OSN 기반 Sybil-resistant trust value 추출 기법들에 대한 성능평가

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

Assessing the performance of extraction methods for OSN-based Sybil-resistant trust values

Kyungbaek Kim*
*Dept. of Electronics and Computer Engineering,
Han-Kook University

요 약

인터넷상에서 다양한 사용자 및 구성요소로 이루어진 분산시스템은 Sybil Attack 에 취약하다. 최근 온라인 소셜 네트워크(Online Social Network)의 그래프 정보를 사용해, Sybil Attack 에 대응하기 위한 Sybil-resistant value 추출 기법들이 제안되었다. 이 논문에서는 이러한 OSN 기반의 Sybil-resistant value 추출 기법들에 대한 성능을 평가한다. 특히 OSN 그래프의 각 노드들의 이웃 노드 개수 정보에 따른 성능과 Sybil 노드들의 Attack Edge 에 따른 성능을 평가한다. Facebook 에서 추출한 샘플 OSN 그래프를 사용한 성능 평가 분석을 통해, 실제 사용자를 위한 Sybil-resistant value 를 정상적으로 추출하기 위해서는 OSN 그래프 상에서 이웃 노드의 개수가 10 개 이상이어야 한다는 점과, Random Route Tail Intersection 기법이 Sybil 사용자 그룹의 Attack Edge 의 영향을 가장 적게 받는 다는 점을 확인 하였다.

1. 서론

오늘날 인터넷 상의 분산 시스템은 빠르게 대중화되어 가고 있다. 인터넷 상의 분산 시스템은 다양한 특성을 가진 구성요소로 이루어져 있을 뿐 만 아니라 다양한 사용자들에 의해서 시스템이 운용된다. 이러한 분산 시스템의 개방적 특성은 임의의 사용자들이 시스템에 정보를 제공하도록 유도한다. 하지만, 개방적 특성은 제공되는 정보의 신뢰성을 판단하기 힘들게 한다. 이러한 문제를 해결하기 위해 Reputation 시스템들이 연구되었다.[1][2] 이 시스템들은 주로 사용자들의 과거의 행동 방식에 기반을 둔 신뢰도 추출방식을 사용한다. 하지만, 이 방식은 과거의 기록이 없는 새로운 사용자와 같은 경우에는 적용하기 힘들다. 이와 같은 이유로, 개방형 분산시스템은 Sybil attack 에 취약하게 된다.[3] Sybil attack 이란 한명의 사용자가 다수의 가상 사용자들을 생성하여 시스템으로 부터 불공정한 이득을 취하거나 시스템을 마비 및 전복시키기 위한 공격을 하는 것이다. 이와 같은 Sybil attack 에서 사용되는 가상의 사용자들을 Sybil 사용자라고 한다. 이와 같은 Sybil 사용자들이 Reputation 시스템에 의해서 시스템 사용에 제재를 받게 되면, 제재된 Sybil 사용자를 버리고 새로운 Sybil 사용자를 생성함으로써 Reputation 시스템의 제재를 피할 수 있다.

이러한 문제를 해결하기 위한 방법으로 온라인 소셜 네트워크(OSN) 그래프를 기반으로 하는 Social-resistant value 추출 기법들이 제안되었다.[4] OSN 그래프는 실제 사용자들의 관계를 나타내게 되고, Sybil 사용자들은 이들을 생성한 사용자를 제외한 다른 실제 사용자들과 OSN 그래프 상에서 잘 연결되지 않는다.[5][6] 이와 같은 OSN 그래프의 특징을 기반으로 임의의 사용자가 실제사용자일 가능성의 정도를 나타내는 Sybil-resistant trust value 를 추출하는 기법들이 연구 되었다. 이 Sybil-resistant trust value 는 사용자의 과거의 행동을 통해 계산되는 것이 아니라, 각 사용자가 시스템에 참여하는 과정, 즉 사용자들 간의 관계를 통해 계산된다. 따라서 Sybil 사용자들이 새롭게 생성되더라도 OSN 기반의 Sybil-resistant trust value 추출 기법을 통해 분산시스템은 효과적으로 이 Sybil 사용자들을 구별할 수 있다.[7][8]

이 논문에서는 OSN 그래프의 특성이 OSN 기반의 Sybil -resistant trust value 추출 기법들의 성능에 미치는 영향을 평가한다. 특히 OSN 그래프의 특성 중 이웃 노드의 개수에 따른 Sybil-resistant trust value 의 분포와 Sybil 사용자들의 Attack Edge 의 변화에 따른 추출 기법들의 변화를 살펴본다. Attack Edge 란 Sybil 사용자

"본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT 연구센터육성 지원사업의 연구결과로 수행되었음" (NIPA-2013-H0301-13-3005)

와 실제 사용자간의 OSN 그래프에서의 연결을 말한다. 성능 평가를 위해, Facebook 에서 수집된 샘플 그래프와 임의로 생성된 Sybil 사용자 그룹 그래프가 사용 되었다. 평가 결과 분석을 통해, 정상적인 실제 사용자가 올바른 Sybil-resistant trust value 를 가지기 위해서는 10 개 이상의 이웃노드를 가져야 한다는 점과, Random Route Tail Intersection 이 Sybil 사용자 그룹의 Attack Edge 의 영향을 가장 적게 받는 다는 점을 확인 하였다.

2. OSN 기반의 Sybil-resistant value 추출 기법

성능평가에 사용된 OSN 기반의 Sybil-resistant trust value 추출기법은 SybilGuard[5]기반의 RRI (Random Route Intersection), SybilLimit[6]기반의 RRTI (Random Route Tail Intersection) 그리고 RWTI (Random Walk Tail Intersection)이다.[4] 이들 기법들은 전체 OSN 그래프를 알고 있는 중앙 서버에서 수행된다. 사용되는 OSN 그래프는 실제 사용자들로 구성된 honest region 과 Sybil 사용자들로 구성된 Sybil region 으로 이루어지고 각각의 region 은 각기 하나의 strongly connected component 이다. Sybil region 과 honest region 은 몇몇 Edge 들로 연결되어 전체 OSN 그래프가 하나의 strongly connected component 가 된다. 이때, Sybil region 과 honest region 을 연결하는 edge 를 Attack Edge 라 한다.

RRI, RRTI, 그리고 RWTI 에서 Sybil-resistant trust value 를 추출하기 위해서 우선 몇몇 임의의 실제 사용자들에 해당하는 OSN 그래프 상의 노드들을 verifier 노드들로 선택한다. 각 verifier 노드들은 각 추출 기법에서 정의된 판별 방법을 통해 OSN 그래프 상의 임의의 노드가 Sybil 사용자인지 실제 사용자인지를 판별한다. 임의의 노드를 실제 사용자에 해당한다고 판별한 verifier 노드를 accepted verifier 노드라 할 때, Sybil-resistant trust value 는 (accepted verifier 노드의 개수)/(전체 verifier 노드의 개수) 와 같이 구할 수 있다. 즉, Sybil-resistant trust value 는 0 에서 1 까지의 값을 가지고 1 에 가까운 값을 가질수록 해당 노드가 실제 사용자에 해당한다고 말 할 수 있다.

RRI 에서는 SybilGuard[5]에서 제안된 random route 를 사용한 판별방법을 사용한다. Random route 는 한 노드의 임의의 이웃 노드를 출발점으로 하여 각 노드에 미리 설정되어 있는 Routing Table 을 따라서 노드들을 방문함으로써 얻을 수 있다. Random route 의 길이가 w 일 경우 w 개의 노드를 방문한 후 random route 프로세스가 멈추게 된다. 각 노드에 설정되어 있는 Routing Table 은 random route 가 들어오는 이웃노드와 나가는 이웃노드를 연결해 놓은 테이블이다. RRI 에서는 verifier 의 random route 가 임의의 노드의 random route 와 만나는 부분이 있을 경우 verifier 는 해당 노드가 실제 사용자에 해당한다고 판별한다. RRI 에서 사용되는 random route 의 길이 w 는 OSN 그래프의 노드의 개수를 n 이라고 할 때 가 되어야 한다.

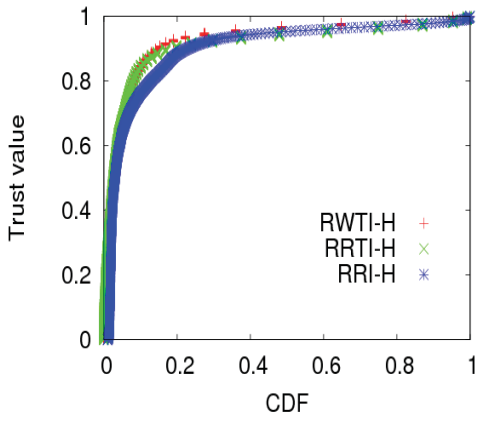
RRTI 에서는 SybilGuard[6]에서 제안된 random route 의 tail 을 비교하는 판별방법을 사용한다. 즉, random route 의 마지막 edge 을 비교한다. 각 노드는 r 개의 tail 을 구하여 저장하고, verifier 의 tail 의 set 과 임의의 노드의 tail 의 set 중에서 같은 tail 이 있을 경우 verifier 는 해당 노드가 실제 사용자에 해당한다고 판별한다. 이때, r 개의 tail 을 구하기 위해서는 r 개의 서로 다른 routing table 을 각 노드가 가지고 있어야 한다. 이때 사용되는 r 은 OSN 그래프의 edge 의 개수를 m 이라 할 때 가 되고, 각 random route 의 길이 w 는 가 되어야 한다.

RRWI 는 random walk 의 tail 을 비교하는 판별방법을 사용한다. RRTI 와 비슷하게 r 개의 tail 을 사용하여 verifier 는 임의의 노드의 실제 사용자 여부를 판별한다. 하지만, RRTI 와는 다르게 각 노드는 routing table 을 준비하지 않고 random walk 기법을 통해서 tail 을 수집한다. 이때 사용되는 r 과 w 는 RRTI 와 같다.

3. 성능 평가

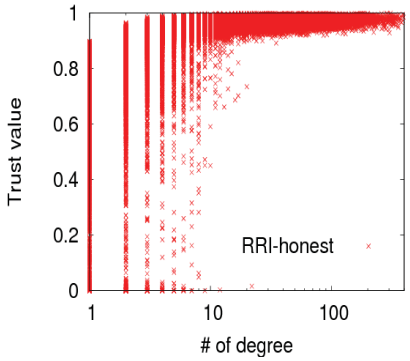
세가지 Sybil-resistant trust value 추출 기법의 성능을 평가하기 위해서 Facebook 에서 추출된 샘플 그래프를 사용하였다. Facebook 에서 추출된 OSN 샘플 그래프는 노드의 개수가 100,000 이고 edge 의 개수는 1,861,360 이고 Diameter 가 18 인 하나의 strongly connected component 이다. 이 샘플그래프를 honest region 으로 가정하였다. Sybil region 을 위해서는 각 노드의 평균 이웃 노드의 개수가 14 인 1000 개의 노드로 이루어진 하나의 strongly connected component 를 생성하였다. Sybil region 의 attack edge 는 임의로 생성하였다. 각 추출기법에서 사용되는 verifier 들은 honest region 에서 random 하게 선택되고 그 수는 100 으로 설정하였다.

우선 각 추출기법들이 honest region 의 노드들을 평가하는 성능을 측정하였다. 이를 위해, RRI 의 w 는 200, RRTI 의 w 는 15, r 은 2000, 그리고 RWTI 의 w 는 20, r 은 2000 으로 설정하였다. 그림 1 에서는 세가지 추출기법들을 통해 추출해낸 honest region 의 노드들의 trust value 의 분포를 Cumulative Distribution Function 형태로 나타내었다. 이 그림에서, RRTI 와 RWTI 를 사용할 경우, honest region 의 약 90%정도는 0.8 이상의 trust value 를 가지는 것을 알 수 있는 반면, RRI 를 사용할 경우 honest region 의 약 85%정도만이 0.8 이상의 trust value 를 가지는 것을 알 수 있다.

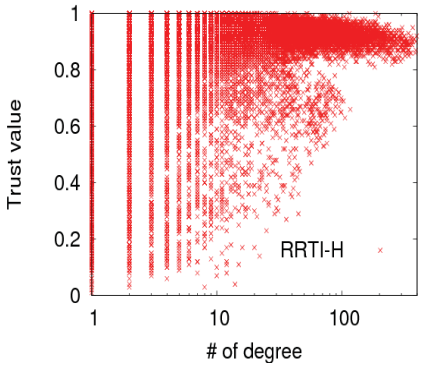


(그림 1) Honest Region 노드들의 Trust Value 분포

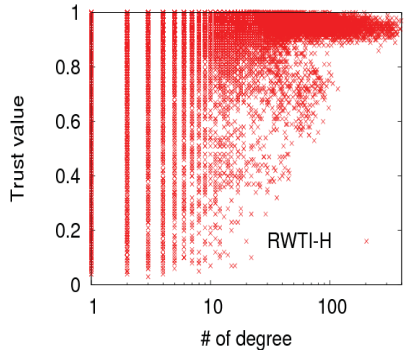
이러한 각 추출 기법들에 따른 honest region 노드들의 trust value 의 분포를 더욱 자세히 보기 위해 그림 2, 그림 3, 그림 4 에서는 각 추출 기법들에 따른 trust value 의 분포를 각 노드의 이웃노드 개수(degree)에 따라서 표현한다. 그림 2 에서는 RRI 를 사용할 경우 노드의 degree 가 10 보다 클 경우 trust value 의 값이 0.9 이상의 값이 되는 반면, 10 보다 작을 경우 아주 작은 trust value 를 가질 수 있음을 확인 할 수 있다. 그림 3 에서는 RRTI 를 사용할 경우 RRI 의 결과와 비슷하게 노드의 degree 가 10 보다 커야 큰 값의 trust value 를 가질 수 있게 된다. 또한 RRTI 의 경우에는 노드의 degree 가 100 보다 작을 경우에는 중간값(0.3 ~ 0.8) 정도의 trust value 를 가질 수 있음을 확인하였다. 그림 4 는 그림 3 과 비슷한 분포를 가지고 있음에 따라 RWTI 와 RRTI 는 비슷한 특성을 가짐을 확인 하였다



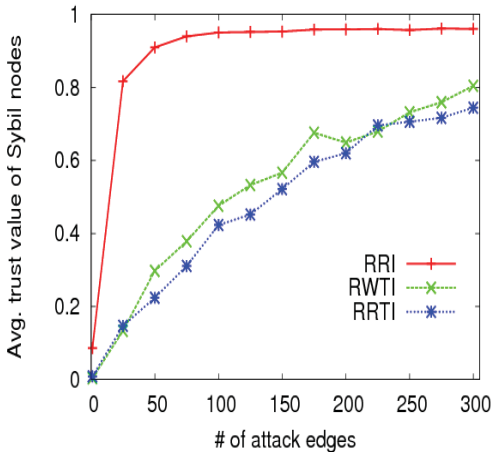
(그림 2) RRI: Trust values of Honest nodes



(그림 3) RRTI: Trust values of Honest nodes



(그림 4) RWTI: Trust values of Honest nodes

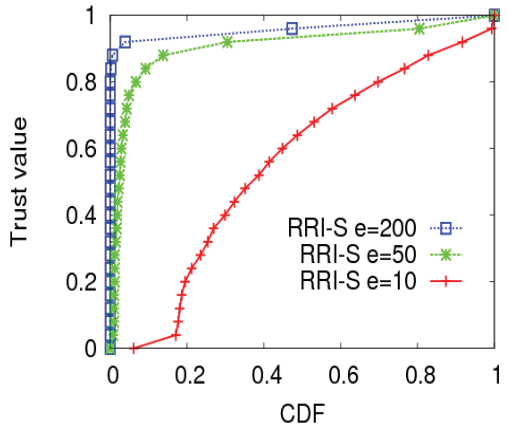


(그림 5) Attack edge 의 개수에 따른 sybil region 노드들의 평균 trust value 변화

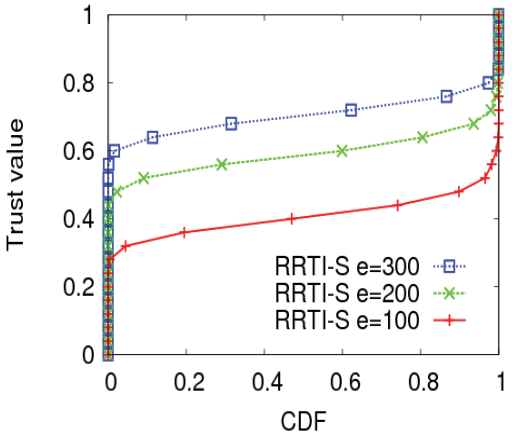
각 추출 기법들이 Sybil region 의 노드들에 대한 trust value 를 추출하는 성능을 평가하기 위해, Sybil region 과 honest region 사이의 attack edge 의 개수를 변화 시켜가며 sybil region 의 노드들의 trust value 를 측정하였다. 그림 5 에서 각 추출기법들에 의해 얻어진 sybil region 노드들의 trust value 의 평균값을 attack edge 의 개수에 따라서 나타내었다. 이 그림에서 RRI 기법의 경우, attack edge 가 20 개 정도 되더라도 Sybil region 노드들의 평균 trust value 의 값이 0.8 이상 됨을 확인 할 수 있다. 즉, RRI 기법으로 추출한 trust value 는 attack edge 에 심각하게 영향을 받는 것으로 확인되었다. 반면, RRTI 와 RWTI 의 경우, attack edge 가 200 개 정도 될 경우 평균 0.6 의 trust value 를 Sybil region 의 노드들이 가지는 것을 확인 할 수 있었다. 즉, 상대적으로 RRI 기법보다는 RRTI 와 RWTI 기법이 attack edge 의 개수에 덜 민감하게 반응함을 확인 하였다. 또한 RRTI 가 RWTI 보다는 미세하게나마 attack edge 의 영향을 덜 받는 것으로 확인 되었다.

이러한 attack edge 의 각 추출 기법들에 대한 영향을 보다 자세히 알아 보기 위해 그림 6, 그림 7, 그림 8 에서는 각 추출 기법들을 사용해서 얻어진 Sybil region 노드들의 trust value 의 분포를 Cumulative Distribution Function 의 형태로 표현하였다. 그림 6 에서는 RRI 를 사용할 경우 attack edge 의 개수가 10 일 경우 Sybil region 노드의 40%가 0.6 이상의 trust value 를 가지게 되고, attack edge 의 개수가 50 일 경우 Sybil region 노드의 90%가 0.8 이상의 trust value 를 가지게 되는 것을 확인 할 수 있다. 즉 RRI 의 경우 attack edge 를 포함하는 Sybil region 노드의 근처 노드들은 높은 trust value 를 가지게 됨으로써 attack edge 의 개수가 늘어남에 따라 급속도로 평균 trust value 가 증가하게 된다. RRTI 를 사용할 경우에도 그림 7 과 같이 attack edge 가 증가할 때 전체 Sybil region 노드들의 trust value 는 평균적으로 증가한다. 하지만, RRI 의 경우와 비교할 때 RRTI 의 경우에서 크게 다른 점은 attack edge 의 개수가 증가함에 따라 모든 노드들의

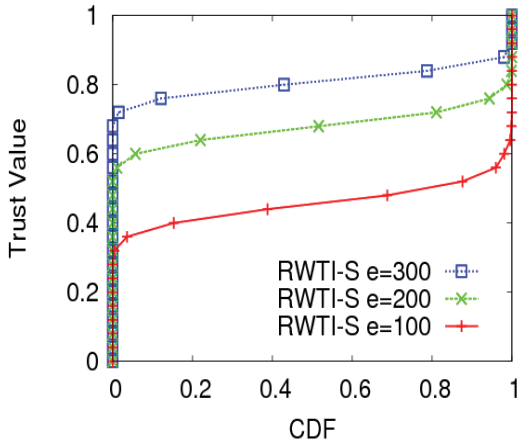
trust value 가 서로 비슷한 값을 가지면서 증가한다는 점이다. 즉 attack edge 가 300 일 경우, Sybil region 노드들의 평균 trust value 의 값은 약 0.7 이지만, 그 최대 값은 0.8 정도로 RRI 의 경우 몇몇 노드가 0.9 이상의 값을 가지는 점과 크게 다르다는 것을 알 수 있다. RWTI 의 경우도 그림 8 에서 나타낸 것과 같이 RRTI 와 비슷한 특성을 가지는 것을 알 수 있다.



(그림 6) RRI : Trust value of Sybil Nodes



(그림 7) RRTI : Trust value of Sybil Nodes



(그림 8) RWTI : Trust value of Sybil Nodes

4. 결론

다양한 특성을 가진 구성요소를 가지고 서로 다른 성향을 가진 사용자들에 의해서 운용되는 분산시스템을 Sybil attack 에 잘 견딜수 있도록 하는 것은 중요하다. 이 논문에서는 OSN 그래프 분석에 기반을 둔 Sybil 노드 탐색 기법의 원리를 사용한 sybil-resistant trust value 를 추출하는 기법들의 성능을 평가 하였다. 성능평가 결과, 제안된 추출 기법들을 통해 높은 (0.8~) trust value 를 얻기 위해서는 OSN 그래프 상에서 이웃노드가 최소한 10 개 이상 되어야 하고, RRTI 나 RWTI 의 경우에는 이웃노드가 100 개 이상 되어야 0.8 이상의 trust value 값을 가질 수 있음을 확인 하였다. 또한, Sybil region 의 attack edge 의 영향은 RRTI 기법이 가장 적게 받는 것으로 확인 되었다

참고문헌

[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