Major Area Exam

Reading List

Stefan Karpinski
Department of Computer Science
University of California, Santa Barbara
sgk@cs.ucsb.edu

May 29, 2006

1 Reference

[1] W. Stevens. *TCP/IP Illustrated, Volume 1: The Protocols*. Addison-Wesley Professional, 1st edition, 1993.

[2] A. Tanenbaum. *Computer Networks*. Prentice Hall, Upper Saddle River NJ, 3rd edition, 1996.

2 General Networking

[3] D. Clark. The Design Philosophy of the DARPA Internet Protocols. *Proceedings of ACM SIGCOMM '88*, 1988.

[4] J. Saltzer, D. Reed, D. Clark. End-to-End Arguments in System Design. *ACM Transactions on Computer Systems*, 1984.

[5] V. Paxson, S. Floyd. Why we don't know how to simulate the Internet. *Proceedings of 1997 Winter Simulation Conference*, Atlanta, GA, December 1997.

[6] M. Faloutsos, P. Faloutsos, C. Faloutsos. On power-law relationships of the Internet topology. *Proceedings of ACM SIGCOMM'99*, Cambridge MA, August 1999.

[7] A. Medina, I. Matta, J. Byers. On the origin of power laws in Internet topologies. *ACM Computer Communication Review*, April 2000.

[8] E. Zegura, K. Calvert, M. Donahoo. A quantitative comparison of graph-based models for Internet topology. *IEEE/ACM Transactions on Networking*, December 1997.

[9] M. Doar. A better model for generating test networks. *Proceedings of IEEE Global Internet (GLOBECOM'97)*, London, UK, November 1997.

[10] M. Satyanarayanan. Fundamental challenges in mobile computing. *Proceedings of ACM Symposium on Principles of Distributed Computing*, Philadelphia PA, May 1996.

[11] J. Macker, M. Corson. Mobile ad hoc networking and the IETF. *ACM Mobile Computing and Communications Review*, 1998.

[12] L. Perrone, Y. Yuan, D. Nicol. Modeling and Simulation Best Practices for Wireless Ad hoc Networks. *Proceedings of the Winter Simulation Conference*, December 2003.

3 Mathematical Foundations

[13] R. Ash. *Information Theory.* Interscience Publishers, New York NY, 1965.

[14] A. Harvey. *Time Series Models.* The MIT Press, Cambridge MA, 2nd edition, 1993.

[15] D. Dubois, H. Prade. Fuzzy sets and probability: Misunderstandings, bridges and gaps. *Proceedings of Second IEEE International Conference on Fuzzy Systems*, Piscataway NJ, 1993.

[16] G. Box, G. Jenkins. *Time Series Analysis: Forecasting & Control.* Holden-Day, 2nd edition, 1994.

[17] P. Kokoszka, M. Taqqu. Fractional ARIMA with Stable Innovations. *Stochastic Processes and their Applications*, 1995.

[18] R. Adler, R. Feldman, M. Taqqu. *A Practical Guide to Heavy Tails: Statistical Techniques and Applications.* Birkhauser, Boston MA, 1998.

[19] J. Carlson, J. Doyle. Highly optimized tolerance: A mechanism for power laws in designed systems. *Physical Review E*, August 1999.

[20] S. Ross. *Introduction to probability and statistics for engineers and scientists.* Harcourt Academic Press, San Diego CA, 2000.

[21] A. Downey. The structural cause of file size distributions. *Proceedings of the Ninth International Symposium in Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS01)*, 2001.

[22] S. Ross. *Probability models for computer science.* Harcourt Academic Press, San Diego CA, 2002.

4 Node Behavior: Analysis, Modeling and Prediction

4.1 Wired Networks

[23] W. Leland, M. Taqq, W. Willinger, D. Wilson. On the self-similar nature of Ethernet traffic. *Proceedings of ACM SIGCOMM'93*, San Francisco CA, 1993.

[24] V. Paxson, S. Floyd. Wide area traffic: the failure of poisson modeling. *IEEE/ACM Transaction on Networking*, 1995.

[25] S. Taqqu, W. Willinger, R. Sherman. Proof of a fundamental result in self-similar traffic modeling. *ACM SIGCOMM Computer Communication Review*, 1997

[26] W. Willinger, V. Paxson. Where mathematics meets the Internet. *Notices of the American Mathematical Society*, August 1998.

[27] R. Wolski. Dynamically forecasting network performance using the network weather service. *Journal of Cluster Computing*, 1998.

[28] M. Fischer, T. Fowler. Fractals, Heavy-Tails and the Internet. *Sigma Technology Summaries, 2 (Summer 2001)*, Mitretek Systems Technical Report, 2001.

[29] N. Brownlee, K. Claffy. Understanding Internet traffic streams: Dragonflies and tortoises. *IEEE Communications Magazine*, October 2002.

[30] S. Saroiu, P. Gummadi, S. Gribble. A Measurement Study of Peer-to-Peer File Sharing Systems. Technical Report UW-CSE-01-06-02, Department of Computer Science & Engineering, University of Washington, 2002.

[31] W. Willinger, R. Govindan, S. Jamin, V. Paxson, S. Shenker. Scaling phenomena in the Internet: Critically examining criticality. *Proceedings of the National Academy of Sciences of the United States of America*, February 2002.

[32] Y. Zhang, L. Breslau, V. Paxson, S. Shenker. On the Characteristics and Origins of Internet Flow Rates. *Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications*, 2002.

[33] C. Estan, S. Savage, G. Varghese. Automatically Inferring Patterns of Resource Consumption in Network Traffic. *Proceedings of SIGCOMM*, Karlsruhe, Germany, 2003.

[34] J. Brevik, D. Nurmi, R. Wolski. Automatic methods for predicting machine availability in desktop grid and peer-to-peer systems. *Proceedings of Fourth International Scientific Workshop on Global and Peer-to-Peer Computing*, Chicago IL, 2004.

[35] T. Karagiannis, M. Molle, M. Faloustos, A. Broido. A nonstationary poisson view of Internet traffic. *Proceedings of IEEE INFOCOM 2004*, March 2004.

[36] S. Uhlig. Non-stationarity and high-order scaling in TCP flow arrivals: a methodological analysis. *ACM SIGCOMM Computer Communications Review*, April 2004.

[37] S. Sen, J. Wang. Analyzing Peer-To-Peer Traffic Across Large Networks. *IEEE/ACM Transactions on Networking*, April 2004.

[38] C. Wright, F. Monrose, G. Masson. HMM Profiles for Network Traffic Classification (Extended Abstract). *Proceedings of VizSEC/DMSEC04*, Washington DC, October 2004.

[39] H. Yin, C. Lin, B. Sebastien, B. Li, G. Min. Network traffic prediction based on a new time series model. *International Journal of Communication Systems*, February 2005.

[40] T. Karagiannis, K. Papagiannaki, M. Faloutsos. BLINC: Multilevel Traffic Classification in the Dark. *Proceedings of ACM SIGCOMM*, August 2005.

[41] N. Brownlee. Some Observations of Internet Stream Lifetimes. *Passive and Active Measurement Workshop*, 2005

[42] F. Harmantzis, D. Hatzinakos. Heavy Network Traffic Modeling and Simulation using Stable FARIMA Processes. Not yet published (see `http://personal.stevens.edu/~fharmant/Publish/Stochastics/HHitc05.PDF`).

[43] W. Gong, Y. Liu, V. Misra. Self-Similarity and Long Range Dependence on the Internet: A Second Look at the Evidence, Origins and Implications. Not yet published (see `http://www-net.cs.umass.edu/~yongliu/papers/cn.pdf`).

[44] W. Shi, M. MacGregor, P. Gburzynski. Synthetic Trace Generation for the Internet: An Integrated Model. Not yet published (see `http://www.cs.ualberta.ca/~pawel/PAPERS/spects04.pdf`).

[45] K. Lan, J. Heidemann. On the correlation of Internet flow characteristics. Not yet published (see `http://www.cse.unsw.edu.au/~klan/paper/flow.pdf`).

4.2 Wireless Networks

[46] T. Liu, P. Bahl, I. Chlamtac. Mobility modeling, location tracking, and trajectory prediction in wireless ATM networks. *IEEE Journal on Selected Areas in Communications*, 1998.

- ▷ Early cell network research on predicting user mobility.
- ▷ Useful for predicting the next cell more accurately, but doesn't provide an overall picture of node mobility.
- ▷ Uses Kalman filter to predict the next cell station.

[47] D. Tang, M. Baker. Analysis of a Metropolitan-Area Wireless Network. *Proceedings of Mobicom 1999*, Seattle WA, Aug 1999.

- ▷ Analysis of seven-week trace of the Metricom metropolitan-area packet radio wireless network.
- ▷ Examines the following features:

  - ◇ number of active nodes over days and hours
  - ◇ frequency and distance of node movement
- ▷ Uses clustering to classify radio stations by proximity, and mobile nodes by:
  - ◇ total number of locations
  - ◇ average number of location changes per active day
  - ◇ total number of active days
  - ◇ average number of active days per active week
  - ◇ average number of queries per day
- ▷ No resulting models, just descriptive analysis

[48] D. Tang, M. Baker. Analysis of a local-area wireless Network. *Proceedings of Mobicom 2000*, Boston MA, 2000.

- ▷ Analysis of a 12-week trace of a WLAN at the Gates Computer Science Building at Stanford.
- ▷ Examines the following aggregate network features:
  - ◇ active users per time of day
  - ◇ active users per day of week
  - ◇ number of active days per user
  - ◇ number of APs visited per user
  - ◇ number of users per AP
  - ◇ number of handoffs per AP
  - ◇ throughput by bytes, packets, APs, applications, and users
- ▷ Examines the following “user traffic characteristics”:
  - ◇ application breakdown by users, packets/bytes, in/out
  - ◇ classes of application mixture profiles (by eye, not clustering)
- ▷ Purely descriptive, no models.

[49] D. Kotz, K. Essien. Analysis of a campus-wide wireless network. *Proceedings of ACM MobiCom 2002*, September 2002.

- ▷ Analysis of traces from WLAN at Dartmouth in the Fall term of 2001.
- ▷ Based on syslog, snmp, and sniffed traffic (node = card).
- ▷ Traffic:
  - ◇ total traffic volume
  - ◇ amount of traffic per card
  - ◇ variance across hours, days, weekdays
- ▷ Users and Mobility:
  - ◇ number of cards
  - ◇ number of active days per card
  - ◇ number of APs per card
  - ◇ number of buildings per card
- ▷ Card (Node) Activity:
  - ◇ number of active cards
  - ◇ times cards are active
  - ◇ session duration
  - ◇ number of sessions per day
  - ◇ distribution of sessions per building
  - ◇ roaming: portion of sessions, how often per session
- ▷ AP Activity:
  - ◇ number of APs
  - ◇ times APs are active

    - ◇ distribution of activity across APs
  - ▷ Buildings...
  - ▷ Protocols...
  - ▷ Basically, it's a laundry list of descriptive questions.

[50] A. Balachandran, G. Voelker, P. Bahl, V. Rangan. Characterizing User Behavior and Network Performance in a Public Wireless LAN. *Proceedings of ACM SIGMETRICS'02*, Marina Del Ray CA, June 2002.

  - ▷ Analysis of public WLAN trace recorded over 3 days at SIGCOMM at UCSD in August 2001.
  - ▷ Aims to "characterize user behavior and network performance in a public WLAN environment."
  - ▷ User behavior:
    - ◇ connection session length
    - ◇ user distribution across APs
    - ◇ mobility patterns
    - ◇ application mixture
    - ◇ bandwidth requirements
  - ▷ Network performance:
    - ◇ overall, individual AP load
    - ◇ packet errors, retransmissions
  - ▷ Major claims/findings:
    - ◇ mobility follows a two-state Markov-Modulated Poisson Process (MMPP).
    - ◇ session length is Pareto-distributed.
    - ◇ load distribution across APs is uneven.
    - ◇ AP workload does not follow number of users.
  - ▷ Major criticisms:
    - ◇ They present *no evidence* for their claim that mobility follows a two-state MMPP model other than stating it repeatedly and referring to previous work. Moreover, the previous work (not available online) is for a model of *aggregate traffic* in mobile networks, not mobility itself. Moreover, it was published in 1996, at which point there was no actual public WLAN data on which to base realistic models.
    - ◇ Their claim that session duration is Pareto is also not backed up—they don't even show the PDF or CDF of the distribution or mention any quality of fit measures.
  - ▷ This paper makes two claims that relate to user behavior modeling, but both claims are unsubstantiated and given without any trace of evidence.

[51] T. Camp, J. Boleng, V. Davies. A survey of mobility models for ad hoc network research. *Wireless Communications and Mobile Computing*, September 2002.

  - ▷ Individual entity models:
    - ◇ random walk (and derivatives)
    - ◇ random waypoint
      - ∗ discusses density waves but not speed decay [52, 54].
    - ◇ random direction
      - ∗ travel to edge, bounce off
      - ∗ avoids density waves (and speed decay)
    - ◇ boundless simulation area
      - ∗ rectangular simulation area topologically transformed in a torus
      - ∗ speed, direction adjusted every unit of time by uniform random delta values
    - ◇ Gauss-Markov mobility model
      - ∗ speed, direction modeled as ARMA(1) random time-series
      - ∗ nodes are forced away from simulation edges
    - ◇ probabilistic random walk
      - ∗ velocity: $(x, y) \in \{-1, 0, +1\}^2$

        - x, y modeled as independent three-state Markov processes
        - the Markov transition matrix is the simulation “control knob”
    - city section model
        - city modeled as uniform grid
        - streets have different speed limits
        - destination chosen at random
        - path chosen is the fastest
- Group mobility models
    - exponential correlated model
    - column model: uniformly moving nodes in some formation
    - nomadic community model
        - group reference point moves according to one model
        - each node has own space and moves within that
    - pursue model: a group follows a moving target
    - reference point model
        - group movements centered on a moving center point
        - not clear on how this differs from nomadic community
        - uses random waypoint for both levels of mobility
- Simulation results: choice of mobility model impacts protocol performance.

[52] J. Yoon, M. Liu, B. Noble. Random Waypoint Considered Harmful. *Proceedings of INFOCOM 2003*, San Francisco, April 2003.

- Discovered that node speed decays in random waypoint model.
- Investigate impact and ways to mitigate the problem.

[53] M. Balazinska, P. Castro. Characterizing mobility and network usage in a corporate wireless local-area network. *Proceedings of ACM MobiSys 2003*, May 2003.

- Analysis of four week trace gathered on a corporate WLAN (1366 users).
- Features examined:
    - distribution of load across APs
    - descriptive analysis of users
    - correlation between users and load
    - correlation between time of day and load
    - distribution of average individual transfer rates
    - number of APs visited per user
    - fraction of time users spend at “home location”
    - AP prevalence and persistence wrt users $\Rightarrow$ this merits more investigation
- Important results:
    - average user transfer rate found to follow a clean power law, possibly Pareto
    - user transfer rates *are* dependent on location, and thus mobility
    - interesting results on prevalence and persistence of APs wrt users

[54] J. Yoon, M. Liu, B. Noble. Sound Mobility Models. *Proceedings of ACM MobiCom'03*, San Diego CA, September 2003.

- Broad class of random mobility models suffer from decay.
- Ways to avoid the decay phenomenon:
    - using a warmup period is a partial solution
    - pick initial parameter values from their marginal distributions

[55] A. Jardosh, E. Belding-Royer, K. Almeroth, S. Suri. Towards Realistic Mobility Models For Mobile Ad hoc Networks. *Proceedings of ACM MobiCom'03*, San Diego CA, September 2003.

- ▷ Obstacle mobility model:
  - ◇ obstacles are polygonal shapes
  - ◇ paths are given by Vornoi graph for obstacle vertices
  - ◇ intersection of graph with obstacles defines "doorways"
  - ◇ other graph intersections define destination points (interior, exterior)
- ▷ Line-of-sight model for transmission through obstacles.

[56] F. Chinchilla, M. Lindsey, M. Papadopouli. Analysis of wireless information locality and association patterns in a campus. *Proceedings of IEEE INFOCOM 2004*, March 2004.

- ▷ Analysis of traces from UNC campus WLAN, February through April 2003 (7,694 clients).
- ▷ Primary focus on data locality and possible caching paradigms.
- ▷ Secondary focus on user association and mobility patterns.
- ▷ Using Markov chains, 86% accuracy in predicting next AP association.
  - ◇ individual transition matrix is maintained *for each mobile node*
  - ◇ Markov state determined by either most recent APs those 24 hours prior
  - ◇ how much of the remaining 14% accuracy is due to inherent randomness?
  - ◇ is it even possible to know this? can we determine inherent randomness?
- ▷ Also looked at revisit probabilities by client and by AP.

[57] D. Schwab, R. Bunt. Characterizing the use of a campus wireless network. *Proceedings of IEEE INFOCOM 2004*, March 2004.

- ▷ Analysis of trace of WLAN at University of Saskatchewan, from one week in January 2003 (134 users).
- ▷ This is another grab-bag of random figures and statistics: no models.

[58] Q. Zheng, X. Hong, S. Ray. Recent Advances in Mobility Modeling for Mobile Ad Hoc Network Research. *Proceedings of ACMSE04*, Huntsville AL, April 2004.

- ▷ Content-free survey of mobility models.
- ▷ Covers the same models as Camp's survey [51], but briefly mentions the obstacle model.
- ▷ Doesn't go into nearly the same depth or detail, except for RWP.

[59] T. Henderson, D. Kotz, I. Abyzov. The changing usage of a mature campus-wide wireless network. *Proceedings of ACM MobiCom 2004*, September 2004.

- ▷ Revisit of Dartmouth campus WLAN previously analyzed by Kotz [49].
- ▷ Utilization and mobility increased; most other characteristics remained static.

[60] X. Meng, S. Wong, Y. Yuan, S. Lu. Characterizing flows in large wireless data networks. *Proceedings of ACM MobiCom 2004*, September 2004.

- ▷ Analysis of multiple wireless traces (Dartmouth '02, UCSD '01, IBM '02, Stanford '00).
- ▷ Analysis of flows rather than packets or bytes.
- ▷ Establishes convincing distribution fits for the following:
  - ◇ Interarrival times: Lognormal
  - ◇ Residing times: Lognormal
  - ◇ Number of APs visited: Weibull/Gamma

[61] M. Musolesi, S. Hailes, C. Mascolo. An Ad Hoc Mobility Model Founded on Social Network Theory. *Proceedings of ACM MSWiM04*, Venezia, Italy, October 2004.

- ▷ Interesting idea, but the actual mobility model is disappointing:
  - ◇ the only thing that's determined according to the social network theory is node groups, and it's not at all clear that the "social network theory" is really accurately modeling anything relevant or why the resulting groups would be more realistic than anything else

  - ◇ nodes move within "the group area" according to good ol' random way point
  - ◇ groups move in some unspecified manner towards "randomly chosen goals"
- ▷ Evaluation is practically non-existent.

[62] C. Tuduce, T. Gross. A Mobility Model Based on WLAN Traces and its Validation. *Proceedings of IEEE INFOCOM*, March 2005.

- ▷ Mobility model derived from two, two-month campus WLAN traces.
- ▷ Details of the mobility model:
  - ◇ simulation space divided into cells
  - ◇ checkpoint set is set of cells a nodes is going to visit
  - ◇ size of checkpoint set follows a power law
  - ◇ additional cells chosen probabilistically (same, neighbor or non-neighbor)
  - ◇ doesn't consider correlation between checkpoint set size and probability values
  - ◇ later, cells are chosen from checkpoint set uniformly
  - ◇ session durations are independently chosen from a Pareto distribution
  - ◇ inter-session times are strangely chosen from a uniform distribution
- ▷ Evaluation:
  - ◇ clever method for choosing the test user set from their data: examine MSE of mobility characteristics (they don't way what) as smaller sets are chosen; choose the smallest set that still will exhibit correct properties.
  - ◇ evaluation of spatial distribution is not so convincing.

[63] A. Jardosh, E. Belding-Royer, K. Almeroth, S. Suri. Real-World Environment Models for Mobile Network Evaluation. *IEEE Journal on Selected Areas in Communications*, March 2005.

- ▷ Same polygonal obstacle model with Voronoi paths, doorways and destinations [55].
- ▷ Mobility model additions and variations:
  - ◇ exponential destination selection
  - ◇ note: it would be good to look at how Markov models of mobility from actual traces relate to the exponential destination selection model; are the transition probabilities really exponentially related to distance?
- ▷ Signal propagation model:
  - ◇ mixed Friis'/two-ray propagation model
  - ◇ use real observations of signal attenuation through building walls
- ▷ Simulation to show differences induced by obstacle model *vs.* other models

[64] M. Papadopouli, H. Shen, M. Spanakis. Characterizing the duration and association patterns of wireless access in a campus. *11th European Wireless Conference*, April 2005.

[65] N. Eagle, A. Pentland. Reality Mining: Sensing Complex Social Systems. *Journal of Personal and Ubiquitous Computing*, September 2005 (to appear; see also `http://reality.media.mit.edu/pdfs/realitymining.pdf`).

[66] R. Jain, A. Shivaprasad, D. Lelescu, X. He. An empirical model for user registration patterns in a campus wireless LAN. *Proceedings of ACM MobiCom'05*, Cologne, Germany, September 2005 (to appear; see also `http://delivery.acm.org/10.1145/1060000/1052877/p59-jain.pdf`).

[67] M. Papadopouli, H. Shen, M. Spanakis. Modeling client arrivals at access points in wireless campus-wide networks. Not yet published (see `http://www.csd.uoc.gr/~maria/papers/lanman05-A.pdf`).

[68] H. Zimmermann , I. Gruber, C. Roman. A Voronoi-Based Mobility Model For Urban Environments. Not yet published (see `http://www.lkn.ei.tum.de/forschung/publikationen/dateien/Zimmermann2005AVoronoi-basedMobilityModel.pdf`).

[69] J. Kim, S. Bohacek. A Survey-Based Mobility Model of People for Simulation of Urban Mesh Networks. Not yet published (see `http://www.eecis.udel.edu/~bohacek/UDelModels/pubs/urbanpedmobility.pdf`).