Mobile Computing

- new paradigm of computing
  - advances in wireless data networking
  - portable information appliances
- Ubiquitous computing / wireless computing / Nomadic computing
- mobile (non-static) and ubiquitous (everywhere)
- connect from different access points through wireless links and might want to stay connected while on the move, despite possible intermittent disconnection
- Applications: Multimedia, Field service, PDA's, Healthcare, Industrial, Managerial, Sales

Challenges and Constraints

- Security Privacy issue - personal information is being captured and stored
- Disconnection, low bandwidth, bandwidth variability, heterogeneity of the network
- Address migration, mobile routing, location dependent information, low power design
- Restricted memory, disk, and user interfaces
- Consistent replicated data, session guarantees
- Cache validation, reduce contention
- Data access model and dynamic network architecture
- Open file system, fault-tolerance, performance
- Distributed control: Distributed object management, Distributed file system (Coda)

Challenges and Constraints

- Mobile host, static hosts as gateways, adhoc network
- Bluetooth - leave and join, active and passive
- Decentralized location query service
- Optimization code, Multiple file systems
- Development
  - Creating new environments and user interfaces for Mobile Computing
  - Agent based remote procedure
  - Web application - WML, Intelligent WAP system, forms manipulation
  - Application-specific knowledge to address mobile resource constraints
  - Video transmission - dynamically adapt the bandwidth

Client-Server Computing in Mobile Environments

- Mobile Vs. Traditional Client-server computing
- Mobility of user and their computers
- Mobile resource constraints
- Goal: a collection of trusted information servers connected via a fixed network to provide information services to a collection of untrusted mobile clients over wireless and mobile networks
- Mobile client-server computing can be categorized into
  - mobile aware adaption
  - extended client-server model
  - mobile data access

Client-Server Computing

- Mobile aware adaption - how systems and applications respond to the environmental changes
  - laisse-faire adaptation - responsibility of individual applications, avoids the need for system support
  - application-transparent adaptation - places responsibility on the system
    * file system proxy, disconnected operations, isolation only transactions
    * automatically detect read/write conflicts, caching, differencing
    * web proxy - client slide intercept, server side intercept, header reduction
  - application-aware adaptation - collaborative adaptation between applications and system
    * applications can decide how to best adapt to the changing environment while
    * the system provides support through the monitoring of resources and the enforcing of resource allocation decisions

Client-Server Computing

- Extended client-server model
  - dynamic partitioning of client-server functionality and responsibilities
  - thin client architecture, flexible client-server architecture - mobile objects
- Mobile data access
  - how data over wireless and mobile networks is structured consistency of client cache
  - type of communication links, connectivity of mobile hosts
  - server data dissemination, client cache management

PCS network architecture

- Personnel communication systems can provide wireless communication services to user on the move
- It is necessary to have an efficient way to locate the mobile user to deliver services
- Cell, Base station, Registration area (RA)
  - An RA consists of an aggregation of a number of cells, forming a contiguous geographical region
- Signaling network
  - used to setup calls and it is distinct from the network used to actually transport the information contents of the calls

Common Channel Signaling (CCS)

- CCS network is used to set up calls which use the signaling systems No. 7 (SS7) protocols
- All base stations are connected via a wire-line network to an end-office switch or Service Switching Point (SSP)
- Each SSP serves as an RA
- All the SSPs of different RAs are connected to higher hierarchical Local Signaling Transfer Points (LSTP), which are connected to a Regional STP (RSTP)
- An RSTP connects to all the LSTPs in one region
- The RSTPs are also connected to a Service Control Point (SCP)
- Each SCP is equipped with a HLR database

- Each VLR is associated with one Mobile Switching Center (MSC)
- MSC connects the BSs and backbone communication infrastructure (PSTN)
- MSC, SSP, and VLR database are associated together to serve to an RA
- Two standards for PCS location management: IS-41 and GSM MAP
- Both use a two-tier system of HLR and VLR databases
- A mobile user performs location update (registration) at the HLR every time the user crosses the boundary of a RA and deregisters at the previous VLR
- Location management does through location registration and paging
- the objective is to reduce signaling traffic
- Methods used - Local Anchor scheme, per-user caching scheme, pointer forwarding scheme

Location Management Schemes

- Local anchor (LA) scheme: a VLR close to the user is selected as the LA for the user and whenever a user moves from one RA to another, it will perform location update to the LA
- A LA for a mobile will change only when a call request to the mobile arrives;
- at the same time HLR is also updated via the registration process
- When a call request terminating at this user is received by the HLR, the user can be traced to the LA
- The LA scheme avoids update to HLR completely at the expense of the increase in local signaling traffic
- The drawback of LA scheme is that when the user keeps moving constantly without receiving any call, the updates to LA may become costly

Location Management Schemes

- per-user pointer forwarding scheme: some updates to the HLR can be avoided by setting up a forwarding pointer from the previous VLR to the new VLR
- When a call request to a mobile user arrives, the PCS network first queries the user's HLR to determine the VLR, which the user was visiting at the previous location update, the follows a chain of forwarding pointers to the user's current VLR to find the mobile user
- The traffic to the HLR is decreased by using the pointer chain; but the penalty is the time delay for tracking the user when a call to the user arrives
- The longer the pointer chain, the less the signaling traffic, the longer the setup delay for finding the user
- To avoid long setup delay, a threshold of the length of the pointer chain is used
- The user needs to perform registration to the HLR after the chain threshold is reached

Location Management Schemes

- Two Level Pointer Forwarding Strategy for Location Management in PCS Networks, *IEEE Transaction on Mobile Computing Vol. 1 No. 1 Jan-Mar 2002*
- Objective: New location management scheme - mitigate the signaling traffic and reduce the tracking delay
- Solution - a set of VLRs traversed by users as the mobility agents (MA), which forms another level of management in order to make some registration signaling traffic localized, instead of updating HLR, MAs are updated
- two kinds of pointers are used; some VLRs are selected as the mobility agents, which will be responsible for location management in a larger area comparing to the RAs and can be geographically distributed
- MA can form the virtual management network (VMN)
- The pointers between MAs are level-1 pointers
- those between VLRs in the charging domain of MAs are level-2 pointers are set

- When the user crosses the boundaries of RAs, the level-2 pointers are set
- If the level-2 pointer chain threshold is reached, the current RA is selected as an MA for this user and a level-1 pointer is set up from previous MA
- Calls to a given user will query the HLR first and follow the level-1 pointer chain to the current MA, then reach the user current VLR by tracking the level-2 pointer chain
- The user does not need to update the HLR until the level-1 pointer chain threshold is reached
- two level pointer scheme chain can be longer than that in simple pointer forwarding scheme, but can have shorter call setup delay due to the level-1 pointer chain
- two operations are defined:
  - MOVE: PCS user moves from one RA to another and
  - FIND: determination of the RA where the PCS user is currently located

- IS-41 Location Management Scheme: Basic MOVE()
  - The mobile terminal detects that it is in a new RA
  - The MS sends a registration message to the new VLR
  - The new VLR sends a registration message to the user's HLR
  - The HLR sends a registration cancellation message to the old VLR
  - The old VLR sends a cancellation confirmation message to the HLR
  - The HLR sends a registration confirmation message to the new VLR
- IS-41 Location Management Scheme: Basic FIND()
  - Call to a PCS user is detected at the local switch, if the called party is in the same RA, then return;
  - Switch queries the called party's HLR; HLR queries the called party's current VLR, V; VLR V returns the called party's location to HLR; HLR returns the location to the calling party;

Two-Level Pointer Forwarding Scheme

- When user moves from one RA to another; it informs the switch (and the VLR) at the new RA about the old RA
- it also informs the new RA about the previous MA it was registered
- the switch at new RA determines whether to invoke the BasicMOVE or the TwolevelFwdMOVE
- In TwolevelMOVE: the new VLR exchange message with the old VLR or the old MA to setup a forwarding pointer from the old VLR to the new VLR; if a pointer is set up from the previous MA, the new VLR is selected as the current MA; it does not involve HLR
- level-2 pointers are built from the old VLR to the new VLR
- when the chain threshold for level-2 pointer is reached so the new RA is selected as the user's new MA
- and a level-1 pointer is set up from the old MA to the new MA and level-2 pointer chain is reset

Performance Analysis

- Analytic model - different parameters for different classes of users
- characterize the classes of users according to their call-to-mobility ratio (CMR)
- CMR of a user is defined as the expected number of calls to a user during the period that the user visits an RA
- if calls are received by the user at a mean rate λ and the time the user resides in a given RA has a mean $1/\mu$ then the CMR denoted as $p = \lambda/\mu$
- For comparison model the basic procedures used in IS-41
  - A user crosses several RAs between two consecutive calls; if basic user location strategy is used the user's HLR is updated every time the user moves to a new RA
  - if two level pointer forwarding strategy is used, the HLR is updated only every $K_1 \cdot K_2$
  - K_1 and K_2 are the level-1 and level-2 pointer chain length threshold, while forwarding pointers are set up for all other moves

- C_B and C_F to be the total costs of maintaining the location information (location updating) and locating the user (location tracking) between two consecutive calls for the basic strategy and two level forwarding strategy
- m the cost of a single invocation of BasicMOVE
- M the total cost of all the BasicMOVES between two consecutive calls F the cost of a single BasicFIND
- M' the expected cost of all TwoLevelFwdMOVEs between two consecutive calls
- F' the average cost of the TwoLevelFwdFIND
- S_1 the cost of setting up a forwarding pointer (level-1 pointer) between MAs during a TwoLevelFwdMOVE
- S_2 the cost of setting up a forwarding pointer (level-2 pointer) between VLRs during a TwoLevelFwdMOVE
- T_1 the cost of traversing a forwarding pointer (level-1 pointer) between MAs during a TwoLevelFwdFIND
- T_2 the cost of traversing a forwarding pointer (level-2 pointer) between VLRs during a TwoLevelFwdFIND

- K_1 the threshold of level-1 pointer chain
- K_2 the threshold of level-2 pointer chain
- $\alpha(i)$ the probability that there are i RA crossings between two consecutive calls
- $C_B = M + F = m/p + F$
- $C_F = M' + F'$
- Say a user crosses i RA boundaries between two consecutive calls - the HLR is updated $\left\lfloor \frac{i}{K_1 K_2} \right\rfloor$ times
- $\left\lfloor \frac{i}{K_2} \right\rfloor - \left\lfloor \frac{i}{K_1 K_2} \right\rfloor$ level-1 pointer creations - every K_2 moves may require a level-1 pointer creation
- level-2 pointers are created for all the rest $i - \left\lfloor \frac{i}{K_2} \right\rfloor$ moves

$$M' = \sum_{i=0}^{\infty} \left\{ \left\lfloor \frac{i}{K_1 K_2} \right\rfloor m + \left(\left\lfloor \frac{i}{K_2} \right\rfloor - \left\lfloor \frac{i}{K_1 K_2} \right\rfloor \right) S_1 + \left(i - \left\lfloor \frac{i}{K_2} \right\rfloor \right) S_2 \right\} \alpha(i)$$

- After the last BasicMove operations, the user traverses $\left\lfloor \frac{i-\left\lfloor \frac{i}{K_1K_2} \right\rfloor K_1K_2}{K_2} \right\rfloor$ level-1 pointers and
- $i - \left\lfloor \frac{i}{K_1K_2} \right\rfloor K_1K_2 - \left\lfloor \frac{i-\left\lfloor \frac{i}{K_1K_2} \right\rfloor K_1K_2}{K_2} \right\rfloor K_2$ level-2 pointers
- To evaluate $\alpha(i)$
  - the call arrivals to a user form a Poisson process with arrival rate λ
  - the residence time of a user at a registration area is a random variable with a general density function
  - the expected residence time of a user at an RA is $\frac{1}{\mu}$
  - for demonstration purpose, RA residence time of a user is Gamma distributed with mean $\frac{1}{\mu}$

$$F' = F + \sum_{i=0}^{\infty} \left\{ \left\lfloor \frac{i - \left\lfloor \frac{i}{K_1K_2} \right\rfloor K_1K_2}{K_2} \right\rfloor T_1 + \left(i - \left\lfloor \frac{i}{K_1K_2} \right\rfloor K_1K_2 - \left\lfloor \frac{i - \left\lfloor \frac{i}{K_1K_2} \right\rfloor K_1K_2}{K_2} \right\rfloor K_2 \right) T_2 \right\} \alpha(i)$$

- updating the HLR and performing a BasicFIND involve the same number of messages between HLR and VLR databases, so we can choose $m = F$
- without loss of generality, normalize $m = 1$
- assume the cost of setting up a forwarding pointer is about twice the cost of traversing it since twice as many messages are involved, $S_1 = 2T_1$ and $S_2 = 2T_2$, $S_2 = \delta$ $\delta < 1$ and
- level-1 point is more expensive than level-2 pointer in terms of setup cost, $S_1 = KS_2$ $K \geq 1$

$$\frac{C_F}{C_B} = \frac{p}{1+p}\left\{1 + \frac{3\delta}{2p} + \frac{(K-1)\delta}{(1+p)^{K_2}-1} + \frac{1-(K+\frac{1}{2}K_1K_2)\delta}{(1+p)^{K_1K_2}-1} + \frac{\delta(K-K_2)[(1+p)^{K_1K_2} - K_1(1+p)^{K_2} + K_1 - 1]}{2[(1+p)^{K_1K_2}-1][(1+p)^{K_2}-1]}\right\}$$

Location Estimation Method

- Location aware computing - personal security, navigation, tourism, and entertainment
- Known as geolocation, location identification, localization, positioning
- geometric approach based on angle and distance estimates from which a location estimate is deduced using standard geometry
- problem of inaccurate measurements
- cellular system - complex propagation of radio waves, obstructions, reflecting objects, interference
- explore the dependency between the location of the receiver and observable signal properties, signal attenuation, reflection, diffraction and interference

Statistical Modeling Approach

- the distribution of received signal power - received power is observed
- other signal properties - angle of arrival, and propagation delay treated as random variables;
- these are statistically dependent on the locations of the transmitter, the receiver and the propagation environment
- A Statistical Modeling Approach to Location Estimation *IEEE Transactions on Mobile Computing Vol.1 No. 1 January-March 2002*
- Signal Propagation Model - predicts some properties of a radio signal at a given location
- Signal Transmitter Model - log loss model - linear regression model

- zero mean Gaussian distribution with a constant variance is used for the error term e
- two regression parameters β_0 and β_1 define the mean value of the received power at a given distance
- the variance of e - σ^2, p is the transmitted power in decibels, d is the transmitter receiver distance
- θ denotes the set of parameters

$$\mu(d, p, \theta) = p + \beta_0 + \beta_1 \ln d$$

- direction of transmission to which the transmitted power is higher than to other directions
- log-loss model can be improved by adding a term which depends on the deviation between the direction of the receiver and the direction of transmission - deviation δ

$$\mu(d, p, \theta) = p + \beta_0 + \beta_1 \ln d + \beta_2 \delta \ln d$$

- the distribution of r received signal power

$$f(r|d,\delta,p,\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{r-\mu(d,\delta,p,\theta)}{\sigma}\right)^2\right)$$

- multiple transmitters: r_j denote the received power of channel j and c_i denote the channel of transmitter i
- each transmitter i has location, denoted by l_i, direction of transmission α_i and transmitted power p_i
- g_j p.d.f. of the received power on channel j, measurement is performed at location l

$$g_j(r|l,\theta) = f(r|d(l,l_i),\delta(l,l_i,\alpha_i),p_i,\theta)$$

the index i is chosen so that it maximizes the mean received power

$$i = argmax_{\{i:c_i=j\}}\mu(d(l,l_i),\delta(l,l_i,\alpha_i),p_i,\theta)$$

- received signal power $\mathbf{r} = (r^{(1)}, \ldots, r^{(n)})$
- distances between the transmitter and the receiver $\mathbf{d} = (d^{(1)}, \ldots, d^{(n)})$
- deviations between the direction of transmission and the direction of the receiver $\delta = (\delta^{(1)}, \ldots, \delta^{(n)})$
- transmitted powers $\mathbf{p} = (p^{(1)}, \ldots, p^{(n)})$
- maximum likelihood from complete data

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} f(r^{(i)} | d^{(i)}, \delta^{(i)}, p^{(i)}, \theta)$$

$$\text{SSE} = \sum_{i=1}^{n} (r^{(i)} - \mu^{(i)})^2$$

- maximum likelihood from incomplete data
- EM can be applied - evaluate the expected value of the logarithm of the complete data log-likelihood

- E-step

$$Q(\theta, \theta_t) = E_{\theta_t} \ln \mathcal{L}(\theta)$$

- M-step: maximize the expected log-likelihood

$$\theta_{t+1} = argmax_{\theta} Q(\theta, \theta_t)$$

- Location Estimation

$$p(l|\mathbf{r}, \hat{\theta}) = \frac{g(\mathbf{r}|l, \hat{\theta})\pi(l)}{\int g(\mathbf{r}|l', \hat{\theta})\pi(l')dl'}$$

- $\mathbf{r}$ vector consisting of the received power values r_j for each channel j
- and $g(\mathbf{r}|l, \hat{\theta})$ is the likelihood function given by

$$g(\mathbf{r}|l, \hat{\theta}) = \prod_j g_j(r_j|l, \hat{\theta})$$

- π is the prior p.d.f. of the location variable