

Understanding Network Traffic from Billing Traces: customers' behavior analysis

Hui Zhang, Luc Aurelien Andriantatsaholiniaina, Milan Nikolic, Slobodan Petrovic, and Ljiljana Trajkovic
 {hzhange, luca, milan, slobodan, ljilja}@cs.sfu.ca, <http://www.ensc.sfu.ca/research/cnl>
 Communication Networks Laboratory, School of Engineering Science, Simon Fraser University

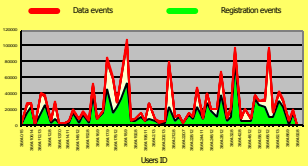
WHAT ARE OUR BILLING TRACES?

- Data summarizing customers' activities for billing purposes
- From Cellular Digital Packet Data network
- Contain Header and Traffic Matrix Segments with registration, de-registration, and data events



DATA EXTRACTION TOOLS

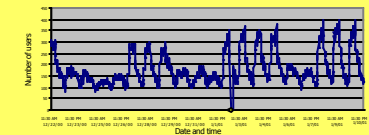
- Structured Query Language Server: MySQL
- Data formatting with Java code
- Automatic classification: AutoClass
- Statistical package: S-PLUS



NETWORK CHARACTERISTICS

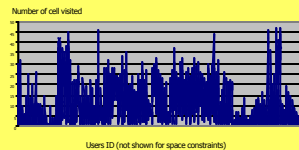
- Static:
- Network topology
 - Network elements
 - Total events
- Dynamic:
- Number of users
 - Data packets
 - Discarded packets

	Number of users	Percentage of discarded
Registration	117248	21.2%
De-registration	17216	21.4%
Data	480227	50.2%
Discarded	480227	100%



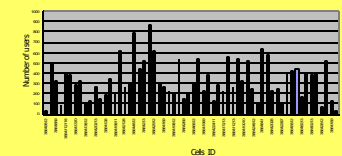
CUSTOMERS' BEHAVIOR

- Mobility: cells visited per user
- Activity: registration, de-registration, and data events per user
- Loss: discarded packets per user
- Clustering: classification of users



CELL ACTIVITY

- Number of users
- Registration events
- De-registration events
- Data events
- Data packets per user



References:

- Java: <http://java.sun.com>
- MySQL: <http://www.mysql.com>
- S-PLUS: <http://www.insightful.com>
- P. Cheeseman and J. Stutz: <http://ic.arc.nasa.gov/ic/bayes-group/group/autoclass>
- D. Tang and M. Baker: <http://mosquitonet.stanford.edu>
- R. Hutchins et al.,: <http://computer.org/cspres>