#### **Projects Using METAL Highway Graphs with Traveler Data**

Jim Teresco

SIENAcollege

Computer Science Loudonville, New York

"Nifty Ideas" Presentation

35th CCSC Eastern Regional Conference Robert Morris University

October 26, 2019

Yet another Powerpoint-free presentation!

# Agenda

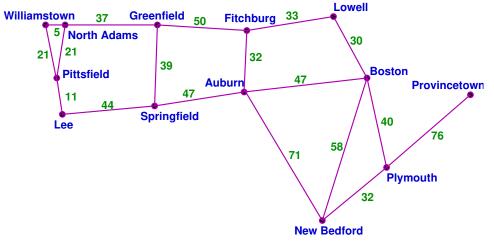
- A quick overview of METAL
  - Motivations
  - Travel Mapping
  - Bringing my "roadgeek" side to work: using Travel
  - Graph data derived from Travel Mapping project
  - Highway Data Examiner
  - METAL's interactive algorithm visualization
- Enhancement: Traveler Data
- Using METAL graphs in an algorithms final project
- Wrapup/acknowledgements

These slides are at http://courses.teresco.org/metal/talk.pdf and the big QR code above will take you there.



## **Early Motivation**

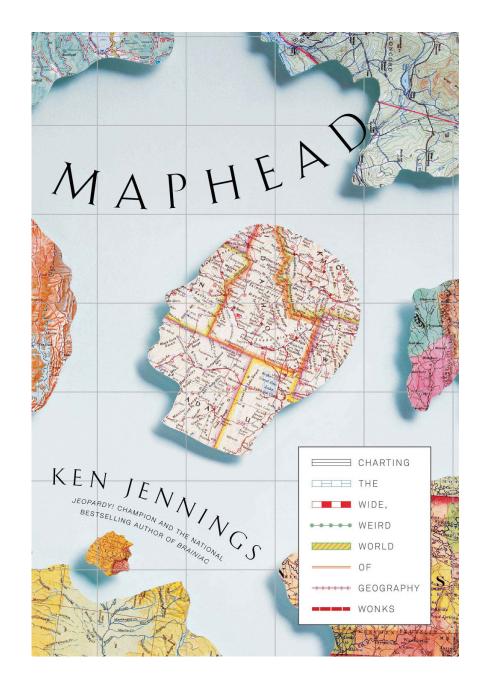
To find a good collection of data sets for teaching graph (and other) algorithms.



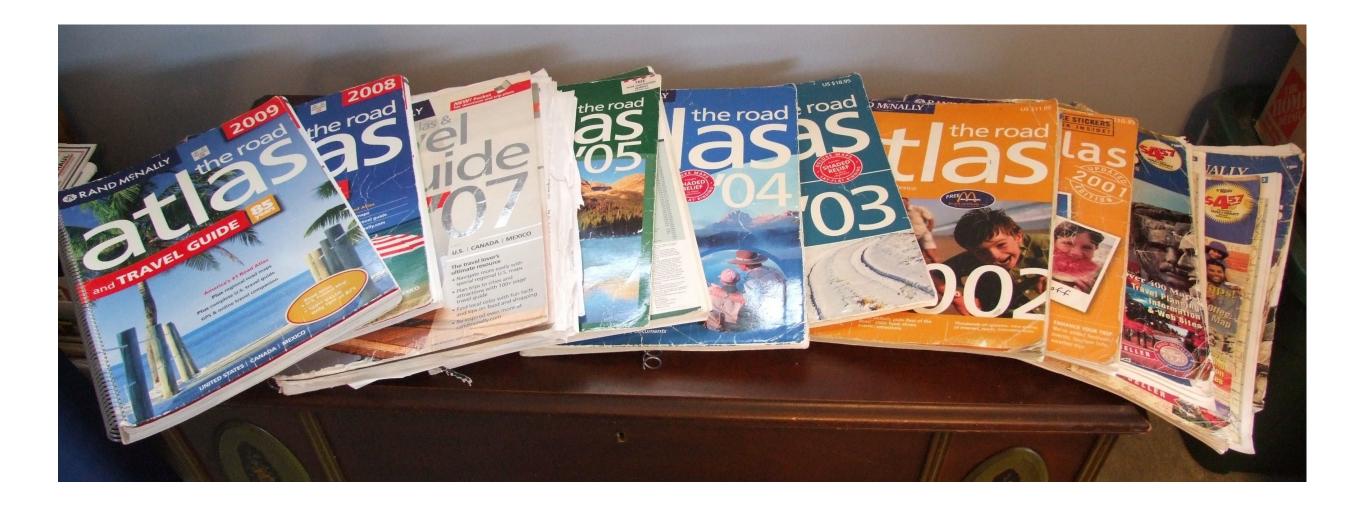
I want an example that's better than this!

- Variety of data sizes
  - small sets to be manageable for testing during development
  - larger sets to be interesting and for performance analysis
- Data source
  - freely and conveniently available data sets
  - a motivating real-world connection
- Meaningful visualizations of results

## Before I was a CS Geek - I was a Map Geek

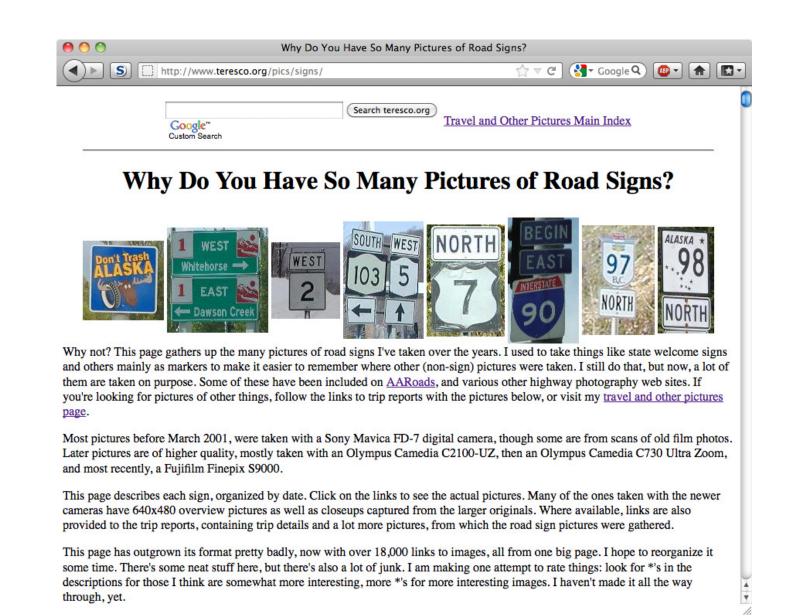


## The Well-Worn Rand McNally Road Atlas Collection



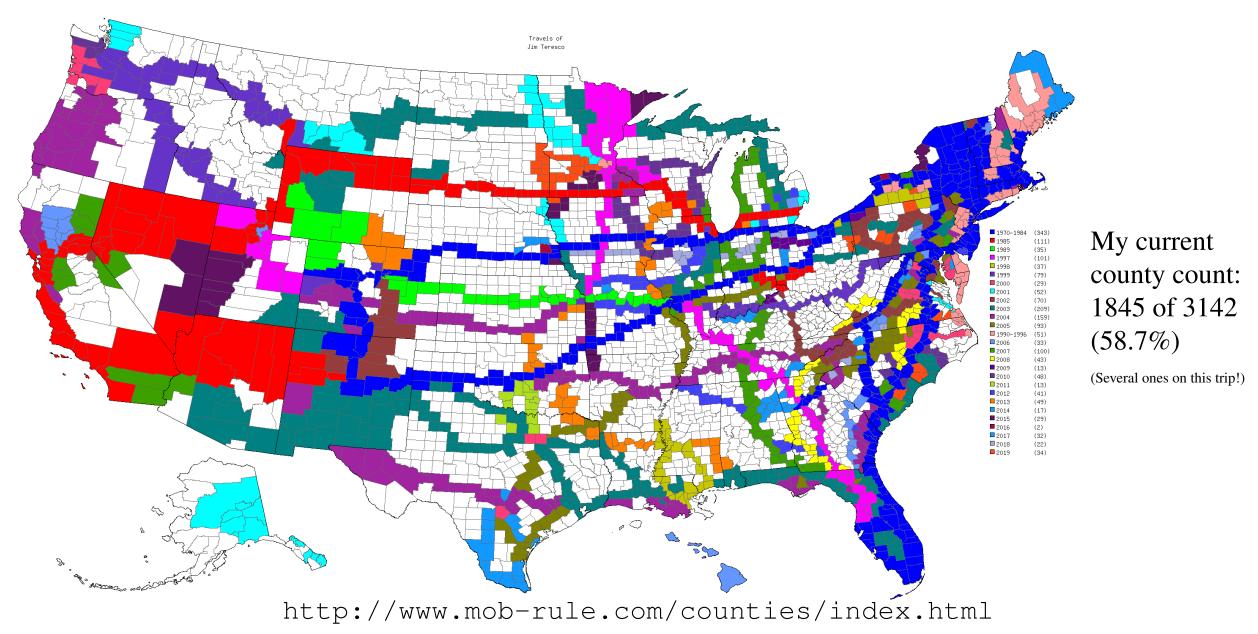
Each marked up with a highlighter to track travels.

#### **And I Used Highway Photography to Document my Travels**



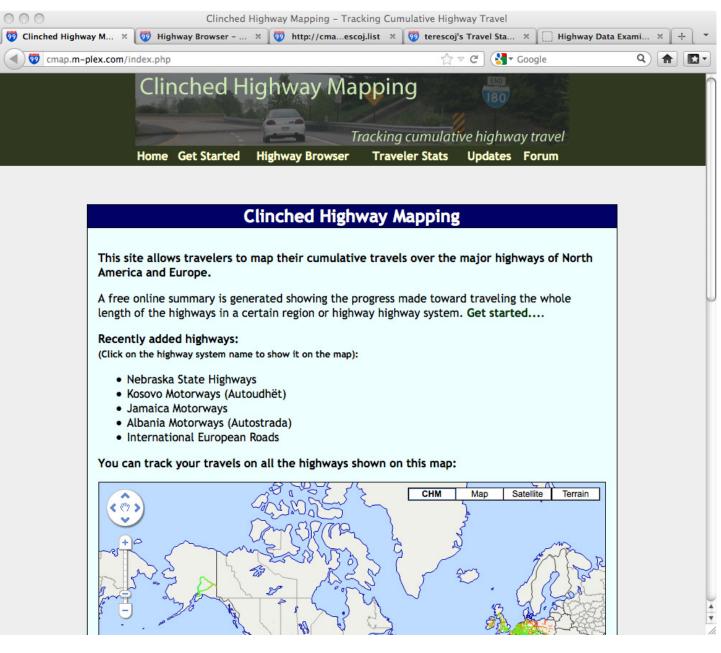
(Knuth has road sign pictures, so why not me?)

### And I Like to "Collect" Things

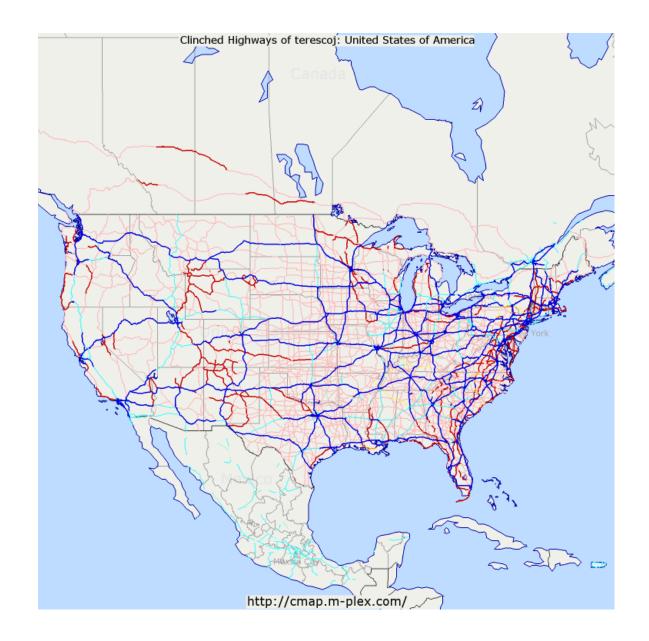


## As a "Roadgeek", I Could Not Resist

#### The Clinched Highway Mapping Project



## **To Track Highway Segments I've Traveled**



## $\mathbf{CHM} \longrightarrow \mathbf{Travel} \ \mathbf{Mapping}$

- Clinched Highway Mapping went dormant in 2014...
  - my role there was as a data contributor and user, but not administrator
  - the site exists, but is not actively maintained
  - CHM users and contributors decided to build a replacement
- Summer 2015: Birth of the Travel Mapping project
  - main site: http://travelmapping.net/
  - GitHub organization: https://github.com/TravelMapping
  - my role is as primary developer and project manager
  - highway data is constantly improving and expanding: now nearly 35,000 commits to the highway data repository by 20+ contributors
- Site users create and submit list of highway segments, e.g.,

https://github.com/TravelMapping/UserData/blob/master/list\_files/terescoj.list

• Enjoy the maps and statistics of your and others' travels!

http://travelmapping.net/user/mapview.php?u=terescoj&rg=PA

## Bringing my "Roadgeek" Side to Work

The observations, questions, and ideas that have led to the development of the tools that have evolved into today's

#### **Map-Based Educational Tools for Algorithm Learning (METAL)**

- Hey, some of my data structures examples and labs are pretty boring...
- Could I use CHM (now TM) data for class examples and assignments?
- If I could do that, could I plot it on maps like CHM does?
- Can I do all this before the Dijkstra's algorithm lab I am not really happy with that I need to assign next week?
- Hmmm, Dijkstra's algorithm. I think I need graph data, not a bunch of individual routes.
- Who needs sleep for the next week? I can do this! I'll create graphs and build a visualization tool too!
- All these years later, I'm still working on it!

### **METAL Graph Data**

METAL generates graphs from TM data for many countries, regions, highway systems, and local areas.

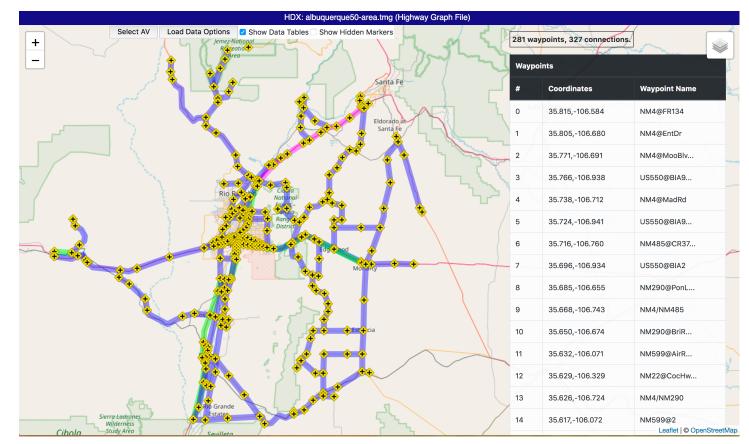
http://travelmapping.net/graphs/

Data set	V	E
ABW-region: All Aruba highways	28	37
DC-region: All District of Columbia highways	93	90
YT-region: All Yukon highways	225	221
albuquerque50-area: All plotted routes within 50 miles of Albuquerque	296	348
siena50-area: All plotted routes within 50 miles of Siena College	1198	1420
west-indies: All plotted routes in the West Indies	4649	5953
PA-region: All Pennsylvania highways	8575	10,053
usai-system: United States Interstate highway system	18,148	18,527
USA-national: All plotted NHS routes in the U.S.	71,392	76,535
USA-country: All plotted (state and national) highways in the U.S.	177,646	212,151
tm-master: All plotted highways Worldwide	421,966	499,353

A selection from among the 1200+ graphs provided by METAL:

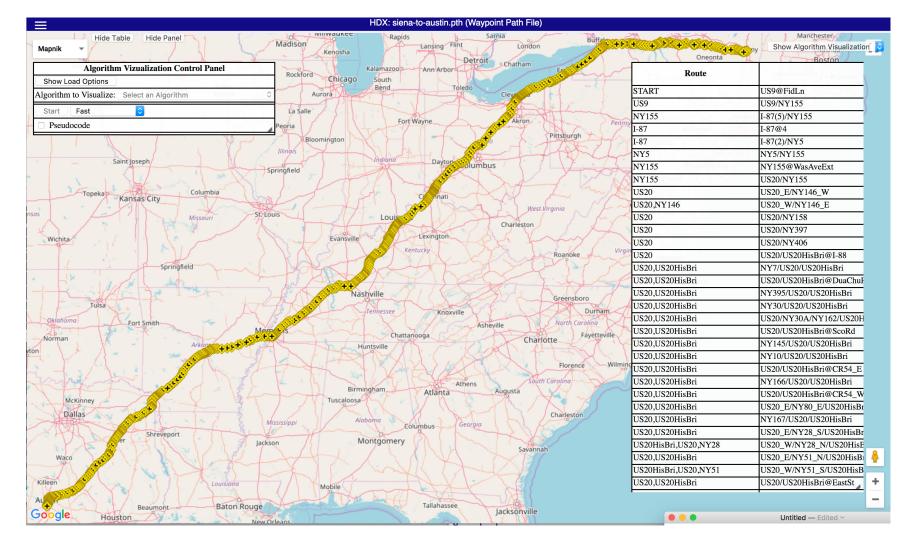
## **Visualizing the Results: Leaflet and OpenStreetMaps**

- Leaflet provides a free application programmer interface (API) to overlay custom data on OpenStreetMap (and other) map tiles
  - API's Javascript classes and methods are fairly easy to learn and use
  - Formerly used Google Maps, but free non-commercial use terminated
- METAL's Highway Data Examiner (HDX) overlays TM/METAL data
  - Graph of plotted highways within 50 miles of Albuquerque:



## **A Shortest Path Result**

HDX can also visualize the results of programs that use METAL data.



Shortest path from Siena College to Austin, Texas, generated using "Dijkstra's Road Trip" reference solution, displayed in HDX as a "pth" file.

## The Latest: METAL's Interactive Algorithm Visualization

- Interactive visualizations of algorithms rather than just the data and results
- The focus of a 2018 SIGCSE paper
- Implemented in HDX

		HDX: DE	region-collapsed.tmg (Highv	vay Graph File) Dijkstra's Alg	gorithm					
+ -	Resume Fast C Trace Pseudocode Reset AV Load Data Options C Show Data Tables Bomboy Hook National						663 waypoints, 814 connections.			
Algorithm Visual			S X 24			#	Coordinates	Waypoint Name		
	30 added to Priority Queue		HILLE		÷ •	#	Coordinates	waypoint Name		
	← new Priority Queue .add(start,null,0)			A State of the second			39.605,-75.745	US40/DE896		
whi	ile not tree.contains(end)					1	38.647,-75.297	DE30/DE24A		
i	if pq.isEmpty		VINY			<u> </u>	36.047,-75.297	DESU/DE24A		
	error: no path					2	39.559,-75.653	US13@BroSt		
	(to,via,d) ← pq.remove()			4			00.050 75.050	DEGLANZ		
i	if tree.contains(to) discard (to,via) // on removal				A. C.	3	38.656,-75.253	DE24AltZoa		
US 50 6	else	JE m			N.	4	39.619,-75.695	US40@PorRd		
US 301	tree.add(to,via,d)				N.					
	for each $e=(to,v) // neighbors$			1 the way	A.	5	38.639,-75.315	DE24AltZoa		
	if tree.contains(v)		at the		$\sim 1$	6	39.606,-75.785	US40@MD/DE		
	discard (v,e) // on discovery	J. C.		COFA	$\sim 10^{-1}$					
	else	Penton			Ă,	7	38.653,-75.368	US113/DE24		
	pq.add(v,e,d+len(e))				-1	8	38.525,-75.357	DE26_E/DE3		
	24_E/DE30_N found via DE30				$-1$ $\lambda$					
Indiscovered: 60			A LANCE	XIX:	$\sim$ D $\sim$	9	38.543,-75.258	DE26/US113		
	ze 18, max 18, avg 11.5) 🛛 limit display to 🧕 🔅				- Leves	10	38.503,-75.567	DE30@BiSta		
ntries		N nor		edden		10	30.000,-70.007			
325→484 7→320 DE24 US113			ANY AN	te Forest	Rehoboth	11	38.632,-75.296	DE30@MouJo		
2.944 11.653	11.526 8.865 8.777 8.709		Ge	ordetowe Gloware	Indian	12	38.502,-75.557	DE30/US13		
panning Tree: 38 iscarded on disc		5		Ainton	River Bay	13	38.664,-75.379	DE404TrkGe		
iscarded on rem		A 44					20.004 75 545			
7 Shortest Paths	Found So Far		A PART	and a second		14	38.604,-75.515	US9@KayeRd		
Pla	ace Distance Arrive From Via	have been	in the t			15	38.570,-75.563	US9/US13		
DE30	/US9 0.000 (START)	5.5		X	Bethan Beach					
DE30@	PreRd 1.266 DE30/US9 DE30					16	38.513,-75.557	US13@SaltB		
DE40	eo_E 1.539 DE30/US9 US9,DE404	US 50				17	38.503,-75.237	US113@DelA		
DE5//	US9 1.873 DE30/US9 US9,DE404	JCK 1			2					
5 mi	Blackwater					18	39.455,-75.675	DE1(136)/D Leaflet © OpenStre		

## **Key METAL Project Components**

We have seen METAL's main components.

1. Large set of graphs for a variety of highway systems generated from TM highway data

- real world connection and visualization on maps
- wide variety of sizes
- 2. HDX to visualize graphs
  - graph data as provided by METAL
  - also can visualize data in other formats generated by programs that use the graphs, such as shortest paths, sets of points, and lists of points
- 3. Interactive algorithm visualizations within HDX
  - Our experience: the *scalability* matters: small graphs to see the details and larger to see big picture behavior (think: BFS vs. DFS)

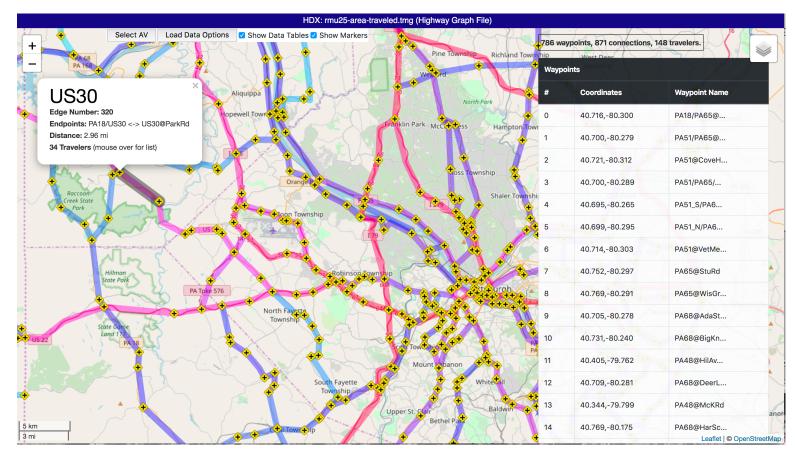
## Use It!

- METAL has been used recently in undergraduate Algorithms courses and Data Structures courses
- See http://courses.teresco.org/metal/using.html for examples including
  - an introductory assignment early in data structures
  - an introductory assignment early in algorithms
  - usage in a "Best Of" data structure
  - convex hull class example
  - assignment to analyze convex hulls
  - Dijkstra's Road Trip
- METAL data and HDX are publicly available (take a flyer)
- Instructor resources available on request
- Please contact me if you want to give it a try!

#### **New Feature: Graphs with Traveler Information**

Spring 2019: Implemented an idea that had been tossed around for a while.

- Idea: take advantage of the traveler data from the TM project
- TM users indicate which highway segments they have traveled
- Those segments correspond directly to the edges in METAL graphs



#### **New Feature: Graphs with Traveler Information**

- Augmented the collapsed format graphs to a new format called traveled
- Header includes the number of unique travelers across all graph edges

TMG 2.0 traveled 786 871 148

This is from rmu25-area-traveled.tmg, indicating 148 unique travelers.

• Each graph edge includes a hexadecimal string with enough bits to represent each traveler

We can see not too many '1' bits on the segment of Route 978, but many more on the segment that carries I-376, US 22, and US 30 through the Pittsburgh area.

• Traveler usernames are included in the last line to allow mapping of bits to users, if desired

# **Spring 2019 Project**

• Part of a final project for *Analysis of Algorithms* 

#### http://courses.teresco.org/cs385\_s19/probsets/ac/

• Each individual or group was required to use METAL data in some form for at least one algorithm studied

- Projects included
  - simple variations on searches: most visited/traveled, least visited/traveled
  - variations on Dijkstra's/Prim's/graph traversals that gave weight to heavily- or lightly-traveled segments
  - convex hull of the places visited by a particular traveler
- Presented as a session in Siena's year-end "Academic Showcase" and was heavily attended!

### **Current Work, Plans, and Ideas for the Future**

- Continue to add new data sets and improve the quality of existing sets
- More examples/assignments/labs/teaching modules built around METAL's data and visualizations
- Expanded and improved interactive algorithm visualization
  - including traveler information
  - recursion
  - parallel algorithms
  - more complex algorithms
  - student algorithm "implementations"
- More studies measuring the effectiveness of METAL

#### **Interested?**

#### http://courses.teresco.org/metal/



"An algorithm must be seen to be believed, and the best way to learn what an algorithm is all about is to try it." - Donald E. Knuth, *The Art of Computer Programming, Volume 1* (1997), p. 4.

Nearly any topic in a course working with data will be more interesting for students if the data has connections to the real world and if they can visualize the data and results in a meaningful way. When using graph data, such as in a data structures or algorithms class, it needs to be small enough to be manageable, but large enough to be interesting. This might consist of a small road system, airline flight schedules, or even the layout of a campus or building. The *Map-based Educational Tools for Algorithm Learning (METAL)* project provides data and visualization capabilities for this purpose. METAL's graph data sets range in size from a few vertices and edges to several hundred thousand. Its data is derived from the <u>Travel Mapping (TM) Project</u>, and represents highway systems from around the world. Its map-based visualization capabilities are built with <u>Leaflet Maps</u>. METAL's interactive code-level algorithm visualization capabilities with debugger-like controls are intended to aid student understanding of graph and other algorithms. Students can implement graph algorithms themselves and display, in map form, the results of computations using those algorithms when applied to METAL data.

Wouldn't it be more fun and interesting to work with this graph



### Acknowledgements

- TM credits at http://travelmapping.net/credits.php
- Thank you to the Saint Rose students who have contributed to the project, especially Razie Fathi (now at RIT and still a project contributor), Dan Priddle, Kevin Bayly, Shipra Goel, Rashmi Reddy Podduturi, and Paul Amodeo.
- Thank you to my Summer 2017, Summer 2018, and Summer 2019 students, MariaRose Bamundo, Arjol Pengu, Clarice Tarbay, Michael Dagostino, Abdul Samad, Eric Sauer, Zach Goodsell, Tyler Gorman, and Alissa Ronca. Their work was made possible through the support of the Summer Scholars Program, Center for Undergraduate Research and Creative Activity (CURCA), at Siena College.



Computer Science



# Acknowledgements

