# Avoiding Gridlock

Globally Scheduling “Rotary-ized” Traffic Flows

Stephen Hamilton | CityTram.org | August 28, 2015

## Introduction

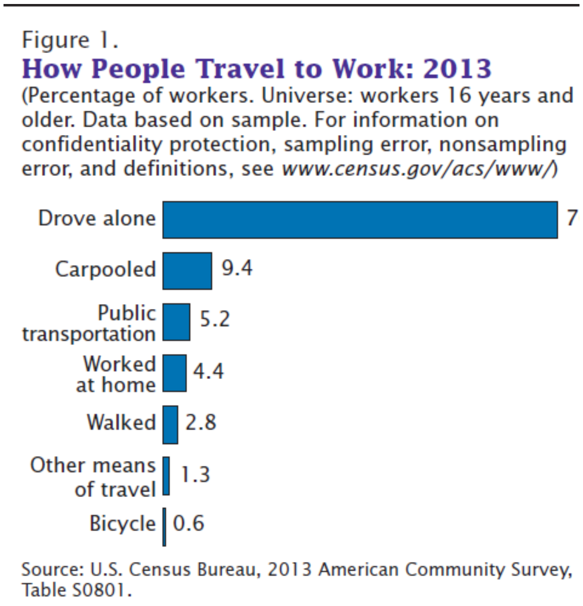
Operation of a personally owned automobile is by far the most common mode of transportation used by urban commuters in America today. It is likely to remain so for some time to come. Coping with traffic congestion on a daily basis is already a routine part of urban life. Congestion is already at such a level that a majority of commuters label it high on their list of complaints about urban life, and actively search for solutions. Since urban populations continue to grow, this pain point is projected to continue and to get worse.

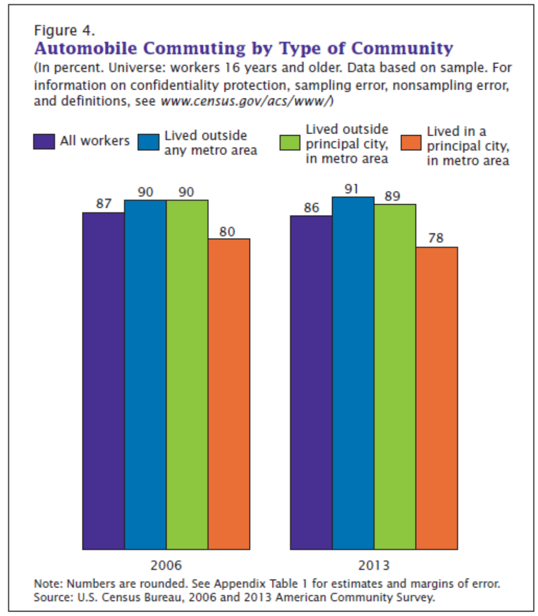
Urban planners and city governments are already experimenting with various techniques to manage this problem, and actively searching for new ones. Most of these techniques either result in higher costs to the commuting consumers, and/or higher taxes for residents, or in loss of convenience to the commuter, or rely upon automotive technology that is not yet proven. They all share in common a motivation to continue to evolve the current car and roadway system in a quantitative fashion, rather than by correcting the fundamental conceptual flaws in the current approach.

This paper identifies what it argues are the fundamental conceptual flaws in the current approach. It proposes solutions to those flaws which can be applied to existing infrastructure. An alternative technology based solution is offered, that has not previously been discussed.

## Congestion is a Fact of Urban Life

According to automotive research firm IHS Automotive there were 253 million cars operating on American roads in 2014 (Hirsch, 2014). With a US population at about 318.7 million (Ortman, 2015) that is roughly 0.794 vehicles per person. With 80.7 percent of the population living in urban areas (rather than rural) (Census, 2012) one might expect that public transit use would lessen the demand for cars to a greater extent. But not so. Truly the personally owned and operated vehicle is the primary way most of us get around, even in cities. The figures below – from the US Census Bureau in 2013 – confirms that. A whopping 85.4 percent use automotive transport to and from work, vs 5.2 percent for public transit. The percentage grows as you get farther from the city center, suggesting conversely that the percentage of public transit use grows as you get closer to the city center.



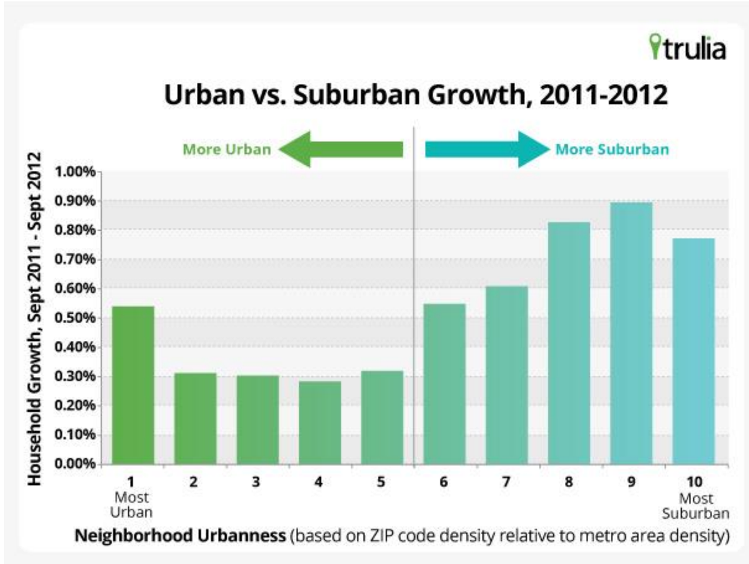




With 219.6 M people (318.7 M x 0.807 x 0.85.4) using cars to commute in metropolitan areas, there is ample load on our roads to create congestion.  The very recent Texas A&M 2015 Urban Mobility Scorecard (Texas A&M Transportation Institute, 2015) quantifies this congestion. It indicates that freeway commutes taking 20 minutes in light traffic conditions require 48 minutes on average. This additional 28 minutes of vehicle operation consumes additional gasoline.  The cost to commuters is the opportunity cost of 42 wasted hours per year, plus $960 in additional fuel costs. In metropolitan areas with more than one million residents the cost is the opportunity cost of 63 wasted hours per year, plus $1440 in additional fuel costs.

The report also shows how these costs have risen since 1982. One conclusion is of particular note. The worst case conditions of “extreme congestion” impacted only 1 in 9 trips in 1982, but impacted 1 in 4 trips in 2014.  This shows how close to the breaking point – over-loaded – our roadways now are, and explains why commuting has become such a pain point associated with urban life.  The report also shows the “reliability” of commute times is under considerable stress.  Congestion can be encountered at any time of day, and with little predictability.  This is causing “planning times” for trips to exceed actual times by growing margins – another indication of how over-loaded the roadways are.

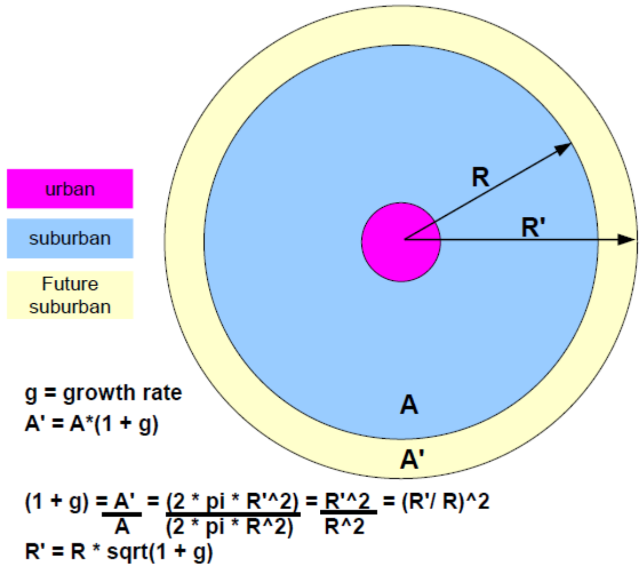
## The Future Is So Bleak I Need To Change Lanes

Every indication exists that the urban commute will only grow worse.  The US population continues to grow, with that growth more in metropolitan areas than in rural areas, and more in suburbs than in city centers.  Therefore metropolitan areas expand physically.  The result is more commuters commuting greater distances, creating greater congestion.




There is a 200 year trend for the US population shifting from rural to metropolitan. As reported by the US Census Bureau (Cohen, 2015) that trend continues. The figure above (Kolko, 2012) shows that this metropolitan growth is occurring faster in the suburban portions of metropolitan areas than the pure urban (city cores). In recent years there has been some variation in this balance, but the longer term trend seems to favor suburbia.

The trend favoring suburban growth over urban growth is reasonably easy to understand. When a number of people move into a metropolitan area there are 3 possibilities for what happens – (1) they move into the city core and increase the density there; (2) they move into the surrounding suburban area and increase the density there; or (3) they move into the area surrounding the suburban area and increase the physical size of the metropolitan area. In reality, what actually happens is a combination of all three of these. But the third of these options is the dominant one.




Since the city center is already surrounded by suburban municipalities it is not possible for it to physically expand. Its only options to manage increased density are to build up, or to shift its land use mixture away from commercial and toward residential. Since commercial properties usually generate higher tax revenues, the choice is most often to support highrise development. The price of real estate tends to rise in response to these changes. The price increases inevitably motivate some businesses and residents to flee the city in favor of the suburbs.

Suburban communities often have zoning restrictions in place which limit the housing density. This has the effect of limiting population density since population density is equal to housing density times family size. With the average family size shrinking, demand for housing units is growing even for a stable population. As housing units become scarce new development is motivated. Once the suburban area is significantly built out, developers are motivated to look outside the suburban ring. Eventually these new developments may be annexed by the suburban municipalities, or incorporated as new municipalities.

So regardless of where new residents to a metropolitan area first reside, equilibrium forces exist. So their entry into the area triggers a set of relocations which ultimately result in the suburban ring physically expanding. People typically seek housing relatively near their place of employment. But after one or more relocations, and possibly one or more job changes, this relationship is no longer assured. So commute distances may be considered to be some fraction of the distance across the city.

The figure above approximates the metro area as a circle. Since area grows by a factor g, so does population (assuming a relatively constant population density in the suburbs), and so does the number of commutes. Since the new communities are almost certainly serviced by extending existing roads, many of the new commutes share the same roads as old ones. Congestion is increased, so average speeds reduce – for everyone. The length of the average new commute increases by the square root of g.

Politics provides natural incentives to feed this negative spiral. Population increases often occur as discrete events. A new business relocates to the area, which results in thousands of new residents.

So the commute load increases discretely also. Road construction on the other hand takes a long time. Road construction also requires funding, which means taxation. For this reason road planning and construction is done in reaction to demand, rather than in anticipation of it. It is not surprising then that road construction is always playing catchup with demand, nor that road construction (which often temporarily reduces lane supply) is done in the presence of increased demand – maximizing the negative impact of the construction.

Understanding the mechanisms that produce congestion make the future of congestion somewhat predictable. Since population growth is predicted for nearly all of the metropolitan areas in the US, it is safe to predict congestion increased for all of those areas. In fact, the 2015 Urban Scorecard referenced earlier has quantified these predictions. The cost to the average commuter is expected to grow from 42 hours to 47 hours per year; and from $960 to $1100 per year, by the year 2020.

## Poor Medicine

Urban planners and city governments are already experimenting with various techniques to manage the congestion problem, and actively searching for new ones. Most of these techniques either result in higher costs to the commuting consumers, and/or higher taxes for residents, or in loss of convenience to the commuter, or rely upon automotive technology that is not yet proven.

### NEW LANES:

Of course road construction scrambles to keep up with the demand. Building new roads would disrupt existing traffic flows the least. But it is more expensive and requires more land use. So adding lane capacity to existing roads is the usual approach. The construction on existing roads often makes congestion worse, at least temporarily. The construction can often last for years. Costs must be borne by the tax payers. Often special bond issues are approved to fund the work. Additionally, it will be shown later why multi-lane roads are still vulnerable to congestion.

### PROMOTING MASS TRANSIT:

Increased use of mass transit can reduce the demand placed upon the roadways, and thus can help ease congestion problems. Incentives are often offered to entice commuters to these alternatives. Unfortunately the incentives are usually financial, while the motivation for car use was convenience and/or transit time. So it is usually the degradation in the automotive commute that motivates the switch to mass transit, more than any financial incentive to do so. The end result is the commuter choosing the lesser of 2 evils – a slower commute, or a much slower commute.

### EXPRESS LANES:

Recently the use of express lanes on freeways has become popular. A portion of the lanes are designated as “express lanes”, and barriers are used to separate them from the normal lanes. These express lanes are converted to toll lanes, often with a variable toll rate based upon the congestion levels of the normal lanes. In essence commuters are asked to pay for the convenience of avoiding congestion delays. It is a cynical admission of defeat by the urban planners. It drives up the cost of automotive commuting, while letting the commuter decide if he pays in dollars or time. It makes congestion worse in the normal lanes on the road the commuter's taxes have already paid for, while extracting additional revenues from the commuter for using the express lanes the commuter's taxes have already paid for.

Express lanes have met with some success in many metro areas. This is evidence of the increased economic stratification in American society – distance between those with and those without wealth. Those with means can afford the fast lanes, while those with lesser means must spend more time / work harder to travel the same path. Thus this approach furthers the stratification. Working multiple part time jobs becomes harder when more time is lost commuting. Meanwhile revenue production incentivizes transit agencies to invest in more toll lanes and fewer normal lanes.

### PROMOTING RIDE SHARING:

Another way to lessen demand on existing roads, and thus ease congestion, is to increase the number of commuters per vehicle. Ride sharing is one way to do that. In many cities the transit agency plays an active role in organizing ride sharing. They manage web sites where sharers can contact each other. They assist local companies in organizing ride sharing among their employees. They designate express lanes on freeways exclusively for multi-occupancy vehicles as a way to incentivize the activity.

Ride sharing has proven to be a resilient and long lasting solution for a small minority of commuters. It is a compromise on the immediacy of commuting in one's own car. It does offer the commuter a reduction in costs, and depending upon the percentage of express lane use, a reduction in travel time. The benefits are only available when at least one endpoint of the commute is in common among the sharers (work at the same place). How well it works depends upon how common the second endpoint is (live pretty close to each other). It also requires reliability in the commute schedule. These requirements have to date precluded ride sharing from being a large scale solution.

### UBERIZING:

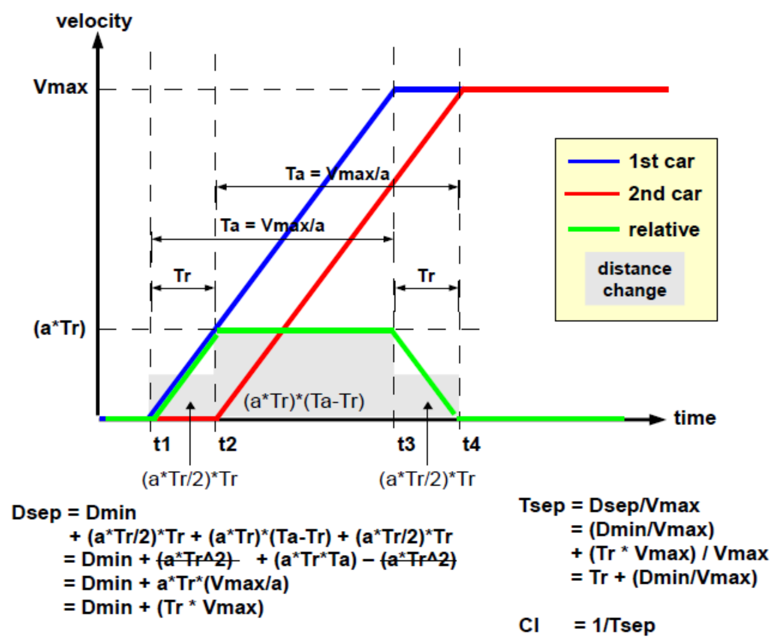
Uber is a company offering a solution they call "ride sharing". But since it is paid ride sharing it might more accurately be called a taxi service. It encourages automotive commuters to offer their commute trips to others. To the extent that it is used this way, it has the benefits of ride sharing. To the extent that it is used as a part time job, it is just a taxi service. Since its use is ambiguous, the company has run afoul of taxi regulations in many metro areas. Some transit agencies are supporting Uber's growth in their areas, some are fighting it. Uber is yet to make a measurable impact on the commute demand the agencies are trying to manage.

## Understanding Congestion

Many of the techniques to reduce congestion being used and/or considered by transit agencies in metropolitan areas share in common a motivation to continue to evolve the current car and roadway system in a quantitative fashion, rather than by correcting the fundamental conceptual flaws in the current approach. That is, they show a fundamental lack of understanding of congestion and its sources (or a lack of boldness to directly address them).

Congestion is simply a high density of cars on the road. This high density is not bad in itself, but it is unbreakably linked with a low rate of traffic flow. Commuters want to get from point A to point B as fast as possible. So the rate of traffic flow is a direct measure of goodness. The roadways themselves are designed to support a certain rate of traffic flow (indicated by the posted speed limits). The degree to which the flow rate is reduced from this designed rate is a measure of congestion.

To understand the relation between density and traffic flow let us consider a trivial example. Two cars stand in line on a single lane road. The second car is immediately behind the first, separated by a minimum distance *Dmin*. At some time *t1* the first car begins accelerating at a constant acceleration rate of *a*. It continues accelerating until it achieves the posted speed limit of *Vmax*, and then cruises down the road at *Vmax* forever. The second car sees the first car begin moving at *t1*, and takes a time *Tr* to react. It then, at *t2* begins accelerating at a constant acceleration rate of *a*. It continues accelerating until it achieves the posted speed limit of *Vmax*, and then cruises down the road at *Vmax* forever.




The figure above depicts this simple example. The blue line shows the velocity of the first car over time. The red line shows the velocity of the second car over time. We are interested in the relative velocity between the cars (green line), as this velocity allows us to calculate the change in the distance between the cars.

Initially (up to *t1*), the distance between the cars is *Dmin*. At *t1* the first car begins accelerating. By *t2* the first car is travelling (*a*Tr*). Since the second car is not yet moving this is also the relative velocity. On average over this period of *Tr* the relative velocity is (*a*Tr*)/2. So the distance between the cars (*Dsep*) increases by (*a*Tr*/2) * *Tr*. The first car accelerates for *Ta = (Vmax – Vmin)/a =*

*Vmax/s*, until $t_3 = t_1 + Ta$. Between $t_2$ and $t_3$ both cars accelerate at the same rate. So their relative velocity remains the same at ($a*Tr$). During this time Dsep continues to increase at this rate, so that by t3 Dsep has increased by another ($a*Tr$)*(Ta – Tr). Between $t_3$ and $t_4$ the velocity of the first car remains constant at *Vmax*, while the velocity of the second car continues to increase. So the relative velocity falls from ($a*Tr$) to zero. On average over this period of *Tr* the relative velocity is ($a*Tr$)/2. So the distance between the cars (*Dsep*) increases by ($a*Tr/2$) * *Tr*. Beyond $t_4$ both cars have velocity of *Vmax*, so the relative velocity remains at zero, and *Dsep* does not change further.

As the figure shows, if we add up all the changes to Dsep between t1 and t4 we see an increase of *(Tr * Vmax)*. Note that it is dependent upon only the reaction time *Tr*, and the final velocity *Vmax*, and not on the acceleration rate *a*. If we take the initial *Dsep* as *Dmin*, then the final (at $t_4$ and beyond) *Dsep = Dmin + (Tr * Vmax)*.

It should be noted that the same calculation apply to braking. If the initial velocity is *Vmax*, and the final velocity is *Vmin = 0*, and the acceleration *a* is negative, then the relative velocity is negative (by the same values), and *Dsep* is <u>reduced</u> by the same amount *(Tr * Vmax)*. To avoid collision we want the final separation to be *Dsep >= Dmin*. So that means the initial separation must be *Dsep = Dmin + (Tr * Vmax)*.

So when accelerating the second car must not accelerate beyond the speed of the first car (*Vmax*), or the relative velocity will go negative and *Dsep* will be reduced below that needed for safely braking without collision.

The end result of his analysis is that the safe following distance is a function of both reaction time and velocity. So the separation between cars must be greater when they are traveling at a greater velocity. Or, unhappily, the number of cars the lane will hold has an inverse relationship with the rate at which those cars are travelling. When you put more cars in a section of lane (reduce *Dsep*) the velocity they travel must be reduced (reduce *Vmax*).

But we don't really care how many cars will fit in a section of lane. We care more about the lane capacity. That is the rate at which cars in that lane will pass by a single point. This lane capacity *Cl* is the inverse of the time separation between cars *Tsep*. If you stand by the side of the road, *Tsep* is the time between two consecutive cars passing by you. This is the time it takes a car to travel *Dsep* at *Vmax*. So *Tsep = Dsep/Vmax = (Dmin + (Tr * Vmax))/Vmax = Tr + (Dmin/Vmax)*. We see that *Tsep* decreases with increasing *Vmax*, so the capacity of a lane to deliver cars (*Cl = 1/Tsep*) increases with the speed of those cars.

The table and figure below shows this relationship explicitly. With the average car length in the 15 to 18 foot range, cars line up in traffic at approximately 30 foot intervals. A human reaction time behind the wheel is on the order of 1.5 seconds – the time it takes to recognize a change in the velocity of the car in front, and react appropriately on the operating pedals. The table shows for these values the separation distances (in feet) for cars travelling at speeds between 0 mph and 70 mph. It also shows the time separation (in seconds) and the lane capacity (in cars per minute). Note that the lane capacity can vary by a factor of 4.

## Lane Capacity

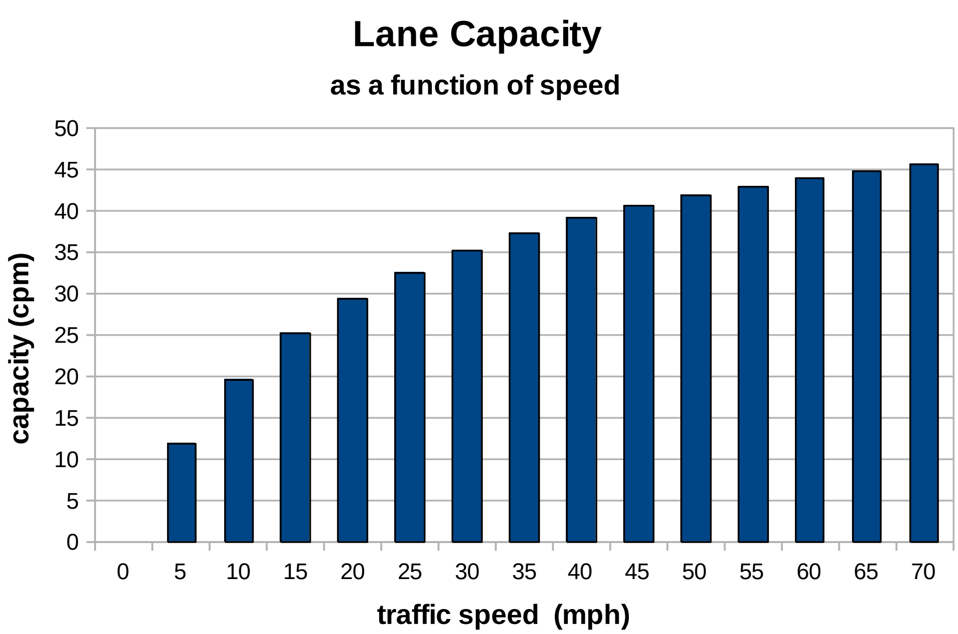
input specs

| | | | |
|---|---|---|---|
| Dmin | 30 | ft | Minimum separation betwe |
| Tr | 1.5 | sec | Reaction time |

conversions

| | | | |
|---|---|---|---|
| | 1.47 | fps/mph | Conversion from mph to ft/ |

| Vmax (mph) | Dsep (ft) | Tsep (sec) | Cl (cpm) |
|---|---|---|---|
| 0 | 30.0 | | 0 |
| 5 | 37.5 | 5.11 | 11.73 |
| 10 | 45.0 | 3.07 | 19.56 |
| 15 | 52.5 | 2.39 | 25.14 |
| 20 | 60.0 | 2.05 | 29.33 |
| 25 | 67.5 | 1.84 | 32.59 |
| 30 | 75.0 | 1.70 | 35.20 |
| 35 | 82.5 | 1.61 | 37.33 |
| 40 | 90.0 | 1.53 | 39.11 |
| 45 | 97.5 | 1.48 | 40.62 |
| 50 | 105.0 | 1.43 | 41.90 |
| 55 | 112.5 | 1.39 | 43.02 |
| 60 | 120.0 | 1.36 | 44.00 |
| 65 | 127.5 | 1.34 | 44.86 |
| 70 | 135.0 | 1.31 | 45.63 |




**This negative feedback effect is crucial to understanding congestion. When the speed of the cars (*Vmax* ) is reduced, the capacity of the lane to move cars (*Cl*) is also reduced**. Whenever *Cl* is less than the demand for the lane *Dl*, accumulation of cars will result. So a lane can be flowing fine, as designed, with *Cl* > *Dl*. Then an event can occur which transiently reduces *Vmax to Vmax'*. This reduces *Cl* to *Cl'* < *Dl*. Accumulation then occurs at a rate of *Dl* – *Cl'*. This accumulation persists until the transient blockage is removed and the lane capacity is restored to *Cl*. The accumulation persists longer while it is reduced at a rate of *Cl* – *Dl*. As long as the accumulation exists it reduces upstream *Vmax* – represents a blockage upstream. How will the time duration of that upstream blockage relate to that of the original transient? If $Dl - Cl' > Cl - Dl$, the upstream transient is longer and deeper than the triggering transient. **So a short interruption in the flow at the front of a line of cars moving down a lane, can grow in duration as it travels down that line of cars, resulting in a long interruption in the flow at the back of the line.** Witness the birth of congestion.

## Sources of Congestion

A clear understanding of the mathematics of congestion leads us to understand that it requires both a trigger event, and a set of conditions that will propagate and amplify this event. The conditions are simple and common enough – the demand rate of cars trying to use a lane (*Dl*) is close to the lane capacity (*Cl*) of the lane. Remember that the lane capacity is set by the posted speed limit, the human reaction time, and the minimum distance interval at which stopped cars line up. Once this load condition exists on a lane, a trigger event can create congestion. The trigger event is always the same – a car at the head of the line reducing velocity. Some occurrences are not avoidable, others are. But looking at all motivations for speed discontinuity events is useful, if for nothing else than to see how much of our pain is self-inflicted.

### PEDAL PUMPING:

It is important to remember that the daily commute is executed by an army of amateur and largely un-monitored drivers. In most states a skills based driving test is given at most once, at the beginning of one's driving career. And the threshold for passing such tests is relatively low, compared to the challenges that will be faced over years of driving. So it should be no surprise that driving skill levels and driving styles vary significantly across the driving population – even when pertaining to the seemingly trivial task of driving in a straight line at a constant speed.

A significant minority of drivers have the tendency to pump the gas pedal, rather than to apply constant pressure. This causes the vehicle to move forward in a series of accelerations separated by coasting decelerations. The velocity of the vehicle, rather than being a constant V, instead follows a somewhat oscillating pattern with a peak-to-peak amplitude of 4 to 7 mph around V. This technique wastes fuel. But worse, in congestion prone conditions it sends wave after wave of speed change events back through the following line of cars, with each wave being amplified by the conditions. This can be enough to start a traffic tie-up.

### SPEED LIMIT CHANGES:

Roads can span considerable distances across a city, and may pass through many different neighborhoods. The characters of those neighborhoods can vary – commercial vs residential, densely vs sparsely populated, etc. The municipality will administer and regulate the road appropriately for the character of each neighborhood. This can result in different sections of the road having different speed limits. As a result, a driver travelling along the road can encounter speed limit reductions. When the demand on the road is high, these speed reductions can be enough to trigger backups upstream of the change.

## LANE CHANGES:

As traffic demand has increased, transit agencies have responded by adding road capacity. The typical expansion project adds lanes to existing roads. As a result, multi-lane roads are quite common in metropolitan areas – both as freeways and as surface streets.

One driving action is unique to multi-lane roads, changing lanes. Lane changes may be legitimately required in order to prepare for a left or right turn, or in order to pass slower moving cars, or in order to avoid blocking faster moving cars. Unfortunately, less skilled drivers also perform lane changes for less legitimate reasons, or for no reason at all other than their own preference for a particular lane.

In congestion prone conditions it is usual that all lanes are near capacity. Therefore it is likely that when one driver desires to change lanes, the lane to which he intends to move is occupied by other vehicles. Often there is another vehicle immediately to the side, or nearly so. In order to execute the lane change, the changing driver must temporarily slow down in order to create a negative relative velocity with that vehicle – to fall behind it. Once the changing vehicle is sufficiently behind the vehicle to the side (Dsep), it can merge from the current lane into the intended lane next to it.

Of course this intentional slow down can operate as a trigger event. Cars behind the trigger car in the original lane will be required to slow down also in order to maintain a safe following distance. But unless cars in the intended lane are already *(2 * Dsep)* behind the car to the side, they will be required to slow down also after the trigger car merges into that lane. So a lane change often produces a speed discontinuity event in two adjacent lanes.

Especially in congestion prone conditions the speed discontinuity event can be larger than minimal. The trigger car may be required to wait longer in order to merge. If several cars are already close behind the car to the side, the trigger car may have to fall behind all of them before changing lanes. Obviously lane changes provide an additional dimension of negative feedback to congestion. The more congested the roadway is, the more likely it is that lane changes are difficult and costly – waiting longer to find the opportunity to change. This creates bigger accumulations, which take longer to clear. This packs more cars into a smaller space, which both creates more opportunities for additional lane change events, and increases the cost of each of those additional lane change events.

## LANE MERGES:

A lane change occurs between 2 lanes running parallel and side by side. Lane merges are similar, but they occur when 2 parallel and side by side lanes are structurally reduced to a single lane. Sometimes lane merges occur in-line along a roadway, when anticipated demand is lower. Perhaps a 4 lane freeway is designed to handle the anticipated demand near the heart of a city, but as that freeway reaches the outskirts of town first 3 and then 2 lanes are expected to be sufficient for the anticipated demand. Lane merges also occur at intersections between roads. Entrance lanes on freeways are lane mergers. But even many surface streets have short acceleration lanes near intersections.

The lane merging action is essentially the same as that of lane changing. Cars in the two merging lanes must match speeds and create alignments with each other prior to the merge point. The same speed reduction and fall back may be required. Therefore the same congestion event triggering impact is possible.

## FEEDER TOPOLOGY:

By now it should be clear that in order to minimize congestion it is desirable to minimize the occurrence of lane change and lane merge events. Unfortunately, current roadway design practices instead maximize occurrence of these events.

As roadways are planned for a metro area, and then evolve as that area grows, the usual approach is to implement a feeder topology. Neighborhoods – especially residential neighborhoods – usually resist having high volume traffic pass through them. Such traffic represents safety risks, inconvenience, and noise. Also, the cost of providing many high volume roads is prohibitive. And expanding roads in such areas to cope with increased demand is very disruptive to the neighborhoods. The best solution to balance all these concerns with the reality of traffic demand is a feeder topology. Neighborhood roads feed larger surface streets, which feed yet larger boulevards, which feed yet larger surface highways, which eventually feed freeways, which finally feed larger super freeways. Much like the blood vessels in a human body, and the branches in a tree, many smaller vessels feed into fewer larger vessels, and the pattern repeats.

While this topology is good for all the reasons listed above, it is bad for other reasons. It adds distance to commutes, since the route is not as direct. But worse, it is far more congestion prone than less hierarchical, more grid-like topologies.

The feeder system is full of lane mergers. When a small road merges into a larger roadway, the number of lanes following the merger is less than the sum of the merged roadways. The reduction is accomplished by a series of lane mergers. In fact, in most cases, the number of lanes in the larger roadway is unchanged before and after the merger or intersection. All lanes of the smaller road are merged into one or more acceleration lanes into the larger road, which eventually are merged into a single entry lane.

The feeder system also maximizes lane changes. Since the topology aggregates the load onto larger highways, those highways require multiple lanes to be capable enough. Typically highways are designed with entrance and exit lanes on the right, and continuation lanes on the left. Occasionally exit or left turn lanes are placed on the far left. This means that cars entering the highway must execute lane changes to the left to reach the continuation lanes, while cars leaving the highway must execute lane changes to either the right or left to reach the exit lanes. When a major exit occurs shortly after a major entrance, the lane changing streams can overlap, which increases their impact.

## ACCIDENTS:

The automotive transport system is designed for vehicles to travel from one place to another without coming into contact with each other, or with any other objects along the roadways. Unfortunately, it does not always operate that way. Occasionally accidents happen. When an accident occurs, the involved vehicle(s) must stop on the roadways for some period of time. In very few exceptional cases, the state permits the involved vehicles to be moved quickly, if they are capable of being moved, and if there is sufficient space on or near the roadway to move the vehicles out of the lanes. It is far more common that the damaged vehicles simply remain in place, blocking one or more lanes, for a considerable period of time. This has the effect of reducing the lane capacity for the blocked lanes to zero, and so significantly reduces the capacity of the road.

Accidents are most likely to occur when congestion prone conditions are already present. Since the distance between cars (*Dsep*) is already small, and speed discontinuity events are more frequent, cars are more likely to come in contact with each other. When an accident occurs in these conditions it is almost certain to reduce capacity of the road below the level of demand, and create a congestion accumulation. The existence of congestion further increases the likelihood for an additional accident (as Dsep is at its minimum, speed change events are at their maximum, and driver patience is strained). This is why clusters of accidents often occur at rush hour.

## UN-REGULATED ENTRY:

Congestion occurs when the lane demand (Dl) for a lane on a road is less than the lane capacity (Cl) for that lane. Lane demand (*Dl)* is a matter of choice, and so is possible any time the total demand on a road, across all lanes, exceeds the capacity of any lane. But lane choice tends to exhibit positive feedback. When one lane appears crowded drivers tend to choose the other lanes. So congestion usually does not result until total demand across all lanes nears total capacity across all lanes.

Total demand across all lanes is much more predictable. It is the aggregate of a large number of commute routes and times which tend to repeat daily. So the origin of congestion is also predictable, and it tends to occur in the same places day after day.

With very limited exceptions, the roadways in metropolitan areas offer un-regulated entry. Any driver anywhere who wishes to enter the roadways is free to do so at any time. There is some use of metered entry on freeways (the limited exceptions). But this metering is applied in the middle of a commute, where it switches from surface streets to freeways, not at the start of a commute. This metered entry does tend to reduce congestion on the freeway, but at the cost of creating congestion on the surface street. No metered entry (or other regulation) exists at the start of commutes.

The reality is that some demand patterns simply exceed the capacity built into the roads. The populations of most metropolitan areas have grown to the level where these excessive demand patterns are commonplace (perhaps daily) occurrences. The choice is whether we allow these demand patterns to be applied to the roads, knowing they will produce congestion; or we somehow convert them into other demand patterns that will produce less (or no) congestion. It would be better for a few drivers to wait less time at the beginning of their trip, than for more drivers to wait more time during their trip. But this would require coordination of all trips. We currently have no such coordination.

## TRAFFIC LIGHTS:

As shown above, the trigger events for congestion are single vehicles performing speed reductions. Our metropolitan roadways are populated with literally thousands of traffic lights whose operation is for no other purpose than to constantly create speed discontinuity events. Every minute or so each light changes, resulting in dozens of cars in each of two directions executing a stop, and dozens of cars in the orthogonal two directions executing a start. Since the word “traffic” is a synonym for the word “congestion”, these devices are appropriately named. They can be properly thought of as congestion producing machines.

Traffic lights are placed at places where roads intersect orthogonally. They enforce time multiplexed sharing of the common pavement, thus avoiding collisions between cars. But the time multiplexed sharing of that common pavement has the effect of severely reducing the capacity of each intersecting road – by a factor of 2 or more. This lowers the threshold – by a factor of 2 or more – at which demand can exceed capacity. It thus makes congestion much easier to create.

Traffic lights also shape the down stream demand into a periodic congestion prone form. When the light turns red a set of cars that were spread out begin to decelerate and pack together maximally (*Dmin*) on the road. When the light turns green those cars begin to accelerate. They are maximally packed together on the road, and will be traveling at approximately the same speed down the road. So for that pack of cars the demand (*Dl*) is exactly equal to the capacity (*Cl*) - a congestion prone collection.

The traffic shaping behavior of traffic lights makes it easier for new cars to enter the road The periodic behavior has an "on" time where the road is at the capacity Cl, followed by an "off" time where the road is essentially empty. This permits new vehicles to safely enter the road to begin their trips. Unfortunately this upsets the equilibrium of the road. If the road has a series of traffic lights with the same cycle times, then roughly the same number of cars should be able to pass each light during green. But if new cars enter the road, that number may not be sufficient to empty the road. The red light may then catch cars not yet through the next light. Over time, this process accumulates enough cars to block the road. A car travelling down the road may require several light cycles to pass each light. New cars have to wait longer to enter. But they still tend to enter as the line creates openings when the light turns green. This continues the overload (it would be best to prohibit new entry until the accumulation clears).

Eventually the surface street congestion can back up sufficiently to block exit lanes from freeways. The exit lanes then accumulate, perhaps sufficiently to eventually block the through lanes. The through lane congestion then manages to block itself, somewhat independently from the exit lane congestion.

## The Flies in the Ointment

The automotive transport system as deployed and in use today in all US metropolitan areas suffers from certain conceptual flaws. These flaws severely limit the practical capacity of the system to a level far below the current demands. This is why those metropolitan areas suffer from present and growing congestion problems. These flaws are discussed here.

To call the flaws "conceptual" may be unfair. In most cases some choices were made when designing the system. These choices were made as a compromise between competing needs, with consideration for current practicalities. These choices created a conceptual framework for how the system works. Many of these choices were made nearly one hundred years ago. The basic framework has been evolved, mostly quantitatively since then. Given today's demands, technology, and other practicalities, the original compromises may look "wrong". If we are to avoid total gridlock of the current system, we must not be afraid to take that look, nor to consider changing these fundamental choices.

## TRAFFIC LIGHTS:

The single greatest impediment to smooth flow of traffic in the automotive transport system is the use of traffic lights. Fundamentally they are a compromise of performance in order to secure safety at low cost. The details and mathematics associated with traffic lights is provided in a separate paper by this author (Hamilton, 2014). The bottom lines is that they reduce the lane capacity for our roads by more than half, and they shape the traffic in ways that are not robust – capacity tends to collapse when under demand pressure.

Freeway systems avoid intersections altogether by using the 3rd dimension (overpasses). Available physical space and construction costs made that an impractical solution for most urban intersections. But at today's levels of congestion that cost trade-off might be different. The cost-benefit of elevating a city street might be better than that of adding extra lanes to freeways or city streets. There are also aesthetic and architectural barriers to elevating some city streets.

The fundamental idea of coordinating the time multiplexed sharing of an orthogonal intersection may still be useful. But the current application of that idea is in the context of human vehicle operators. The mathematics reveals that human compatible reaction times result in traffic light cycle times that are far too slow for city layout/roadway design and commute demand. It is this slow reaction time that requires the coordination on groups of cars which then must stop and restart, rather than on individual cars. But technology is now available to support semi-automated vehicles. This makes faster reaction times a possibility, so a solution based upon vehicle synchronization rather than stop-start is feasible. Multiplexing of the intersection can occur at a much faster rate, on a per-vehicle basis rather than on groups of vehicles, so those vehicles never need to stop. Each vehicle passes through the space created between vehicles in the orthogonal stream (the following distance). This is a considerably more controlled version of key feature you might have seen figure 8 racing! The level of automation control this requires is quite high, and it would be required on all vehicles (in order for all vehicles to stay safe). So while it is technically feasible today (and will be more easily so with the coming V2X technologies), the practicalities of deployment seem problematic.

Recently the idea of eliminating traffic lights, road lane markings, and even street signs in urban centers has gained some credibility (Badger, 2011). In pedestrian dominated areas where vehicle speeds are low this may have merit. But it does not likely offer a solution to the general commute problem.

Rotary intersections are well proven in Europe. Since they operate as lane mergers, potentially with multiple lanes and requiring lane changes and interleaving, they are a source of congestion. But, there is strong evidence they are a far better solution than traffic lights. They see some use in America – usually in residential neighborhoods with light traffic and lower speeds, or in transit hubs such as Airports. In order to support higher vehicle speeds and vehicle capacities, quite large rotaries are required.

## UN-REGULATED ENTRY:

In today's system entry regulation, if it exists at all, is general and soft. The transit agency may work with local employers to encourage flex-time work hours in an attempt to spread out the demand. But there are no specific localized and hard entry barriers. So most metropolitan areas are subject to daily occurrence of peak demand periods when demand far outstrips capacity in multiple locations.

The current solution chooses to pack as many cars as possible onto the roads, which lowers the distance between them (decreasing safety) and slows them down (increases travel time). It also increases the velocity sensitivity - size of event that can cause a slow down, and the amount of the slow down (further decrease in safety). Queuing (congestion accumulation) occurs on the road (usually multiple times per trip).

An alternative would not stop egress to the roadways, but rather control its rate. Egress rate would be matched to the capacity of the system. Queuing would occur once, at entry. After entry, velocity would be higher (less travel time) and more stable, with a greater vehicle separation distance (more safety). In such a system congestion would be rare. Travel times would be relatively predictable, although total transit times would vary as a function of the demand rate. The variable delay would be spent queued for egress at the start of the commute.

## LACK OF CENTRALIZED SCHEDULING:

When a commuter enters the roadway system to begin a trip, he is in essence spending commute resources. In the following minutes his vehicle will require a space in a lane all along the route he intends to travel. There is no way of knowing in advance that resource will be available when the vehicle arrives. Lack of entry regulation makes it likely multiple vehicles will spend that resource, so that when all but the first of them arrives the resource will not be available. Since the resources are self-restoring, the vehicles will simply wait until the resource is available. All during those waits the vehicles will be consuming resources needed by other commuters.

The political right wing in America has been very successful at embedding into the populous the meme that an unfettered competitive (economic) system will naturally evolve to a globally optimum (efficient) solution. Since the 1920s this is mathematically known to be false. Competitive participants lack global visibility. They will optimize based upon their local and short term visibility. Local winners will lock out more distant competitors. This locks the system into a Nash equilibrium which prevents it from achieving the global optimum. This mathematical truth applies to our automotive transport system (Baker, 2009) (Gastner, 2008). Letting commuters do what they believe gets them the best progress actually slows the system down for everyone.

Congestion is a local phenomenon. Even a regulated entry system will have trouble avoiding congestion if it is not aware of the intended route. But with today's technology it is easy to imagine an app for cell phones and GPS devices to support centralized scheduling and entry regulation. The commuter enters the desired destination, and selects among suggested routes. The scheduling system then reserves system resources for the trip and based upon their availability at the appropriate time gives the driver authorization to begin his trip. Assuming reasonably predictable travel times for each segment of the trip, the vehicle should arrive at each resource at the predicted time and find it available for use.  To help ensure those predictable travel times, the app can provide feedback to the commuter to adjust his pace during the trip.

## FEEDER TOPOLOGY:

The number of paved lane miles in each of our metropolitan areas is astronomical. But the feeder system attempts to concentrate load on a small subset of those lane miles.  Given our understanding of congestion and its sources, this is counter productive.  Travel times are negatively impacted by traffic lights.  Therefore longer commutes are biased towards major trunks/freeways.  If traffic lights were synchronized to avoid stopping, this would not be true.  Spreading the demand across the maximum number of parallel roads minimizes congestion opportunities.   So while wholesale abandonment of the feeder strategy is impractical – too many existing roads and neighborhoods - a small reversal adjustment is clearly called for.  The key making that adjustment possible is maximization of capacity on existing trunks.

## LEFT TURNS:

Roads are typically layed out in a topology approximating an irregular two dimensional grid.  So point to point trips (which look like the hypotenuse of a triangle) are executed by following a Manhattan geometry (along the edges of the triangle).  The route goes first along one dimension (north-south), and then turns 90 degrees and proceeds along the other dimension (east-west).  Often multiple such turns are taken, with each only covering a portion of the distance along that dimension.  There are two such Manhattan paths for each hypotenuse.  For example, one path goes north and then east, while the other path goes east and then north.  The first uses exclusively right turns, while the second uses exclusively left turns.  So our roadway systems must support turns in both directions.

But turns in the two directions have drastically different costs to the road system.  Given the American standard of driving on the right hand side of the road, left hand turns are very problematic and costly.  Left hand turns require the vehicle to pass across traffic lanes coming in the opposing direction.  Therefore they have much higher safety risk (head on collisions).  In high demand conditions there are also far fewer opportunities for this crossing, so left turning cars wait longer, creating more blockage.  To mitigate these costs left turn lanes and left turn signals are often used.  But these lanes are at the expense of additional through lanes, and these signals further reduce the duty cycles of the traffic signals.  Both hurt potential capacity of the roads.

The problem case is left hand turns from bi-directional roads (lanes in both directions).  So the problems can be avoided either by eliminating the left hand turns, or by eliminating the bi-directionality of the roads.  Given that each trip has a Manhattan path requiring only right hand turns (unless the grid is incomplete), this would seem the cheapest solution.  But both options are available for use.

# Proposing a Globally Scheduled Rotary Based Network

The existing automotive transit system in any single metropolitan area is huge! It encompasses thousands of roads, thousands of traffic lights, hundreds of thousands of cars, and most importantly hundreds of thousands of drivers. It touches dozens of neighborhoods. Changing it in any significant way will be a monumental undertaking. The changes must occur incrementally over time, and will encounter resistance to change on all fronts by many parties.

Yet the existing systems are at their limit. It is past time to design our way into a better, more scalable solution. The previous "Flies" section was primarily to consider the fundamental attributes of the current system, and to decide pragmatically what can and what cannot be changed, and to consider what replacement options exist. Now, hard choices must be made to move toward a better future. It seems impractical to abandon the feeder system of roads we have; while it seems desirable and essential to eliminate traffic lights and traffic crossing left turns. Making much greater use of rotary intersections seems the most practical way to achieve these goals. Then relying upon the current and coming driver assistance technology to help facility the merging activity inherent in rotary intersections further minimizes the congestion creation attributes that remain. Eventually an entry regulated, globally scheduled road system is the goal.

What is proposed here is a goal or vision of what could be, and some suggestions of what might be useful steps toward obtaining it. Likely any attempt to travel the journey from today to this vision in a specific urban area will have some unique challenges not covered here, and will require some local creativity. Some of the specific technological components described here do not yet exist. But they are well within the capability of technology today, and so would require only a short interval to develop and deploy – perhaps as much as several years, but likely less.

## OVERVIEW:

The goal is to promote and ensure traffic **flow**, which is the same thing as ensuring velocity – the antithesis of traffic lights. So the goal is to eliminate as many traffic lights as possible, and specifically to eliminate **all** traffic lights along certain trunks through a development zone. What is proposed is a "continuous motion" system. Freeway systems are the continuous motion system with which you are familiar. Vehicles enter the freeway system at controlled points, accelerate to match the speed of through traffic, and then merge into the traffic flow. The vehicle remains in continuous motion all the way to its destination. Near its destination the vehicle moves into a deceleration lane and then exits at a controlled point. What is proposed is a continuous motion system for surface streets as well.

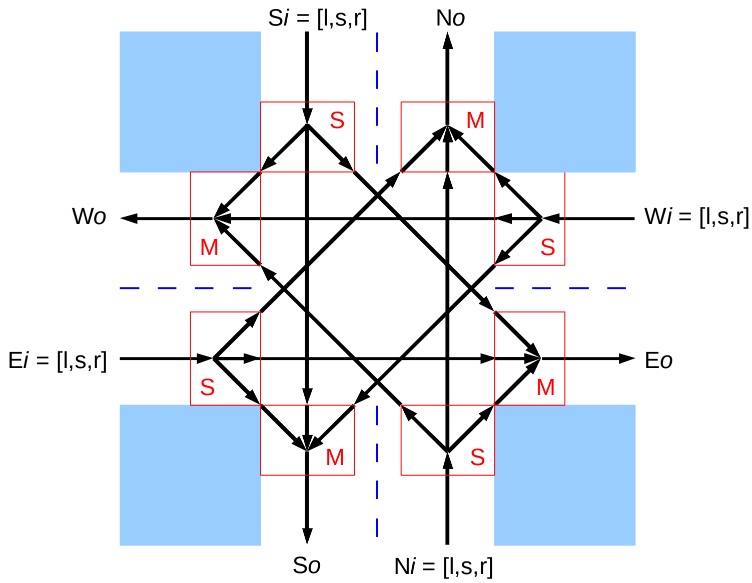
Freeway systems do not allow flows of traffic to cross one another – only traffic splits and merges are used. Left turns are forbidden, and instead are accomplished either with 3 right turns (or an equivalent loop), or with an overpass. Orthogonally crossing flows are separated with either overpasses or underpasses at the point of intersection.

This proposal likewise makes heavy use of splits and merges to separate flows. Large rotary flows are created where major orthogonal trunks intersect, by the pairing of uni-directional roads. This eliminates left turns from bi-directional roads. It also provides substantial distances between the rotary merger and split points, so merging and lane changing can be accomplished. This also limits any use of in-plane crossing of orthogonal flows to feeder streets, with none on the major trunks. Because it is a continuous motion system on the major trunks, in both dimensions, pedestrian (and potentially bicycle traffic) must be grade separated crossing those trunks.

## THE ROTARY CONCEPT:

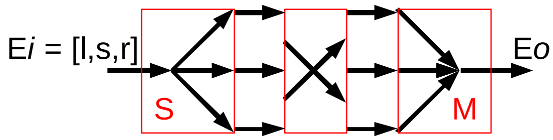
The current road network architecture is what is known in networking theory as a “full crossbar”. Parallel streets running along one dimension are crossed by parallel streets running along an orthogonal dimension. At each point of intersection the crossing streets are connected to each other in all possible directions. This architecture is also sometimes called a “fanout-fanin” network, because that is how it is constructed. An intersection with this architecture is depicted in the drawing below. Traffic enters the intersection from 4 directions: north (Ni), south (Si), east (Ei), and west (Wi). Each of those traffic input streams comprise sub-streams of cars destined for one of three exits from the intersection: turn left (l), straight through (s), turn right (r).

Traffic flowing into the intersection along any dimension in either direction first fans out into multiple parallel streams, each stream having a unique destination (exit from the intersection). This fanout function is graphically depicted with a “splitter” component (S in the drawing). Then all traffic streams destined to each exit from the intersection is collected together, or fanned in, into a single stream. This fanin function is graphically depicted with a “merger” component (M in the drawing). The splitting and merging functions are easily mapped to the driving experience – turning out of a flowing stream of cars; and merging into one.

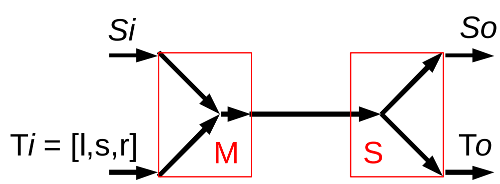



It is well understood from network theory that these architectures create the highest number of crossings. Our road design pushes these split and merge points in close to the intersections, concentrating all these crossings so they can share the same pavement and control. It is clear to see from the drawing that the crossings require either vertical separation (overpasses) or time multiplexing to be safe.

The complexity of this architecture comes from the order in which the routing is performed – split then merge. This causes the two dimensions to interact (cross) at their widest points. The same routing function can be performed with the reverse order – merge then split. This serializes the routing decisions, but always maintains a minimally narrow interaction (crossing) of the dimensions.




Split then merge

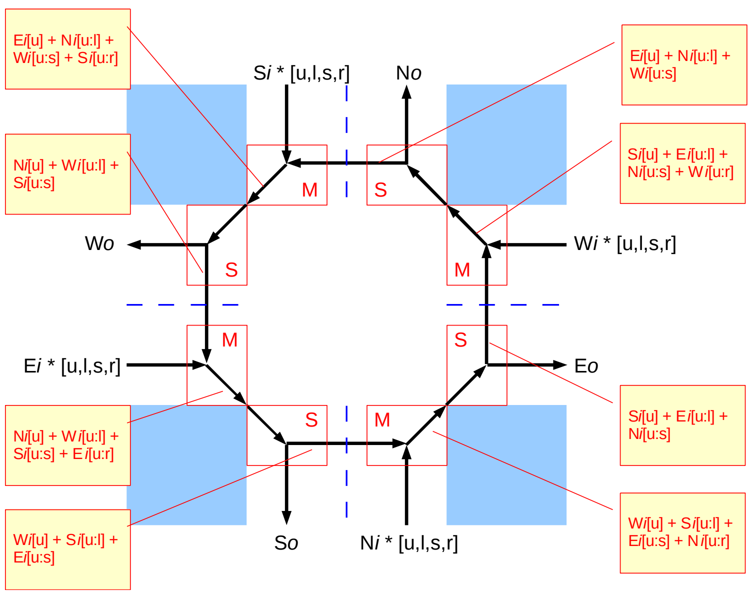



Merge then split

This is the nature of a rotary intersection. It is a series of splitter and merger components. Traffic flows around in a circle. When a road heads off in a direction, traffic can split off from the circle at that point. When a road comes in from a direction, traffic from that road can merge into the circular stream.

## OCTAGONS AND APPROXIMATIONS:

The serial nature of the routing function in merger-splitter networks can be mapped very conveniently to the bi-directional two dimensional nature of our roadway intersections. Since we have 4 directions, and each has both incoming and outgoing traffic, then a series of 4 merger-splitter pairs performs the function neatly. Since we have a series of 8 functions – 4 pairs that merge then split – and those functions cover all 360 degrees of the compass, an octagon is the perfect shape.

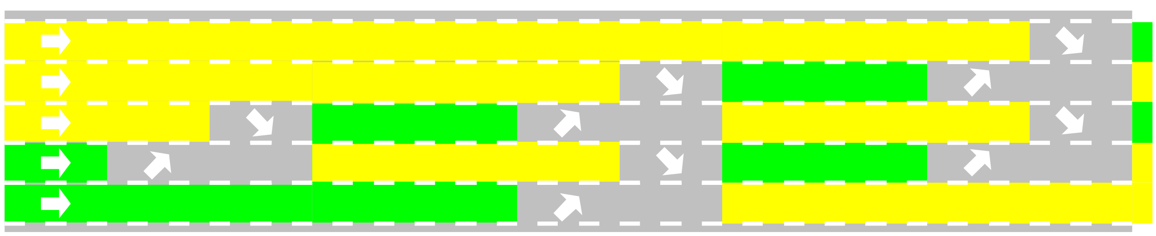



We can quantify the traffic coming into the intersection from all 4 directions, as Ni, Si, Ei, Wi. We can also quantify what percentage of that traffic is destined for each output of the intersection, as u (u-turn), l (turn left), s (straight through), r (turn right), with u + l + s + r = 100%. We include the u-turn traffic for completeness of the model, even though we may expect it to be negligible.

Lets follow the eastbound traffic through the intersection. All of the eastbound traffic (Ei) enters the intersection from the left. It merges with the traffic already in the octagon, so the diagonal on the lower left carries all of the eastbound traffic (we use the colon notation to represent the ordered sum Ei[u:r] = Ei[u] + Ei[l] + Ei[s] + Ei[r]). At the lower splitter the eastbound traffic turning right splits off at the So exit. So the bottom side of the octagon carries only the remaining eastbound traffic (Ei[u:s]). The northbound traffic merges in, and then the straight through eastbound traffic splits off at Eo. So the right side of the octagon carries only the remaining eastbound traffic (Ei[u:l]). The westbound traffic merges in, and then the left turn eastbound traffic splits off at No. So the top side of the octagon carries only the remaining eastbound traffic (Ei[u]). Finally the southbound traffic merges in, and then the u-turn eastbound traffic splits off at Wo. So the left side of the octagon carries no eastbound traffic. While the rotary makes it possible for traffic to loop around the octagon multiple times, we ignore this in our model (judging it to be truly negligible).

Each of the other inlets to the intersection can be modelled similarly. In this way we can sum up the demand loading for each side of the octagon (as shown in the callout boxes in the drawing). If we wish to avoid the intersection becoming a blockage in our roadway system, then these sides each need sufficient lane capacity for the calculated loading, in the same way that the roads leading into the intersection need sufficient lane capacity.

It is important to recognize that the loadings on the octagon sides may be high, high enough to require multiple lanes. This makes another distinction important. The diagonals in the octagon, which are produced by the "mergers", are in fact "swaps" rather than pure mergers. As the eastbound traffic merges into the traffic already in the octagon, it enters from the right. Only the eastbound right turn traffic is positioned for exit at So. All of the other eastbound traffic (Ei[u:s]) will continue on past the So exit, and so must merge to the left. Meanwhile some of the traffic already in the octagon is destined for the So exit (Si[s] + W[l] + N[u]), and so must merge to the right. These multiple lane swaps are clearly the weak point in the rotary architecture. **It is critical that the diagonal sides have sufficient length to allow the lane swapping.** It is also possible for lane markings and/or lane barriers to assist in organizing or controlling the swapping. This is an opportunity for the expression of philosophy. Some urban planners will believe in a maximalist approach that makes heavy use of barriers to control the swapping. Other planners will believe in a minimalist approach that provides very little restriction, and relies upon the perceptive power of the drivers to best manage the challenge.



The drawing above depicts a slightly minimalist middle of the road approach. It implements a 3 lane across 2 lane swap. It provides lane coloring and merge arrow markings as hint guides to the drivers. It conforms to the "yield to the right" paradigm which drivers should already have habitualized, while strongly suggesting a "1 left then one right" sharing, and while hinting that left-most lanes should wait. But there are no barriers to prevent "creativity" for any given situation. You can see that it attempts to coordinate the swap as multiple steps, with each step swapping adjacent lanes. Only 3 of these steps is shown, while 4 are needed in this case (off the right of the drawing). To swap N vs M lanes with this approach requires N + M – 1 steps (3 + 2 – 1 = 4). This should underline the importance of having sufficient distance between the merge point and the split point to allow the rotary to work well.

## Goodness and Benefits

Note in the octagon drawing that there are **no traffic crossings** at the intersection. It is possible (although undesirable) that traffic may have to slow in order to negotiate the turning and merging required to pass through the intersection. But the time multiplexing takes place at a fine grained level, as streams of traffic headed to different destinations interleave into a single lane. There is no need to stop east-west traffic to avoid collisions with north-south traffic. In short, there is **no need**

**for a traffic light** ! As a result, all of the ugly congestion causing attributes of traffic lights are eliminated.

Note that time multiplexing still occurs at the intersection, so there still is a duty cycle associated with it. Consider the simple case where all traffic passes straight through the intersection ( Si[s] = Ei[s] = Ni[s] = Wi[s] = 100% ). The diagonal leading up to the So exit carries both the Ei[s] and Si[s] traffic. So use of the roadway on that link is shared across the two dimensions. The sharing is still done across time, but the time granule for the interleaving is very small – the time for a single car to pass. In fact the sharing duty cycle is dynamic, and naturally follows the relative demand. So the eastbound duty cycle on that link is Ei[s] / ( Ei[s] + Si[s] ). This makes the rotary intersection much more efficient than one controlled by a traffic light. No efficiency is lost for safety intervals between states. No efficiency is lost because instantaneous traffic loading does not fit the programmed traffic light cycle.

Far less accumulation occurs leading into the intersection. Since cars only slow into the intersection, rather than stop, the relative velocity between the cars up front and those behind is much smaller. Therefore less accumulation occurs. Further, when accumulation does occur, it drives the duty cycle up so the accumulation is removed more quickly. The duty cycle can go above 50%. It is not artificially limited by a pre-programmed light pattern, and no time reservation need be made for left turning traffic. There is no risk of accumulation at one intersection falling into a negative resonance cycle with the traffic light pattern at an up-stream intersection. So traffic flow across the area is more stable than with light controlled intersections.

## Pedestrian separation

One good thing that traffic lights do is to stop all traffic to provide a safe environment for pedestrians to cross the road. This comes at a cost to traffic flow and congestion, but we value pedestrian safety over commuting delays. As the whole purpose of rotary based intersections is to keep cars moving, they pose a real barrier to pedestrian crossing.

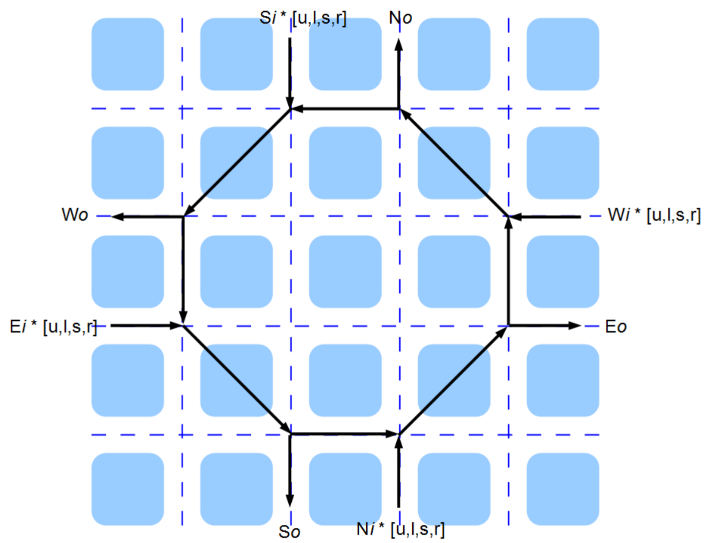
While it is theoretically possible to equip roads at rotary crossings with pedestrian crossing lights, this would be the worst location to do so. If such crossing lights were to be provided at all, the better place to implement them is along the trunk roads connected to rotaries, rather than right at the rotaries themselves. This would provide the traffic leaving such a pedestrian crossing stop the chance to accelerate and spread out before entering the intersection, thus avoiding a demand spike that would clog the “swaps”.

The more desirable solution to pedestrian crossing of trunks is grade separation. Either you elevate the pedestrians, or you elevate the cars. Pedestrian bridging is far smaller and cheaper to build.

## Approximations

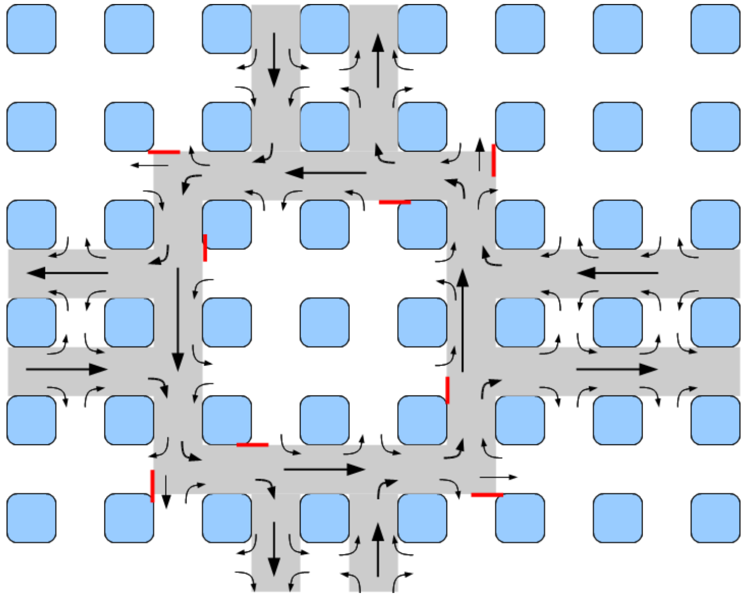
In a new suburban development being built on the edge of a city it is possible to plan and build a feeder roadway system designed around rotary based trunk intersections. The rotaries can be round, or octagons, or some other appropriate shape. The critical thing is that rather than being small rotaries localized to a single pair of crossing roads, they are larger. The intersection between a pair of roads, even a pair of trunk roads in a feeder system is limited in size – to perhaps 120 ft x 120 ft. A rotary circle with 120 ft diameter represents a fairly tight turning radius for speeds much about 30 mph. For trunk roads with a posted speed limit of 45 mph, the tight rotaries would provide a barrier. Comfort alone would define them as a slowing point. Further, the merge length would be only a quarter of the circumference. 90 ft would not support much swapping. So multi-lane rotaries would be difficult to implement. Single lane rotaries would prove to be a capacity bottleneck for trunk roads.

Fitting a larger rotary polygon onto the existing city plan is more likely to achieve a positive result. The drawing below maps an octagonal shaped rotary onto a 3x3 city block grid.




Note that the plan implements the trunk roads as parallel pairs of one-way roads leading to and from the rotary. Note also that the rotary mergers cut diagonally through the 4 corner blocks. For a sub-urban plan having 12 blocks per mile, the rotary diameter is ¼ mile. This should provide ample length for multi-lane lane swaps to work effectively, and comfortably support reasonably high speeds. A single pedestrian bridge across the middle of each block provides adequate access (with half as many as placing them at the corners), which is conveniently placed.

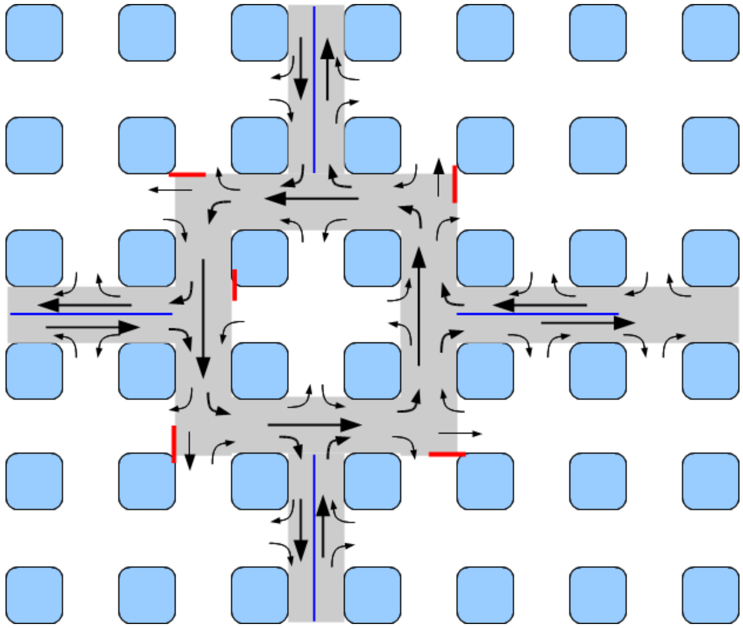
The reality however is that many of the locations suffering the worst congestion are not in new suburbs being built, but in existing cities and suburbs. It is possible to map this same approach onto an existing city block footprint, while preserving most of its benefit.



In this case a square having 3 block long sides is used. The corner mergers therefore are not diagonal. Instead they cover the manhattan diagonal (around the corner). Trunk roads are still implemented as one-way pairs. Trunks support u-turns only as left turn exits which lead 1 block away to left turn entries to the trunk in the reverse direction. There are no street lights anywhere on the trunk roads or the rotary. Off the trunks either street lights or stop signs can be used. There is no orthogonal crossing of the trunk roads. To cross a trunk one would enter in its direction, proceed 1 or more blocks, perform a u-turn, and then exit on the other side of the trunk or its pair.

Since the “diagonals” are already the location for much lane swapping, it seems safest to avoid excessive merging into them. Barriers can be placed to prevent entry to them from the non-trunk streets at dangerous locations. Similarly, since the “diagonals” must turn a corner, and since we wish to promote velocity to increase vehicle separation making merging easier, it is unhelpful to have any reasons for vehicles to slow. Especially near corners, exits are a reason for vehicles to slow. Barriers can be placed to eliminate those exits posing the greatest risk. As long as entry to the rotary center is possible somewhere along each edge, the burden placed on route lengths is acceptable. Dedicated left side exit lanes along the post-corner side of the diagonals is good idea where possible.

There are many safety and efficiency benefits yielded by heavy use of one-way roads. But in many places in many cities this has not been done. Re-working an existing road network to fit a one-way pairs trunk approach may not always be practical. In those cases a 2x2 rotary block can still accomplish much. The “turning radius is still 440 ft, and the diagonals are still 880 ft long.

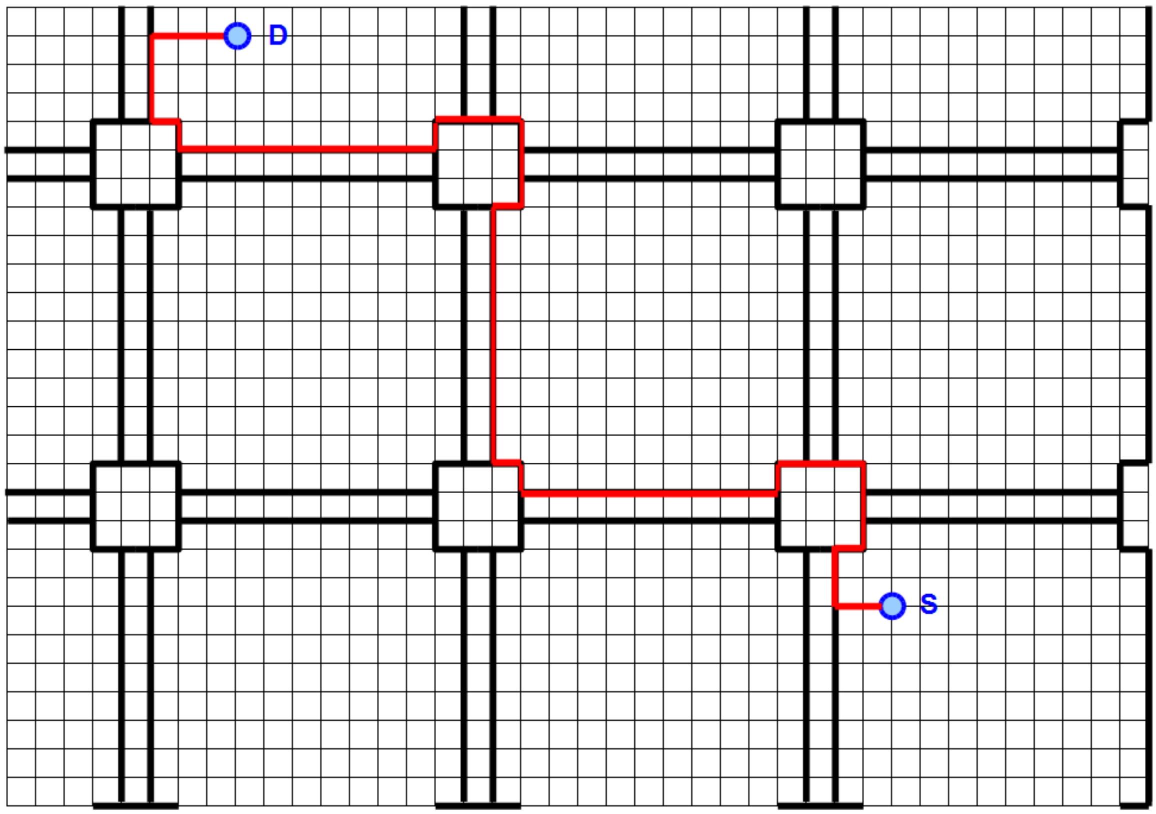


## The Bigger Picture

Converting a single intersection to a rotary will have minimal impact on commute times. The real benefit accrues from implementing the pattern across all trunks over a wide area. This provides north-south and east-west paths for non-stop travel across the city. The lack of stops raises the average commute speed dramatically, even without increasing the maximum velocity of the vehicles. This is more than enough to offset the slight increase in route distances.

The rotary scheme comes with a cost in terms of route lengths. An eastbound vehicle must travel 1 block south and then 1 block north in order to pass straight through the rotary in an eastbound direction. That makes its route length 2 blocks longer than necessary. While this is true, it must be remembered that this does not happen at every city block. It only happens at trunk intersections. So if north-south trunks cross east-west trucks every 12 blocks, the cost in additional route length is 2 blocks every 12 blocks traveled ( 2/12 = 16.7% ). As long as average speed is increased by at least 16.7%, then commute times will be shorter. Additionally, lower stress and better fuel economy will improve the driver's commute experience.

But the typical route length cost may actually be slightly lower or slightly greater than this. Few commute trips are strictly north-south, or strictly east-west. Most are diagonal across the city, with the manhattan topology of our roads forcing them to be a combination of north-south and east-west segments. The very nature of the rotary promotes the use of right turns. Right turns are the easiest since they avoid any lane swapping. This will promote routes that are diagonal across the city, turning first right at one trunk, and then left at the next. The route cost of each rotary depends upon the path taken. Right turns have no cost. Proceeding straight through costs 2 blocks of length. Left turns cost 4 blocks of length. So the overall impact on route length will depend upon how many of each traversals of the rotaries along the way were used. Figuring there will often be one more or one less right turns than left turns, then the average cost will be slightly less or slightly more than the 2 blocks per rotary (16.7%).




## CONCLUSION:

Typically twice per day a defenseless roadway system with no entry regulation is overwhelmed as millions of cars begin their journeys at nearly the same time. The feeder architecture of the roads funnels the overload onto hyper-overloaded trunk roads, where an in-efficient traffic light based flow control system fails to keep up with the demand. Negative feedback in that control loop tips the system into instability. A wave of stall sweeps up nearly all of the commuters. And we know it will continue to get worse.

Transit districts spend millions of dollars trying to keep up. Roads are reworked to make more lanes, sophisticated monitoring equipment is installed, and complex light control patterns are tried. All this makes money for some suppliers, and employs many hard working well intentioned government employees. But is does not, and will not solve the problem.

The demand we place on our roads, versus the capacity they were designed for, has past the point where the inherent in-efficiencies in our control system are tolerable. It is time for an objective scientific re-thinking of the fundamental concepts in our design.

This paper has looked at the basics of what congestion is, and exposed in some particulars the mechanisms that cause it. It has offered one perspective on the alternatives available to design a new system, within the requirement that the new system must be evolved from the current system. It has offered some steps that can practically and affordably be taken along that path. It has hinted at how some proprietary new technology can fit into this new world to achieve its maximum potential.

---

## References


Badger, E. (2011, September 22). *Removing Signals and Signs from Intersections Just Might Make Us Safer.* Retrieved from www.citylab.com: http://www.citylab.com/commute/2011/09/shared-spaces/116/

Baker, E. (2009, January 19). *Removing Roads and Traffic Lights Speeds Urban Travel.* Retrieved from Scientific American: http://www.scientificamerican.com/article/removing-roads-and-traffic-lights/

Census, U. (2012, MARCH 26). *Growth in Urban Population Outpaces Rest of Nation.* Retrieved from www.census.gov: https://www.census.gov/newsroom/releases/archives/2010_census/cb12-50.html

Cohen, D. T. (2015, March). *Population Trends in Incorporated Places: 2000 to 2013.* Retrieved from US Census Bureau: https://www.census.gov/library/publications/2015/demo/p25-1142.html

Gastner, M. (2008, August 7). *The Price Of Anarchy in Transportation Networks.* Retrieved from arxiv.org/pdf/0712.1598

Hamilton, S. (2014, December 15). *The Trouble With Cars.* Retrieved from www.cityram.org: www.citytram.webnode.com/

Hirsch, J. (2014, June 9). *253 Million Cars and Trucks on US Roads.* Retrieved from Los Angeles Times: http://www.latimes.com/business/autos/la-fi-hy-ihs-automotive-average-age-car-20140609-story

Kolko, J. (2012, 10 12). *Even After the Housing Bust, Americans Still Love the Suburbs*. Retrieved from newgeography: http://www.newgeography.com/content/003139-even-after-housing-bust-americans-still-love-suburbs

Martinez, M. (2015, January 5). *2014 car sales best in U.S. since '06.* Retrieved from Detroit News: .detroitnews.com/story/business/autos/2015/01/05/automakers-report-december-sales/21277199/

Ortman, S. L. (2015, March). *Projections of the Size and Composition of* . Retrieved from www.census.gov: https://www.census.gov/content/dam/Census/library/publications/2015/demo/p25-1143.pdf

Texas A&M Transportation Institute. (2015, August 25). *2015 Urban Mobility Scorecard.* Retrieved from http://mobility.tamu.edu/:

http://d2dtl5nnlpfror.cloudfront.net/tti.tamu.edu/documents/mobility-scorecard-2015-wappx.pdf

## About the Author

Stephen Hamilton is a newly retired individual who just concluded a 40 year career as a computer and integrated circuit designer. He has had a lifelong interest in physics, cars, and transit technology in general. Believing that climate change is a species threatening reality, and that governments have failed to face it, he has decided to spend his remaining time and talents trying to use commerce to impact it. Stephen has decided that personal travel in America is the area where his impact can be greatest. He has spent the past 18 months studying this area, and thinking about its challenges and possible/likely solutions. CityTram.org was created to encapsulate these efforts.