Compromised 노드 특성에 따른 Sybil Attack의 성능 평가

김경백
전남대학교 전자컴퓨터공학부
e-mail : kyungbaekkim@jnu.ac.kr

Assessing the performance of Sybil Attack based on the characteristics of Compromised nodes

Kyungbaek Kim
Dept. Electronics and Computer Engineering,
Chonnam National University

요 약

인터넷의 발달의 주요 요인 중 하나로 익명성을 생각할 수 있다. 이러한 익명성은 분산시스템에서 다양한 사용자들이 자유롭게 자신의 의사를 표현하는데 많은 도움들 주어 협업 시스템 등 다양한 분산시스템의 활성화에 큰 도움이 되어 왔다. 하지만, 이러한 익명성은 시빌 어택(Sybil Attack)과 같은 위협을 제공하기도 한다. 시빌 어택이란 한명의 사용자가 다수의 가짜 사용자를 조정함으로써 시스템에서 불공정한 이득을 취하는 공격이다. 최근 OSN(Online Social Network)상의 그래프를 분석하여 시빌 어택에 사용되는 시빌 노드(Sybil Node)들을 알아내기 위한 연구들이 수행 되어 왔다. 이 중 시빌 어택의 성능에 가장 큰 영향을 미치는 요소 중 하나는 어택 에지(Attack Edge)의 개수라는 것이 밝혀졌다. 어택 에지란 시빌 노드를 일반 사용자에 해당하는 노드와 연결시키는 OSN 그래프상의 에지를 의미하고, 어택 에지에 연결된 일반 사용자 노드를 Compromised 노드라고 한다. 이 논문에서는 어택 에지를 연결하는 Compromised 노드의 특성에 따라 그 성능이 어떤 식으로 달라질 수 있는지를 실험을 통하여 살펴본다. Facebook에서 추출한 OSN 그래프를 기반으로, 서로 다른 특성을 가지는 Compromised 노드에 어택 에지가 생성 되었을 경우, 시빌 노드를 확인하는 기법중 하나인 RRTI (Random Route Tail Intersection)의 성능을 측정하였다. 이를 통해, OSN상에서 이웃노드가 51개에서 100개 이하인 노드에 어택 에지가 생성될 경우 보다 효과적인 시빌 어택을 수행할 수 있음을 확인하였다.

1. 서 론

오늘날 인터넷 상의 여러 사용자들에 의해서 운용되는 다양한 구성 요소들로 이루어진 분산 시스템은 빠르게 대중화 되어 가고 있다. 이를 촉진 시키는 요인 중 하나로는 익명성을 생각 할 수 있다. 즉 분산 시스템의 개방적 특성에 따라 임의의 사용자들이 시스템을 사용하며 정보를 제공하도록 유도함으로써 사용자간의 협업을 촉진 시킨다. 하지만, 개방적 특성은 제공되는 정보의 신뢰성을 판단하기 힘들게 만들기도 한다. 이러한 문제를 해결하기 위해 사용자들의 과거의 행동 방식에 기반을 둔 신뢰도 추출 방식을 사용하는 Reputation 시스템들이 연구 되었지만 [1][2], 이들 시스템들은 과거의 기록이 없는 새로운 사용자와 같은 경우에 신뢰도 측정이 힘들게 된다. 이와 같은 이유로 개방형 분산 시스템은 시빌 어택에 취약하게 된다 [3]. 시빌 어택은 한명의 사용자가 다수의 가상/가짜 사용자들을 생성하여 분산 시스템에서 불공정한 이득을 취하거나 시스템을 마비시키는 공격이다. 시빌 어택에서 사용되는 가상의 사용자들은 시빌 사용자라고 불리고, 이 시빌 사용자들은 비교적 손쉽게 가공/폐기가 가능하다. 이와 같은 이유로 시빌 어택을 일반적인 과거 행동 방식 기반의 Reputation 시스템에서 방어한다는 것은 쉽지 않다.

이러한 문제 해결을 위해 온라인 소셜 네트워크 (OSN) 그래프를 기반으로 하는 Social resistant value 추출 기법들이 제안되었다[5][6]. OSN 그래프는 실제 사용자들의 관계를 나타내는데, 가상의 시빌 사용자들은 이 실제 사용자들과 잘 연결되지 않는 특성을 가진다[5][6]. 이러한 특징을 사용하여, OSN그래프 상의 임의의 사용자가 실제 사용자일 가능성의 정도를 나타내는 시빌-저항 신뢰도 (Sybil-resistant trust value)를 추출하는 기법인 RRTI (Random Route Tail Intersection)이 연구 되었다[4]. 이 시빌-저항 신뢰도는 사용자가 시스템에 참여 되는 과정, 즉 시스템 사용자들 간의 관계에 의해서 그 값이 계산 된다. 따라서 시빌 노드가 새롭게 생성 되더라도 생성 후 일반 사용자들과의 관계가 없게 되면 높은 시빌-저항 신뢰도를 가질 수 없게 된다. 이렇게 계산된 시빌-저항 신뢰도는 인터넷 증명서 발급 시스템[7], 스팸 메일 필터링 시스템[8]등에 응용되어 사용 될 수 있다.

이러한 OSN기반의 시빌 노드 확인 기법들은 시빌 어택

† "본 연구는 미래창조과학부 및 정보통신산업진흥원의 대학IT연구센터육성 지원사업의 연구결과로 수행되었음" (NIPA-2013-H0301-13-3005)

에 사용되는 어택 에지 (Attack Edge)의 개수에 크게 영향을 받는 것이 확인되었다[9].즉 어택 에지의 개수가 늘어남에 따라 시빌 어택이 더욱 위력을 발휘하게 된다는 점이다. 어택 에지란 시빌 노드를 일반 사용자에 해당하는 노드와 연결시키는 OSN 그래프상의 에지를 의미하고, 어택 에지에 연결된 일반 사용자 노드를 Compromised 노드라고 한다. RRTI의 경우, 시빌 노드 1000개로 구성된 집단의 어택 에지의 개수가 300개 이상일 경우 더 이상 시빌 노드를 구분할 수 없게 된다. 따라서 시빌 어택을 방지하기 위해서는 Compromised 노드의 개수를 최소화 하는 노력이 필요하다. 반면 시빌 노드 입장에서는 이렇게 제한된 개수의 Compromised 노드만이 가능한 상황에서 어떤 노드를 Compromised 노드로 만들어야 시빌 어택의 성능을 최대화 할 수 있을지를 생각하게 된다.

이 논문에서는 제한된 어택 에지의 개수를 가정하였을 때 다양한 Compromised 노드의 특성이 시빌 어택에 미치는 성능에 대해서 실험적으로 분석하였다. 고려된 노드의 특성은 각 노드의 이웃 노드의 개수이다. 이는 최근 연구된 논문에서 각 노드가 할당 받는 시빌-저항 신뢰도 값이 이웃 노드의 개수와 관계가 있는 것으로 확인되었기 때문이다.[9] 실험은 Facebook에서 추출한 OSN 그래프 상에서 수행되었고, 시빌 노드들은 임의의 평균 이웃 노드의 개수를 가정한 클러스터를 구성한다고 가정한 후 어택 에지를 추가 하여 수행 되었다. 결론적으로, 단순히 이웃 노드 개수가 많은 노드가 Compromised 노드가 되는 것보다 이웃노드의 개수가 51에서 100개 이하인 노드들이 Compromised 노드가 되는 것이 보다 효과적인 시빌 어택을 가능하게 함을 확인 하였다.

이 논문의 구성은 다음과 같다. 2장에서는 OSN기반의 시빌-저항 신뢰도 추출을 위한 기법인 Random Route Tail Intersection기법에 대한 간략한 소개를 한다. 3장에서는 Facebook에서 추출된 그래프를 사용한 다양한 Compromised 노드에 따른 시빌 어택의 성능을 평가하고, 4장에서 이 논문의 결론을 내린다.

2. Random Route Tail Intersection 기법

성능평가에 사용된 OSN기반의 시빌-저항 신뢰도 추출 기법으로 SybilLimit[6]기반의 RRTI (Random Route Tail Intersection)을 고려할 수 있다.[4] 이 기법은 전체 OSN 그래프를 알고 있는 중앙 서버에서 수행된다. 사용되는 OSN 그래프는 실제 사용자들로 구성된 실제사용자 리젼(Honest Region)과 시빌 사용자들로 구성된 시빌 리젼(Sybil Region)으로 이루어지고 각각의 리젼은 각기 하나의 strongly connected component이다. 시빌 리젼과 실제 사용자 리젼은 주어진 개수의 어택 에지들로 연결되어 전체 OSN 그래프가 하나의 strongly connected component가 된다.

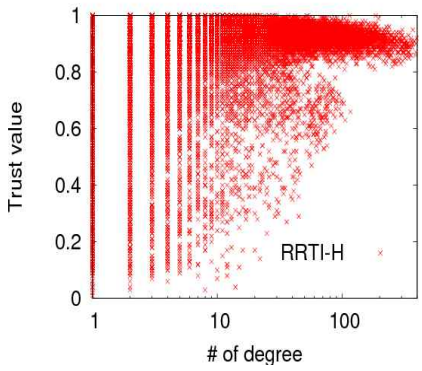
RRTI기법에서 시빌-저항 신뢰도를 추출하기 위해서 우선 몇몇 임의의 실제 사용자들에 해당하는 OSN 그래프 상의 노드들을 검증 노드(Verifier Node)들로 선택한다. 각 검증 노드들은 RRTI에서 정의된 판별 방법을 통해 OSN 그래프 상의 임의의 노드가 시빌 노드인지 실제 사용자 노드인지를 판별한다. 임의의 노드를 실제 사용자에 해당한다고 판별한 검증 노드를 accepted 검증 노드라 할 때, 시빌-저항 신뢰도는 (accepted 검증 노드의 개수)/(전체 검증 노드의 개수) 와 같이 구할 수 있다. 즉, 시빌-저항 신뢰도는 0에서 1까지의 값을 가지고 1에 가까운 값을 가질수록 해당 노드가 실제 사용자에 해당한다고 말 할 수 있다.

RRTI에서는 SybilGuard[6]에서 제안된 랜덤 라우트(random route)의 테일(tail)을 비교하는 판별방법을 사용한다. 즉, 랜덤 라우트의 마지막 에지를 비교한다. 각 노드는 r개의 테일을 구하여 저장하고, 검증 노드의의 테일의 집합과 임의의 노드의 테일의 집합 중에서 같은 테일이 있을 경우 검증 노드는 해당 노드가 실제 사용자에 해당한다고 판별한다. 이때, r개의 테일을 구하기 위해서는 r개의 서로 다른 라우팅 테이블(routing table)을 각 노드가 가지고 있어야 한다. 이때 사용되는 r은 OSN 그래프의 edge의 개수를 m이라 할 때 $\Theta(\sqrt{(m)})$ 가 되고, 각 random route의 길이 w는 OSN그래프의 노드의 개수가 n이라고 할때 $\Theta(\sqrt{(\log(n))})$ 가 되어야 한다

3. 다양한 Compromised 노드에 따른 시빌 어택의 성능 평가

다양한 Compromised 노드에 따른 시빌 어택의 성능을 평가 하기 위해서 Facebook에서 추출된 샘플 그래프를 사용하였다. Facebook에서 추출된 OSN 샘플 그래프는 노드의 개수가 100,000 이고 에지의 개수는 1,861,360이고 지름(Diameter)가 18인 하나의 Strongly Connected Component이다. 이 샘플 그래프를 실제사용자 리젼으로 가정하였다. 시빌 리젼은 각 노드의 평균 이웃 노드의 개수가 14인 100개의 노드로 이루어졌다고 가정하고, 각 시빌 리젼은 두 개의 어택 에지를 가지고 있다고 가정한다. 각 어택 에지에서 다양한 Compromised 노드들을 고려하기 위하여, 실제사용자 리젼의 노드중, 이웃노드가 5개 이하, 이웃노드가 6개 이상 10개 이하, 이웃노드가 11개 이상 50개 이하, 이웃노드가 51개 이상 100개 이하, 이웃노드가 101개 이상 150 이하인 노드들로 구분하였다. 따라서 총 5가지의 다른 특성을 가진 Compromised 노드들과 연결된 시빌 리젼을 고려하기 위해, 각 타입에 대해 5개씩, 총 25개의 시빌 리젼을 생성하였다. RRTI기법에서 사용되는 검증 노드들은 실제사용자 리젼에서 임의로 선택된다고 가정하고, 그 수는 1600개로 설정하였다. 또한 RRTI기법에서 사용하는 r은 2000으로 설정하였고, w는 20, 40, 80, 160으로 설정하였다.

RRTI를 통해 측정된 실제사용자 리젼의 노드들의 시빌-저항 신뢰도 분포는 그림 1과 같다. 이 그림에서, RRTI를 사용할 경우, 이웃 노드의 개수가 10보다 작을 경우 아



(그림1) RRTI : 이웃 노드 개수에 따른 시빌-저항 신뢰도 분포

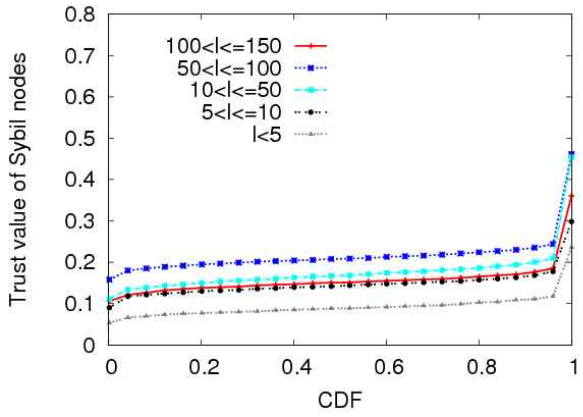
주 낮은 시빌-저항 신뢰도를 가질 수 있음이 확인 되었고, 이웃 노드의 개수가 100 보다 작을 경우 중간정도(0.3~0.8)의 시빌-저항 신뢰도를 가지는 것을 볼수 있다.

이와 같은 결과에 따라, 우리는 Compromised 노드가 이웃노드의 개수가 많을수록 시빌 어택의 성능이 향상 될 것으로 예측하였다. 이를 확인하기 위해 그림 2, 그림 3, 그림 4에서는 서로 다른 w하의 RRTI에서 각 시빌 노드들에 할당한 시빌-저항 신뢰도에 대한 CDF(Cumulative Density Function)를 보여준다. 이 그림들에서, 이웃노드의 개수가 10개 이하로 너무 작을 경우에는 높은 시빌-저항 신뢰도를 얻지 못한다는 것을 확인 할 수 있었다. 또한, 예측했던 것과는 반대로 이웃노드의 개수가 51개 이상, 100개 미만인 Compromised 노드들과 연결된 시빌 리전의 시빌 노드들이 평균적으로 가장 높은 시빌-저항 신뢰도를 가지는 것으로 확인 되었다. 즉 이웃노드가 아주 많은 Compromised 노드들에 어택 에지가 생성되는 경우, 시빌 어택의 효과가 그리 높지 않음을 확인 할 수 있었다.

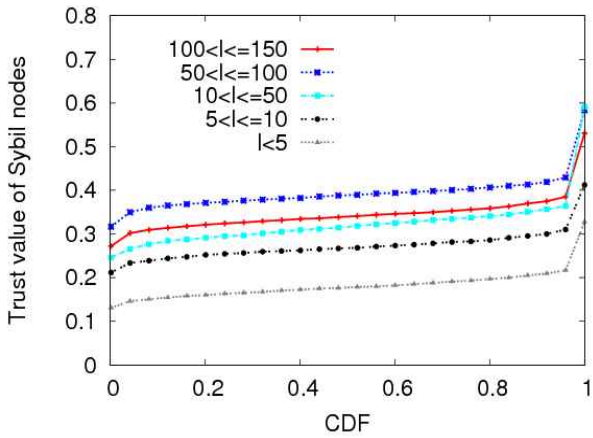
이러한 결과는 w의 변화에 의해서 더욱 뚜렷하게 나타난다. 일반적으로 w의 크기가 커질수록 RRTI에서 할당하는 시빌-저항 신뢰도의 평균적인 값이 증가한다. 이러한 영향은 시빌 노드에게도 직접적인 영향을 미치는데, 그림 4와 같이 w가 아주 클 경우, Compromised노드의 특성에 따라 시빌 노드들이 얻을 수 있는 시빌-저항 신뢰도에 많은 차이가 나는 것을 확인 할 수 있었다.

4. 결론

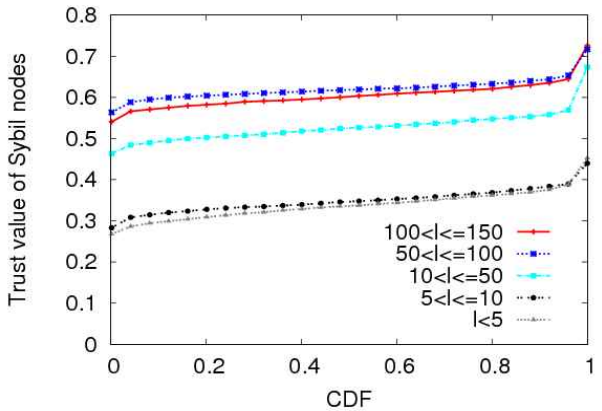
시빌 어택에 대한 방어 기법은 분산시스템을 운용하는데 있어서 정보의 신뢰도를 유지하는데 있어서 필수적이다. 일반적으로 시빌 어택에서 사용되는 어택 에지의 개수를 줄일 수 있도록 사용자들에게 홍보함으로써 분산시스템에서 신뢰도를 보장 할 수 있다. 이러한 상황에서 시빌 리전은 제한된 Compromised 노드들을 집중적으로 공략하여 어택 에지를 생성해야 한다. 이 논문에서는 Facebook에서 추출한 OSN그래프를 사용한 실험을 수행하여, 보다 효과적인 시빌 어택을 위해서는 이웃노드의 개수가 51개



(그림2) CDF of Sybil-Resistant Trust value of Sybil nodes. w=40.



(그림3) CDF of Sybil-Resistant Trust value of Sybil nodes. w=80.



(그림4) CDF of Sybil-Resistant Trust value of Sybil nodes. w=160.

이상 100개 이하인 노드를 Compromised 노드로 만들어야 한다는 점을 확인 할 수 있었다. 이러한 점을 통해, 분산시스템에서 시빌 어택의 가능성을 줄이고 그 영향력을 최소화하기 위해서는 51개 이상 100개미만의 이웃노드를 가지는 사용자들에게 집중적으로 시빌 어택에 대해 홍보를 하는 것이 필요함을 알 수 있었다.

참 고 문 헌

[1] S. D. Kamvar, M. T. Schlosser and H. Garcia-molina. The EigenTrust Algorithm for Reputation Management in P2P Networks. In Proceedings of 12th WWW, 2003

[2] A.G.P. Rahbar and O. Yang. PowerTrust: A Robust and Scalable Reputation System for Trusted Peer-to-Peer Computing. In IEEE TOPDS, Vol.18, Issue.4, April 2007

[3] J. R. Douceur. The sybil attack. In Proceedings of IPTPS 2002, pages 251-260, 2002.

[4] Kyungbaek Kim. Sybil-Resistant Trust Value of Social Network Graph. In Proceedings of the First International Conference on Smart Media and Applications (SMA 2012), August 21-24, 2012, Kunming, Yunnan, China

[5] H. Yu, M. Kaminsky, P. B. Gibbons and A. Flaxman. SybilGuard: Defending Against Sybil Attacks via Social Networks. In Proceedings of ACM SIGCOMM, August 2006

[6] H. Yu, P. B. Gibbons, M. Kaminsky and F. Xiao. SybilLimit:A Near-Optimal Social Network Defense against Sybil Attacks. In Proceedings of IEEE Symposium on Security and Provacy 2008, pp. 3-17, 2008

[7] Michael Sirivianos, Kyungbaek Kim, Jian Wei Gan and Xiaowei Yang. Assessing the Veracity of Identity Assertions via OSNs. In Proceedings of COMSNETS 2012, January 3-7, 2012, Bangalore, India

[8] Michael Sirivianos, Kyungbaek Kim and Xiaowei Yang. SocialFilter: Introducing Social Trust to Collaborative Spam Mitigation. In Proceedings of IEEE INFOCOM 2011, April 10-15, 2011, Shanghai, China

[9] 김경백, OSN기반 Sybil-Resistant Trust Value 추출 기법들에 대한 성능평가, 2013년 정보처리학회 춘계학술대회, 5월, 2013, 부산