특징을 알고 그것을 사용한 분류를 할 수 있는 능력이 먼저 갖추어져야 한다.

본 논문의 범위는, 네트웍 트래픽으로부터 공격 트래픽을 발견이 가능하도록 하는 특징 파악이다. 본 연구에서 구분할 수 있는 공격의 유형은 응용프로그램 수준의 데이터 내용에 무관하게 네트웍 트래픽 동태 정보만을 바탕으로 탐지할 수 있는 서비스 거부 공격류(DoS: Denial of Service)와 공격을 위한 사전 조사 행위들(probing)로 제한된다.

기존의 네트웍 트래픽의 특징 파악 및 군집화에 관한 연구들이 제안하는 방법은 패킷 크기, 패킷 시간 간격, 패킷 방향성을 사용하는 것이며, 공격 트래픽에 관한 연구의 경우 연결을 맺는 양 호스트 간 포트를 포함한 고유 연결 수, 포트 수를 사용하는 방법을 제안하고 있다. 본 연구에서는 이보다 섬세한 수준에서 공격 유형들 사이의 네트웍 흐름 상 차이점을 발견하기 위해 추가적인 요소들을 찾고자 하였다. 공격이 일어나지 않은 네트웍 트래픽 데이터와 공격 흔적이 나타나는 네트웍 트래픽 데이터 각각에 대하여 전송된 IP 패킷들의 흐름과 상태를 분석한다. 분석된 결과로부터 차이점들을 찾아, 그것들을 바탕으로 어떠한 네트웍 흐름이 공격인지 아닌지를 판단할 수 있는 특징들로 삼는다. 이들 특징을 사용하여 공격 트래픽들을 유사한 것끼리 군집화할 수 있는지 실험을 수행하고 결과를 보인다.

2. 공격 트래픽의 특징 추출

공격 트래픽을 설명할 수 있는 특징을 파악하기 위해 다음과 같은 과정을 거쳤다. 우리가 이미 그 상세한 내용을 알고 있는, 공격을 포함하는 네트웍 트래픽 데이터를 공격 세션과 정상 세션으로 나누었다. 공격 세션들과 정상 세션들에 대해 캡처한 네트웍 트래픽 데이터로부터 구할 수 있는 모든 정보를 알아보기 쉬운 형태로 정리하였다. 정보 요소들 중 우리의 눈으로 보기에 공격 세션들과 정상 세션들 사이에서 차이점을 가지는 요소들을 추려 공격 트래픽의 특징을 나타낼 수 있는 요소로 삼았다.

* 이 논문은 2003년도 두뇌한국21사업에 의하여 지원되었음.

네트웍 트래픽 데이터를 공격 유형별로 분류할 수 있도록 하는 특징을 파악하기 위해, 공격을 포함하고 있으며 분석되고 포함된 공격 트래픽이 파악된 데이터 집합이 필요하다. 이를 위하여 본 연구에서는 MIT Lincoln 연구소에서 침입탐지 시스템 평가를 위해 제공한 데이터 집합들 중 하나를 사용하였다.

2.1 정상 트래픽과 공격 트래픽의 비교

공격 트래픽을 설명하는 특징을 찾기 위해, 먼저 모든 선택한 정보 요소들을 바탕으로 정상 트래픽과 공격 트래픽의 세션들을 비교하였다. 그 결과 정상 트래픽을 공격 트래픽에 대해 설명할 수 있었는데, 그 중 대표적인 인터넷 서비스들에 대한 내용은 다음과 같다. 실제 정보의 일부를 표 1에서 볼 수 있다.

- HTTP : 연속된 시간동안 클라이언트의 여러 포트로부터 오고 가는 패킷 수가 각각 10개 내외인 짧은 연결들이 생성된다. 패킷들 사이의 도착 간격은 경우에 따라 다르다. 패킷의 크기는 100 바이트를 중심으로 다양하다. 패킷의 방향은 서버로부터 계속 클라이언트로 전송되는 경우가 많고, 양 방향으로 오가는 경우도 많다.
- SMTP : 패킷 개수는 대개 15개 내외이며 패킷 사이의 도착 간격은 다른 서비스의 경우와 비슷하다. 패킷 클라이언트가 서버로 보내는 패킷의 크기는 비교적 크며, 서버가 클라이언트로 보내는 패킷의 크기는 그보다 작다. 패킷의 방향은 연속하여 일정방향으로 전달되는 경우보다 교대로 오고 가는 경우가 많다.
- FTP : 명령을 전달하는 트래픽은 연결이 지속되며, 패킷은 교대로 오고가는 경우가 아주 많다. 서버에서 클라이언트로 패킷이 연속으로 전달되는 경우도 많은 편이다. 서버가 클라이언트로 보내는 패킷의 크기가, 보통 텔넷의 패킷 크기와 비슷한 반대방향으로 전달되는 패킷의 크기보다 보통 크다. 데이터를 전달하는 트래픽은 연속된 시간동안 계속 클라이언트의 여러 포트로부터 새로운 연결이 생성된다. 한 연결에서 오고 가는 패킷 수는 HTTP의 경우와 같이 적다. 클라이언트로부터 서버로 전달되는 패킷 수가 보통 더 적다. 패킷들 사이의 도착 간격은 매우 짧다. 클라이언트로부터 서버로 전달되는 패킷들의 사이즈는 40바이트 내외이며 서버로부터 클라이언트로 전달되는 패킷들의 사이즈는 대부분 매우 크고 다양하다. 패킷은 서버로부터 연속적으로 클라이언트에 전달되는 경우도 많고, 양 방향으로 교대로 오고 가는 경우도 그보다 적지만 많다.
- telnet : 한번의 연결이 매우 오랜 시간 지속되며 오고 가는 패킷 수도 타 서비스에 비해 현저히 많다. 클라이언트로부터 서버에 전달되는 패킷수가 반대 방향의 1.5배 정도이다. 클라이언트로부터 서버로 전달되는 패킷들의 평균 도착 간격은 상당히 짧고 반대 방향 패킷들의 평균 도착 간격은 긴

표 1 정상 트래픽 정보

a->b	A pkt num	A avg pkt interval	A avg pkt size	B pkt num	B avg pkt interval	B avg pkt size	AA	AB	BA	BB	service
172.16.114.148:20 -> 135.13.216.191:11759	5	0.0032	192.4000	3	0.0111	41.3333	3	2	2	1	ftp (data)
172.16.114.148:20->135.13.216.191:11760	5	0.0029	89.8000	3	0.0103	41.3333	3	2	2	1	
172.16.114.148:20->135.13.216.191:11761	5	0.0131	117.4000	3	0.0508	41.3333	3	2	2	1	
172.16.114.148:20->135.13.216.191:11762	5	0.0025	79.2000	3	0.0086	41.3333	3	2	2	1	
172.16.114.148:20->135.13.216.191:11763	5	0.0038	169.0000	3	0.0137	41.3333	3	2	2	1	
172.16.114.148:20->135.13.216.191:11816	6	0.0067	335.0000	3	0.0313	41.3333	4	2	2	1	
172.16.114.148:20->135.13.216.191:11878	5	0.0012	115.4000	3	0.0031	41.3333	3	2	2	1	
...											
135.13.216.191:11752->172.16.114.148:21	126	0.2208	51.8889	110	13.8115	77.7818	21	105	105	5	ftp (cmd)
135.13.216.191:1559->172.16.114.148:21	100	0.2930	50.4300	85	14.4311	76.2941	19	81	81	4	
135.13.216.191:4106->172.16.114.148:21	40	0.5969	47.2500	30	7.4984	70.2333	14	26	26	4	
135.8.60.182:1104->172.16.114.169:25	15	0.0756	260.6670	14	1.0356	69.1429	3	13	12	1	smtp
135.8.60.182:2111->172.16.114.169:25	13	0.0096	137.8460	12	0.0894	68.0833	3	11	10	1	
135.8.60.182:22632->172.16.114.169:25	12	0.0858	115.0000	12	0.8454	68.0833	3	10	9	2	
135.8.60.182:25270->172.16.114.169:25	11	0.0112	140.0910	12	0.0889	68.0833	2	10	9	2	
...											
172.16.112.100:1047->172.16.114.50:80	11	645.6360	119.1820	9	1616.4000	563.1110	4	8	7	1	http
172.16.112.100:1048->172.16.114.50:80	11	13.1236	95.4545	12	13.4339	1028.4200	4	8	7	4	
172.16.112.100:1049->172.16.114.50:80	9	16.4104	120.0000	10	30.9873	574.2000	4	6	5	4	
172.16.112.100:1050->172.16.114.50:80	8	358.1240	79.8750	8	360.4040	667.7500	3	6	5	2	
172.16.112.100:1055->172.16.114.50:80	8	1950.0800	112.7500	8	1952.2300	737.5000	3	6	5	2	
172.16.112.100:1056->172.16.114.50:80	6	3268.8300	154.6670	5	4089.8600	63.2000	3	4	3	1	
...											
135.13.216.191:1489->172.16.112.50:23	1394	0.6516	40.5337	745	405.0780	74.1973	651	743	743	2	telnet
135.13.216.191:23556->172.16.112.50:23	321	0.8952	40.6947	179	109.1170	63.6592	143	178	178	1	
172.16.114.207:1107->172.16.113.50:23	146	0.0821	41.0753	93	6.1010	47.0753	56	90	90	3	
172.16.114.207:1221->172.16.112.50:23	469	0.6196	40.6269	259	118.6150	58.6602	212	257	257	2	

편이다. 클라이언트로부터 서버로 전달되는 패킷들의 평균 크기는 40바이트 남짓이며 반대방향으로 전달되는 패킷들의 평균 크기는 그보다 약간 크다. 패킷들은 양 방향으로 교대로 전달되는 경우가 가장 많고, 클라이언트로부터 서버로 연속적으로 전달되는 경우도 자주 있다.

이와 비교하여 공격 트래픽의 성격을 다음과 같이 이야기할 수 있다. 추출한 공격 트래픽의 실제 정보의 일부를 표 2에서 볼 수 있다.

- probe : 대개의 경우 공격 대상의 여러 포트에 한 개 정도의 패킷을 짧은 시간 간격으로 두고 보낸다. 보내는 패킷의 사이즈는 40바이트로 고정되어 있다. 공격 대상은 공격자에게 패킷을 보내지 않을 수도 있다.
- chrashiis(DoS) : 공격 대상과 공격 자 사이에 40 바이트 이상의 크기를 가지는 패킷들이 대량 전송된다.
- mailbomb(DoS) : 공격 대상의 25번 포트에 공격자의 매우 많은 포트로부터 여러 패킷들이 대개 약 0.15초미만의 시간 간격으로 발신되며, 발신되는 패킷들의 크기는 약 53바이트로 일정하다. 공격 대상의 25번 포트에서 공격자의 각 포트로 대개 약 70바이트 미만의 사이즈의 패킷들이 약 0.2초 정도의 시간 간격으로 발신된다. 패킷들은 대개 양 방향으로 교대로 전송된다.
- processtable(DoS) : 공격 대상의 22번 포트에

표 2 공격 트래픽 정보

a->b	A pkt num	A avg pkt interval	A avg pkt size	B pkt num	B avg pkt interval	B avg pkt size	AA	AB	BA	BB	attack	category
153.107.252.61:-->172.16.112.100:-	3	1.4936	106.0000	0	0.0000	0.0000	3	0	0	0	portsweep	probe
153.107.252.61:33457->172.16.112.100:22	1	0.0000	40.0000	1	0.0000	40.0000	1	1	0	0		
153.107.252.61:33490->172.16.112.100:25	2	0.0030	40.0000	1	0.0000	44.0000	1	1	1	0		
153.107.252.61:38014->172.16.112.100:21	2	0.0029	40.0000	1	0.0000	44.0000	1	1	1	0		
153.107.252.61:62583->172.16.112.100:80	2	0.0011	40.0000	1	0.0000	44.0000	1	1	1	0		
172.16.118.20:62519->172.16.113.50:25	2	0.0026	40.0000	1	0.0000	44.0000	1	1	1	0	portsweep	probe
172.16.118.50:36580->192.168.1.1:25	1	0.0000	40.0000	1	0.0000	40.0000	1	1	0	0	portsweep	probe
172.16.118.50:44450->192.168.1.1:513	1	0.0000	40.0000	1	0.0000	40.0000	1	1	0	0		
172.16.118.50:46444->192.168.1.1:22	1	0.0000	40.0000	1	0.0000	40.0000	1	1	0	0		
172.16.118.50:46819->192.168.1.1:21	1	0.0000	40.0000	1	0.0000	40.0000	1	1	0	0		
172.16.118.50:51996->192.168.1.1:80	1	0.0000	40.0000	1	0.0000	40.0000	1	1	0	0		
172.16.118.70:1109->172.16.112.100:80	4	1.6085	43.5000	3	2.4796	41.3333	2	3	2	0	crashiis	DoS
202.77.162.213:41523->172.16.114.169:143	1	0.0000	40.0000	0	0.0000	0.0000	1	0	0	0	portsweep	probe
202.77.162.213:41524->172.16.114.169:143	1	0.0000	40.0000	0	0.0000	0.0000	1	0	0	0		
202.77.162.213:51344->172.16.114.169:21	1	0.0000	40.0000	0	0.0000	0.0000	1	0	0	0		
202.77.162.213:51345->172.16.114.169:21	1	0.0000	40.0000	0	0.0000	0.0000	1	0	0	0		
202.77.162.213:53859->172.16.114.169:23	1	0.0000	40.0000	0	0.0000	0.0000	1	0	0	0		
202.77.162.213:53860->172.16.114.169:23	1	0.0000	40.0000	0	0.0000	0.0000	1	0	0	0		
202.77.162.213:55661->172.16.114.169:79	1	0.0000	40.0000	0	0.0000	0.0000	1	0	0	0		
202.77.162.213:55662->172.16.114.169:79	1	0.0000	40.0000	0	0.0000	0.0000	1	0	0	0		
202.77.162.213:58245->172.16.114.169:25	1	0.0000	40.0000	0	0.0000	0.0000	1	0	0	0		
202.77.162.213:58246->172.16.114.169:25	1	0.0000	40.0000	0	0.0000	0.0000	1	0	0	0		
6.238.105.108:-->172.16.112.50:-	266	0.0579	329.0000	0	0.0000	0.0000	266	0	0	0	smurf	DoS

공격자의 매우 많은 포트로부터 7개의 패킷들이 약 42초의 시간 간격으로 발신된다. 이 때 패킷의 사이즈는 각 연결마다 다르다. 공격 대상에서 공격자로 향하는 패킷들의 수는 연결마다 6개씩이며 그 사이즈는 반대 방향으로 전달되는 패킷의 세 배 이상이며 패킷 도달 간격은 약 89초로 일정하다. 패킷들의 전달은 주로 양방향 교대로 이루어진다.

2.2 네트워 트래픽 특성 요소 선택

표 1과 표 2에서 보듯이 하나의 세션에 대한 속성의 종류는 많다. 이 중 유형별 공격 트래픽 무리의 특징을 나타낼 수 있을 특징을 찾는 것은, 후에 그것을 사용하여 새로운 네트워 트래픽에서 공격을 발견하는 작업의 정확성에 직접 영향을 미치는 중요한 작업이다. 본 연구에서는 네트워 상의 데이터에서 내용에 무관하게 트래픽의 행태만을 바탕으로 공격 트래픽을 수행하고자 하며, 위에서 설명한 바와 같이 파악한 정상 트래픽과 공격 트래픽의 연결과 패킷 전달에 있어서 차이점을 바탕으로 네트워 트래픽의 행태를 설명할 수 있을 것으로 예상되는 요소들을 다음과 같이 선택하였다.

두 호스트 중 한 쪽을 호스트 a, 다른 한 쪽을 호스트 b라 하고 호스트 a에서 호스트 b 방향을 방향 A, 반대를 방향 B라고 하자.

표 3 네트워 트래픽의 특성 요소

양 호스트 사이의 모든 연결에 대해
• 호스트 a가 사용한 포트 수 • 호스트 b가 사용한 포트 수
양 호스트 사이에 맺어진 각 연결에 대해
• 총 전달된 패킷 수 • 총 전달된 패킷들 중 방향 A로 전달된 패킷들의 비율 • 방향 A로 전달된 패킷들의 평균 크기 • 방향 A로 전달된 패킷들의 평균 시간 간격 • 총 전달된 패킷들 중 방향 B로 전달된 패킷들의 비율 • 방향 B로 전달된 패킷들의 평균 크기 • 방향 B로 전달된 패킷들의 평균 시간 간격 • 방향 A로 전달된 패킷 후 다시 같은 방향의 패킷이 전달된 비율 • 방향 A로 전달된 패킷 후 방향 B로 패킷이 전달된 비율 • 방향 B로 전달된 패킷 후 방향 A로 패킷이 전달된 비율 • 방향 B로 전달된 패킷 후 다시 같은 방향의 패킷이 전달된 비율

3. FCM을 사용한 공격 트래픽 분류

3.2절에서 설명한 바와 같이 표 3의 네트워 트래픽 특성 요소들을 선택하였다. 이를 바탕으로 트래픽을 클러스터링 하는 경우 같은 유형의 공격끼리 군집화되는지 살펴보았다. 이를 위하여 본 연구에서는 퍼지 c-means 클러스터링(FCM) 기법을 사용하였다.

3.1 퍼지 c-means 클러스터링

퍼지 c-means법은 데이터 벡터 x_k와 클러스터 i의 중심 v_i와의 거리와 x_k의 클러스터 i로 멤버쉽 값 u_{ik}에 관한 다음 식을 최소화 하는 퍼지 분할을 행하는 것이다.

$$J=\sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^q d_{ik},\ d_{ik}=\|x_k-v_i\|^2$$

단, n은 데이터 벡터 수, c는 클러스터 수, $q(>1)$는 스무징 · 파라미터이다. 멤버쉽 값은 아래의 제약조건을 만족시킨다.

$$0\le u_{ik}\le 1,\ i=1,2,\cdots,c,\ k=1,2,\cdots,n$$

$$\sum_{k=1}^{n}u_{ik}>0,\ i=1,2,\cdots,c$$

$$\sum_{i=1}^{c}u_{ik}=1,\ k=1,2,\cdots,n$$

3.2 공격 트래픽 분류 실험

공격 트래픽 데이터에 대해 두 과정을 거쳐 분류를 수행한다. 첫 번째 분류 과정에서는 연결의 포트를 무시하고 연속적인 호스트 대 호스트의 모든 연결을 하나의 데이터로 삼는다. 이때 데이터를 표현하는 벡터는 호스트 a가 사용하는 포트 수와 호스트 b가 사용하는 포트 수, 총 전달된 패킷 수로 구성된다. 두 번째 분류 과정에서는 양 호스트 사이에 맺어진 각 연결, 즉 포트의 구분을 고려한 연결들 각각을 하나의 데이터로 삼는다. 하나의 데이터를 표현하는 벡터를 구성하는 요소들은 표 3의 가장 아래 컬럼에 나열된, 양 호스트 사이에 맺어진 각 연결에 대한 특성 요소들이다. 두 번째 분류 과정에서는 분류 결과 첫 번째 분류 과정에서 데이터로 사용된 하나의 입력이 여럿으로 나뉘지 않도록 클러스터링 되도록 한다. 두 분류의 결과를 모두 고려하여 최종적으로 공격 트래픽들의 군집화를 마친다. 본 연구에서 사용한 데이터에 대해서는 DoS의 mailbomb, crashiis, smurf, probe가 각각 따로 군집을 이룰 수 있었다.

4. 결 론

본 연구에서는 네트워 트래픽에서 공격 트래픽을 특성화하는 요소들을 발견하였고, 이것이 공격 트래픽의 군집에 사용가능한지 검증하였다. 이 논문에서 선택한 네트워 트래픽 특성 요소들 사이의 연관관계를 파악하고 효과적으로 특성 요소를 간략화하여, 군집화과정의 효율성과 정확도를 높이는 연구가 앞으로 필요하다. 또한 현재까지의 결과를 바탕으로 새로운 네트워 트래픽에 대해 분류를 할 수 있는 규칙을 포함하는, 트래픽 분류와 공격 판별 시스템의 설계가 앞으로의 연구 방향이다.

[참 고 문 헌]

[1] R. Caceres, P. Danzig, S. Jamin, and D. Mitzel. Characteristics of Wide-Area TCP/IP Converstations. *ACM SIGCOMM*, September 1991.

[2] W. Cleveland and D. Sun. Internet Traffic Data. *Journal of the American Statistical Association*, pp. 979-985, September 2000.

[3] J. Doak. An Evaluation of Search Algorithms for Feature Selection. *UCD/LANL*, January 1994.

[4] T. Dunigan and G. Ostrouchov. Flow Characterization for Intrusion Detection. Technical report, Oak Ridge National Laboratory, Oak Ridge, July 2001.ORNL/TM-2001/115.

[5] S. M. Weiss and N. Indurkhya. Predictive Data Mining: a Practical Guide. Morgan Kaufmann Publishers, Inc. 1998.

[6] J.E. Dickerson, J. Juslin*, O. Koukousoula*, J.A. Dickerson, Fuzzy intrusion detection, *IFSA World Congress and 20th North American Fuzzy Information Processing Society (NAFIPS) International Conference*, Vancouver, British Columbia, Volume 3, 1506-1510, July, 2001.

[7] J.E. Dickerson, J.A. Dickerson, Fuzzy Network Profiling for Intrusion Detection. *Proceedings of NAFIPS 19th International Conference of the North American Fuzzy Information Processing Society*, Atlanta, July, 301-306, 2000.