Transit Networks

by
William Lieberman, AICP

Based upon a paper submitted to the Institute of Transportation Engineers
February 2008

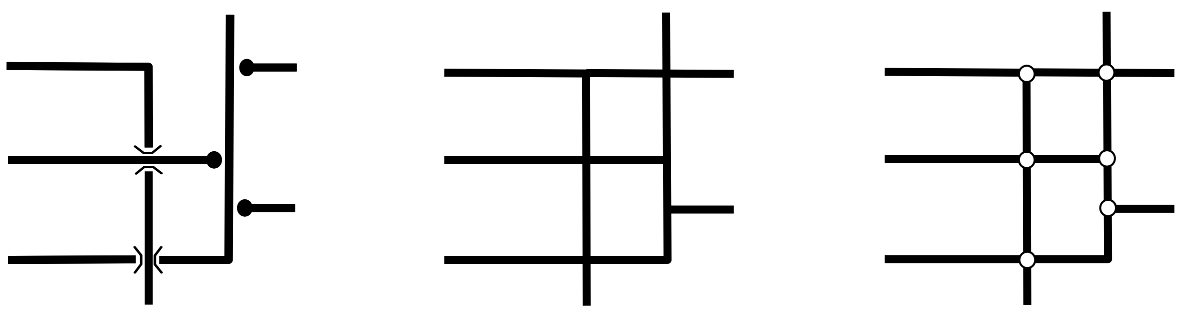
Abstract: The networks used by transit systems have evolved over time. In their purest form, they can be classified as amorphous, radial, grid, and hub-and-spoke. Each has its unique advantages and drawbacks, which is why larger transit systems generally employ more than one form of network. Various metrics can be tapped to analyze and improve transit networks. The configuration of a transit network is greatly influenced by the configuration of the community's roadway network. The two must be considered in concert when making changes to either one. Moreover, certain roadway network types are more supportive than others for facilitating the movement of transit vehicles and allowing pedestrians convenient access to transit. While flexible modes of transit can partially compensate for restrictive roadway patterns, an interconnected street network is the most effective way of facilitating the movement of transit vehicles and pedestrians. Design practitioners should be aware of network issues in order to maximize the effectiveness of both new development and the transit lines that serve it. Moreover, transit lines should always be modified in the context of their networks, as opposed to using an isolated, case-by-case approach as an expedient to solve an immediate problem.

Background

When the subject of transit service arises, there is a tendency for the public and practitioners alike to focus on individual transit *lines*. This is natural, because lines are the basic building blocks of transit service and are what we focus on when using public transportation. Two or more lines constitute a *system*. The spatial configuration of lines in a system is termed a *network*. The network forms the underlying context of transit service. It can determine which kinds of trips are time consuming and which are rapid, or which are economical to operate and which are expensive. For these reasons, transit networks should be understood by those dealing with the built environment, even if they are not directly involved in transit service planning.

The purpose of a transit network is to collect and distribute passengers around a larger area than would be possible with a single line. Good network design has two principal objectives, the primary one being the creation of patterns of transit lines that serve the travel desires of the public. The closer the correlation between this network and local travel patterns, the more convenient transit will be and the greater the likelihood that people will use it. The second objective is the accommodation of passenger flows in the most cost-effective manner possible. While capital costs are often the focus of transit development decisions, it is the day-to-day operating costs that, over time, can have the more significant financial impact. Operating costs are strongly influenced by the shape and function of the transit network. These two objectives--convenience and cost-effectiveness--are sometimes mutually reinforcing and sometimes at odds with each other, but they are part of the balancing act that constitutes transit planning and network design.

Transit lines that do not cross paths and therefore are not conducive to passenger interchange between them do not constitute a network.[1] They are simply unrelated lines. In contrast, when lines cross or join, passengers are able to transfer from one to another, increasing the area in which they can travel by transit. Their travel is facilitated even further if the network is fully integrated: that is, the cost of making transfers is free or nominal and the schedules of the lines involved are synchronized, reducing the time that passengers must wait at transfer points.



"Non-Network" **Network** **Network (fully integrated)**

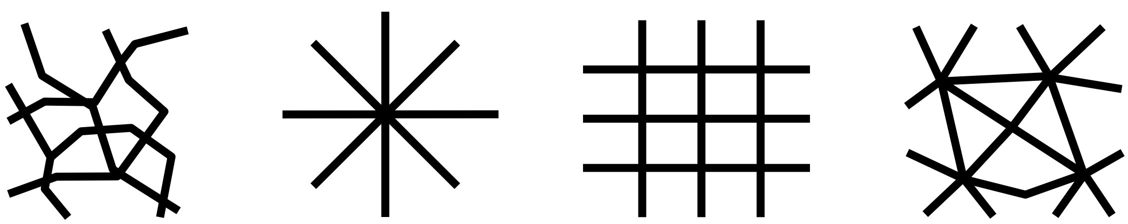
Historical Development. Transit networks have evolved since public transportation first arose almost two centuries ago. In the earliest years, transit was provided by horse-drawn coaches operated by individuals or small companies, each line administered independently of the other. The network, as it existed, was simply the aggregation of lines that had been developed incrementally over time. Little forethought was given to its overall shape or function. Particularly when fares and schedules were not coordinated, the power of a network to satisfy mobility problems was not tapped.

Later in the 19th century, street railways appeared, at first powered by horses and later by electricity. Rapid transit lines emerged at about the same time, operating rail vehicles above or below the street. The infrastructure of tracks, structures, and wires needed by rail transit forced their operating companies to consider the networks that they were creating. With a well designed, interconnected network, a company could minimize its capital and operating investments. It could also attract more business if passengers were able to transfer easily from one line to another. Transit operators therefore began looking at the big picture as well as the details.

That big picture is often lost in today's focus on cutting corners to stay within shrinking operating budgets, or fulfilling short-term political commitments. Nevertheless, our most effective endeavors are generally those that fit specific actions into a broader context. Transportation professionals considering additions, deletions, or modifications of transit lines would do well to consider the network in which these changes must function. If that network is ill suited to the demands to be satisfied, then the modifications should be reconsidered. In more extreme situations, changes to the network itself may be warranted.

Types of Transit Networks

Essentially, there are four "pure" types of transit networks: amorphous, radial, grid, and hub-and-spoke. There are also many hybrids. The pure forms are illustrated by the diagrams below:

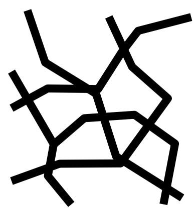


Amorphous **Radial** **Grid** **Hub-and-Spoke**

It must be stated at the outset that a network, no matter what its form, is just the spatial dimension of transit service. Other dimensions are just as important in the functioning of a system. To illustrate, a transit system map might give the impression of very rich service when transit lines are closely spaced and cover extensive areas of a city. However, the frequency of service on these lines may be low, the equipment antiquated, and the rates of fare high, resulting in a system which is really not very effective or valued by local residents. For this reason, the reader should be cautioned that this discussion focuses on just one dimension of transit service. Real world planning must be tempered by consideration of many others.

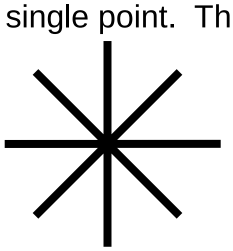
Amorphous Networks. The first type of network, *amorphous*, is one without a regular geometric pattern. It is generally the result of one of two situations. The first is that separate routing decisions have been made incrementally, rather than in concert. The end result is merely a collection of lines which overlap or cross each other haphazardly. Amorphous networks were typical in the early stages of many cities' transit histories but were soon supplanted by more systematic ways of reconciling individual lines. However, even today amorphous networks can be discerned in some metropolitan areas when all modes of public transportation are plotted together on a map. The individual networks of publicly-operated buses, privately-operated buses, trains, ferries and shuttles each may have rational layouts in themselves. However, with little or no schedule coordination among them, and with separate fare structures, the combined network as a whole essentially makes no sense. As a result, travel across amorphous networks can be difficult and confusing if it involves more than one line or one transit operator.

The other situation leading to amorphous transit networks is more deliberate. It occurs when the urban street system does not have much regularity in its



pattern. In this case, transit lines that follow the main roadways may form an interconnected grid, but it is one which can appear chaotic. If this grid provides good coverage to the areas the public wishes to access, it is sometimes referred to as a *ubiquitous network*. A good example of this is the network of the Paris Metro. This subway system was developed largely by one company, but its network defies description. While some of the lines follow a general north-south or east-west orientation, others change direction and cross each other for no apparent reason. Such abrupt changes in the direction of individual lines can make it difficult for infrequent users to orient themselves and get a grasp of the whole system. Fortunately, there are transfer opportunities at most points where two lines cross. Moreover, the Paris Metro is equipped with extensive wayfinding aids, including illuminated route maps at stations, supplemented by directional signs that facilitate navigating through an often complex maze of passageways. As a result, passengers can be oblivious to the network itself and still find their destinations. Because they are difficult to visualize, amorphous networks are best reserved for situations like this where there are no practical alternatives. However, as shown in Paris, even a potentially hopeless network configuration can be made highly functional with conscious effort.

<u>Radial Networks</u>. In a *radial* network, all or most transit lines converge upon a single point. There are frequently two sorts of lines such networks: *radials*,

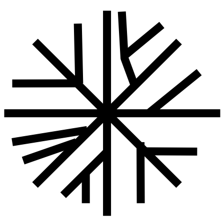


which terminate at the central point; and *diametricals*, which pass through it and provide "one-seat" service from one outlying area to another. In either case, the appearance of a radial network when examined in plan view is "star-shaped." Travelers on any line may access the central point directly, without the need for a transfer. Moreover, they can reach any other point on the network with just a single transfer. However, this transfer, being at the center of the system, may entail considerable out-of-direction travel.

It is worth noting that radial networks need not focus on a city's downtown. They may be oriented to a university campus, shopping center, or commuter railroad station, for example. Such systems may be free-standing, or they may constitute a small portion of a much larger network. When examined in isolation, nevertheless, they function as radials. Moreover, a radial network need not focus on a single point. In large cities, each line destined for the CBD often serves different streets within it. The lines essentially form a grid of sorts within the downtown area. Such systems are most effective when they are arrayed so that travel between any two lines can be made with a single transfer, allowing the network as a whole to still function as a simple radial.

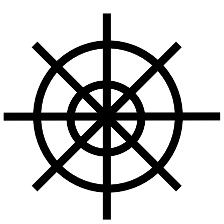
Radial networks work well where the majority of trips have a central orientation, but they are not very adaptable to shifts of activity to other areas. Radials were a natural form for early transit systems, since most of the employment, shopping, and other activities prior to World War II were clustered downtown in a city's central business district (CBD). Many transit systems still exhibit a radial pattern, even though these activities long ago dispersed to other points in the urban area.

There are a couple of notable variations of the radial network. One is the

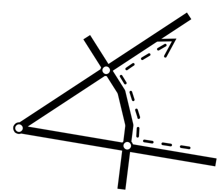


branched radial, in which some of the lines radiating out from the central point act as a "trunk", with "branches" forking off further from the center. This increases the density of service in outlying areas, allowing more people to live within walking distance of a transit line. It also results in higher service frequencies on the trunks, where both residential densities and travel volumes are generally greater.

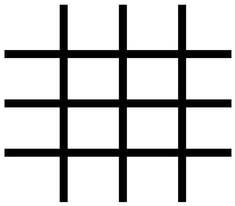
A second variation is the *radial web*, in which the lines radiating out from the



central point are intersected by one or more circular routes, called *circumferential* or *orbital* lines. Circumferentials reduce the necessity of transferring at the central node when traveling from one outlying point to another. In some cases, two transfers may be required to reach a destination when using a circumferential line. This may

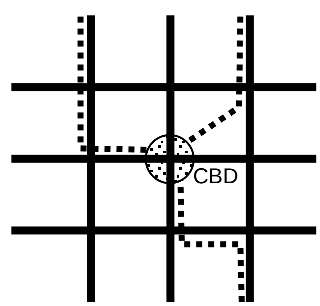


actually take longer than going to the central point to transfer, depending on individual circumstances of line location and frequency of service. Riders usually discover which travel pattern suits them best. In a radial web, the more circumferential lines there are, the more convenient the system. However, most such networks have only one completely circular circumferential line. Due to difficult terrain or the need to avoid large bodies of water, more often semi-circular lines are employed. Another technique is to use a tangential line to cut across the radials in one part of the system. These methods increase the connectivity of the network in some districts though not throughout the urban area as a whole. By and large, radial web networks with multiple circumferential lines, though often seen in freeway systems, are rarely found in transit.



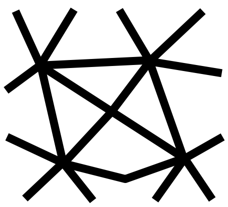
Grid Networks. *Grid* networks could be said to have any configuration in which there are intersecting lines with stations or stops that allow transfers between them, including the radial web described above. However, the term is most often used to denote a *rectilinear grid*, in which transit lines follow the north-south/east-west pattern of perpendicular streets found in most "newer" cities. Actually, grid street systems can trace their roots back to ancient settlements established almost three millennia ago. This street pattern is familiar because it is so prevalent in North American cities. In spite of this fact, many American transit systems in cities with grid street networks actually operate radial transit networks. Their individual transit lines may follow north-south and east-west routings in outlying areas, but as they approach the downtown, these lines change direction and converge upon the CBD. Pure grid transit networks do not do this; their transit lines remain true to their grid orientation. This necessitates transfers by downtown-bound passengers to those grid lines passing through the central area. However, transferring is part and parcel of a grid network, and grids are excellent for serving dispersed origins and destinations. They also are the easiest network to understand, as each line tends to follow one street for all or much of its route.

Theoretically, grid transit networks, like radial networks, can be designed to accommodate travel from any one point in the system to any other with just a single transfer. However, to realize this benefit, lines must be closely spaced (one-half mile apart allows everyone to be within a quarter-mile walking distance of a line) and frequent (10- or 12-minute headways make the average wait at transfer points just 5 or 6 minutes). In actuality, these conditions can be difficult to achieve. Terrain and the street configuration may force transit lines to be too far apart in some areas. Moreover, the expense of maintaining frequent service cannot be afforded by many transit systems. This latter circumstance results in some grid transit systems that operate at low frequencies, resulting in long waits at transfer points. Travel is tedious on these systems and not attractive to riders with a choice of modes. Even systems with frequent headways on the grid during the day may have to reduce them when patronage drops in the evening and on weekends, leading to longer transfer times. For these reasons, the most effective grid transit networks are found in larger urban areas that have closely spaced arterial streets and volumes of travelers sufficient to warrant high service frequencies throughout the day.



Grid with peak-hour radials (dashed lines)

As mentioned, a downtown trip in a grid network often requires a transfer. It may entail some out-of-direction movements, further adding to travel time. To avoid this inconvenience, many urban areas have superimposed radial lines over their basic grid transit networks. In some, limited-stop buses operate directly downtown during peak hours, sharing the street with grid buses in outlying areas but diverting to reach the downtown more directly. In other areas, downtown trips are accommodated on a radial rapid transit or commuter rail network. With either technique, travel to the CBD is expedited, and overloading is avoided at key transfer points in the grid.



Hub-and-Spoke Networks. *Hub-and-spoke* networks consist of transit lines that converge upon several sub-regional nodes of activity, such as shopping centers, town centers, and rail transit stations. Each activity node is like the hub of a wheel, and the lines converging on it are like the spokes. In a sense, the network as a whole comprises many small radial networks, each centered upon an activity area or transfer point. While grid transit networks offer a way of accommodating dispersed travel patterns in dense urban areas, hub-and-spoke networks are especially effective in just the opposite kind of environment. They work well where high frequencies cannot be justified and where street patterns may not exhibit a grid pattern. These conditions typically occur in the suburbs of many North American cities.

When transit lines converge upon a hub, they allow three kinds of travel to take place:

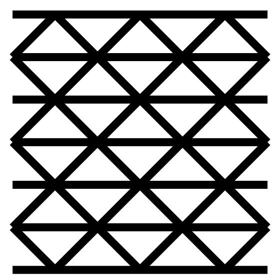
- Riders are brought directly to the activities located in the vicinity of the hub;
- Riders can transfer at the hub from a suburban line to one serving the central business district;
- Riders can transfer at the hub from one suburban line to another, and by this means reach other outlying destinations.

Hub-and-spoke networks have been used for many years in the airline industry, where they were found to be the most practical way of carrying travelers between many dispersed origin and destination points. In the 1970s, this network typology began to be applied to transit lines. It is most effective when combined with *timed transfers*. Rather than converging on the hub randomly, the transit vehicles (generally buses) all arrive there at about the same time, say, on the hour and half hour. There is a short window of time (typically 3 to 5 minutes) in which the buses wait for passengers to transfer, and then the buses leave the hub simultaneously for their separate destinations. Hub-and-spoke networks are often utilized on rail transit systems, with the buses providing the local “feeder” function and the rail line providing the “trunk line” connection to the CBD.

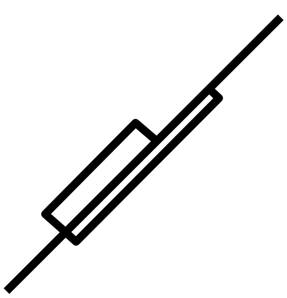
With the type of system described above, transfer time is minimized. Moreover, the number of buses converging on the hub can justify the development of a station to accommodate the buses and passengers there. Such stations are often located off-street and are typically furnished with sheltered waiting areas, transit information and, in some cases, parking facilities, rest rooms and food concessions. This is considerably more attractive to riders than dashing across wide arterial streets to make transfers. Moreover, the assurance of a timed meet at the transfer station decreases their anxiety in using the system. For these reasons, the hub-and-spoke network, when combined with timed transfers, has been found to be the most effective way to serve dispersed travel patterns in low-density environments.

(It should be noted that timed transfer scheduling is sometimes used at the center of radial transit networks. Its application is restricted mainly to small systems, where a manageable number of buses can be accommodated at the focal point and service frequencies are relatively low. Larger transit systems have occasionally used timed transfers downtown to facilitate travel at night or on weekends when low service frequencies are offered. Timed transfers are not applied to grid networks because there are too many crossing points of individual lines. Moreover, they are not needed when service frequencies are high. On low-frequency grids, timed meets are sometimes scheduled at one or perhaps two crossing points of certain lines, but that is usually the limit of what can be operated reliably.)

<u>Theoretical Networks</u>. There are other network typologies besides those mentioned thus far, but these have not generally been used for either street configurations or transit systems and must therefore be considered purely theoretical. They are mentioned here for the sake of completeness. For example, the *triangular* (or *delta*) *grid* features straight lines in three directions of travel. While some urban areas have strong diagonal boulevards superimposed on their rectilinear grid street networks (Washington, DC is prototypical) or have one or more diagonal streets which break their rectilinear grids into separate orientations (as does Market Street in San Francisco), no city has a total street or transit network with a delta grid orientation. While it has a very high degree of connectivity, a delta grid would create complex intersections and be extremely disorienting to negotiate.



A *linear* network could be useful where all trip origins or destinations are located along parallel lines in one main corridor of travel. Linear cities have been suggested in the past but never built, although linear corridors are commonly found along highways at the edges of, or between, urban areas. Most of these corridors consist of a single main street served by a single transit line, and a single line does not constitute a network—only two or more connected lines do. (Using this same rationale, the linear automated people movers found in many airports do not operate on linear "networks.")



Sometimes an isolated portion of another network type can resemble a linear network. For example, the famous bus rapid transit system of Curitiba, Brazil has arms on its radial network that, when viewed in isolation, replicate the operation of linear networks. Each transit "spine" street is flanked on both sides by a busier one-way street with express bus lines, the three streets constituting a *"trinary."* Viewed in their entirety, however, these trinaries are interconnected components of a radial web rather than true, freestanding linear networks.

Network Evolution and the Hybrid Network

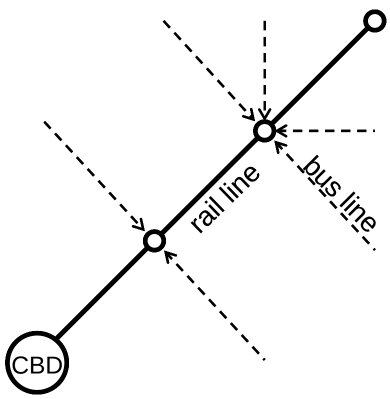
As urban travel patterns have shifted, transit network preferences have shifted, as well. The earliest amorphous networks were soon reorganized into radial patterns as public transportation became more sophisticated. Radials were well suited to the centralized location of activities in nineteenth and early twentieth century cities. A few of the larger cities adopted grid networks. In the past thirty years, many radially-oriented systems have given way to hub-and-spoke networks, which are more suited to accommodating the dispersed travel patterns experienced today. Not all transit systems have made this adjustment. Some, especially those serving small cities, have retained radial networks because of the difficulty in attracting non-radial trips to transit; it's just too difficult for them to be competitive with autos for such travel. Other systems, however, cling to radial networks even though their primary travel markets long ago abandoned them. Perhaps resistance to change, both by the remaining transit riders and by transit managers themselves, is the driving factor. Whatever the reasons, these systems continue their traditional radial operations at a great price. Several studies over the past 30 years have documented the loss in both ridership and cost-effectiveness of such systems compared to those that have adopted multi-destination networks.[2]

One way of adapting radial networks to handle more diverse ridership patterns is through the creation of hybrid networks that have the attributes of several network typologies. In fact, most larger transit systems today do not operate one of the "pure" network typologies described above but, rather, utilize several types simultaneously. It can be said that the end result, when examined as a whole, is a hybrid network. For example, Portland, OR has a grid bus system on its east side, where the land is mostly flat and a grid street pattern prevails. In other parts of its metropolitan area, terrain has forced the street system into more irregular patterns, so both radial and hub-and-spoke bus networks are employed. Superimposed on these bus networks is a radial light rail network.

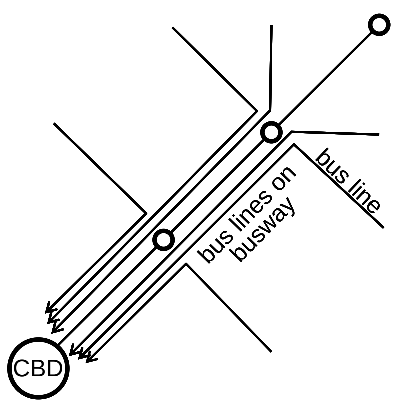
Rather than disparaging such systems because they do not stay true to form, we should carefully examine them because they use what works best in their individual circumstances. The patterns of travel in today's urban areas are complex, and they warrant transit networks that are more complex than those developed in earlier years. Utilizing several network configurations for a transit system is one response to accommodating these travel patterns in a way that is attractive to the public.

Coupled and Uncoupled Transit Lines

One aspect of transit network design that is often overlooked has a very decisive impact on the nature of transit operations and passenger convenience. This aspect is whether the lines in a network are "coupled" to each other or "uncoupled" (the terms *integrated* and *independent* are sometimes used). *Coupled lines* are those which are dependent upon each other because what happens to one influences the operation of the others. *Uncoupled lines* operate largely independently of each other. While not directly related to network typology, this matter does impact the service characteristics of a network.



Examples of coupled versus uncoupled lines can often be seen on radial transit networks. When a major radial line is operated by rail transit—whether light rail, rail rapid transit, or commuter rail—it is, by necessity, uncoupled from bus lines in its vicinity.[3] Local bus lines are commonly configured as "feeders" to the rail stations, requiring passengers to transfer between the two modes. However, what happens on the bus lines does not directly affect the operation of the rail line, and vice versa; hence, the two types of lines are uncoupled.



Some cities employ a busway instead of rail line in a major radial corridor (a *busway* being a separated road for the exclusive use of buses). Such busways often operate with coupled lines. There is typically a core bus service that operates along the length of the busway in the same manner that a rail line would. However, this core bus line is supplemented by local bus lines from outlying areas. The local buses converge upon the stations, as with a rail line, but then continue onto the busway itself, sharing the busway with core buses to provide through service to the central business district. In some instances, this sharing occurs only during peak hours, while in others, local lines run through to the CBD all day long. Local bus lines in the corridor are thus "coupled" to the trunk line on the busway.

Examples of coupled lines can be seen in other situations, as well, for both bus and rail. The branched radial network described earlier is based on coupled lines, as one or more radial “branches” converge to form a “trunk” line into the central core. Even grid networks sometimes have lines that branch at their outer ends. Since the various branches share the same route on the trunk, their scheduling should be coordinated to try to achieve an even interval of service.

Coupled networks of the sort described above have the advantage of offering a “one seat ride” to passengers traveling to the central focus of the trunk line, generally the CBD. Convenience is high for these riders because no transfer is required between the branch and trunk services. On the other hand, it is not uncommon for less than 5% of travel within metropolitan areas in the US to have downtown origins or destinations. With its focus on downtown travel, coupled radial networks aren’t as well suited to serving dispersed destinations. Moreover, when a problem arises on one line of a coupled system, it can cascade into problems on the whole network. For example, traffic congestion on an arterial street that blocks the flow of buses on a branch could result in fewer buses and insufficient capacity on the trunk line.

In contrast, coupled systems do impose penalties of time and inconvenience on passengers wishing to transfer between any two lines in the network. The compensation is that they accommodate a wider range of travel patterns. As discussed earlier, the stations on uncoupled rail transit lines often function as the hubs of hub-and-spoke networks utilizing timed transfers. Passengers may arrive there by one feeder line and leave by another, destined to a non-downtown location. These passengers may never use the trunk line, yet the stations serve to optimize the local feeder network for more diverse travel.

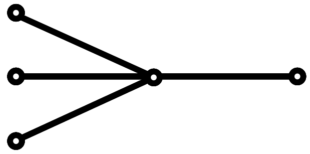
It is important to balance service levels on coupled lines. Otherwise, the trunk may become overloaded with transit vehicles or the branches too lightly served. Since rail vehicles cannot easily pass each other without the provision of switches and sidings, every effort must be made to ensure that traffic from each rail branch is managed properly when it arrives at the junction point with the section of track used in common. However, even busways have to grapple with the consequences of heavy flows on their trunk segments. Moreover, as mentioned earlier, any disruption of service on the trunk portion of a coupled network affects the branches, as well. The entire corridor can be impacted, rather than just one portion of it. With an uncoupled network, a service disruption on one section inconveniences fewer riders, as the other sections operate independently. Some transferring riders may be affected, but the network as a whole, still functions.

Whether coupled or uncoupled transit lines are favored in any corridor depends upon the travel objectives to be served (e.g., focused versus dispersed trip patterns), the expected volume of riders to be accommodated, the capacity of the lines in question, the degree of reliability to be achieved, and the mechanics of how to handle passenger flows. Unfortunately, decision makers are often unaware of these considerations or even the fact that coupled and uncoupled lines exist. As a result, debates about alternative transit corridor improvements often center on capital costs and the convenience of traveling downtown. As indicated in this section, those are simply two of many factors that should be considered.

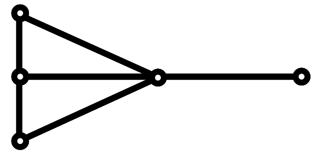
Network Metrics

All networks, whether they are used for transit systems, streets, pipes, or electrical circuits, have certain attributes in common. For example, *connectivity* is the degree to which a network has many interconnections between the points that it joins together. A well connected network is said to be *redundant* when alternative pathways are provided. This term need not have a negative connotation for transit, since redundant networks give passengers a choice of route and are particularly useful for avoiding breakdowns or overloads on one segment of a system. *Circuity*, on the other hand, is a quality to be avoided in transit. It is the property of networks that are not well connected and, hence, require roundabout paths of travel.

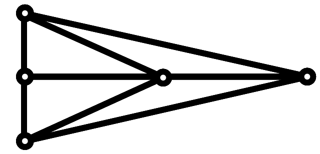
Mathematicians working with graph theory have devised ways to measure networks and compare them, so that their various attributes can be made explicit. These metrics are rarely used in the transit industry and seem to be relegated to the realm of academia. However, a simple understanding of the basics is useful to determine what sorts of changes (such as more stations, more transfer points, longer lines, etc.) might be considered to improve an existing network or to incorporate in a new network. While a full discussion of these metrics is beyond the scope of this paper, a simple example may suffice here to show their usefulness. The complexity of a network can be measured by the indicator β, which, for a transit system, is the ratio of the number of line segments between stations (called *arcs* or *edges* in graph theory) divided by the number of stations (called *nodes* or *vertices*). The illustrations below show how β differs among four networks with the same number of stations but varying numbers of lines. It could be said that the network to the far right offers twice the convenience as the network to the far left, as it includes more paths of travel between the stations and has a β value which is twice as high.



β = 0.8



β = 1.2



β = 1.6

Depending upon the objectives being pursued and the amount of funding available, one of the above networks would likely be more suitable than the others for a particular application. Using similar analytic techniques, alternative network configurations can be compared in terms of their directness of service, utilization of line capacity, area coverage, density of travel, and many other attributes. Readers desiring a fuller discussion of this topic are referred to Vuchic[4] or to any of several mathematical texts on this subject.

Street Networks and Transit Networks

Transit networks are reflective of--and constrained by--the street networks on which they operate. As might be expected, this is true to a greater extent for transit modes operating on the surface of streets and highways (such as buses, streetcars, and some light rail lines) than those utilizing off-street rights-of-way (such as busways, light rail in exclusive alignments, subways, elevated railways, commuter rail, and automated guideway technologies). Of course, the latter group may be *influenced* by the street grid. Subway lines, for example, customarily follow the alignment of streets above because this avoids impacts to the foundations of buildings. Even considering this fact, however, it is obvious that transit lines operating in their own rights-of-way are simply not as dependent upon the street network as are surface transit lines, so the discussion that follows applies mainly to the latter.

Urban Street Networks. To a large extent, public transportation has been able to adapt to the vagaries of urban street networks. Where necessary, transit alignments make turns to follow the streets needed to achieve the desired paths of travel for buses and surface rail. The example was already given of radial bus systems that operate on grid street networks by simply turning onto streets headed directly toward the central business district. This flexibility is generally more achievable in the central portion of cities, where streets suitable for transit are numerous and closely spaced. In a few cases, short sections of exclusive transit streets have been created to fill in missing links in the street network, but in general the streets themselves provide the pathways needed.

Suburban Street Networks. Out in the suburbs, it’s often a very different story. There, street patterns are often dendritic: they resemble the organic form of a tree, with a trunk, branches, and twigs. Typically, a large number of local streets feed into a smaller number of collectors, which feed even fewer arterials. Moreover, the local and collector streets—or “twigs and branches” in this analogy—often have circuitous alignments, and cul-de-sacs abound. These patterns make it hard for surface transit (mainly buses) to penetrate the neighborhoods. It’s also difficult for residents of those areas to find direct walking paths to the nearest transit line.

In spite of these drawbacks, both developers and homeowners have tended to favor dendritic street networks. This pattern is perceived as reducing through traffic and appearing more bucolic than a rectilinear street grid. Such attributes come at a price, however. That price includes the dependence of residents on autos for even the most minor trips. This dependence, in turn, results in higher transportation costs because of the need for multiple-car ownership. It also generates significant traffic congestion, energy consumption and air pollution in a community.

There are other challenges, as well, to transit in the suburbs. Because arterial streets are spaced farther apart than in denser inner city areas, it's often impossible to maintain the quarter-mile distance that has generally been accepted as the limit people are willing to walk to a transit stop. In addition, fewer arterials in the street grid mean that each arterial must be wide in order to accommodate traffic demand. Anyone who has waited for a bus along one of these wide, busy arterials, or had to cross such a street in order to access a bus stop, can attest to the fact that it is not a pleasant experience. These design issues are among the reasons that transit tends to be less attractive to prospective riders in the suburbs than in the inner city.

The drawbacks of dendritic street networks to transit use will become more obvious as the population ages. Given our demographic profile, the number of suburbanites dependent upon public transportation is projected to increase dramatically over the next twenty years as more baby boomers have difficulty continuing to drive. These boomers will learn first-hand how difficult it is to have effective transit service when the street network discourages it. Given all of the recent interest in sustainability, street system design needs to shift toward more flexible and interconnected networks. They work better for transit and pedestrians, yet still accommodate the movement of other forms of street transportation.

Adaptive Solutions. Fortunately, there are several ways of overcoming the challenges of suburban street patterns. One is by adapting transit to the street network by offering flexible modes instead of those with fixed routes. Demand-responsive "dial-a-ride" services are able to penetrate even the most challenging street networks with vans or minibuses that offer door-to-door service. Unfortunately, this kind of transit is expensive to operate, averaging eight times higher per passenger served than conventional fixed-route buses.[5] Moreover, many passengers don't like to make advance reservations for everyday travel and then wait until a particular time for the vehicle to arrive. A less expensive technique is flex-route buses, which have conventional routings on some portions of their route but may deviate off-line in certain neighborhoods upon request.

Another way of dealing with circuitous street systems is by adapting the street network to transit. This can be done in areas with circuitous streets by developing off-street walkways that create shortcuts for pedestrians. These allow for more direct walking paths between transit stops and homes (and, in some cases, offices and shops). Needless to say, these walkways are relatively easy to incorporate while a development is under design but very difficult to retrofit into an established neighborhood.

Interconnected Streets. In the solutions mentioned above—flexible transit services and shortcuts to bus stops—public transportation is trying to serve a street network that was not designed with transit riders (or pedestrians) in mind. A better alternative for new construction is not to use dendritic street patterns at all. The traditional pre-World War II interconnected street network (whether a rectilinear or irregular grid) is adaptable to more users of streets, including transit riders, bicyclists, and pedestrians, as well as motorists. For this reason, it has been adopted by New Urbanist and transit-oriented developments over the past two decades.

Within an interconnected street network, the ideal spacing for major streets suitable for bus operation (i.e., *collectors* and *arterials* in conventional street design, and *connectors, streets, avenues* and *boulevards* in New Urban design) is one half mile in a rectilinear grid pattern. The half-mile spacing, as mentioned earlier, allows everyone to be within a quarter mile of transit, and this is the accepted walking distance for surface transit lines. (In fact, the number of people who are willing to walk to a bus stop drops off precipitously after about two blocks. However, research has confirmed that catchment areas of one quarter mile--approximately 400 meters, or about a 5-minute walk--around bus stops encompass about 75% of those willing to walk to transit on relatively level terrain.)

This need not mean that newly developing areas must be created around "monotonous grids." Variations in the spacing of the major grid streets can be introduced, especially to accommodate features of terrain. Moreover, the smaller local streets encapsulated within this grid needn't have a rectilinear orientation. As long as they are interconnected to shorten travel by walking and bicycling, creative local street patterns can be devised to give each neighborhood a unique flavor. The important point is that the major streets upon which buses or other surface transit can operate should be relatively straight and spaced at approximately half-mile intervals. These features will allow for both efficient transit operation and convenient walking access for the majority of residents in contiguous areas.

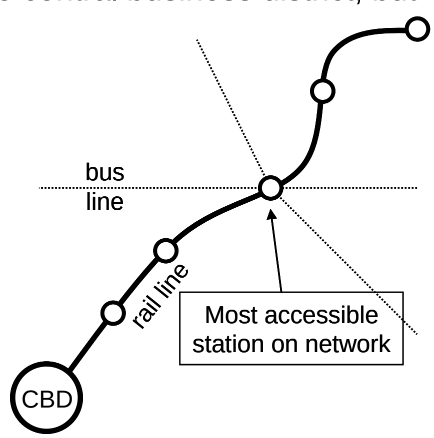
Transit Networks and the Design Practitioner

Generally, land use planners, traffic engineers, urban designers, and developers have little knowledge of the transit network typologies described thus far, and little control over them. If these professionals deal with transit at all, their focus is generally on one or two transit lines, not the entire network. This poses no problem if the principle previously discussed remains uppermost: to develop street networks that are interconnected. Such street networks will be able to accommodate all of the commonly used transit networks and still be useful for other forms of transportation, as well. This simple maxim is all that really need be understood.

Engineers. Transportation engineers have an additional dimension to deal with besides street networks, and that relates to the design of individual thoroughfares. It has already been mentioned that both circuitous local streets and very wide arterials are difficult for transit to serve. There are many other design considerations that could be cited in order to make individual roadways more transit-friendly.[6] From the perspective of network design, however, a high degree of roadway connectivity is still the main principle.

Developers. Land developers will profit by understanding not just the optimal street patterns within their proposed developments, but how the locations of those developments should be assessed vis-à-vis the larger transit network. For example, a development site located near the junction or crossing point of two or more transit lines is significantly more accessible than one located adjacent to a single line. In fact, a location served by multiple bus lines may be superior to one served by a single rail line.

This situation is often overlooked, particularly in the zeal to develop a parcel of land near rail transit. A potential development site may offer excellent access to the locations served by the rail line, especially the central business district, but what of other, more dispersed work and shopping locations? The preponderance of households with two-wage-earner families makes it likely that even if one family member works downtown, the other will not. Moreover, people change employers more frequently than they used to, and today's CBD job may soon be traded for one in an outlying location.



Housing that is located near the junctions of transit lines can more easily accommodate these situations. It allows residents to take advantage of travel opportunities in many directions, increasing the likelihood that transit patrons can change jobs without necessarily having to change travel modes…or to change homes. The same kind of locational advantage holds true for non-residential activities situated near junction points on the transit network.

Transportation Planners. For the transportation planner, consideration of the network should be uppermost when developing individual transit lines or making changes to existing service. Do the changes being considered reinforce or undermine the larger transit network? If the latter, is it time to reformulate the network itself? What travel patterns predominate—trips to a central point or those to many dispersed locations? Is frequent service economically sustainable, or is a timed-transfer-based system more realistic? Should the lines in the network be coupled or uncoupled? The answers to these and similar questions can lead to service changes that are positive not only for those using the individual lines that are involved, but for riders throughout the larger transit system.

Summary

Transit networks are “big picture” concepts, of which most people are only dimly aware. Network patterns and functions nevertheless contribute a very great deal to the utility of transit in a community. Effective transit networks attract more riders and, as a result, reduce the impacts of the automobile. They also give residents a choice of travel mode and route. If well designed, they keep the cost of transit affordable. They play a large but generally unrecognized part in the sustainability of a metropolitan area.

Several network typologies have been utilized for transit over the years. The traditional radial pattern has given way, in many instances, to grid and hub-and-spoke networks, which can more easily serve the geographically diverse locations involved in today’s urban travel. Many systems now rely upon hybrids that contain elements of more than one network type. Ideally, each is used to advantage in handling the unique nature of the passenger flows to be accommodated. Indeed, urban travel patterns should be the ultimate determinant of transit network design.

A transit network can become the armature for the vision that guides the growth of a community. Roadways and land development can be designed and modified in concert with this network. In some cases, land use and roadway planning will follow a transit investment decision. In most others, land uses and roads will have been set in advance and transit must follow. In either case, the process should be iterative, with modifications made to the land uses and the transportation networks in order to arrive at optimum configurations of both.

While transportation planners hold the primary responsibility for transit network design, other planning and design practitioners can help foster viable transit networks—and benefit from them—if they understand the underlying principles involved. In this way, public transportation is strengthened and the community as a whole is enhanced.

References

The author is indebted to the writings of many people. Among them, the works of Kevin Lynch, Edward K. Morlock, Jr., and Vukan R. Vuchic were particularly helpful. The following are explanations and specific references for passages that were footnoted in the text:

---

[1] The idea that lines which do not cross paths do not constitute a network is true in a practical sense for transit networks but is not true for all networks. Those described in graph theory as *non-planar* could have lines passing above or below each other and still be considered part of the same network.

[2] Among such studies, the reader is referred to the following:

- Schumann, John W. (1997). Rail in Multimodal Transit Systems. *Transportation Research Record 1571.* Washington, DC: National Academy Press (Transportation Research Board).
- Thompson, Gregory L. (1977). Planning Considerations for Alternative Transit Route Structures: Concept for Improving Urban Mobility by Increasing Choices for Travel and Lifestyle. *Journal of the American Institute of Planners 43, 2:* 158-168
- Thompson, Gregory L. and & Matoff, Thomas G. (2003). Keeping Up with the Jonses: Radial vs. Multidestinational Transit in Decentralizing Regions. *Journal of the American Planning Association 69, 3:* 296-312

[3] Generally, rail and bus lines in the same network are uncoupled because they utilize different types of running ways. There are notable exceptions, however, as in cases where light rail tracks are embedded in pavements that allow buses to share their alignment. This occurs where Pittsburgh's light rail lines join the South Busway to enter the downtown area, for example, or in many European cities where buses operate on active streetcar reservations in public streets.

[4] Vuchic, Vukan R. (2005). *Urban Transit Operations, Planning, and Economics.* Hoboken, NJ: John Wiley & Sons, Inc. See especially pp. 248-259.

[5] According to the American Public Transportation Association's *2007 Public Transportation Fact Book*, American transit bus lines served 5.855 million passenger trips at an operating cost of $16,786.8 million, while paratransit services transported 125 million passenger trips at an operating cost of $2,828.4 million. The cost per passenger trip of paratransit, at $22.63, was almost 8 times higher than the cost per passenger trip of buses at $2.87.

[6] Numerous references provide the physical parameters needed by transit vehicles on city streets, such as:

- Metropolitan Transit Development Board (1993). *Designing for Transit: A Manual for Integrating Public Transportation and Land Development in the San Diego Metropolitan Area*.
- Santa Clara Valley Transportation Authority (2003). *Community Design and Transportation: A Manual of Best Practices for Integrating Transportation and Land Use*