차분 프라이버시 히스토그램 공개 알고리즘의 개선

An Improved Differentially Private Histogram Publication Algorithm

구 한 준, 정 우 환, 심 규 석
서울대학교

Hanjun Goo, Woohwan Jung, Kyuseok Shim
Seoul National Univ.

요약

최근 공격자의 사전 지식에 상관없이 개인 정보를 보호할 수 있는 차분 프라이버시 보호 기법에 대한 연구들이 진행되고 있다. 본 논문에서는 차분 프라이버시를 만족시키는 적은 수의 버킷을 가지는 히스토그램 공개 알고리즘을 소개하고 기존 알고리즘이 사용한 휴리스틱 방법의 문제와 개선 방법을 소개한다. 또한, 실험을 통해 개선한 방법이 기존의 알고리즘에 비하여 더 좋은 영역 합 질의 성능을 가지는 것을 보인다.

I. 서론

개인 정보가 담긴 데이터를 공개하는 경우 개개인의 프라이버시는 지키면서 데이터의 특성을 유지하여 공개하는 기법들이 연구되고 있다. 차분 프라이버시 기법은 공격하는 사람의 사전 지식에 상관없이 개인 정보를 확률적으로 보호해 줄 수 있기 때문에[1] 다양한 종류의 데이터 형태에 대해 연구가 진행되고 있다.

히스토그램은 많은 양의 데이터에 대한 통계 정보를 나타내기 위하여 빈번하게 활용되는 데이터로써, 이를 공개하기 위한 차분 프라이버시 기법도 활발히 연구되고 있다. 히스토그램의 버킷을 숫자를 줄여 공개하는 연구로는, 다이나믹 프로그래밍 기법을 활용하여 버킷을 줄이면서 발생하는 오차의 제곱 합을 최소화하는 히스토그램을 찾는 연구가 있다[5]. 그리고 [3]에서는 이를 활용하는 차분 프라이버시를 만족하는 히스토그램 공개 알고리즘을 제안하였다. 하지만 [3]에서는 원본 히스토그램에 노이즈를 더하고 이를 [5]의 연구를 활용하여 적은 버킷의 히스토그램으로 바꾸는데, 원본 히스토그램의 버킷 숫자가 클수록 노이즈가 많이 삽입되는 단점이 있다. 이를 2단계의 과정으로 나누어서 히스토

그램의 구간을 찾는 부분과 히스토그램의 도수를 찾는 부분으로 나누어 원본 히스토그램의 버킷 수와 상관없이 노이즈 삽입을 줄일 수 있는 연구[4]가 제안되었지만, 히스토그램의 구간을 찾는 부분에서 등도수로 나누는 휴리스틱을 사용하였다.

본 연구팀은 차분 프라이버시를 만족하는 히스토그램 공개를 위하여 기존의 2단계로 히스토그램을 생성하는 알고리즘[4]과 최적 히스토그램[5] 생성 알고리즘을 조합하여 더 적은 오차를 가지는 히스토그램을 생성한다.

* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음. (IITP-2018-2013-0-00881) 또한, 이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.2018-0-00498, 차분 프라이버시 기반 비식별화 기술 개발).

Ⅱ. 관련 연구

차분 프라이버시[1]가 제안된 이후 다양한 데이터를 공개하기 위한 여러 알고리즘들이 제안 되었다. 이 중 히스토그램을 공개하는 과정에서 오차를 줄이기 위한 대표적인 알고리즘인 Privelet[2]은 웨이블릿 변환을 이용하여 삽입하는 오차의 수준을 낮추었다. 하지만 히스토그램의 구간 크기가 고정되어 있어 오차를 줄이는데 한계가 있었다. 이를 극복하기 위해 구간의 크기를 동적으로 변경하여 오차를 더욱 줄이기 위한 연구들이 진행되었다[3,4].

Ⅲ. 배경 지식

1. 히스토그램

버킷이 N개인 히스토그램은 $H=(r_1,f_1),...,(r_N,f_N)$으로 나타낼 수 있고 r_i와 f_i는 다음과 같다.

r_i : i번째 버킷의 구간의 최대값 ($r_1 < ... < r_N$)

f_i : i번째 버킷의 도수 (r_{i-1}과 r_i사이에 존재하는 데이터 수)

2. 차분 프라이버시

어떤 양수 ϵ에 대하여 ϵ-차분 프라이버시를 만족시키는 알고리즘 A는, 1개의 레코드만 다르고 나머지는 모두 같은 데이터를 가진 임의의 데이터 쌍 D_1, D_2에 대하여 알고리즘 A를 수행하였을 때 아래의 부등식을 항상 만족시킨다.

$$\Pr(A(D_1)=S) \le \exp(\epsilon)*\Pr(A(D_2)=S) \quad (1)$$

3. 차분 프라이버시 순차 구성

알고리즘 1. HistoDP+(H_0, B, ϵ, R)
Input: 원본 히스토그램 $H_0 = (r_{o,1}, f_{o,1}), ..., (r_{o,N}, f_{o,N})$ 버킷 개수 B, 정보 보호 수준 ϵ ϵ_1, ϵ_2를 나누는 비율 R Output: 버킷 B개를 가지는 히스토그램 H_o''
1. $H_o' = H_0$의 각 버킷에 $Lap(1/\epsilon_1)$의 노이즈 삽입 2. H_o'' = OptHist(H_o', B) 3. H_o''의 구간 정보를 이용하여 실제 도수 계산 4. H_o''의 각 버킷에 $Lap(1/\epsilon_2)$의 노이즈 삽입 5. $return\, H_o''$

임의의 양수 ϵ_1, ϵ_2에 대하여 (ϵ_1, ϵ_2)-차분 프라이버시를 만족시키는 알고리즘의 결과를 동시에 공개하면 이는 $(\epsilon_1 + \epsilon_2)$-차분 프라이버시를 만족한다.

4. 문제 정의

버킷이 N개인 히스토그램이 주어질 때, 버킷이 B개 ($\langle N$)인 ϵ-차분 프라이버시를 만족시키는 히스토그램을 찾는 문제이다.

Ⅳ. 히스토그램 공개 알고리즘

1. 기존 알고리즘

노이즈를 원본 히스토그램에 삽입한 뒤, 이를 오차가 최소가 되는 최적의 히스토그램 계산을 통해 바꾸는 차분 프라이버시 알고리즘이 제안되었다[3]. i부터 j번째 버킷을 하나의 버킷으로 합치면서 발생하는 오차를 $SSE([i,j])$라고 하면, 식 (2)는 버킷을 줄이면서 발생하는 오차를 최소화하는 B개의 버킷을 가지는 히스토그램을 계산할 때 $(B-1)$개의 버킷을 가지는 최적 히스토그램 계산을 이용하는 재귀 식을 나타낸다.

$$SSE^*(i,B) = \min_{1 \le j < i} SSE^*(j, B-1) + SSE([j+1, i]) \quad (2)$$

[3]의 경우, 원본 히스토그램의 버킷의 숫자가 클 때 노이즈 삽입 회수가 많아지는 단점이 있어, 이를 개선하기 위하여 [4]에서는 차분 프라이버시의 순차 구성 원리를 적용하여, 정보 보호 수준을 2가지로 나누어, 우선 원본 히스토그램의 노이즈를 더하여 B개의 버킷으로 등도수가 되도록 나누고, 찾아진 버킷의 구간 정보를 이용하여 원래 데이터의 도수를 다시 세고, 다시 노이즈를 삽입하는 방법이 제안되었다.

2. 제안하는 알고리즘

기존 알고리즘[4]은 순차 구성을 적용하여 정보 보호 수준을 2가지로 나누어 히스토그램의 구간과 도수를 계산하는데 사용한다. 이 중, 히스토그램의 구간을 계산하기 위하여 등도수를 만족하는 B개의 구간을 찾는 휴리스틱 알고리즘을 사용하였는데, 데이터의 분포가 불균일한 경우 데이터가 몰려있는 곳을 합치게 되어 영역 합 질의의 성능이 나빠질 수 있다. 이를 위하여 [3]에서 사용하였던 최적 히스토그램 공개 알고리즘(OptHist)[5]를 사용하여 (line 2) 이를 개선한다.

Ⅴ. 실험

1. 실험 환경

데이터: American Community Survey에서 31,104,288명의 나이와 주당 근로 시간 데이터를 사용하였다. 각각 100개의 버킷이 존재한다.

실험 방법: 1000개의 랜덤한 영역 합 질의를 생성하여 성능을 측정하였고, 실험 결과는 10번씩 반복하여 평균 내었다.

실험 파라미터: 정보 보호 수준은 $\epsilon = 0.01$, 두 정보 보호 수준의 비율은 $R = 0.05$을 사용하였다.

평가 기준: 평가 기준은 [4]에서 사용한 평균 상대 오차를 사용하였으며 계산하는 식은 아래와 같다.

$$\frac{1}{|Query|} \Sigma_{q \in Query} \frac{|rangeSum(q) - rangeSumDP(q)|}{rangeSum(q)} \quad (3)$$

2. 실험 결과

표 1은 원본 히스토그램을 B개의 버킷으로 바꾸었을 때, 상대 오차를 나타낸 것이다. B값이 커질수록 더욱 많은 정보를 담은 히스토그램을 공개하기에 두 알고리즘 다 오차가 줄어드는 경향을 볼 수 있다. 또한, 제안하는 HistoDP+가 HistoDP에 비해 늘 좋은 상대오차를 보이는 것을 확인할 수 있다.

표 1. B에 따른 알고리즘별 상대 오차

B		20	30	40	50
나이	HistoDP	0.0163	0.0125	0.0049	0.0063
	HistoDP+	0.0038	0.0025	0.0011	0.0007
근로 시간	HistoDP	1.337	0.455	0.501	0.485
	HistoDP+	0.423	0.414	0.235	0.115

■ 참 고 문 헌 ■

[1] Dwork, C. "Differential privacy: A survey of results". In International Conference on Theory and Applications of Models of Computation (pp. 1-19). Springer, Berlin, Heidelberg. 2008.

[2] Xiao, X., Wang, G., & Gehrke, J. "Differential privacy via wavelet transforms". IEEE Transactions on Knowledge and Data Engineering, Vol. 23, No. 8, pp1200-1214. 2011.

[3] Xu, J., Zhang, Z., Xiao, X., Yang, Y., Yu, G., & Winslett, M. "Differentially private histogram publication". The VLDB Journal, Vol. 22, No. 6, pp797-822. 2013.

[4] 구한준, 오성웅, 정우환, & 심규석. "차분 프라이버시를 적용한 히스토그램 공개 알고리즘", 한국정보과학회 학술발표논문집, pp278-280. 2016.

[5] Jagadish, H. V., Koudas, N., Muthukrishnan, S., Poosala, V., Sevcik, K. C., & Suel, T. "Optimal histograms with quality guarantees", VLDB, Vol. 98, pp. 24-27. 1998.