Joshua Weiland
Space Syntax and Transit Networks

Statement:
This paper looks at nineteen subway systems from around the world, and the relationship between two sets of spatial properties and system ridership. The first hypothesis is subway systems' *intelligibility* (Hillier, 1996, p. 129) is correlated to boardings. Second, a proof-of-concept will be attempted, hypothesizing that in the Washington Metro, stations' *integration*, *connectivity* and *total depth* (Hillier, 1996) are correlated to boardings. The method of analysis was successful, but only one set of hypotheses was supported: intelligibility is not shown to be correlated to ridership, but all three syntactic attributes were shown to be correlated to the Washington Metro' station boardings.

Network analysis of metros pre-Derrible

Derrible & Kennedy (2009) cite much research linking graph theory and transportation, but say that there has been very little linking graph theory and public transportation networks, and where it has been applied it has not been linked directly to ridership (p. 3).

Derrible & Kennedy (2009) applied graph theory to nineteen metro systems of the world: Athens, Berlin, Chicago, London, Lyon, Madrid, Mexico, Montreal, Moscow, New York, Osaka, Paris, Seoul, San Francisco, Singapore, Stockholm, Tokyo, Toronto, and Washington. For each system, they compared ridership, in terms of annual boardings per capita, to three network analyses: *transit coverage*, "based on the total number of stations and land area"; *directness*, based upon "the maximum number of transfers necessary to go from one station to another"; and *connectivity*, which "attempts to get an overall view on the transfer possibilities to travel network so as to appreciate a sense of mobility" (p. 2). They found that both transit coverage and directness had 95% statistically significant correlations to boardings per capita, while the multivariate regression produced a 90% correlation.

Space Syntax

Another form of graph-based analysis is *space syntax* analysis. Developed first by Hillier & Hanson (1984), and elaborated in great degree in Hillier (1996), it is a network-based method of methodically analyzing buildings' and cities' forms. At first glance it appears to lend itself to looking at public transportation systems that are likewise network-based. During the past thirty years of research, there are many properties of space syntax analysis that have been shown to correlate well with how humans use the built environment. Two of the more basic are *integration* and *intelligibility*. Both have been shown to correlate with pedestrian and vehicle presence and so it is reasonable to hypothesize that they might correlate with transit system boardings (Hillier, 1996).

In the *Shorter Oxford English Dictionary* (2003), the second definition of "syntax" is "Orderly or systematic arrangement of parts or elements" (p. 3155). Typically, space syntax analysis looks at the network relationships of lines of sight, either represented by axial lines or isovists (Hillier, 1996). This analysis will look at the relationships between metro stations in the network that is called a metro system, and they will be represented by nodes. The syntactical measurements can apply to any network, so we can apply it here as well. It is difficult to reword a good definition of these properties, so instead I will turn to good previously written definitions, changing them so that they refer to nodes instead of axial lines.

Integration is "an indicator of how easily one can reach a specific [node] of the [system]. Mathematically, integration is an algebraic function of the number of [nodes] that must be traversed if one were to move from every [node (metro stop)] to every other [node (metro stop)] in the [node] map. The higher the integration value of a [node], the lower the number of [nodes] needed to reach that [node]" (Baran, Rodriguez & Khattak, 2008, p. 9). Integration can be analyzed globally, that is considering the entire system, or locally, considering only the few nodes that are closest. Local integration is typically used to counter the edge effect that occurs when a small section of a larger city is analyzed. In analyzing a metro system, the entire network is being looked at, so edge effect is not a worry. Additionally, the program that is used here, AGRAPH, displays global integration, but not local. Therefore, this paper shall look at global integration.

Connectivity (Hanson Hillier, or HH) is defined in space syntax differently than in Darlington: In space syntax, it is simply the number of elements, in our case nodes (metro stations) that an element (node, metro station) is connected to.

Connectivity and integration are components of *Intelligibility*, which "is defined as the correlation between the connectivity and global integration values of a [node]" within a system (Baran, Rodriguez & Khattak, 2008, p. 17). In other words, "An intelligible system is one in which well-connected spaces also tend to be well-integrated spaces" (Hiller, 1996, 129).

Total Depth "of a node n, TD(n), is the total of the shortest distances from node n to the other nodes in the systems" (Manum, Rusten & Benze, n.d., p. 98).

Method

The computer program *AGRAPH* (Manum, Rusten & Benze, 2005) was used to analyze metro systems' syntactical attributes. There are several very good programs for analyzing space syntax, notably *UCL Depthmap* (Turner, 1998) and *Confeego* (Space Syntax Limited, n.d.). *AGRAPH* is most useful for this project, however, because it has the ability to conduct analysis of a graph constructed from nodes, like transit stops, while others only will conduct analyses of graphs constructed of axial lines and isovists. Unfortunately, it is not a perfect program: as I will discuss below, the graphical representation of the nodes, each of which is rather large, potentially jeopardized the precision of the analysis (*figure 1)*.

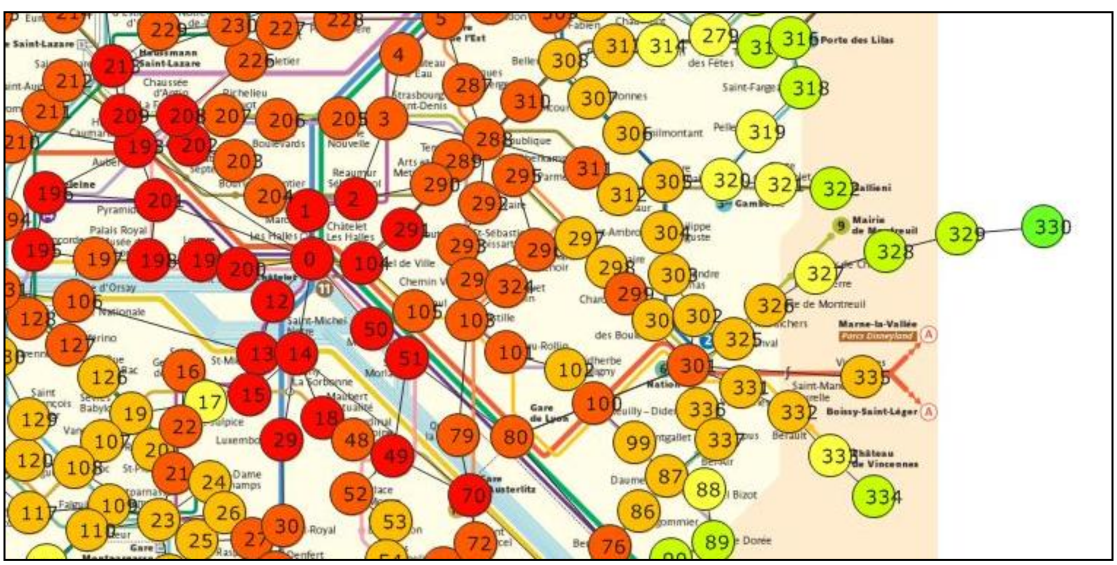

Figure 1: Showing large nodes created by AGRAPH (Paris map) and the imprecision that may have resulted.

As can be seen in *figure 1*, a node was placed (approximately) upon each metro station, and a connecting line was drawn to the corresponding node for each station that the station connected to. As opposed to Derrible & Kennedy's project, this basic syntactical analysis did not consider transfers.

Once this was complete for the entire system, AGRAPH produced several charts of information on each node, including connectivity, integration, and total depth. The resulting data was charted in Excel. And a regression was run to establish the level of correlation between the two attributes. The resulting R^2 is the level of intelligibility of the system.

The outcome variable, following Derrible & Kennedy (2009), is annual boardings per capita. They provided the data which they gathered from various sources, including the transportation systems themselves, and it will be used here as well (*figure 2)*. The data are from 2005 to 2007 (pp. 5-6).

City	Boardings per capita	Intelligibility
Athens	27.23	0.29188701
Berlin	130.42	0.24082298
Chicago	33.53	0.21504335
London	136.49	0.22353388
Lyon	106.28	0.40445243
Madrid	171.81	0.25636408
Mexico	109.57	0.20736448
Montreal	115.05	0.12285012
Moscow	235.3	0.47376518
New York	138.69	0.21085514
Osaka	100.85	0.39168219
Paris	221.46	0.22932944
San Francisco	49.97	0.08704453
Seoul	165.13	0.12395204
Singapore	101.68	0.10892772
Stockholm	156.32	0.17123953
Tokyo	197.35	0.30256498
Toronto	77.23	0.14973958
Washington	114.88	0.23295268

Figure 2: Boardings per capita and Intelligibility for nineteen metro systems. Boardings per capita from Derrible & Kennedy (2009).

Linear regression[1] was used to determine whether there is a relationship between boardings per capita and intelligibility. The results are show in *figure 3*.

[1] Thanks to: Tab Combs, Amanda Dwelley and Eric Schultheis for help with reading the Stata output.

Source	SS	df	MS
Model	5871.86631	1	5871.86631
Residual	53708.1495	17	3159.30291
Total	59580.0158	18	3310.00088

Number of obs =	19
F(1, 17) =	1.86
Prob > F =	0.1906
R-squared =	0.0986
Adj R-squared =	0.0455
Root MSE =	56.208

boardingsp~a	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
rsq	173.7589	127.4543	1.36	0.191	-95.14625	442.6641
_cons	85.10478	32.48257	2.62	0.018	16.57255	153.637

Figure 3: Stata output for regression of rsq (Intelligibility) vs. boardingsp~a (Boardings per Capita)

T = 1.36 and p = .1906. We cannot say then that the correlation is particularly strong. Nevertheless, the scatter-plot (figure 4) does have some sense about it that there is a positive relationship, though not a strong one. As nineteen is not a particularly large set, and given a larger analysis of systems, a significant relationship might be discovered.

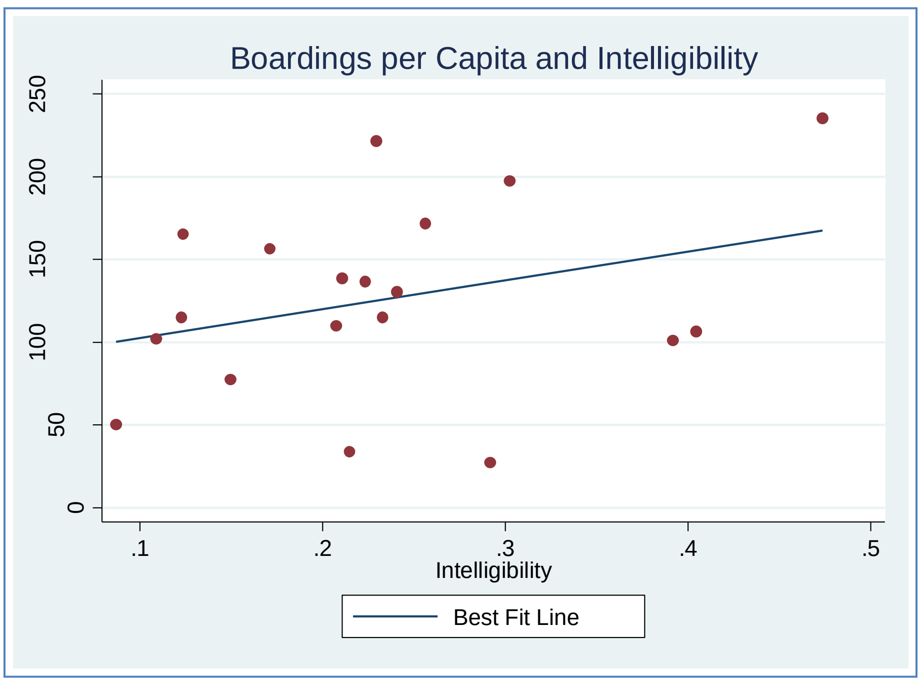

Figure 4: Graph of Boarings per Capita and Intelligibilty with best fit line.

The second hypothesis is that within a particular system, boardings will be higher at stations that are more integrated. We will run a proof of concept by applying this to the Washington Metro. Data on boardings at individual stations is difficult to come by, but that as recent as 2007 is available at Swivel.com (rob.goodspeed, 2007). The average annual boardings from 2003 to 2007 was compared against individual stations' integration, connectivity (HH) and total depth (all of which are available outputs of AGRAPH).

Connectivity (HH) and boardings

Source	SS	df	MS
Model	370167980	1	370167980
Residual	3.1999e+09	84	38093933.3
Total	3.5701e+09	85	42000686.8

Number of obs	86
F(1, 84)	9.72
Prob > F	0.0025
R-squared	0.1037
Adj R-squared	0.0930
Root MSE	6172

avg~20022007	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
connectivi~h	3308.655	1061.402	3.12	0.002	1197.94	5419.369
_cons	1273.159	2271.846	0.56	0.577	-3244.656	5790.973

Figure 5: Stata output for regression of connectivi~h (Connectivity HH) vs. boardingsp~a (Boardings per Capita)

For connectivity, R^2 = .1037; t = 3.12 and p = .0025. We therefore reject the null hypothesis that there is no relationship between connectivity (HH) and boardings. The regression results show a positive relationship between connectivity (HH) and boardings, though the low R^2 suggests that it is not a very important factor.

Integration and boardings

Source	SS	df	MS
Model	669751259	1	669751259
Residual	2.9003e+09	84	34527465.7
Total	3.5701e+09	85	42000686.8

Number of obs	86
F(1, 84)	19.40
Prob > F	0.0000
R-squared	0.1876
Adj R-squared	0.1779
Root MSE	5876

avg~20022007	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
integration	2142.148	486.3793	4.40	0.000	1174.93	3109.367
_cons	-574.0516	2056.857	-0.28	0.781	-4664.337	3516.234

Figure 6: Stata output for regression of Integration vs. boardingsp~a (Boardings per Capita)

For integration, R^2 = .1876; t = 4.40 and p = .0000. We therefore reject the null hypothesis that there is no relationship between integration and boardings. The regression results show a positive relationship between integration and boardings, though the low R^2 suggests that it is not a very important factor.

Total Depth and boardings

Source	SS	df	MS
Model	459646451	1	459646451
Residual	3.1104e+09	84	37028713.4
Total	3.5701e+09	85	42000686.8

Number of obs	86
F(1, 84)	12.41
Prob > F	0.0007
R-squared	0.1288
Adj R-squared	0.1184
Root MSE	6085.1

avg~20022007	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
totaldepth	-9.540713	2.707935	-3.52	0.001	-14.92574	-4.155688
_cons	17017.73	2630.076	6.47	0.000	11787.54	22247.93

Figure 7: Stata output for regression of Total Depth vs. boardingsp~a (Boardings per Capita)

For Total depth, R^2 = .1288; t = -3.52 and p = .0007. We therefore reject the null hypothesis that there is no relationship between total depth and boardings. The regression results show a positive relationship between total depth and boardings, though the low R^2 suggests that it is not a very important factor.

Therefore, in the case of the Washington Metro, ridership is higher in systems with higher at stops with higher connectivity, integration and lower total depth.

Discussion and Conclusion

This paper has been a trial to see whether space-syntax analysis can be used constructively with analysis of metro systems. When comparing systems, intelligibility is not a particularly good indicator of boardings per capita. Nevertheless, other syntactical measurements might have higher correlation, and a further study is warranted. As the Washington system shows, space syntax can be a good tool to analyze stations within a system. This could be explored further especially comparing between different types of systems: larger and smaller, spoke-hub or more evenly distributed. Finally, in both cases, multi-variant regressions could be run to see how highly correlated syntactical measurements are compared to other factors, including Derrible and Kennedy's, density, wealth, urban design and automobility.

Better software, or an update of AGRAPH would be useful. Most GIS programs allow zooming, which in turn allows higher precision. Smaller graphic "nodes" which better showed connections between nodes would be helpful as well for error-checking. When a complex metro system map is being analyzed, much less a bus system, it is very difficult to be accurate with the current software. Additionally, more comrehensive programs such as Depthmap and Confeego, that compute many more variables than AGRAPH, and allow zooming, could add the ability to make node-based graphs, such that metro-system analysis could occur. These hurdles may have made my analysis less accurate, especially for larger systems such as Paris'(*figure 1)*.

Works Cited

Baran, P., Rodriguez, D. & Khattak, A. (2008). Space syntax and walking in New Urbanist and Suburban neighbourhoods. *Journal of Urban Design,* Vol. 13. No. 1, 5–28.

Derrible, S. & Kennedy, K. (2009). A network analysis of subway systems in the world using updated graph theory. *Transportation Research Board 2009 Annual Meeting*.

Hillier, B. (1996). *Space is the Machine*. Cambridge : Cambridge University Press.

Hillier, B. & Hanson, J. (1984). *The Social Logic of Space*, Cambridge : Cambridge University Press.

Manum, B., Rusten, E. & Benze, P. (2005) AGRAPH, Software for Drawing and Calculating Space Syntax "Node-Graphs" and Space Syntax "Axial-Maps", in *Norwegian University of Science and Technology.* Retrieved December 3, 2009 from http://www.ntnu.no/ab/spacesyntax.

Manum, B., Rusten, E. & Benze, P. (n.d.) AGRAPH, Software for Drawing and Calculating Space Syntax "Node-Graphs" and Space Syntax "Axial-Maps", in *Norwegian University of Science and Technology.* Retrieved December 5, 2009 from http://www.spacesyntax.tudelft.nl/media/Long%20papers%20I/agraph.pdf.

rob.goodspeed. (December 06, 2007). Swivel. In WMATA Average Weekday Rail Ridership by Year. Retrieved November 6, 2009, from http://www.swivel.com/data_sets/show/1010686.

Shorter Oxford English Dictionary. (5th ed.). (2003). New York : Oxford University Press.

Space Syntax Limited (n.d.). Confeego, in *Space Syntax Laboratory*. Retrieved December 7, 2009 from http://www.spacesyntax.org/software/newtools.asp.

StataCorp. (2007). *Stata Statistical Software: Release 10*. College Station, TX: StataCorp LP.

Turner, A. (1998). UCL Depthmap, in *University College London Bartlett School of Graduate Studies.* Retrieved August 17, 2009 from http://www.vr.ucl.ac.uk/depthmap/; http://www.spacesyntax.org/software/depthmap.asp.