A Study of Contact Durations for Vehicle to Vehicle Communications

Quynh Nguyen and Bhaskar Krishnamachari

Abstract Emerging vehicular networks will take advantage of vehicle to vehicle short range communications. In such networks the V2V link is active only so long as the two nodes stay within communication range of each other. This contact duration which is a statistical phenomenon affected by the mobility of vehicles on the road network will have a significant impact on the network performance, and thus merits deeper understanding. We undertake an intensive study of contact duration using real trace data from taxis in Shanghai assuming a fixed-range radio. We quantify the aggregate contact duration statistics, as well as the contact duration statistics conditioned on factors such as time, location and vehicle directions.

1 Introduction

Vehicular network (VANETs) has become an emerging area attracting extensive research effort in the past few years. In VANETs, it is challenging to create an end-to-end path between any pair of nodes due to either extremely dynamic and node-behaviour dependent mobility or very sparse network architecture [1].

In these opportunistic VANETs, a contact (an encounter) between any pair of nodes is generated when the pair is within communication range of each other. However, nodes in the networks keep moving in their own ways so communication links among mobile nodes are on and off continuously; the pair will no longer be in contact when one node moves out of the other node's vicinity area or its link quality fluctuates. Therefore, there are hardly any existing complete path from sources to

---

Quynh Nguyen
Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, California 90036 e-mail: `quynhngu@usc.edu`

Bhaskar Krishnamachari
Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, California 90036 e-mail: `bkrishna@usc.edu`

destinations, and even discovered complete paths are very unstable and of short term duraton. In those VANETs whose links among nodes are generated highly dynamically and intermittently, store-carry-forward techniques and opportunistic data dissemination framework are introduced as a way to overcome intermittent connectivity, although relative high delay is sometimes the cost to get information from one source to another destination.

Understanding how the contact process really works and knowing related encounter statistics can improve data distribution demand in terms of either reducing the overall delay, saving relaying bandwidth, local storage buffer and energy cost at intermediate nodes or increasing data throughput and tranfer reliability, etc. A pair of nodes having higher meeting frequency and shorter inter-contact time [6–8] compared to the others will have better opportunity to transmit data to each other. Any nodes having higher encounter frequency with larger group of neighboring nodes within a shorter time duration can become better candidates as intermediate nodes for relaying data in routing algorithms. Any pair that has contact duration longer than those of other pairs, which is a sign that contact link between them is more stable and data transferred between them, is also more a reliable candidate to reach a given destination successfully.

Among multiple crucial parameters of the contact process, contact duration is a critical parameter [9, 10]. We consider in this work an intensive study about vehicular contact duration based on the data obtained by Shanghai Jiao Tong University [15]. We make two sets of contributions:

- We analyze the aggregated contact duration distribution based on the real taxi data trace, and
- We quantify the contact duration conditioned on different parameters including time, location and vehicle directions.

The rest of this paper is organized as follows: section 2 lists related work; section 3 gives an introduction about the data set as well as the methodology for the vehicular contact duration study; section 4 shows observed fact and draws conclusions about vehicle contact duration distribution and statistics and impact of time, location and direction on vehicle contact. And finally, we present a concluding discussion in section 5.

2 Related Work

Contact duration has become an important parameter applied by multiple researchers in different protocols and techniques to enhance data access, data transfer and network connectivity in opportunistic in general delay tolerant networks as well as vehicular networks. The authors in [11] presented Link Contact Duration-based Routing Protocol to deliver as many messages as possible within a short time. In [12], PRoPHET, a probabilistic routing protocol for Delay Tolerant Networks, both intermeeting time and contact duration are taken into account to compute the delivery probability to improve performance of the proposed routing protocol. Another work applying contact duration and meeting frequency so as to estimate message delivery probability and present a novel routing algorithm is [13]. X. Zhuo *et al.* also used contact duration to enhance the traditional cooperative caching protocol to improve the performance of data access in DTNs. In [10], the same technique is used to improve data replication for data sharing in DTNs.

Furthermore, there are also other works trying to characterize contact duration patterns in a vehicular network. Y. Li *et al.* [9] has carried out experiments using Beijing and Shanghai traces to study the contact duration characteristics. They concluded that the contact duration obeys an exponential distribution, while beyond a characteristic time point it decays as a power law one. While many works such as this have been utilizing contact duration as a crucial factor to help enhance the performance of their protocols in DTNs and VANETs, our work mainly brings into focus the impact of multiple factors (time, location and direction) on contact duration distribution after intensively analyzing and studying contact information obtained from the Shanghai taxi trace.

3 Dataset introduction and approach methodology

3.1 Shanghai Dataset

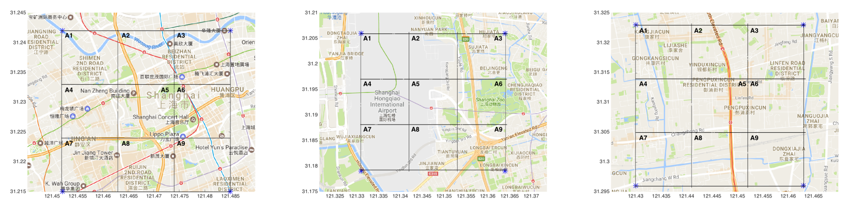
SUVnet-Trace Data, obtained from Wireless Sensor Networks Lab (WnSN) Shanghai Jiao Tong University includes GPS information of roughly 2400 taxis in Shanghai city from January 31 to February 27th. The coverage of the area is 6340 km square. Figure 1 shows the total region covered by the Shanghai data set.

We have studied contact duration for 10 weekdays from Feb 5 to Feb 16 at two different time slots 7am-9am and 11am-1pm. The first slot is the rush period while the second one represents the normal hours. After dividing the whole region into 1kmx1km grid, we compute densities of all cells, and have picked some cells for further investigation. Figure 3 show corresponding maps for 3 types of areas. We chose region 1 as a representative for a densly populated residence area including the center of residential district, Shanghai concert hall and Jin Jiang tower. Region 2 is mainly parts of the Shanghai Hongqiao International Airport (A1, A2, A4, A5, A7 and A8); A6 covers the Shanghai Zoo; and A3 and A9 are civil areas.. Finally, region 3 is a sparser residence area (Pengpuxincun and Linfen Road Residential district).

Table 1 shows the average density of both time slots (7am-9am and 11am-1pm) of all areas in corresponding regions over 10 days. The density unit is [number of vehicles/km^2/minute]. We have made the standard assumption that the effective V2V communication radio range is 300m using 802.11p [3].

From the table, we could see clearly that Region 1 has very high density (ranging between 9.75 and 22.9), including many residential districts, Shanghai Concert Hall

Fig. 1: Shanghai Covered Region



Dense residence Region   Dense airport Region   Sparse residence Region

Fig. 3: Shanghai Data Coverage Description

Area	Region 1	Region 2	Region 3
A1	10.9238	0.1388	1.2625
A2	12.7233	3.6996	2.6158
A3	9.7842	5.2100	4.9737
A4	19.2858	0.5513	2.5329
A5	22.8887	21.0396	4.5896
A6	17.3758	2.3792	2.1292
A7	16.7371	0.0558	2.9500
A8	15.5487	0.9517	3.9387
A9	15.7575	1.8817	2.4517

Table 1: Average Density of all areas in each region

and Jin Jiang Tower. The traffic in Region 3 is much less dense (less than 5). In region 2, there are almost no traffic in A1, A4, A7 (less than 1); and A5 is the densest area compared to all other areas (almost 21) because it is the entry of Shanghai Hongqiao Internaltional Airport Terminal 1 and has all the major airport shuttle stations, airport shopping malls... which is the major places to pick up and drop off for taxis services.

3.2 Methodology

There are a few parameters which impact contact duration that we want to analyze, including time, location and direction. Analyzing the selected 3x3 cells area described in 3.1, we can collect all empirical contact duration giving GPS and timestamp information for each cell within the interesting period. In the study, we used two sample Kolmogorov Smirnov test and two sample Anderson Darling test which are nonparametric tests for validation if the two samples generated from the same distribution. Both of them are statistic tools used to test whether two underlying one-dimensional probability distributions differ. However, according to [4, 5], Anderson-Darling test is used more reliably for smaller number of samples while KolmogorovSmirnov test is suitable for larger number of samples. Based on number of samples (larger or smaller than 10000), we performed correspondingly suitable validation test to decide if any two empirical contact duration set at different time, location or direction are generated from the same distribution meaning that they share the same contact duration characteristics. Based on those results, we quantified time, location and direction impact on contact duration by estimating portion of which they have the same distribution given corresponding conditions.

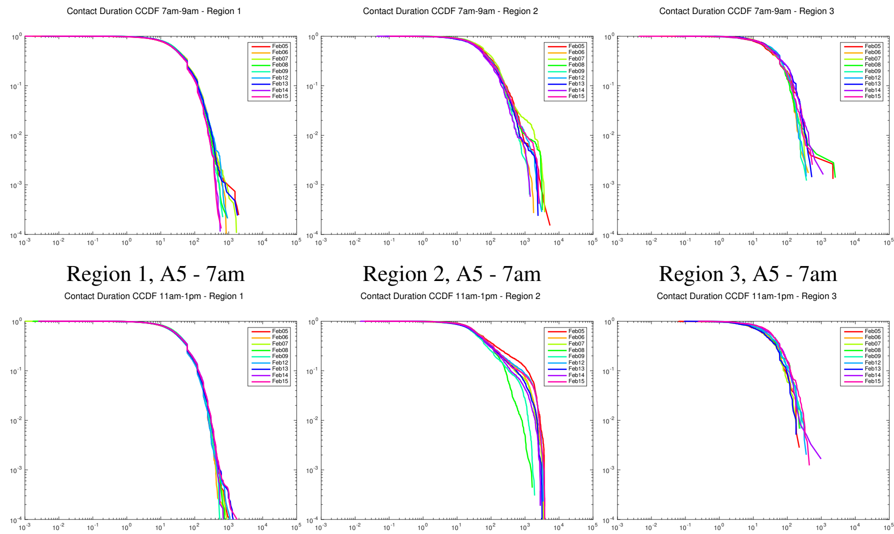
We will make the contact duration statistics dataset publicly available at http://anrg.usc.edu/www/Downloads/

4 Factual observations

This section shows the observation of encounter duration statistics varying time, location and direction conditions, and examines their impact.

4.1 Time

Figures 5 describe the CCDF (complimentary cumulative density function) result of aggregated contact duration over 10 days at 7am and 11am correspondingly for area 5 of 3 regions . Different colors represent different days. For both periods of week days, we can see that even though there is little variation for larger duration, the encounter duration information at the same area on different days tend to match each other quite well, especially for smaller values. It appears that there are stable traffic flow characteristics within a specific area even for different weekdays, resulting in a stable contact duration pattern for each area.

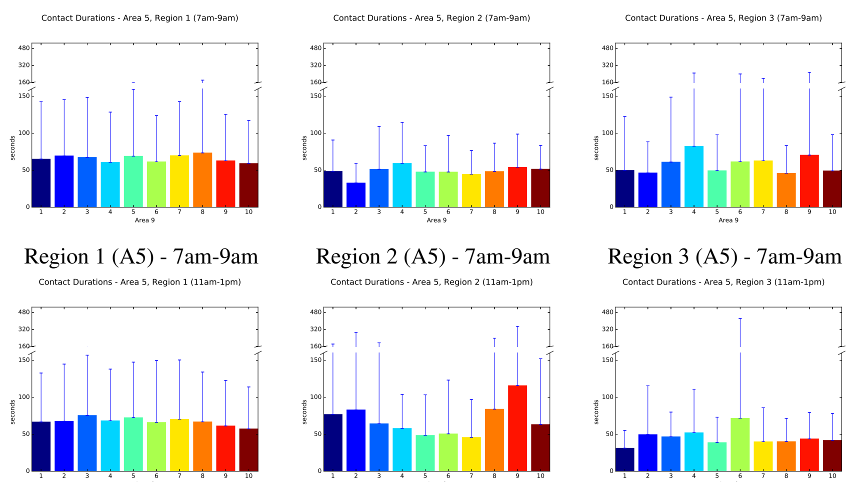


Region 1, A5 - 7am | Region 2, A5 - 7am | Region 3, A5 - 7am

Region 1, A5 - 11am | Region 2, A5 - 11am | Region 3, A5 - 11am

Fig. 5: Contact Duration CCDF - Area 5 of 3 regions over 2 time slots - Time comparison

Furthermore, the bar charts in Figures 7 illustrate the changing of encounter duration through 10 days for Area 5. We can still see the steady mean and very little variation across different days within the same area even for both periods.



Region 1 (A5) - 7am-9am | Region 2 (A5) - 7am-9am | Region 3 (A5) - 7am-9am

Region 1 (A5)- 11am-1pm | Region 2 (A5) - 11am-1pm | Region 3 (A5) - 11am-1pm

Fig. 7: Mean and variance - Time comparison

Finally, Table 2 shows an average of results from applying 95%-confidence validation test (Kolmogorov-Smirnov and Anderson Darling, depending on the number of samples available in each case) which indicate whether the contact duration distributions are similar, over different pairs of days. Higher numbers indicate more similar distributions for a given area across different pairs of days.

Location	Region1-I1	Region1-I2	Region2-I1	Region2-I2	Region3-I1	Region3-I2
A1	0.6000	0.2222	0.75	0.9556	1	0.7556
A2	0.2667	0.0667	0.5333	0.5556	0.4889	0.6000
A3	0.1778	0.2444	0.2444	0.2667	0.3556	0.2667
A4	0.1333	0.0222	0.3778	0.4222	0.4667	0.9556
A5	0.0889	0.0667	0.0667	0	0.3111	0.3778
A6	0.2667	0.0444	0.5333	0.1778	0.6222	0.7333
A7	0.1778	0.0222	0.5	0.3929	0.6667	0.9333
A8	0.1778	0.0667	0.5333	0.8000	0.2889	0.3556
A9	0.2000	0.1556	0.7111	0.6667	0.6222	0.8000

Table 2: KS and AD test result: portion of pair of days over varying location

In summary, we can give some conclusions:

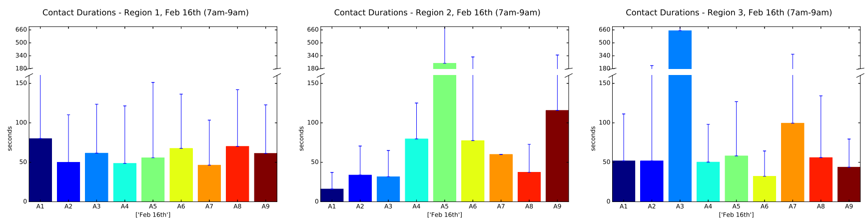
- Each area has its own contact distribution pattern on diferrent days. Different days do not have large impact on contact distribution within the same area.
- Some areas show a greater variation in the contact distribution over time than others.
- The variation in contact distribution can also depend on particular time of the day

4.2 Location

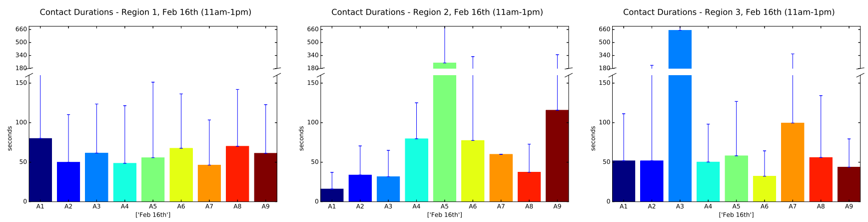
Similarly, the bar charts in Figures 9 illustrate the changing of encounter duration through 9 areas for one specific day Feb 16th. In these figures, we can see clear variation across different areas within the same day for both periods. In region 2, the mean encounter duration for Area 5 (the densest area having entry of Shanghai Hongqiao Internaltional Airport Terminal 1) and Area 9 (the major intersection Longbai'ercun) is much higher compared to all other areas. There is also a peak in Area 3 in region 3.

Table 3 shows proportion of pairs of areas where a 95%-confidence KS and AD-test indicating that the contact duration distributions are identical. Higher numbers indicate more similar distributions across areas on a given day. There is little difference in portions of similarity between the two periods. However, portions of similarity among different locations are quite low compared to the calculated portions among days. Therefore, we can see a weak impact of location on contact duration. There is low coherence in terms of contact duration at different locations.

In summary, we can give some conclusions:



Region 1, Feb 16th - 7am-9am Region 2, Feb 16th - 7am-9am Region 3, Feb 16th - 7am-9am



Region 1, Feb 16th - 11am-1pm Region 2, Feb 16th - 11am-1pm Region 3, Feb 16th - 11am-1pm

Fig. 9: Mean and variance - Location comparison

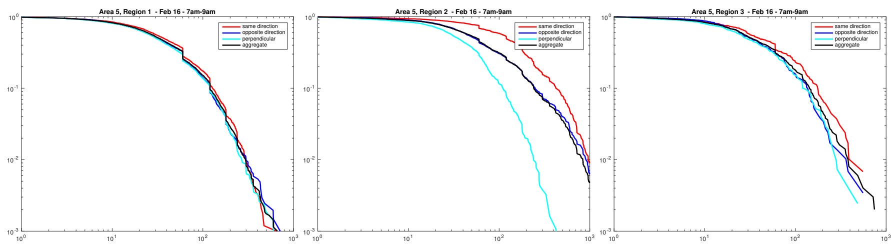
Day	Region1-I1	Region1-I2	Region2-I1	Region2-I2	Region3-I1	Region3-I2
Feb 5th	0.0556	0	0.3214	0.1944	0.1944	0.4167
Feb 6th	0.0278	0.0278	0.5	0.2500	0.3611	0.2500
Feb 7th	0.0278	0.0278	0.2143	0.3056	0.3611	0.3889
Feb 8th	0.0556	0.0278	0.1071	0.3571	0.4167	0.3889
Feb 9th	0.0278	0	0.3571	0.4722	0.3611	0.5556
Feb 12th	0.0278	0	0.4643	0.2778	0.4167	0.4722
Feb 13th	0.0278	0	0.1389	0.4444	0.3333	0.4444
Feb 14th	0.0278	0.0278	0.4167	0.2222	0.1389	0.5278
Feb 15th	0.0278	0.0278	0.2500	0.3611	0.4444	0.3056
Feb 16th	0.0278	0	0.2857	0.4167	0.2500	0.1944

Table 3: KS and AD test result: portion of 9 areas over 10 days

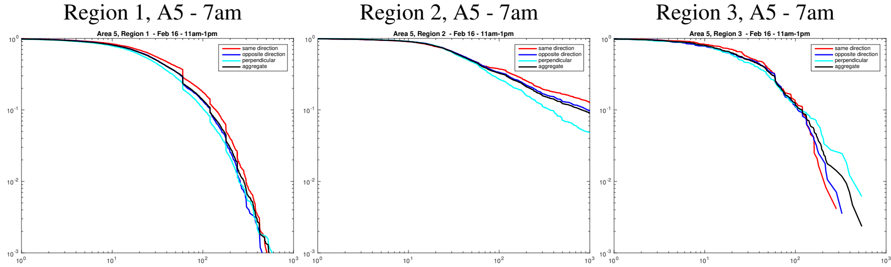
- Each area has its own contact distribution pattern. Different locations do have large impact on contact distribution within the same day.
- The variation of the mean contact duration across areas depends on the particular day.

4.3 Direction

Figures 11 describe the CCDF (complimentary cumulative density function) result of aggregated contact duration as well as contact distribution generated for same direction, opposite direction and perpendicular direction at Area 5 over 10 days at 7am and 11am correspondingly. Vehicles travelling on different directions East, West, South and North are identified and contact duration is estimated for vehicles going on the same, opposite or perpendicular direction. For both time intervals, CCDF for different types of relative directions are quite different. The CCDF generated from the opposite direction flow of vehicles seems most close to the aggregated CCDF compared to the other two types.



Region 1, A5 - 7am Region 2, A5 - 7am Region 3, A5 - 7am



Region 1, A5 - 11am Region 2, A5 - 11am Region 3, A5 - 11am

Fig. 11: Contact Duration CCDF - Area 5 of 3 regions over 2 time slots - Direction comparison

Table 4 shows proportion of pairs of areas and days having the majority direction (same, opposite, perpendicular direction) where a 95%-confidence KS and AD-test indicating that the contact duration distributions are identical. Higher numbers indicate more similar distributions across areas on a given day.

Direction	**Region1-I1**	**Region1-I2**	**Region2-I1**	**Region2-I2**	**Region3-I1**	**Region3-I2**
Same	0.0667	0.0111	0.4444	0.4103	0.5955	0.7000
Opposite	0.9778	0.8889	0.9333	0.8767	0.9775	0.9773
Perpendicular	0.0667	0.0222	0.5200	0.4247	0.6854	0.7386

Table 4: KS and AD test result: portion of 9 areas over 10 days

In summary, we can give some conclusions:

- The contact duration distributions are sensitive to the pairwise direction of the two vehicles contacting each other. The shortest contact duration is for vehicles going in perpendicular direction to each other.
- The contact durations corresponding to opposite direction contacts have the least variation across days. The contact durations for same direction contact have the most variation across days.

5 Conclusion

This is one of the first studies in the literature to undertake a fine-grained analysis of contact duration from a real set of vehicular traces. Besides presenting the aggregate contact duration, we have examined and presented how contact durations vary with area, time of day, different days, as well as relative pairwise vehicular directions. We find that the contact durations tend to be somewhat stable over different days, but vary over areas. The pairwise direction of vehicles is shown to have a significant impact on contact duration: vehicles going in the same direction tend to have the longest contact duration, while vehicles going in a perpendicular direction to each other were found to have the shortest contact duration.

References

1. F. W. Scholz; M. A. Stephens, ”K-Sample Anderson-Darling Tests”, Journal of the American Statistical Association, Sep 1987
2. A. Jamalipour and Y. Ma, ”Intermittently Connected Mobile Ad Hoc Networks From Routing to Content distribution”, Springer-Verlag, New York 2011.
3. J. Yin, T. ElBatt, G. Yeung, B. Ryu, S. Habermas, H. Krishnan, and T. Talty, ”Performance evaluation of safety applications over DSRC vehicular ad hoc networks”, *ACM*, 2004
4. Sonja Engmann, Denis Cousineau, ”Comparing distributions: the two sample Anderson-Darling Test as an alternative to the Kolmogorov-Smirnoff Test”, Journal of Applied Quantitative Methods, Fall 2011
5. https://stat.ethz.ch/R-manual/R-patched/library/stats/html/ks.test.html
6. Hongzi Zhu, Minglu Li, Luoyifu Guangtao xue, Yanmin Zhu, ”Impact of Traffic Influxues: Revealing Exponential Inter-Contact Time in urban VANETs”, IEEE Transactions on Parallel and Distributed Systems
7. Augustin Chaintreau, Pan Hui, Jon Crowcroft, Christophe Diot, Richard Gass, and James Scott, ”Impact of Human Mobility on the Design of Opportunistic Forwarding Algorithms”, IEEE Infocom, 2006.
8. Xiaolan Zhang, Jim Kurose, Brian Neil Levine, Don Towsley, Honggang Zhang, ”Study of a Bus-based Disruption-Tolerant Network: Mobility Modeling and Impact on Routing”, IEEE Infocom, 2007.
9. Yong Li, Depeng Jin, Lieguang Zen, Sheng Chen, ”Revealing Patterns of opportunistic contact durations and intervals for Large Scale Urban vehicular Mobility”, ICC 2013.
10. X. Zhuo, Q.Li, W.Gao, G.Cao, Y.Dai, ”Contact duration aware data Replication in Delay tolerants network”, 19th IEEE International Conference on Network Protocols, 2011.
11. Kyoung-Hak Jung , Wan-Seon Lim, Jae-Pil Jeong, Young-Joo Suh, ”A link contact duration-based routing protocol in delay-tolerant networks”, Springer Wireless Journal, January 2013.
12. H. J. Lee, J. C. Nam, W. K. Seo, Y. Z. Cho and S. H. Lee, ”Enhanced PRoPHET routing protocol that considers contact duration in DTNs,” 2015 International Conference on Information Networking (ICOIN), Cambodia, 2015.
13. Chen Yu, Zhongqiu Tu, Dezhong Yao, Feng Lu andHai Jin, ”Probabilistic routing algorithm based on contact duration and message redundancy in delay tolerant network”, International Journal of Communication Systems,2015.
14. X. Zhuo, Q. Li, G. Cao, Y. Dai, B. Szymanski and T. La Porta, ”Social-Based Cooperative Caching in DTNs: A Contact Duration Aware Approach,” 2011 IEEE Eighth International Conference on Mobile Ad-Hoc and Sensor Systems, 2011.
15. http://wirelesslab.sjtu.edu.cn/taxi_trace_data.html