Privacy

Félix Robitaille

Challenges

- AI requires large datasets.
- Research and development of new AI technology need access to similarly large datasets.
- A lot of the data needed consists of sensitive or private information.
- The data needs to be protected before it can be used widely.
- Protection has a cost.

Challenges

- Idea:

  Access to a statistical database should not enable one to learn anything about an individual that could not be learned without access.

- Impossible:

  Auxiliary information.

  Ex: Suppose one's exact height were considered a sensitive piece of information.

  Assume that the database yields the average heights of a certain group of people.

C. Dwork, "Differential privacy," in ICALP, 2006.

Protecting the data – Differential privacy

- the risk to one's privacy should not substantially increase as a result of participating in a statistical database.

- A randomized function K gives ε-differential privacy if for all data sets D_1 and D_2 differing on at most one element, and all $S \subseteq \text{Range}(K)$

$$\Pr[K(D_1) \in S] \leq e^{\varepsilon} \Pr[K(D_2) \in S]$$

C. Dwork, "Differential privacy," in ICALP, 2006.

Protecting the data – Differential privacy

- Other variants of the definitions exists:

- A randomized function K gives (ε, δ)-differential privacy if for all data sets D_1 and D_2 differing on at most one element, and all $S \subseteq Range(K)$

$$Pr[K(D_1) \in S] \leq e^{\varepsilon} Pr[K(D_2) \in S] + \delta.$$

- Preferably, $\delta < 1/|D_1|$

C. Dwork, F. McSherry, K. Nissim, and A. Smith. "Calibrating noise to sensitivity in private data analysis". In TCC, pages 265–284. Springer, 2006.

Protecting the data – Differential privacy

- Data utility will eventually be consumed:
- The Fundamental Law of Information Recovery states that overly accurate answers to too many questions will destroy privacy in a spectacular way.

- Differential privacy ensures that the same conclusions will be reached, independent of whether any individual opts into or opts out of the data set.

C. Dwork, F. McSherry, K. Nissim, and A. Smith. "Calibrating noise to sensitivity in private data analysis". In TCC, pages 265–284. Springer, 2006.

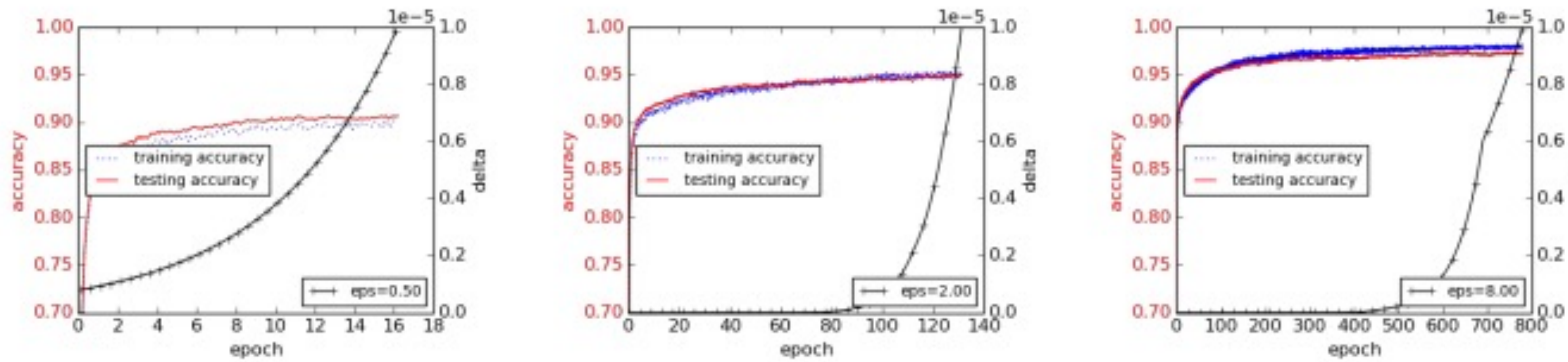
SGD with Differential privacy

- Protect against model-inversion attacks.
- Assumes that the trained model is exposed.
- Compute SGD while clipping and adding noise to each step
- In this case, Gaussian noise is used
- Computing the cost:
- $\sigma = \frac{\sqrt{2log\frac{125}{\delta}}}{\varepsilon}$
- is $(O(q\varepsilon\sqrt{T}), \delta)$- differentially private

- Input:
  - Examples $\{x_1, \dots, x_N\}$,
  - loss function: $L(\theta) = \frac{1}{N}\sum_i L(\theta, xi)$.
- Parameters:
  - learning rate η_t,
  - noise scale σ,
  - group size L,
  - gradient norm bound C.

- Initialize θ_0 randomly
- for $t \in [T]$ do
  - Take a random sample L_t with sampling probability L/N
  - For each $i \in L_t$, compute $g_t(xi) \leftarrow \nabla\theta_t L(\theta_t, x_i)$
  - $\bar{g}_t(x_i) \leftarrow g_t(x_i)/ \max(1, \frac{||g_t(x_i)||_2}{C})$
  - $\tilde{g}_t \leftarrow \frac{1}{L}(\sum_i \bar{g}_t(x_i) + N(0, \sigma^2 C^2 I))$
  - $\theta_{t+1} \leftarrow \theta_t - \eta_t \tilde{g}_t$
- Output θ_T and compute the overall privacy cost (ε, δ) using a privacy accounting method.

Abadi et al. "Deep Learning with Differential Privacy" in arXiv:1607.00133v2, 2016



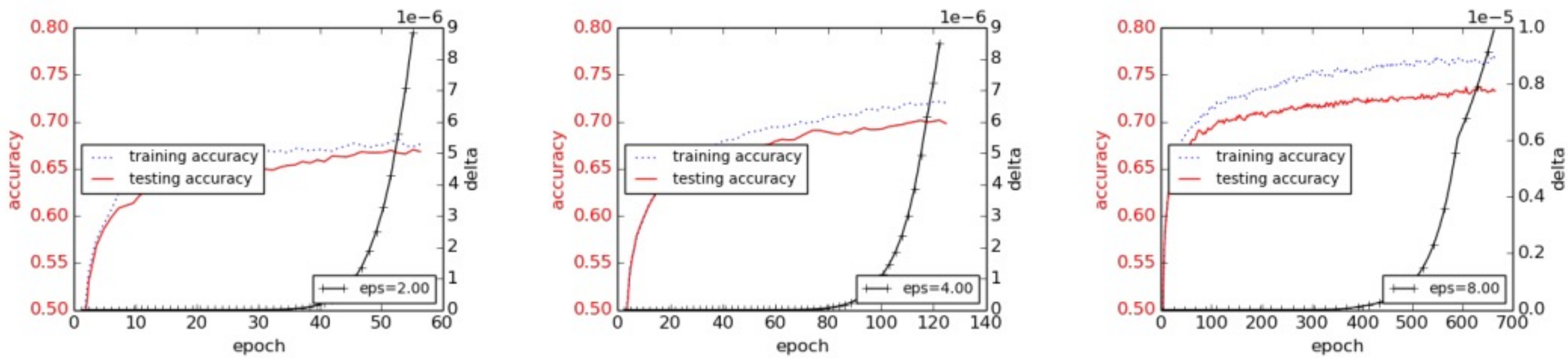
(1) Large noise

(2) Medium noise

(3) Small noise

Results

- Tested on the MNIST dataset
- Baseline model:
  - a 60-dimensional PCA projection layer and a single hidden layer with 1,000 hidden units.
  - Using the lot size of 600, we can reach accuracy of 98.30% in about 100 epochs
- Differential privacy: same architecture with three levels of noise



(1) $\varepsilon = 2$ (2) $\varepsilon = 4$ (3) $\varepsilon = 8$

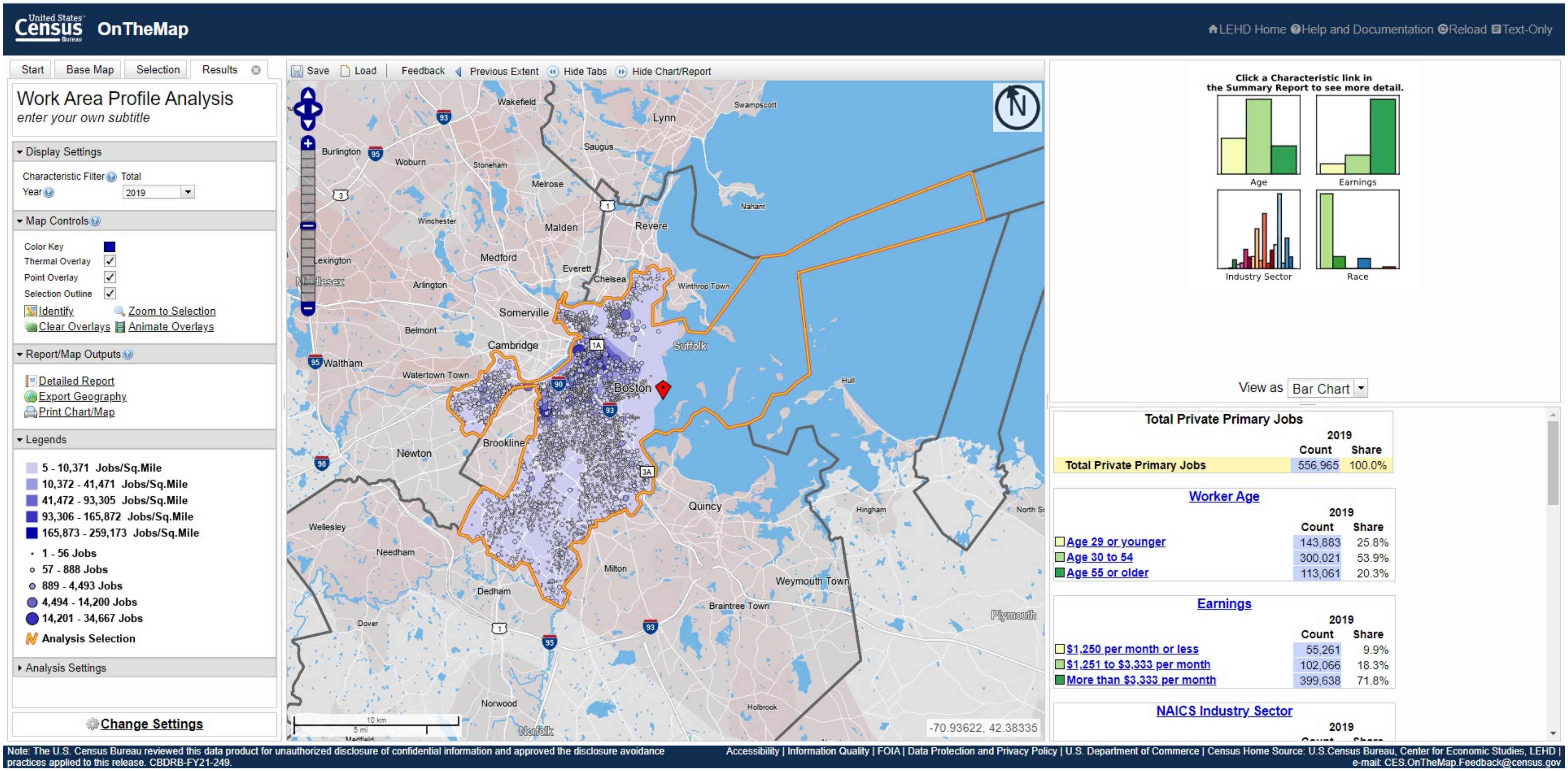
Results

- Tested on the CIFAR-10 dataset
- Baseline model:
    - Network architecture from the TensorFlow convolutional neural networks tutorial.
    - Two convolutional layers followed by two fully connected layers.
    - Reaches about 86% accuracy in 500 epochs.
- Differential privacy: same architecture with three levels of noise

In practice: OnTheMap

- Mapping program to show the commuting patterns of the USA population.
- Shows where American workers are employed and where they live.
- Data from the US census bureau.
- The data can’t be used directly because of privacy concerns.

- https://onthemap.ces.census.gov/


United States Census Bureau
OnTheMap
LEHD Home
Help and Documentation
Reload
Text-Only
Start
Base Map
Selection
Results
Save
Load
Feedback
Previous Extent
Hide Tabs
Hide Chart/Report
Work Area Profile Analysis
enter your own subtitle
Display Settings
Characteristic Filter Total
Year 2019
Map Controls
Color Key
Thermal Overlay
Point Overlay
Selection Outline
Identify
Zoom to Selection
Clear Overlays
Animate Overlays
Report/Map Outputs
Detailed Report
Export Geography
Print Chart/Map
Legends
5 - 10,371 Jobs/Sq.Mile
10,372 - 41,471 Jobs/Sq.Mile
41,472 - 93,305 Jobs/Sq.Mile
93,306 - 165,872 Jobs/Sq.Mile
165,873 - 259,173 Jobs/Sq.Mile
1 - 56 Jobs
57 - 888 Jobs
889 - 4,493 Jobs
4,494 - 14,200 Jobs
14,201 - 34,667 Jobs
Analysis Selection
Analysis Settings
Change Settings
Wakefield
Lynn
Swampscott
Burlington
Woburn
Stoneham
Saugus
Melrose
Nahant
Winchester
Malden
Revere
Lexington
Medford
Everett
Chelsea
Winthrop Town
Middlesex
Arlington
Somerville
Belmont
Cambridge
Suffolk
Waltham
Watertown Town
Boston
Hull
Brookline
Newton
Quincy
Hingham
Wellesley
Needham
Milton
Weymouth Town
Dedham
Braintree Town
Plymouth
Dover
Norwood
Holbrook
Norfolk
10 km
5 mi
-70.93622, 42.38335
Click a Characteristic link in the Summary Report to see more detail.
Age
Earnings
Industry Sector
Race
View as Bar Chart
Total Private Primary Jobs
2019
Count Share
Total Private Primary Jobs 556,965 100.0%
Worker Age
2019
Count Share
Age 29 or younger 143,883 25.8%
Age 30 to 54 300,021 53.9%
Age 55 or older 113,061 20.3%
Earnings
2019
Count Share
$1,250 per month or less 55,261 9.9%
$1,251 to $3,333 per month 102,066 18.3%
More than $3,333 per month 399,638 71.8%
NAICS Industry Sector
2019
Note: The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and approved the disclosure avoidance practices applied to this release. CBDRB-FY21-249.
Accessibility | Information Quality | FOIA | Data Protection and Privacy Policy | U.S. Department of Commerce | Census Home Source: U.S.Census Bureau, Center for Economic Studies, LEHD | e-mail: CES.OnTheMap.Feedback@census.gov

In practice: OnTheMap

- Data points contain id, origin block, destination block.
- Destination block are public data.
- Origin block is treated as the sensitive attribute.
- Looked at different criterias for privacy.
- Used original Differential privacy criteria.
- Uses synthethic data generation to anonymize the data.
- Privacy comes from the bias from the model and noise from random sampling.

Machanavajjhala, Ashwin, Daniel Kifer, John Abowd, Johannes Gehrke, and Lars Vilhuber, "Privacy: Theory Meets Practice On the Map," International Conference on Data Engineering (ICDE) 2008, pp. 277-286.

Problems

- Differential Privacy requires consider adversaries who know all but one data point.
- Needs more than 913 people in each block to satisfy the criteria.

- Cause: worst case scenario.
- Generates completely unrepresentative synthetic data.
- Extremely unlikely

- Variation to relax Differential Privacy: Probabilistic Differential Privacy.

Problems

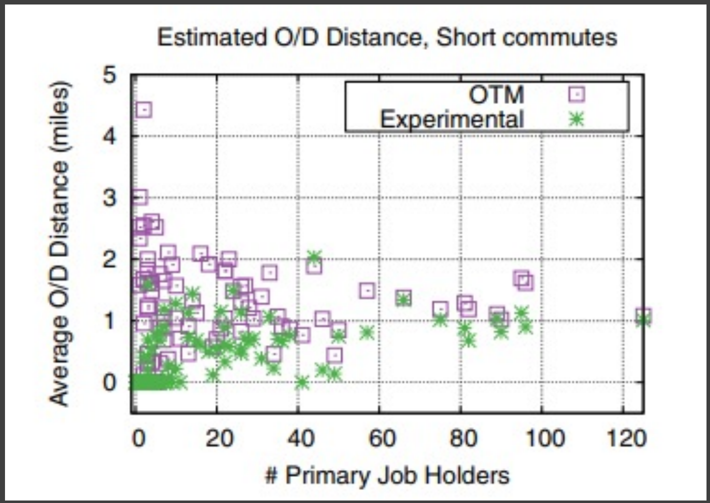
- Unrepresentative data can be generated.
  - the accept/reject method, is to choose a "representativeness" metric and rerun the algorithm until we get an output which is representative of the input
  - Needs an acceptance metric that is compatible with Probabilistic Differential Privacy.

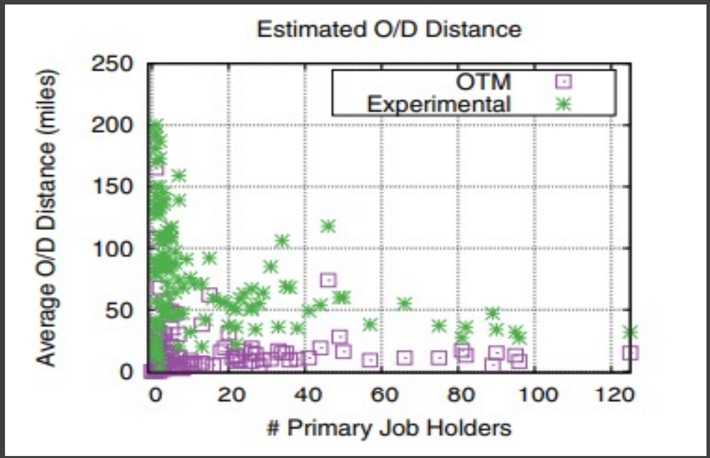
- Needs noise in each block to satisfy Probabilistic Differential Privacy, but the data is sparse.
  - Clustering.
  - Needs a clustering algorithm that is compatible with Probabilistic Differential Privacy.

Results

- Compared the average commute distance for each destination block and compared it to the ground truth from the data.

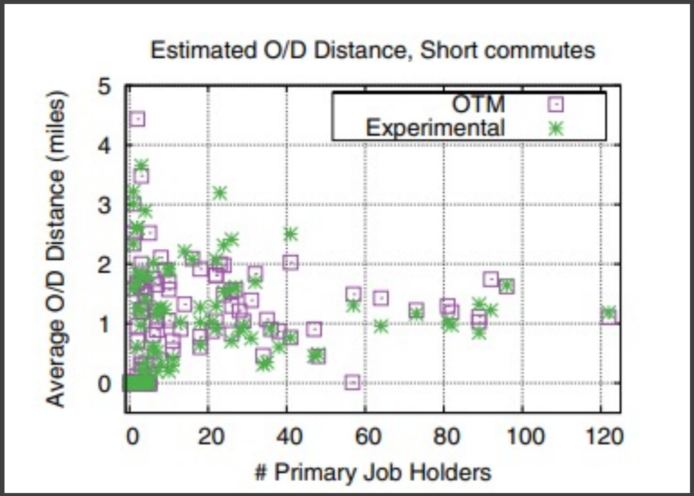
- Long commutes are overestimated.

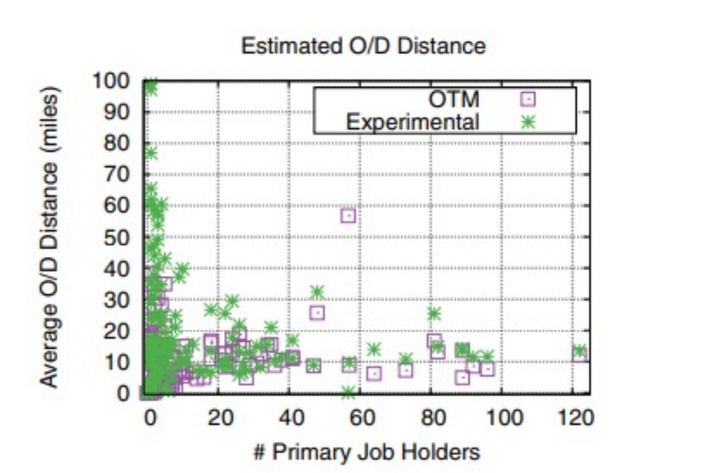




Results

- Compared the average commute distance for each destination block and compared it to the ground truth from the data.
- Long commutes are overestimated.
- Added more restrictions to prevent Outliers.





Improvements

- Preventing drops in accuracy when integrating Differential privacy to a model.
- Dealing with scarcity of data to prevent outliers.

Questions?