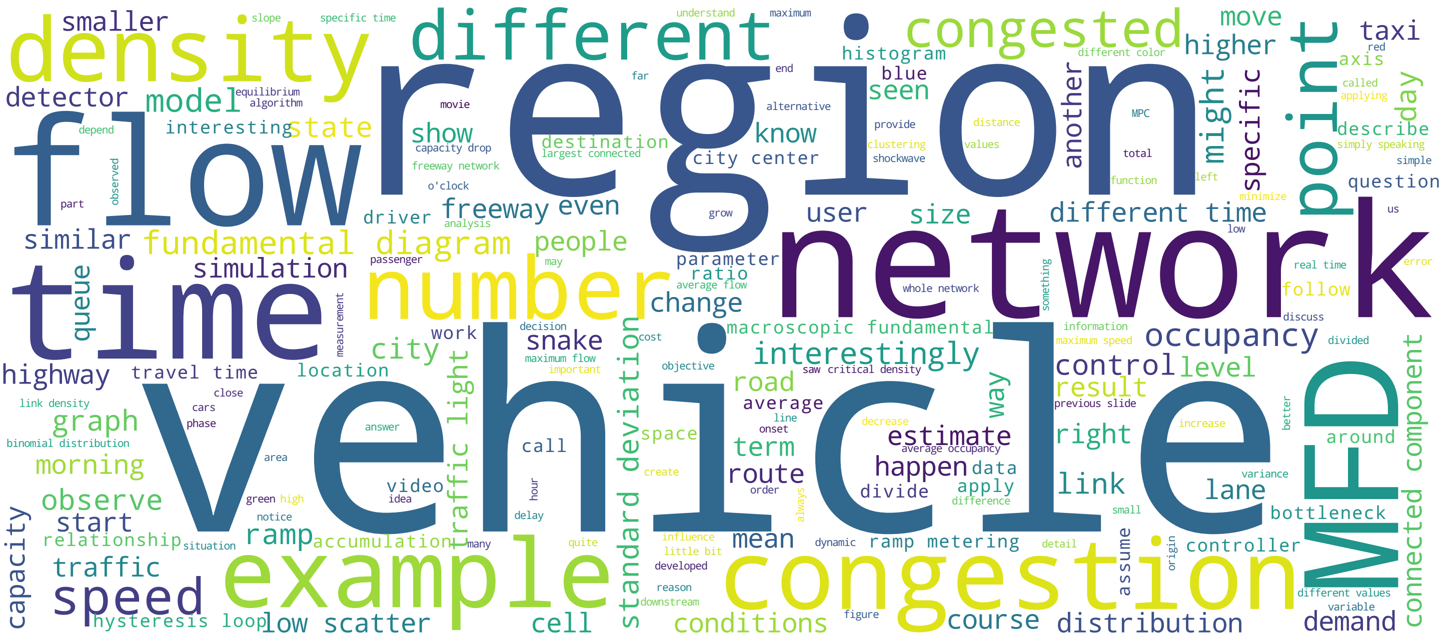
Macroscopic Fundamental Diagram:
Existence, Physical Properties and Dynamic Modeling

Intro to traffic flow modeling and ITS

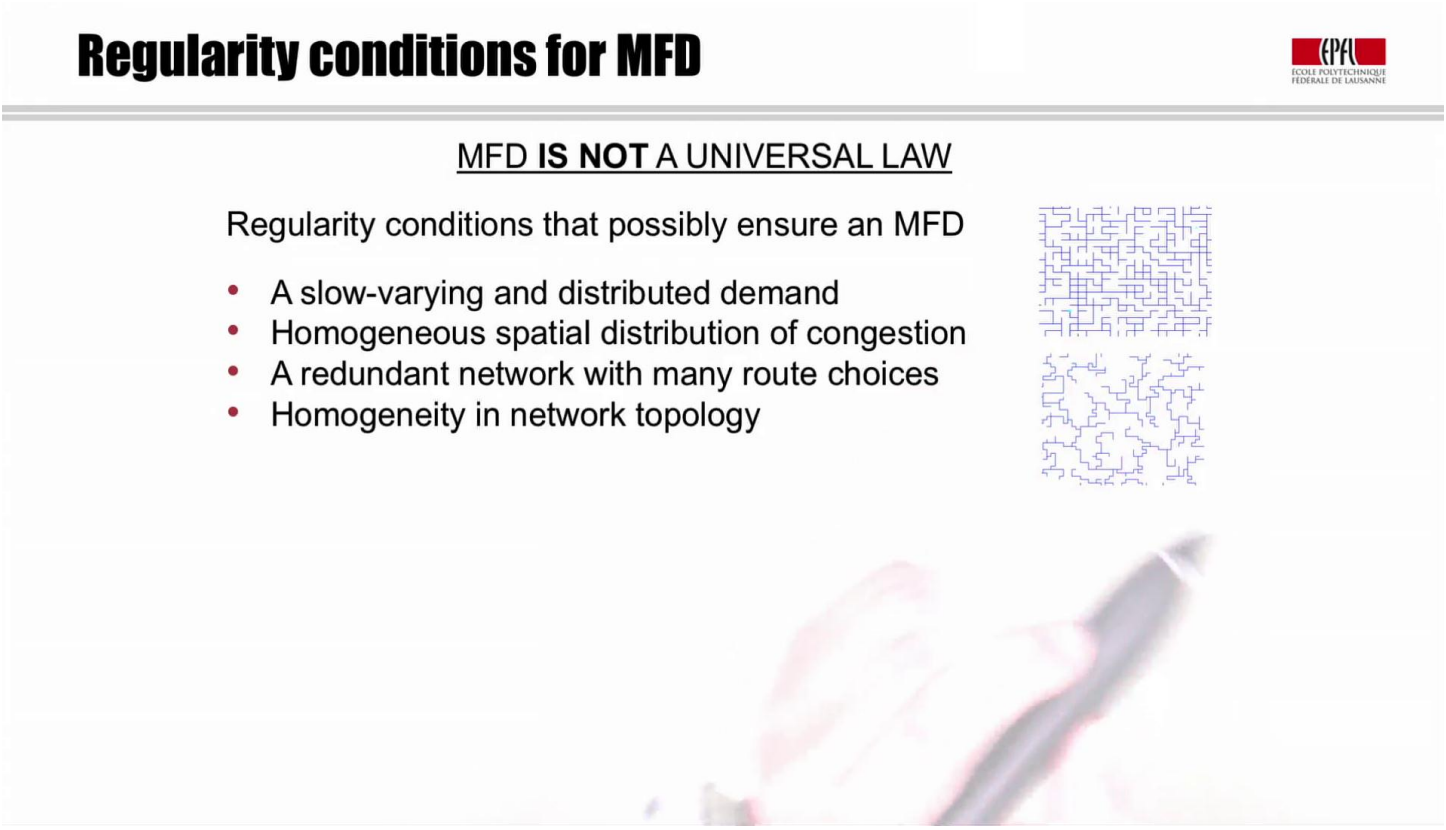
Prof. Nikolas Geroliminis



Search MOOC

Video

Regularity conditions for MFD



Welcome. So far we have seen that in some specific cases, a macroscopic fundamental diagram between Network flow and Network accumulation or density exists and it has low scatter, but this is not necessarily universal law. So during this video, we will investigate under what properties and what conditions an MFD with low scatter exists and we will go deeper in the analysis of this interesting tool. An MFD is not a universal law. It doesn't mean that anytime we aggregate at the network level flow and density, we will get an MFD with low scatter. There are some regularity conditions that possibly ensure a low scatter macroscopic fundamental diagram. So first of all, we need to have a slow wiring in distributed demand. So if demand changes very rapidly, this might influence and create significant scatter in the MFD. Like for example, when a city is evacuated. The second important regularity condition that we will see in details is, we want the spatial distribution of congestion to have a good degree of homogeneity. Also, in terms of the network topology, we would like to have a redundant network with many route sources. This means that when you go from point A to point B.

Notes

Summary

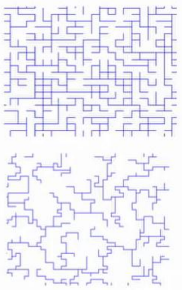
0m 05s

Regularity conditions for MFD

MFD **IS NOT** A UNIVERSAL LAW

Regularity conditions that possibly ensure an MFD

- A slow-varying and distributed demand
- Homogeneous spatial distribution of congestion
- A redundant network with many route choices
- Homogeneity in network topology



- An MFD with low scatter
  - locally heterogeneous but macroscopically regular networks (e.g. cities with multiple modes)
- An MFD with high scatter
  - Networks with uneven and inconsistent distribution of congestion (e.g. freeways)

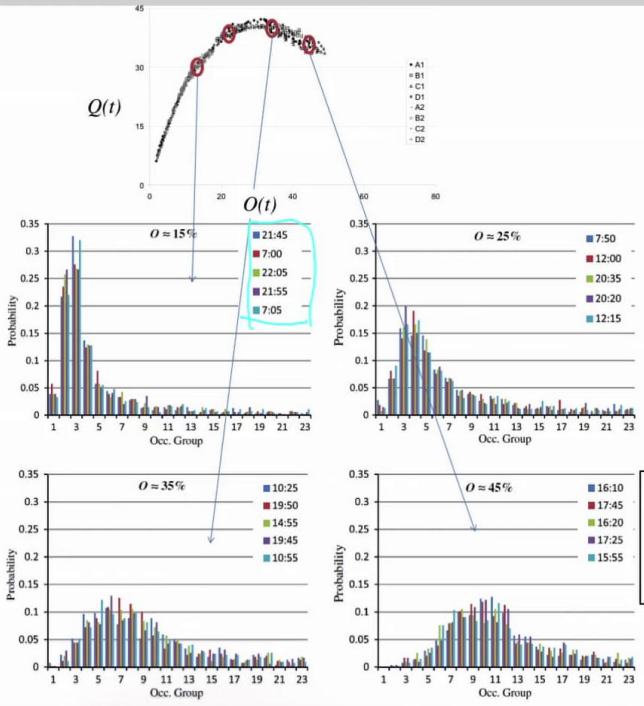
You might have plenty of alternatives in this network. While if you have a non redundant network, there are only very few alternatives to go from one point to another. And this actually influences the spatial distribution of congestion. Because if you want to go from A to B and there is a bottleneck in between, then people can update their routes in real-time and these will have as a result that more roads get congested. And also the last regularity condition is we want to have some level of homogeneity in the network topology. Which means that for example, we don't want to have in the same aggregation arterial roads with freeway roads. Summarizing, an MFD with low scatter can possibly exist when we have local heterogeneity, but microscopically regular networks, as for example when we have city with multiple modes of transfer. But an MFD with high scatter might happen one you have networks that have an even and inconsistent distribution of congestion, like is the example of freeways as we will see later.

Notes

Summary

1m 36s

Properties of well-defined MFDs



$d_r(t)$: pdf of individual detectors' density in region r
$Q(t)$ and $O(t)$: Space-Mean network flow and occupancy

$$\{Q(t_1) = Q(t_2) \quad \text{and} \quad O(t_1) = O(t_2)\}$$
$$\iff d_r(t_1) \sim d_r(t_2).$$

Variance much higher than binomial's

WHY?

Correlation of link density (propagation)

Geroliminis and Sun (2011) – Tr. Res. Part B

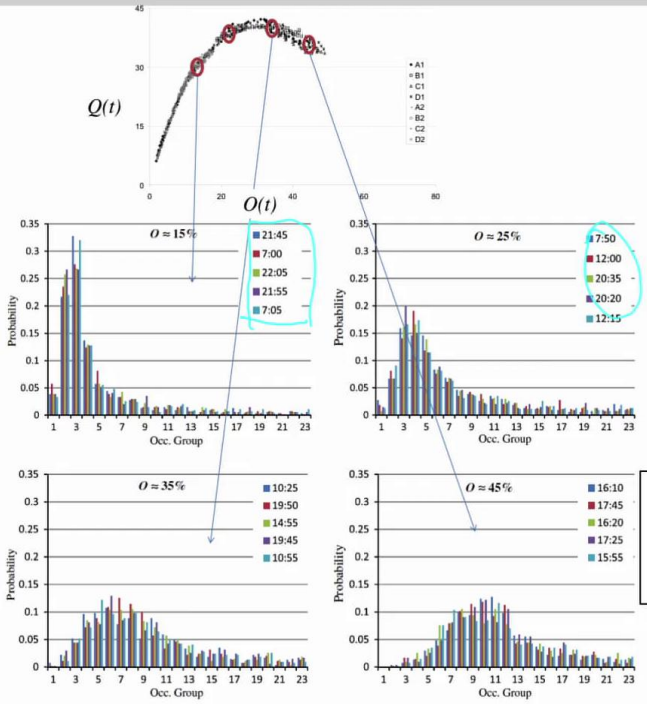
So let's try to understand properties of well defined MFDs. And the first important one is the spatial distribution of congestion. So we showed before in this graph, this is the MFD from Yokohama where each of these points represent the space mean of occupancy and flow from 500 detectors for a 5 minute period. So what we can do, we can collect all the snapshot of the network that have a similar level of occupancy. So over here, we have 5 different times where the average occupancy of the network is around 15%. And we can do that also for different values of occupancy. So 15% is quite and congested. So we do it also for 25% where the network is a little bit before maximum flow, a little bit after maximum flow, so 35% and when the network is really congested then we have an occupancy of 45%. So what we do, we take all these 500 detectors and for each of these different times that are represented in this graph with a color, we plot a histogram of the occupancy. So the occupancy varies between 0 and 100. So what we have, we divide this in 23 occupancy groups. So the first group represents very low occupancy, the last group represents very, very high ultimately, close to 100 and we see what fraction of detectors are in this ultra pants group.

Notes

Summary

2m 57s

Properties of well-defined MFDs



$d_r(t)$: pdf of individual detectors' density in region r
$Q(t)$ and $O(t)$: Space-Mean network flow and occupancy

$$\{Q(t_1) = Q(t_2) \quad \text{and} \quad O(t_1) = O(t_2)\}$$
$$\iff d_r(t_1) \sim d_r(t_2).$$

Variance much higher than binomial's

WHY?

Correlation of link density (propagation)

Geroliminis and Sun (2011) – Tr. Res. Part B

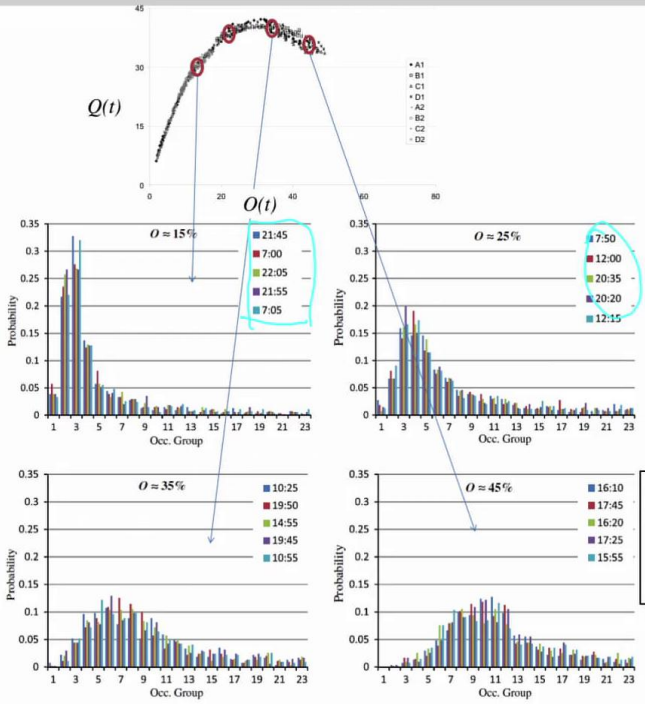
So very interestingly, even if these are very different times of the day. So you see morning, but also evening and also for different days. What we see, we see that the histogram of the distribution of occupancy of individual detectors is very similar. And if we repeat the same analysis for the different values of occupancy, we see that the trend remains the same. So clearly, when average occupancy increases, what we see, we see that there is larger probability that some detectors are more congested. And how to observe this? We observe it by showing that these histograms are moving from the left to the right. But at the same time in all situations the different times of the days for the same level of occupancy, we have the same histogram. So if you'd like to quantify things mathematically what we could say is that if during two different times t1 and t2 networks have the same point in the fundamental diagram in terms of macroscopic fundamental diagram,so the occupancy and the network flow are the same. What we observe here is that the distribution of congestion is very, very similar during these times. So another point I would like to make is that, the variance of this distribution is much higher than the binomial distribution.

Notes

Summary

4m 48s

Properties of well-defined MFDs



$d_r(t)$: pdf of individual detectors' density in region r
$Q(t)$ and $O(t)$: Space-Mean network flow and occupancy

$$\{Q(t_1) = Q(t_2) \quad \text{and} \quad O(t_1) = O(t_2)\}$$
$$\iff d_r(t_1) \sim d_r(t_2).$$

Variance much higher than binomial's

WHY?

Correlation of link density (propagation)

Geroliminis and Sun (2011) – Tr. Res. Part B

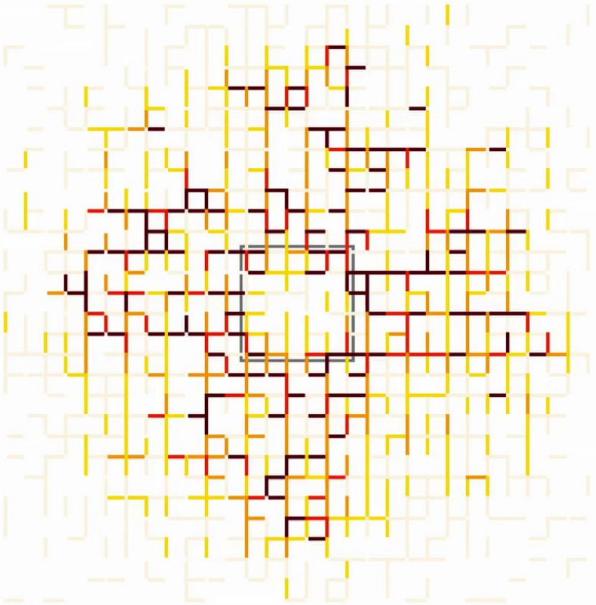
And what is the binomial distribution? Binomial distribution is for example when vehicles are randomly distributed in the network. So if for any small part of the network, there is the same probability that it has a vehicle or not, then we can easily show that the distribution of the number of vehicles the network follows a binomial distribution. It actually follows a hyper geometric distribution, but because you have a large number of points, it tends to be a binomial. So what we observe, we observe that there is a much higher variance than the binomial distribution and the reason for this is because you have strong correlations of link density. So this means that if a road is congested, there is a high probability that the road next to it will also be congested. And actually this happens in many physical systems that have special correlations and what we observe is that this system have much higher variance.

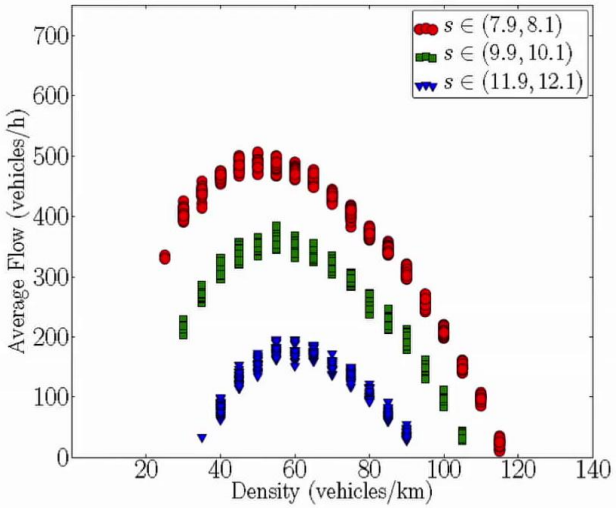
Notes

Summary

6m 35s

Spatial heterogeneity and MFD shape





Mazloumian, Geroliminis and Helbing (2010) – Ph.Trans.Roy. Soc. A

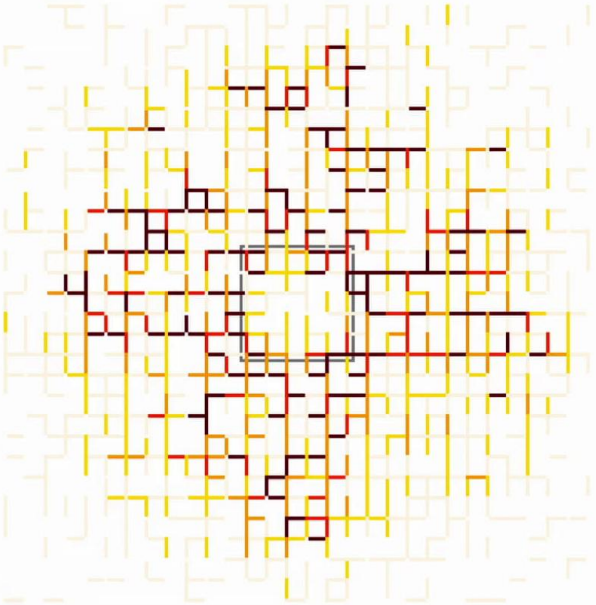
So let's continue more our analysis about the spatial heterogeneity in the shape of the MFD. And now we move from real data back to simulation, because we were able to simulate many, many different cases. What we see on the left, we see a grid Network. Where we have a perimeter and a city center. And vehicles are generated and are moving around. And what we do, we make many different applications with different random seats. So vehicles are generated in a different way but following the same distributions. And we extract the state of its link during the course of the simulation. So here dark color means roads that are congested, while light color means roads that are un-congested. And then Westmead, the spaceman flow, the spaceman density, but we also estimate a third variable which is the standard deviation of a link density. So simply speaking, during a specific time, during a specific snapshot, we take the density of each individually. So let's say here we have 1,000 roads and wasting made the standard deviation of these 1000 roads. And then we can make a plot between these three variables. Average flow, average density and the standard deviation of link density, which is represented in this graph with a different color.

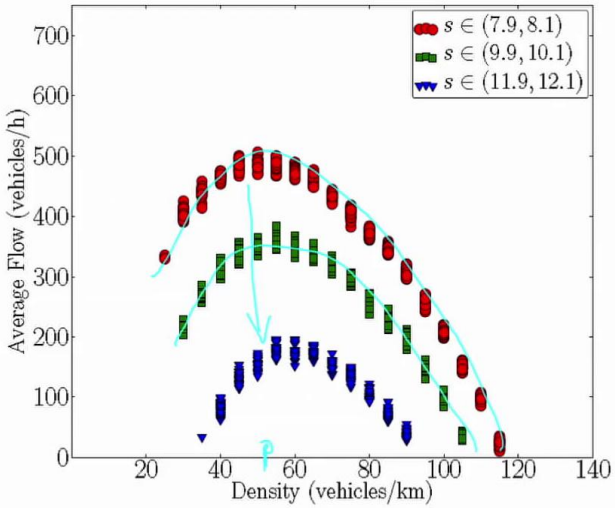
Notes

Summary

7m 40s

Spatial heterogeneity and MFD shape





Mazloumian, Geroliminis and Helbing (2010) – Ph.Trans.Roy. Soc. A

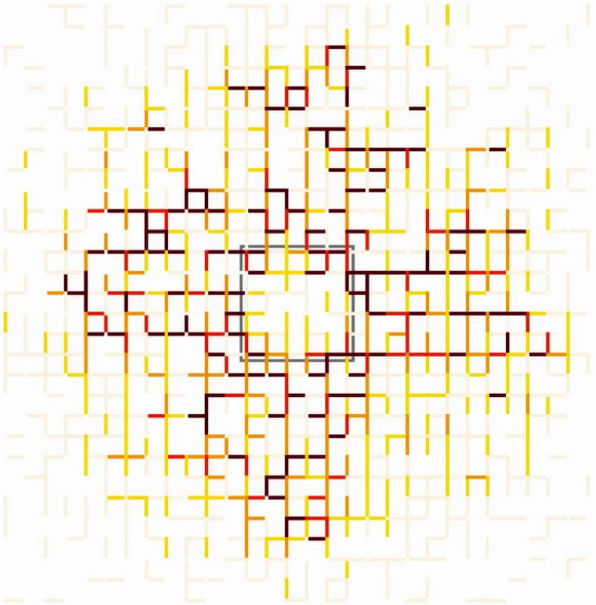
So interestingly and this actually represents results from some hundreds of replications. When the standard deviation is smaller, what we see, we see that we have an MFD which has higher flow compared with an MFD where the standard deviation of link density is higher. So this is a very interesting observation, because this means that even if we have the same number of vehicles in the network which is represented by the density, the higher the standard deviation, the lower the flow of vehicles. And actually, if we take the extreme, we will probably reach almost this point because if the standard deviation is the highest possible, when you have only full and empty roads. Full roads because of gridlock the flow will be close to zero. Empty road, the flow again will be close to zero. So when the standard deviation is maximized, what we have, we have gridlock and vehicles cannot move.

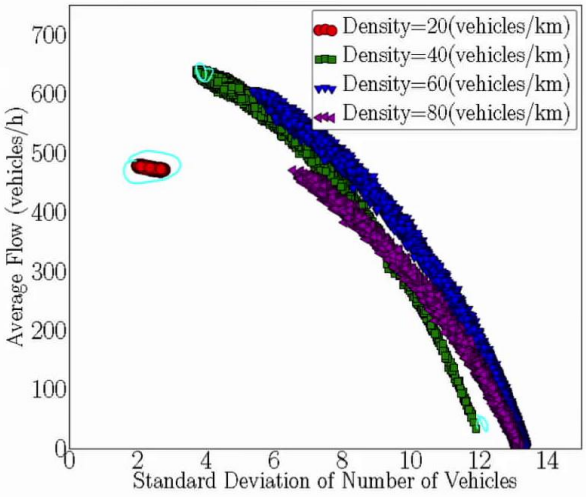
Notes

Summary

9m 15s

Spatial heterogeneity and MFD shape





Mazloumian, Geroliminis and Helbing (2010) – Ph.Trans.Roy. Soc. A

Interestingly, we can see this graph from a different point of view. So here we see the standard deviation of the number of vehicles versus the average flow and the different colors represent the density in vehicles per kilometer during these replications, and of course its point represents an average, let's say of a couple of minutes. When the network is un-congested, so when the density is low, what we see, we see that flow varies very little between a value of 416 and 480 vehicles per hour. But when the density is higher we see that, even when we repeat the same simulations with similar demand patterns, the flow can rate significantly from very high values to very low values. Which actually shows the stochastic characteristics of large-scale transportation networks. And interestingly this stochastic flow is not a completely random event, but there is a very good trend between the average flow and the standard deviation of the number of vehicles at its individual link.

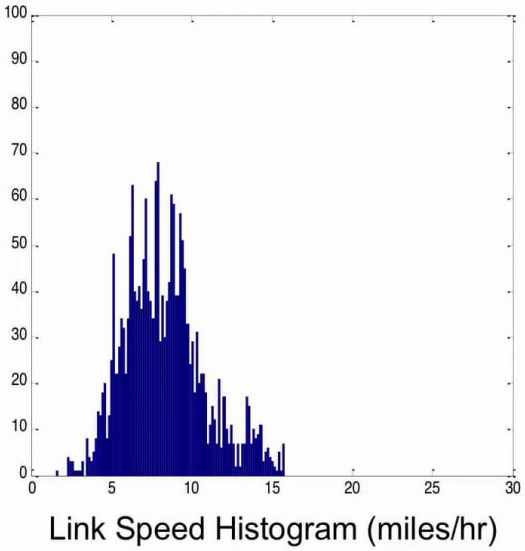
Notes

Summary

10m 27s

Congestion propagation and MFD – Empirical analysis for a megacity

20000 taxis (25M points/day) | 9000 links | 12M population | Shenzhen, China



Ji, Y., Luo, J., Geroliminis, N. (2014). Empirical Observations of Congestion Propagation and Dynamic Partitioning with Probe Data for Large-Scale Systems, in Transportation Research Record, 2422 (2), 1-11. ***(Greenshields' prize)***

So we're interested to see how the propagation of congestion influences the macroscopic fundamental diagram. And to answer this question, we'll do some empirical analysis from a megacity. This is the Centum city in China that has a population of about 12 million and this network has more than 10,000 links and we were able to obtain data from more than 20,000 taxis. That give us their location every 15 to 20 seconds and also stamp if the tax is carrying a passenger or not. And we would like to understand if an MFD can be observed by looking at the taxi data for this very large network. So initially by using this taxi data and after applying some filters, we're able to estimate the speed in every route. We see an animation for a two-hour period between 6:00 in the morning and 8:00 in the morning, where what you see, you see how the speed sends in different parts of the network and if you notice the histogram of the link speed, this moves from the right to the left because the network becomes more and more congested. So clearly the image here becomes darker but at the same time, we don't really understand exactly where is congestion. And one additional reason is because tax data is noisy and when we estimate average speeds on its road of this network, we have also experienced some errors.

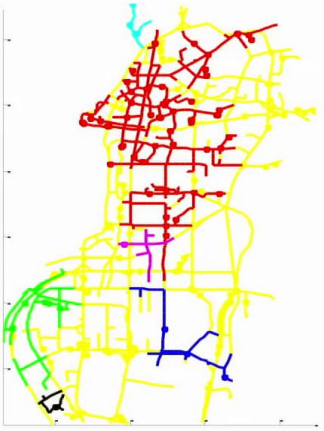
Notes

Summary

11m 43s

Taxi data processing

- Location and speed of taxis every ~20sec
- Estimate link speeds from taxi data
  - 6am - 8am, 15-min interval every 5 min
- Identify congested links
  - Congested link speed <= 1/3 of max speed
- Estimate the largest connected components
  - breadth-first search algorithm



So the way we process the data is that as we have the location the speed of the taxes, we estimate the link speed from the taxi data during the morning peak and actually we apply an average of 15 minutes interval every 5 minutes. What we do afterwards, we give a stamp every time interval for each link and it's linked can be either congested or un-congested. And the simple criterion we use is that we consider a link it is congested if it's link speed is less than one third of the maximum speed which is actually a good approximation that we also observed when we look at individual fundamental diagrams. And then what we do by applying a known algorithm in graph theory which is called breadth-first search, we estimate the largest connected components of the network. So in the picture of the right we see with yellow color the geometry of the network, all the links, and then with the different colors we see the six largest connected components of congestion. So all the red links are congested and interestingly, all these create a connected sub-graph. And then the blue is also another connected sub-graph of congested links but as it is not connected with this roads, with the pink or the red world is a separate. We would like to see how these connected components change over time in space.

Notes

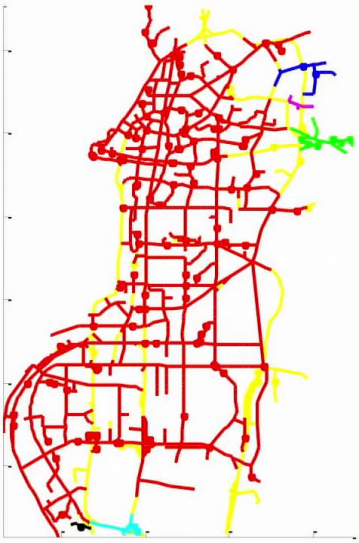
Summary

13m 35s

Evolution of congested regions

Small number of critical pockets of congestion

Dynamic partitioning is feasible with CC



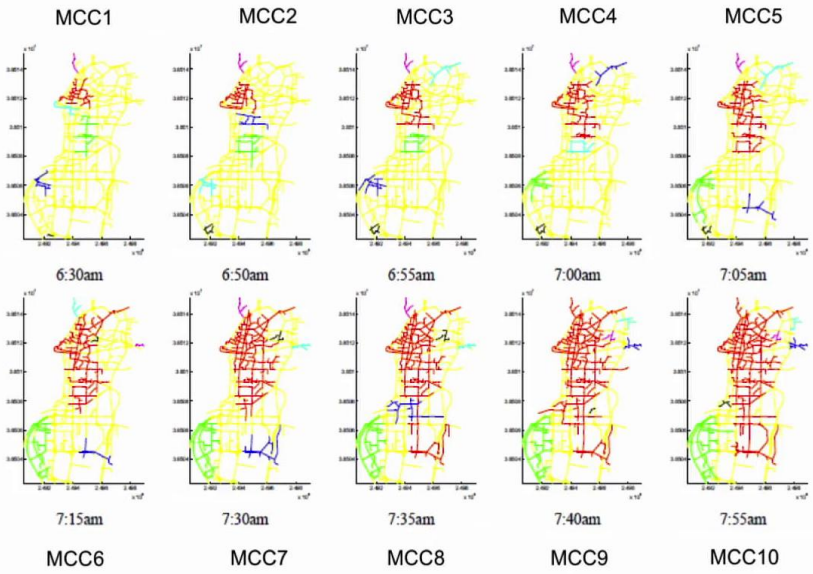
And this is what we have in this movie. So this again shows the congestion propagation between 6 o'clock in the morning and 8 o'clock in the morning for the six largest components. And if we see the movie again, this is a snapshot at 6 o'clock in the morning. So what we see, we see that we'll have only a few congested links and if we start the movie again, what we see, we see that these connected components merged by creating larger ones. And actually in the middle of the pickle around 8 o'clock in the morning, you have a gigantic component of congestion which covers almost the whole network.

Notes

Summary

15m 15s

Snapshots for different time periods



MCC: Maximum Connected Component (Red regions)

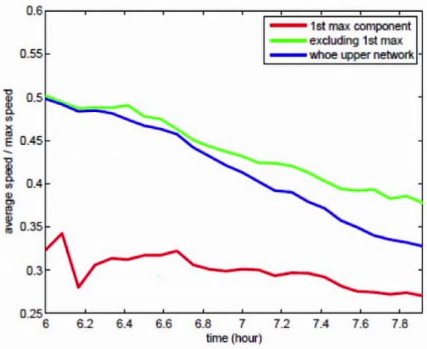
Let's look now at some snapshots of the previous movie for different time periods and here we have 10 periods. And with different colors as we mentioned, so the different clusters of congestion and please keep in mind that the red regions represent the maximum connected component. So we would like to do this quantitative analysis and see how the size of the largest component stay change during the course of the day. And also what we would like to do, we would like that for these 10 red regions which are the largest component during different time periods, we would like to estimate an MFD for each of these regions during the two-hour period and see how they look like.

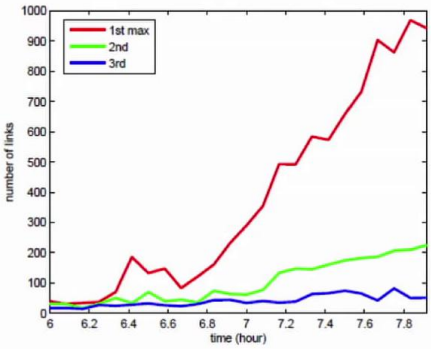
Notes

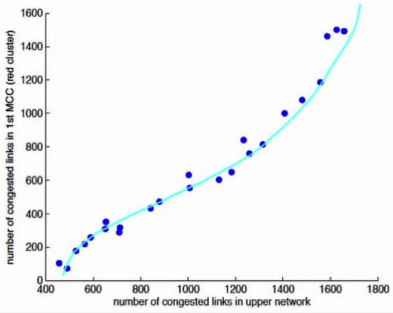
Summary

16m 02s

Congestion evolution of MCC







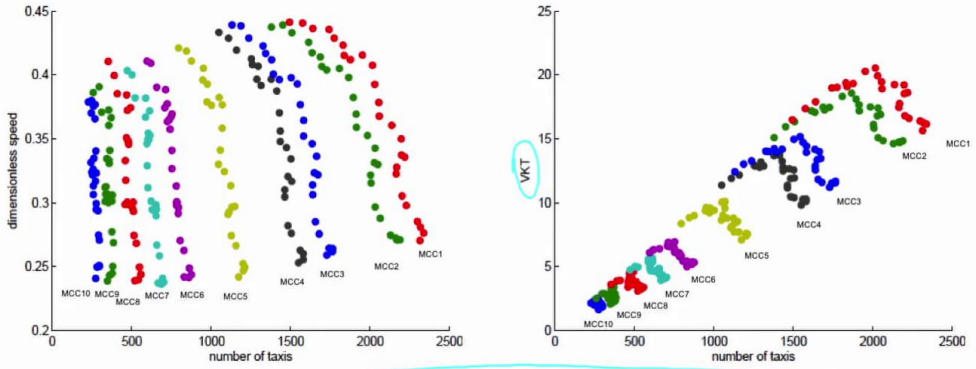
First, let's look the congestion evolution of the maximum connected component. So in this graph, we have the ratio of the average speed over the maximum speed for the largest connected component for the whole network and for all the components included the largest one. So interestingly, what we see, we see that the largest connected component has significantly lower speed than the rest of the network. And what we also see, we see that with the course of the day, the size of the component significantly increases as we also notice in the movie while the second and the third component still remain significantly smaller. And very interestingly, we see a very relationship between the number of congested links in the network and also the number of links and the size is actually of the maximum connected component.

Notes

Summary

16m 58s

MFDs for MCCs



$$n_j = n_u \times (l_j/l_u) \times \left(\left(n_j^f / n_j^e\right) / \left(n_u^f / n_u^e\right)\right)$$

number of taxis is proportional

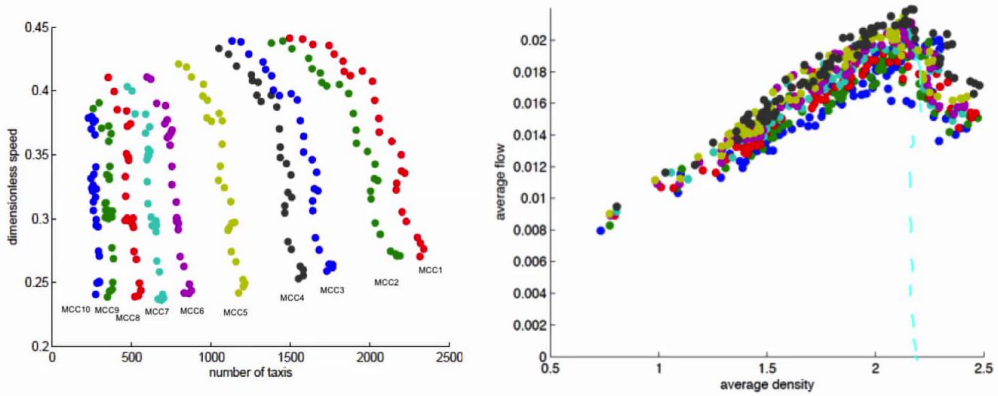
- to the size of the region
- to the ratio of full over empty taxis

It will be really interesting to try to plot macroscopic fundamental diagrams for the different largest connected components that they have been defined in the previous slide. So MCC 1 up to MCC 10, which are the red regions that have different sizes. So in the left figure over here the y-axis is the dimensional split speed. So actual speed divided by the maximum speed. And on the x-axis is the number of taxis which is estimated based on this formula. So while more information is provided in the class material, simply speaking, what we see here is that the number of taxis in region J is proportional to the size of the region and also to the ratio of full of empty taxis. So Nu is the number of taxis in the whole region in the big one. Lj over Lu is actually the ratio of the kilometers of roads between the regions J and U. And here what we see Nf and Ne is the ratio full over empty taxes for regions A and for region U. We see that each region has a well-defined an MFD where the speed decreases as the number of taxes increases. And if we see the same graph but actually instead of the speed, we monitor the vehicle kilometers traveled, what we see, we see that its region has an MFD which is unimodal, but of course because the size of the regions is different, MFD is not overlapping.

Notes

Summary

18m 04s



$$n_j = n_u \times (l_j/l_u) \times \left(\left(n_j^f / n_j^e \right) / \left(n_u^f / n_u^e \right) \right)$$

number of taxis is proportional

- to the size of the region
- to the ratio of full over empty taxis

But interestingly, if we also try to make dimensionless the y axis, the x-axis, sorry, by dividing the number of taxis by the Lane kilometers of roads for this specific region, this is what happens. So the MFD is overlap the one in the top of the other and actually what we can see, we can see that all regions have about the same value of critical density. For which the average flow for this region, but also for all regions is maximized. We'll observe some level of scatter but let's not forget that the penetration rate of taxes compared to the total number of vehicles the network is quite small, probably less than half percent. So this is a very nice result, because then this means that was somebody could use this estimation to develop traffic management techniques. So when you see that the density of taxis is in a region exceeds this critical density, it is a sign that the whole region is congested for all the vehicles and some actions have to be made.

Notes

Summary

19m 48s

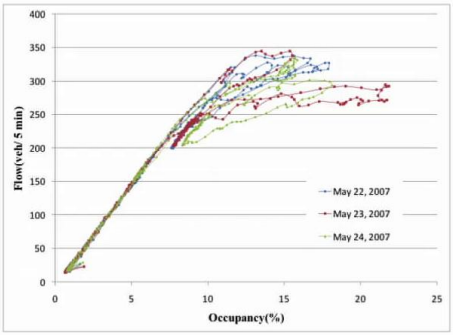
An example of a "bad" MFD

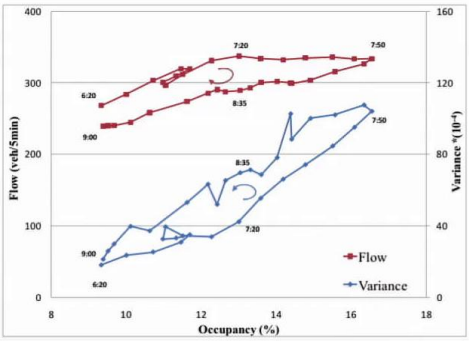
- **Strong hysteresis phenomena in freeway MFDs**

EXPLANATION

- Different distribution of congestion (onset vs. offset)
- Synchronized hysteresis of individual locations (due to capacity drop)

Freeway network of Minneapolis (USA)





Geroliminis and Sun (2011) – ISTTT 19

So as we said in the beginning an MFD is not a universal law. So we need to identify and show some examples that actually by aggregating flow and density for a large network and MFD with low scatter might not occur and here is an example. This is the case of freeway networks. And what we see here,we see a real analysis of a large area with about 500 detectors around the Twin Cities Minneapolis in United States. So what we see in the Left graph. We see the location of the detectors and somewhere here is the downtown city center. So let's see, first, how the MFDs look like for three different days which is the top graph. So spaceman flow versus space mean occupancy. Each day is very different than the other one, but also within a day we observe strong hysteresis loop. So let's take one of the days and let's zoom in so here. We look at May 22, and if we zoom in and we look at flow versus occupancy between 6:20 in the morning and 9 o'clock in the morning, we see these clockwise hysteresis loop. But also what we see in the same graph which is on the other y-axis, we see the variance of the occupancy of individual detectors at a specific time. And what is interesting is that here we follow an anti-clockwise hysteresis loop.

Notes

Summary

21m 03s

An example of a "bad" MFD

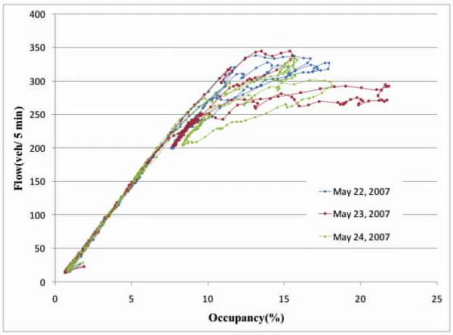
- **Strong hysteresis phenomena in freeway MFDs**

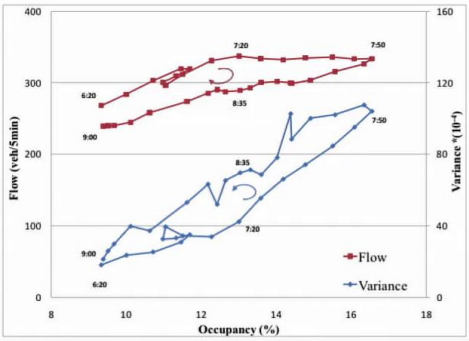
EXPLANATION

- Different distribution of congestion (onset vs. offset)
- Synchronized hysteresis of individual locations (due to capacity drop)

Freeway network of Minneapolis (USA)

Geroliminis and Sun (2011) – ISTTT 19





So look for example a time 7:20 and 8:35, 7:20 during the onset of congestion and flow is much higher and at 8:35 for the same level of occupancy, flow is much smaller. While the story for variance is the opposite. So low-flow means higher variance, which is actually is consistent what we also seen with the simulation results and also empirical studies for urban networks. So let's look a little bit more carefully during different times of the day to better explain the phenomenon.

Notes

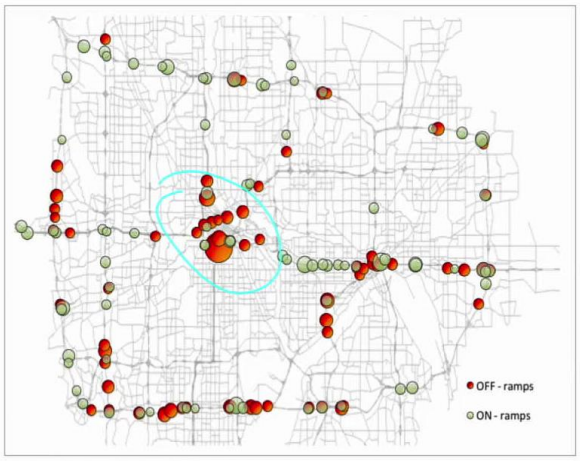
Summary

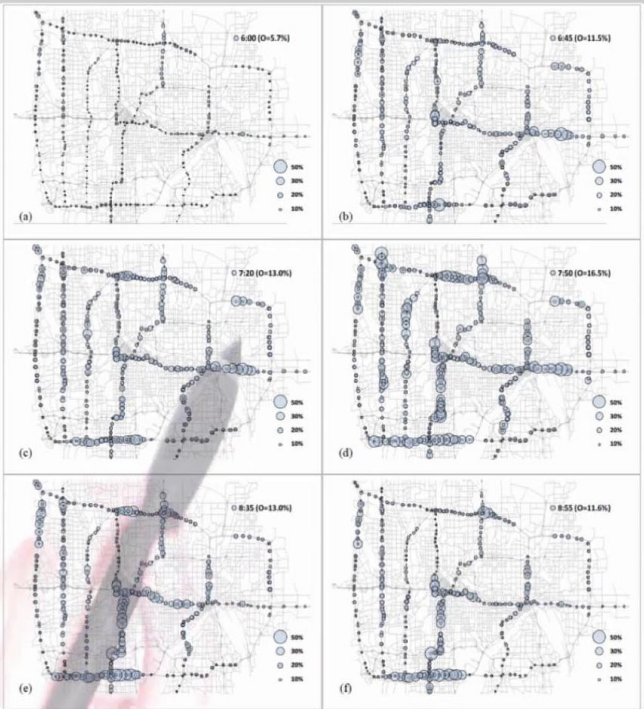
22m 47s

An example of a "bad" MFD

EXPLANATION 1

- Different distribution of congestion (onset vs. offset)





So what we see in the left graph, we see the distribution of origins and destinations in the freeway network which is estimated by Counting how many vehicles are entering the freeway network from the on-ramps and how many are exiting from the off-ramp. And these are aggregated results during the course of the day. So interestingly, what we see, we see that origins are evenly distributed around the city, because actually this is where a lot of people are living, while many, many destinations are around the city centre because actually this is where most of people have their activities and work. And interestingly, what we see, we see that there is a different distribution between the onset and the offset of congestion. And I would like to look more carefully at specific times of the day at the values of the occupants of specific detectors. And this is what we see on the right graph. So here we have six different times and in parentheses you see the average occupancy and the size of the bubble identifies what is the value of the occupancy for this specific detector.

Notes

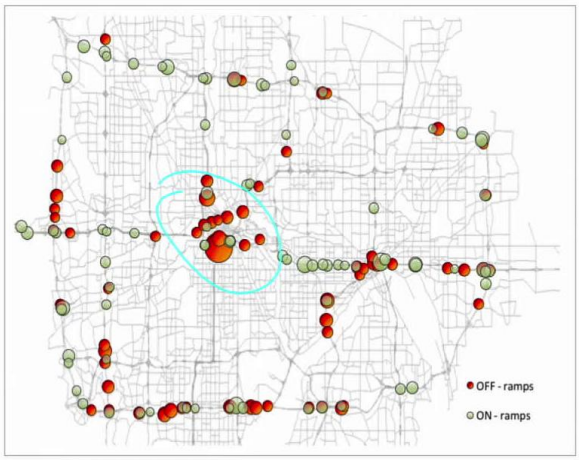
Summary

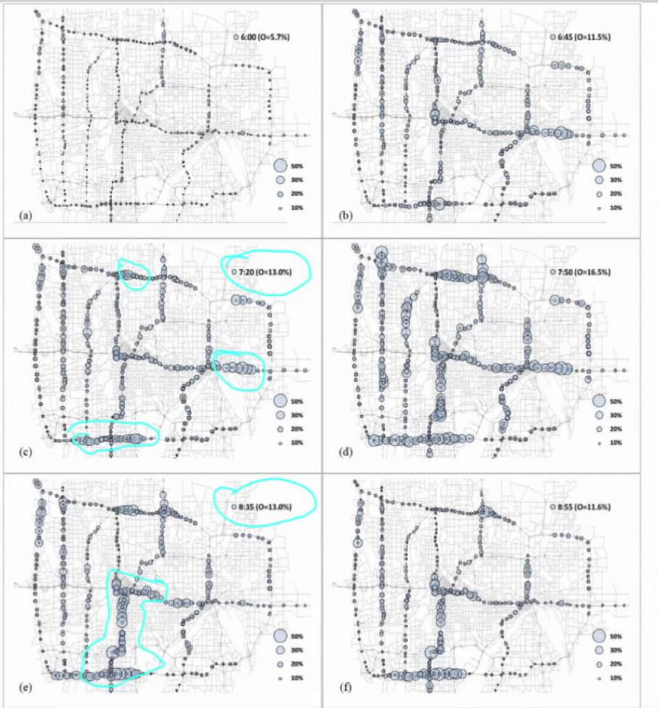
23m 25s

An example of a "bad" MFD

EXPLANATION 1

- Different distribution of congestion (onset vs. offset)





So if we focus at two different times, 7:20 and 8:35 please refer to the previous slide that they have the same value of average occupancy 13%, what we see is that 7:20 there are some regions around the city center that have high values of occupancy which means that these are the active bottlenecks during the onset of congestion. But if we look during the offset of congestion, the location of the active bottlenecks is around the city center. So this very different distribution is one important reason about vehicle hysteresis loop and it is explained through also the different values of variance of the link occupancy of the individual detector occupancy.

Notes

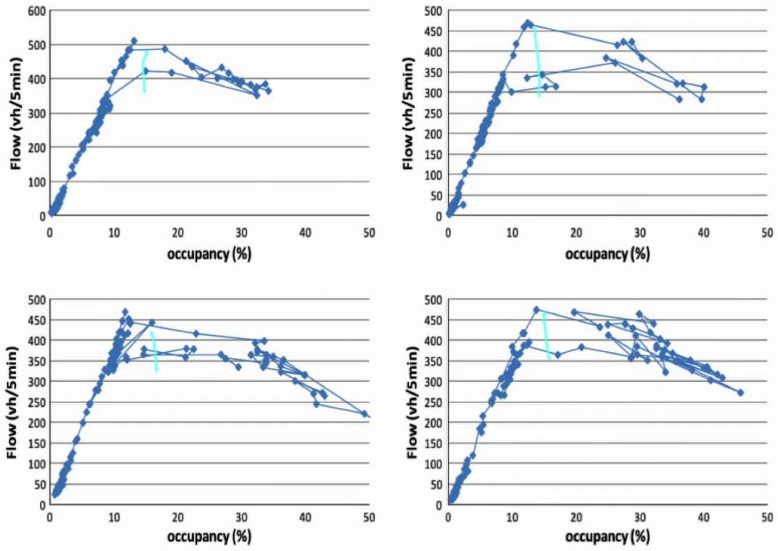
Summary

24m 39s

An example of a "bad" MFD

EXPLANATION 2

- Different distribution of congestion (onset vs. offset)
- Synchronized hysteresis of individual locations (due to capacity drop)



FDs for 4 individual detectors (5min data)

But there is one additional reason why MFD for freeway networks are not always having loss cutter. And actually this region is related with the individual traffic phenomena at single locations. And there is a known phenomenon which is called capacity drop where the flow decreases in the active bottleneck after the breakdown occurs. And what we see in these graphs, we see individual fundamental diagram for four different detectors and what we see is that in all of them there is some hysteresis loop and some type of capacity drop. And very interestingly, what we have observed is that these individual histories and capacity drops happen in the synchronized manner. So most of the detectors become congested during a similar time or with a small time lag and because studies, effects, or Q in many of them this has as a result the flow of the MFD to go down.

Notes

Summary

25m 30s

An intro to MFD clustering

- Objectives Detection of directional congestion
  - –Static/Dynamic
  - –Minimize Spatial regional heterogeneity
  - –Deal with missing data

- Three-step algorithm
  - –Find distinct local components (Snake)
    - Running snakes
    - Defining similarities
    - Reduce search-space
  - –Snake segmentation (optimization approach)
  - –Fine-tuning

M. Saeedmanesh and N. Geroliminis. Clustering of heterogeneous networks with directional flows based on "Snake" similarities, in Transportation Research Part B Methodological, vol. 91, p. 250-269, 2016.

So in the last part of this lecture, I would like to give you an introduction to clustering. So what we have seen, we have seen that MFDs have well-defined relationships when a part of a city a region is homogeneously congested. But as we know in principle, cities are heterogeneously congested, because people have activities distributed all over the city. So the idea is that if you have enough data we can divide the city in different zones, in different regions that each region has a small level of heterogeneity which has a good chance to guarantee a well-defined loss scatter MFD. So our objectives when we do class during are, first of all, we want to minimize the spatial regional heterogeneity. But what we also want, we want to deal with missing data. And another important objective is that we would like to identify directional congestion. So consider for example that during the morning, a lot of trips and activities are towards the center of the city. While during the afternoon, a lot of people are living from the city center. So what we might see, we might see that the two directions of a road might be the one congested and the other un-congested, and in that case we might want to consider that these two directions belong in different clusters.

Notes

Summary

26m 37s

An intro to MFD clustering

- Objectives Detection of directional congestion
  - –Static/Dynamic
  - –Minimize Spatial regional heterogeneity
  - –Deal with missing data
- Three-step algorithm
  - –Find distinct local components (Snake)
    - Running snakes
    - Defining similarities
    - Reduce search-space
  - –Snake segmentation (optimization approach)
  - –Fine-tuning

M. Saeedmanesh and N. Geroliminis. Clustering of heterogeneous networks with directional flows based on "Snake" similarities, in Transportation Research Part B Methodological, vol. 91, p. 250-269, 2016.

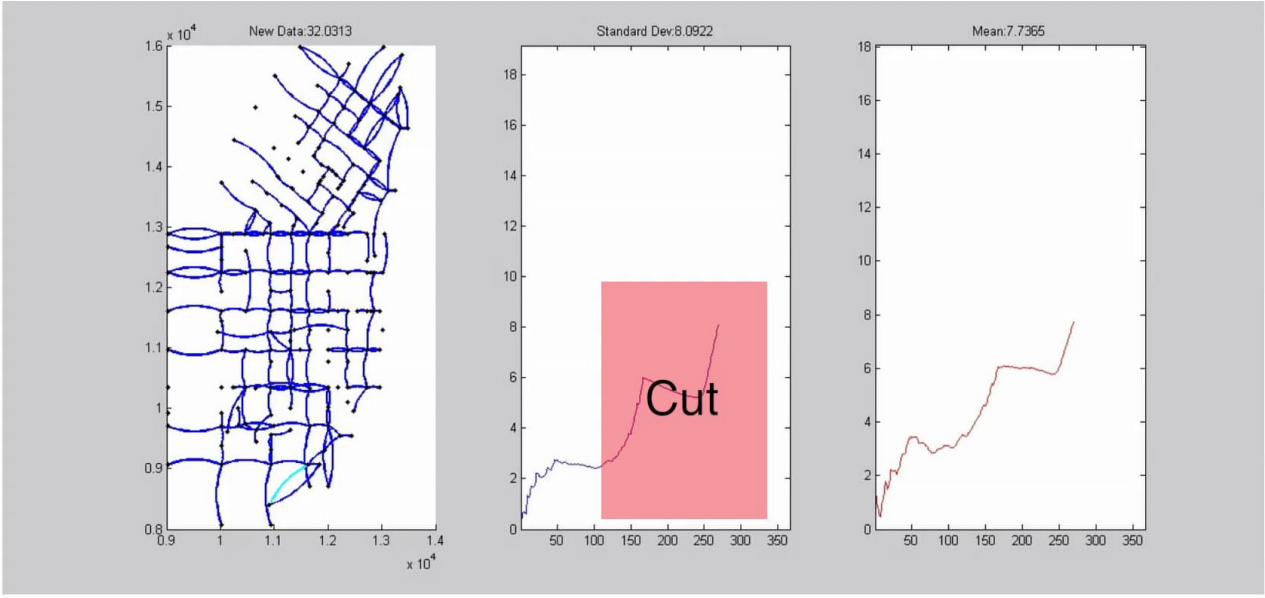
And the last important part of objectives for a clustering is that we also might need to consider dynamic cases, because as we observe from the sensing data, congestion propagates. So here we will not provide a lot of technical details. You can refer if you are interested in some publication that we have included in the class material, but here, I will conceptually present an algorithmic we call it a snake algorithm that has been developed in our team that will utilize to divide the city in homogenous zones.

Notes

Summary

28m 06s

Illustration of Snake Algorithm



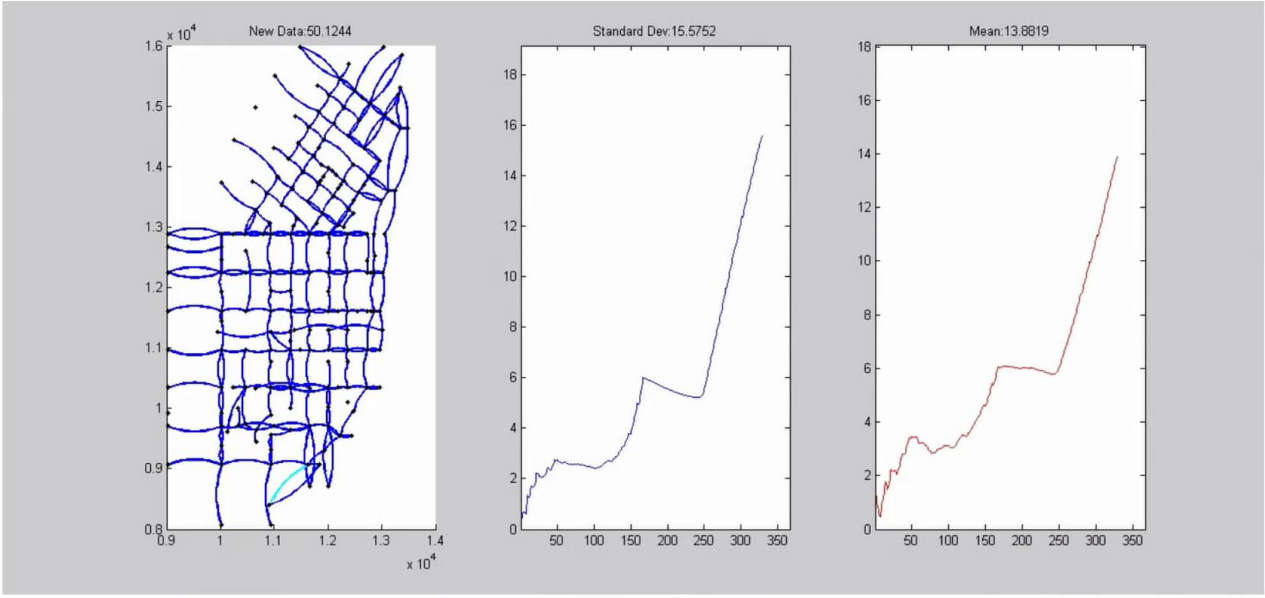
So let's see how the snake algorithm works. A snake has two properties. Property number one, the snake is very hungry. And property number two, the snake does not like different variety of food. So now just place food with congestion and we'll have a good story to say. So snake will start from a specific link that has some special density and as it grows, it will try to attract or let's say eat links that have density as similar as possible to the average of the existing density of the snake and the snake will try to grow in the whole network and let's see how this looks like in the specific video. So the snake starts from one link in the south of the network and it grows. And what we see in the other two graphs, here we see, the size of the snake versus a standard deviation of the link density or speed and here we see the mid. So interestingly what we see, we see that the snake has grown and it has a size of about 120 and the standard deviation remains small, but after this value, the snake cannot identify other routes in the network that have similar level of congestion. And as a result, what it happens, the standard deviation grows. So the idea of this algorithm is that we try to run snakes by starting from different parts of the network and we try to cut some snakes at a level where the standard deviation will be small.

Notes

Summary

28m 40s

Illustration of Snake Algorithm



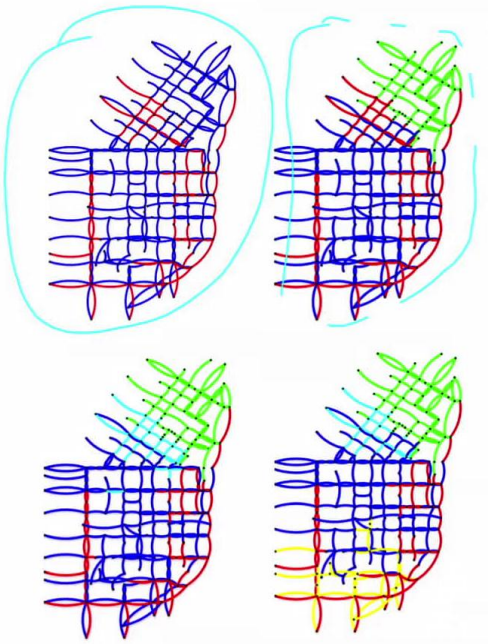
And by construction, if we also consider that we want to have number of snakes so that we cover the whole network, this defines some specific clustering.

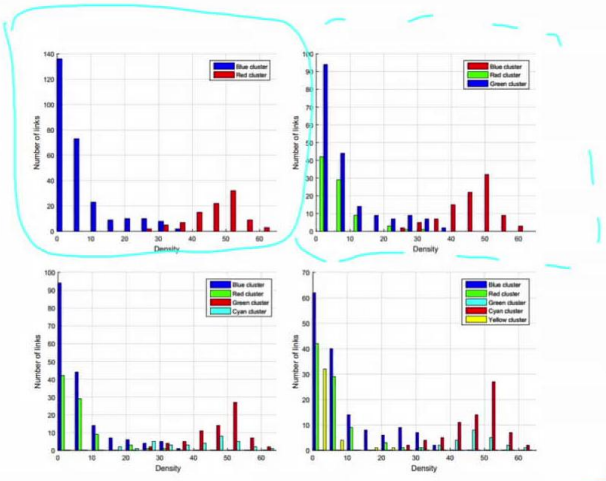
Notes

Summary

30m 15s

Partitioning results – Static case





$$\mathrm{TV_n} = \frac{\sum_{i=1}^{N_s} N_{A_i} \times var(A_i)}{N \times var(A)}$$

(μ/σ)	Blue	Red	Green	Cyan	Yellow	TV_n
2	7.73/8.28	47.43/7.22	-	-	-	0.175
3	8.30/9.16	47.43/7.22	6.48/5.71	-	-	0.174
4	7.11/7.94	47.32/7.38	6.48/5.71	41.15/11.56	-	0.166
5	9.57/9.64	47.32/7.38	6.48/5.71	47.80/6.68	3.36/4.33	0.165

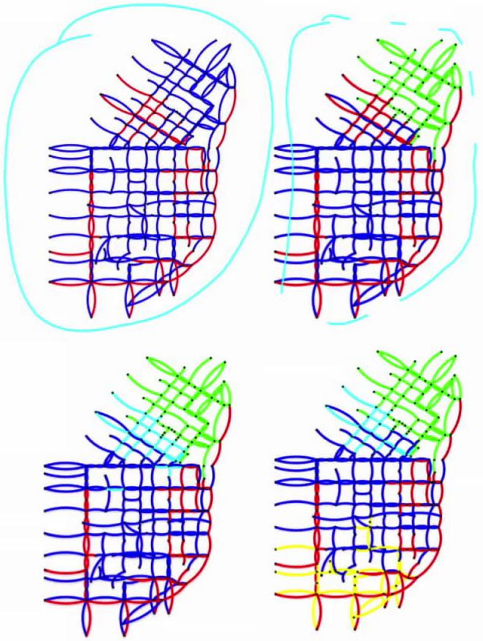
So let's see now some partitioning results for one time instance. And what we see we see the results of the algorithm if we use two, three, four or five snakes. And what we also see in these four graphs, we see the histogram of the speed or density actually here is density of all the rows that belong to a specific snake. So for example, here, I have two clusters, red and blue. And over here we see the distribution of congestion for the links that belong in the blue and in the red cluster. And interestingly, what we see, we see that the red cluster contains links with higher intensity while the blue cluster contains mainly links with lower density. But if we see the example of three clusters, what is interesting over here is that again, we'll have a red cluster which is very different, but the green and the blue have a very similar distribution of congestion. So question is, why we might need to have three clusters? We could survive with only two. The answer on this is that if you notice the green region and the blue region are far from each other. So the one covers the north part of the network and the other covers the middle and the south part.

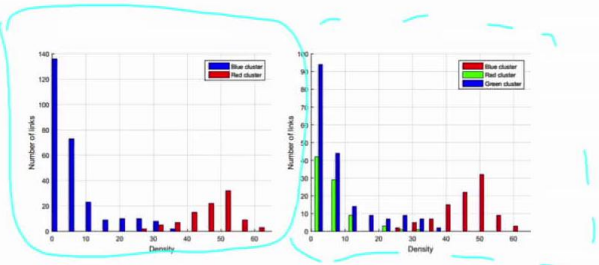
Notes

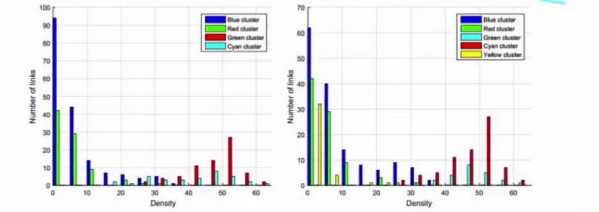
Summary

30m 27s

Partitioning results – Static case







$$\mathrm{TV_n} = \frac{\sum_{i=1}^{N_s} N_{A_i} \times var(A_i)}{N \times var(A)}$$

(μ/σ)	Blue	Red	Green	Cyan	Yellow	TV_n
2	7.73/8.28	47.43/7.22	-	-	-	0.175
3	8.30/9.16	47.43/7.22	6.48/5.71	-	-	0.174
4	7.11/7.94	47.32/7.38	6.48/5.71	41.15/11.56	-	0.166
5	9.57/9.64	47.32/7.38	6.48/5.71	47.80/6.68	3.36/4.33	0.165

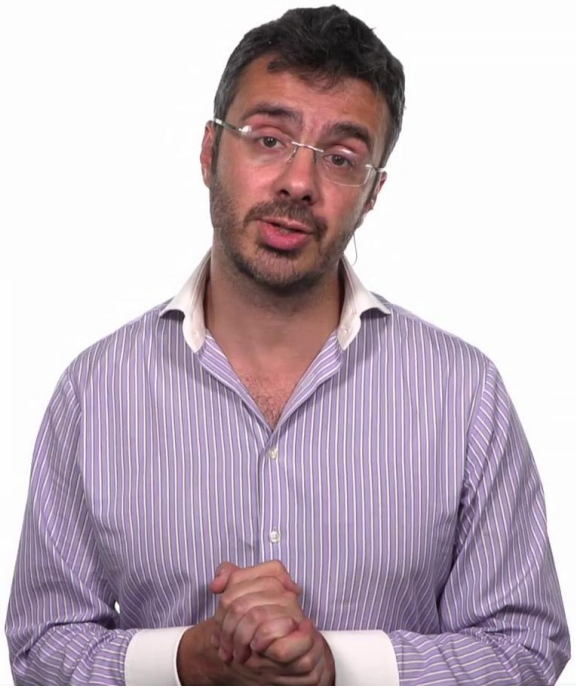
So as these are two different regions and in between you have some red zone which is congested, the algorithm prefers to divide the city in three zones, because what is important is that we have compact regions of congestion. And what we see in the table below, we see the total variance which is a dimensionless quantity, a value that has a value between 0 and 1 that shows what fraction of the total variance remains after clustering. So the smaller the value, the better the clustering. And interestingly what we see is by doing proper clustering, we're able to decrease the variance of the network by more than 80%. And what you see over here, you see for that for the case for two, three, four and five clusters the mean and the standard deviation for each cluster.

Notes

Summary

31m 57s

MFD properties - Summary



So in this video, we saw what are the physical properties that influence the existence in the shape of a macroscopic fundamental diagram and we also saw some examples under what conditions an MFD might could with low scatter and might not acute like we saw in the freeways. And then in the last part, we gave an introduction on how we can divide with different algorithms in city in a number of zones that have homogeneity in the level of congestion. So in the next video, we will deal with the dynamic modeling of traffic congestion with MFDs and we will be together soon.

Notes

Summary

32m 48s