QARTA: An ML-based System for Accurate Map Services

Mohamed Mokbel

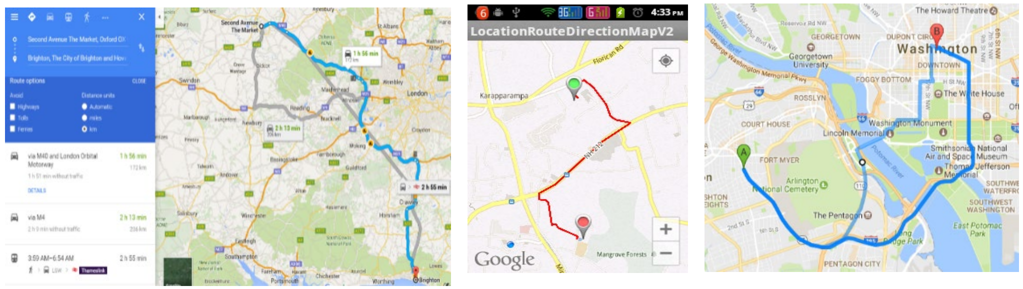
University of Minnesota

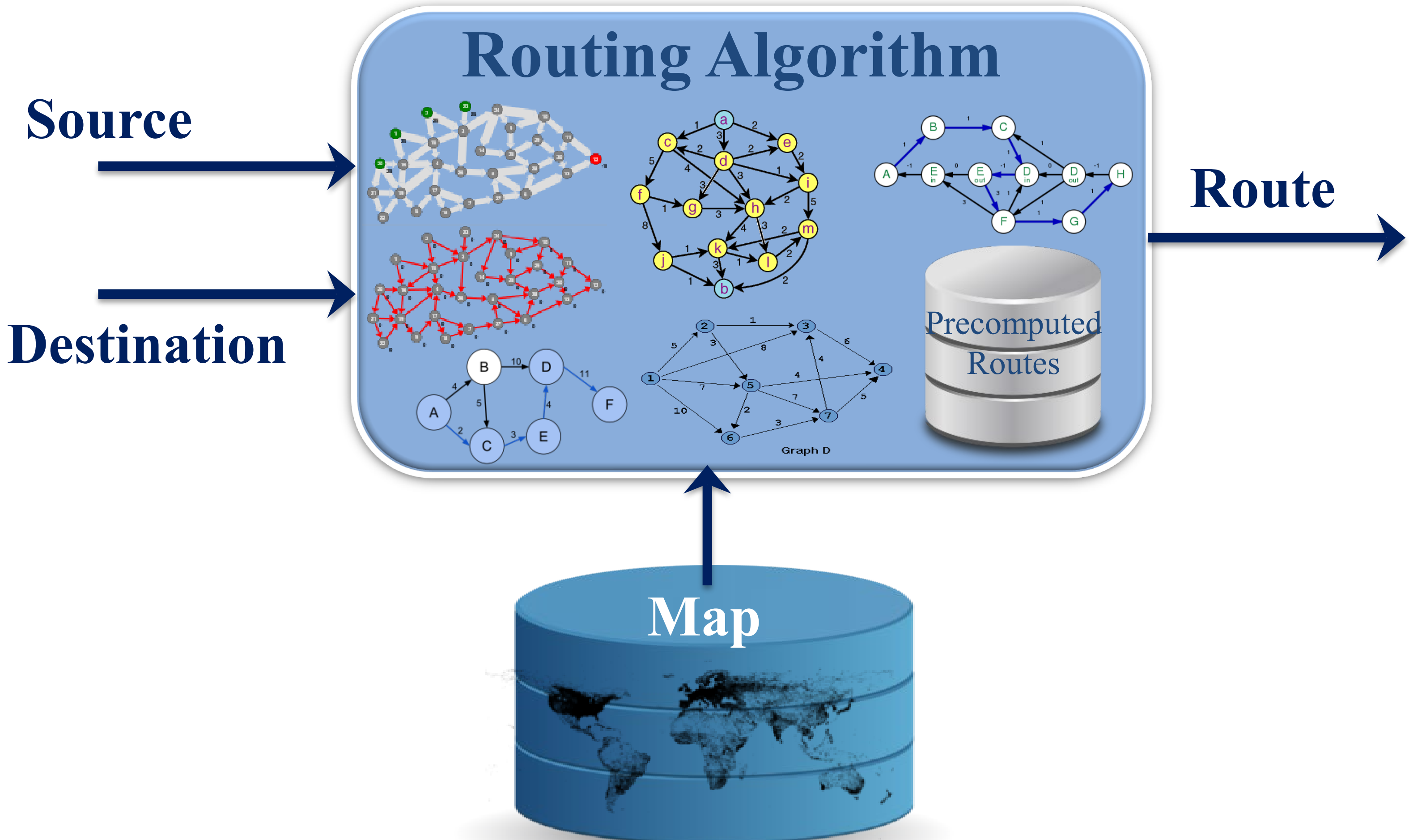
Work with:

Sofiane Abbar, Mohamed El-Shrif, Rade Stanojevic (QCRI) and Mashaal Musleh (UMN)

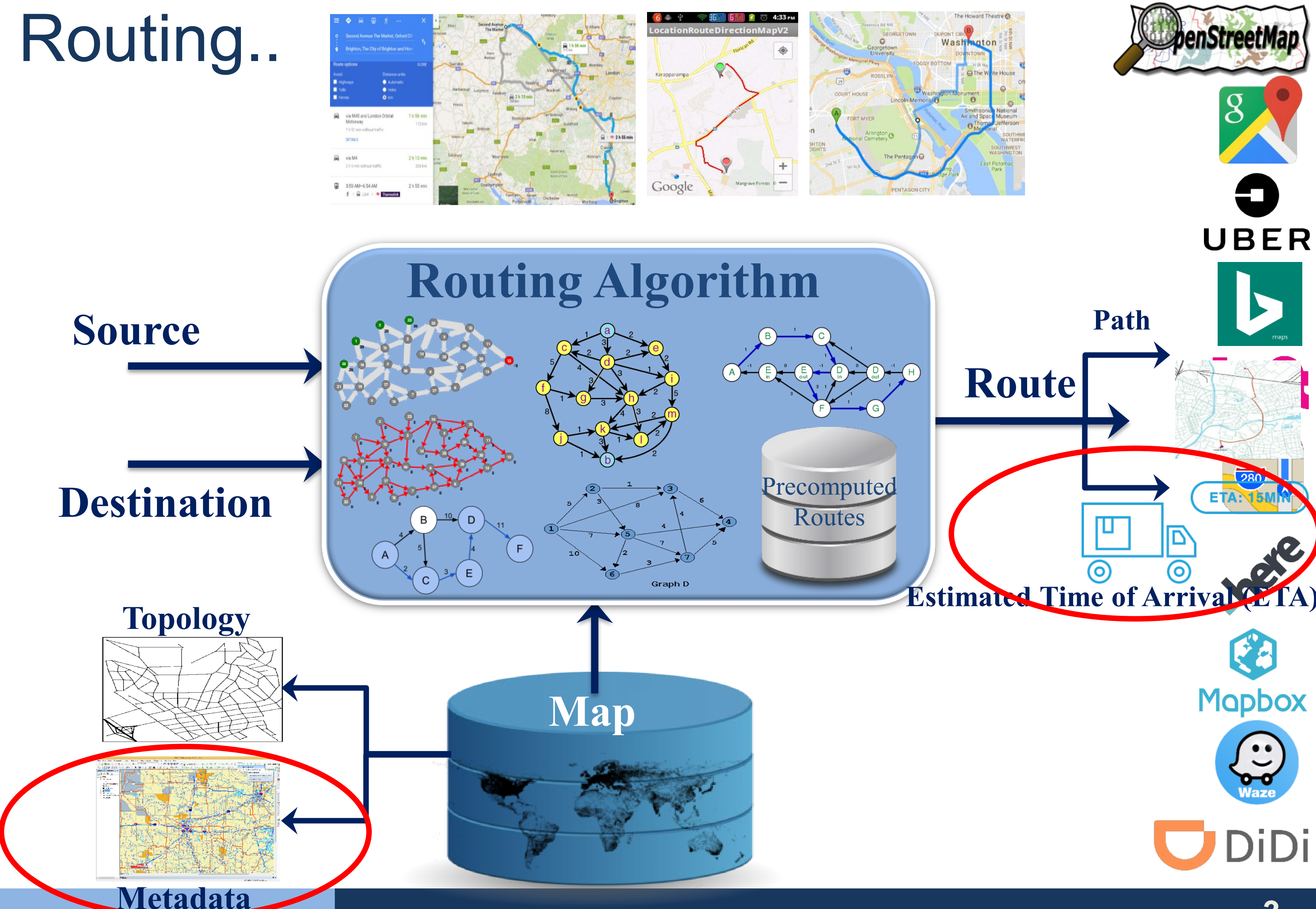
QCRI
معهد قطر لبحوث الحوسبة
Qatar Computing Research Institute
جامعة حمد بن خليفة
HAMAD BIN KHALIFA UNIVERSITY

Routing..




Source
Destination
Routing Algorithm
Graph D
Precomputed Routes
Route
Map

Routing..



Routing Algorithm
Source
Destination
Precomputed Routes
Route
Path
Estimated Time of Arrival (ETA)
ETA: 15MIN
Topology
Metadata
Map
Graph D
OpenStreetMap
UBER
Mapbox
Waze
DiDi

The War Over Maps

GRAND VIEW RESEARCH

Reports Services About Us Insights Contact

Home » Next Generation Technologies » Digital Map Market Size, Share, Global Industry Report, 2027

PDF Market Analysis Report

Digital Map Market Size, Share & Trends Analysis Report By Type (GIS, LiDAR, Aerial Photography, Digital Orthophotography), By Usage, By Services, By End Use, By Region, And Segment Forecasts, 2020 - 2027

Published Date: Mar, 2020 | Base Year for Estimate: 2019 | Report ID: GVR-2-68038-021-7

Format: Electronic (PDF) | Historical Data: 2016 - 2018 | Number of Pages: 163

Report Summary | Table of Contents | Segmentation | Methodology | Request a Free Sample Copy

Report Overview

The global digital map market size was estimated at USD 6.19 billion in 2019 and is expected to expand at a CAGR of 13.0% from 2020 to 2027. Increasing adoption of novel technologies in map making and surveying along with the rising adoption of digital maps across various industry verticals is propelling the growth. The increasing internet penetration and rise in connected devices such as smartphones, tablets, and interactive displays for adoption of

CNN BUSINESS Markets Tech Media Success Video

The billion dollar war over maps

by Seth Fiegerman @sfiegerman

June 7, 2017: 3:50 PM ET

The future of cars: Self-driving and electric

During a test drive near Ford's Michigan headquarters, the team noticed something strange with its self-driving cars.

Each car shifted slightly at the same point in the lane "as if they were avoiding a pothole," says Jim McBride, Ford's senior technical leader for autonomous cars.

The problem wasn't the cars -- it was the map.

GPS WORLD GNSS POSITION NAVIGATION TIMING

Follow Us

Search the Site... subscrib

GNSS OEM UAV Survey Mapping Transportation Defense Mobile Machi

Google to improve urban GPS accuracy for apps

December 9, 2020 - By Frank Van Diggelen Est. reading time: 5 minutes

Originally posted in the Android Developers Blog, the following is reprinted with permission from authors Frank van Diggelen, principal engineer, and Jennifer Wang, product manager, Google.

FOX 40 NEWS MORNING WEATHER SPORTS COMMUNITY STUDIO40 LIVE ABOUT US

YOUR LOCAL ELECTION HEADQUARTERS

Google Maps leads voters to wrong polling place in Arden-Arcade

by: Lonnie Wong

Posted: Nov 4, 2020 / 05:42 PM PST / Updated: Nov 4, 2020 / 05:42 PM PST

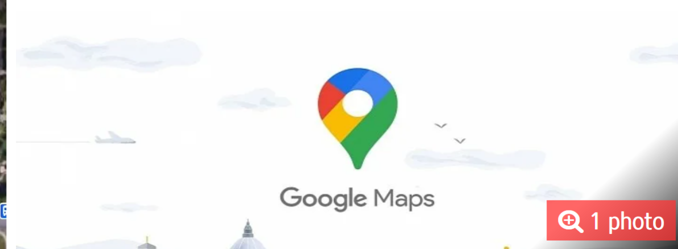
SACRAMENTO, Calif. (KTXL) – Technology has helped election officials register voters and tabulate their ballots during election cycles.

FEATURED autoevolution's Toyota Month →

The Nightmare a Simple Google Maps Error Can Create

Home > News > Technology 21 Dec 2020, 10:33 UTC · by Bogdan Popa

Google Maps isn't only the world's most used navigation app out there, but also the solution plenty of companies turn to for their services.



For example, food delivery companies adopt Google Maps to better optimize their services, which means the accuracy of data provided by Google is critical for their customers.

A Londoner whose story was recently highlighted by The Guardian reveals just how troublesome simple Google Maps errors can become, explaining that in the last three years, Uber drivers, food delivery services, and pretty much anybody who wanted to reach his flat ended up in the middle of the park.

npr

NEWS | ARTS & LIFE | MUSIC | SHOWS & PODCASTS | SEARCH

TECHNOLOGY

Google Maps Leads About 100 Drivers Into A 'Muddy Mess' In Colorado

June 27, 2019 · 11:35 AM ET

MERRIT KENNEDY

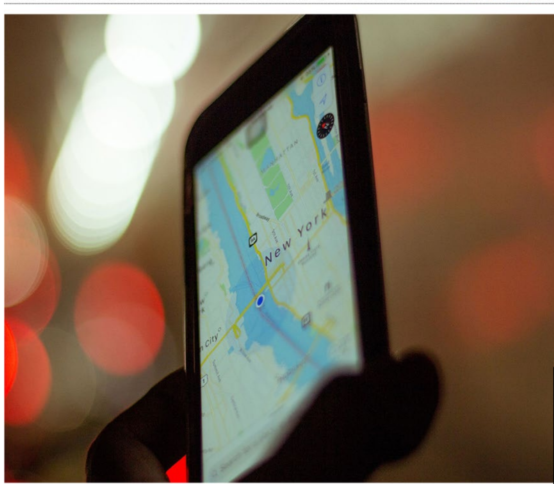
Nearly 100 drivers were recently led astray on a Google Maps-suggested detour near Denver's airport.

POPULAR SCIENCE

Google and Apple Maps have plenty of errors. Here's how to fix them.

Flag missing roads, update restaurants' opening hours, and more.

By David Nield | November 1, 2018



GADGET HACKS

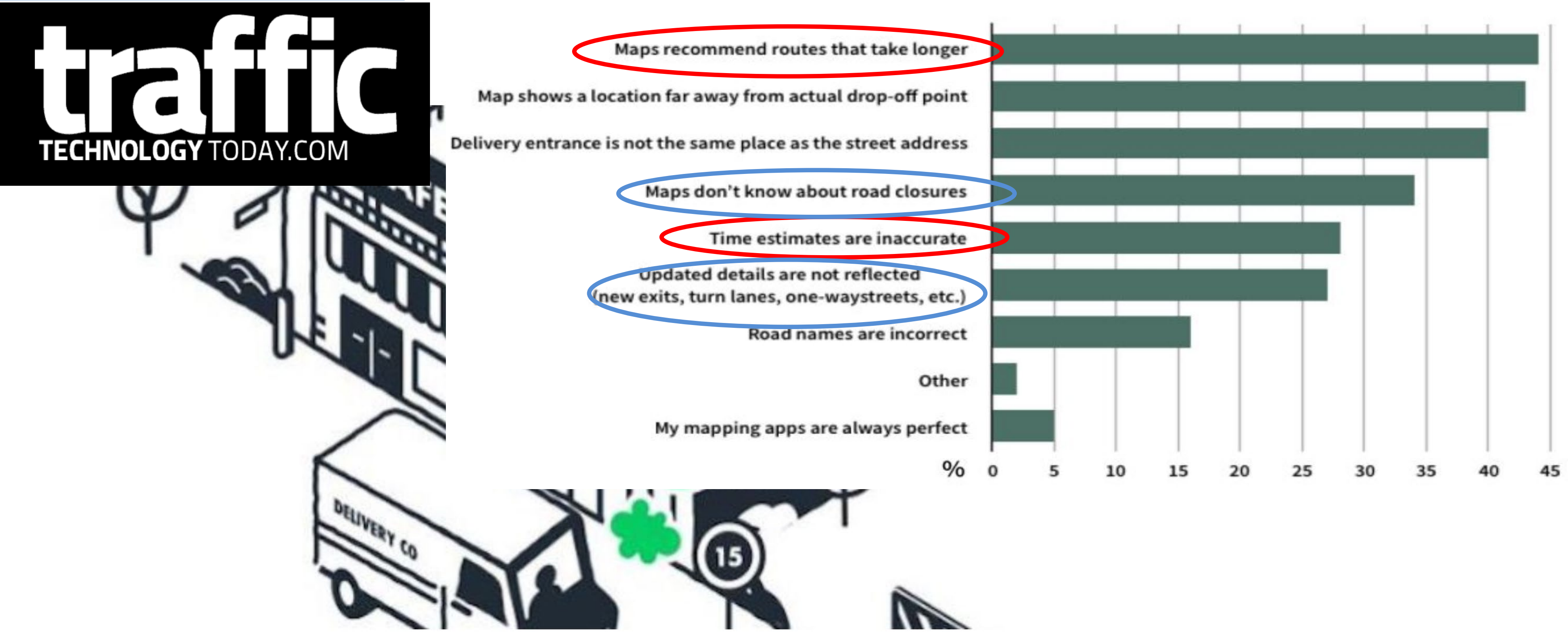
BEST NAVIGATION APPS

Google Maps vs. Apple Maps vs. Waze vs. MapQuest

Best Maps & Navigation Apps for Mobile

GADGET HACKS	Google Maps	Apple Maps	Waze	MapQuest
General				
Platform	Android, iOS, macOS, Windows	iOS, macOS	Android, iOS, Windows	Android, iOS, macOS, Windows
Map Features				
Countries & Terriories Mapped	266	181	72	2[illegible]2
Countries & Territories with Driving Directions	256	101	72	252
Street View	Yes	No	No	No
Overlays	Satellite, Terrain, Transit, Traffic, Bicycling	Satellite, Transit	None	Satellite
3D View	3D Structures	3D Renderings	No	No
Live Location Sharing	Yes	Yes	No	No
Location History	Yes	Yes	Yes	Yes
Cultural Hotspot Indicators	Yes	No	No	No
Weather Data	None	Weather, Temperature, Air Quality	None	Weather, Temperature
Indoor Maps	Airports, Malls, Museums	Airports, Malls	No	No
Offline Maps	Yes	Yes	No	No
Navigation Features				
Traffic Data	Alternate Routes, Accidents, Road Work, Speed Traps	Alternate Routes, Accidents, Road Work	Accidents, Alt. Routes, Road Work, Potholes, Police, Speed Traps	Accidents, Road Work, Traffic Cameras
Traffic Data Source	In-House, User Curated	In-House, Third-Party	In-House, User Curated	Third-Party, User Curated
High Traffic Warnings	Yes	Yes	No	No
Speed Limits	Yes	Yes	Yes	Yes
Lane Guidance	Yes	Yes	Yes	Yes
Add Toll & HOV Passes	No	No	Yes	No
Avoid Tolls & Highways	Yes	Yes	Yes	Yes
Choose Different Routes	Yes	Yes	Yes	Yes
Add Pit Stops	Unlimited	1	1	Unlimited
Show Gas Prices	Yes	Yes	Yes	Yes
Hands-Free Control in App	Yes	Yes	Yes	No
Directions Using Other Modes of Transport	Transit, Biking, Walking, Ride Share	Transit, Walking, Ride Share	Motorcycles, Taxis	Biking, Walking
Re-Center	Yes	Yes	Yes	Yes
Accessible Navigation	Yes	Yes	No	No
Save Parking Spot	Yes	Yes	No	No
Offline Navigation	Yes	Yes	No	No
Works With Screen Off	Yes	Yes	Yes	Yes
App Features				
Dark Mode	Yes	Yes	Yes	Yes
Ride Share Integration	Uber, Lyft, Lime,	Uber, Lyft	None	None
Picture In Picture	Yes (Android Only)	No	No	No
Lock Screen Navigation	Yes	Yes	Yes	Yes
Show Festivals & Protests	No	No	Yes	No
Personalized Recommendations	Yes	No	No	No
Book Dinner Reservations	Via OpenTable	Via OpenTable	No	No
Report Traffic Issues	No	No	Yes	No
Post Reviews	Yes	No	No	No
Car Support	Android Auto, CarPlay	CarPlay	Android Auto, CarPlay	No
AR Features	Interactive Street View	Flyover	None	None
Widgets	Yes	Yes	Yes	No
Music Integration	Yes	Yes	Yes	No

https://smartphones.gadgethacks.com/how-to/best-navigation-apps-google-maps-vs-apple-maps-vs-waze-vs-mapquest-0194591/



Poor maps costing delivery companies US$6bn annually

0

BY ADAM FROST ON FEBRUARY 19, 2020 MAPPING

Based on a survey of delivery drivers in the USA and conducted by an independent research firm, the first 'Mapping in Logistics Report' has revealed that 'broken maps' are costing the logistics sector an estimated US$6bn annually.

https://www.traffictechnologytoday.com/news/mapping/poor-maps-costing-delivery-companies-us6bn-annually.html

Meanwhile, in Academia..

- **Shortest Path Queries:**
  - Too many algorithms: Implicitly assume road network is correct
  - Either consider Free Flow (max speed) or given edge weights a
  - Recent techniques use ML for path recommendation

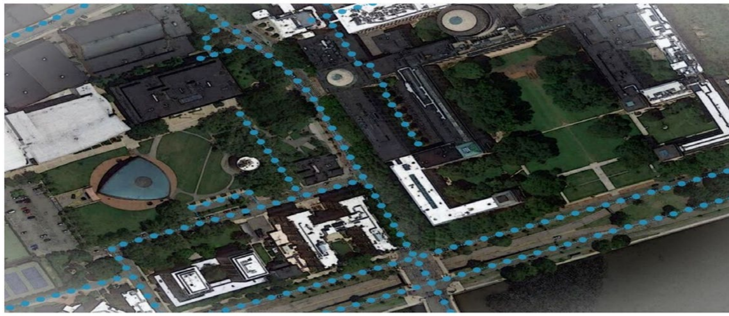
- **Map Making**
  - Focus on discovering road network topology from satellite data or GPS traces

MIT News
ON CAMPUS AND AROUND THE WORLD

Using artificial intelligence to enrich digital maps

Model tags road features based on satellite images, to improve GPS navigation in places with limited map data.

Rob Matheson | MIT News Office
January 23, 2020



A model invented by researchers at MIT and Qatar Computing Research Institute (QCRI) that uses satellite imagery to tag road features in digital maps could help improve GPS navigation.

The Science of Algorithmic Map Inference

KDD 2018 Tutorial

Abstract:

A necessary condition for autonomous vehicles to become "mainstream" is the availability of highly accurate and updatable geographical road network maps. Several larg commercial mapmaking efforts by automobile manufacturers and technology companies have been recently announced . The race to build the most accurate maps is "truly on." societal concern is that, in the near future, the most accurate maps may not be a public good but a property of private stakeholders. A concerted effort is required to democratiz mapmaking, and collaborative efforts like OpenStreetMap (OSM) demonstrate the power of the community coming together to create and maintain maps as a public good.

In this tutorial we will review the emerging area of algorithmic map inference (AMI), i.e., the design of algorithms to automatically build and update maps using diverse data source , primarily GPS data and satellite images. A substantial body of research has now emerged around AMI primarily in KDD and related communities. Thus it is an opportune time t organize the AMI literature in a proper context and introduce it to a wider audience in the research and applications community.

Program:

We will be in ICC Capital Suite Room 2 (Level 3) on 19 August 2018 1:00 pm - 5:00 pm.

- Introduction (1:00 pm)
- Map Inference with GPS Trajectories (1:20 pm)
- Map Inference with Satellite Imagery (2:15 pm)
- Map Inference with In-Vehicle Cameras (3:00 pm)
- Coffee break (3:30 pm)
- Quality of Maps (4:00 pm)

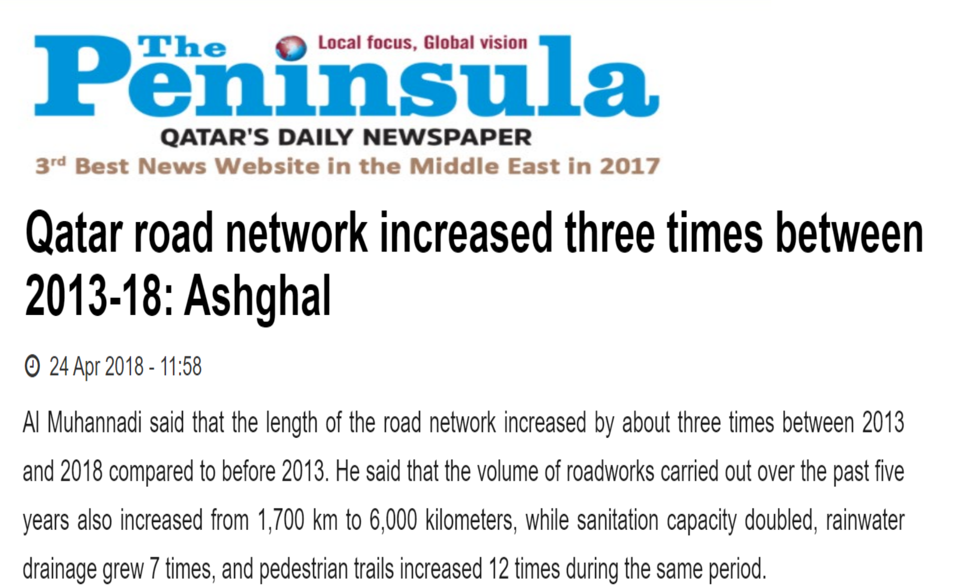
QARTA: An ML-based System for Accurate Map Services

QARTA: Why..??

- Problem came up from the Taxi company working in Qatar

Too much construction and road changes in town (in preparation to FIFA 2022)

Commercial maps cannot cope with such changes in road networks, and are not cheap



The Peninsula
Local focus, Global vision
QATAR'S DAILY NEWSPAPER
3rd Best News Website in the Middle East in 2017

Qatar road network increased three times between 2013-18: Ashghal

24 Apr 2018 - 11:58

Al Muhannadi said that the length of the road network increased by about three times between 2013 and 2018 compared to before 2013. He said that the volume of roadworks carried out over the past five years also increased from 1,700 km to 6,000 kilometers, while sanitation capacity doubled, rainwater drainage grew 7 times, and pedestrian trails increased 12 times during the same period.

Raya Daily (Sept. 8, 2020), 20

QARTA: Main Concepts

- Supports dynamics in the map topology and metadata
- Map-Centric
  - QARTA *learns* its own map in terms of topology and metadata
- Query Calibration
  - QARTA *learns* the error margins of various algorithms and use it to calibrate its answer

CACM, April 2021

Digital Mapping | DOI:10.1145/3447731

Traffic Routing in the Ever-Changing City of Doha

BY SOFIANE ABBAR, RADE STANOJEVIC, SHADAB MUSTAFA, AND MOHAMED MOKBEL

ON DECEMBER 2, 2010, Qatar was announced to host 2022 FIFA World Cup. That was time for celebrating the first-ever Middle Eastern country to organize the tournament. The 1.8M population of Qatar then (2.8M today) never imagined the journey their country was about to embarked. Indeed, in less than 10 years, the population grew by more than a half, pushing the available urban resources and services to their limit. At the same time, the country undertook an ambitious investment plan of $200B on various infrastructural projects including a brand new three-line metro network, six new stadiums, several new satellite cities, and an astonishing 4,300km of new roads, which tripled the size of the road network in only five years.[3]

While this enterprise boosted the socio-economical life of people in Qatar, it did disrupt the way they navigate the urban space and their mobility patterns in general. Simple commutes to work, drops and pickups of kids to and from schools, became challenging and impossible to plan with daily changes in the road layout, including temporary and permanent closures, deviations, new connections, conversions of roundabouts into signaled intersections, turn restrictions, to name but a few. A commute to school that lasted 10 minutes yesterday, could last 25 minutes today. Cab drivers in the city of Doha (Qatar's capital), who are mostly foreigners, also wish they could rely on popular navigation services such as Google Maps, Here, or Tomtom.

Yet, all such systems fall short in coping up with the rapid urbanization and the ever-changing roads in Doha. This was actually depicted in a very popular caricature in one of the most widely distributed daily local newspapers showing Google maps as a limping turtle that is helplessly trying to catch a bunny representing the road changes in the city of Doha.[4]

Besides the general public who is not happy with the routes offered by navigation systems, other stakeholders from public and private sectors were struggling with the poor quality of existing digital maps. For example, the Ministry of Transport and Communication was facing issues getting access to the most accurate map of the road network, needed for their traffic modeling. Also, transportation, delivery, and logistics companies that heavily rely on accurate maps, routes, and travel time estimates were tired of the many lost drivers and missed rendezvous.

Early work: Silent maps are not enough. The issue of inaccurate local maps has triggered an early work at Qatar Computing Research Institute (QCRI) in collaboration with Qatar Mobility Innovation Center (QMIC) to come up with an accurate map for the city of Doha, Qatar.[7] The idea was to use data collected from a fleet of vehicles that are continuously tracked, for accurate and timely detection of road changes, such as new roads, road closures, and detours. Though that early work was successful in coming up with a more accurate map than what navigation systems have, it was not enough to address the main problem of routing. Accurate topological maps do not say much about the time needed to go through each road segment—a main functionality needed for any routing application.

Data access and collaboration. To address the routing problem in the ever-changing roads of Doha, we partnered with the national taxi company Karwa. The collaboration gives us access to all taxi data (both historic and live) that took place in the country, including pick-up and drop-off locations, time, duration, speed, fare, route, as well as sampled GPS points for each trip—a gold mine for our research agenda. But most importantly, we also learned from our partners about the real challenges they face, which helped us prioritize our projects.

Map enrichments for traffic-aware routing. Our first project with Karwa was to enrich the topological maps with traffic information, that is, accurate edge weights for each road segment for each hour of the day. Inferring traffic information from a large number of vehicles can be relatively straightforward. However, the problem is much more challenging when the data is sparse and does not cover many roads with large frequency. We tackle these problems in Stanojevic et al.[5,6] and derive a traffic layer with an accuracy comparable to the commercial maps using only sparse data available to us either from Karwa Taxi data as in Stanojevic et al.[5] or from using commercial map APIs as in Stanojevic et

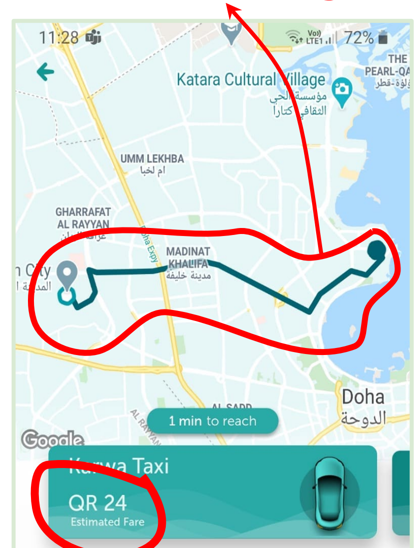
QARTA in Deployment

QARTA is deployed in *all* Taxis in Qatar $\sim 4K$ vehicles

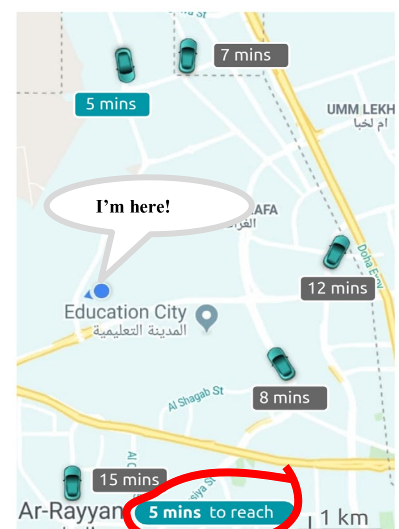
A local food delivery company $\sim 3K$ motorbiks

- QARTA receives:
  - ❑ $\sim 235K$ daily API calls
  - ❑ $\sim 1\ Million$ daily GPS tracks
- APIs & Services:
  - ❑ In-traffic routes
  - ❑ Travel time estimation
  - ❑ Complex route planning
  - ❑ OD matrices
  - ❑ Search & addresses

Routing



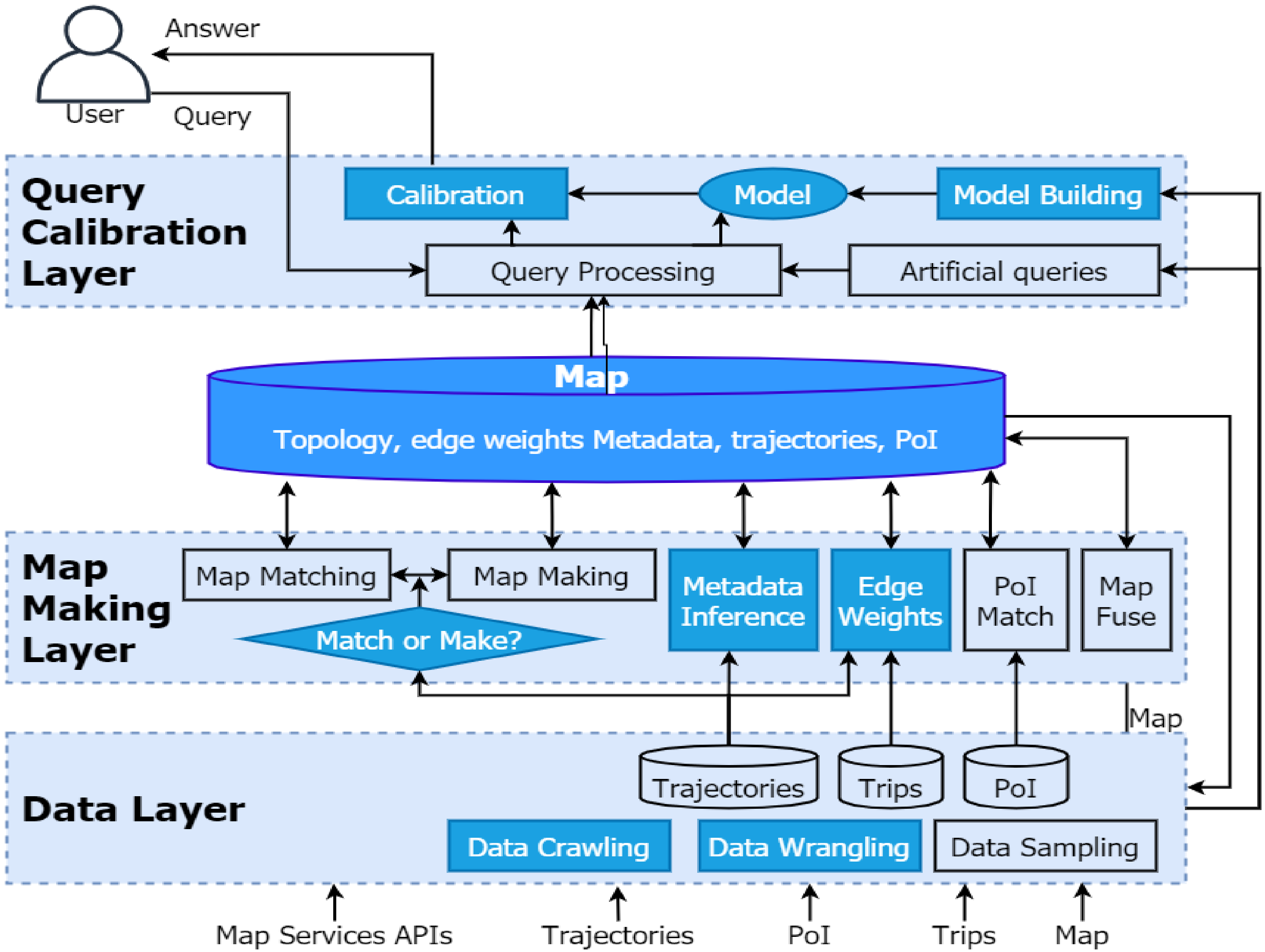
Fare estimation



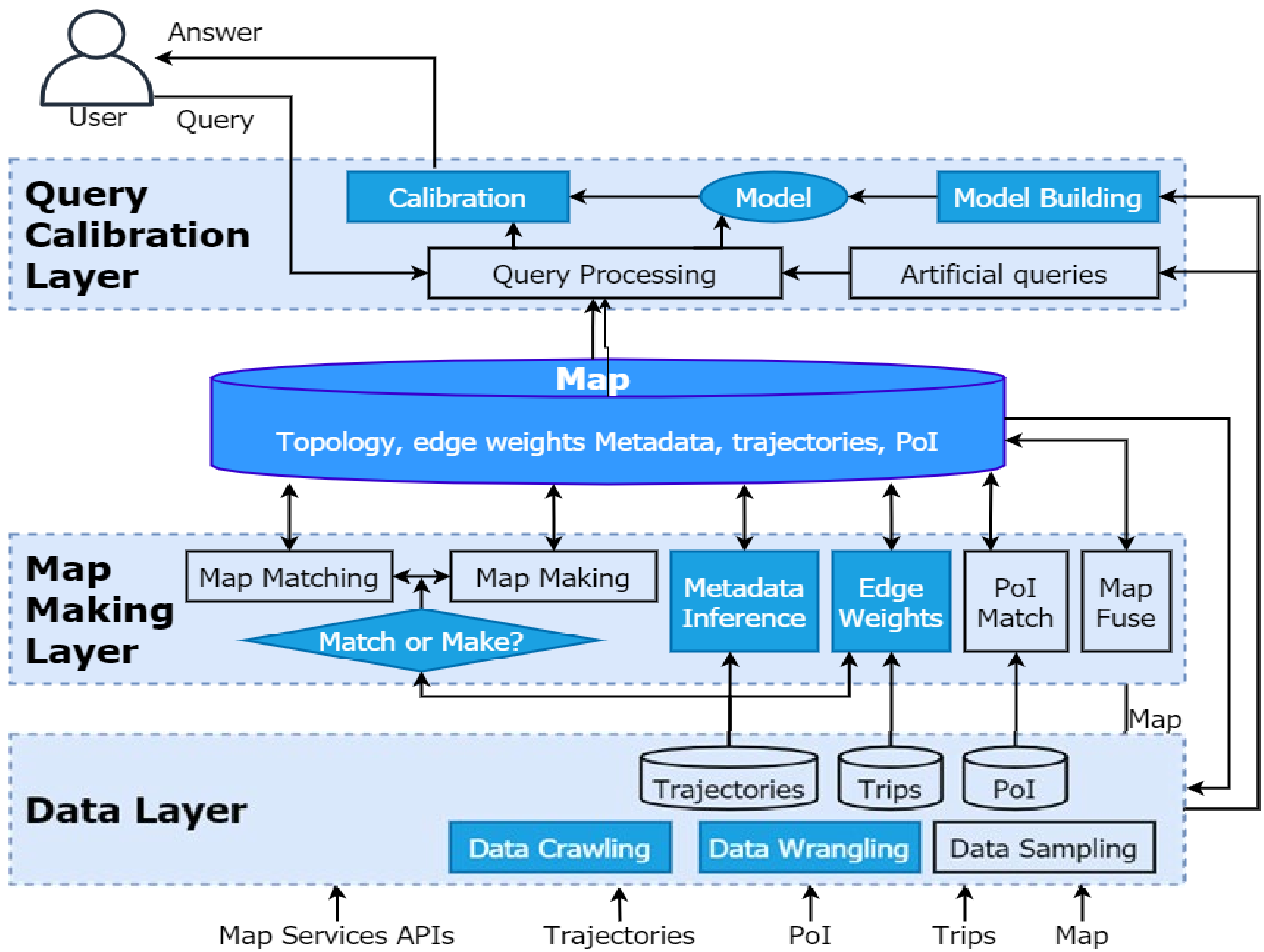
Taxi Dispatching

Link: https://qarta.io

QARTA Architecture



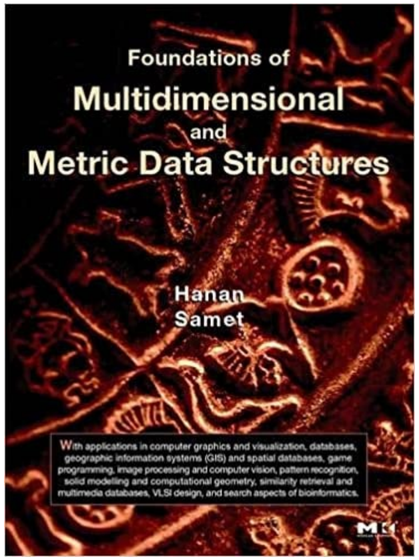
Data Layer

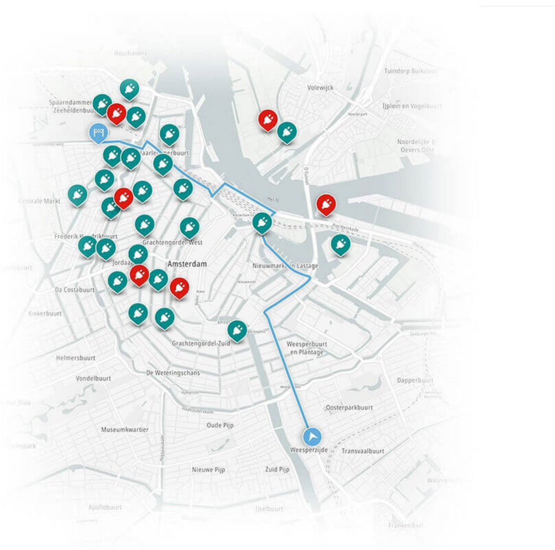


Answer
User
Query
Query Calibration Layer
Calibration
Model
Model Building
Query Processing
Artificial queries
Map
Topology, edge weights Metadata, trajectories, PoI
Map Making Layer
Map Matching
Map Making
Match or Make?
Metadata Inference
Edge Weights
PoI Match
Map Fuse
Map
Data Layer
Trajectories
Trips
PoI
Data Crawling
Data Wrangling
Data Sampling
Map Services APIs
Trajectories
PoI
Trips
Map

Data Layer

- Storage and retrieval infrastructure for Points of Interest (PoI), trip information, and complete trajectories
- Includes off-the-shelf spatial indexing methods
- Digests incoming live high traffic data
- Smart Data Crawling
  - Can bootstrap its data by crawling some map services with limited number of API calls that
    - Weekday/Weekend future days
    - Different times of the day
    - Short trips
    - PoI trips

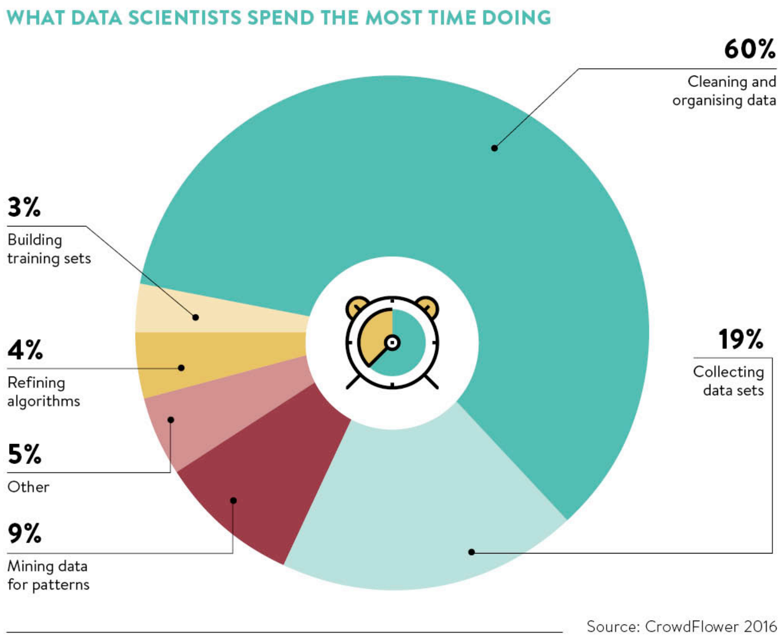




Rule-based Data Cleaning & Wrangling

- Existing efforts for data cleaning and wrangling do not support spatial and spatio-temporal data

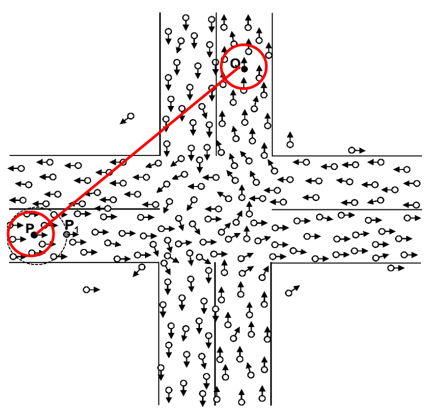
"After analyzing all your data, I think we can safely say that none of it is useful."

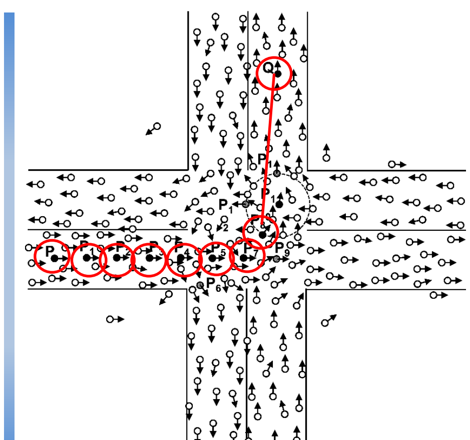


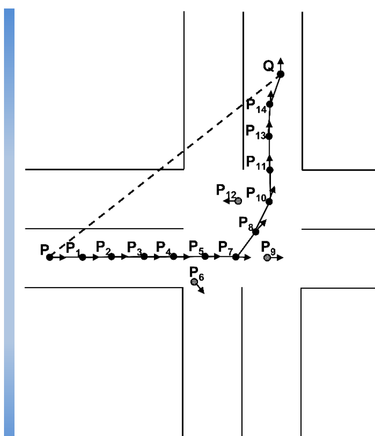
- Deployed Rules in QARTA
  - Trajectories with a stop
    - Split the trajectory
  - Unrealistic points
    - Remove the point
  - Missing points
    - Split the trajectory

Trajectory Imputation

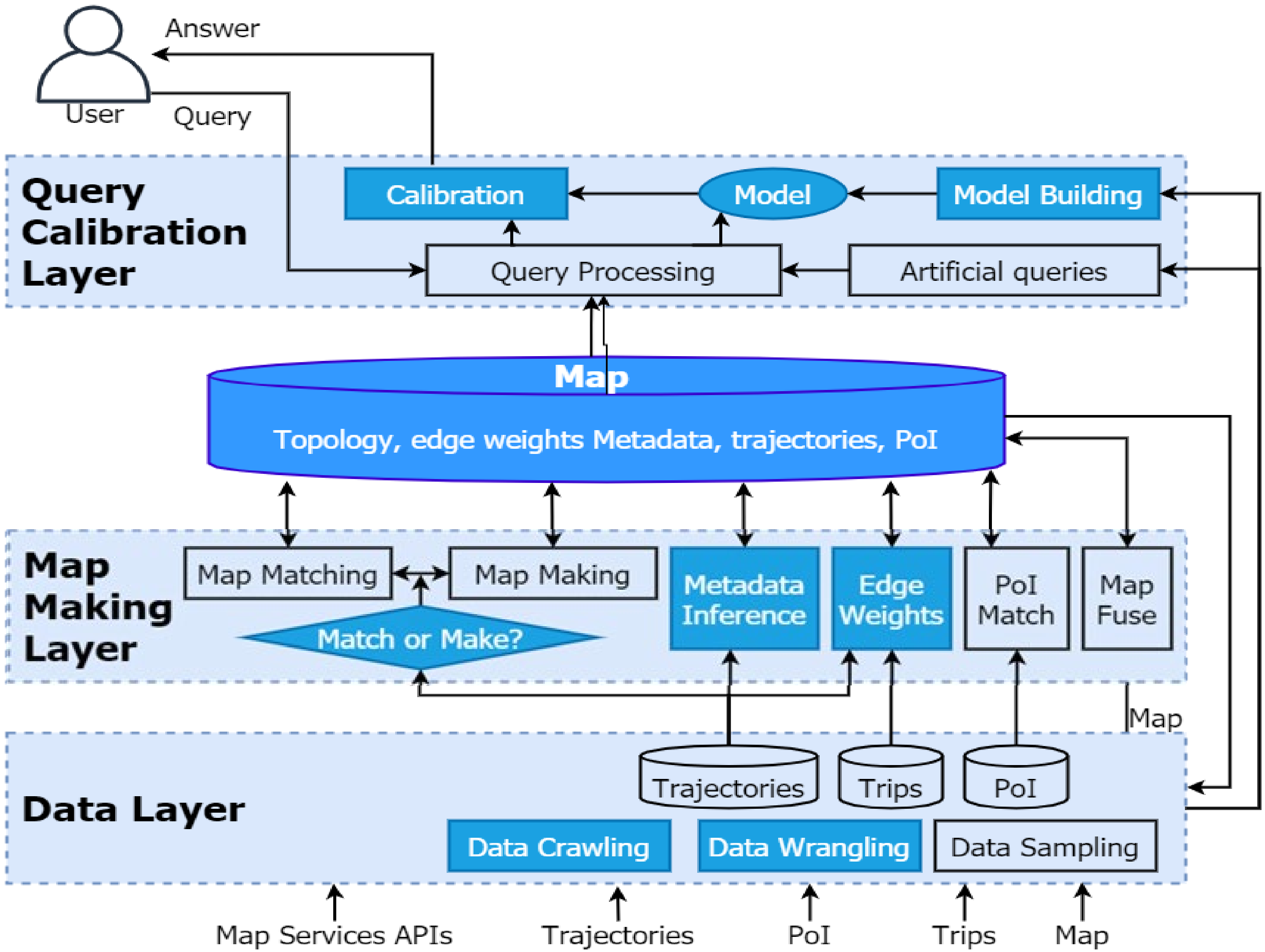
- Dense trajectory data is needed for further analysis
- Need to densify our trajectories
  - Use the wisdom of the crowd to impute each trajectory







Map Making Layer

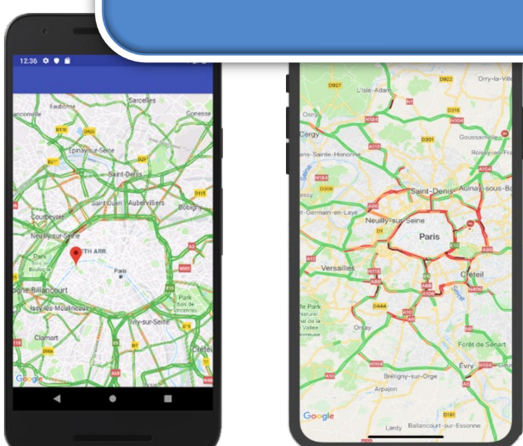


Edge Weight Inference: Who is doing it?

- Traffic departments: Loop detectors or plate recognition

Edge Weights are considered as proprietary information, not to be shared

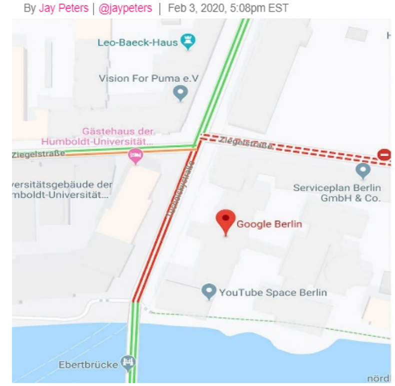
- Co



ORE
with
99 phones and a little red wagon

The streets were mostly empty, but the map showed a traffic jam

By Jay Peters | @jaypeters | Feb 3, 2020, 5:08pm EST



Edge Weight Inference in QARTA: High level Idea

- Learning Edge weights per time granularity (e.g., hour)
- **Input:** Trips (Pickup time/location, Drop off time/location)

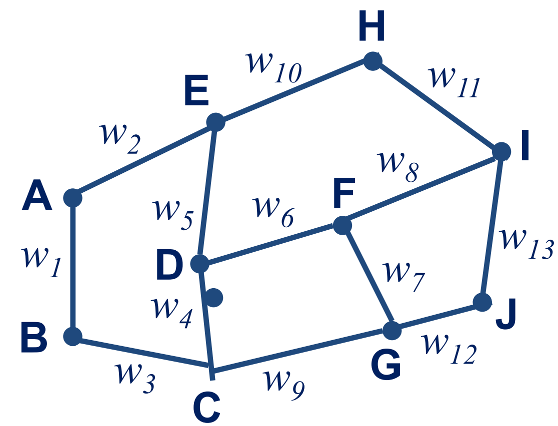
(A, F, 15) ➔ $w_2 + w_5 + w_6 = 15$

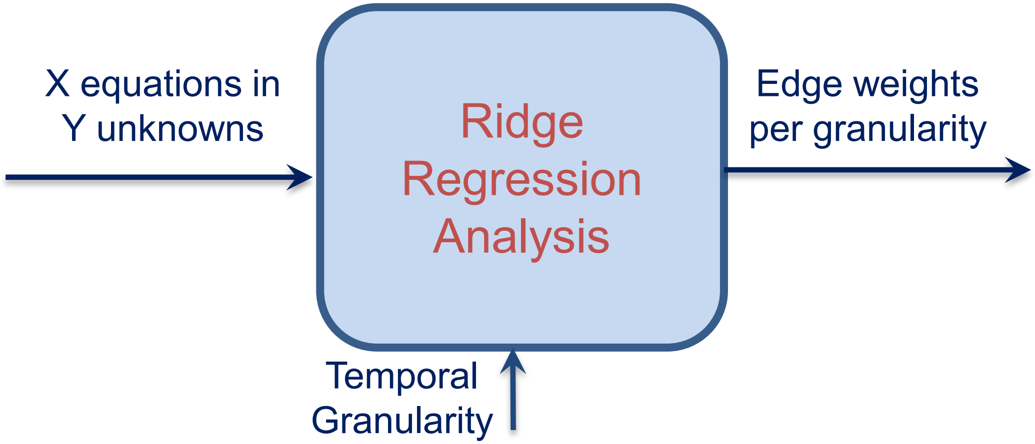
(B, H, 28) ➔ $w_3 + w_7 + w_8 + w_9 + w_{11} = 28$

(A, I, 19) ➔ $w_1 + w_3 + w_7 + w_8 + w_9 = 19$

…

…





Problem Formulation

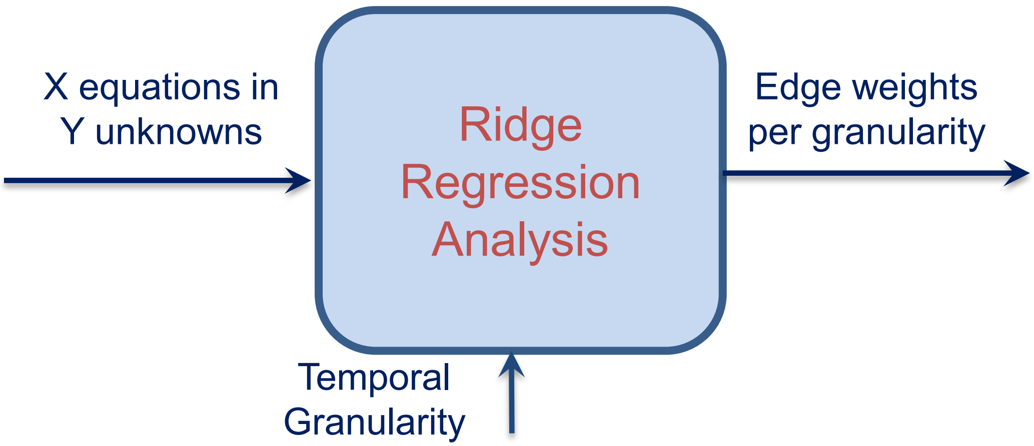
- **Input:** Trips (Pickup time/location, Drop off time/location)
  - The lowest common denominator of publicly available trajectory datasets
  - For each trip t, δ_t is the actual trip time as (Drop off time – Pickup time)
  - For each trip t, a trip path $P_t\,[\,e_0, e_1, \ldots..e_l\,]$ is a sequence of edges from an off-the-shelf routing engine
  - Each edge e has a *known* length l_e (in meters) and *unknown unit* edge weight W_e (sec/meter).
  - Time to travel through e is $W_e l_e$. *Time to go through* P_t *is* $\sum_{e \epsilon Pt} W_e l_e$
- **Objective:** Given a set of trips T, find the weights W_e that would minimize:

$$\sum_{t \epsilon T} \left(\sum_{e \epsilon Pt} W_e l_e - \delta_t \right)^2$$

Basic Solution

- **Objective:** Given a set of trips T, find the weights W_e that would minimize:

$$\sum_{t \epsilon T} \left(\sum_{e \epsilon Pt} W_e l_e - \delta_t \right)^2$$



- **Challenges:**
  - A direct solution may result in zero or negative weights
  - Scalability is a major issue: Hundreds of thousands of edges with millions of trajectories
  - Over-fitting for unreliable edges

Tuning Step 1: Heavy Edges Inference

- **Problem:** Each edge in the graph, regardless of its popularity, acts as a regression feature ➔ Over-fitting & unnecessarily expensive computations.
- **Idea:** Distinguish between heavy (popular) edges *H*, where high accuracy is needed and light edges. *H* includes the *top-k* edges in terms of number of trips covering them (default K=10,000).
- **Objective:** All light edges have the same weight W_0 . Need to only find weights of heavy edges:

$$\sum_{t \epsilon T}\left(\sum_{e \,\epsilon\, (P_t \cap H)} W_e l_e + W_0 \sum_{e \,\epsilon\, (Pt \setminus H)} l_e - \delta_t\right)^2$$

- Number of regression features is reduced by two orders of magnitude ➔ higher scalability and less over fitting

Tuning Step 2: Heavy Road Detection

- **Problem:** A long road is usually represented by multiple connected edges, that are highly likely to have the same weight

- **Idea:** Group each of such edges together as one Heavy Road with one weight W_g. Split heavy edges H into r disjoint sets H_1 to H_r

- **Objective:**

$$\sum_{t \epsilon T} \left(\sum_{g : P_t \cap Hg \neq \emptyset} W_g L_g + W_0 \sum_{e \, \epsilon \, (Pt \setminus H)} l_e - \delta_t \right)^2 , \qquad L_g = \sum_{e \, \epsilon \, Hg} l_e$$

- Number of unknowns is reduced to $r+1$, number of model features is reduced by 75% ➔ higher scalability

Tuning Step 3: Enforcing Physical Constraints

- To avoid having $W_g \leq 0$, we use Ridge regression regularization, where we add a regularization term that penalizes weights deviating from the average speed:

- **Objective:**

Regularization strength

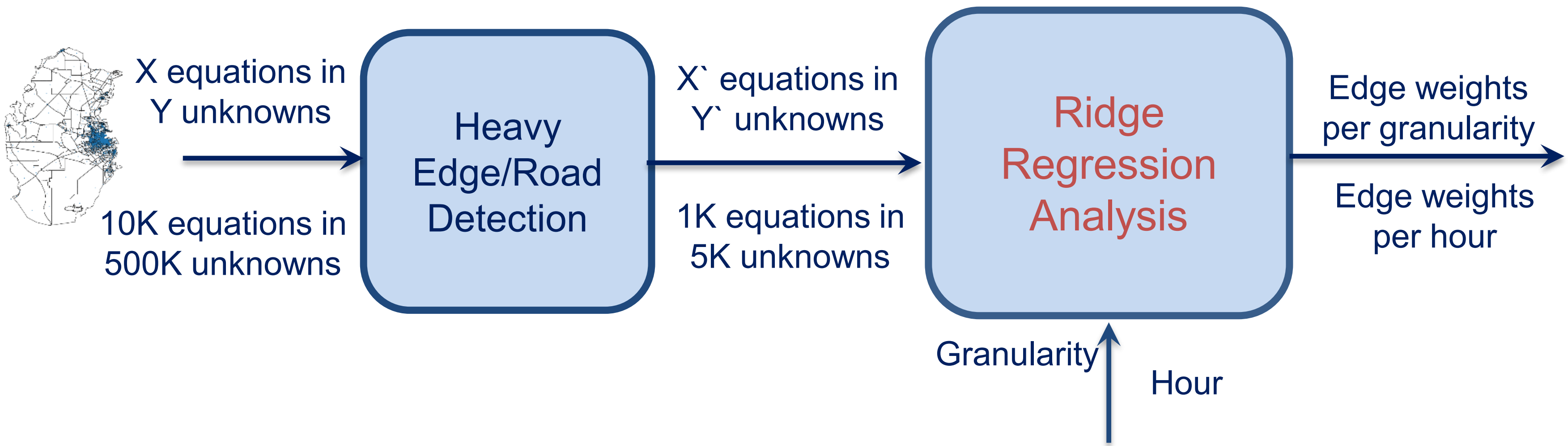
Average speed

$$\sum_{t \epsilon T} \left(\sum_{g: P_t \cap Hg \neq \emptyset} W_g L_g + W_0 \sum_{e \,\epsilon\, (Pt \setminus H)} l_e - \delta_t \right)^2 + \alpha \sum_g (W_g - \sigma)^2$$

$$L_g = \sum_{e \,\epsilon\, Hg} l_e$$

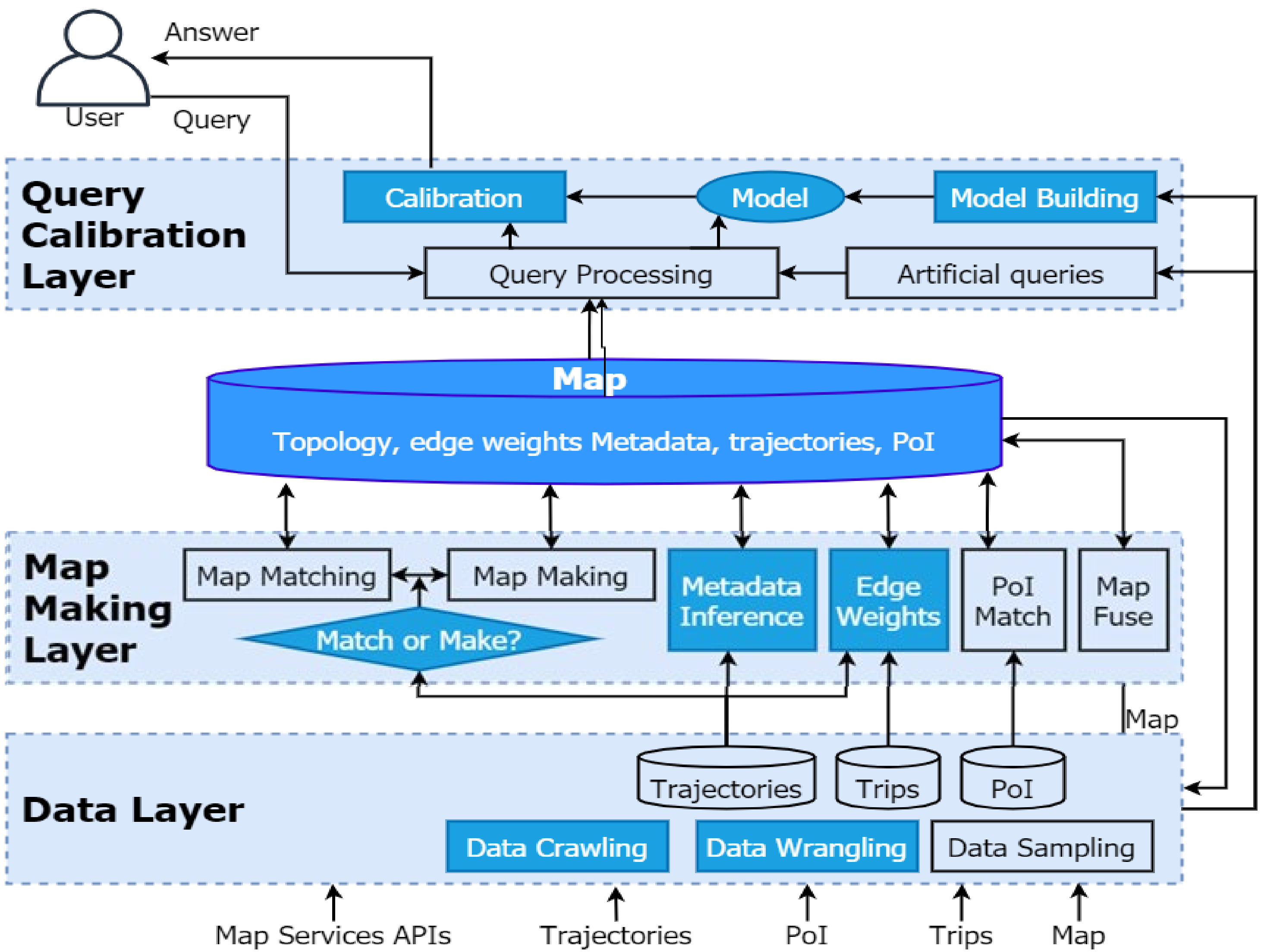
Edge Weight Inference in QARTA

- Divide all trips into time granularity (e.g., one hour)
- Solve the regression equation per time granularity



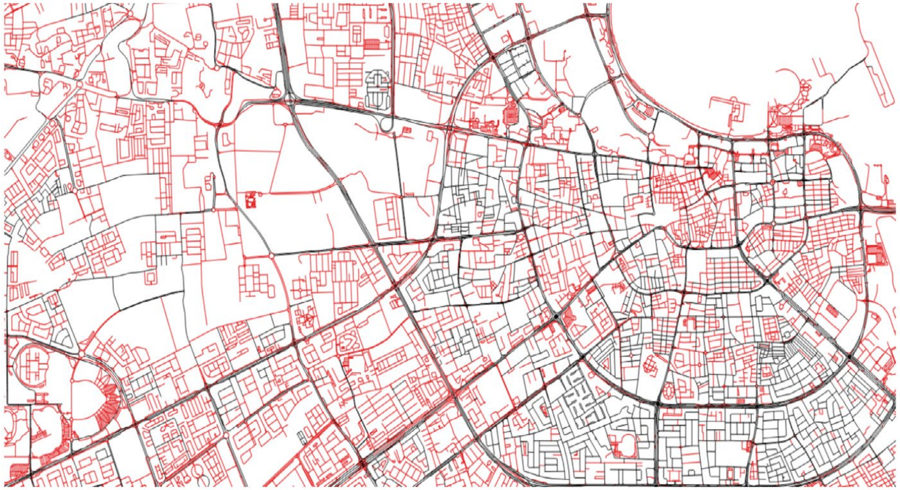
- More than 99% of the roads satisfy the physical constraint. For those that do not, we set edge weight to the minimum possible value (*1 / maxspeed)*

Map Making Layer



Metadata Inference

- Need rich metadata (annotation) for road networks
  - Speed limit
  - Number of lanes
  - Road type
  - Traffic lights
  - Safety
  - Eco-friendly



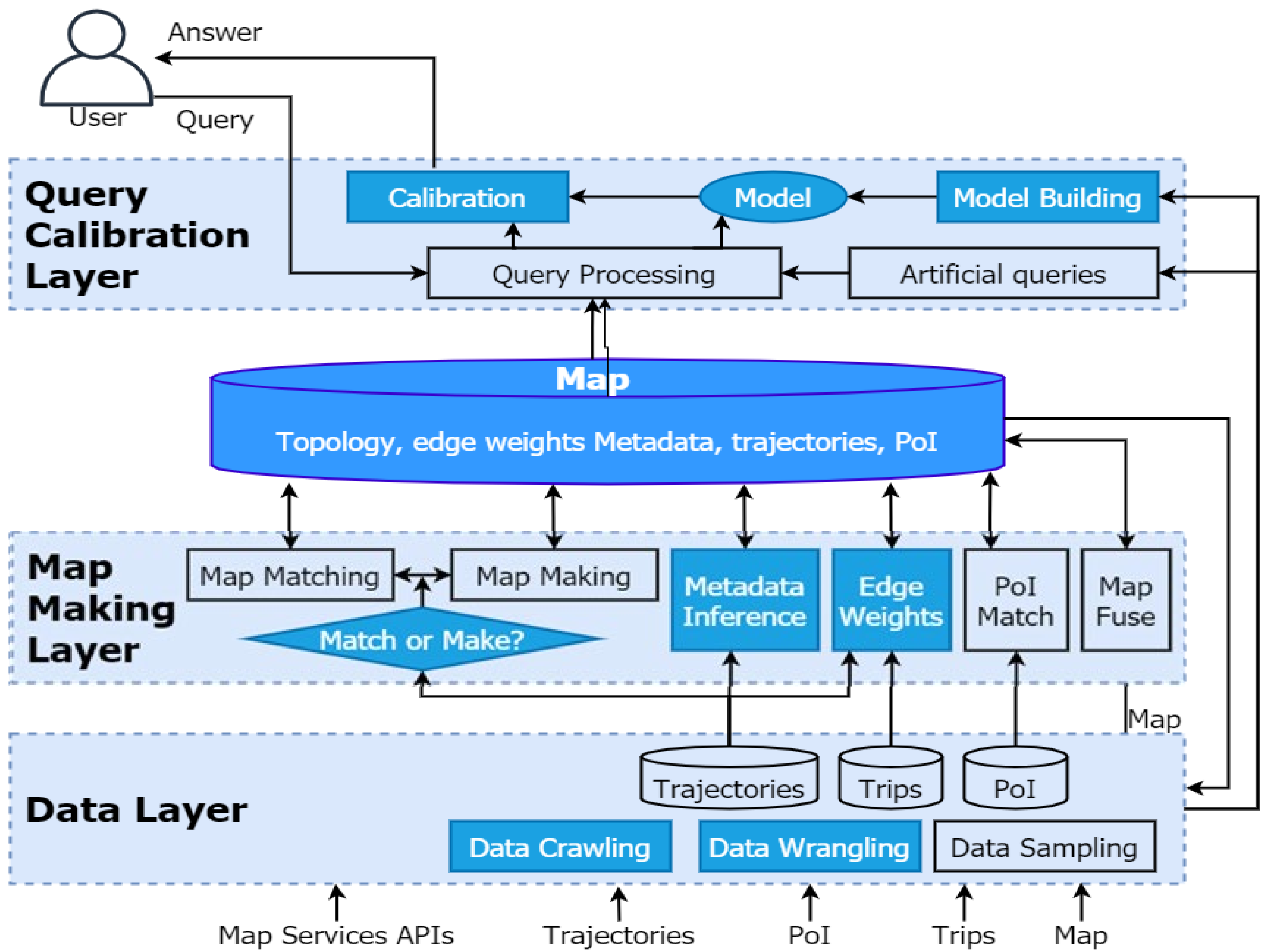
- Public maps have very poor metadata coverage
- Metadata inference in QARTA is framed as a supervised learning problem
  - Step 1: Find the best models that would map road features to certain metadata
  - Step 2: Use these models to predict the missing metadata values

Metadata Inference

- Feature Engineering
  - Structural features (one per road)
    - Road length, Number of in/out junctions, Road curvature
  - Functional features (one per time granularity per road)
    - Speed average, Speed standard deviation, Density, Distance to center line

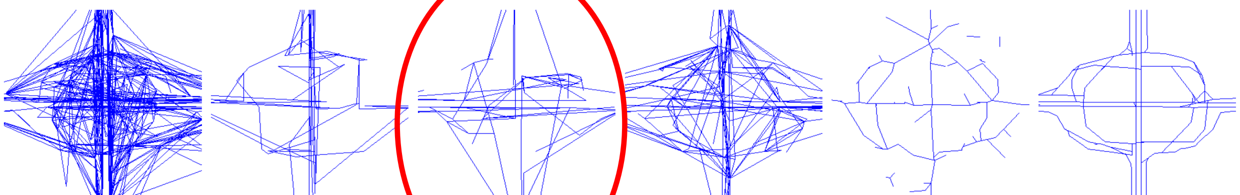
- Learning Kernel
  - A list <*Vi, Li*> *f*or each metadata *L* and road segments with known *L*
  - Partition the list to: *Training*, *Validation*, and *Testing* datasets
  - We experiment with different machine learning techniques
    - Logistic regression, support vector machines, random forests, boosting gradients, and deep neural networks,
  - Models are stored with their Key Performance Indicators

Map Making Layer



Map Making

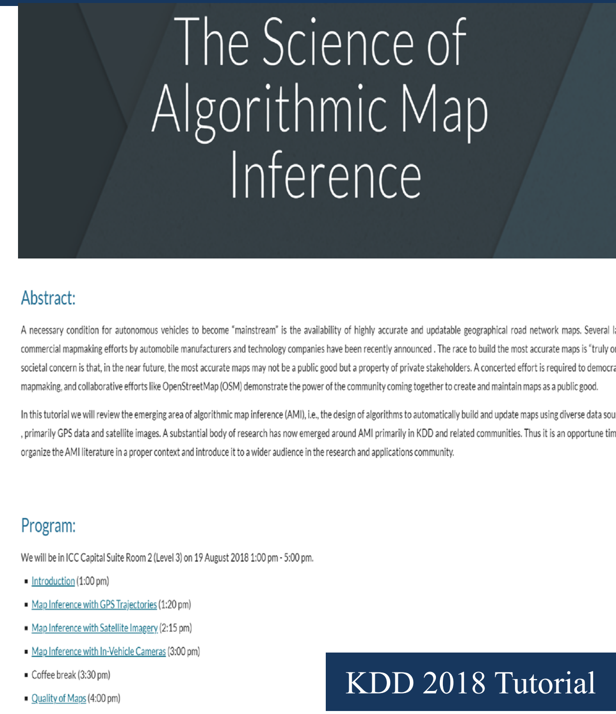
- QARTA deploys off-the-shelf techniques for constructing map topology from GPS traces



(a) Edelkamp. (b) Chen. (c) Kharita. (d) Cao. (e) Biagioni. (f) DeepMG.

Figure 5: Seven algorithm at a roundabout near Gongzhufen area in TaxiBJ.

Sijie Ruan, *et al.* "***Learning to Generate Maps from Trajectories***". AAAI 2020



The Science of Algorithmic Map Inference

Abstract:

A necessary condition for autonomous vehicles to become "mainstream" is the availability of highly accurate and updatable geographical road network maps. Several large commercial mapmaking efforts by automobile manufacturers and technology companies have been recently announced . The race to build the most accurate maps is "truly on." A societal concern is that, in the near future, the most accurate maps may not be a public good but a property of private stakeholders. A concerted effort is required to democratize mapmaking, and collaborative efforts like OpenStreetMap (OSM) demonstrate the power of the community coming together to create and maintain maps as a public good.

In this tutorial we will review the emerging area of algorithmic map inference (AMI), i.e., the design of algorithms to automatically build and update maps using diverse data sources , primarily GPS data and satellite images. A substantial body of research has now emerged around AMI primarily in KDD and related communities. Thus it is an opportune time to organize the AMI literature in a proper context and introduce it to a wider audience in the research and applications community.

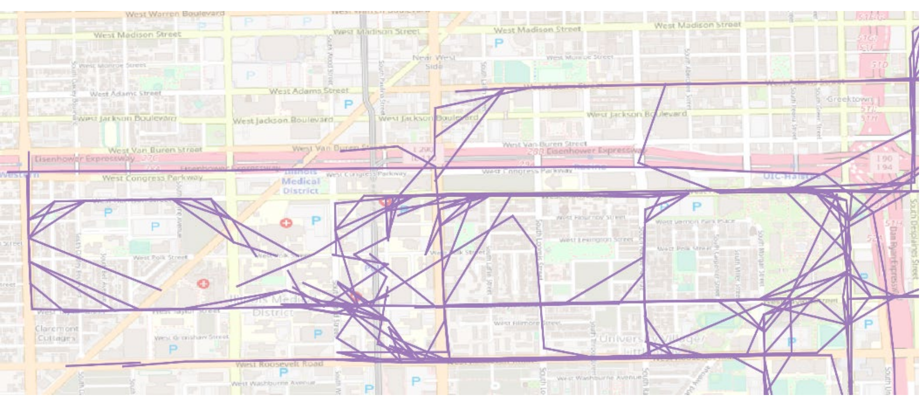
Program:

We will be in ICC Capital Suite Room 2 (Level 3) on 19 August 2018 1:00 pm - 5:00 pm.

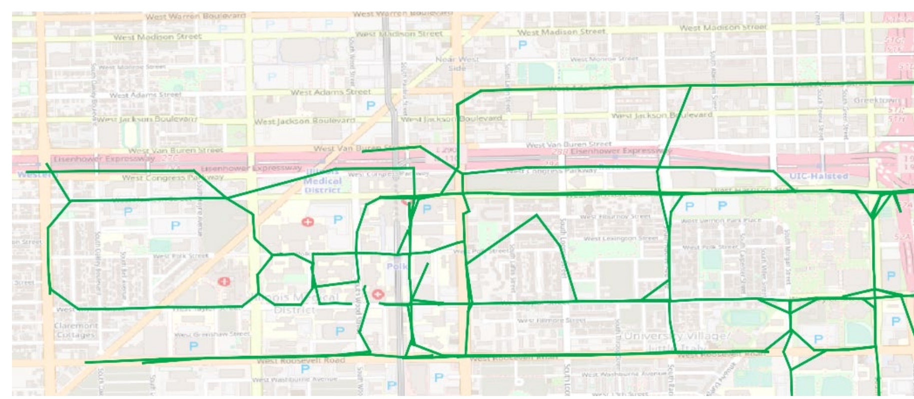
- Introduction (1:00 pm)
- Map Inference with GPS Trajectories (1:20 pm)
- Map Inference with Satellite Imagery (2:15 pm)
- Map Inference with In-Vehicle Cameras (3:00 pm)
- Coffee break (3:30 pm)
- Quality of Maps (4:00 pm)

KDD 2018 Tutorial

- Trajectory Imputation significantly boost the accuracy



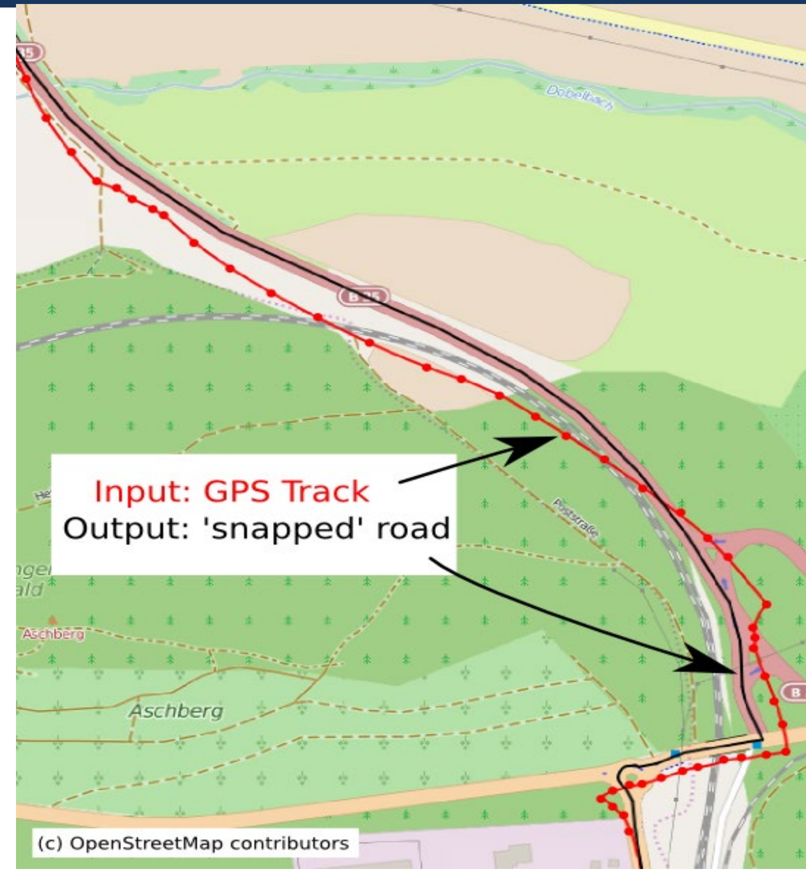
Kharita with no Trajectory Imputation



Kharita + Trajectory Imputation

Map Matching

- Classical topic in transportation
- QARTA deploys off-the-shelf techniques for
  - Snapping GPS tracks on underlying road network
  - Removing outlier GPS points that do not match the road network



SURVEY

A Survey of Measures and Methods for Matching Geospatial Vector Datasets

Authors: Emerson M. A. Xavier, Francisco J. Ariza-López, Manuel A. Ureña-Cámara

Authors Info & Affiliations

Publication: ACM Computing Surveys • August 2016 • Article No.: 39 • https://doi.org/10.1145/2963147

ADC 2020: Databases Theory and Applications pp 121-133 | Cite as

A Survey on Map-Matching Algorithms

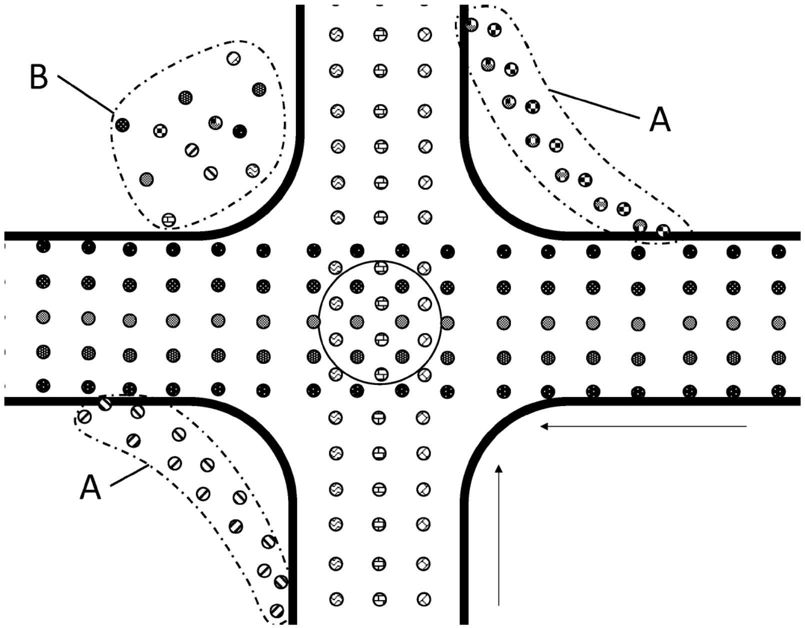
Authors | Authors and affiliations

Pingfu Chao, Yehong Xu, Wen Hua, Xiaofang Zhou

Match or Make

- Map Making implicitly assumes that GPS traces are correct

- Map Matching implicitly assumes that the road network is accurate



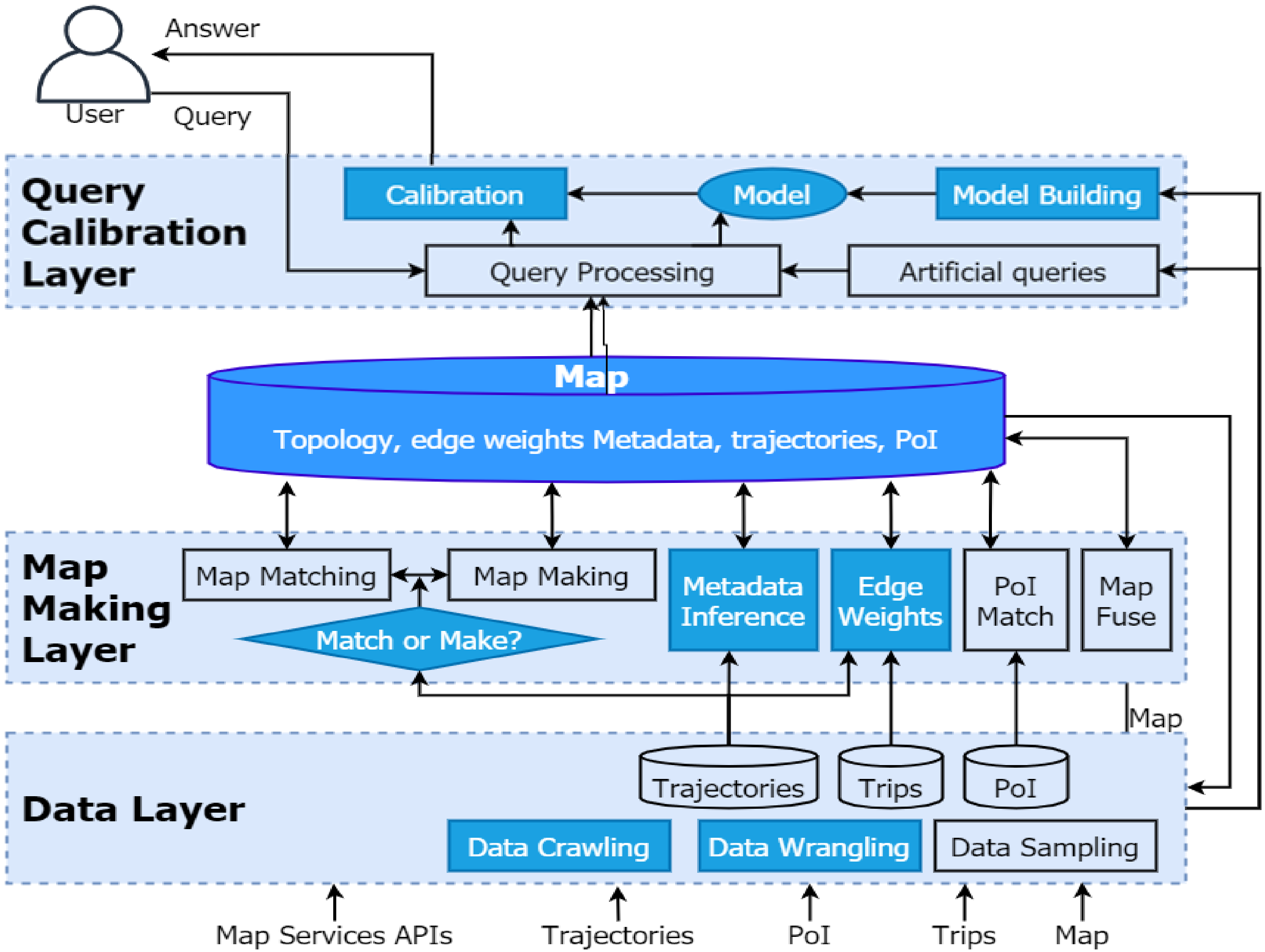
- Example: A roundabout that is recently converted to a bridge with two new exits:
  - Applying Map making on GPS traces would make new roads for points A and B
  - Applying Map Matching would snap points A and B to a wrong map

Match or Make

- Given a road network *R* and GPS points *P*, decide on the part(s) of the map where *R* is more accurate than *P* and vice versa.
  - For parts where *R* is more accurate ➔ Do *Map Matching* to match *P* on *R*.
  - For parts where *P* is more accurate ➔ Do *Map Making* from *P*

- Idea: Use ML supervised learning to learn the features of accurate maps and points, then use it to classify the map and points
  - Step 1: Finding accurate points and roads
    - Those that match perfectly
  - Step 2: Inject various forms of errors in some correct roads and points
    - Remove a segment, shift road/point coordinates, reduce road resolution
  - Step 3: Feature extraction & model building
    - Build ML classifier model that maps point features to how good/bad it is.
  - Step 4: Decide to Match or Make
    - Use the developed ML model to decide about the points we are doubtful about

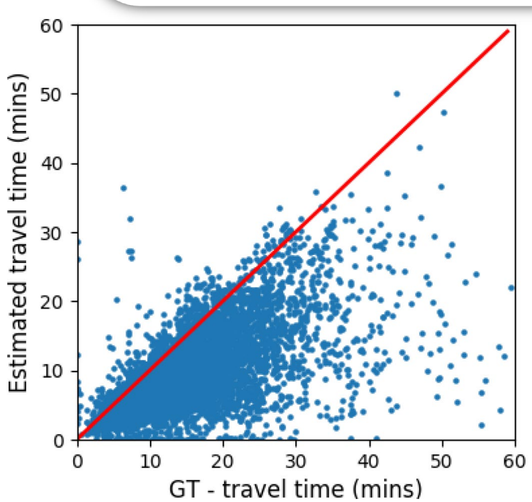
Query Calibration Layer



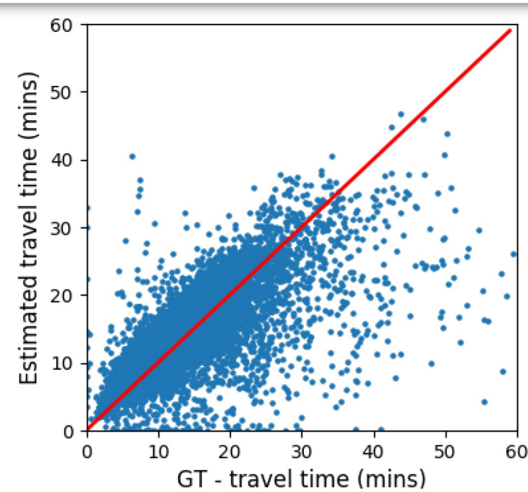
Estimated Time of Arrival (ETA)

- The accuracy of query answers heavily rely on Estimated time of Arrival

Idea: Can we study the error patterns of each algorithm under various context, and use to adjust the query answer.



OSRM



Google Maps

VB

Uber taps ClimaCell to improve ETA estimates with hyper-local weather data

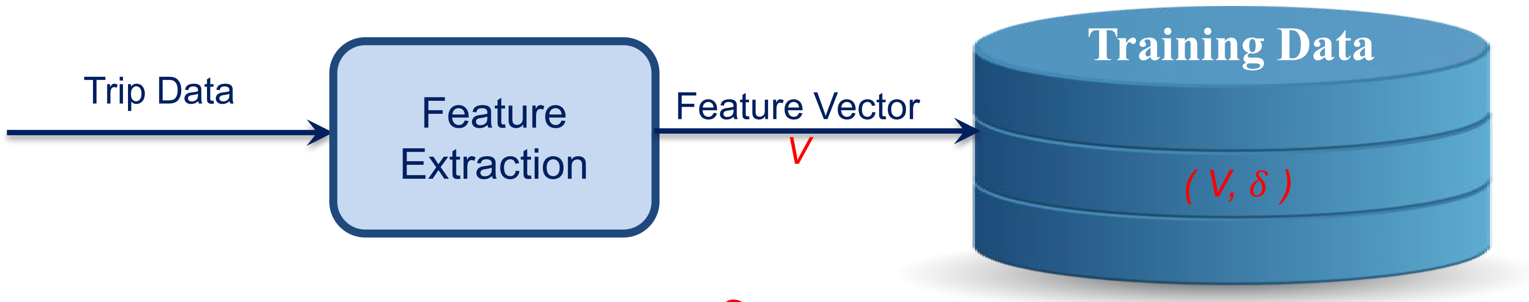
PAUL SAWERS @PSAWERS FEBRUARY 6, 2020 8:00 AM

Uber is partnering with weather technology company ClimaCell to enable more accurate estimated time of arrival (ETA) predictions for drivers and riders.

Founded in 2016, Boston-based ClimaCell specializes in real-time weather forecasts. Rather than relying on government data typically garnered from

Model Building: Feature Engineering

- Trip: (Pickup time/location, Drop off time/location, δ)
  - δ is the difference between actual and estimated time of the trip

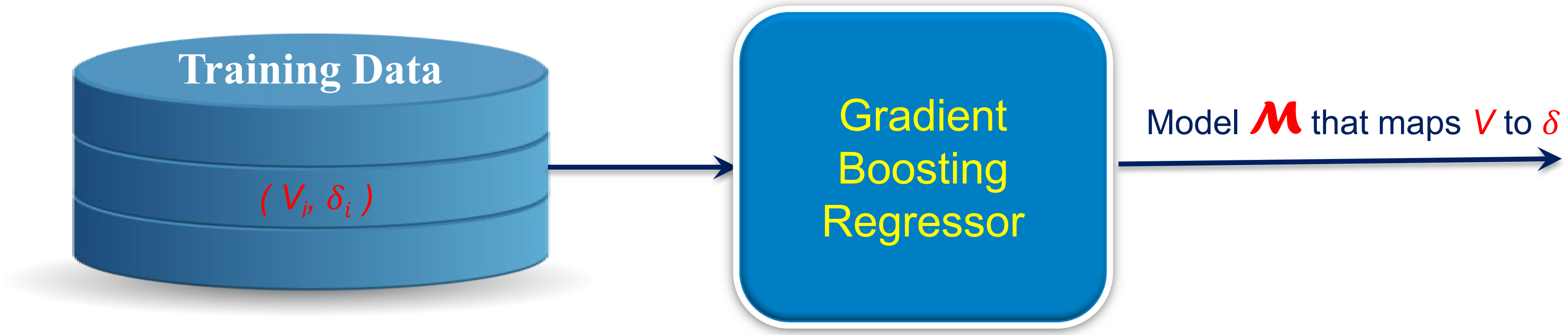


- Features in V that would impact δ
  - Spatial Zoning
    - Origin
    - Destination
  - Temporal Zoning
    - Pickup time
    - Drop off time
  - Trip Characteristics
    - Trip distance
    - Trip duration

Model Building: Training

- Need to find a function *F* that maps V_i *to* δi *while minimizing a loss function L for all trips in T*

$$\frac{1}{|T|} \times \sum_{t_i \in T} L(\delta_i, F(Vi))$$

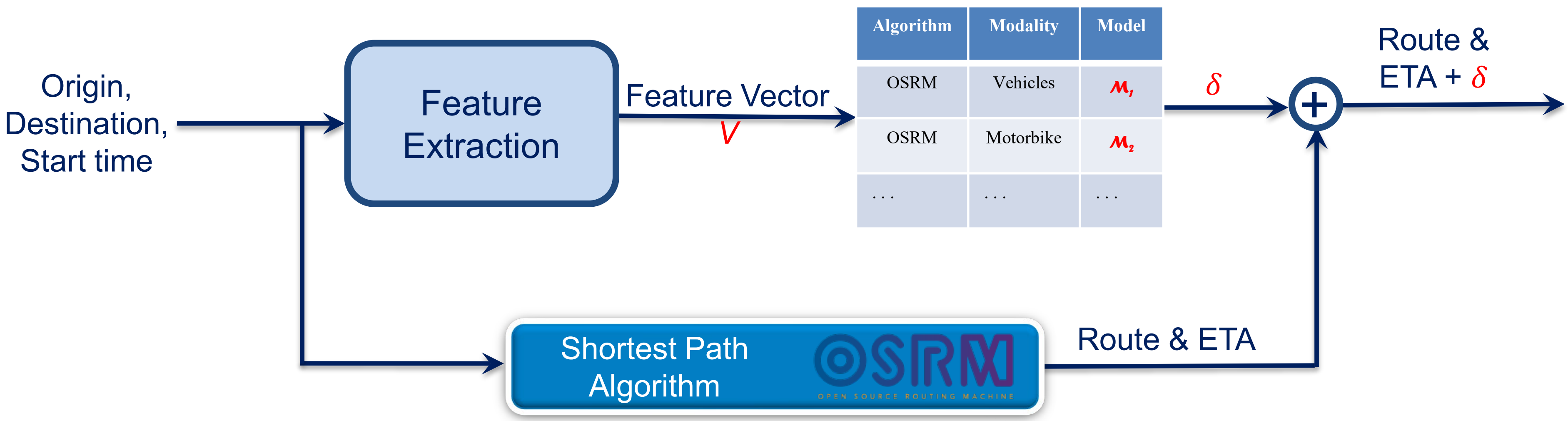


- A model $\mathcal{M}$ will be built for each ETA algorithm and driving modality →

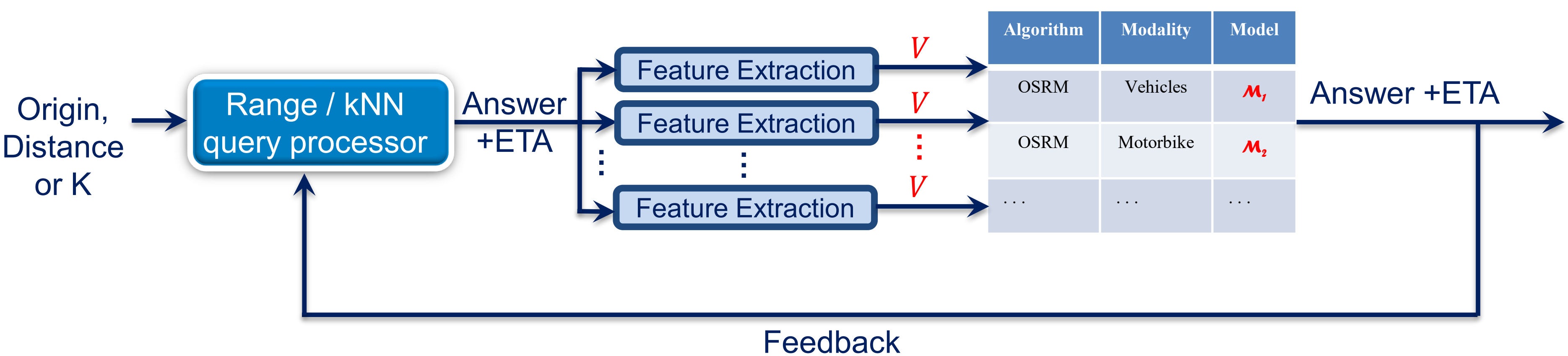
Algorithm	Modality	Model
OSRM	Vehicles	$\mathcal{M}_1$
OSRM	Motorbikes	$\mathcal{M}_2$
. . .	. . .	. . .

Query Calibration in QARTA

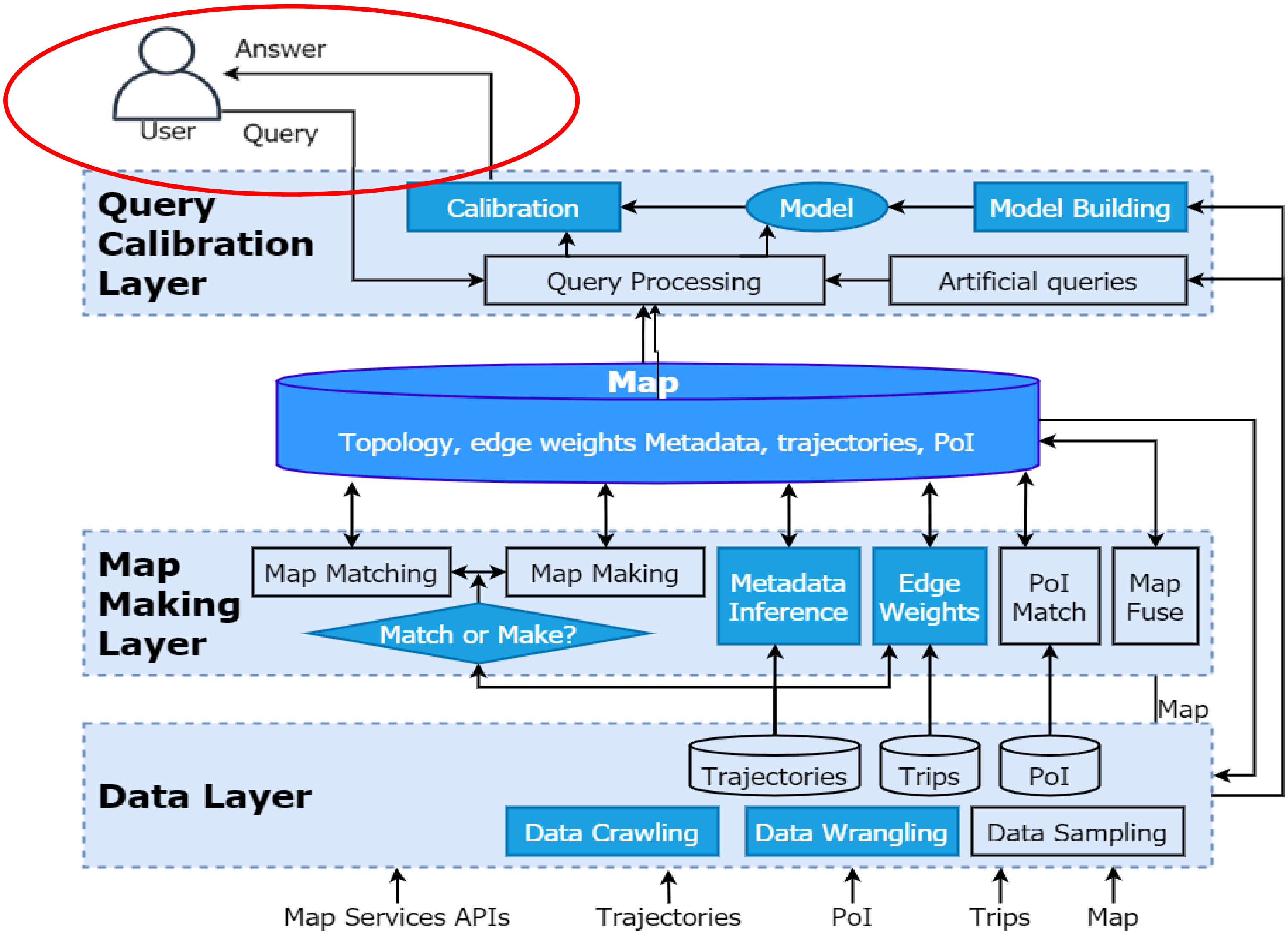
- Shortest Path queries



- Range and kNN queries



User Interface & Performance Evaluation



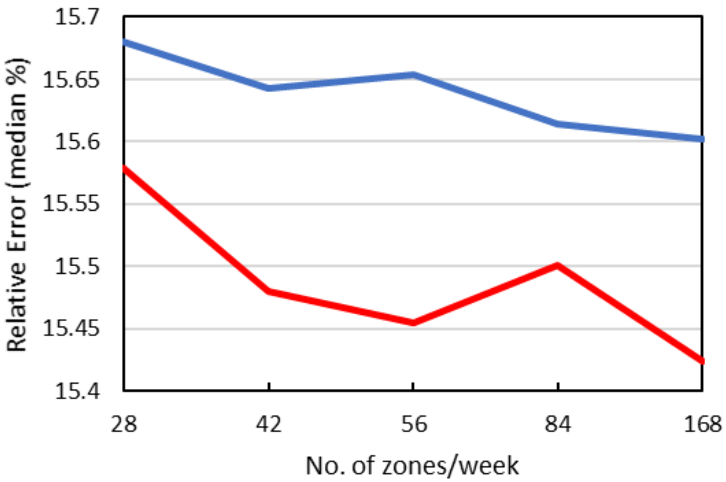
Experimental Evaluation

- Based on real deployment of QARTA
  - All taxis in Qatar $\sim 4K$ vehicles
  - A local food delivery company $\sim 3K$ motorbikes
  - ~235K daily API calls
  - ~1 $Million$ daily GPS tracks

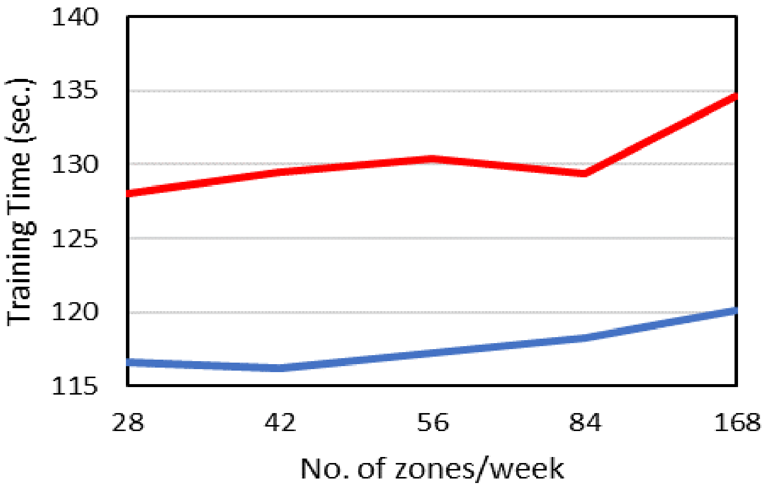
- ML Model Building
  - 250K trips: 75% for building the model, 25% for testing

- Underlying Algorithms
  - OSRM for shortest path
  - OSM Map
  - Off-the-shelf algorithms for range and k-NN queries

Spatial and Temporal Zoning for Edge Weight Inference



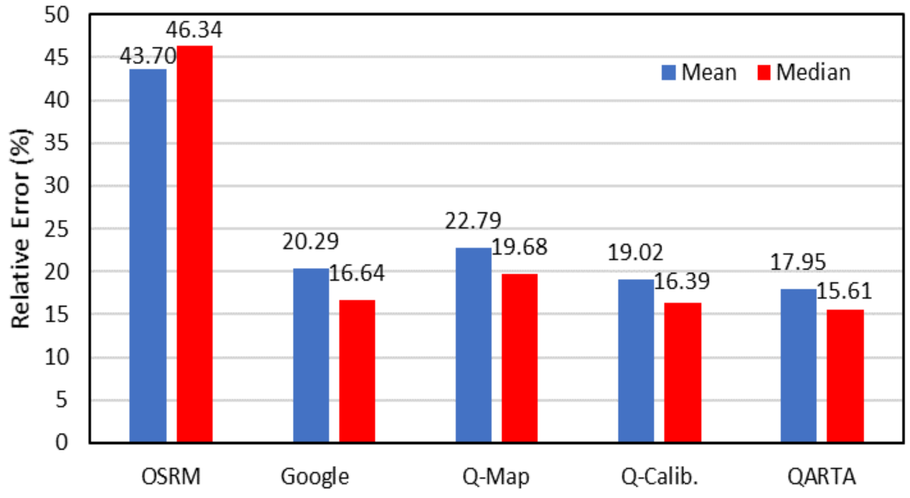
Administrative Zones (Sparse)



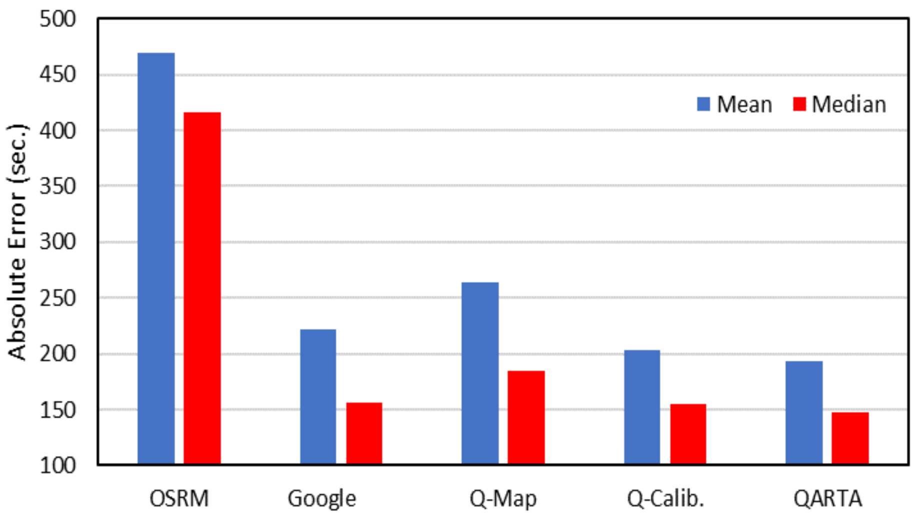
Transportation Zones (Dense)

QARTA vs Other Map Services: Shortest Path Query

- Q-Map: Runs QARTA Map Making layer without any calibration
  - OSRM on QARTA map

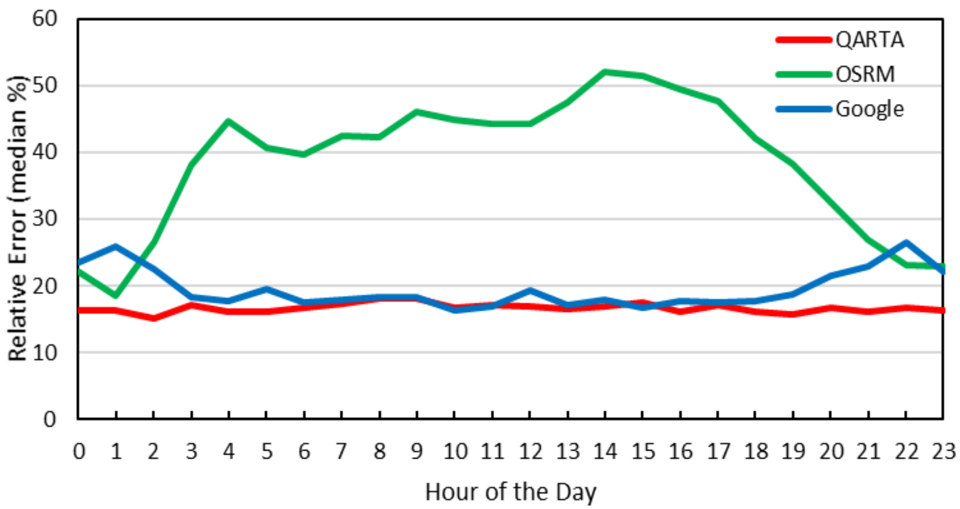


- Q-Calib: Runs QARTA calibration without Map Making layer
  - Calibrating OSRM engine

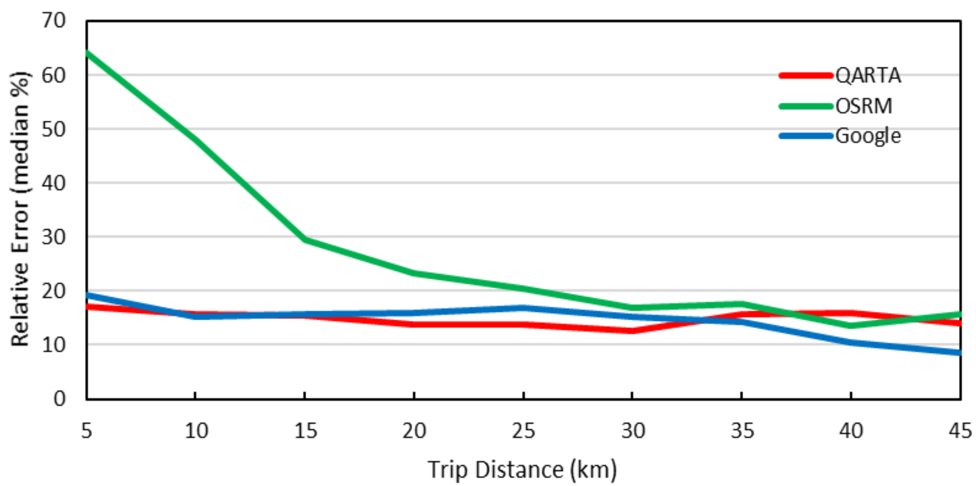


QARTA vs Other Map Services: Shortest Path Query

- Daytime trips are the ones affected by traffic, and QARTA is consistently better

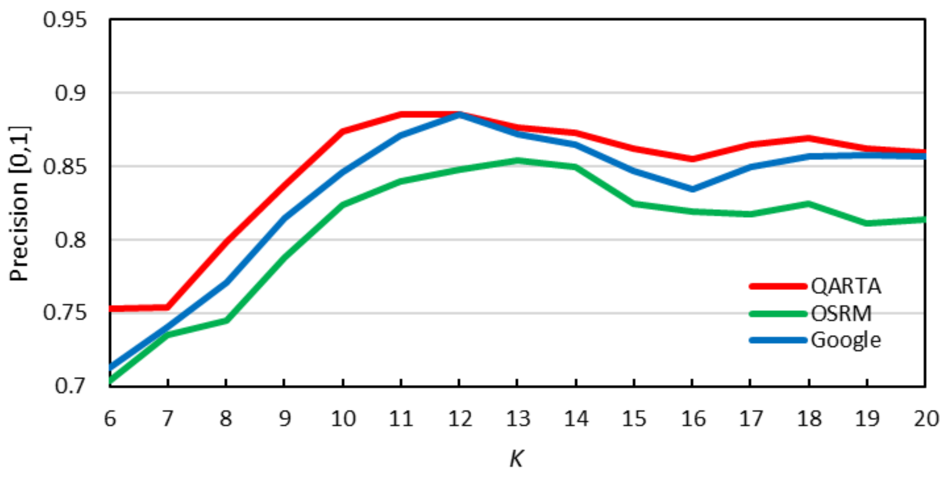


- Shorter trips are mostly affected by traffic
- Number of long trips is low, hence QARTS is not doing good
- In Doha, a 30KM trip is too long, across the whole city

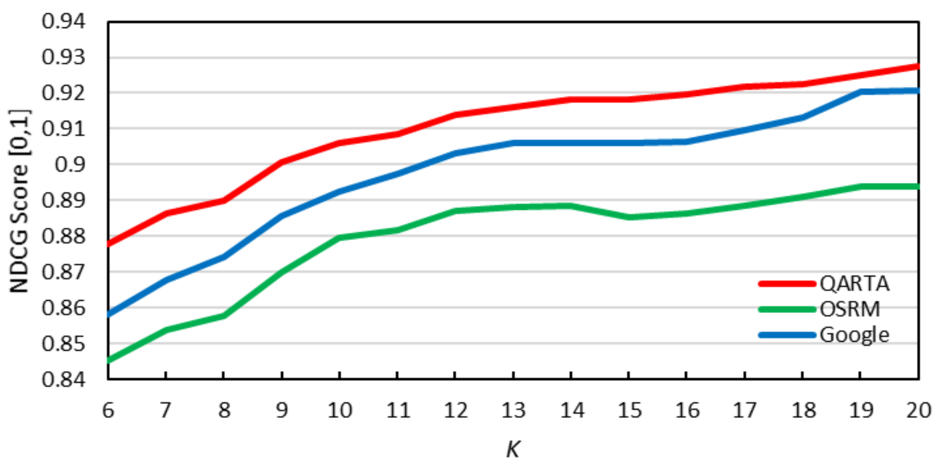


QARTA vs Other Map Services: *K*NN Query

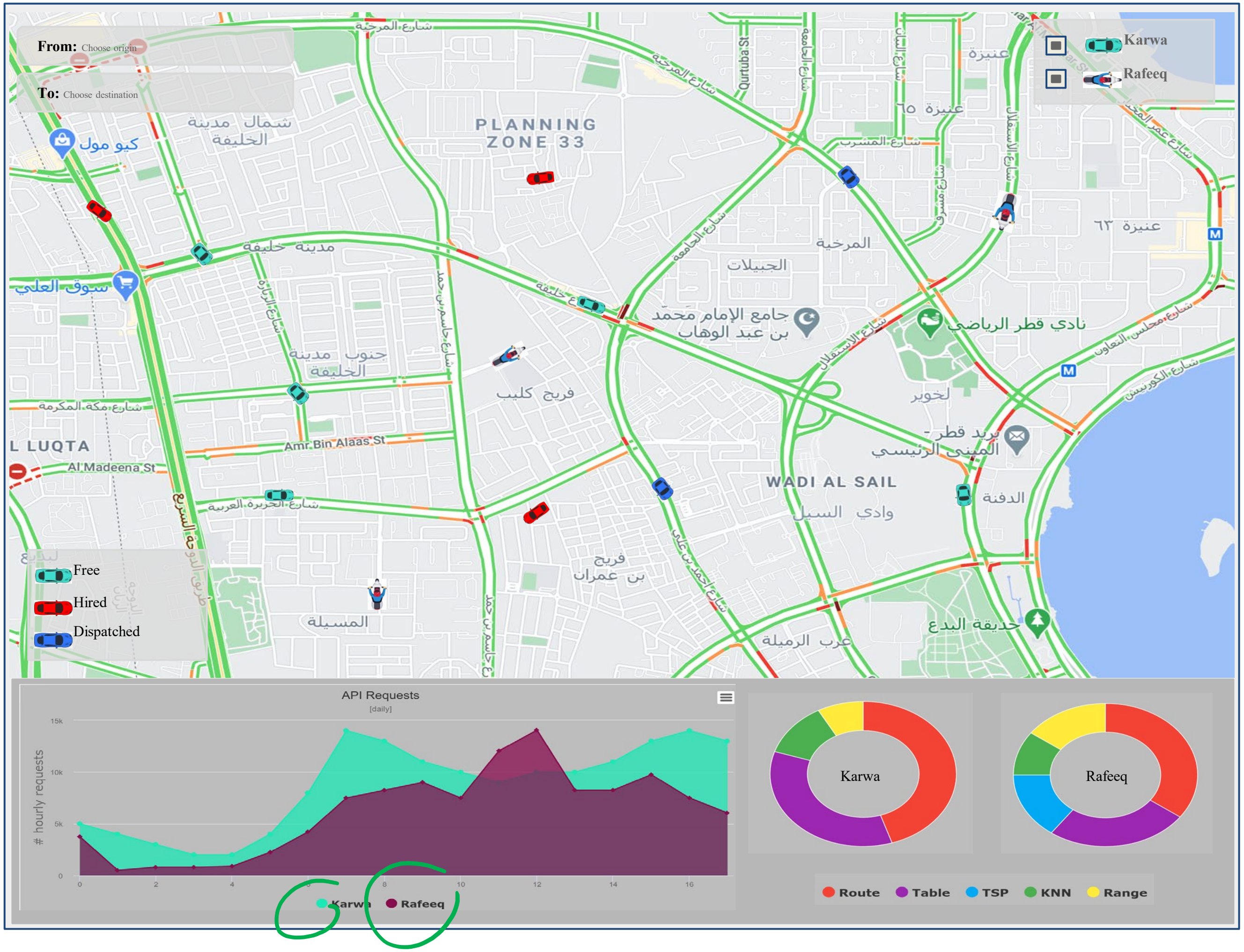
- *Precision*: Number of items in *K*NN list that overlap with ground truth
  - All very similar performance



- *NDCG*: A ranking quality measure that takes into account the order if items in the list



QARTA Dashboard



QARTA Admin Panel

QARTA MAP

Run Every : 7 days

Nb. Models : 168

Training Win : 30 days

Last Run: 14-03-2021 16:24:51 [OK]

QARTA Calibration

Run Every : 1 day

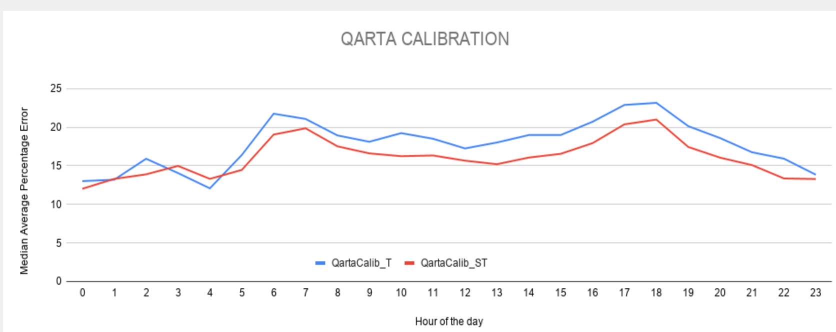
Training Win : 7 days

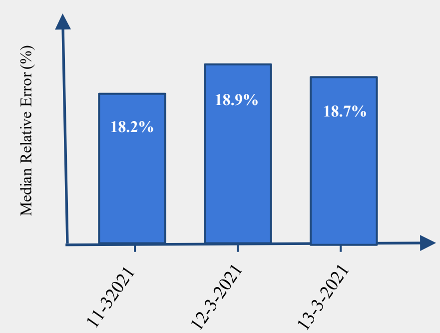
Temp. Partition : 24

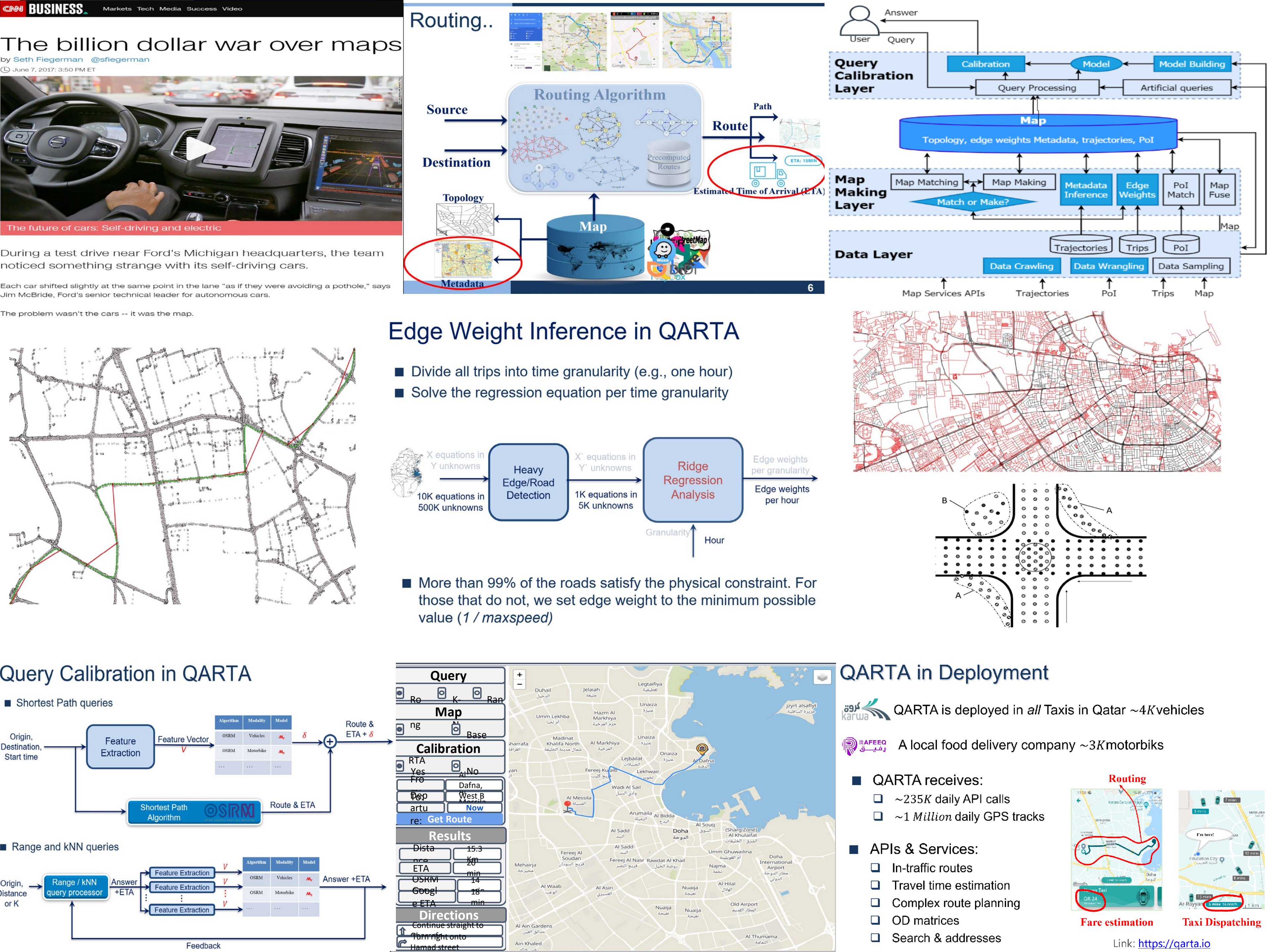
Spatial Partition : Upload shape file

Last Run: 12-03-2021 01:15:38 [OK]

Component: QARTA Cal **Date:** 14-03-2021 **Metric:** MAPE






CNN BUSINESS
Markets Tech Media Success Video
The billion dollar war over maps
by Seth Fiegerman @sfiegerman
June 7, 2017: 3:50 PM ET
The future of cars: Self-driving and electric
During a test drive near Ford's Michigan headquarters, the team noticed something strange with its self-driving cars.
Each car shifted slightly at the same point in the lane "as if they were avoiding a pothole," says Jim McBride, Ford's senior technical leader for autonomous cars.
The problem wasn't the cars -- it was the map.
Routing..
Routing Algorithm
Source
Destination
Path
Route
Precomputed Routes
Estimated Time of Arrival (ETA)
ETA: 15MIN
Topology
Map
Metadata
6
Answer
User
Query
Query Calibration Layer
Calibration
Model
Model Building
Query Processing
Artificial queries
Map
Topology, edge weights Metadata, trajectories, PoI
Map Making Layer
Map Matching
Map Making
Match or Make?
Metadata Inference
Edge Weights
PoI Match
Map Fuse
Map
Data Layer
Trajectories
Trips
PoI
Data Crawling
Data Wrangling
Data Sampling
Map Services APIs
Trajectories
PoI
Trips
Map
Edge Weight Inference in QARTA
Divide all trips into time granularity (e.g., one hour)
Solve the regression equation per time granularity
X equations in Y unknowns
10K equations in 500K unknowns
Heavy Edge/Road Detection
X` equations in Y` unknowns
1K equations in 5K unknowns
Ridge Regression Analysis
Edge weights per granularity
Edge weights per hour
Granularity
Hour
More than 99% of the roads satisfy the physical constraint. For those that do not, we set edge weight to the minimum possible value (1 / maxspeed)
B
A
A
Query Calibration in QARTA
Shortest Path queries
Origin, Destination, Start time
Feature Extraction
Feature Vector V
Algorithm Modality Model
OSRM Vehicles
OSRM Motorbike
δ
Route & ETA + δ
Shortest Path Algorithm
OSRM
Route & ETA
Range and kNN queries
Origin, Distance or K
Range / kNN query processor
Answer +ETA
Feature Extraction
Answer +ETA
Feedback
Query
Map
Calibration
Get Route
Results
Directions
Doha
QARTA in Deployment
QARTA is deployed in all Taxis in Qatar ~4K vehicles
A local food delivery company ~3K motorbiks
QARTA receives:
~235K daily API calls
~1 Million daily GPS tracks
APIs & Services:
In-traffic routes
Travel time estimation
Complex route planning
OD matrices
Search & addresses
Routing
Fare estimation
Taxi Dispatching
Link: https://qarta.io


CNN BUSINESS Markets Tech Media Success Video

The billion dollar war over maps

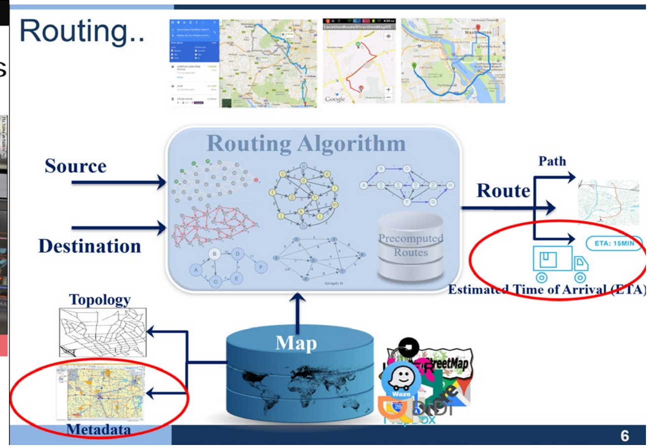
by Seth Fiegerman @sfiegerman
June 7, 2017: 3:50 PM ET

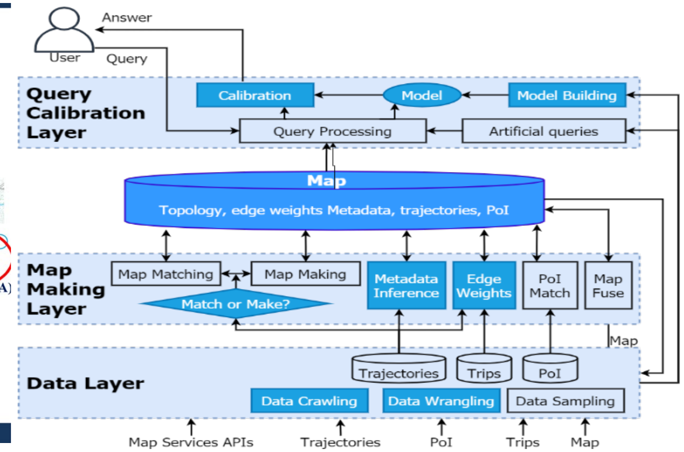
The future of cars: Self-driving and electric

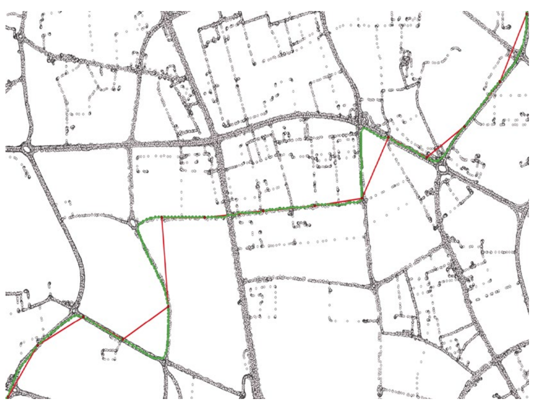
During a test drive near Ford's Michigan headquarters, the team noticed something strange with its self-driving cars.

Each car shifted slightly at the same point in the lane "as if they were avoiding a pothole," says Jim McBride, Ford's senior technical leader for autonomous cars.

The problem wasn't the cars -- it was the map.

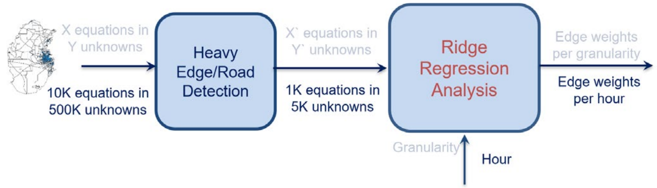




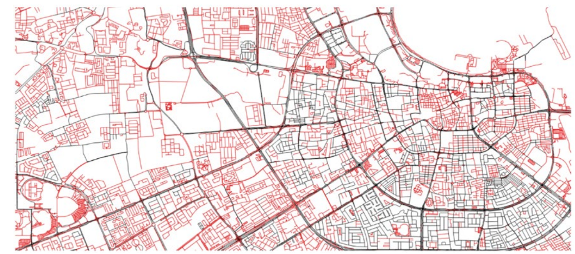


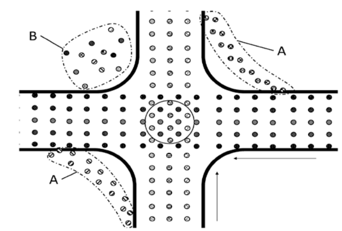
Edge Weight Inference in QARTA

- Divide all trips into time granularity (e.g., one hour)
- Solve the regression equation per time granularity



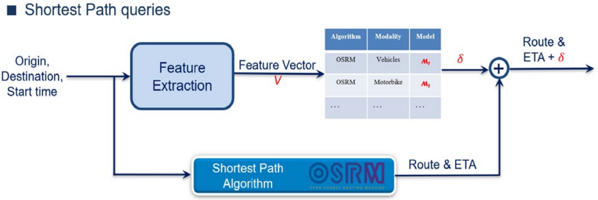
- More than 99% of the roads satisfy the physical constraint. For those that do not, we set edge weight to the minimum possible value (*1 / maxspeed*)



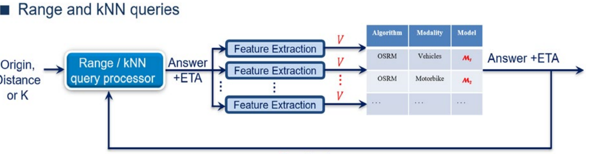


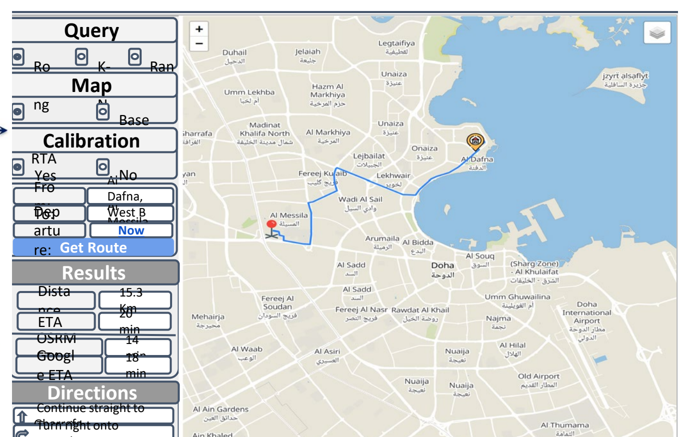
Query Calibration in QARTA

- Shortest Path queries



- Range and kNN queries



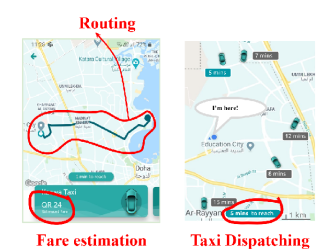


QARTA in Deployment

QARTA is deployed in *all* Taxis in Qatar $\sim 4K$ vehicles

A local food delivery company $\sim 3K$ motorbiks

- QARTA receives:
  - $\sim 235K$ daily API calls
  - ~ 1 *Million* daily GPS tracks
- APIs & Services:
  - In-traffic routes
  - Travel time estimation
  - Complex route planning
  - OD matrices
  - Search & addresses



Link: https://qarta.io


Thank you