Australasian Transport Research Forum 2017 Proceedings
27 – 29 November 2017, Auckland, New Zealand
Publication website: http://www.atrf.info

"This is an abridged version of the paper originally submitted for ATRF 2017. For further information about this research please contact the authors."

Insights into geographical window based SSIM for comparison of OD matrices

Krishna N. S. Behara[1], Ashish Bhaskar[1] Edward Chung[1]

[1]Smart Transport Research Centre, School of Civil Engineering and Built Environment, Queensland University of Technology, Brisbane, QLD 4000, Australia

Email for correspondence: krishnanikhilsumanth.behara@hdr.qut.edu.au

Abstract

The rich structural information of travel patterns in terms of number of trips distributed to different destinations within a region is an essential element that cannot be ignored during any statistical assessments such as comparison of Origin-Destination (OD) matrices. While most of the traditional indicators fail to compare the structural differences between OD matrices, Structural Similarity (SSIM) index outperforms them by comparing the structure, mean and standard deviations within one single formulation. In the existing literature with respect to OD matrices comparison, the application of SSIM is still theoretical in nature and there is no clarity on the level of acceptability of SSIM values due to its sensitivity towards the local window size. Thus SSIM in the context of OD matrices has to be further refined especially by emphasising on the size and physical significance of local window and local SSIM values, so that it can be applied best in practice. In this light, the paper proposes a practical approach for computing SSIM based on geographical windows that has a physical significance in terms of size and shape of windows, geographical correlation and its potential to analyse local travel patterns due to different travel distributions in different sections of the network. Also, using real Bluetooth zonal OD matrices from Brisbane network, the study also demonstrates the potential of SSIM over traditional indicators.

Keywords: structural similarity (SSIM), OD matrices comparison, geographical window, Bluetooth OD, local travel patterns

1. Introduction and literature review

Estimation of Origin-Destination (OD) matrices has long been a topic of research and researchers have focused on both the static and dynamic estimation of OD matrix (Antoniou et al. 2016). Traditionally, matrix estimation techniques have used loop detector data only for optimisation. With technology advancements, other data sources such as Bluetooth (Barceló Bugeda et al. 2010) have also been explored. Most of the research has focused on the estimation of OD matrix, and there is a limited research on the use of indicators for OD matrices comparison.

The 'structure' of an OD matrix is defined as distribution of the demand between different OD pairs within the matrix. Most of the traditional indicators namely, Root Mean Square Error (RMSE), Mean Square Error (MSE), are widely used as goodness of fit measures in OD matrix estimation because of simplicity in their formulations. However, these traditional indicators compare individual cells of the OD matrices and fail to capture the overall distribution of the demand (structure) in the OD matrices. This is further illustrated using real data in Section 5. This structural property cannot be ignored during any statistical assessment such as comparison of OD matrices. Thus there is need for a potential indicator in addition to existing traditional metrics for a better comparison of OD matrices.

Limited indicators such as Structural Similarity Index (SSIM) (Djukic, Hoogendoorn and Van Lint 2013) and Wasserstein metric (Ruiz de Villa, Casas and Breen 2014) focus on assessing structural similarity of OD matrices. Wasserstein metric is based on optimisation formulation and is computationally intensive for large scale networks The concept of SSIM was originally developed in the context of images comparison (Wang et al. 2004). However, the state-of-the-art application of SSIM in transportation perspective is theoretical in nature and needs further exploration of its potential in more realistic settings by emphasising on the physical meaning of it, so that it can applied best in practice (Pollard et al. 2013).

This research considers SSIM and investigates its appropriateness and demonstrates it as a potential tool to compute the structural similarity between OD matrices.

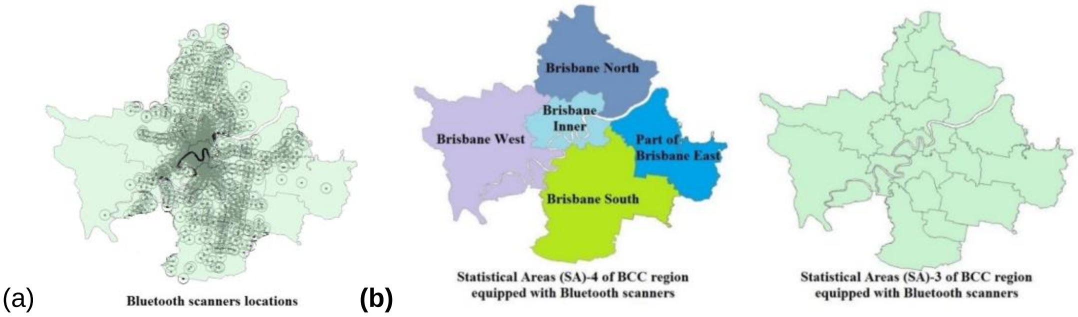
The rest of the paper is structured as follows. Description of study site and data is presented in section 2. Sensitivity of SSIM to local window size is illustrated in section 3; Geographical window based SSIM is proposed in section 4; SSIM potential over traditional indicators is investigated using real Brisbane data in section 5 and finally the paper is concluded in section 6.

2. Study site and Data

Brisbane City Council (BCC) region is chosen as the study site. Raw Bluetooth data, representing temporal detections of MAC IDs (Bhaskar and Chung 2013), is collected by BCC from over 845 Bluetooth MAC Scanners installed along many key corridors and intersections within the BCC region (see Figure1 (a)). Based on population distribution, BCC region is divided into four Statistical Areas (SA) namely SA4, SA3, SA2 and SA1 (order from higher to lower) respectively (see Figure 2(b) for SA3 and SA4 zones for BCC region).

Trips identified from Bluetooth detections (Michau et al. 2014) are critical construct for the bOD matrices (of size 845 x 845) between scanner pairs. The dimensions of bOD matrices are then reduced to 20 x 20 by geographically integrating the trips from Bluetooth detections with Statistical Area-3 (SA3) obtained from Australian Bureau of Statistics (ABS). This provides a 20 x 20 OD matrix where each cell represents an OD pair at SA3 level. In this study, Brisbane East refers to a portion of entire Brisbane East that is equipped with Bluetooth scanners.

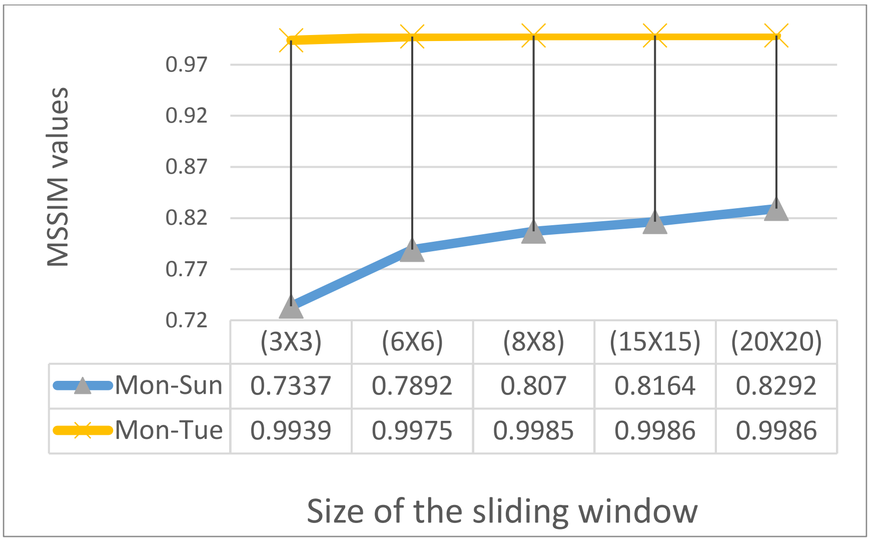
Figure 1: (a) Location of Bluetooth Scanners within BCC region; (b) SA4 and SA3 levels in BCC region



3. Sensitivity of SSIM to local window size

To demonstrate the sensitivity of SSIM with window size, consider mean SSIM (MSSIM) values computed using different window sizes (3x3 to 20x20) for Monday-Sunday (Blue line) and Monday- Tuesday (Orange line) OD matrix pairs as shown in Figure 2. Note: Here the order of OD pairs is same in both the matrices. It is observed in this study that, larger the size of sliding window, lesser is the sensitivity of SSIM towards fine correlation distortions within the OD matrix. In Figure 2, x-axis represents the size of the local window and y-axis is the MSSIM value. The MSSIM values increase as the sliding window size increases. The rate of increment of MSSIM values is less for Monday- Tuesday pair as compared to Monday- Sunday pair. This attributes to similar travel patterns between Monday-Tuesday (both of them being working days) and less similar patterns between Monday-Sunday pair. Thus, if a sliding window is used, then there is no clear consensus on the level of acceptability of the window size and its corresponding SSIM values. To overcome this limitation, SSIM based on geographical window is proposed in the next section.

Figure 2: MSSIM value increases with increase in the size of the window



	(3X3)	(6X6)	(8X8)	(15X15)	(20X20)
Mon-Sun	0.7337	0.7892	0.807	0.8164	0.8292
Mon-Tue	0.9939	0.9975	0.9985	0.9986	0.9986

4. Proposed approach-Geographical window based SSIM

Here, we propose to apply SSIM by first arranging the origin and destinations of the OD matrix in order of the geographical similarity. Thereafter, define the window for SSIM analysis as per the geographical boundaries. The origin and destinations are first grouped into respective SA4 level and the matrix is rearranged accordingly so that its cells can be grouped to form SSIM windows, where the window is having a geographical representation at SA4 level. By doing so the SSIM window has a physical significance. Note: For this the size of the window need not to be a square matrix. Since the number of SA4 OD pairs are 25, the number of geographical windows for the OD matrix is also 25 (see Figure 3).

The SA4 zones for BCC region are Brisbane East, Brisbane North, Brisbane South, Brisbane West and Brisbane Inner (see Figure 1 (b)). Figure 3 demonstrates the application of SA4 based geographical windows for comparing SA3 (20 x 20) OD matrices of Monday (Figure 3(a)) and Sunday (Figure 3(b)), respectively. For example consider a geographical window of SA4 OD pair “Brisbane East” and “Brisbane North”. It consists of SA3 OD pairs i.e. 30101-30201, 30101-30202, 30101-30203, 30101-30204, 30103-30201, 30103-30202, 30103-30203, and 30103-30204. These SA3 OD pairs are geographically correlated because they have same SA4 origin i.e. “Brisbane East” and SA4 destination i.e. “Brisbane North”. Since “Brisbane East” and “Brisbane North” consist of 2 and 4 SA3 zones respectively, the size of the local geographical window is 2 x 4. The local SSIM values are then calculated for all geographical windows exclusively and the overall MSSIM is the average of all local SSIM values. For example, MSSIM for Sunday-Monday matrices pair, computed based on geographical window is 0.7231 (see Table 1).

By averaging, it implies that, the overall SSIM value is obtained by smoothing over all local values. Although mean SSIM values are used in this study, the local SSIM values based on geographical windows have practical significance in their own respects. For example, local SSIM computed for any local window provides valuable insights towards local travel patterns between different suburbs of the region. If the purpose is to compute the similarity of Sunday and Monday travel patterns between major suburbs, then the concept of sliding window of any size will not work. From Figure 3, it can be observed for suburb pair- “Brisbane South to Brisbane North”, Sunday travel patterns are less similar to that of Monday with a local SSIM value of 0.4653 (bold in Table 1). On the other hand, for another major suburb pair- “Brisbane South to Brisbane West”, Sunday travel patterns have a better similarity value of 0.8037 as compared to that of 0.4653 (of suburb pair- “Brisbane South to Brisbane North”). This can also be justified from the fact that, local trips are more dominant during weekends. Since the South and West suburbs are closer to each other as compared to other pair, it is obvious that the total number of trips are higher, resulting in higher SSIM value as compared to their corresponding Monday trips. Thus this approach provides some valuable insights towards local travel patterns by identifying the sections of network that experience different demand distributions.

Figure 3: Splitting Monday (a) and Sunday (b) OD matrices into geographical (SA4) windows

MONDAY

Origin \ Dest		Brisbane East		Brisbane North				Brisbane South						Brisbane West				Brisbane Inner			
		30101	30103	30201	30202	30203	30204	30301	30302	30303	30304	30305	30306	30401	30402	30403	30404	30501	30502	30503	30504
Brisbane East	30101	1369	2375	9	17	79	25	157	61	36	14	18	6	0	19	13	5	106	49	127	20
	30103	2184	18080	79	82	416	215	1308	491	438	192	207	135	20	216	89	58	635	720	496	140
Brisbane North	30201	8	103	5565	1637	804	1973	27	84	29	46	32	10	14	4	155	286	206	32	522	107
	30202	20	85	1534	9195	756	814	38	194	58	78	38	26	20	3	158	483	523	38	1708	234
	30203	90	618	716	882	12274	1669	192	282	170	187	213	76	54	11	367	146	782	73	2044	310
	30204	26	405	1690	868	1787	14200	118	113	80	71	56	44	17	3	115	90	246	36	568	99
Brisbane South	30301	141	1589	26	54	206	122	8296	1641	1004	272	122	89	23	371	117	48	966	740	430	207
	30302	55	504	74	178	312	93	1661	15291	1847	2096	808	315	135	65	594	228	2978	789	1916	678
	30303	38	494	42	54	195	85	974	1951	19523	1355	679	1134	51	37	231	106	1245	192	722	309
	30304	16	220	55	104	238	76	266	2621	1343	8889	1398	1636	71	25	443	163	1643	294	1187	363
	30305	19	194	32	40	219	65	126	844	648	1509	17893	2017	184	9	505	60	563	80	434	120
	30306	4	115	11	25	100	36	105	341	1234	1486	2137	7717	38	8	90	26	455	40	260	74
Brisbane West	30401	2	29	10	16	52	14	23	160	49	76	221	32	6050	34	564	51	192	21	254	223
	30402	16	134	2	3	17	3	224	55	35	24	16	10	53	805	50	4	103	28	36	19
	30403	14	88	135	163	365	102	114	657	192	401	474	99	482	55	11424	382	1230	114	1765	1809
	30404	9	68	286	496	134	96	37	277	71	98	76	20	44	5	364	6153	640	51	1174	942
Brisbane Inner	30501	114	651	198	540	727	264	959	3370	1020	1107	539	399	181	104	1273	616	34177	980	5077	2015
	30502	63	783	19	57	73	39	732	931	178	278	94	47	20	39	120	46	1094	3571	557	210
	30503	116	597	501	1693	2099	643	435	2180	633	908	467	223	232	39	1757	1166	5310	593	26786	1919
	30504	25	135	99	241	270	97	225	783	290	219	131	68	221	21	1875	995	2040	219	2090	10052

(a)

SUNDAY

Origin \ Dest		Brisbane East		Brisbane North				Brisbane South						Brisbane West				Brisbane Inner			
		30101	30103	30201	30202	30203	30204	30301	30302	30303	30304	30305	30306	30401	30402	30403	30404	30501	30502	30503	30504
Brisbane East	30101	841	1515	7	10	19	15	122	31	42	14	6	4	1	21	6	2	45	34	77	10
	30103	1495	11690	22	36	88	94	856	211	232	63	65	71	12	201	59	25	340	508	253	90
Brisbane North	30201	10	41	3765	1014	290	1509	13	39	15	12	11	12	14	1	120	214	105	12	199	75
	30202	9	37	989	6576	403	454	21	110	46	36	11	13	13	5	115	305	278	24	1040	157
	30203	19	120	297	453	4042	918	56	112	47	57	41	15	38	3	135	47	353	34	862	121
	30204	25	248	1393	468	892	10043	72	66	57	35	25	20	10	5	72	56	140	20	327	60
Brisbane South	30301	128	1086	15	32	50	63	7150	1068	837	177	58	93	16	289	82	45	553	578	326	133
	30302	27	191	46	163	163	79	1045	11067	1172	1122	410	194	86	26	473	218	2254	498	1505	502
	30303	34	221	11	33	56	53	811	1175	13909	807	291	1071	44	25	156	75	666	114	411	193
	30304	7	65	6	36	76	35	178	1188	815	5392	504	1090	54	15	193	34	651	122	470	109
	30305	5	67	8	24	43	18	60	437	294	529	7403	1289	102	7	263	24	266	37	157	53
	30306	4	66	6	14	36	24	78	187	1013	1085	1277	6534	31	4	75	21	355	34	210	68
Brisbane West	30401	2	11	9	13	25	10	14	84	39	53	103	28	4484	23	395	35	118	16	150	130
	30402	9	108	2	2	5	2	168	25	23	5	9	6	38	657	19	7	37	15	18	14
	30403	7	49	90	120	168	80	85	422	144	200	246	62	330	42	7752	251	768	96	1177	1121
	30404	4	31	190	323	45	78	50	160	56	26	27	19	25	5	223	4647	411	36	701	609
Brisbane Inner	30501	71	363	87	322	361	175	606	2262	705	669	261	335	122	49	901	436	24048	635	3507	1297
	30502	39	503	11	32	37	22	618	523	112	161	37	36	16	22	113	47	694	2899	375	146
	30503	82	336	190	982	884	381	347	1385	427	458	157	226	150	31	1178	694	3481	407	18112	1190
	30504	16	64	66	131	120	60	151	417	198	99	35	76	121	15	1201	624	1296	141	1191	7060

(b)

Table 1: Local SSIM values based on geographical windows computed for BCC region

	Brisbane East	Brisbane North	Brisbane South	Brisbane West	Brisbane Inner
Brisbane East	0.8319	0.2437	0.7650	0.9517	0.7755
Brisbane North	0.3311	0.7353	0.4034	0.7378	0.6299
Brisbane South	0.7771	**0.4653**	0.8062	**0.8037**	0.8117
Brisbane West	0.8340	0.7754	0.7562	0.8884	0.8165
Brisbane Inner	0.7716	0.6265	0.8257	0.8385	0.8750
Mean SSIM (MSSIM)					0.7231

5. MSSIM vs Traditional indicators using real OD matrices

To demonstrate the potential of MSSIM over traditional indicators, real Bluetooth OD matrices for 88 days (consisting of Sundays, Saturdays, Public holidays and School holiday weekdays) are compared with a typical working weekday using all five indicators (see Table 2). A typical working weekday OD matrix is computed by taking average of 75 regular working weekdays from the year 2016. Before comparing, with other distance metrics, MSSIM is converted into a dissimilarity measure (i.e. 1-MSSIM). SSIM can be an effective tool when traditional indicators are not capable of capturing structural differences. For example, Table 2 presents the OD matrices comparison results for 27th March 2016 (Easter Sunday) and 26th December 2015 (Saturday following Christmas day) as compared to aforementioned typical working weekday.

Table 2: MSSIM overcomes limitation of traditional indicators- real OD matrices comparisons

Date (Day)	1-MSSIM	MSE	RMSE	GU	MAE
27th March 2016 (Easter Sunday)	0.3352	1.31*10^9	3.63*10^4	0.35	0.55
26th December 2015 (Saturday following Christmas day)	0.4296	1.31*10^9	3.63*10^4	0.35	0.53

It is expected that both these days (Easter Sunday and Christmas break) should have higher non-work related trips such as private travel towards different destinations (parks, relatives or friends' homes' etc.) as compared to a regular working weekday. However, since, 26th December 2015 is amidst of a long 15-day Christmas and New Year Holidays, it can have different travel pattern as compared to that of Easter Sunday (from a short (4-day) holiday span). This difference is not captured by MSE, RMSE, GU and MAE but it is predominant in the MSSIM values.

6. Conclusion

The paper proposes geographical window based SSIM to address the sensitivity issues of SSIM to local window size. This is ensured by adjusting the local window boundaries such a way that it captures only those lower level zonal OD pairs (SA3) belonging to a particular higher level zonal (SA4) OD pair. This approach allows SSIM to adapt to local sub-matrix statistics rather than computing on the entire size of OD matrix. SSIM computed based on geographical window has physical significance especially when local SSIM values throw some light on the local travel patterns comparison. Also it seems to be more appropriate than choosing a sliding window of random size without any physical understanding of what the size represents.

Although many statistical indicators exist for OD matrices comparison, we choose, to select SSIM as it is robust in accounting the structural properties of OD matrices i.e. travel patterns. Using real OD matrices, SSIM values are compared with that of traditional indicators' and it is observed that, when traditional indicators fail to distinguish the structural differences between OD matrices, SSIM can be used as a potential metric.

As a part of future scope, travel patterns shall also be analysed from the route choices perspective and shall be compared with current study SSIM interpretations.

7. References

Antoniou, Constantinos, Jaume Barceló, Martijn Breen, Manuel Bullejos, Jordi Casas, Ernesto Cipriani, Biagio Ciuffo, Tamara Djukic, Serge Hoogendoorn and Vittorio Marzano. 2016. "Towards a generic benchmarking platform for origin–destination flows estimation/updating algorithms: Design, demonstration and validation." *Transportation Research Part C: Emerging Technologies* 66: 79-98.

Barceló Bugeda, Jaime, Lídia Montero Mercadé, Laura Marqués and Carlos Carmona. 2010. "A Kalman-filter approach for dynamic OD estimation in corridors based on bluetooth and Wi-Fi data collection." In *12th World Conference on Transportation Research WCTR, 2010*, edited.

Bhaskar, Ashish and Edward Chung. 2013. "Fundamental understanding on the use of Bluetooth scanner as a complementary transport data." *Transportation Research Part C: Emerging Technologies* 37: 42-72.

Djukic, Tamara. 2014. "Dynamic OD demand estimation and prediction for dynamic traffic management." PhD Thesis, TU Delft, Delft University of Technology.

Djukic, Tamara, Serge Hoogendoorn and Hans Van Lint. 2013. "Reliability assessment of dynamic OD estimation methods based on structural similarity index." In *Transportation Research Board 92nd Annual Meeting*, edited.

Michau, Gabriel, Alfredo Nantes, Edward Chung, Patrice Abry and Pierre Borgnat. 2014. "Retrieving trip information from a discrete detectors network: The case of Brisbane Bluetooth detectors." Paper presented at the 32nd Conference of Australian Institutes of Transport Research (CAITR 2014), University of New South Wales, Sydney,NSW, 17-18 February 2014.

Michau, Gabriel, Nelly Pustelnik, Pierre Borgnat, Patrice Abry, Alfredo Nantes, Ashish Bhaskar and Edward Chung. 2017. "A primal-dual algorithm for link dependent origin destination matrix estimation." *IEEE Transactions on Signal and Information Processing over Networks* 3 (1): 104-113.

Pollard, Tim, Nick Taylor, Tom van Vuren and Mott MacDonald. 2013. "Comparing the Quality of OD Matrices in Time and Between Data Sources." In *Proceedings of the European Transport Conference*, edited.

Ruiz de Villa, Aleix, Jordi Casas and Martijn Breen. 2014. "OD matrix structural similarity: Wasserstein metric." In *Transportation Research Board 93rd Annual Meeting*, edited.

Wang, Zhou, Alan C Bovik, Hamid R Sheikh and Eero P Simoncelli. 2004. "Image quality assessment: from error visibility to structural similarity." *IEEE transactions on image processing* 13 (4): 600-612.