Technische Universität Hamburg

ComNets
Institute of Communication Networks
TUHH

PARAMETERIZATION OF THE SWIM MOBILITY MODEL USING CONTACT TRACES

OMNeT++ Summit 2017, Bremen

Zeynep Vatandas, Manikandan Venkateswaran, Koojana Kuladinithi, Andreas Timm-Giel

Hamburg University of Technology

Date: 07.09.2017

Contents

- ✓ **Small Worlds in Motion (SWIM) Mobility Model**

Small Worlds In Motion (SWIM)

- SWIM is a model based on purely location preferences
- It is a simple model with few parameters to tune
- SWIM is based on two intuitions of human mobility
  - People prefer nearby locations to their homes
  - If it is a far away place, then it will most likely be popular

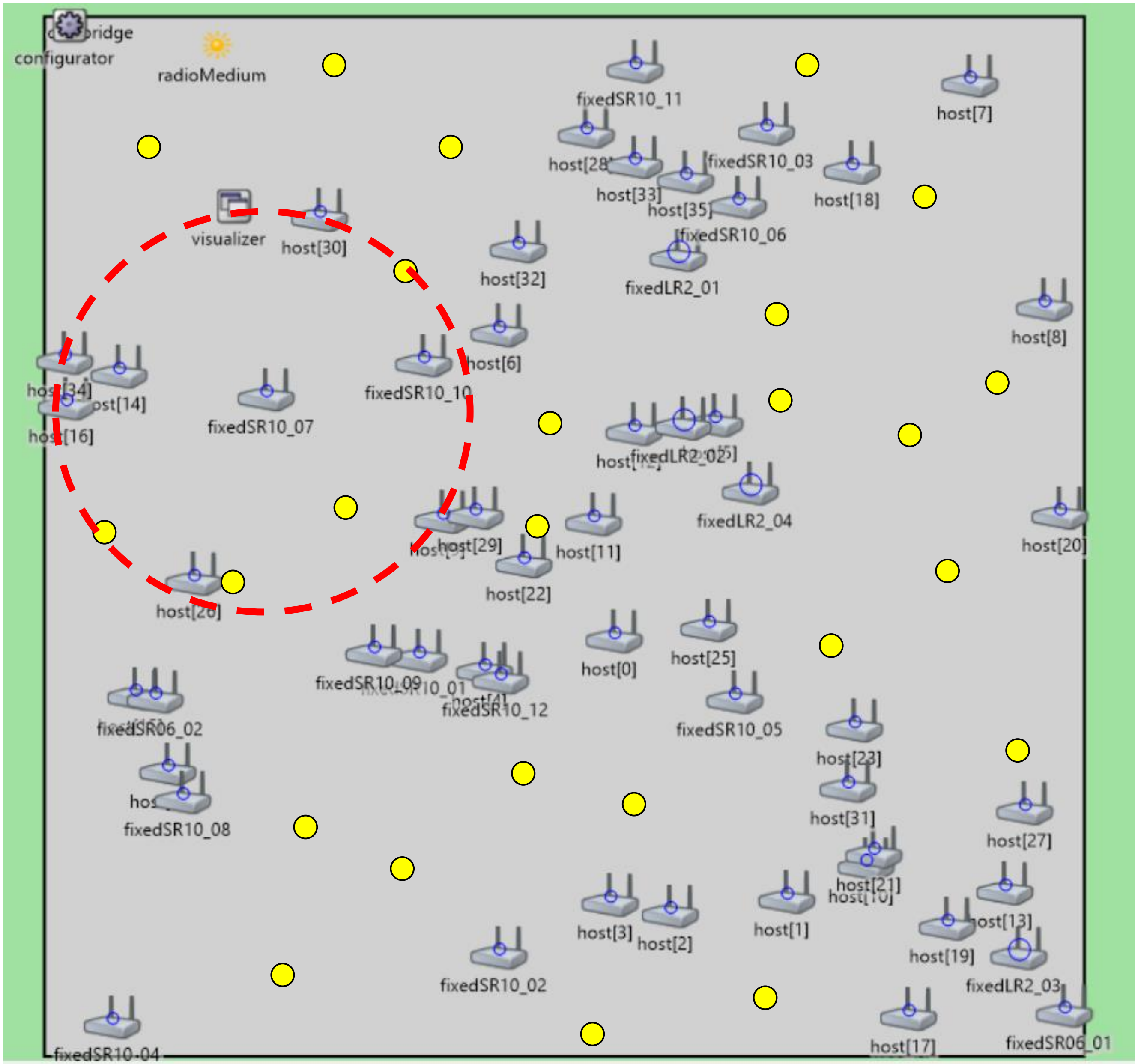
Small Worlds In Motion (SWIM)

- Each location C is assigned with a weight by each node

$$\mathbf{w(C)} = \alpha \cdot \mathbf{distance(h,C)} + (1-\alpha) \cdot \mathbf{seen(C)}$$

- distance(h,C) is a function which decays as a power law of distance
- Seen(c) is the number of nodes seen by a node when it last visited the location C
- $\alpha \in [0,1]$ is a constant
  - **When α is large, places nearby are prefered**
  - **When α is small, popular locations are prefered**
- **Note**: A popular location need not be far away

The map in SWIM model



Alpha (α)	Same
Waiting time	Different for each movement
Speed	Same
Neighborhood Radius	Same

Motivation

- How to decide SWIM parameters using real contact traces?
  - Alpha value and neighborhood area

Motivation

- How to decide SWIM parameters using real contact traces?
  - Alpha value and neighborhood area
- The pairwise contact probabilities obtained from the real traces are used to tune the parameters of the SWIM mobility model.
- The traces and the SWIM model are compared in terms of
  - contact durations,
  - inter-contact times and,
  - number of pairwise contacts.

Deciding SWIM Parameters Based on Traces

(INFOCOM 2005 - 2006 & CAMBRIDGE 2005)

TUHH Technische Universität Hamburg

Calculating pairwise contact probability

- The calculation starts with counting the number of pairwise contacts between each node pair
- Let us consider a scenario with 10 mobile nodes
- We calculate a matrix A which has the number of pairwise contacts between each node pair

A =

	1	2	3	4	5	6	7	8	9	10
1	0	**13**	**23**	**20**	**26**	**40**	**18**	**42**	**35**	**10**
2	13	0	**27**	**20**	**15**	**15**	**13**	**12**	**24**	**17**
3	23	27	0	**37**	**25**	**11**	**10**	**24**	**21**	**30**
4	20	20	37	0	**16**	**27**	**22**	**38**	**32**	**41**
5	26	15	25	16	0	**18**	**16**	**32**	**42**	**22**
6	40	15	11	27	18	0	**27**	**37**	**26**	**21**
7	18	13	10	22	16	27	0	**28**	**26**	**24**
8	42	12	24	38	32	37	28	0	**23**	**39**
9	35	24	21	32	42	26	26	23	0	**25**
10	10	17	30	41	22	21	24	39	25	0

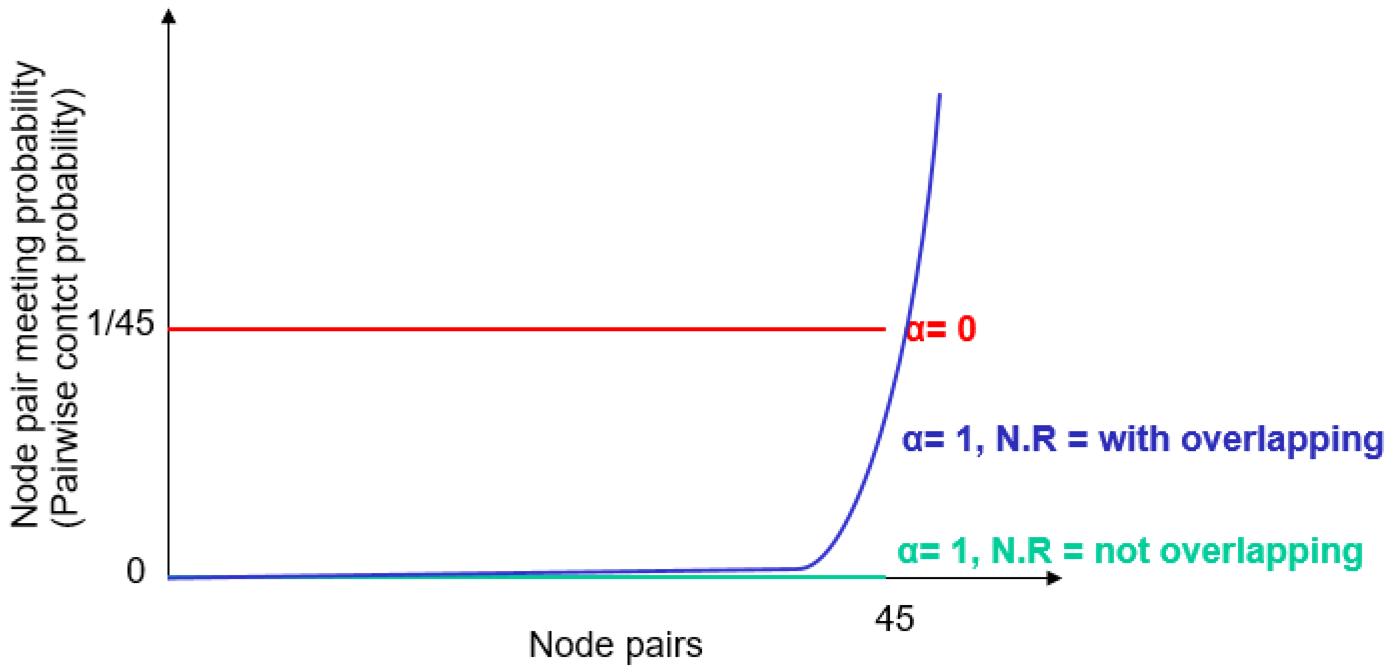
Calculating pairwise contact probability

- From the matrix A, the pairwise contact probabilities are calculated by normalising each element in the matrix A by the sum of the upper triangle elements of the matrix A
- Only the upper (or lower) triangle is chosen because the matrix A is symmetric

	0	**0.011712**	**0.020721**	**0.018018**	**0.023423**	**0.036036**	**0.016216**	**0.037838**	**0.031532**	**0.009009**
	0.011712	0	**0.024324**	**0.018018**	**0.013514**	**0.013514**	**0.011712**	**0.010811**	**0.021622**	**0.015315**
	0.020721	0.024324	0	**0.033333**	**0.022523**	**0.00991**	**0.009009**	**0.021622**	**0.018919**	**0.027027**
	0.018018	0.018018	0.033333	0	**0.014414**	**0.024324**	**0.01982**	**0.034234**	**0.028829**	**0.036937**
	0.023423	0.013514	0.022523	0.014414	0	**0.016216**	**0.014414**	**0.028829**	**0.037838**	**0.01982**
P=	0.036036	0.013514	0.00991	0.024324	0.016216	0	**0.024324**	**0.033333**	**0.023423**	**0.018919**
	0.016216	0.011712	0.009009	0.01982	0.014414	0.024324	0	**0.025225**	**0.023423**	**0.021622**
	0.037838	0.010811	0.021622	0.034234	0.028829	0.033333	0.025225	0	**0.020721**	**0.035135**
	0.031532	0.021622	0.018919	0.028829	0.037838	0.023423	0.023423	0.020721	0	**0.022523**
	0.009009	0.015315	0.027027	0.036937	0.01982	0.018919	0.021622	0.035135	0.022523	0

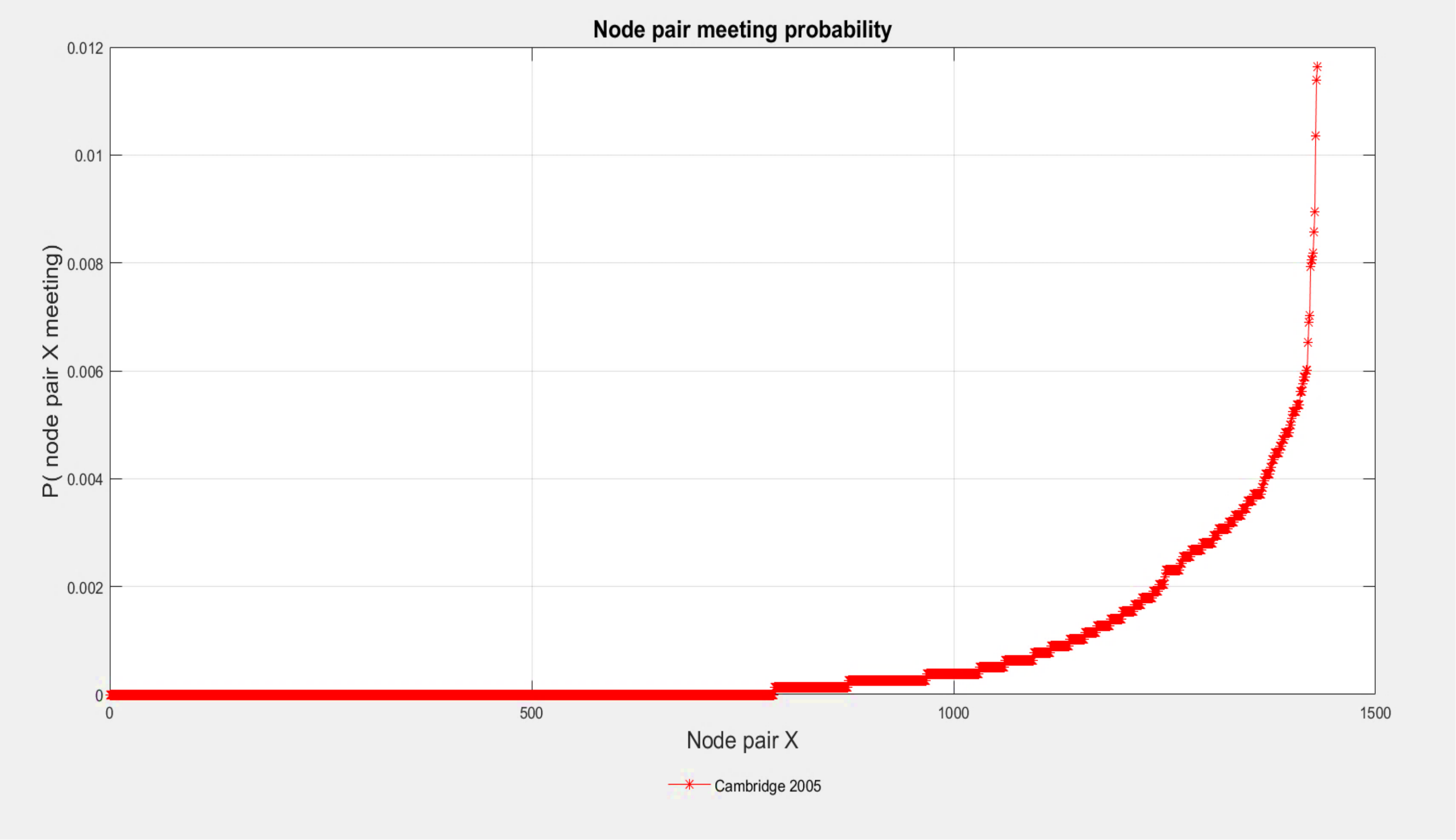
Calculating pairwise contact probability

- The upper (or lower) triangle elements of the matrix P are sorted in ascending order to get a 1 x $\frac{N*(N-1)}{2}$ matrix called as $\mathbf{P}_{\mathbf{pair}}$



N.R : Neighborhood Radius

Deciding the α value for simulating the SWIM model



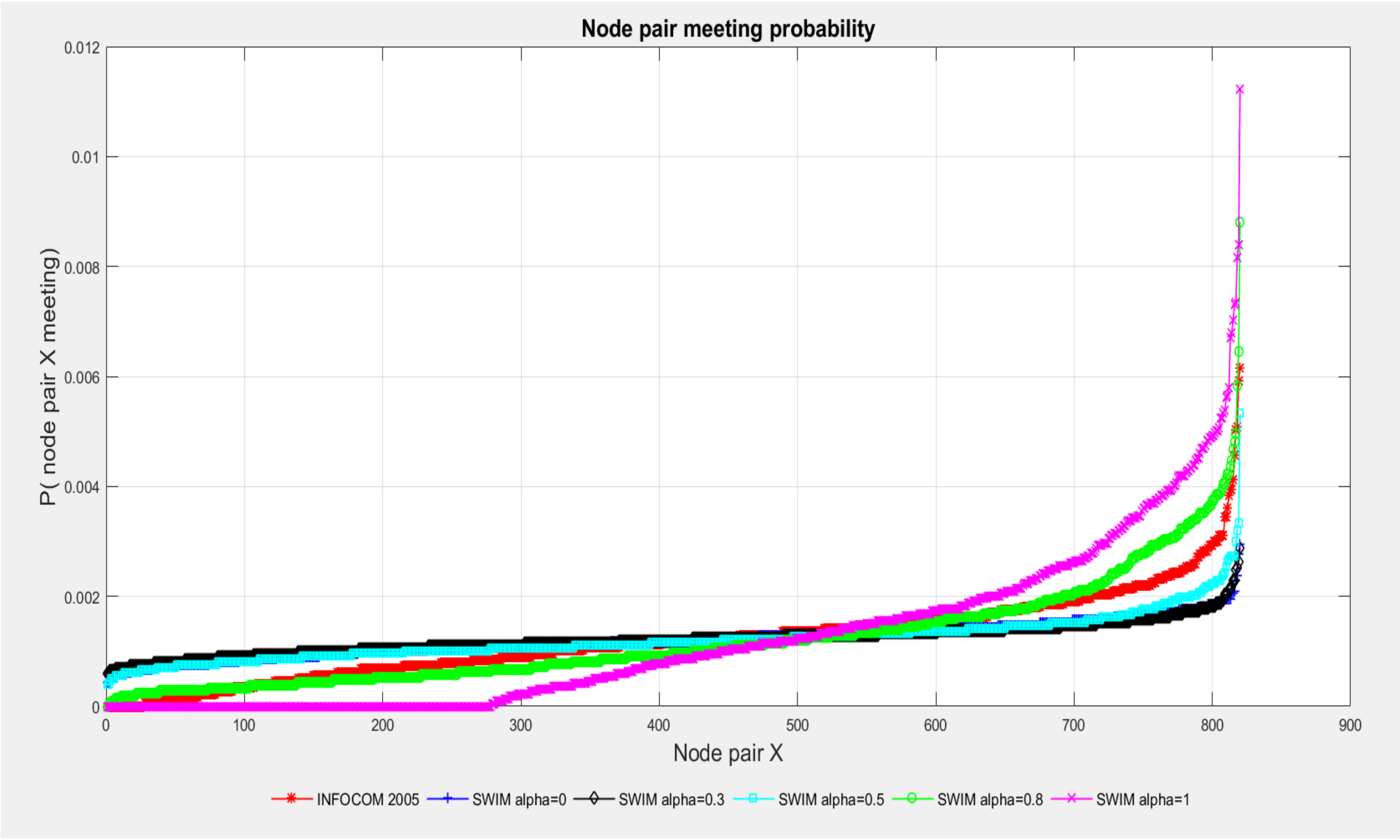
Node pair meeting probability P_{pair} of Cambridge 2005

✓ Simulation Results

Simulation parameters

Parameter	INFOCOM 2005	INFOCOM 2006	Cambridge 2005
Simulation area	300m x 300m	2000m x 2000m	2000m x 2000m
Number of nodes	41 mobile nodes	78 mobile nodes + 20 stationary nodes	36 mobile nodes + 18 stationary nodes which include 4 long range, 14 short range nodes
Number of locations	48	40 (stationary nodes must be placed in the locations)	38 (stationary nodes must be placed in the locations)
Mobility speed (m/s)	equal to the distance in metres		
Radio range	11 m	Mobile nodes: 30 m Stationary: 100 m	Mobile nodes: 11 m Short range nodes: 11 m Long range nodes: 22 m
Beacon Interval	100 seconds	120 seconds	Mobile nodes: 10mins 4 long range nodes: 2mins 2 short range nodes: 6mins 12 short range nodes: 10mins
Neighbourhood radius	100m	200m	500m
Swim Alpha (α)	**0.8**	**0.75**	**0.9**
Waiting time	exponential(500seconds)	exponential(6 minutes)	exponential(30 minutes)
Simulation time	4 days	5 days	11 days

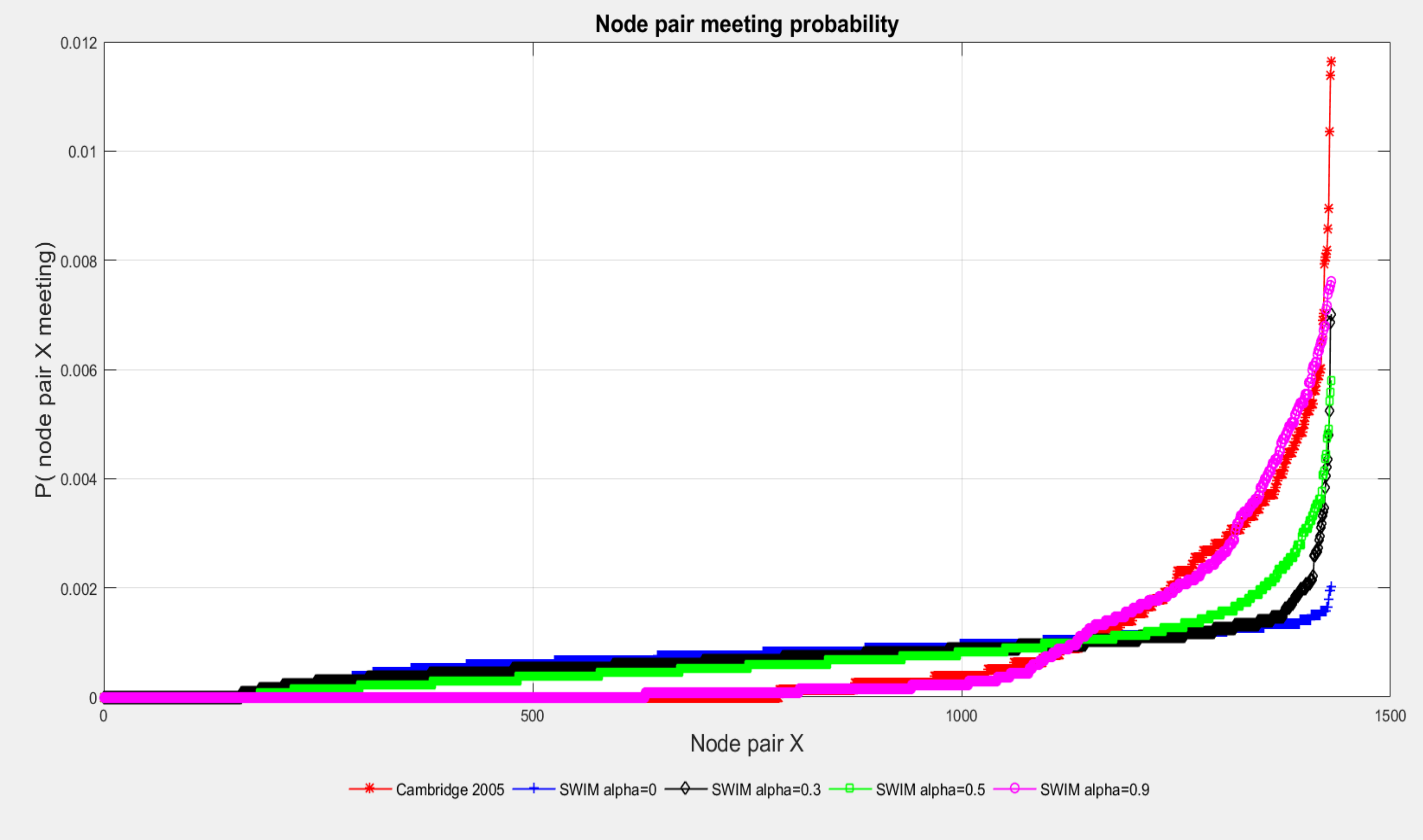
Simulation results



Node pair meeting probability P_{pair} for INFOCOM 2005 and SWIM model for various alpha values

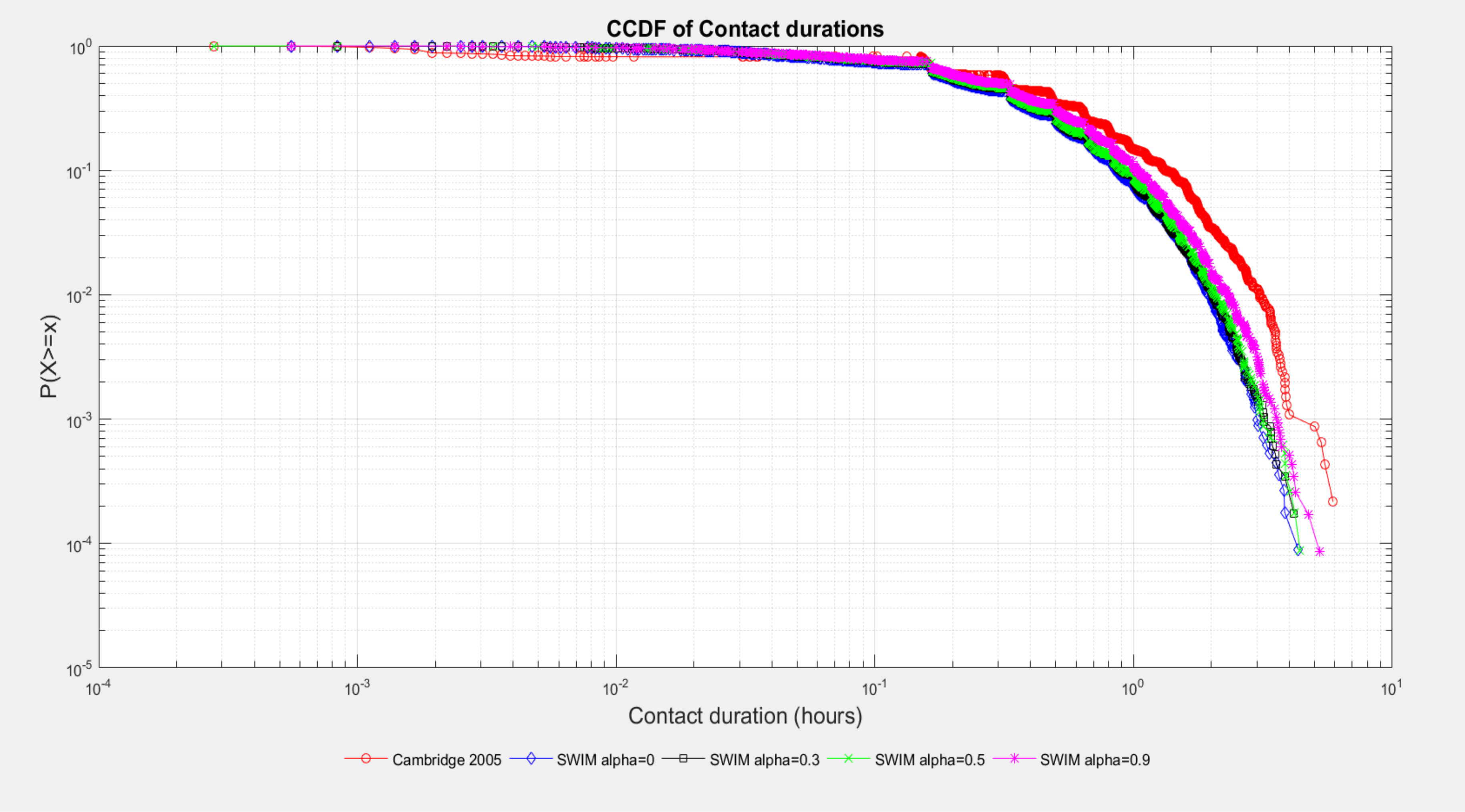
Simulation results

TUHH
Technische Universität Hamburg



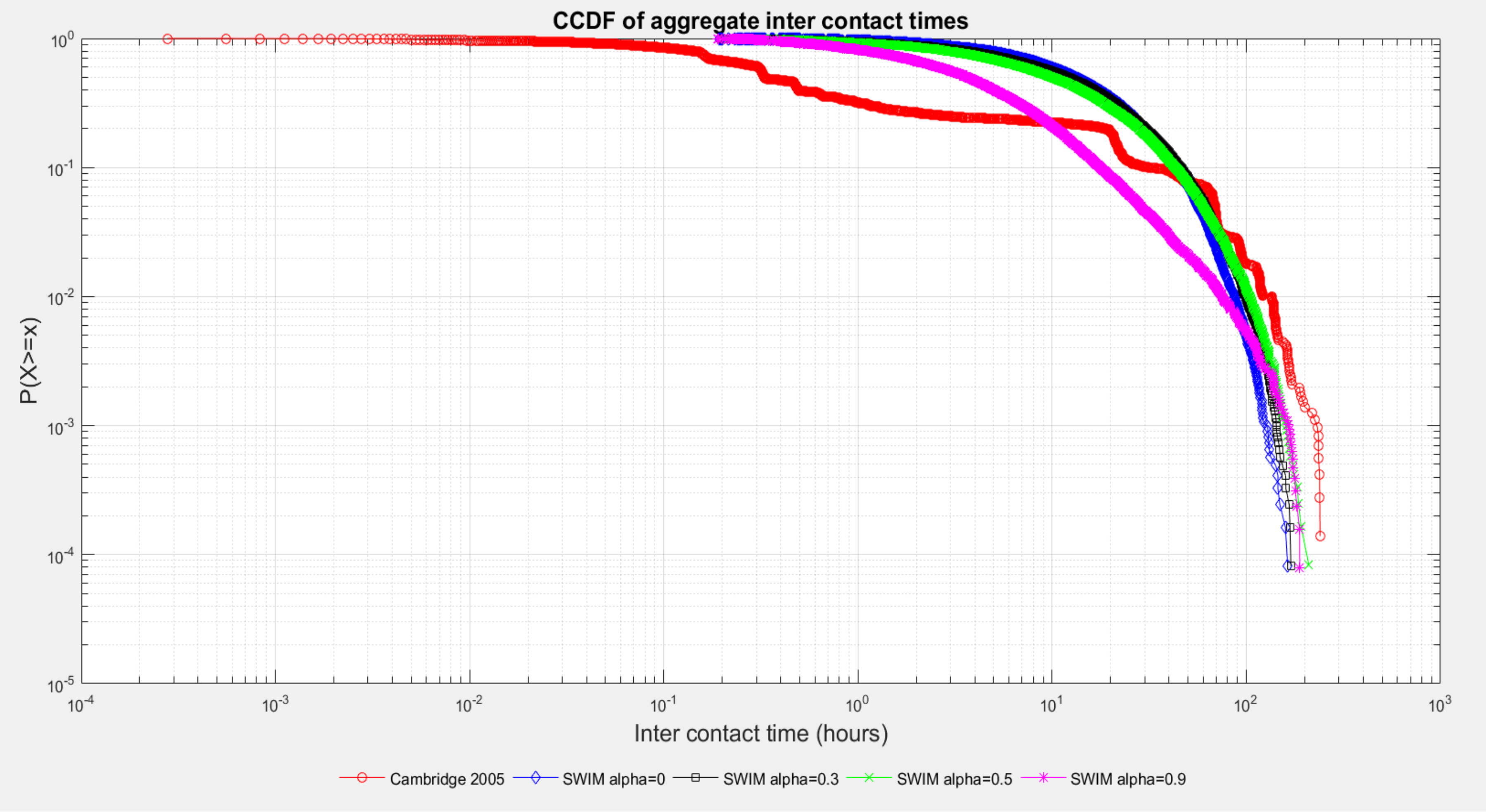
Node pair meeting probability P_{pair} for Cambridge 2005 and SWIM model for various alpha values

Comparison of the traces and the SWIM model



Inter-contact times of **Cambridge 2005** compared with SWIM model (log-log axis)

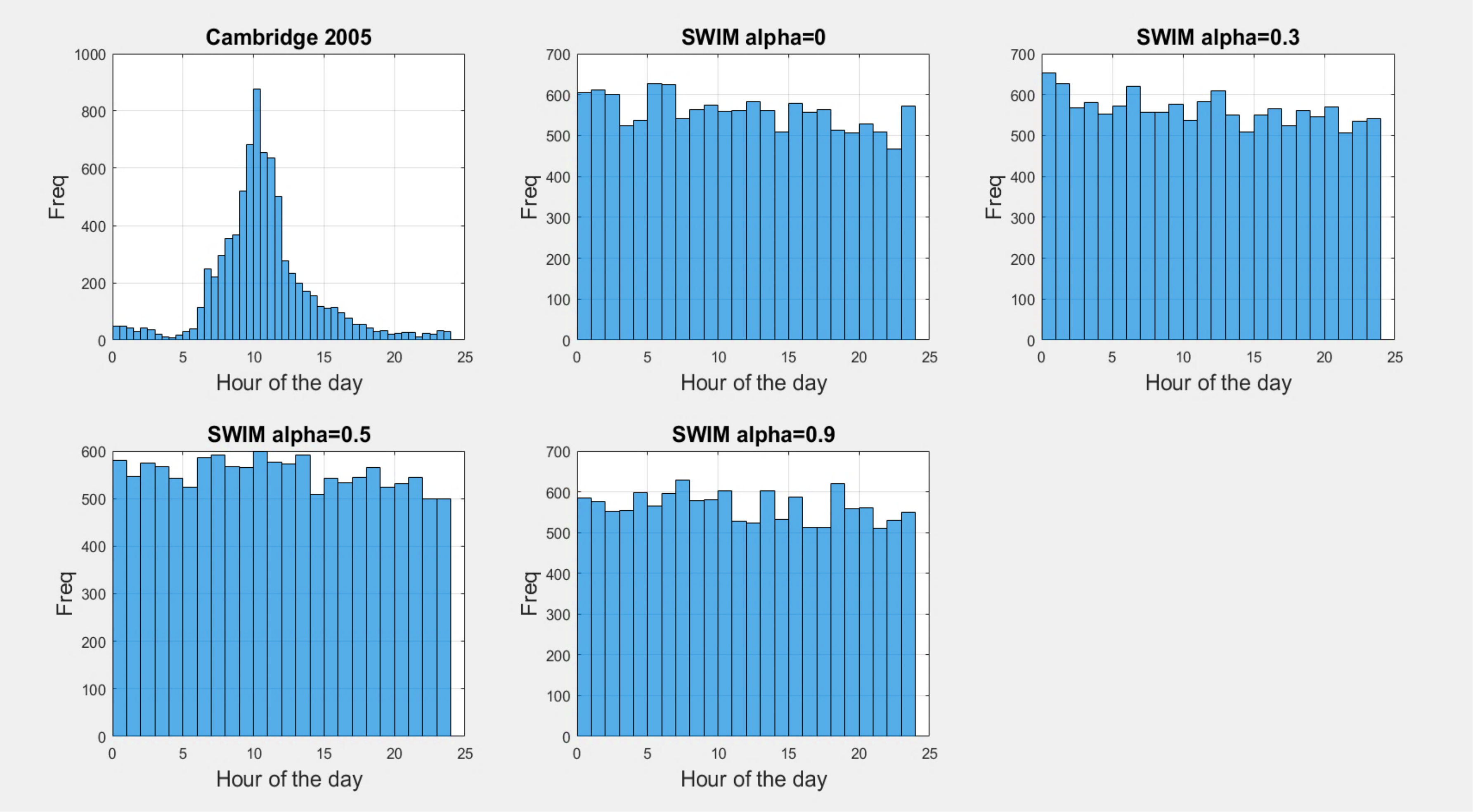
Comparison of the traces and the SWIM model



Contact durations of **Cambridge 2005** compared with SWIM model (log-log axis) [3] [4]

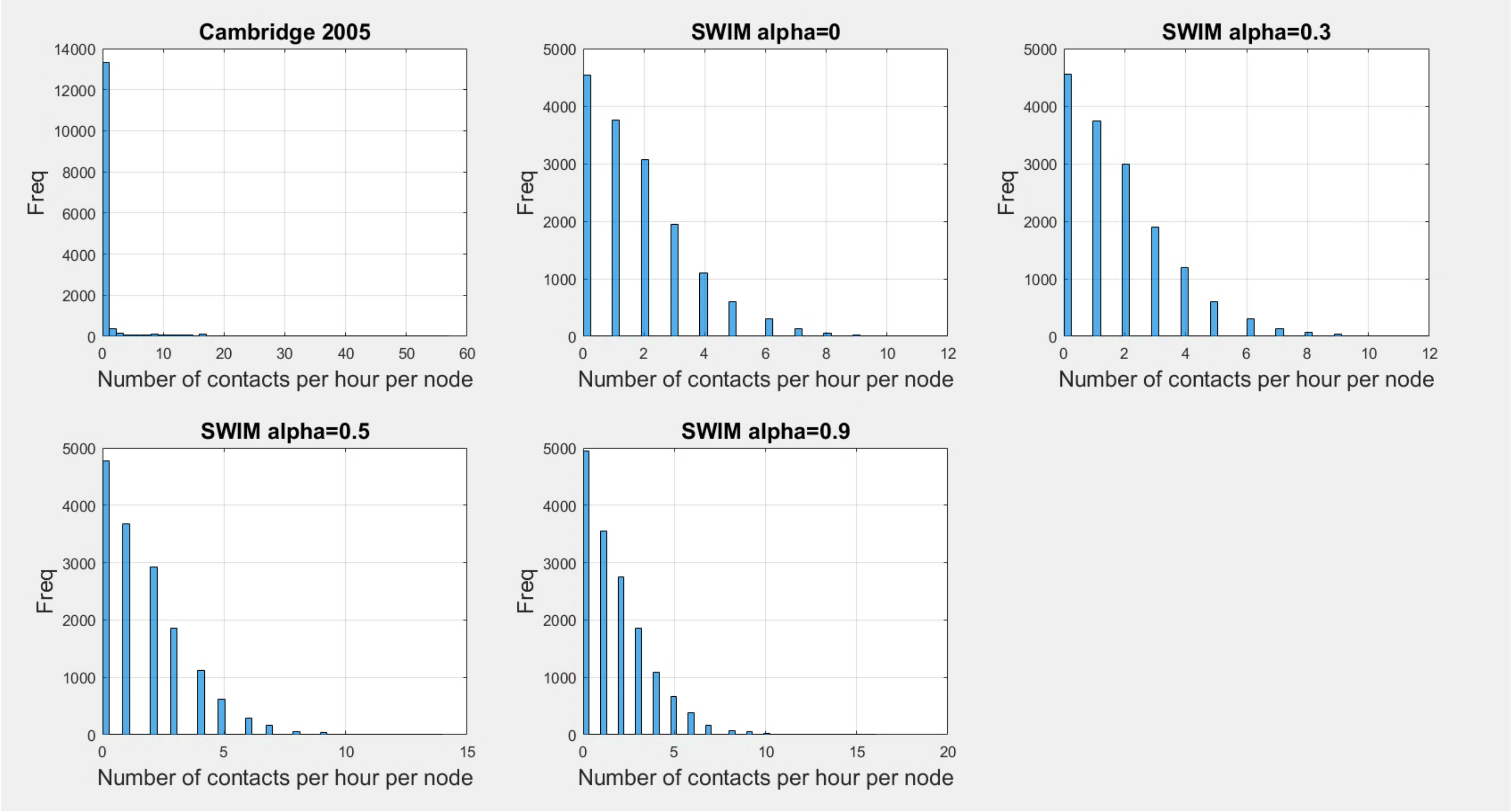
Comparison of the traces and the SWIM model

TUHH
Technische Universität Hamburg



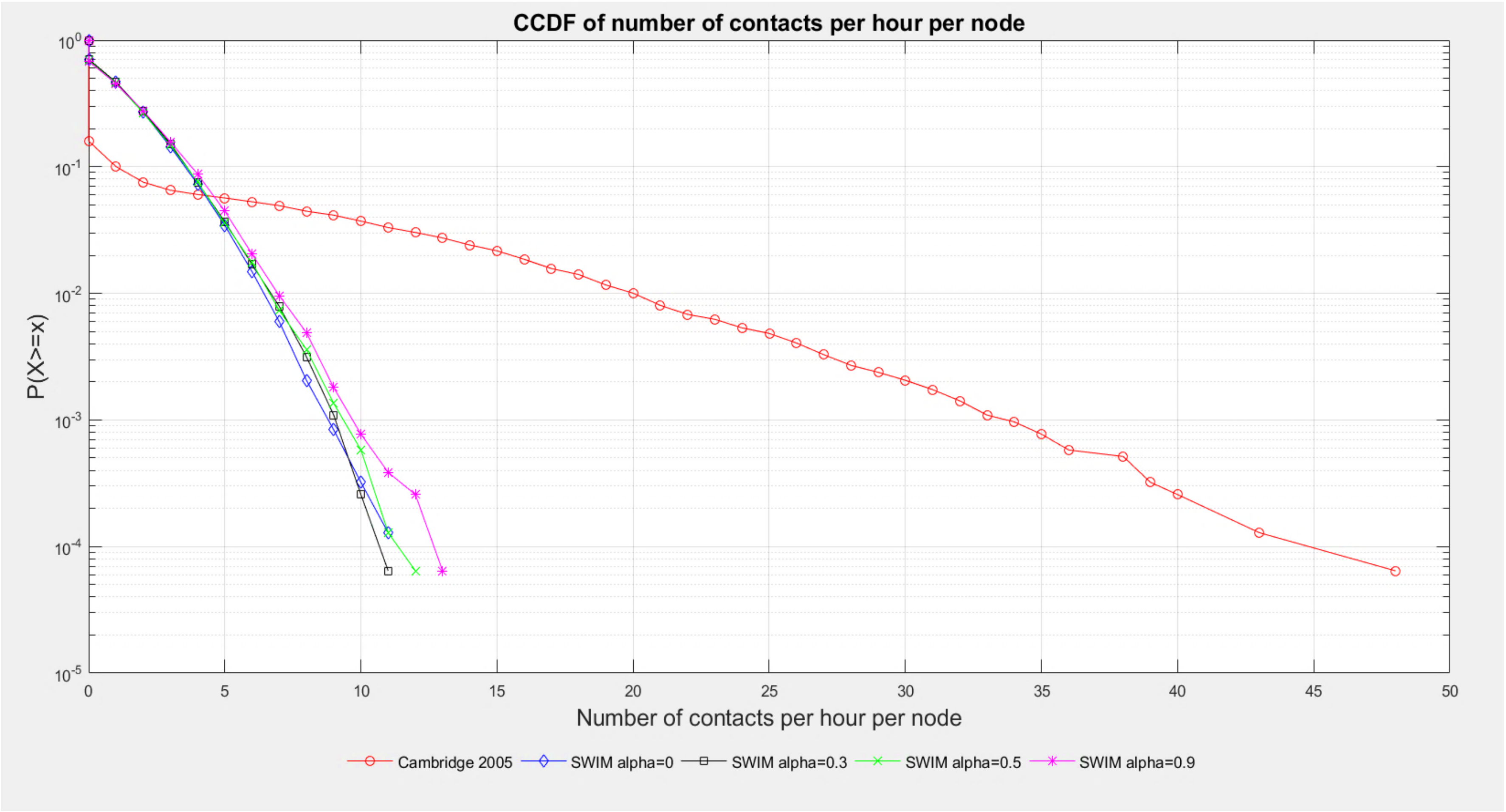
Number of overall pairwise contacts based on hour of the day for **Cambridge 2005** and SWIM

Comparison of the traces and the SWIM model



Aggregate number of contacts per hour per node for **Cambridge 2005** compared with SWIM

Comparison of the traces and the SWIM model



CCDF of number of contacts per hour per node for **Cambridge 2005** compared with SWIM (Y axis in log)

✓ Conclusion and Future Work

TUHH
Technische Universität Hamburg

Conclusions and future works

- Pairwise contact probabilities can be used to map the location preferences in SWIM model
- The traces are very heterogeneous in nature and the parameters in SWIM need heterogeneity
- Future work
  - Representing day and night times in SWIM model
  - A mathematical model for predicting alpha value from P_{pair}
  - Fine divisions in neighbourhood radius for fine tuning P_{pair}

References

[1] A. Foerster et al, A Novel Data Dissemination Model for Organic Data Flows, MONAMI 2015, September 1618, 2015, Santander, Spain

[2] A. Mei and J. Stefa, SWIM: A Simple Model to Generate Small Mobile Worlds, IEEE INFOCOM, 2009.

[3] T. Camp and J. Boleng and V. Davies, A Survey of Mobility Models for Ad Hoc Network Research, Wireless Communications and Mobile Computing, August 2002

[4] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, CRAWDAD data set cambridge/haggle (v. 2006-01-31), http://crawdad.org/cambridge/haggle/20090529/

[5] https://github.com/ComNets-Bremen/OPS

[6] A.Udugama, B.Khalilov, A.b.Muslim, A.Foerstrer, K.Kuladinithi, Implementation of the SWIM Mobility Model in OMNET++ Proc. of the 3rd OMNeT++ Community Summit, Brno University of Technology – Czech Republic - September 15-16, 2016

[7] K.Garg, S. Giordano, M. Jazayeri, How Well Does Your Encounter- Based Application Disseminate Information?, Conference: The 14th IFIP Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net) - 2015

Thank you! Questions?

BACKUP

- distance(h,C)= $\frac{1}{(1+k\|x-y\|)^2}$, k=0.05, $\| x - y \|$ is the euclidean distance between the node and the centre of the location C

Conclusions and Future Works

- Different values: can be used to represent day and night time mobility behavior. Nodes can be divided into clusters representing a particular behavior with regard to (e.g., behavior of students vs teachers)

- Different neighborhood radius: multiple sectors with different radii making each sector having a different priority of visiting. This allows the possibility of expanding the neighborhood into fine divisions (e.g. Kitchen and lab area in Cambridge traces)

- A mathematical model: This work is starting point for using pairwise contact probabilities for parameterizing the SWIM model. Therefore, a mathematical model to predict the and other parameters of SWIM model based on existing properties of real life traces (e.g. pairwise contact probability) is a part of future work. The approach presented in this paper is purely graphical in guessing the for simulating the SWIM model.

Small Worlds In Motion (SWIM)

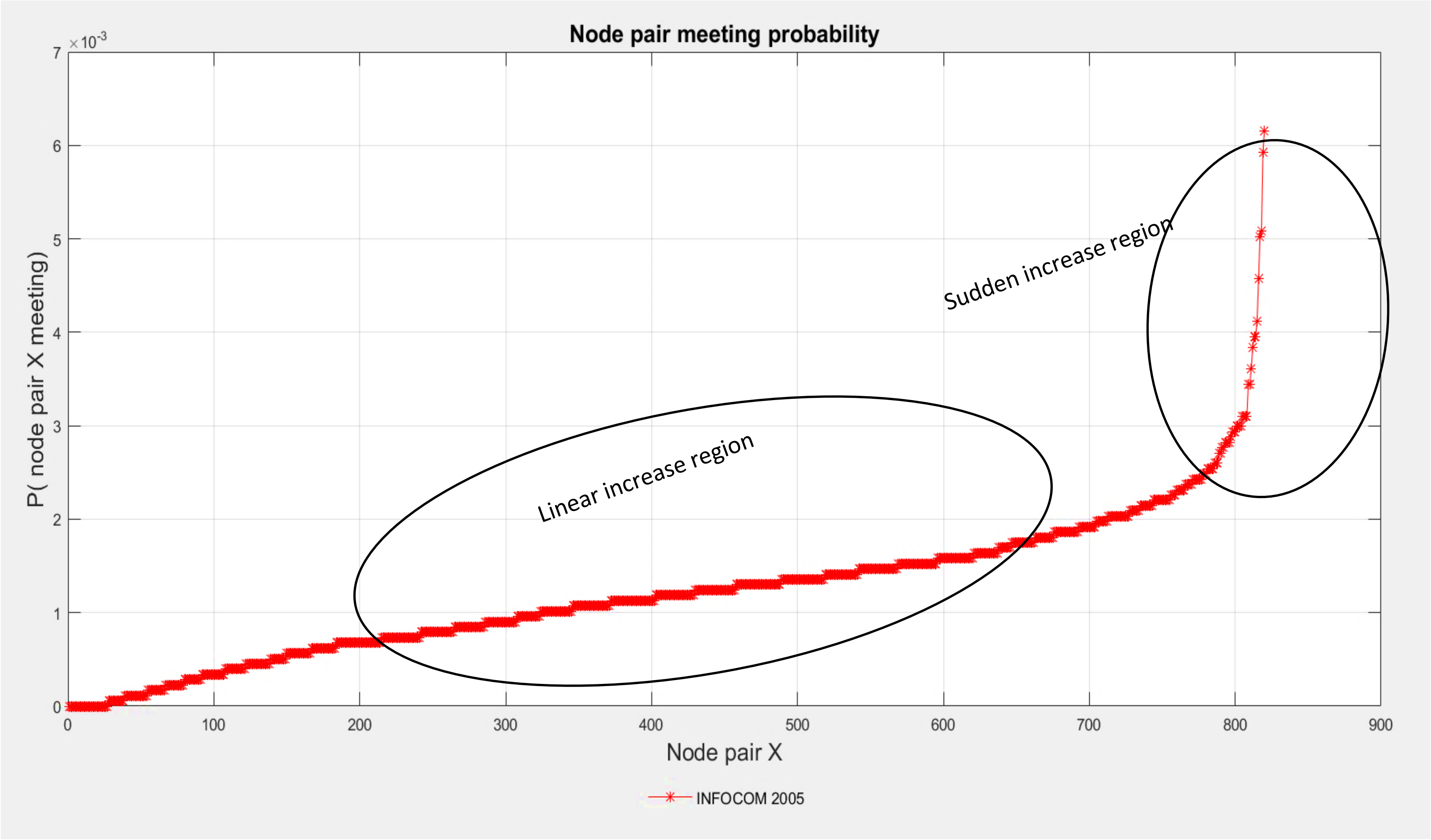
- Each node maintains a set of weights for all the locations
- Each node remembers only the weights of the locations it has visited
- Locations which are not visited will have weight=0 initially
- **Speed = distance between the locations**

- The decision by each node depends on all the past decisions made by itself and other nodes
- The current decision of a node affects the future decision of all the other nodes
- The node can return to the home location with a certain probability

Calculating pairwise contact probability

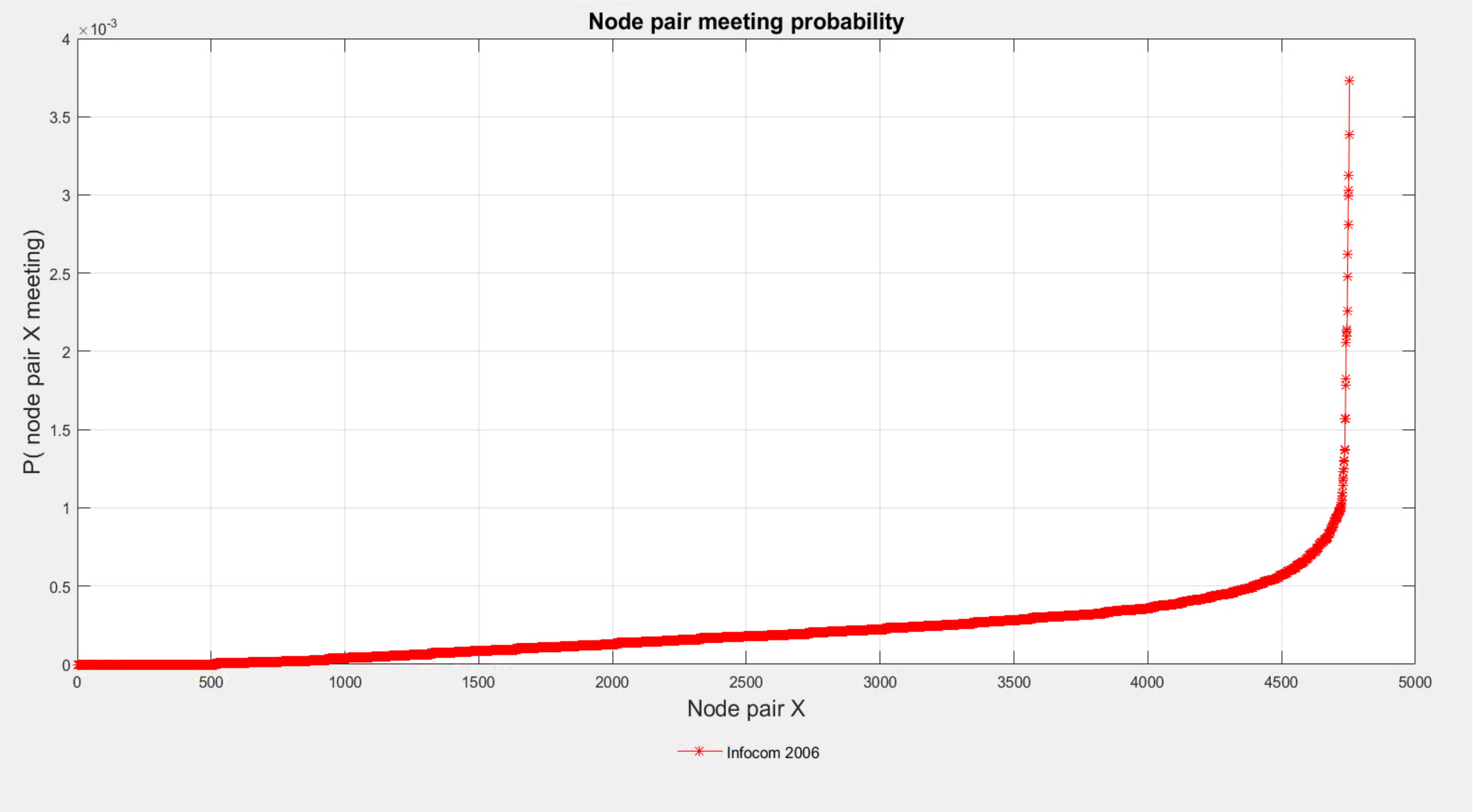
- The upper (or lower) triangle elements of the matrix P are sorted in ascending order to get a 1 x $\frac{N*(N-1)}{2}$ matrix called as $\mathbf{P_{pair}}$
- P_{pair} is a **modified form** of the **pairwise contact probability matrix P** specially for the SWIM model, for the following reasons
  - The SWIM model is a pure location based model
  - There is no control over node attractions
  - The pairwise contact probabilities cannot be used to directly program the network
  - There is no way to exactly match the pairwise contact probabilities between the traces and the SWIM model
  - Thus, there is a need to see an **overall pattern** of the pairwise contact probabilities
- The P_{pair} **removes the identity** of the node pairs. It just gives the pattern in which the probabilities increase.
  - This is very useful for SWIM model as the identities of node pairs cannot be matched with the matrix P

Deciding the α value for simulating the SWIM model



Node pair meeting probability P_{pair} of INFOCOM 2005

Deciding the α value for simulating the SWIM model

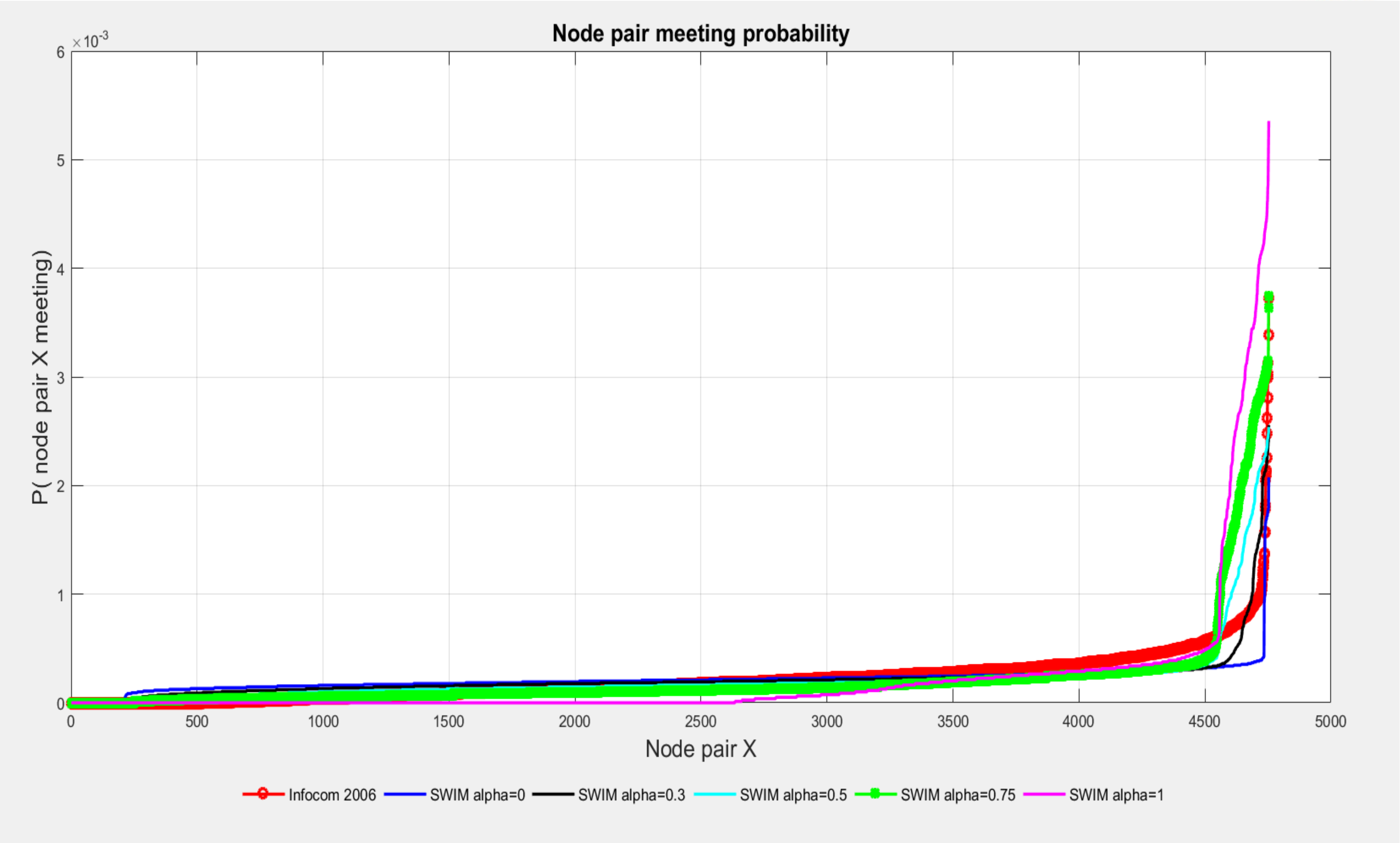


Node pair meeting probability P_{pair} of INFOCOM 2006

Deciding the α value for simulating the SWIM model

- **Thumb rule:** The more linear the plot increases, lower is the alpha value
- A **high alpha value** (say 0.9) will make sure that a node visits nearby locations 90% of the time and the other locations outside the neighbourhood are visited only 10% of the time.
  - For a alpha=1, a lot of node pairs will never meet making P_{pair} zero (or very low) for those node pairs
- A **low alpha value** allows a node to visit outside the neighbourhood and hence increasing the probability of meeting all nodes at popular locations.
  - For alpha=0, the nodes will only visit popular locations and hence increasing possibility of meeting all the nodes
- A fully linear increasing P_{pair} plot nearly parallel to the X axis needs an alpha value of zero to make sure that there is a chance to meet every node

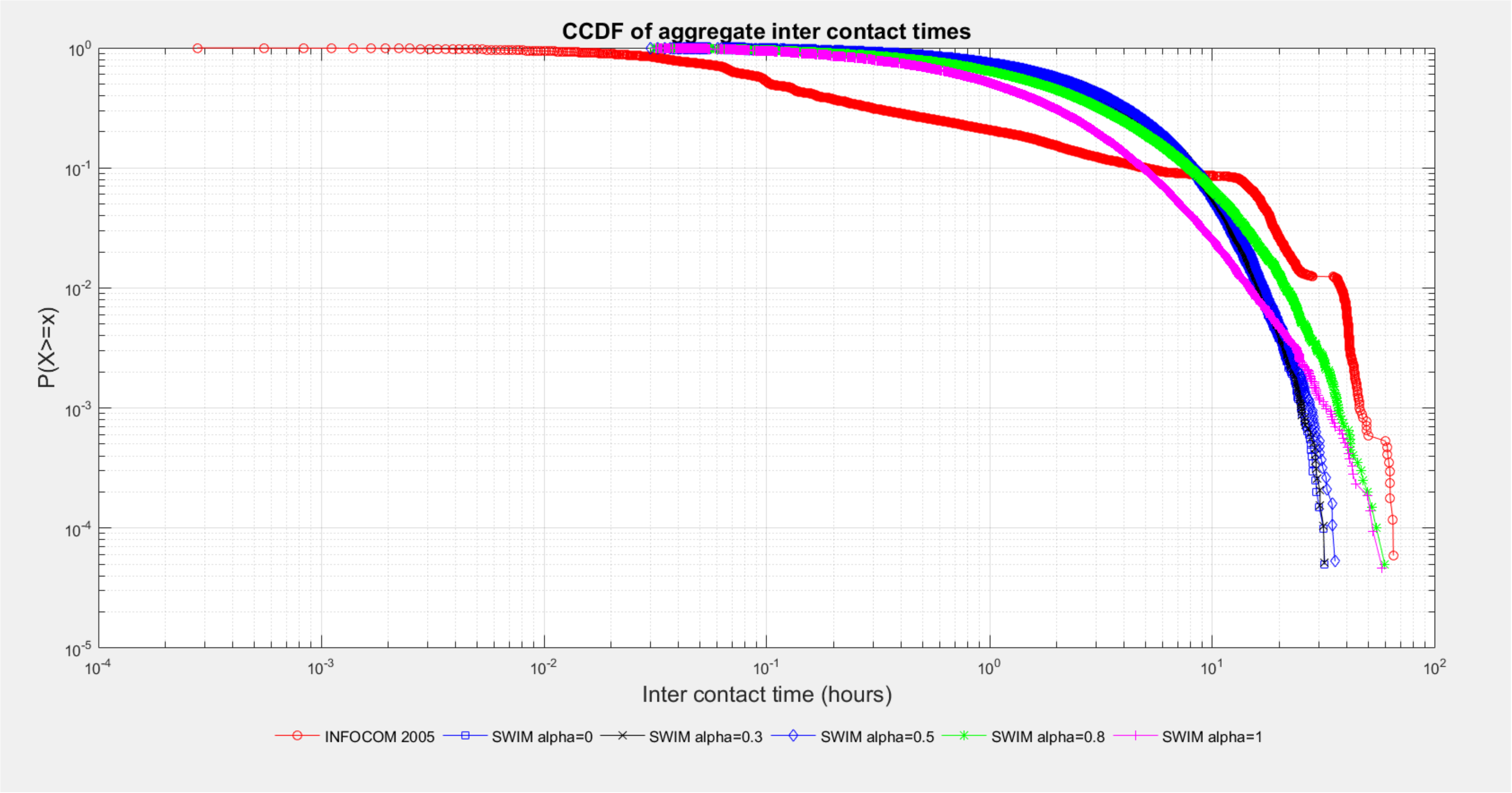
Simulation results



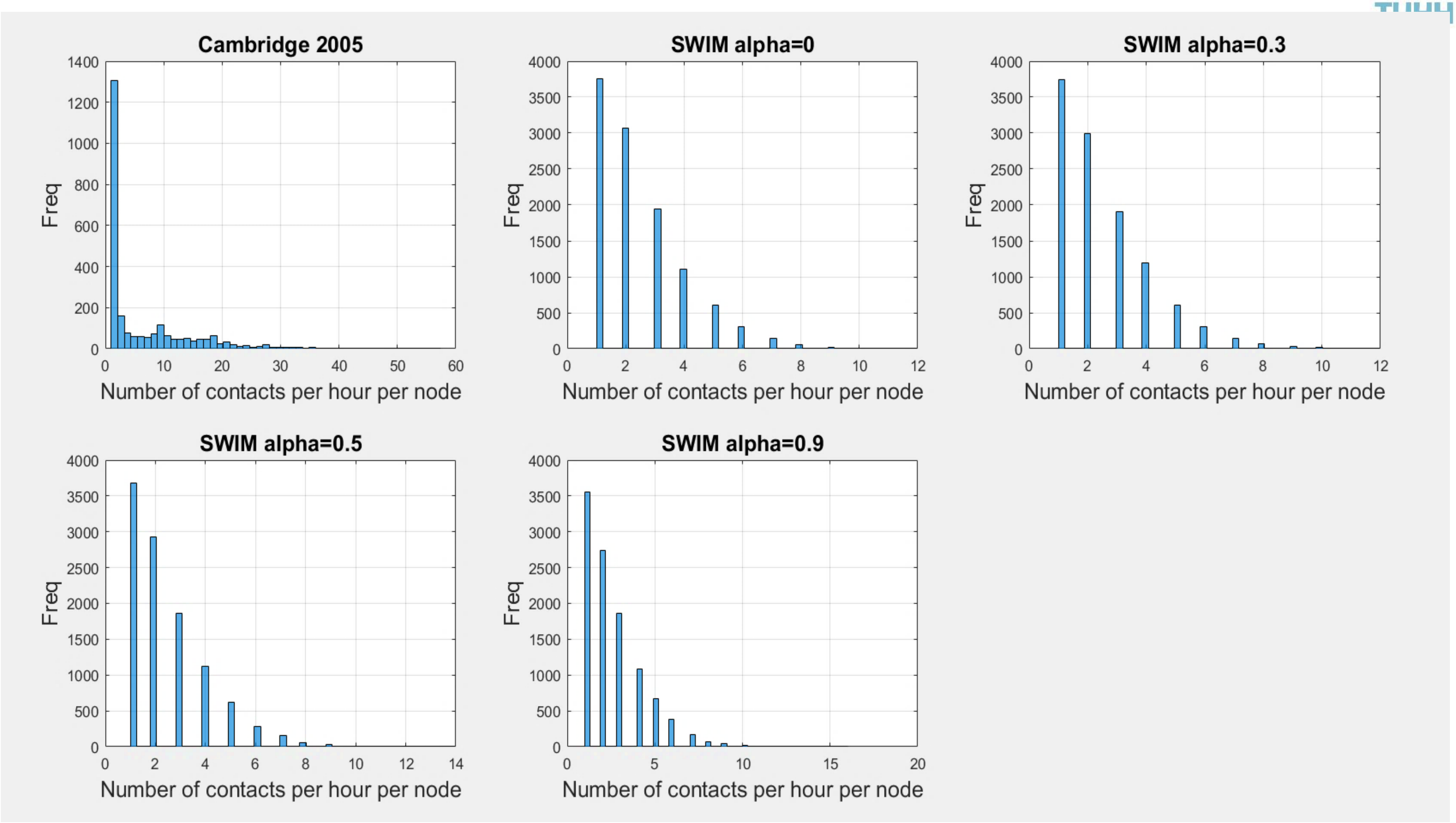
Node pair meeting probability P_{pair} for INFOCOM 2006 and SWIM model for various alpha values

Observations and inferences from P_{pair}

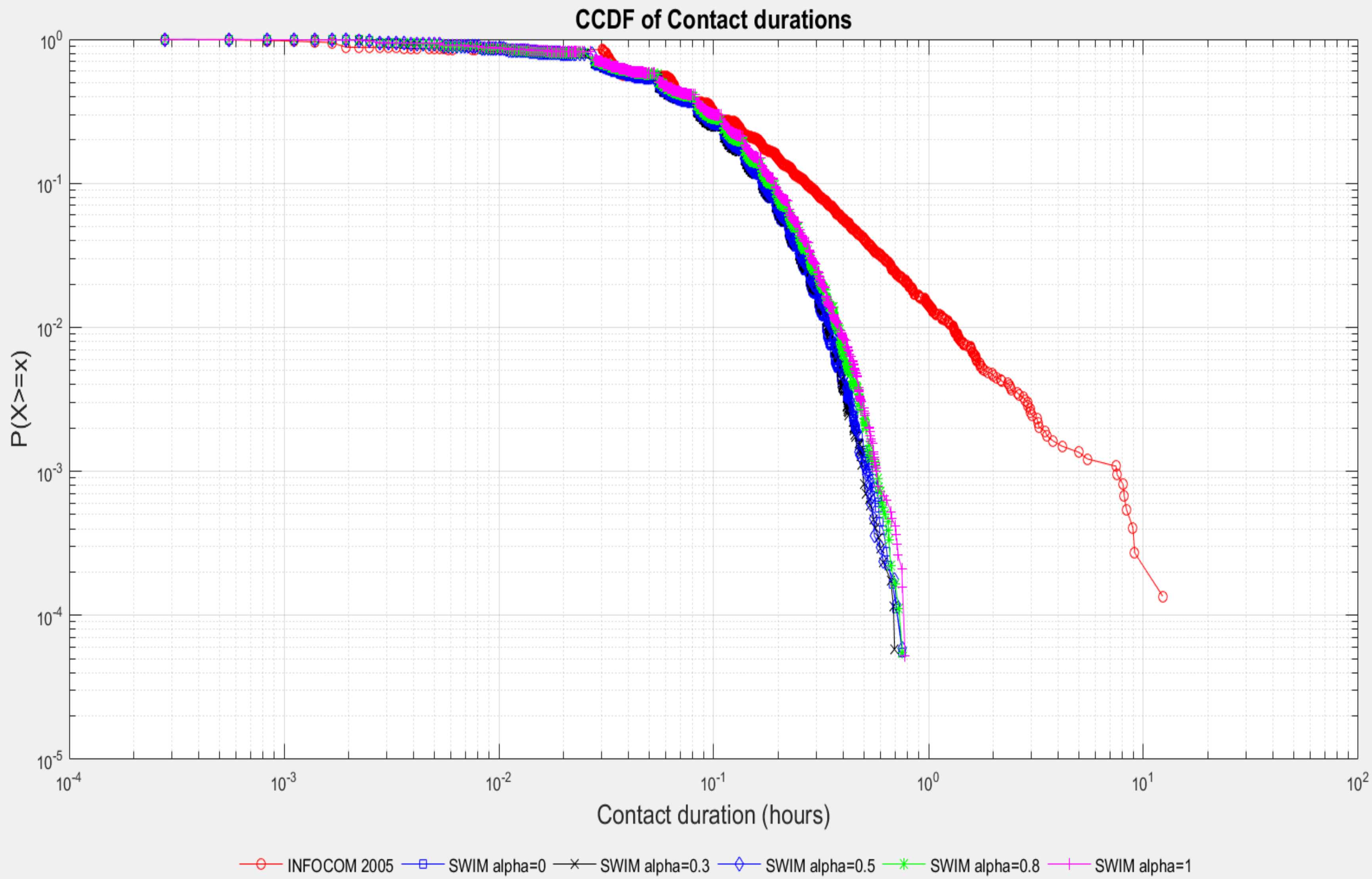
- Higher the alpha value, higher the curve reaches
- Lower the alpha value, more is the linearity in the curve
- Higher the end of the curve reaches, lower the start of the curve goes
  - This is because sum of all values of P_{pair} is 1
- When alpha is near 1, a lot of node pairs will never meet
  - The number of node pairs which will never meet depend on the neighbourhood radius, communication range, number of locations within the neighbourhood and, how near alpha is to 1
- When alpha is near 0, there is a possibility that all the nodes will meet with the same probability in the long run
  - The number of node pairs for which the curve is linear depends on the communication range and how near the alpha value is to 0

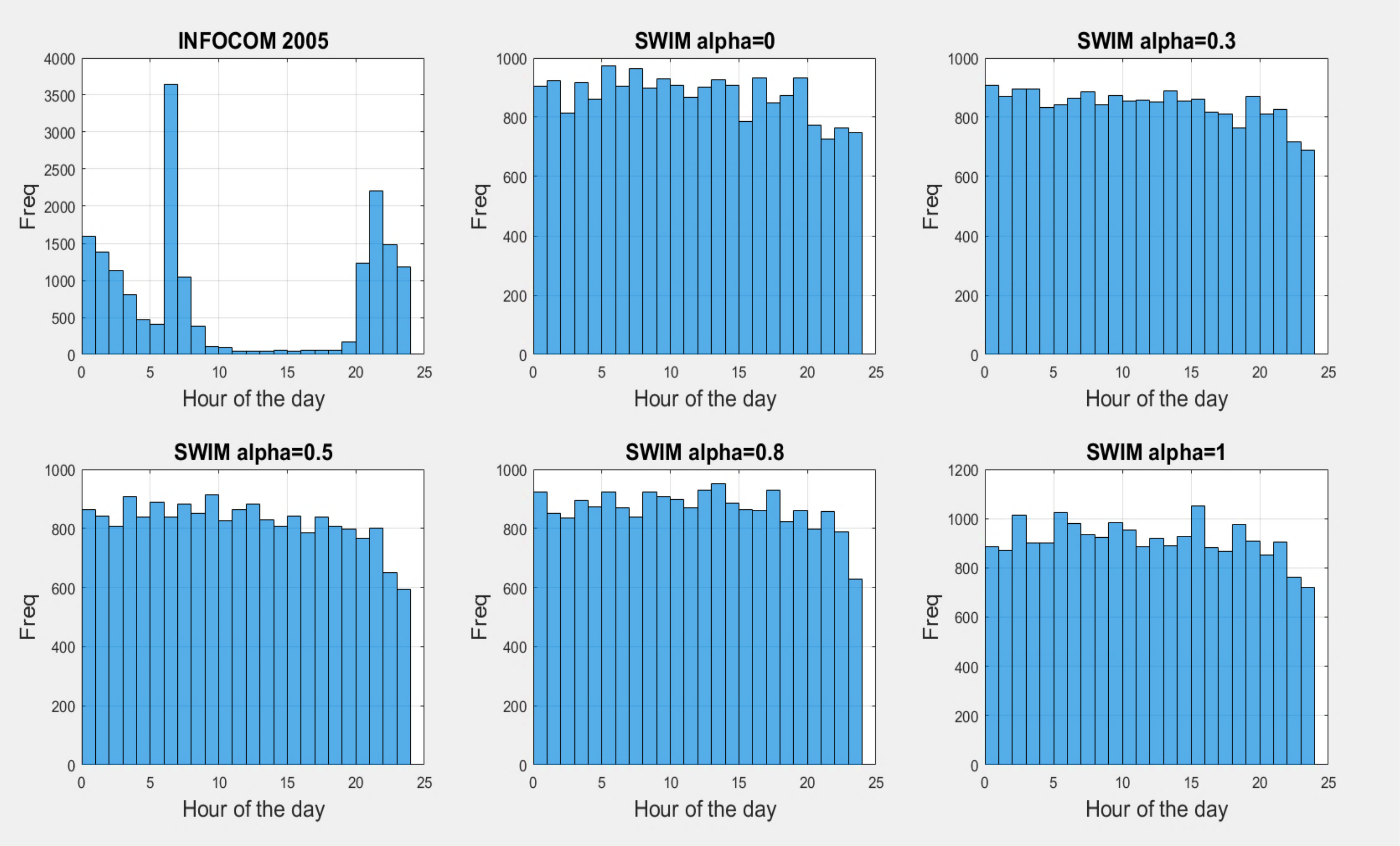

CCDF of aggregate inter contact times
P(X>=x)
Inter contact time (hours)
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
INFOCOM 2005
SWIM alpha=0
SWIM alpha=0.3
SWIM alpha=0.5
SWIM alpha=0.8
SWIM alpha=1

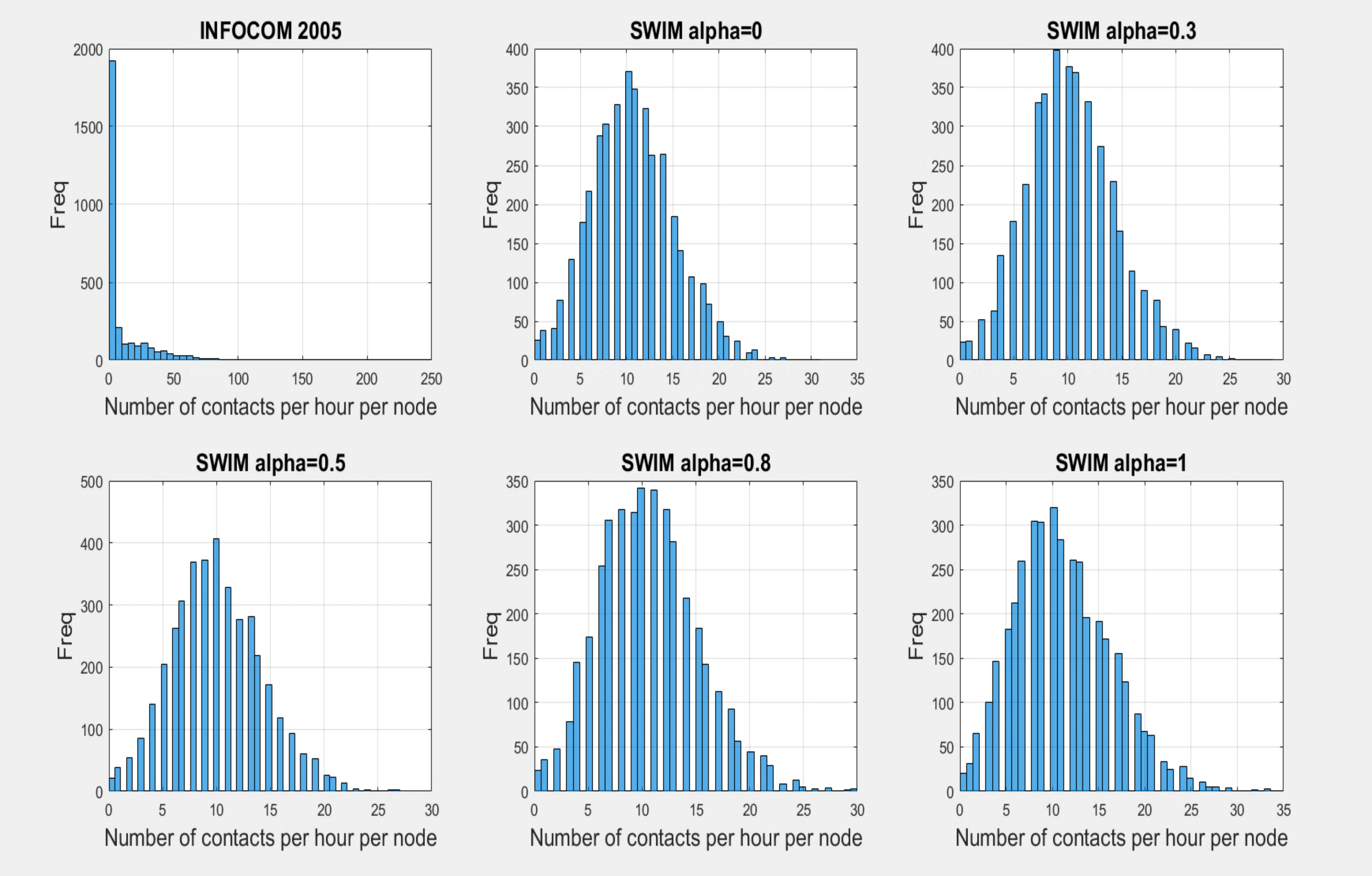
Comparison of the traces and the SWIM model

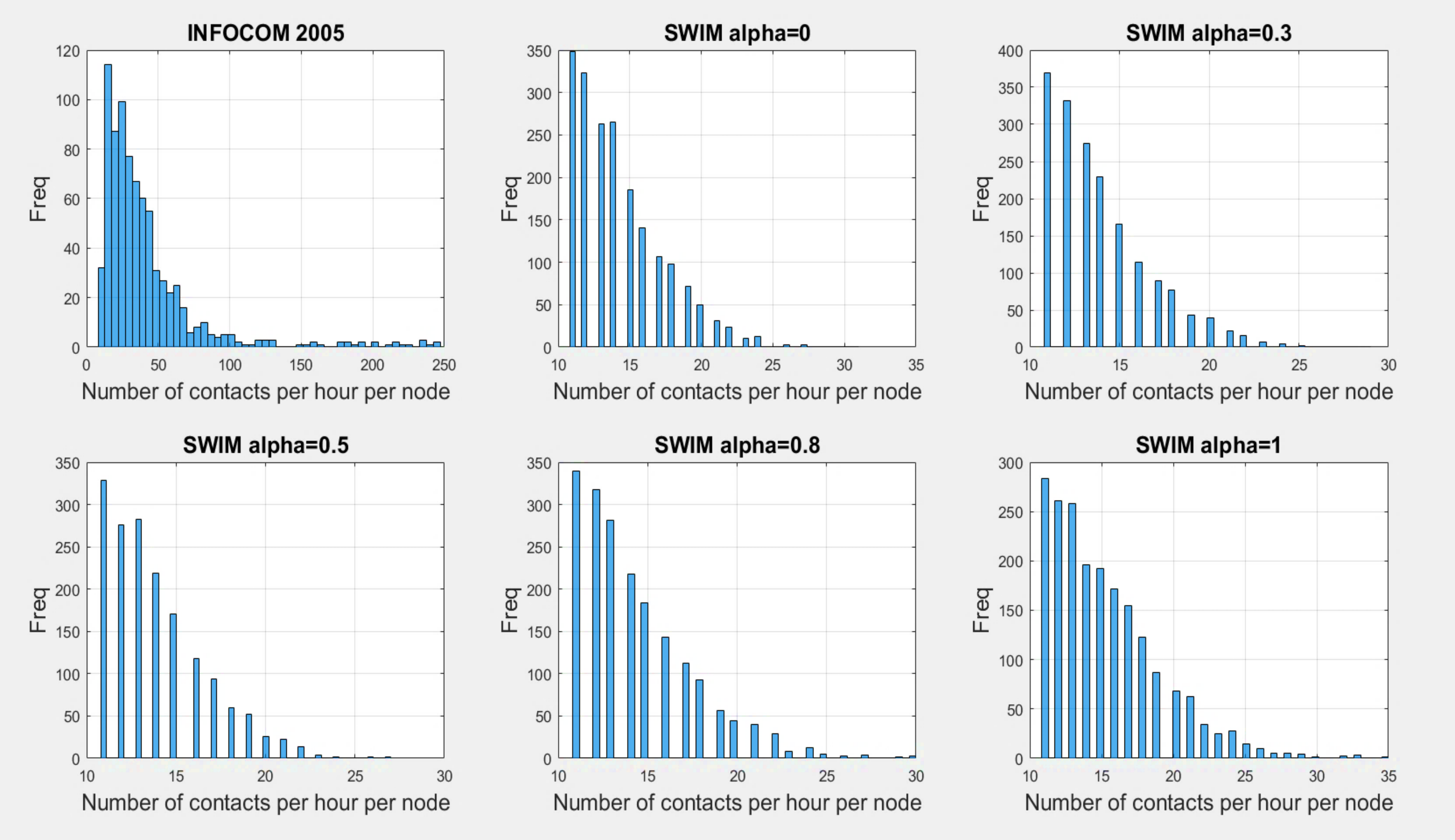


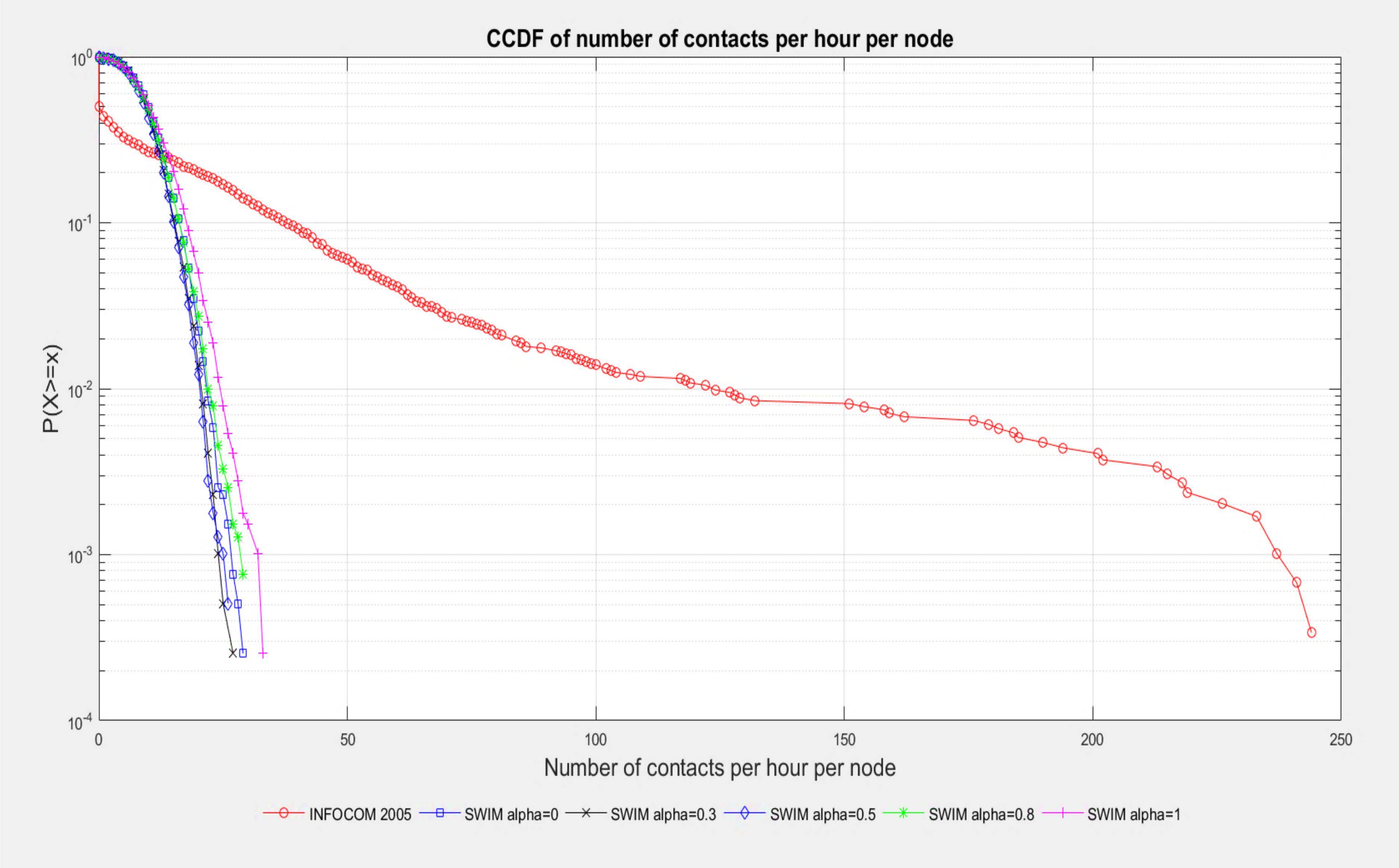
Truncated aggregate number of contacts per hour per node for **Cambridge 2005** compared with SWIM with lower bound of 1 contacts per hour per node

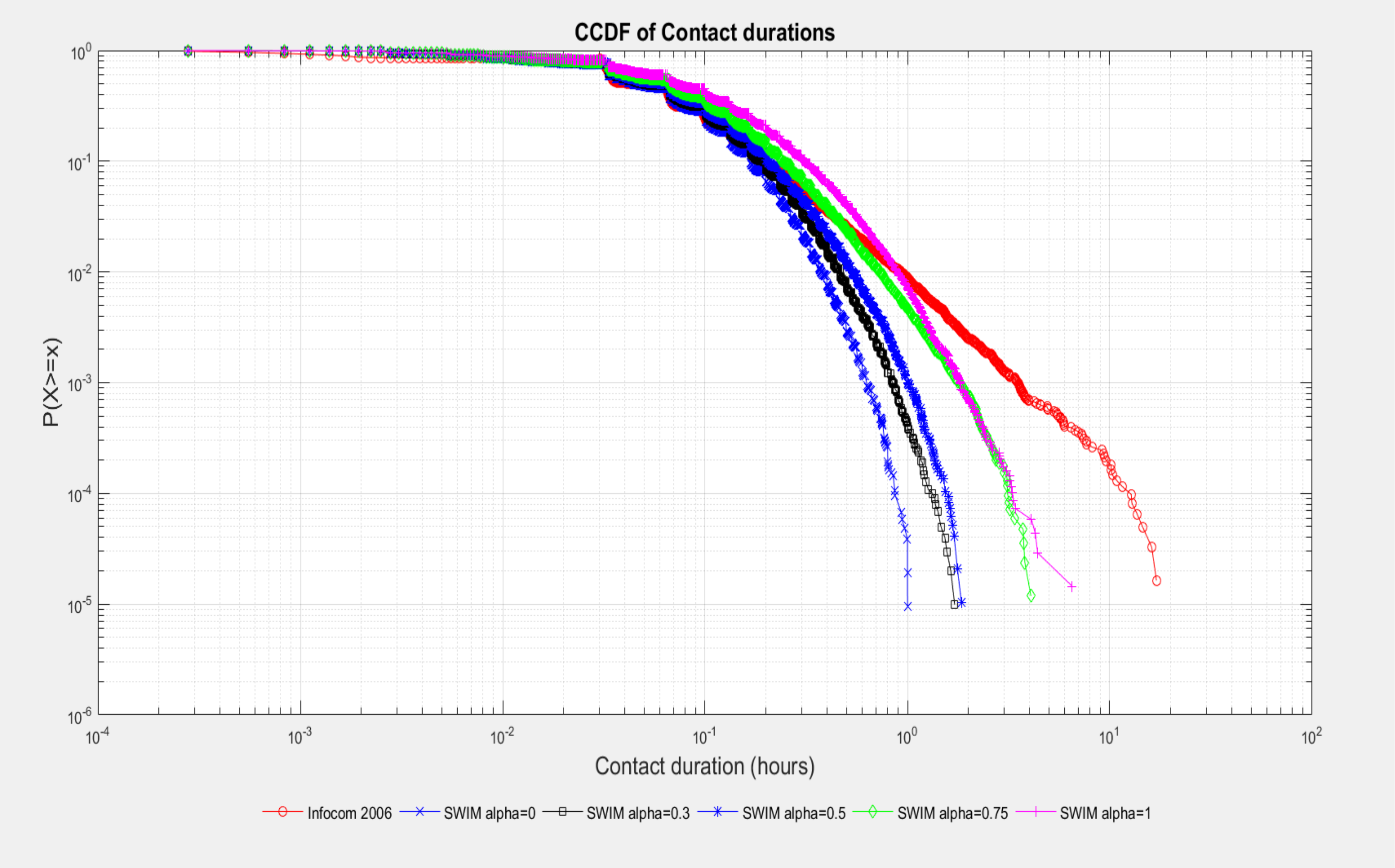

CCDF of Contact durations
P(X>=x)
Contact duration (hours)
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^0
10^1
10^2
INFOCOM 2005
SWIM alpha=0
SWIM alpha=0.3
SWIM alpha=0.5
SWIM alpha=0.8
SWIM alpha=1

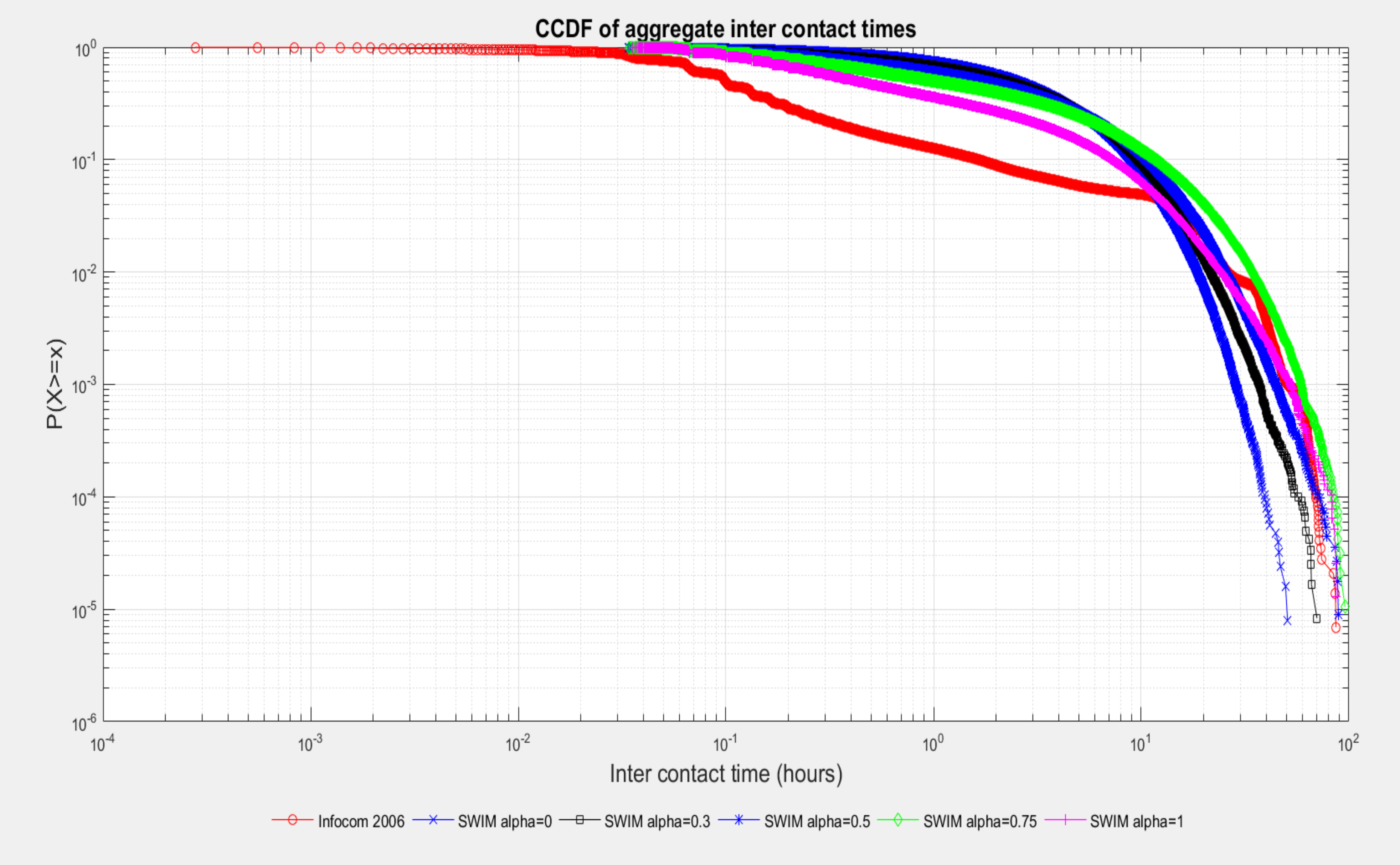

INFOCOM 2005
SWIM alpha=0
SWIM alpha=0.3
SWIM alpha=0.5
SWIM alpha=0.8
SWIM alpha=1
Freq
Hour of the day
0
5
10
15
20
25
500
1000
1500
2000
2500
3000
3500
4000
200
400
600
800
1200

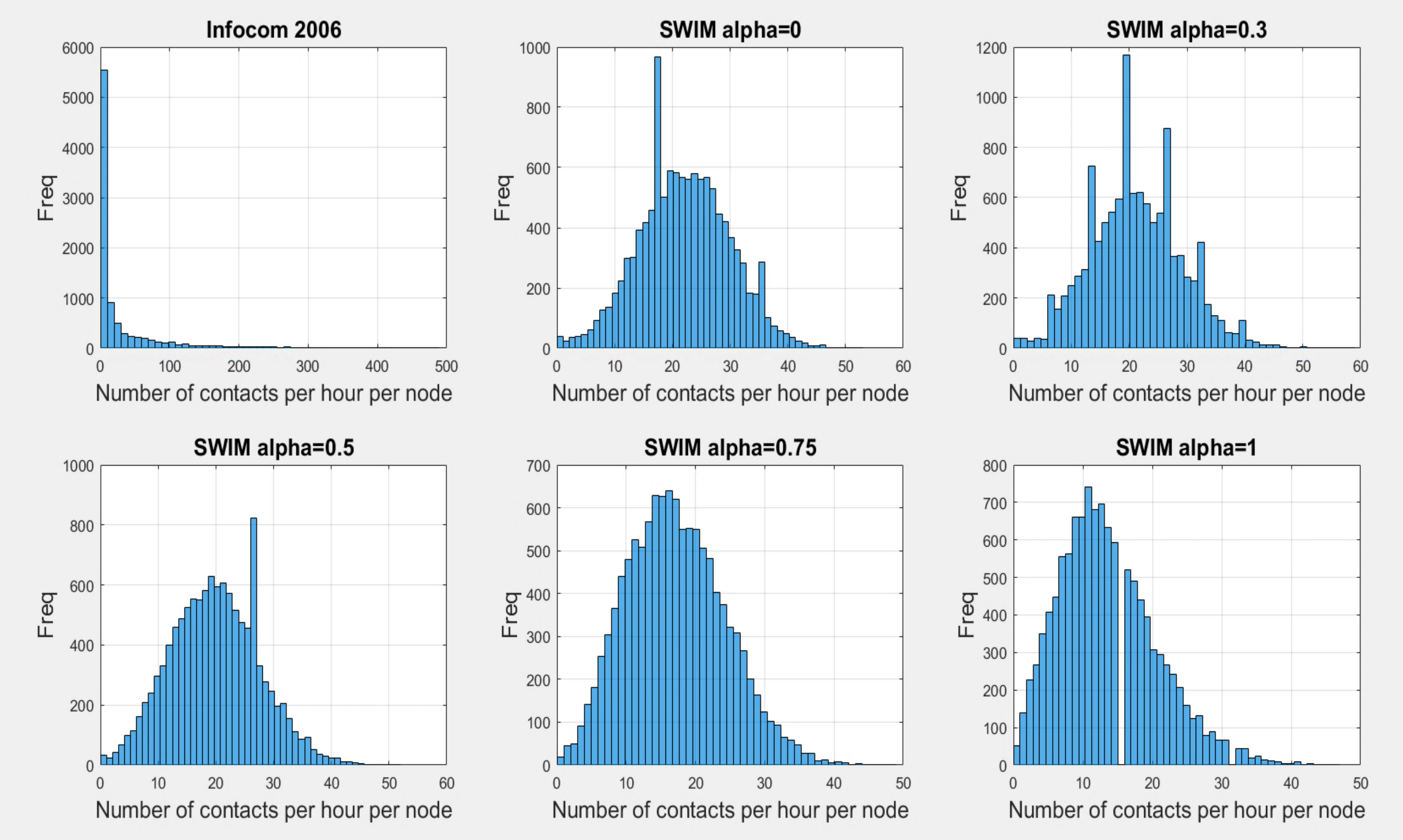

INFOCOM 2005
SWIM alpha=0
SWIM alpha=0.3
SWIM alpha=0.5
SWIM alpha=0.8
SWIM alpha=1
Freq
Number of contacts per hour per node


INFOCOM 2005
SWIM alpha=0
SWIM alpha=0.3
SWIM alpha=0.5
SWIM alpha=0.8
SWIM alpha=1
Freq
Number of contacts per hour per node

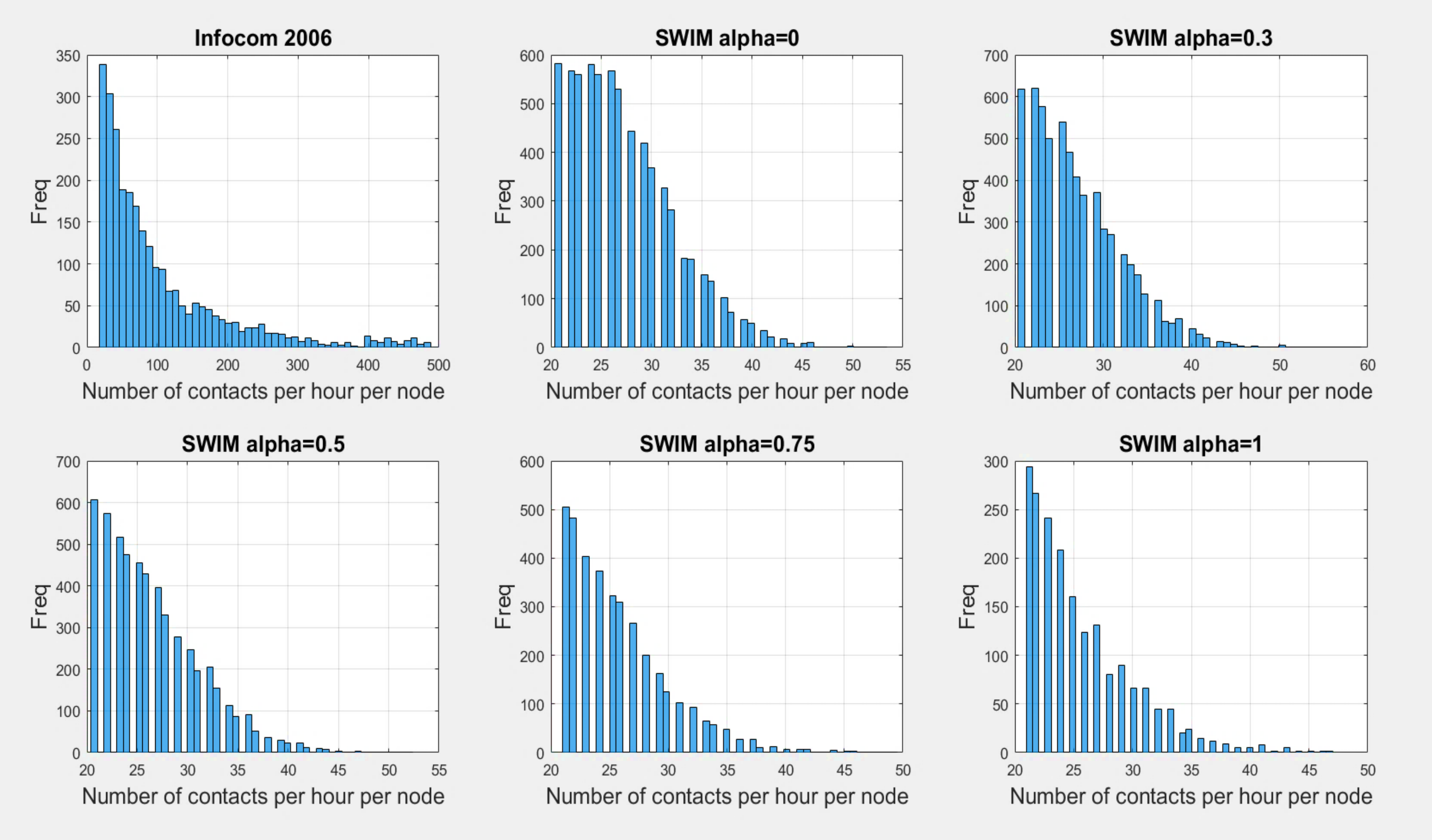

CCDF of number of contacts per hour per node
P(X>=x)
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0
50
100
150
200
250
Number of contacts per hour per node
INFOCOM 2005
SWIM alpha=0
SWIM alpha=0.3
SWIM alpha=0.5
SWIM alpha=0.8
SWIM alpha=1

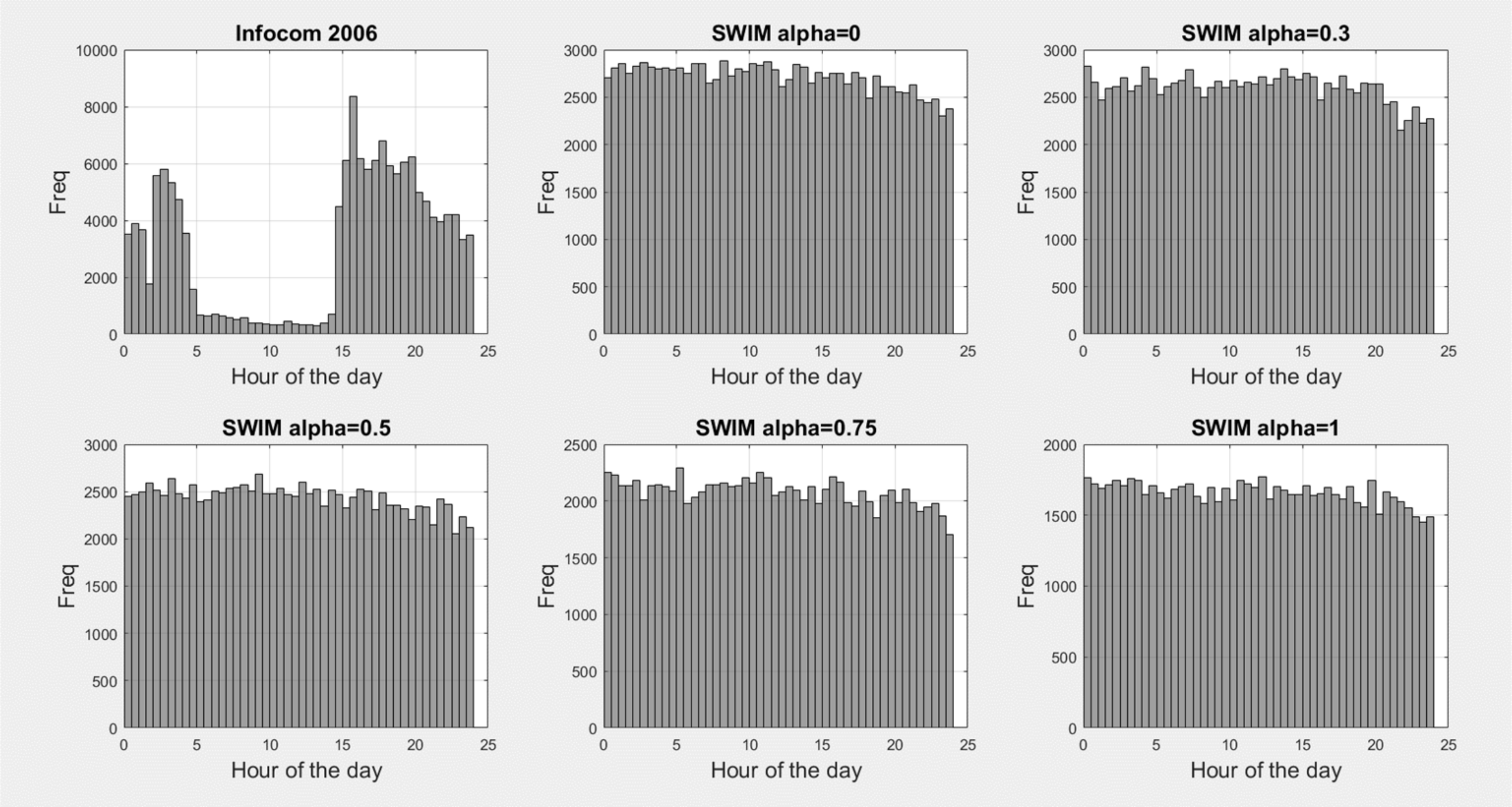

CCDF of Contact durations
P(X>=x)
Contact duration (hours)
Infocom 2006
SWIM alpha=0
SWIM alpha=0.3
SWIM alpha=0.5
SWIM alpha=0.75
SWIM alpha=1

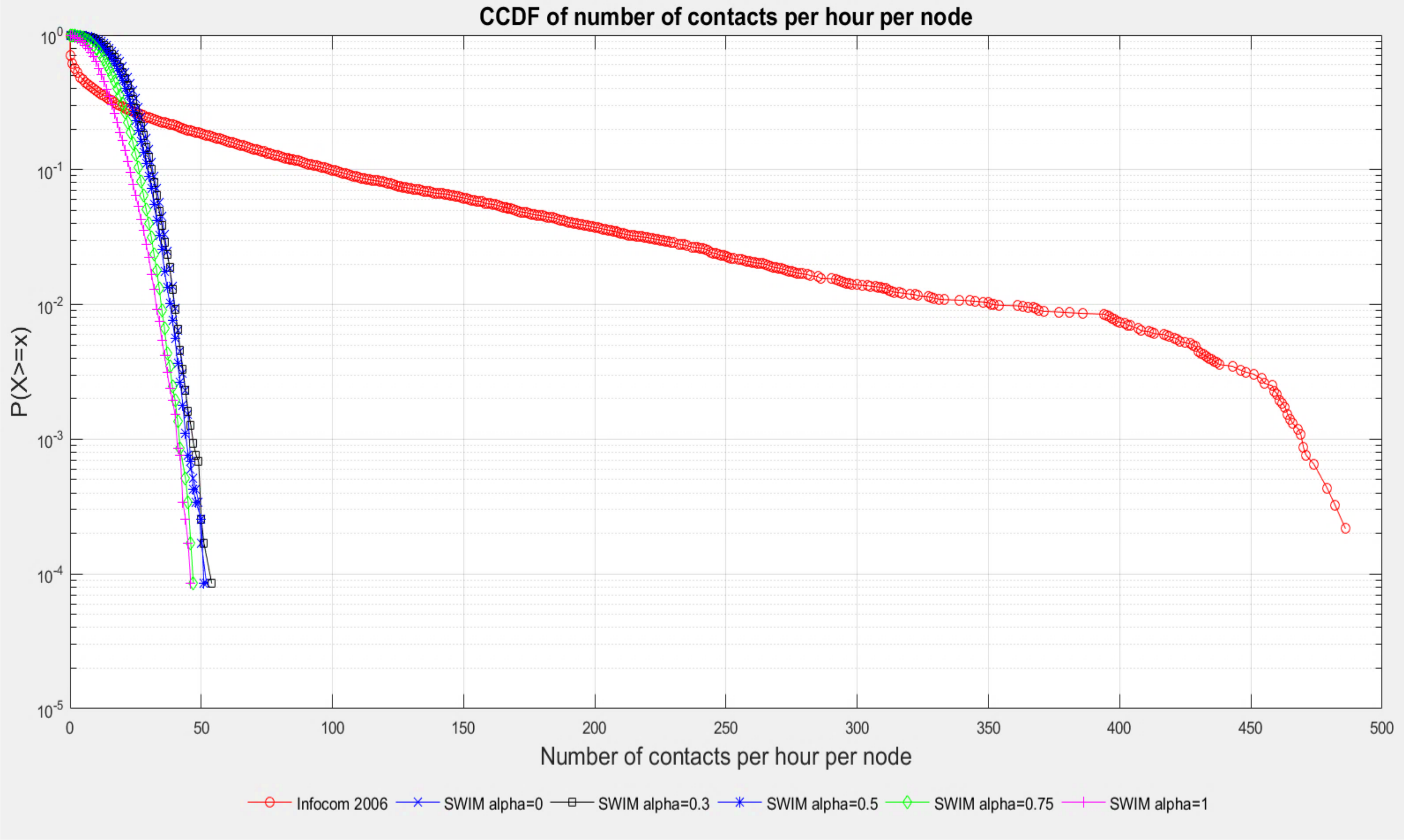

CCDF of aggregate inter contact times
P(X>=x)
Inter contact time (hours)
Infocom 2006
SWIM alpha=0
SWIM alpha=0.3
SWIM alpha=0.5
SWIM alpha=0.75
SWIM alpha=1



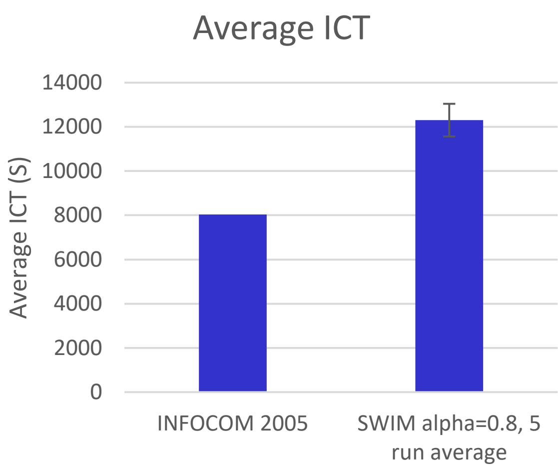
Infocom 2006
Freq
Number of contacts per hour per node
SWIM alpha=0
Freq
Number of contacts per hour per node
SWIM alpha=0.3
Freq
Number of contacts per hour per node
SWIM alpha=0.5
Freq
Number of contacts per hour per node
SWIM alpha=0.75
Freq
Number of contacts per hour per node
SWIM alpha=1
Freq
Number of contacts per hour per node

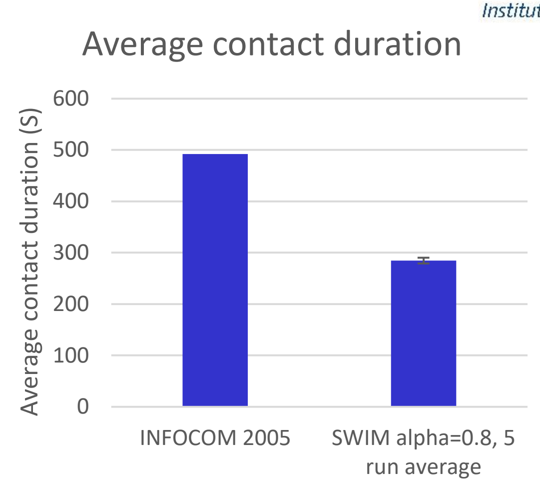

Infocom 2006
SWIM alpha=0
SWIM alpha=0.3
SWIM alpha=0.5
SWIM alpha=0.75
SWIM alpha=1
Freq
Number of contacts per hour per node

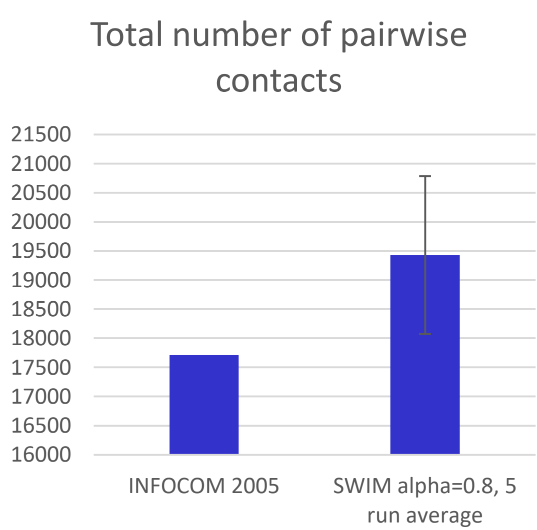

Infocom 2006
SWIM alpha=0
SWIM alpha=0.3
SWIM alpha=0.5
SWIM alpha=0.75
SWIM alpha=1
Freq
Hour of the day

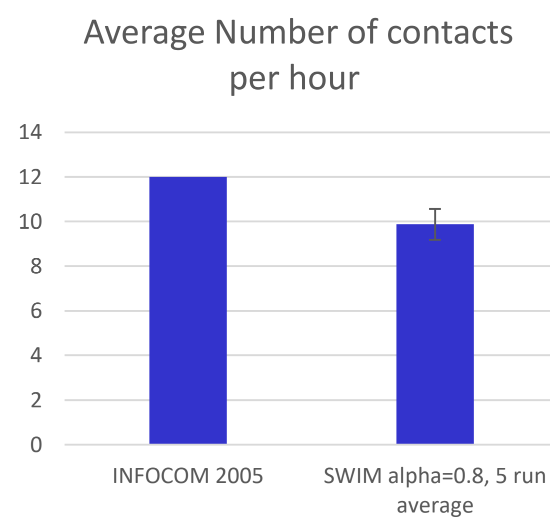


CCDF of number of contacts per hour per node
P(X>=x)
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
0
50
100
150
200
250
300
350
400
450
500
Number of contacts per hour per node
Infocom 2006
SWIM alpha=0
SWIM alpha=0.3
SWIM alpha=0.5
SWIM alpha=0.75
SWIM alpha=1


Average ICT
Average ICT (S)
14000
12000
10000
8000
6000
4000
2000
0
INFOCOM 2005
SWIM alpha=0.8, 5 run average


Average contact duration
Average contact duration (S)
600
500
400
300
200
100
0
INFOCOM 2005
SWIM alpha=0.8, 5 run average


Total number of pairwise contacts
21500
21000
20500
20000
19500
19000
18500
18000
17500
17000
16500
16000
INFOCOM 2005
SWIM alpha=0.8, 5 run average


Average Number of contacts per hour
14
12
10
8
6
4
2
0
INFOCOM 2005
SWIM alpha=0.8, 5 run average

Where were the cambridge traces obtained?

- Experiment 1 and 2 were conducted in Cambridge during the month of January 2005 with 8 and 12 participants respectively.
- **Experiment 3** was conducted in the Conference INFOCOM 2005 in March 2005 in Miami with 41 participants
- **Experiment 4** were conducted in and around cambridge (2005), UK with students of the cambridge University with 36 mobile nodes and 18 stationary nodes
- **Experiment 6** was conducted in the conference INFOCOM 2006 in April in Barcelona with 78 participants and 20 stationary iMotes

Mobility models

- Mobility models are used to model movement
- Mobility models are often simulated to obtain synthetic traces of mobility
- The real life traces cannot be always used because
  - It takes a lot of time and resources to record them
- Mobility models can be simulated to obtain similar behaviour like the real life traces

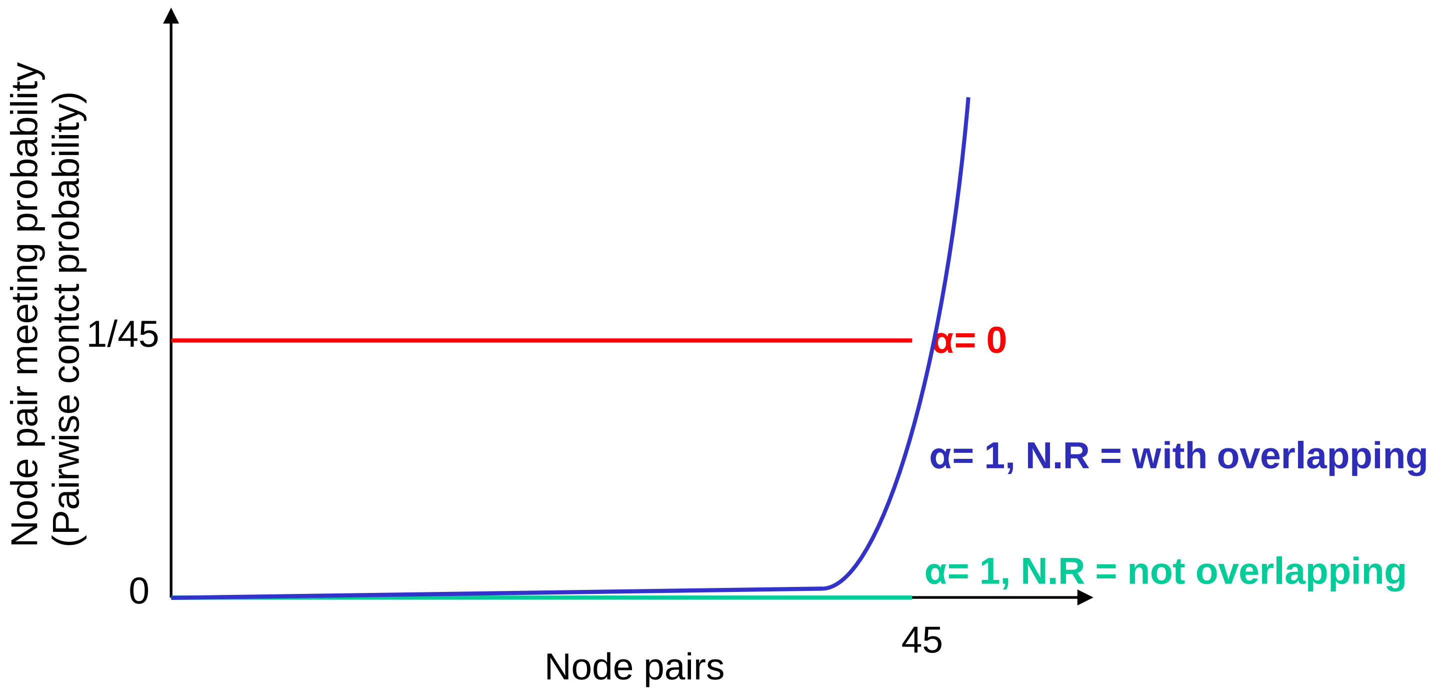
Motivation

- The future is going to be booming with the Internet of things
- Some data are urgent and some data are not
- Why not send the „not so urgent“ data using moving humans?
- Humans will take the data from place to place, acting as the enablers of „Opportunistic Networks“

Aim of the project

- State of Art (SoA) analysis of trace based human movement patterns in opportunistic networks and delay tolerant networks (DTNs)
- **Investigate on procedures to find pairwise contact probabilities from the Cambridge iMote traces**
- **Represent the pairwise contact probabilities as a matrix**
- **Comparison of results with the SWIM [1] (Small Worlds in Motion) mobility model**

[1] A. Mei and J. Stefa, "SWIM: A Simple Model to Generate Small Mobile Worlds", IEEE INFOCOM,2009


Node pair meeting probability
(Pairwise contct probability)
1/45
0
45
Node pairs
α= 0
α= 1, N.R = with overlapping
α= 1, N.R = not overlapping