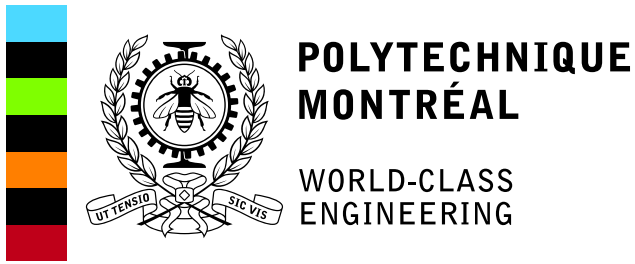
Open Source Tools for Trajectory Data Analysis

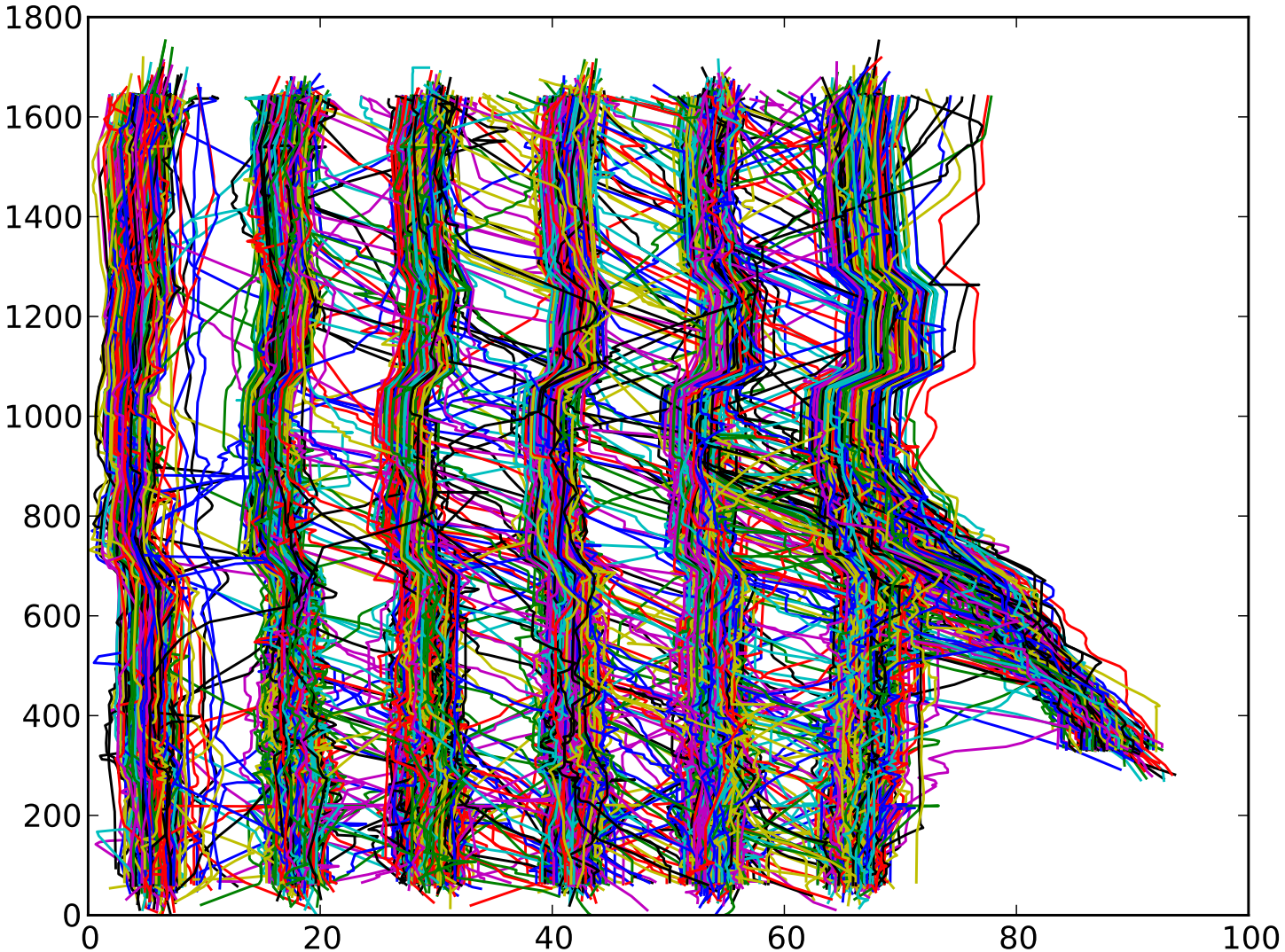
Workshop on Big Data Informatics
Transportation Research Board 92th Annual Meeting

Nicolas Saunier
nicolas.saunier@polymtl.ca



January 13th 2013

Transportation Data



Trajectory Data

$$(t_i, x(t_i), y(t_i))$$

Issues

- Different sampling rates/speeds

Issues

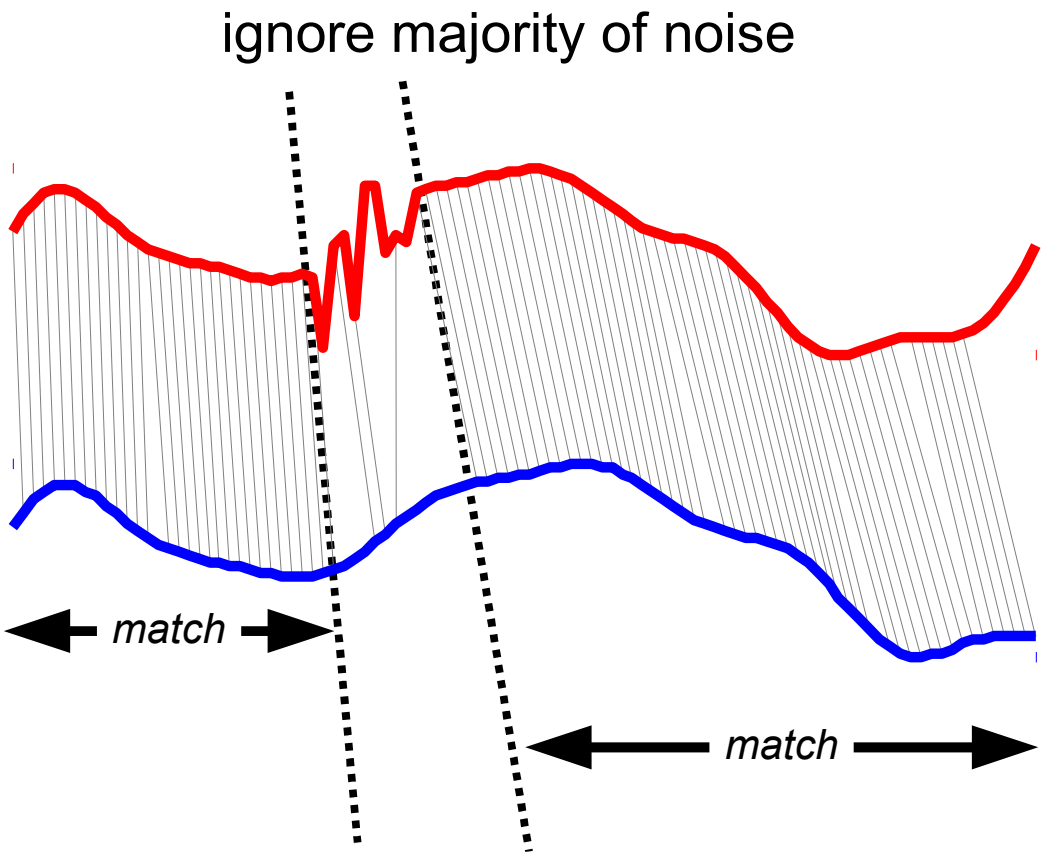
- Different sampling rates/speeds
- Outliers and uncertainty

Issues

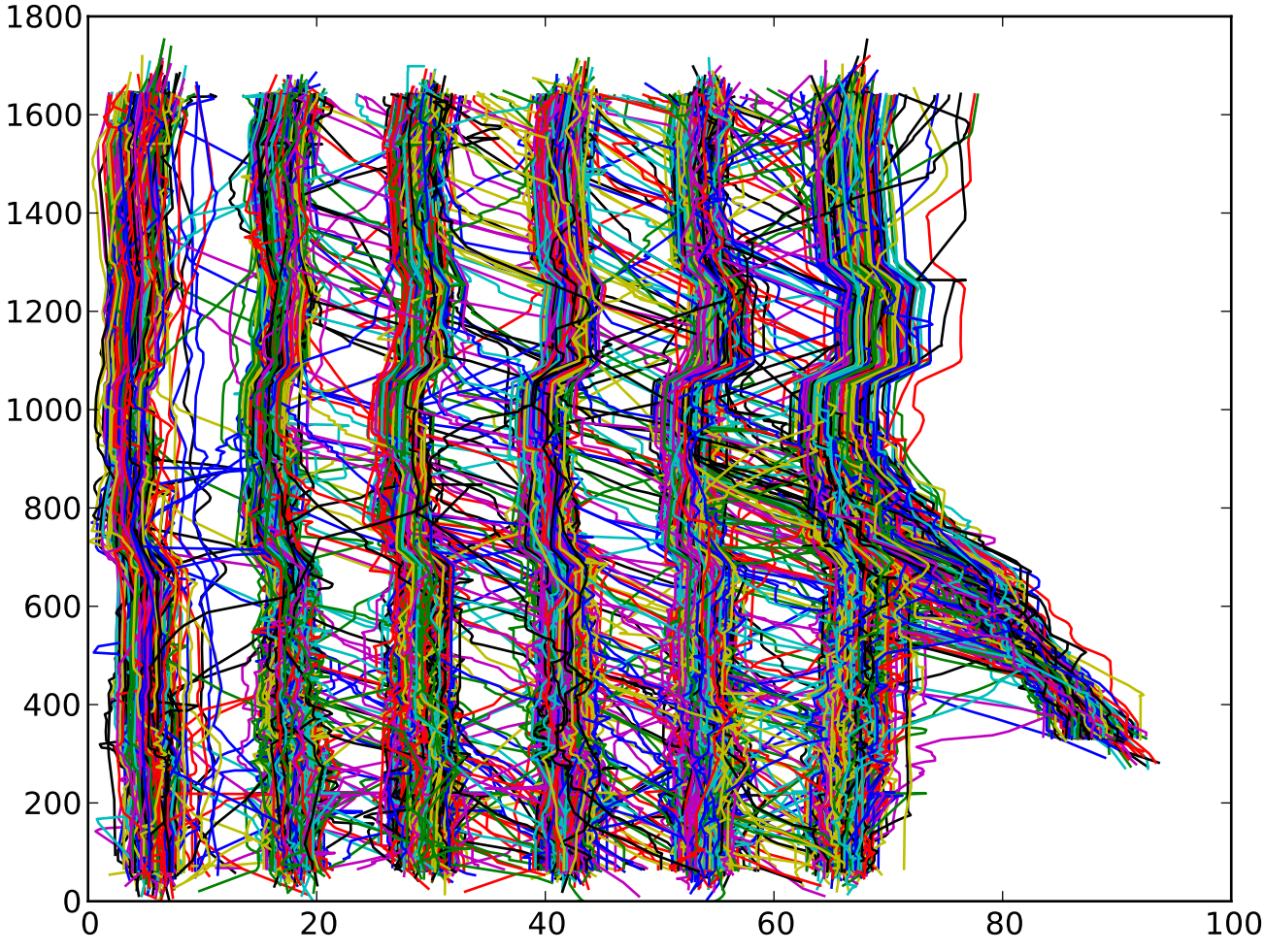
- Different sampling rates/speeds
- Outliers and uncertainty
- Different lengths: trajectories cannot be processed in fixed-size tables (e.g. spreadsheets), re-sampling loses information, actual positions

Solutions

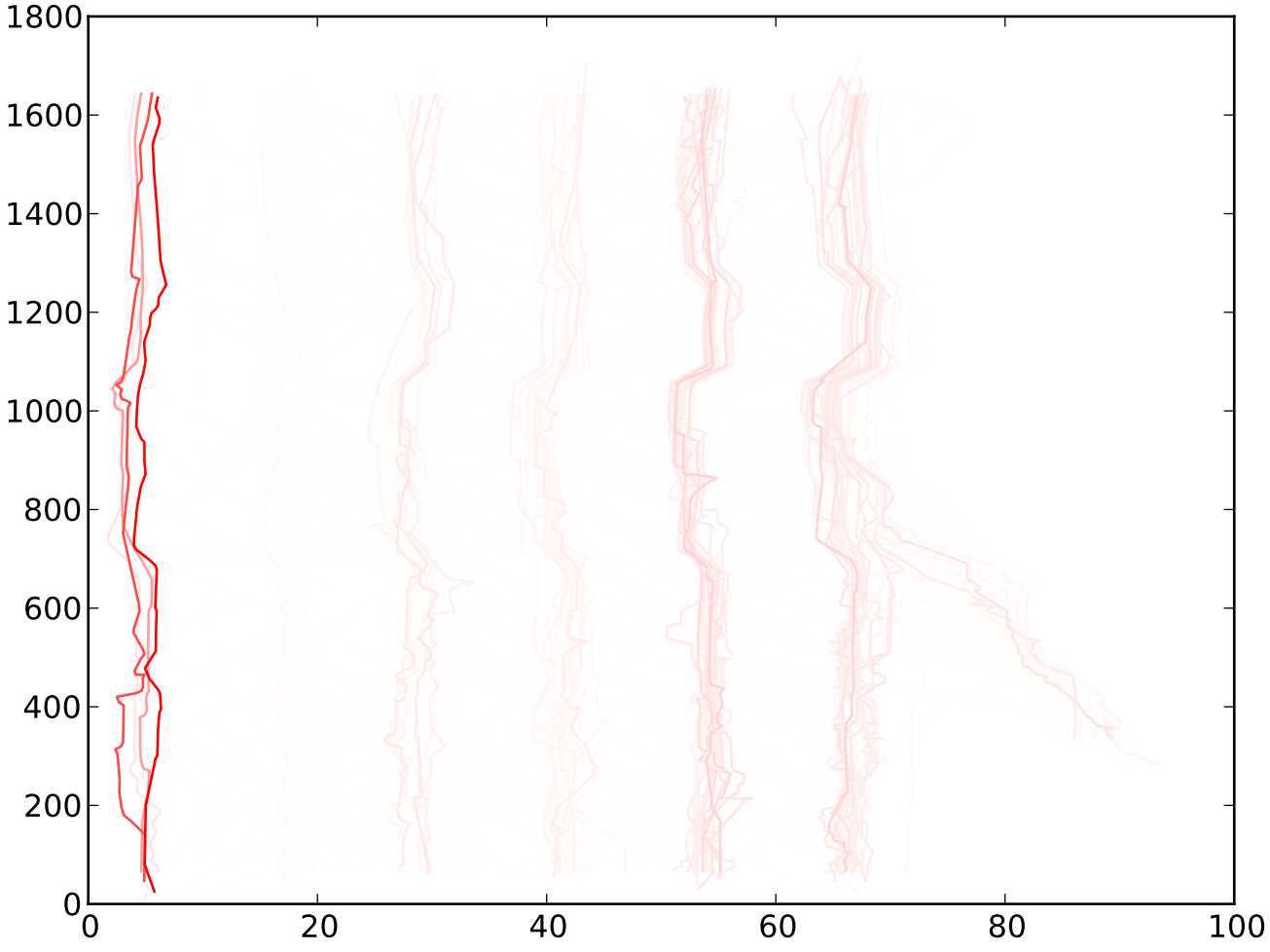
Use adapted distances/similarities such as the longest common subsequence similarity (LCSS)



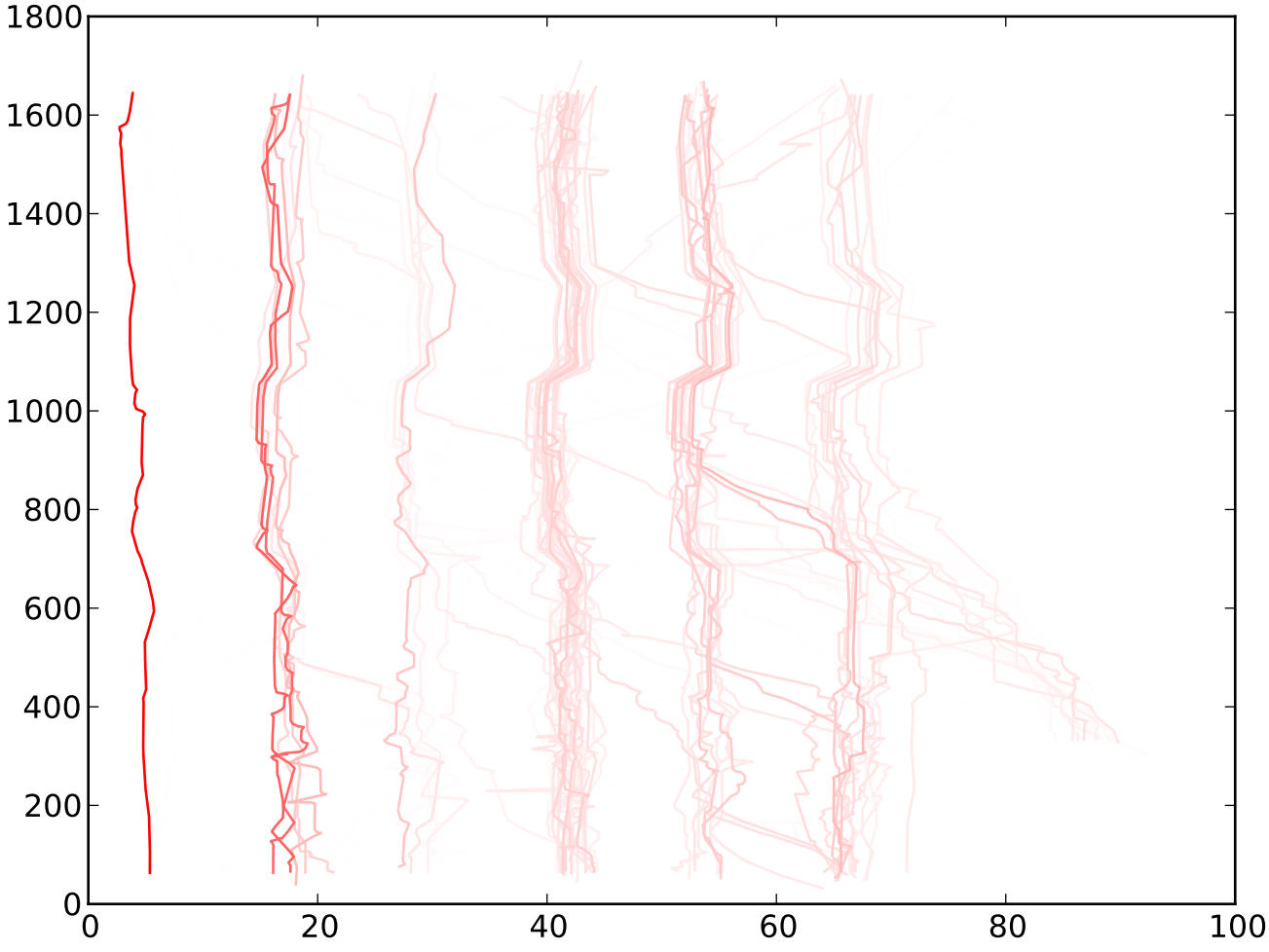
Clustering Examples: NGSIM Dataset (2052)



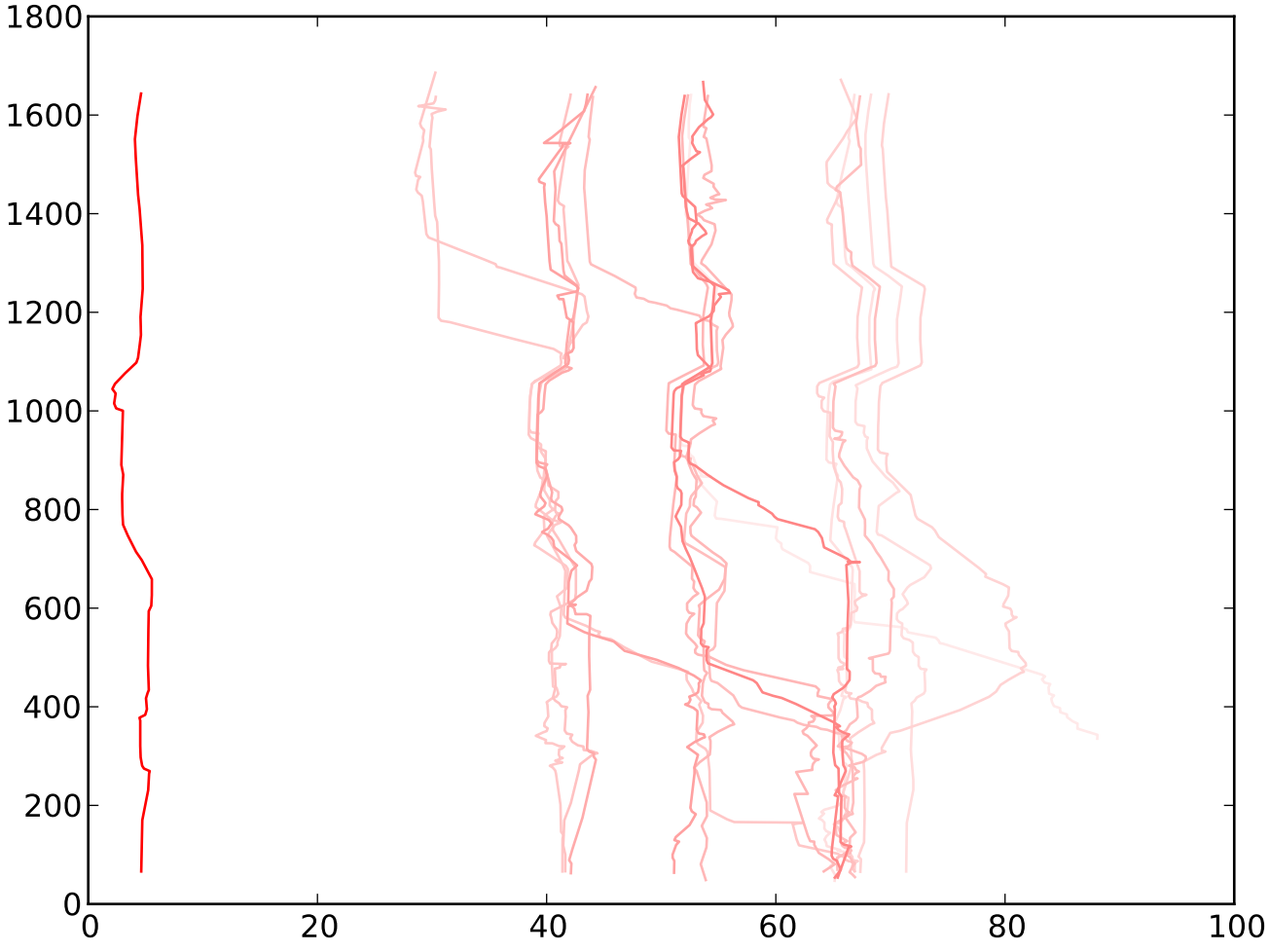
Clustering Examples: NGSIM Dataset (333)



Clustering Examples: NGSIM Dataset (96)

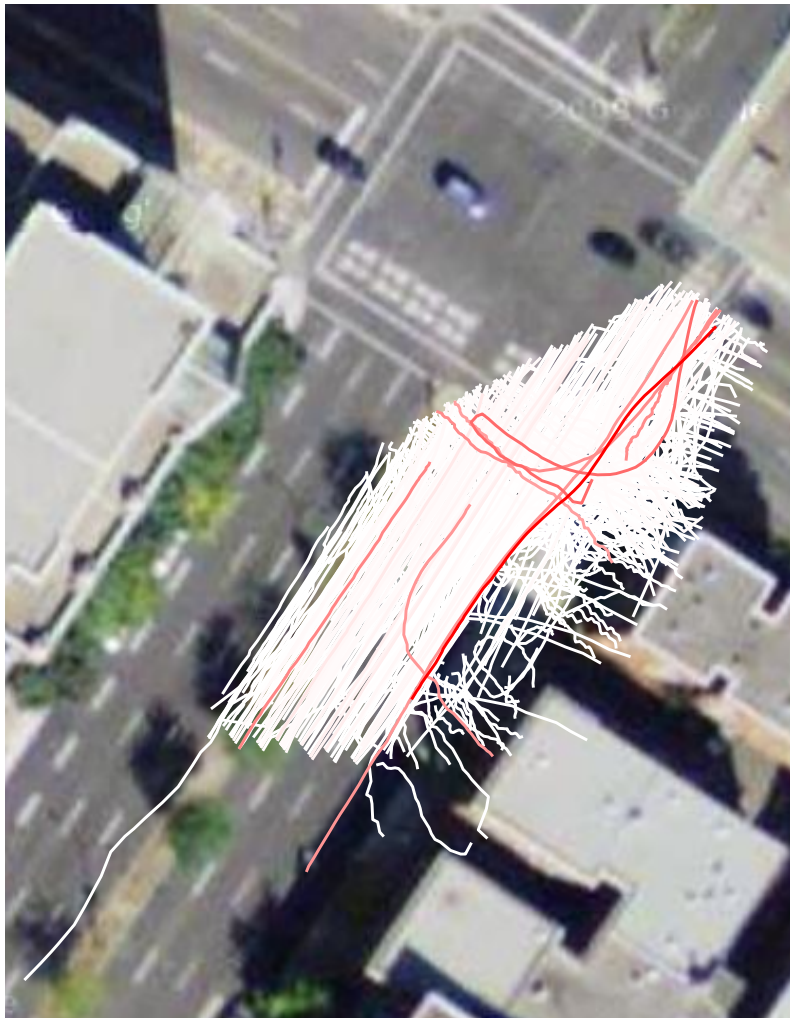


Clustering Examples: NGSIM Dataset (19)

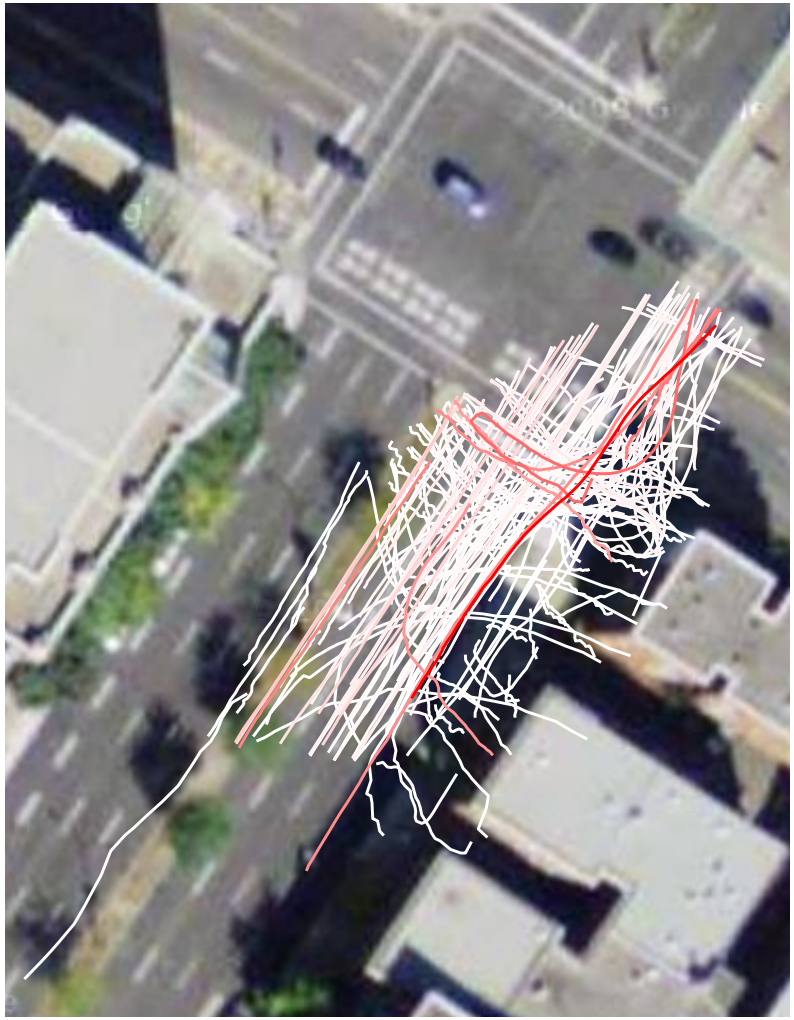


Clustering Examples: Montréal Intersection (6777)

Clustering Examples: Montréal Intersection (587)



Clustering Examples: Montréal Intersection (168)



Clustering Examples: Montréal Intersection (9)

Ongoing Open Source Software Development

- Trajectory management and analysis library
  `https://bitbucket.org/trajectories/trajectorymanagementandanalysis`
- Video-based road user tracking tool
  `https://bitbucket.org/Nicolas/trafficintelligence`

Under BSD/MIT License

Conclusion

- Trajectory data is everywhere and we need the right tools to process it

Conclusion

- Trajectory data is everywhere and we need the right tools to process it
- Open Source Software development in progress at Polytechnique Montréal

Conclusion

- Trajectory data is everywhere and we need the right tools to process it
- Open Source Software development in progress at Polytechnique Montréal
    - opportunities for collaboration

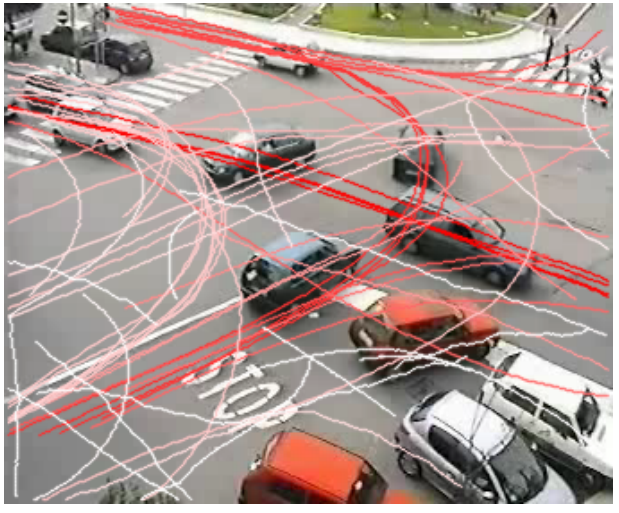
- Acknowledgments: Piotr Bilinski (Université Sophia-Antipolis), François Bélisle (Polytechnique)
- Funding: Google Summer of Code 2010, NSERC

- Acknowledgments: Piotr Bilinski (Université Sophia-Antipolis), François Bélisle (Polytechnique)
- Funding: Google Summer of Code 2010, NSERC

Questions?

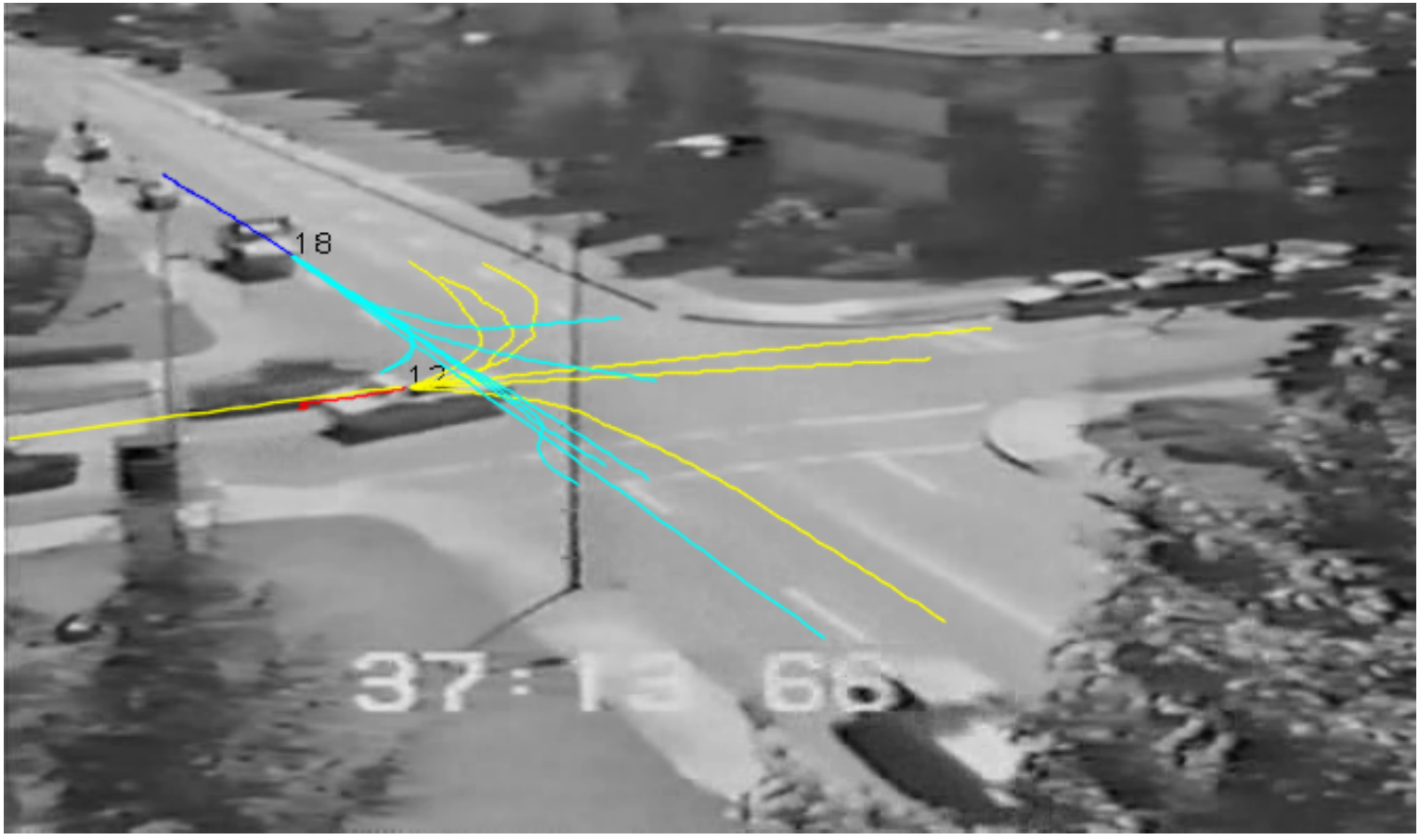
`http://nicolas.saunier.confins.net`

Motion Pattern Learning



Traffic Conflict Dataset, Vancouver	Reggio Calabria, Italy
58 prototype trajectories (2941 trajectories)	58 prototype trajectories (138009 trajectoires)

Application to Road Safety Diagnosis



Conflict data (Vancouver)

Open Source Software (OSS)

- OSS defining characteristics (Open Souce Initiative)
  - Free redistribution
  - Source code
  - Derived work
- OSS is everywhere and you are using it daily
  - Google, Linux web servers, Android, Facebook...
- OSS often generates strong reactions: this is not about giving away software for free, or being anti-profit, but about a superior software engineering method
  - for example, The Apache foundation is suported by Microsoft, Facebook, Yahoo!, Google, IBM, HP, AMD, etc.

Benefits of Open Source Software

1. Reproducibility of scientific results and fair comparison of algorithms
2. Uncovering problems
3. Building on existing resources (rather than re-implementing them)
4. Guaranteed access to software and tools
5. Combination of advances
6. Faster adoption of methods in different disciplines and in industry
7. Collaborative emergence of standards

Benefits of Open Source Software

- OSS should be an obvious choice for academia (being publicly funded) and considered by industry
- Buyers should be very careful about standards and continued access to technology, and open source is an important part of the solution
- There are successful mixed business models with open source core libraries and paid graphical interfaces, technical support, consulting services, etc.