ISSN 2383-630X(Print) / ISSN 2383-6296(Online)
Journal of KIISE, Vol. 44, No. 2, pp. 195-207, 2017. 2
https://doi.org/10.5626/JOK.2017.44.2.195

프라이버시 보호 데이터 배포를 위한 모델 조사

(Models for Privacy-preserving Data Publishing : A Survey)

김 종 선 † (Jongseon Kim)
정 기 정 † (Kijung Jung)
이 혁 기 † (Hyukki Lee)
김 수 형 †† (Soohyung Kim)
김 종 욱 ††† (Jong Wook Kim)
정 연 돈 †††† (Yon Dohn Chung)

요 약 최근 다양한 분야에서 데이터들이 활발하게 활용되고 있다. 이에 따라 데이터의 공유나 배포를 요구하는 목소리가 높아지고 있다. 그러나 공유된 데이터에 개인과 관련된 민감한 정보가 있을 경우, 개인의 민감한 정보가 드러나는 프라이버시 유출이 발생할 수 있다. 개인 정보가 포함된 데이터를 배포하기 위해 개인의 프라이버시를 보호하면서 데이터를 최소한으로 변형하는 프라이버시 보호 데이터 배포(privacy-preserving data publishing, PPDP)가 연구되어 왔다. 프라이버시 보호 데이터 배포 연구는 다양한 공격자 모델을 가정하고 이러한 공격자의 프라이버시 유출 공격으로부터 프라이버시를 보호하기 위한 원칙인 프라이버시 모델에 따라 발전해왔다. 본 논문에서는 먼저 프라이버시 유출 공격에 대해 알아본다. 그리고 프라이버시 모델들을 프라이버시 유출 공격에 따라 분류하고 각 프라이버시 모델들 간의 차이점과 요구 조건에 대해 알아본다.

키워드: 데이터 프라이버시, 프라이버시 모델, 익명화, 프라이버시 보호 데이터 배포

Abstract In recent years, data are actively exploited in various fields. Hence, there is a strong demand for sharing and publishing data. However, sensitive information regarding people can breach the privacy of an individual. To publish data while protecting an individual's privacy with minimal information distortion, the privacy- preserving data publishing(PPDP) has been explored. PPDP assumes various attacker models and has been developed according to privacy models which are principles to protect against privacy breaching attacks. In this paper, we first present the concept of privacy breaching attacks. Subsequently, we classify the privacy models according to the privacy breaching attacks. We further clarify the differences and requirements of each privacy model.

Keywords: data privacy, privacy model, anonymization, privacy-preserving data publishing

· 이 논문은 2015년도 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임(No.R0190-15-2019, 빅데이터 환경에서 비식별화 기법을 이용한 개인정보보호 기술 개발)

† 비 회 원 : 고려대학교 컴퓨터학과
jongseon9312@gmail.com
jungkj9400@gmail.com
sq5727@gmail.com
†† 비 회 원 : 고려대학교 IT융합학과
soohyungdkim@gmail.com
††† 정 회 원 : 상명대학교 미디어소프트웨어학과 교수
jongwook.kim1004@gmail.com
†††† 종신회원 : 고려대학교 컴퓨터학과 교수(Korea Univ.)
ydchung@korea.ac.kr
(Corresponding author임)

논문접수 : 2016년 8월 17일
(Received 17 August 2016)
논문수정 : 2016년 10월 19일
(Revised 19 October 2016)
심사완료 : 2016년 10월 25일
(Accepted 25 October 2016)

정보과학회논문지 제44권 제2호(2017. 2)

1. 서 론

정부와 기업에서 데이터로부터 새로운 가치를 창출하기 위해 데이터를 활발하게 공유하고 있다. 그러나 이러한 데이터에 병명과 같은 개인의 민감한 정보가 있는 경우, 프라이버시를 고려하지 못한 데이터의 배포는 큰 문제를 일으킬 수 있다. 예를 들어, 미국 Netflix사에서는 효율적인 영화 추천 알고리즘 개발을 위해 사용자의 ID를 제거한 50만 건의 영화 평점 데이터를 배포하였다. Narayanan의 연구에서 이 데이터와 외부의 영화 평점 사이트의 데이터를 연결하여 특정한 개인을 식별하였으며 개인의 정치적 성향까지 알아내었다[1]. 또 다른 예로, Sweeney의 연구에서 그림 1처럼 선거인명부와 병원의 의료기록을 우편번호, 생일, 성별 속성으로 연결하여 매사추세츠 주지사의 병명을 알아내었다[2]. Sweeney의 연구에 따르면, 공개된 데이터를 사용하여 우편번호, 생일, 성별 3가지의 속성을 연결하면 미국 인구의 87%를 식별할 수 있다[2]. 위의 사례들과 같이 단순히 특정한 개인을 직접적으로 식별하는 속성을 제거하여 데이터를 배포하는 것만으로는 개인의 프라이버시를 보호할 수 없다. 이는 직접적으로 개인을 식별할 수 있는 속성을 제외하더라도 나머지 속성들을 서로 조합하거나 외부의 데이터와 연결하면 특정한 개인을 식별할 수 있기 때문이다.

데이터에 존재하는 특정한 개인이 식별되는 것을 막으면서 데이터의 유용성을 보존하기 위한 일련의 과정을 프라이버시 보호 데이터 배포(privacy-preserving data publishing, PPDP)라 부른다. 일반적인 프라이버시 보호 데이터 배포 과정은 그림 2와 같다.

프라이버시 보호 데이터 배포 과정은 데이터 배포자가 데이터 소유자로부터 데이터를 모으는 데이터 수집 과정과 데이터 배포자가 모은 데이터를 익명화하여 데이터 수령자에게 배포하는 데이터 배포 과정으로 이루어진다[3]. 예를 들어, 병원에서 환자 데이터를 외부의 의료 연구원에 배포한다고 하자. 이 때 데이터 소유자는 환자이며 배포자는 병원 그리고 수령자는 의료 연구원이다. 데이터 배포자가 데이터 수령자에게 원본 데이터를 배포하는 것은 프라이버시 유출의 위험이 있으므로 데이터 배포자가 데이터를 익명화하여 배포해야 한다.

이러한 프라이버시 보호 데이터 배포는 단순히 데이터를 암호화하여 특정인에게만 데이터를 제공하는 데이터베이스 암호화와는 다르다. 데이터베이스 암호화는 복호화 키를 가진 대상을 제외하고는 데이터를 볼 수 없지만, 프라이버시 보호 데이터 배포는 데이터를 받는 사람이 불특정 다수인 상황을 가정하기 때문이다. 불특정 다수 중에 선의의 데이터 사용자에게는 충분히 데이터

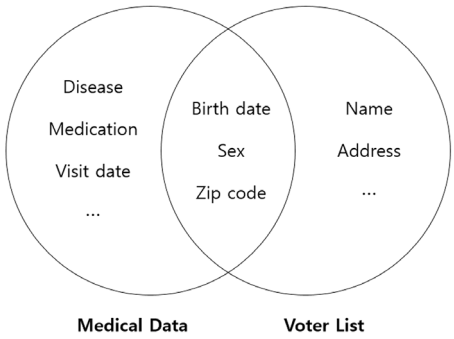

그림 1 선거인명부와 의료 데이터를 연결
Fig. 1 Linking voter list and medical data

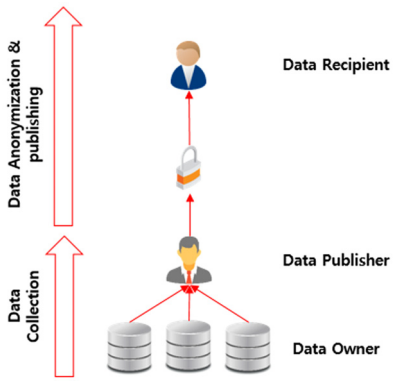

그림 2 데이터 수집 및 프라이버시 보호 데이터 배포 과정
Fig. 2 Process of data collection and privacy-preserving data publishing

의 유용성을 제공함과 동시에 개인의 프라이버시를 유출하려는 공격자로부터 데이터에 존재하는 개인의 프라이버시를 보호할 수 있어야 한다.

프라이버시 보호 데이터 배포를 위한 프라이버시 보호 원칙은 매우 다양하다. 어떠한 프라이버시 유출 공격에 대해 일정한 프라이버시 보호를 보장하는 원칙 또는 기준을 프라이버시 모델(privacy model)이라 부른다. 현재까지 다양한 프라이버시 유출 유형이 연구되었고 이러한 프라이버시 유출 공격으로부터 개인의 프라이버시를 보호하기 위해 수많은 프라이버시 모델이 제안되었다. 앞으로도 프라이버시 연구가 지속적으로 진행될 것이라 기대되는 바이므로 향후 프라이버시 보호 연구의 발전을 위해 지금까지 제안된 프라이버시 모델의 정리가 필요하다. 따라서 본 논문에서는 현재까지의 프라이버시 모델들이 가정하는 다양한 상황에 대해 논의하고, 모델들 간의 차이점 및 요구조건을 분석하며, 마지막으로 프라이버시 유출 공격에 따라 프라이버시 모델들을 분류한다.

본 논문의 구성은 다음과 같다. 2장에서는 프라이버시 보호 데이터 배포를 이해하기 위한 배경지식을 소개한다. 3장에서는 관련 연구에 대해 논의하며 4장에서는 관계형 데이터 프라이버시 모델에 대해 알아본다. 마지막으로 5장의 결론을 통하여 논문을 마무리한다.

2. 배경지식

본 논문에서 프라이버시 보호 데이터 배포는 다른 언급이 없다면 관계형 데이터에서 이루어진다고 가정한다. 관계형 데이터의 스키마는 특징에 따라 다음과 같이 분류할 수 있다.

D(식별자, 준식별자, 민감한 속성)

• 식별자(identifier)

식별자는 속성값 자체로 특정한 개인을 식별할 수 있는 속성이다 식별자의 예시로는 이름, 주민등록번호, 사회보장번호, 전화 번호, 이메일 주소 등이 있다.

• 준식별자(quasi-identifier)

준식별자는 속성값 자체로는 특정한 개인을 식별할 수 없지만 여러 개의 준식별자 속성이 조합되면 특정한 개인을 식별할 수 있는 속성이다. 준식별자의 예시로는 성별, 우편번호, 연령 등이 있다.

• 민감한 속성(sensitive attribute)

민감한 속성은 특정한 개인과 연결되면 프라이버시가 유출되는 속성으로 병명, 급여, 소득 분위 등이 있다.

• 식별자 제거(identifier removal)

식별자 제거는 식별자를 없애는 과정이다. 식별자를 제거하면 직접적으로 특정한 개인을 식별할 수 없지만 앞의 예시[1,2]처럼 준식별자를 통한 재식별이 가능한 상태이다. 식별자를 제거한 관계형 데이터의 스키마는 다음과 같다.

D'(준식별자, 민감한 속성)

표 1의 의료 데이터에서 식별자는 이름이며 준식별자는 우편번호, 연령, 성별이고 민감한 속성은 병명으로 볼 수 있다. 식별자인 이름을 제거하면 표 2와 같다.

• 익명화(anonymization)

익명화는 식별자가 제거된 데이터를 익명화 연산에 따라 테이블을 변형하여 프라이버시 모델을 만족시켜 특정한 개인의 식별이나 민감한 속성에 연결되는 것을 막는 기법이다. 익명화 과정에서 민감한 속성은 분석의 대상이 되는 속성이므로 삭제하지 않는다고 가정한다.

• 정보 손실(information loss)

정보 손실은 익명화 과정에서 준식별자가 얼마나 손상되었는지를 의미한다. 준식별자는 분석에서 사용될 수 있는 정보를 담고 있으며, 준식별자를 제거하면 데이터의 유용성이 크게 줄어들 수 있다. 그러므로 익명화 과정에서 프라이버시 보호 수준과 준식별자를 포함하여 얻는 유용성을 고려하여 준식별자를 최소한으로 변형해야 한다.

표 1 원본 환자 테이블

Table 1 Original table of patients

Identifier	Quasi-identifier			Sensitive Attribute
Name	Zipcode	Age	Sex	Disease
Amy	13053	38	Female	Diabetes
Bob	13068	49	Male	MERS
Clerk	13053	29	Female	Flu
David	13068	49	Male	MERS
Elvis	17583	70	Male	Pneumonia

표 2 식별자를 제거한 표 1

Table 2 Table 1 with identifier removal

Quasi-identifier			Sensitive Attribute
Zipcode	Age	Sex	Disease
13053	38	Female	Diabetes
13068	49	Male	MERS
13053	29	Female	Flu
13068	49	Male	MERS
17583	70	Male	Pneumonia

3. 관련 연구

2010년 Fung 등은 처음으로 프라이버시 보호 연구를 정리하였다[4]. Fung 등이 작성한 조사 논문은 2010년까지 연구된 프라이버시 모델과 익명화 알고리즘을 정리하였다.

2014년 Xu 등이 작성한 조사 논문[5]은 다양한 정보 손실 측정 기준을 바탕으로 익명화 알고리즘을 정리하였다. 그러나 Xu 등은 프라이버시 모델은 간단하게 언급하는 정도에서 끝냈다. 2014년에 Gkoulalas 등이 작성한 조사 논문[6]은 EHR(Electronic Health Record)의 프라이버시 보호 배포를 위한 조사로 확장하여 정리하였다. 그러나 Gkoulalas 등은 Xu 등과 마찬가지로 주로 익명화 알고리즘에 대해 정리하였다.

프라이버시 모델을 만족시키는 익명화 알고리즘은 프라이버시 모델에 따라 매우 다양하다. 대표적인 k-익명성 모델[2]을 만족시키는 알고리즘으로는 Incognito[7]와 Mondrian[8]이 있다. Incognito는 k-익명성 모델을 만족시키면서 정보 손실을 최소로 하는 최적의 해를 찾아낸다. Mondrian은 각 준식별자를 좌표축으로 하는 공간 내에서 k-익명성 모델을 만족하게 공간을 분할해나가면서 해를 찾아낸다. 두 알고리즘 모두 공통적으로 값을 더 상위 개념의 값으로 치환하는 방법을 사용한다. 이 외에도 테이블을 분리시켜 익명화하는 알고리즘인 Anatomy[9], Slicing[10], Disassociation[11] 등과 값에

노이즈를 더하여 익명화하는 기법인 Laplace Mechanism[12], Exponential Mechanism[13] 등이 있다.

본 논문에서는 이전 조사 논문에서 다뤄지지 않았던 2010년 이후에 연구된 프라이버시 모델을 추가하여 각 프라이버시 모델의 정의와 어떻게 프라이버시 유출 공격을 막는지를 중점적으로 논의하였다.

4. 관계형 데이터 프라이버시 모델

프라이버시 유출 공격은 공격자가 가지고 있는 배경지식을 통해 특정한 개인을 식별하거나 또는 특정한 개인의 민감한 속성을 유추하는 방법으로 이루어진다. 이러한 공격에는 신원 노출, 속성 노출, 귀속 노출, 확률 공격의 4가지 방법이 있다. 본 장에서는 프라이버시를 유출하는 공격 유형과 이를 막기 위한 프라이버시 모델에 대해 논의한다. 참고로, 본 논문은 동일한 데이터가 반복되어 배포되는 상황과 다수의 데이터 제공자가 결합하여 데이터를 배포하는 상황은 고려하지 않는다.

4.1 신원 노출(identity disclosure)을 막기 위한 프라이버시 모델

신원 노출은 임의의 레코드가 누구를 의미하는지 알아내는 프라이버시 유출 공격이다. 이러한 신원 노출은 공격자가 특정한 개인의 준식별자 속성값을 알고 있을 때 발생한다. 예를 들어, 공격자가 자신의 옆집에 사는 사람이 Amy이고 Amy의 우편번호와 연령, 성별이 {13053, 38, Female}라는 것을 알고 있다 가정하자. 만약 공격자가 병원으로부터 배포된 표 2를 참조하면 배경지식을 통해 첫 번째 레코드가 Amy를 의미한다는 것을 알 수 있다.

4.1.1 단일 테이블에서 신원 노출을 방어하는 프라이버시 모델

단일 테이블에서 신원 노출을 제한하는 대표적인 프라이버시 모델로는 k-익명성(k-anonymity) 모델[2]이 있다. k-익명성 모델은 준식별자를 익명화하여 각각의 레코드가 적어도 서로 구분되지 않는 k-1개의 레코드를 가지게 하는 모델이다.

예제 1. k-익명성을 통한 신원 노출 방어

그림 3의 범주 트리를 사용하여 준식별자 속성값을 일반화 할 수 있다.

표 3 2-익명성을 만족하는 표 2 (Elvis의 레코드 삭제)
Table 3 2-anonymity Table 2 (Remove Elvis's record)

	Zipcode	Age	Sex	Disease
t_1	[13060-13570]	[40-50]	Male	MERS
t_2	[13060-13570]	[40-50]	Male	MERS
t_3	[13050-13560]	[20-40]	Female	Flu
t_4	[13050-13560]	[20-40]	Female	Diabetes

표 4 4-익명성을 만족하는 표 2 (Elvis의 레코드 삭제)
Table 4 4-anonymity Table 2 (Remove Elvis's record)

	Zipcode	Age	Sex	Disease
t_1	[13050-13570]	[20-50]	*	Diabetes
t_2	[13050-13570]	[20-50]	*	MERS
t_3	[13050-13570]	[20-50]	*	Flu
t_4	[13050-13570]	[20-50]	*	MERS

표 3과 표 4는 각각 2-익명성과 4-익명성을 만족하게 표 2를 익명화한 테이블이다. 만약 공격자가 특정한 개인의 준식별자 속성값을 알아도 각각 2명과 4명의 레코드에 대응하므로 특정한 개인을 식별할 수 없다. □

정의 1. 동질 클래스(equivalence class)

동질 클래스는 동일한 준식별자 속성값을 가지는 레코드의 집합이다. 표 3에서 위에서부터 레코드를 $t_1,\cdots,t_4$라 할 때, $\mathrm{EC}_1=\{t_1,t_2\}$, $\mathrm{EC}_2=\{t_3,t_4\}$가 되며, 존재하는 모든 동질 클래스의 집합은 $\mathbb{EC}=\{\mathrm{EC}_1,\mathrm{EC}_2\}$이다. □

정의 2. k-익명성

테이블 T를 익명화하여 모든 동질 클래스의 크기가 자연수 k 이상일 때($\forall|\mathrm{EC}|\geq k$,), k-익명성을 만족한다. □

Wang 등은 원본 테이블의 스키마에 새로운 준식별자가 추가되었을 때, 이전에 익명화하여 배포한 테이블과 연결하여 신원 노출이 발생할 수 있음을 발견하였다. 이러한 문제를 해결하기 위하여 k-익명성 모델을 일반화한 (X,Y)-익명성((X,Y)-anonymity) 모델[14]을 제안하였다.

정의 3. (X,Y)-익명성

X와 Y는 서로 공통된 속성을 가지지 않는 분리 집합($X\cap Y=\varnothing$)이라 가정한다. x를 X 속성에 해당하는

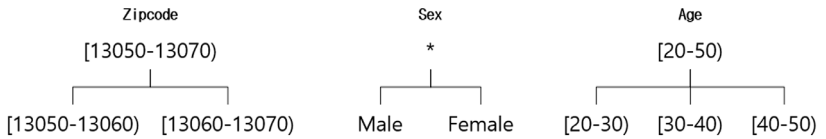

그림 3 주소, 성별, 연령의 범주 트리(Taxonomy tree)
Fig. 3 Taxonomy tree of Zipcode, Sex and Age

값이라 할 때, x값을 가지는 레코드의 Y 속성값의 종류의 수를 $a_{Y(x)}$라 정의한다. $A_Y(X)=\min\{a_Y(x)|x\in X\}$라 할 때, 테이블 T는 자연수 k에 대하여 $A_Y(X)\geq k$일 때, (X,Y)-익명성을 만족한다. □

4.1.2 복수 테이블에서 신원 노출을 방어하는 프라이버시 모델

k-익명성 모델과 (X,Y)-익명성 모델은 단일 테이블을 익명화하는데 초점을 두었다. 그러나 현실에서 데이터베이스는 여러 개의 관계형 테이블로 구성되어 있는 경우도 있다. 이런 경우 여러 개의 관계형 테이블을 조인하면 신원 노출이 발생할 수 있다.

예제 2. 복수 테이블에서 프라이버시 유출

예를 들어, 공격자가 George가 이번 학기에 History와 Religion 과목을 들었고 History 과목에서 American History 책을 쓰고 있다는 사실을 알고 있다고 가정하자. 만약 표 5~7의 테이블이 그대로 배포된다면 테이블에 이름이 없어도 표 6에서 George의 SID가 S3이며

표 5 학생 학점 테이블
Table 5 Student GPA table

SID	GPA
S1	3.72
S2	2.34
S3	4.00

표 6 각 학생이 들은 과목과 성적 테이블
Table 6 Course and grade table

SCID	SID	Course	Grade
SC1	S1	Math	93
SC2	S1	Physics	91
SC3	S1	History	85
SC4	S2	CS	78
SC5	S2	Physics	62
SC6	S2	Religion	42
SC7	S3	History	98
SC8	S3	Religion	96

표 7 과목별 수업용 책과 가격 테이블
Table 7 Course books and price table

SCID	Book	Price
SC1	Discrete	$63
SC2	Calculus	$89
SC3	Religion History	$33
SC4	Discrete	$65
SC5	Dynamics	$51
SC6	Buddhism	$38
SC7	American History	$54
SC8	Religion History	$39

각 과목에서 받은 성적을 알 수 있게 된다. 뿐만 아니라 표 5와 표 7에서 각각 SID와 SCID로 조인하여 George의 GPA와 각 과목에서 쓰는 책을 사는데 얼마를 썼는지도 알 수 있게 된다. □

여러 개의 테이블이 조인된 테이블에서 신원 노출을 막기 위해 Nergiz 등은 다중 테이블 k-익명성(MultiR k-anonymity) 모델[15]을 제안하였다.

정의 4. 다중 테이블 k-익명성

특정한 개인을 식별할 수 있는 기본 키(primary key)와 민감한 정보로 이루어진 PT 테이블이 있고, PT와 조인할 수 있는 테이블 $T_1,\cdots,T_n$이 있을 때, 모든 테이블을 자연 조인(natural join)한 테이블 $JT=PT\bowtie T_1\bowtie\cdots\bowtie T_n$에 임의의 레코드와 서로 구별되지 않는 k-1개의 레코드가 있다면 다중 테이블 k-익명성을 만족한다. □

4.2 속성 노출(attribute disclosure)을 막기 위한 프라이버시 모델

속성 노출은 알 수 없었던 특정한 개인의 민감한 속성값이 드러나는 프라이버시 유출 공격이다. 속성 노출 공격은 공격자가 대상의 준식별자 속성값을 모르는 경우에도 발생할 수 있다. 예를 들어, 공격자가 표 3을 참조했을 때 표 3이 2-익명성을 만족해도 우편번호와 연령, 성별이 {[13060-13570], [40-50], Male}인 사람을 모두 MERS에 걸렸다고 확신할 수 있다. 만약 외부의 테이블이 동일한 준식별자 속성을 가지고 있다면, 두 테이블을 연결하여 프라이버시 유출이 발생할 수 있다. 이러한 문제는 동질 클래스 내의 민감한 속성값의 다양성이 충분하지 않다는 점에 기인한다.

4.2.1 민감한 속성값의 다양성(diversity)을 고려하는 프라이버시 모델

속성 노출을 막기 위한 대표적인 프라이버시 모델은 Machanavajjhala 등이 제안한 l-다양성(l-diversity) 모델[16]이다. l-다양성 모델은 각각의 동질 클래스가 적어도 l개의 서로 다른 민감한 속성값을 가지게 하여 속성 노출을 방어한다.

정의 5. l-다양성

모든 동질 클래스에 서로 다른 l개의 민감한 속성값이 있을 때, 테이블 T는 l-다양성을 만족한다. □

l-다양성 모델이 효과적으로 속성 노출을 방어하지만 한 가지 문제점이 있다. l-다양성 모델은 모든 민감한 속성값의 민감한 정도를 동일하게 취급한다는 점이다. 예를 들어, 같은 병명이라도 감기와 MERS의 프라이버시 위협의 정도는 다르게 받아들여질 수 있지만 l-다양성 모델에서는 이를 반영하지 못한다. 이러한 문제를 해결하기 위해 Wong 등은 (α,k)-익명성((α,k)-anonymity) 모델[17]을 제안하였다. (α,k)-익명성 모델은 k-

익명성 모델에 α-연결조건(α-association criteria)을 추가한 모델이다. α-연결조건은 각각의 동질 클래스 내에서 특정 민감한 속성값의 빈도가 $\alpha(0 \leq \alpha \leq 1)$를 넘지 않게 하는 제약 조건이다. α-연결조건을 통해 배포자가 민감한 속성값에 따라 프라이버시 보호 정도를 다르게 설정할 수 있다.

정의 6. (α,k)-익명성

동질 클래스를 EC라 하자. 민감한 속성값 s에 대하여 (EC,s)를 EC안에 민감한 속성값으로 s를 가지는 레코드의 집합이라 한다. 이 때 α-연결조건은 $\frac{|(EC,s)|}{|EC|} \leq \alpha$ 으로 정의한다. 모든 동질 클래스에서 k-익명성과 α-연결조건을 만족할 때, (α,k)-익명성을 만족한다. □

그러나 민감한 속성값이 숫자 값일 때(예 : 급여)는 l-다양성 혹은 (α,k)-익명성 모델을 적용해도 속성 노출이 발생할 수 있다. 예를 들어, 어떤 동질 클래스의 월급 속성값이 좁은 범위 안에 몰려있다면, 특정한 개인의 월급을 높은 확률로 유추할 수 있다.

예제 3. 민감한 속성값이 숫자 값일 때의 속성 노출

표 8의 회사원의 급여 테이블이 있다고 하자. 회사에서 표 8을 3-다양성을 만족하게 익명화하여 표 9처럼 3개의 동질 클래스로 이루어진 테이블을 배포하였다. 만약 공격자가 Andy의 우편번호와 나이를 알고 있을 때 표 9를 참조하여 공격자는 Andy가 첫 번째 동질 클래스에 속한다는 것을 알 수 있다. 그러나 급여 속성값이 좁은 범위([$3,400-$3,600]) 내에 몰려 있어 공격자는 Andy의 급여를 높은 확률로 유추할 수 있다. □

표 8 원본 급여 테이블
Table 8 Original salary table

Name	Zipcode	Age	Salary
Andy	37258	55	$3,600
Brown	37231	35	$3,500
Carol	37244	40	$3,400
Dennis	37222	50	$7,000
Edwin	37216	47	$7,300
Fiona	37214	36	$7,700
Galen	37266	45	$11,000
Harry	37220	42	$15,000
Isaac	37213	33	$20,000

표 9 3-다양성을 만족하는 표 8
Table 9 3-diversity Table 8

Name	Zipcode	Age	Salary
Andy	372**	[35-55]	$3,600
Brown	372**	[35-55]	$3,500
Carol	372**	[35-55]	$3,400
Dennis	372**	[36-50]	$7,000
Edwin	372**	[36-50]	$7,300
Fiona	372**	[36-50]	$7,700
Galen	372**	[33-45]	$11,000
Harry	372**	[33-45]	$15,000
Isaac	372**	[33-45]	$20,000

예제 3의 프라이버시 유출 원인은 l-다양성 모델이 민감한 속성값 간의 거리를 고려하지 않는 모델이기 때문이다. 이러한 문제를 해결하기 위해 Zhang 등은 (k,e)-익명성((k,e)-anonymity) 모델[18]을 제안하였다.

정의 7. (k,e)-익명성

모든 동질 클래스에서 민감한 속성값의 종류가 k개 이상이고, 민감한 속성값의 범위(최댓값 - 최솟값)가 최소 e이상일 때, (k,e)-익명성을 만족한다. □

예제 4. (k,e)-익명성을 통한 속성 노출 방어

표 8을 익명화하여 표 10처럼 (4, 7,600)-익명성을 만족하게 익명화할 수 있다. 각 동질 클래스는 4개의 다른 민감한 속성값을 가지고 있으며, 범위는 최소 7,600 이상이므로 (4, 7,600)-익명성을 만족한다. 예제 3과 동일한 공격자가 표 10을 참조했을 때, Andy의 급여값은 [$3,400-$11,000] 범위 안에 있게 되므로 [$3,400-$3,600]에 비해 정확한 급여값을 유추할 확률이 감소한다. □

(k,e)-익명성 모델은 민감한 속성값이 숫자 값일 때, 민감한 속성값의 범위에 대한 조건을 추가하여 효과적으로 속성 노출을 방어한다. 그러나 (k,e)-익명성 모델은 숫자 값의 범위만 고려할 뿐, 숫자 값의 분포를 고려하지 않는다. 표 10의 준식별자 속성값이 {Zipcode, Age} = {372**, [35-55]}인 동질 클래스에서 Salary 속성값의 종류는 $3,400, $3,500, $3,600, $11,000이다. 공격자가 특정한 개인이 이 동질 클래스에 속한다는 것을 알 때, 공격자는 여전히 높은 확률(75%)로 대상의 Salary가 [$3,400 - $3,600] 범위 안에 있다는 것을 유추할 수 있다. 이러한 프라이버시 유출은 (k,e)-익명성 모델이 민감한 속성값의 범위 내에 속하는 민감한 속성값의 분포를 고려하지 않기 때문에 발생한다. Li 등은 민감한 속성값의 범위 내에 포함된 값의 분포도 제한하는

표 10 (4, 7,600)-익명성을 만족하는 표 8
Table 10 (4, 7,600)-anonymity Table 8

Name	Zipcode	Age	Salary
Andy	372**	[35-55]	$3,600
Brown	372**	[35-55]	$3,500
Carol	372**	[35-55]	$3,400
Galen	372**	[35-55]	$11,000
Edwin	372**	[33-47]	$7,300
Fiona	372**	[33-47]	$7,700
Dennis	372**	[33-47]	$7,000
Harry	372**	[33-47]	$15,000
Isaac	372**	[33-47]	$20,000

(ϵ,m)-익명성((ϵ,m)-anonymity) 모델[19]을 제안하였다.

정의 8. (ϵ,m)-익명성

테이블 T에서 각 레코드 t에 대해, t의 민감한 속성값 $t.S$가 실수 ϵ에 대하여, 범위 $I(t) = [t.S-\epsilon, t.S+\epsilon]$를 가지고 있다고 하자. 모든 동질 클래스 $\mathbb{EC}=\{EC_1,\cdots,EC_m\}$에서, 각 레코드 t가 자연수 m에 대하여 $\frac{x}{|EC_i|}\leq\frac{1}{m}$을 만족할 때, (ϵ,m)-익명성을 만족한다. x는 동질 클래스 EC_i 안에서 레코드 t의 민감한 속성값의 범위 $I(t)$ 내의 값을 민감한 속성값으로 가지는 레코드의 수를 의미한다. □

예제 5. (ϵ,m)-익명성을 통한 속성 노출 방어

표 8을 익명화하여 표 11처럼 (3000, 3)-익명성을 만족하게 익명화할 수 있다. 마찬가지로 Andy의 준식별자 속성값을 알고 있는 공격자가 표 11을 참조했을 때, Andy가 세 번째 동질 클래스에 속한다는 것을 알 수 있다. 세 번째 동질 클래스의 급여값은 $3,600, $7,700, $20,000이며 각각의 값의 차이가 $3,000보다 크다. 따라서 공격자는 $6,000 범위 이내로 Andy의 급여값을 유추할 확률은 최대 1/3이 된다. □

표 11 (3000, 3)-익명성을 만족하는 표 8
Table 11 (3000, 3)-anonymity Table 8

Name	Zipcode	Age	Salary
Carol	372**	[40-50]	$3,400
Dennis	372**	[40-50]	$7,000
Galen	372**	[40-50]	$11,000
Brown	372**	[35-47]	$3,500
Edwin	372**	[35-47]	$7,300
Harry	372**	[35-47]	$15,000
Andy	372**	[33-55]	$3,600
Fiona	372**	[33-55]	$7,700
Isaac	372**	[33-55]	$20,000

4.2.2 민감한 속성값의 분포와 의미(semantics)까지 고려하는 프라이버시 모델

Li 등은 l-다양성 모델을 만족하는 테이블이 동질성 공격(similarity attack)에 약하다는 것을 발견하였다. 동질성 공격은 동질 클래스를 구성하는 레코드의 민감한 속성값이 서로 비슷한 의미를 가질 때 발생한다. 예를 들어, 3-다양성을 만족하는 동질 클래스의 민감한 속성값이 {폐암, 폐렴, 기관지염}이라 가정하자. 공격자가 대상이 이 동질 클래스에 속한다는 것을 알 수 있다면 대상의 정확한 병명은 모르지만 폐와 관련된 질병을 가지고 있다고 확신할 수 있다. 동질성 공격은 민감한 속성값의 의미(semantics)가 서로 비슷하기 때문에 발생한다.

이러한 l-다양성 모델의 취약점을 보완하고자 Li 등은 t-근접성(t-closeness) 모델[20]을 제안하였다. t-근접성 모델은 *Earth Mover's Distance* (*EMD*)[21]를 사용하여 원본 테이블과 동질 클래스의 민감한 속성값의 분포가 얼마나 가까운지(closeness)를 계산하며, *EMD*가 $t(0\leq t\leq 1)$ 이하일 때 만족하는 프라이버시 모델이다.

정의 9. t-근접성

원본 테이블을 T, 익명화한 테이블의 어떤 동질 클래스를 *EC*라 하자. 모든 동질 클래스에서 $EMD(T,EC)\leq t$ 일 때, t-근접성을 만족한다. □

t-근접성 모델은 모든 동질 클래스에서 *EMD*를 t 이하로 강제하기 때문에 익명화 과정에서 준식별자와 민감한 속성간의 관계가 크게 손상될 가능성이 높다.

예제 6. t-근접성을 만족시키기 위한 익명화 과정에서 정보 손실

표 12는 각 환자의 준식별자인 우편번호, 연령과 민감한 속성인 병명과 각 레코드의 수를 나타내는 속성으로 구성되어 있다. 만약 표 12를 익명화하여 0.1-근접성을 만족시키기 위해서는 하나의 동질 클래스로 익명화되어야 한다.

표 13처럼 하나의 동질 클래스로 익명화해야 원본 테이블과 동질 클래스 간의 *EMD*가 0.1 이하가 될 수 있다. 하나의 동질 클래스를 구성하기 위해 우편번호와 연령이 모든 값을 포함하게끔 일반화되었다. 이에 따라 준식별자와 민감한 속성간의 관계가 크게 약화되어 유용성이 감소한다. □

표 12 원본 환자 테이블
Table 12 Original patients table

ZIP	Age	Disease	Count
47673	29	Cancer	100
47674	21	Flu	100
47605	25	Cancer	200
47602	23	Flu	200
47905	43	Cancer	100
47904	48	Flu	900
47906	47	Cancer	100
47907	41	Flu	900
47603	34	Cancer	100
47605	30	Flu	100
47602	36	Cancer	100
47607	32	Flu	100

표 13 0.1-근접성을 만족하는 표 12
Table 13 0.1-closeness Table 12

ZIP	Age	Disease	Count
47***	[21-48]	Cancer	700
47***	[21-48]	Flu	2300

이를 해결하고자 Li 등은 t-근접성 모델의 제약 사항을 완화한 (n,t)-근접성((n,t)-closeness) 모델[22]을 제안하였다.

정의 10. (n,t)-근접성

익명화된 테이블 T의 어떤 동질 클래스 EC에 대하여 EC를 포함하는 동질 클래스 $G(EC \subseteq G)$가 적어도 n개의 레코드를 가지고 있으며, 두 동질 클래스 간의 $EMD(EC, G) \leq t$일 때, (n,t)-근접성을 만족한다. □

(n,t)-근접성 모델은 민감한 속성값의 분포를 비교하는 대상을 원본 테이블이 아닌 더 큰 집합의 동질 클래스로 설정하면서 적어도 n개의 레코드가 있는 동질 클래스로부터 민감한 속성값에 대한 정보를 얻는 것을 허용하는 모델이다.

표 14는 (1000, 0.1)-근접성을 만족하게 표 12를 익명화한 테이블이다. EC_2는 레코드의 개수가 2,000개 이므로 더 큰 동질 클래스를 자기 자신으로 설정하면 (1,000, 0.1)-근접성을 만족한다. EC_1과 EC_3을 포함하는 동질 클래스 G를 {ZIP, Age} = {476**, [20,39]}라 가정하면 G는 1,000개의 레코드를 가지며 EC_1과 EC_3와 민감한 속성값의 분포가 {Cancer, Flu} = {0.5, 0.5}로 동일하므로 (1,000, 0.1)-근접성을 만족한다. t-근접성 모델을 위한 익명화 알고리즘으로는 $SABRE$[23]이 있다.

표 14 (1,000, 0.1)-근접성을 만족하는 표 12
Table 14 (1,000, 0.1)-closeness Table 12

	ZIP	Age	Disease	Count
EC_1	476**	2*	Cancer	300
	476**	2*	Flu	300
EC_2	479**	4*	Cancer	200
	479**	4*	Flu	1800
EC_3	476**	3*	Cancer	200
	476**	3*	Flu	200

4.3 귀속 노출(membership disclosure)을 막기 위한 프라이버시 모델

귀속 노출은 테이블에 개인의 포함 여부가 드러나는 프라이버시 유출 공격이다. 예를 들어, 익명화된 AIDS 환자 테이블을 배포했을 때, 공격자가 특정한 개인이 이 테이블에 존재하는 사실을 안다면 특정한 개인이 AIDS 환자임이 드러나 프라이버시가 유출된다. 다음의 예시는 귀속 노출 공격을 보여준다.

예제 7. 귀속 노출 공격

배포자가 원본 테이블 표 15를 가지고 있다. 배포자는 표 15를 4-익명성을 만족하게 익명화한 표 16을 배포하였으며 또한 표 17을 바탕으로 MERS 환자 테이블 표 17을 배포하였다. 공격자는 표 17을 가지고 있으며, Alice

표 15 원본 모집단 테이블
Table 15 Original population table

Name	Zip	Age	Sex	MERS
Alice	21842	51	Male	O
Bob	21884	30	Female	X
Catrine	21858	52	Male	O
Duggan	21847	54	Female	O
Frank	21879	36	Male	X
Gary	22861	38	Female	X
Harry	22845	57	Male	O
Iria	22867	30	Female	X
James	22874	53	Male	O

표 16 4-익명성을 만족하는 표 15
Table 16 4-anonymity Table 15

Name	Zip	Age	Sex	MERS
Alice	2*	>=50	*	O
Catrine	2*	>=50	*	O
Duggan	2*	>=50	*	O
Harry	2*	>=50	*	O
James	2*	>=50	*	O
Bob	2*	<50	*	X
Frank	2*	<50	*	X
Gary	2*	<50	*	X
Iria	2*	<50	*	X

표 17 표 16의 부분 연구 테이블
Table 17 Research subset of Table 16

Name	Zip	Age	Sex	MERS
Alice	2*	>=50	*	O
Catrine	2*	>=50	*	O
Duggan	2*	>=50	*	O
Harry	2*	>=50	*	O
James	2*	>=50	*	O

의 준식별자 속성값 {Zip, Age, Sex} = {21842, 51, Male}을 알고 있다. 공격자가 표 17을 참조했을 때 모든 레코드가 Alice에 해당될 수 있으므로 공격자는 Alice가 MERS 환자임을 확신할 수 있다. □

이러한 귀속 노출을 막기 위한 대표적인 프라이버시 모델은 δ-존재성(δ-presence) 모델[24]이다. δ-존재성 모델은 익명화한 테이블에서 특정한 개인의 레코드가 존재한다고 유추할 확률을 $(\delta_{\min}, \delta_{\max})$ 사이의 값으로 묶는 모델이다.

정의 11. δ-존재성

원본 테이블을 E라 하고 E에서 배포용으로 사용할 레코드가 있는 테이블 T와 T를 익명화한 테이블 V가 있다고 하자. T에 존재하는 레코드 t에 대하여 t가 V에 존재할 확률 $\Pr(t \subseteq T|V)$이 $\delta_{\min} \leq \Pr(t \subseteq T|V)$

$\leq \delta_{max}$을 만족할 때, $(\delta_{min},\delta_{max})$-존재성을 만족한다. □

4.4 확률 공격(probabilistic attack)을 막기 위한 프라이버시 모델

확률 공격은 공격자가 테이블을 참조한 후 특정한 개인의 민감한 속성값의 유추 확률을 증가시키려는 프라이버시 유출 공격이다. 확률 공격은 익명화된 테이블을 보기 전과 후에 특정한 개인의 민감한 속성에 대한 확률적 믿음(probabilistic belief)의 차이가 커질 때 발생한다[16]. 공격자는 확률 공격을 통해 익명화된 테이블을 참조하여 배경지식 이상의 정보를 얻는 것이 목표이다. 따라서 확률 공격을 막기 위한 프라이버시 모델은 익명화된 테이블을 참조한 후에 공격자의 배경지식의 증가를 최소화하는 것을 목표로 한다.

Chawla 등은 (c,t)-고립성((c,t)-isolation) 모델[25]을 제안하였다. (c,t)-고립성 모델에서 레코드는 각 준식별자 속성을 좌표축으로 하는 좌표공간상의 점으로 정의한다. 공격자는 좌표공간에서 특정한 하나의 점을 찾아내려는 *isolator*로 정의한다. (c,t)-고립성 모델은 이러한 *isolator*가 유일한 점을 식별하지 못하게 하는 모델이다.

정의 12. (c,t)-고립성

각 준식별자 속성을 좌표축으로 하는 좌표공간이 있으며 각 레코드는 좌표공간에서 한 점을 의미한다고 가정한다. *isolator*가 식별하려는 특정한 개인의 좌표공간상의 점을 p라고 하자. q를 공격자가 배경지식을 통해 유추한 공격하려는 대상의 좌표공간상의 점이라고 하자. 이 때, $\delta_p = \|q-p\|$는 p와 q간의 거리이다. $B(q,c\delta_p)$를 q를 중심으로 하고 반지름이 $c\delta_p$인 좌표공간상의 구라고 하자. 구 안의 점의 개수가 t개 이하일 때, 점 q는 점 p를 (c,t)-고립한다고 정의한다. □

(c,t)-고립성 모델은 단순히 공격자가 배경지식으로부터 t개 이하의 레코드를 식별하지 못하게 할 뿐, 확률 공격으로부터 프라이버시 보호 수준을 정량적으로 제시하지 못하였다. 이에 Rastogi 등은 프라이버시 보호 수준을 확률적으로 정의한 (d,γ)-프라이버시((d,γ)-privacy) 모델[26]을 제안하였다. 원본 테이블 T에서 어떤 한 레코드를 t라 할 때, $\Pr(t)$를 공격자가 t가 T에 존재한다고 유추할 확률이라 하고 $\Pr(t|V)$를 익명화된 테이블 V를 참조한 후에 t가 T에 존재한다고 유추할 확률이라 하자. (d,γ)-프라이버시 모델에서 공격자가 가지고 있는 배경지식은 모든 레코드 t에 대해, $\Pr(t) \leq d (0 < d < 1)$ 또는 $\Pr(t)=1$이며 확률은 독립적이라 가정한다. $\Pr(t)=1$은 공격자가 이미 레코드 t가 원본 테이블에 존재한다고 확신하고 있다는 것을 의미한다.

정의 13. (d,γ)-프라이버시

모든 레코드 t에 대해 공격자의 배경지식 중 $\Pr(t)=1$인 레코드를 제외한 모든 레코드가 $\Pr(t) \leq d$라고 가정한다. 이러한 레코드 t에 대하여 공격자가 익명화된 테이블 V를 참조하고 나서 $\frac{d}{\gamma} \leq \frac{\Pr(t|V)}{\Pr(t)}$과 $\Pr(t|V) \leq \gamma$ $(0 \leq \gamma \leq 1)$를 만족하면 V는 (d,γ)-프라이버시를 만족한다. □

(d,γ)-프라이버시 모델은 공격자가 익명화된 테이블을 참조한 후의 배경지식의 차이를 효과적으로 제한하는 모델이다. 그러나 가정하는 공격자 모델이 현실적으로 성립하기 어렵다는 문제가 있다. 이를 해결하고자 Li 등은 (B,t)-프라이버시((B,t)-privacy) 모델[27]을 제안하였다. (d,γ)-프라이버시 모델이 공격자가 가질 수 있는 모든 배경지식에 대해 익명화된 테이블을 참조하여 발생하는 확률 공격을 제한하는 것과 다르게 (B,t)-프라이버시 모델은 오직 배포하려는 익명화된 테이블과 관련된 배경지식을 가지고 있는 공격자의 확률 공격을 제한하는 것이 차이점이다. 따라서 (d,γ)-프라이버시 모델에 비해 가정하는 공격자 모델이 비교적 현실적이다. (B,t)-프라이버시 모델은 커널 추정(kernel estimation)[28]의 평활모수(bandwidth)를 공격자의 배경지식 B로 정의하였다. B값이 클수록 공격자가 공격하려는 대상의 준식별자와 민감한 속성 사이의 관계에 대한 배경지식이 적다는 것을 의미한다.

예제 8. 베이지안(Bayesian) 추론 기법을 이용한 확률 공격

표 18처럼 준식별자가 동일한 3개의 레코드로 구성된 동질 클래스가 있다고 하자. 표 19의 배경지식을 가지고 공격자는 표 20처럼 3가지 경우에 대해 t_3가 MERS일 확률을 베이지안 추론 기법으로 계산할 수 있다.

$$P(Case\ 1) = P(none|t_1) \times P(none|t_2) \times P(MERS|t_3)$$
$$= 0.95 \times 0.95 \times 0.3 = 0.271$$
$$P(Case\ 2) = P(none|t_1) \times P(MERS|t_2) \times P(none|t_3)$$
$$= 0.95 \times 0.05 \times 0.7 = 0.033$$

표 18 3개의 레코드로 이루어진 동질 클래스
Table 18 Equivalence class with three records

Record	Disease
t_1	none
t_2	none
t_3	MERS

표 19 공격자의 배경지식 테이블
Table 19 The adversary's prior belief table

t_1	t_2	t_3
$P(MERS\|t_1)=0.05$	$P(MERS\|t_2)=0.05$	$P(MERS\|t_3)=0.3$
$P(none\|t_1)=0.95$	$P(none\|t_2)=0.95$	$P(none\|t_3)=0.7$

표 20 공격자가 공격 가능한 3가지 경우
Table 20 Three possible cases of probability attack

	t_1	t_2	t_3
Case 1	none	none	MERS
Case 2	none	MERS	none
Case 3	MERS	none	none

$$P(Case\ 3) = P(MERS|t_1) \times P(none|t_2) \times P(none|t_3)$$
$$= 0.95 \times 0.05 \times 0.7 = 0.033$$

따라서 t_3가 MERS를 가질 확률은

$$P(Case\ 1) = \frac{0.271}{0.271 + 0.033 + 0.033} = 0.8$$

즉, 공격자가 t_3가 MERS를 가지고 있다고 유추할 확률이 0.3에서 0.8로 증가하게 된다. □

모든 공격자의 배경지식을 고려하여 민감한 속성값을 유추할 확률을 계산하는 것은 시간 복잡도가 지수 형태이다. 따라서 레코드가 많아질 경우 계산하는데 매우 많은 시간이 걸린다. 이를 해결하기 위해 Li 등은 익명화된 테이블을 참조한 후에 민감한 속성값을 유추할 확률 분포를 계산하는 Ω-추정(Ω-estimate)이라는 근사 기법을 제안하였다. Ω-추정은 다음과 같이 계산한다.

$$\Omega(s_i|t_j) = \frac{P(s_i|t_j)}{\sum_{i=1}^{k} P(s_i|t_i)}$$

$P(s_i|t_j)$는 동질 클래스 안의 레코드 t_i가 민감한 속성값으로 s_i를 가질 확률을 의미한다. 이를 이용하여 (B,t)-프라이버시 모델을 다음과 같이 정의한다.

정의 14. (B,t)-프라이버시

공격자가 가지고 있는 배경지식을 커널 추정의 평활 모수 B라 하고, 공격하려는 대상의 레코드 r의 준식별자 값을 q라 할 때, 공격자가 배경지식으로 가지고 있는 r의 민감한 속성값을 유추할 확률을 $P_{pri}(B,q)$라고 하자. 공격자가 익명화된 테이블 V를 참조하고 난 후의 r의 민감한 속성값을 유추할 확률을 $P_{pos}(B,q,V)$라 한다. 이 때, 준식별자 속성값으로 q를 가지는 모든 레코드에 대해 다음의 식을 만족하면 (B,t)-프라이버시를 만족한다.

$$\max \mathrm{D}[P_{pri}(B,q), P_{pos}(B,q,V)] \leq t$$

두 확률 분포 간의 거리 $\mathrm{D}[P,Q]$는 젠슨-샤논 발산(Jensen-Shannon divergence)[29]로 계산한다. □

(B,t)-프라이버시 모델은 공격자가 익명화된 테이블을 참조한 후의 배경지식의 차이를 t값으로 제한하지만, 그 차이가 상대적이 아닌 절대적이라는 문제가 있다. 예를 들어, 어떤 테이블의 민감한 속성값이 Flu와 HIV로만 이루어져 있고 민감한 속성값의 분포가 (Flu, HIV) = (0.99, 0.01)라고 가정하자. 만약 공격자가 익명화된 테이블을 참조하고 공격하려는 대상이 있는 동질 클래스의 민감한 속성값의 분포가 (Flu, HIV) = (0.97, 0.03)라는 것을 알았을 때, (B,t)-프라이버시 모델에서 두 분포간의 거리는 0.0133이다. 그러나 이는 공격하려는 대상이 HIV에 걸렸다고 유추할 확률이 상대적으로 200% 증가(0.01 -> 0.03)했다는 것을 반영하지 못한다. 확률 분포간의 상대적인 차이를 반영하지 못하는 문제점을 해결하기 위해 Cao 등은 β-가능도(β-likeliness) 모델[30]을 제안하였다.

정의 15. β-가능도

원본 테이블의 민감한 속성값이 $V=(v_1,\cdots,v_m)$이며, 민감한 속성값의 분포가 $P=(p_1,\cdots,p_m)$이다. 익명화된 테이블의 어떤 동질 클래스의 민감한 속성값의 분포가 $Q=(q_1,\cdots,q_m)$라 하자. 모든 동질 클래스에서 다음의 식을 만족하면 β-가능도를 만족한다.

$$\max[\mathrm{D}(p_i,q_i)|p_i \in P, p_i < q_i] \leq \beta \ (\beta > 0)$$

두 분포간의 거리는 $\mathrm{D}(p_i,q_i) = \frac{q_i - p_i}{p_i}$로 계산한다. □

5. 결 론

본 논문에서는 공격자의 다양한 배경지식으로부터 발생할 수 있는 프라이버시 유출 공격의 유형과 이를 막기 위한 프라이버시 모델들에 대해 논의하였다. 또한 프라이버시 모델들 간의 차이점(표 21)과 요구 조건에 대해 알아보았다.

최근 우리나라 정부에서는 각 기업들이 개인 정보가 포함된 데이터를 안전하게 사용하기 위한 개인 정보 비식별 조치 가이드라인을 발간하였다. 이 가이드라인에는 k-익명성, l-다양성, t-근접성 모델들을 설명하면서 데이터를 어떻게 익명화하여 사용할지를 안내하고 있다. 또한 일본에서도 데이터의 활용 확대를 위해 비식별 가이드라인 발간을 준비하고 있다. 이 외에도 미국이나 EU에서는 프라이버시 보호 법률을 제정하여 익명화된 데이터는 연구나 통계 목적으로 사용할 수 있음을 명시하고 있다.

이렇듯 각국에서 개인의 프라이버시를 보호하기 위해 활발한 움직임을 보이고 있다. 그러나 현재까지 진행된 프라이버시 보호 연구는 이론적인 측면에 집중하여 현실의 요구를 반영하지 못하고 있다. 이러한 문제를 해결하고자 UTD에서는 Incognito와 Mondrian 알고리즘을 구현하여 데이터를 익명화할 수 있는 UTD 익명화 툴[32]을 제공하고 있다. 또한 Flash 알고리즘[31]을 바탕으로 하는 ARX 익명화 툴[33]도 공개되어 있다. 그러나 이러한 익명화 소프트웨어는 주로 연구용으로 쓰이

표 21 프라이버시 모델 요약
Table 21 Summary of privacy models

Privacy models	Privacy attacks			
	Identity disclosure	Attribute disclosure	Membership disclosure	Probabilistic attack
k-anonymity	○			
MultiR k-anonymity	○			
(X, Y)-anonymity	○	○		
l-diversity	○	○		
(α, k)-anonymity	○	○		
(k, e)-anonymity		○		
(ϵ, m)-anonymity		○		
t-closeness		○		
(n, t)-closeness		○		
δ-presence			○	
(c, t)-isolation	○			○
(d, γ)-privacy			○	○
(B, t)-privacy		○		○
β-likeliness		○		○

고 있으며 현실에서 프라이버시 보호 데이터 배포는 전무한 상황이다. 따라서 본 논문을 통해 연구자들에게는 지금까지 연구된 모델을 정리하여 새로운 연구 방향을 모색하는데 도움이 되는 한편 현실의 요구를 반영한 프라이버시 보호 연구도 필요하다.

References

[1] Narayanan A, Shmatikov V, "Robust de-anonymization of large datasets (how to break anonymity of the Netflix prize dataset), 2008," *University of Texas at Austin*, 2008.

[2] Sweeney L, "k-anonymity: A model for protecting privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 10, No. 3, pp. 557-570, 2002.

[3] Gehrke J, "Models and methods for privacy preserving data publishing and analysis," *Proc. of the 22nd International Conference on Data Engineering (ICDE)*, Vol. 105, 2006.

[4] Fung B, Wang K, Chen R, Yu PS, "Privacy-preserving data publishing: A survey of recent developments," *ACM Computing Surveys (CSUR)*, Vol. 42, No. 4, pp. 14, 2010.

[5] Xu Y, Ma T, Tang M, Tian W, "A survey of privacy preserving data publishing using generalization and suppression," *Applied Mathematics & Information Sciences*, Vol. 8, No. 3, pp. 1103, 2014.

[6] Gkoulalas-Divanis A, Loukides G, Sun J, "Publishing data from electronic health records while preserving privacy: a survey of algorithms," *Journal of biomedical informatics*, Vol. 50, pp. 4-19, 2014.

[7] LeFevre K, DeWitt DJ, Ramakrishnan R, "Incognito: Efficient full-domain k-anonymity," *Proc. of the 2005 ACM SIGMOD international conference on Management of data*, pp. 49-60, 2005.

[8] LeFevre K, DeWitt DJ, Ramakrishnan R, "Mondrian multidimensional k-anonymity," *22nd International Conference on Data Engineering (ICDE'06)*, pp. 25-25, 2006.

[9] Xiao X, Tao Y, "Anatomy: Simple and effective privacy preservation," *Proc. of the 32nd international conference on Very large data bases*, pp. 139-150, 2006.

[10] Li T, Li N, Zhang J, Molloy I, "Slicing: A new approach for privacy preserving data publishing," *Knowledge and Data Engineering, IEEE Transactions on*, Vol. 24, No. 3, pp. 561-74, 2012.

[11] Terrovitis M, Mamoulis N, Liagouris J, Skiadopoulos S, "Privacy preservation by disassociation," *Proc. of the VLDB Endowment*, Vol. 5, No. 10, pp. 944-955, 2012.

[12] Dwork C, "Differential privacy: A survey of results," *International Conference on Theory and Applications of Models of Computation*, pp. 1-19, 2008.

[13] McSherry F, Talwar K, "Mechanism design via differential privacy," *Foundations of Computer Science, 2007 FOCS'07 48th Annual IEEE Symposium on*, pp. 94-103, 2007.

[14] Wang K, Fung B, "Anonymizing sequential releases," *Proc. of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 414-423, 2006.

[15] Nergiz ME, Clifton C, Nergiz AE, "Multirelational k-anonymity," *Knowledge and Data Engineering, IEEE Transactions on*, Vol. 21, No. 8, pp. 1104-1117, 2009.

[16] Machanavajjhala A, Kifer D, Gehrke J, Venkitasub-

ramaniam M, "l-diversity: Privacy beyond k-anonymity," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, Vol. 1, No. 1, pp. 3, 2007.
[17] Wong RC-W, Li J, Fu AW-C, Wang K, "(α, k)-anonymity: an enhanced k-anonymity model for privacy preserving data publishing," *Proc. of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 754-759, 2006.
[18] Zhang Q, Koudas N, Srivastava D, Yu T, "Aggregate query answering on anonymized tables," *Data Engineering, 2007 ICDE 2007 IEEE 23rd International Conference on*, pp. 116-125, 2007.
[19] Li J, Tao Y, Xiao X, "Preservation of proximity privacy in publishing numerical sensitive data," *Proc. of the 2008 ACM SIGMOD international conference on Management of data*, pp. 473-486, 2008.
[20] Li N, Li T, Venkatasubramanian S, "t-closeness: Privacy beyond k-anonymity and l-diversity," *Data Engineering, 2007 ICDE 2007 IEEE 23rd International Conference on*, pp. 106-115, 2007.
[21] Rubner Y, Tomasi C, Guibas LJ, "The earth mover's distance as a metric for image retrieval," International journal of computer vision, Vol. 40, No. 2, pp. 99-121, 2000.
[22] Li N, Li T, Venkatasubramanian S, "Closeness: A new privacy measure for data publishing," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 22, No. 7, pp. 943-956, 2010.
[23] Cao J, Karras P, Kalnis P, Tan K-L, "SABRE: a Sensitive Attribute Bucketization and REdistribution framework for t-closeness," *The VLDB Journal*, Vol. 20, No. 1, pp. 59-81, 2011.
[24] Nergiz ME, Atzori M, Clifton C, "Hiding the presence of individuals from shared databases," *Proc. of the 2007 ACM SIGMOD international conference on Management of data*, pp. 665-676, 2007.
[25] Chawla S, Dwork C, McSherry F, Smith A, Wee H. "Toward privacy in public databases," *Theory of Cryptography*, pp. 363-385, 2005.
[26] Rastogi V, Suciu D, Hong S, "The boundary between privacy and utility in data publishing," *Proc. of the 33rd international conference on Very large data bases*, pp. 531-542, 2007.
[27] Li T, Li N, Zhang J, "Modeling and integrating background knowledge in data anonymization," *Data Engineering, 2009 ICDE'09 IEEE 25th International Conference on*, pp. 6-17, 2009.
[28] Friedman J, Hastie T, Tibshirani R, "The elements of statistical learning," Springer series in statistics Springer, 2001.
[29] Lin J, "Divergence measures based on the Shannon entropy," *IEEE Transactions on Information theory*, Vol. 37, No. 1, pp. 145-151, 1991.
[30] Cao J, Karras P, "Publishing microdata with a robust privacy guarantee," *Proc. of the VLDB Endowment*, Vol. 5, No. 11, pp. 1388-1399, 2012.
[31] Kohlmayer F, Prasser F, Eckert C, Kemper A, Kuhn KA, "Flash: efficient, stable and optimal k-anonymity," *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*, pp. 708-717, 2012.
[32] [Online]. Available: http://www.cs.utdallas.edu/dspl/cgi-bin/toolbox/index.php
[33] [Online]. Available: http://arx.deidentifier.org/anonymization-tool/

김 종 선
2016년 고려대학교 컴퓨터학과 졸업(학사). 2016년~현재 고려대학교 컴퓨터학과 석박사 통합과정. 관심분야는 데이터 프라이버시

정 기 정
2016년 고려대학교 컴퓨터학과 졸업(학사). 2016년~현재 고려대학교 컴퓨터학과 석박사 통합과정. 관심분야는 데이터 프라이버시

이 혁 기
2015년 고려대학교 컴퓨터학과 졸업(학사). 2015년~현재 고려대학교 컴퓨터학과 석박사 통합과정. 관심분야는 데이터 프라이버시

김 수 형
2012년 고려대학교 컴퓨터·통신공학부 졸업(학사). 2012년~현재 고려대학교 IT융합학과 석박사 통합과정. 관심분야는 데이터 프라이버시

김 종 욱

2000년 고려대학교 전산과학과(학사) 2002년 한국과학기술원 전산학과(석사) 2009년 Arizona State University, Computer Science(박사). 2010년~2013년 Teradata, Software Engineer. 2013년~현재 상명대학교 미디어소프트웨어학과 조교수. 관심분야는 Database System, Data Mining

정 연 돈

1994년 고려대학교 전산과학과 졸업(학사). 1996년 한국과학기술원 전산학과 졸업(석사). 2000년 한국과학기술원 전산학전공 졸업(박사). 2000년~2003년 한국과학기술원 전산학전공 Post-Doc. 연구원 및 연구교수. 2003년~2006년 동국대학교 컴퓨터공학과 교수. 2006년~현재 고려대학교 컴퓨터학과 교수. 관심분야는 Database Privacy, Spatial Databases, Mobile Databases, Graph Databases, Data-Intensive Systems, Database Systems