무선 센서 네트워크에서의 소스 위치 프라이버시 : 익명성 제공 방법 및 수학적 분석*

이송우, 박영훈, 손주형, 서승우, 강유*, 문호건*, 이명수*

서울대학교 전기컴퓨터공학과, *KT연구소 정보보호단

Source-Location Privacy in Wireless Sensor Networks : Countermeasures and Analysis

Song-Woo Lee, Young-Hun Park, Ju-Hyung Son, Seung-Woo Seo, Yu Kang*, Ho-Kun Moon*, Myuong-Soo Lee*

School of Electrical Engineering and Computer Science, Seoul National University, *KT Information Security Center, Korea Telecom

요 약

최근 센서 네트워크 분야를 비롯해 네트워크 분야에서 내용 보호, 인증뿐만 아니라 Source의 위치 프라이버시와 관련하여 많은 연구가 이루어지고 있다. 특히 센서 네트워크에서의 Source 위치는 실제 센서의 지리적 위치이기 때문에 Source의 위치를 노출하지 않는 것이 매우 중요하다. 그리고 인터넷과 Ad-Hoc 네트워크에서 익명성(Anonymity)을 제공하기 위한 기법들이 많이 제안되었지만, 이러한 기법들은 센서 네트워크에 적합하지 않기 때문에 센서 네트워크의 특성에 맞는 익명성 제공 방법이 요구되고 있다. 본 논문에서는 Source의 익명성과 관련해 센서 네트워크에서 나타날 수 있는 Eavesdropper의 유형을 정의하고, 이러한 Eavesdropper의 유형에 따라 Source의 익명성을 제공할 수 있는 방법을 제안하였다. 그리고 제안한 방법이 얼마만큼의 익명성을 제공하는가를 정량화하기 위해 엔트로피(Entropy) 성질을 이용해 수학적으로 분석하였다. 그 결과, 제안하는 방법이 보다 높은 익명성을 제공하고, 센서의 전송 거리가 Source의 익명성 제공에 있어 매우 중요한 요소임을 확인하였다.

I. 서론

저비용, 고성능의 센서 제작 기술의 발달로 사용 목적에 맞게 다양한 센서 네트워크 Application이 등장하고 있다. 하지만 센서의 주 사용 목적은 주변 환경이나 중요한 자산 등을 모니터링 하는 것이 될 것이다. 이러한 센서의 모니터링을 통해 얻어진 정보는 중요도에 따라 다르겠지만 제 3자에게 노출이 되어서는 안 될 것이다. 그리고 여기서 주목해야 할 것은 센서가 모니터링한 환경이나 자산 자체에 대한 정보, 즉 내용(Contents)이 중요할 수도 있지만, 정보를 전송한 센서의 위치(Source-location)가 더 중요할 수 있다는 것이다. 따라서 암호화를 통한 정보 보호 이외에 센서의 위치 노출을 방지하기 위한 추가적인 대책이 필요하다.

인터넷과 같은 유선 네트워크에서는 Sender의 익명성을 보장하기 위해 여러 가지 방법이 제안되어 왔다. 첫 번째로 Chaum이 제안한 "Mix-net"[1]은 공개키 방식으로 중간에 있는 Mix 서버의 공개키로 주소를 비롯한 데이터를 암호화하여 Mix 서버를 통해 목적지까지 보냄으로써 Sender의 익명성을 보장해 주는 방법이다. 두 번째로 "Crowds"[2]는 AT&T에서 제안한 것으로 WWW 환경에서 웹 서버에 "Request"를 보낼 때 사용자의 익명성을 보장해 주는 방법이다. 이 방법에서는 사용자가 "Crowds"라는 익명의 사용자 그룹에 가입하게

* 본 연구는 (주)한국통신 및 ITRC (대학정보통신연구센터) 지원으로 수행되었음.

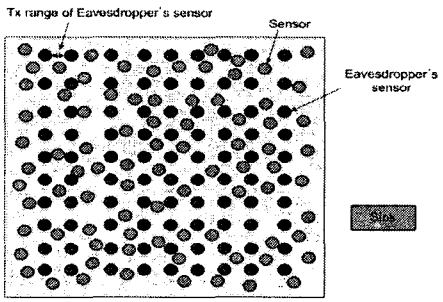

그림 1. Global Eavesdropper

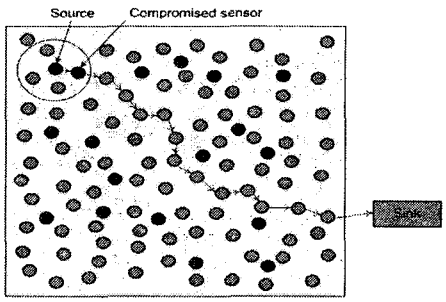

그림 2. Compromising Eavesdropper

되고, 사용자의 "Request"는 그룹 내 멤버들 간의 무작위 경로를 통해 웹 서버에 전달된다. 이러한 방법들은 환경 자체가 다르고 센서 네트워크에서 적용하기 어려운 공개키 방식을 사용하기 때문에 센서 네트워크에 적용하기 어렵다.

따라서 본 논문에서는 센서 네트워크의 특성을 고려하여 Source의 익명성을 제공하기 위한 방법을 제시하였다. 이를 위해 익명성을 해칠 수 있는 Eavesdropper의 유형을 정의하고 각 Eavesdropper의 유형에 따른 대책을 제시하였고, 그에 따른 Source의 익명성 정도(Degree of Anonymity)를 수학적으로 분석하여 정량화하였다. 익명성 정도는 Eavesdropper의 입장에서 판단할 수 있는 각각의 센서가 source가 될 확률을 이용해 구할 수 있었다. 본 연구는 센서 네트워크에서 Eavesdropper의 유형에 따라 Source 익명성 제공을 위한 방법을 처음으로 제시하였고, 무선 환경에서 Source의 익명성 정도를 수학적으로 분석했다는 점에서 중요한 의미를 갖는다.

본 논문의 다음과 같이 구성되어 있다. 2장에서는 센서 네트워크에서 Source의 익명성을 제공하기 위해 제시한 방법을 제시하고, 3장에서는 Source의 익명성 정도를 수학적으로 분석하고 그 결과를 정량화 한다. 마지막으로 4장에서는 본 논문의 결론을 제시한다.

II. Source 익명성 제공 방법

이 장에서는 센서 네트워크에서 나타날 수 있는 2 가지의 Eavesdropper 유형에 대해 정의하고, Source의 익명성을 제공하는 방법에 대해 구체적으로 설명하고자 한다.

2.1 Eavesdropper의 유형

Eavesdropper는 데이터의 내용이 아닌 데이터를 보낸 Source의 위치를 알아내려고 하는 Attacker의 형태이다. Eavesdropper의 유형은 크게 두 가지로 나누었다. 첫 번째는 네트워크 지역에서 발생하는 모든 신호를 감지할 수 있는 Global Eavesdropper이고, 두 번째는 배치된 일부 센서를 compromise하여 그 센서들을 이용하는 Compromising Eavesdropper이다.

Eavesdropper 유형 분류 시 무선 환경에서는 유선 환경에서와는 달리 전송 신호가 감지되어 Source의 위치가 노출될 수 있다는 특성을 고려하였다. 그리고 Eavesdropper는 Resource가 충분하고, 발견되지 않기 위해 데이터를 조작하거나 데이터 전송을 방해하는 등의 악의적인 행동은 하지 않으며, 센서의 ID와 위치를 알고 있다고 가정하였다.

2.1.1 Global Eavesdropper

이 Eavesdropper는 네트워크 지역에서 발생하는 모든 신호를 감지하기 위해 네트워크 지역에 자신의 센서를 설치한다. 그리고 Eavesdropper는 자신이 배치한 센서에 의해 감지된 신호를 종합하여 최초의 신호 발생 위치와 전송 경로를 파악할 수 있다. 그러나 전송되는 데이터의 내용은 알 수 없다. 그림 1과 같이 일반 센서는 랜덤하게 배치되고, Eavesdropper는 네트워크 전 지역을 모니터링하고 최대의 효과를 얻기 위해 센서를 균일하게 배치한다고 가정하였다. 그리고 Eavesdropper의 센서는 일반 센서의 전송 거리와 동일하다고 가정하였다.

2.1.2 Compromising Eavesdropper

이 Eavesdropper는 네트워크 지역에 뿌려진 센서 중의 일부를 compromise할 수 있으며, Eavesdropper에 의해 compromise된 센서들이 수신하는 데이터를 통해서 Source의 위치를 파

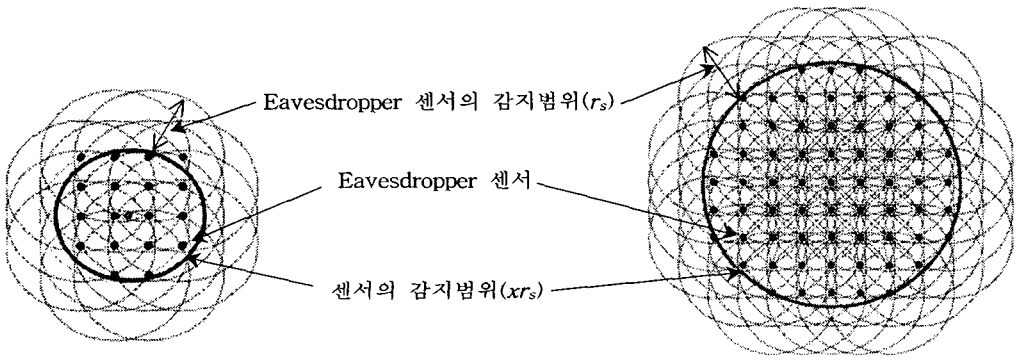

그림 3(a). 전송 power level이 1인 경우

그림 3(b). 전송 power level이 2인 경우

악할 수 있다. 단, compromise된 센서가 전송 경로 상에 있으면서 Source와 이웃해 있을 때만 위치를 알아낼 수 있는 유형이다.

2.2 익명성 제공 방법

이 절에서는 Source의 익명성을 제공하기 위한 방법에 대해 제안한다. 제안하는 방법의 기본 개념은 다음과 같다.

1) 위장 ID(Forged ID) 사용
2) 데이터 프레임 수정
(Source ID 생략/ Message Authentication Code 사용)
3) 동적인 전송 power level 조절

2.2.1 위장 ID

익명성 보장을 위해서는 기본적으로 위장 ID를 사용해야 한다. 위장 ID를 사용함으로써 ID만으로는 위치가 노출되지 않는다. 이러한 위장 ID는 최초 센서가 뿌려진 후 이웃 Discovery 과정에서 이웃 센서들 간에 교환되며, 각 센서는 특정 범위 내에서 임의의 숫자를 선택하여 위장 ID로 사용한다. 만약 이웃 센서들 중 동일한 ID가 있을 경우에는 다시 임의의 숫자를 선택한다. 각 센서는 실제 ID와 위장 ID의 테이블을 저장하며 위장 ID는 주기적으로 갱신된다.

2.2.2 데이터 프레임 수정

Source ID	Destination ID	Payload	CRC

그림 4(a). original 데이터 프레임

Source ID	Destination ID	Payload	MAC

그림 4(b). 수정된 데이터 프레임

데이터 전송 시 Source ID를 생략하고 Destination ID만을 사용하고, CRC대신에 MAC(Message Authentication Code)을 추가한다. 그리고 데이터는 Sink와 공유하고 있는 대칭키로 암호화된다. 이를 위해 센서들은 최초 배치되기 전후에 Sink와 이웃 센서들의 대칭키를 공유하게 된다. 데이터 전송 시 Source ID를 생략함으로써 Destination 센서만이 누가 Source인 지 알 수 있게 되는데, 그 이유는 MAC을 확인하기 위해서는 Source와 동일한 키를 가지고 있어야 하기 때문이다. 즉, MAC 확인이 가능하다는 것은 누가 보낸 데이터인가를 알 수 있다는 것을 의미한다. 반대로 동일한 키를 갖고 있지 않은 이웃 센서들은 Source가 누구인지를 알 수 없을 뿐만 아니라 예측하기도 어렵다. 앞서 언급한 위장 ID를 사용하고 데이터 프레임을 수정함으로써 Eavesdropper가 Source를 찾아내는데 혼란을 줄 수 있다.

2.2.3 동적인 전송 Power Level 조절

Eavesdropper의 유형에 따라 전송 power level을 동적으로 조절함으로써 데이터 전송 시 Source의 익명성 정도를 높일 수 있다. 이 절에서는 Eavesdropper의 유형에 따라 어떻게 전송 power level를 조절해야 하는 지를 설명한다.

2.2.3.1 Global Eavesdropper의 경우

이 경우에는 Source의 전송 power level을 어떻게 정하느냐가 중요한 요소가 된다. 즉, Source가 최초 전송 시에 전송 power level을 높임으로써 그림 3(b)와 같이 더 많은 Eavesdropper의 센서가 신호를 감지하도록 해

Eavesdropper가 Source의 위치를 유추하기 어렵게 할 수 있다. 다시 말해 Eavesdropper가 유추할 수 있는 지역의 범위를 넓게 함으로써 Eavesdropper가 Source라고 예측할 수 있는 센서의 수가 증가되어 Source의 익명성 정도를 높일 수 있게 되는 것이다. 하지만 전송 power level을 높이는 것은 센서의 전력 소모를 동반하기 때문에 최적의 전송 power level을 결정하는 것이 매우 중요하다.

2.2.3.2 Compromising Eavesdropper의 경우

이 경우에는 Source로부터 Sink 노드까지의 경로길이에 따라 익명성의 정도가 결정 된다. 다시 말해, 경로 길이가 길어질수록 Source의 익명성 정도가 높아지게 된다. 따라서 센서의 전송 power level을 낮추는 것이 중요하다. 센서의 전송 power level을 줄임으로써, 즉 센서의 전송거리를 짧게 하여 경로길이를 길게 함으로써 Source의 익명성 정도를 높일 수 있다. 경로 길이를 줄임으로써 익명성의 정도가 높아지는 이유에 대해서는 3장에서 확인할 수 있다.

III. 익명성 정도에 대한 수학적 분석

이 장에서는 위에서 제시한 방법을 적용하였을 때 Source의 익명성 정도가 어떻게 변화되는지를 [3]에서 적용한 바와 같이 엔트로피의 관계식을(식 1) 이용하여 분석하였다. 익명성의 정도는 Eavesdropper의 입장에서 판단할 수 있는 각각의 센서가 Source가 될 확률로 구할 수 있는데, 예를 들어, 모든 센서가 Source가 될 확률이 동일하다고 Eavesdropper가 판단할 때 Source의 익명성 정도가 최대가 된다. 그리고 바로 이때, 네트워크의 엔트로피의 값이 최대가 된다. 이러한 엔트로피의 성질을 이용하여 신호 발생 이전의 최대 엔트로피 값과 신호가 발생된 이후의 엔트로피 값과의 관계를 통해 센서 네트워크에서 Source의 익명성 정도를 정량화 하였다.

$$H_M = -\sum_{k=1}^{n} p_k \log_2 p_k, \quad p_k = p \text{ for } all \qquad (1)$$

3.1 Global Eavesdropper

표 1. Parameters

• # of sensors : n
• initial tx range : r_0
• tx power level : x
• tx range : $r=r_0x$
• sensing range : $r_s=2.2r$
• total area : $A(r)$
※ Eavesdropper의 센서들이 감지 가능한 총 면적
• sensor density : e

$$H_M = -\sum_{k=1}^{n} \frac{1}{n} log_2 \frac{1}{n} = \log_2 n \qquad (2)$$

H_M은 모든 센서가 Source일 확률이 동일한 경우로 최대 엔트로피 값이다. 신호가 발생한 이후에는 식 3과 같이 엔트로피 값이 변화한다. 왜냐하면 데이터를 전송하게 되면 Eavesdropper의 센서가 그 신호를 감지하여 Source의 위치를 유추할 수 있는 지역의 범위가 변하기 때문이다. 즉, 신호를 감지한 Eavesdropper의 센서가 감지할 수 있는 범위 내에 있는 센서가 Source일 확률이 더 높아지게 되고, 그 외의 지역에 있는 센서는 Source일 확률은 '0'이 되는 것이다. 그리고 앞서 제안한 익명성 제공 방법을 사용함으로써 범위 내에 있는 센서들이 Source일 확률은 모두 동일하게 된다.

$$H_X = -\sum_{k=1}^{\rho A(r)} \frac{1}{\rho A(r)} log_2 \frac{1}{\rho A(r)} = \log_2 \rho A(r) \qquad (3)$$

따라서 Eavesdropper는 감지 범위 내에 있는 센서 중에 하나가 Source일 것이라 판단하게 된다. 이를 바탕으로 Source의 익명성 정도를 구하면 식 4와 같다. 그리고 Eavesdropper가 예측할 수 있는 전체 범위(그림 3)의 면적은 식 5와 같이 원의 면적으로 근사할 수 있다.

$$D(x) = \frac{\log_2(\rho A(r))}{\log_2 n} = \frac{\log_2(\rho\pi(xr_s + r_s)^2)}{\log_2 n} \qquad (4)$$

$$A(r) \approx \pi(xr_s + r_s)^2 \qquad (5)$$

식 4에서 보듯이 Source의 익명성 정도는 전

송거리 r이 증가할수록 커짐을 알 수 있다. 그러나 2장에서도 언급했듯이 전송거리 r을 증가하게 되면 센서의 전력 소모량도 증가하기 때문에 최적의 r을 결정하는 것이 중요하다. 이를 위해 전력의 소모량은 전송거리의 제곱에 비례하는 성질을 이용하고, 익명성 증가에 따른 이득이 익명성 정도에 비례한다고 가정하였다. 그러면 전송거리 증가에 따른 전력 소모로 인한 손실과 익명성 정도의 증가로 인한 이득과의 차(net benefit)를 식 6과 같이 표현할 수 있다.

$$B(x) = \alpha D(x) - \beta r^2 = \alpha \frac{\log_2(\rho\pi(xr_s + r_s)^2)}{\log_2 n} - \beta r^2 \quad (6)$$

이때, 이 값을 미분하고 그 값이 '0'이 되게 하는 r을 구하면 그 r이 바로 최적의 전송 거리가 된다.

$$\frac{d}{dx}B(x) = \frac{d}{dx}\left(\alpha \frac{\log_2(\rho\pi(xr_s + r_s)^2)}{\log_2 n} - \beta r^2\right) = \frac{\alpha}{(\frac{r}{r_0}+1)\ln n} - \beta r r_0$$

$$r = \frac{-r_0 + \sqrt{r_0^2 + 4\alpha/(\beta \ln n)}}{2} \quad (7)$$

3.2 Compromising Eavesdropper

표 2. Parameters

• # of sensors : n
• # of compromised sensors : c
• Average path length : l
※ 센서의 tx range가 r일 때

전체 센서 중 c개가 eavesdropper에 의해 compromise 되었다고 가정하면 최대 엔트로피 값은 식 8과 같이 구할 수 있다.

$$H_M = -\sum_{k=1}^{n-c} \frac{1}{n-c} log_2 \frac{1}{n-c} = \log_2(n-c) \quad (8)$$

Eavesdropper는 식 9·10과 같이, compromise된 센서에 이웃하고 있는 센서 중에 하나에게 가장 높은 확률을 할당하고 나머지 센서들에게는 동일한 확률을 할당할 것이다. 식 9는 평균경로길이가 L인 전송 경로 상에 compromise된 센서가 하나 이상 존재할 때, 첫 번째 홉에 위치한 센서가 compromise된 센서일 확률을 의미한다. 따라서 Eavesdropper가 compromise한 이후의 엔트로피 값을 구하면 식 11과 같다.

$$p_{c+1} = \frac{c/(n-1)}{1 - \prod_{x=1}^{L-1}(1 - \frac{c}{n-x})} \quad (9)$$

$$p_i = \frac{1 - p_{c+1}}{n-c-1} \qquad c+2 \le i \le n \quad (10)$$

$$H_X = p_{c+1}\log_2(p_{c+1}) + (1-p_{c+1})\log_2\left[\frac{n-c-1}{1-p_{c+1}}\right] \quad (11)$$

마찬가지로, $D(x)=H_x/H_M$이므로 Source의 익명성 정도를 구할 수 있다.

3.3 분석 결과

3.3.1 Global Eavesdropper

그림 5는 센서의 전송거리에 따른 Source의 익명성 정도를 보여주고 있다. 여기서는 네트워크 지역의 면적을 2km×2km, 센서의 밀도를 320개/km^{2}로 가정하여 익명성 정도를 구하였다. 그림에서 보듯이 전송거리가 증가함에 따라 익명성 정도는 높아짐을 알 수 있다. 전송 거리를 100m에서 400m로 증가시켰을 경우 Source의 익명성 정도는 약 0.26이 높아지게 되는데, 이는 최초 Source가 데이터를 전송하는 경우 Eavesdropper가 Source라고 예측할 수 있는 센서의 개수가 150여개에서 900여개로 증가하여 Eavesdropper가 Source를 알아낼 확률이 현저히 낮아짐을 의미한다.

3.3.2 Compromising Eavesdropper

그림 6은 compromise된 센서 수에 따른 Source의 익명성 정도의 변화를 보여주고 있다. 그림에서 보듯이 compromise된 센서 수가 증가함에 따라 익명성 정도는 낮아짐을 알 수 있다.

그림 7은 전송 경로 길이에 따른 Source의 익명성 정도를 나타내고 있다. 경로 길이가 증가함에 따라 익명성 정도도 높아짐을 알 수 있다. 그리고 초기에는 경로 길이가 길어질수록 익명성 정도가 급격히 높아지나 경로 길이가

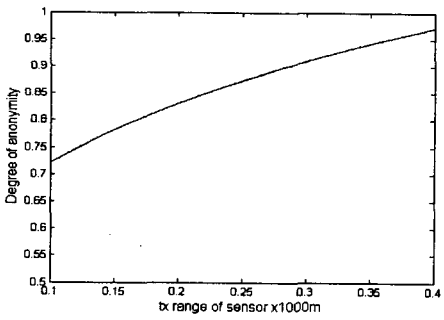

그림 5. 전송 거리에 따른 익명성 정도 변화 (n=1280)

16이상에서는 익명성 정도는 거의 같아진다. 따라서 최적의 경로 길이도 이 결과를 통해서 얻어낼 수 있다.

IV. 결론

본 논문에서는 센서 네트워크의 특성을 고려하여 센서 네트워크에서 익명성을 제공하기 위한 방법을 제시하였다. 그리고 각 방법에 대해 수학적으로 분석하여 각 방법이 얼마만큼의 익명성을 제공하는가에 정량화하였다. 또한 이러한 수학적 분석을 통해 제안하는 방법이 보다 높은 익명성을 제공함을 확인할 수 있었고, 센서의 전송 거리가 Source의 익명성 제공에 중요한 요소임을 알 수 있었다.

결론적으로 본 연구는 센서 네트워크에서 Eavesdropper의 유형에 따라 Source의 익명성 제공을 위한 방법을 처음으로 제시하였고, 무선 환경에서 Source의 익명성 정도를 수학적으로 분석했다는 점에서 중요한 의미를 갖는다.

[참고문헌]

[1] D. Chaum, Untraceable Electronic Mail, Return addresses and digital Pseudonyms, Communications of the ACM, Vol. 24, No. 2, pp. 84-88, Feb, 1981.

[2] Michael K. Reiter, Aviel D. Rubin, Crowds : Anonymity for Web Transactions, ACM Transactions on information and system security, 1998.

[3] Claudia Diax, Stefaan Seys, Joris Claessens, and Bart Preneel, Towards measuring anonymity, Appeared in Proceedings of PET, April, 2002.

[4] Andrei Serjantov, On the Anonymity of Anonymity Systems, Dissertation for the degree of Doctor of Philosophy in University of Cambridge, March, 2004.

[5] 이현숙, 변진욱, 박현아, 이동훈, 임종인, 익명 통신로에 관한 최근 연구 동향, 정보보호학회지, Vol. 14, No. 6, pp. 53-61, December, 2004.

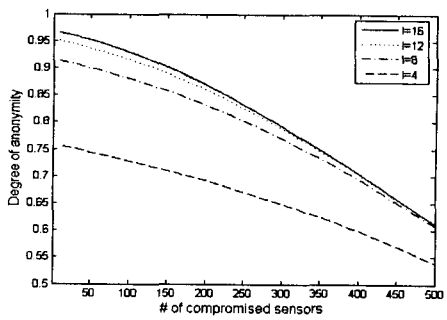

그림 6. compromise된 센서 수에 따른 익명성 정도 변화 (n=1000)

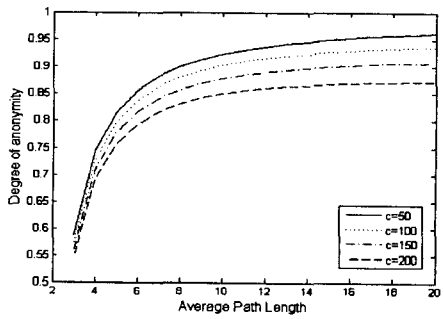

그림 7. 경로 길이에 따른 익명성 정도 변화 (n=1000)