Technical Report 129

Project Title:

Imputing Missing Data via Sparse Reconstruction Techniques

Constantine Caramanis
WNCG

June 2017

Data-Supported Transportation Operations & Planning Center (D-STOP)

A Tier 1 USDOT University Transportation Center at The University of Texas at Austin

CENTER FOR TRANSPORTATION RESEARCH

Wireless Networking & Communications Group

D-STOP is a collaborative initiative by researchers at the Center for Transportation Research and the Wireless Networking and Communications Group at The University of Texas at Austin.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
D-STOP/2017/129		

4. Title and Subtitle
Executive Summary: Imputing Missing Data via Sparse Reconstruction Techniques

5. Report Date
June 2017

6. Performing Organization Code

7. Author(s)
Constantine Caramanis

8. Performing Organization Report No.
Report 129

9. Performing Organization Name and Address
Data-Supported Transportation Operations & Planning Center (D-STOP)
The University of Texas at Austin
1616 Guadalupe Street, Suite 4.202
Austin, Texas 78701

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
DTRT13-G-UTC58

12. Sponsoring Agency Name and Address
Data-Supported Transportation Operations & Planning Center (D-STOP)
The University of Texas at Austin
1616 Guadalupe Street, Suite 4.202
Austin, Texas 78701

13. Type of Report and Period Covered

14. Sponsoring Agency Code

15. Supplementary Notes
Supported by a grant from the U.S. Department of Transportation, University Transportation Centers Program.

16. Abstract
The State of Texas does not currently have an automated approach for estimating volumes for links without counts. This research project proposes the development of an automated system to efficiently estimate the traffic volumes on uncounted links, in the event of rare disturbances of the typical traffic flow. The idea we plan to leverage is that the road network provides a mixing effect, whereby localized disturbances (accidents, flooding, road damage, etc.) have an impact whose effect can be measured at many places across the city. This forms the important analog to the well-known uncertainty principle, whereby a signal cannot be sparse in both the time and frequency domains—a result that is critically utilized in the signal reconstruction algorithms for fMRI.

17. Key Words	18. Distribution Statement
	No restrictions. This document is available to the public through NTIS (http://www.ntis.gov): National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161

19. Security Classif.(of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified		

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Acknowledgements

The authors recognize that support for this research was provided by a grant from the U.S. Department of Transportation, University Transportation Centers.

Executive Summary: Imputing Missing Data Via Sparse Reconstruction Techniques.

In this line of work, we consider problems with potentially large amounts of missing data. In planning problems, a central reason for missing data comes from the simple fact that sensors measuring, for example, capacity reduction along certain routes, are limited and may not be deployed in all areas of interest. Examples of capacity reduction we care about include minor flooding or other adverse weather-induced conditions, such as ice or snow. Our point of departure in this line of work is the hypothesis that interconnectivity of the transportation system can allow us to use point measurements of other quantities more readily measured, to infer local measurements where there are no sensors. Our work has only validated this conceptual framework on synthetic data, and we hope to validate extensively on real data in a second phase of the project. Thus far, we have developed highly scalable algorithms that succeed in imputing missing data from measurements that satisfy a particular *incoherence* property. The inspiration to the connection to transportation is that the interconnection of the transportation system (a capacity reduction in one location has impacts that can be measured in other locations) is a physical analog to the incoherence property that fMRI measurements exhibit. These have been used in sparse signal reconstruction in completely different application domains.

The validation to real data is an important goal, but beyond the scope of this first phase of the project. We expect that we will use our initial results as a proof of feasibility of the concept, in order to pursue appropriate scale funding for further exploration of this direction.

Note: A full report is forthcoming.