Lecture 1: Mobility Management in Mobile Wireless Systems

Ing-Ray Chen

CS 6204 Mobile Computing
Virginia Tech
Fall 2005

Mobility Management

- Location Management
  - Search: find a mobile user's current location
  - Update (Register): update a mobile user's location
  - Location info: maintained at various granularities (cell vs. a group of cells called a registration area)
  - Research Issue: organization of location databases
    - Global Systems for Mobile (GSM) vs. Mobile IP
- Handoff Management
  - Ensuring that a mobile user remains connected while moving from one location (e.g., cell) to another
  - Packets or connection are routed to the new location

2

Handoff Management

- Decide when to handoff to a new access point (AP)
- Select a new AP from among several APs
- Acquire resources such as bandwidth channels (GSM), or a new IP address (Mobile IP)
  – Channel allocation is a research issue: goal may be to maximize channel usage or revenue generated
- Inform the old AP to reroute packets and also to transfer state information to the new AP
- Packets are routed to the new AP

3

Tradeoff in Location Management

- Network may only know approximate location
- By location update (or location registration):
  – Network is informed of the location of a mobile user
- By terminal paging or search:
  – Network is finding the location of a mobile user
- A tradeoff exists between location update and search
  – When the call arrival rate is low, resources are wasted with frequent updates
  – If not done and a call comes, bandwidth is wasted in paging

4

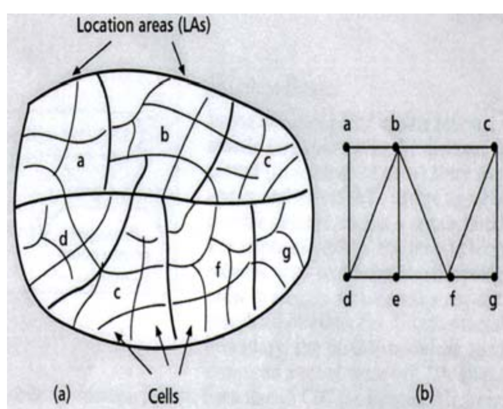
Registration Area (RA) and the Basic HLR-VLR Scheme

- Current Personal Communication Service (PCS) networks (i.e., cellular networks such as GSM) use RA-based basic HLR-VLR schemes:
  - The service coverage area is divided into registration areas (RAs) or location areas (LAs)
  - Each RA covers a group of cells
  - A user has a permanent home location register (HLR)
  - Base stations within the same RA broadcast their IDs
  - If ID is sensed different by the mobile terminal, then a location update is sent to the visitor location register (VLR) of the current RA.
  - When crossing a RA boundary, an update is sent to the HLR.
  - A search goes by HLR->VLR->cell->paging (by the base station)

5

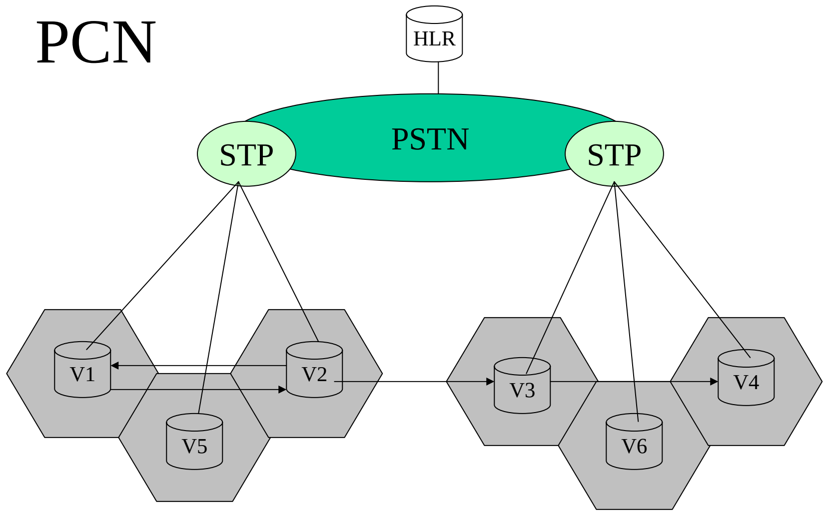
Registration Areas in a PCN

- Large number of cells uses much bandwidth in paging
- Figure shows a PCS network
  - RA topology
  - RA graph model



6

PCN



PSTN: Public Switched Telephone Network

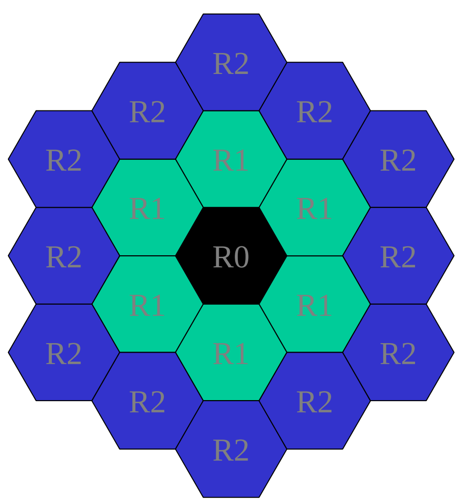
STP: Service Transfer Point

HLR: Home Location Register

VLR: Visitor Location Register

7

Hexagonal Network Coverage Model for PCN



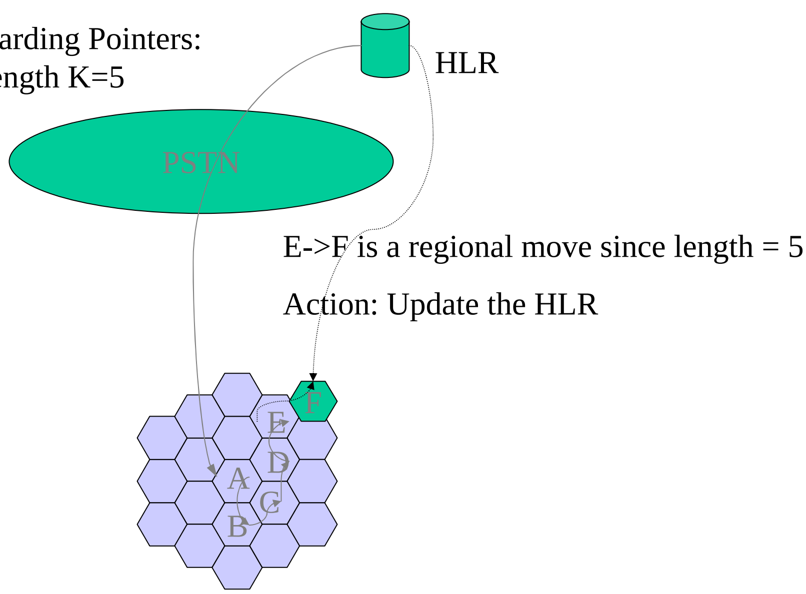
8

Forwarding Pointers

- Update
  - When the length of forwarding pointers < K:
    Set up a pointer between the two involved VLRs
  - When the length of forwarding pointers = K:
    Update information in the HLR
- Search
  - HLR -> VLR0 -> …. -> VLRi -> cell ->paging

9

Forwarding Pointers:
Set length K=5



A->B->C->D->E are all local movements since the length of the forwarding chain is less than K=5

Action: Put a forwarding pointer between two involved VLRs

10

How Big Should K be for Forwarding Pointers?

- The cost "saving" due to forwarding for a location update operation is τ where τ is the cost of accessing a remote registrar (approximately).
- The "increased" cost per search operation is $K\tau$ to follow the forwarding pointers of length K.
- Let λ be the call arrival rate (incurring search) and σ be the mobility rate (incurring location update). Then the increased cost due to search operations per unit time is $\lambda K\tau$, while the cost saving due to update operations per unit time is $\sigma\tau$
- When $\sigma\tau > \lambda K\tau$, or $\sigma/\lambda > K$, it makes sense to have forwarding pointers. In other words, K should be bounded by σ/λ, the reciprocal of λ/σ, or the reciprocal of the call to mobility ratio (CMR)

11

Dynamic Location Update

- Location update algorithms can be static or dynamic
  - With static, an update is triggered because of crossing of RA boundaries, e.g., the basic HLR-VLR scheme
  - With dynamic, update or not depends on a user's call and mobility patterns
- Dynamic Location update schemes:
  - Time-Based, Movement-Based, Distance-Based

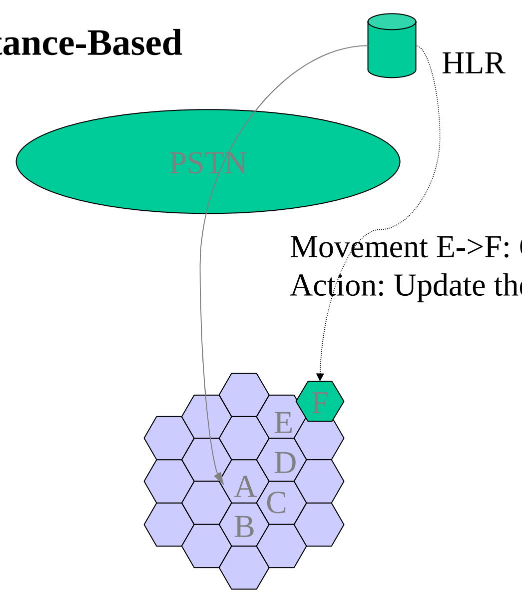
12

Dynamic Location Update Schemes

- Time-Based: A mobile terminal updates in every T time units
- Movement-Based: A mobile terminal counts the number of boundary crossings and performs the update when a threshold is exceeded (e.g. M=6)
  - Forwarding pointers can be considered as a variation of it
- Distance-Based: A mobile terminal tracks the distance D (in terms of RAs) it has moved since the last update
  - Update is performed when a distance threshold is exceeded
  - Mobile terminal needs some knowledge of the network topology
  - Local Anchor can be considered as a variation of it.

13

Distance-Based



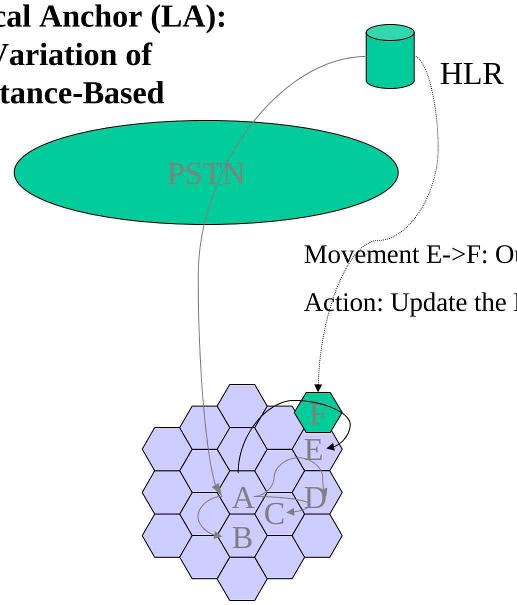
Movement E->F: Outside of 2-RA Distance
Action: Update the HLR

Movements A->B, B->C, C->D, D->E: within 2-RA Distance

Update Action: None

14

Local Anchor (LA): A Variation of Distance-Based



Movement E->F: Outside of a 2-RA Anchor Area

Action: Update the HLR

Local Movement: within the 2-RA Anchor Area

Update Action: Update the LA (VLR A)

Note the difference in the update action for a local movement here vs. a pure distance-based scheme

15

Update Time Interval for Time-Based Schemes

- T is the time interval for performing a location update
- Assume a search operation performs an expanding ring search
- Let λ be the call arrival rate and σ be the mobility rate. Then the maximum area to be searched is a circle of radius $\sigma * \min(1/\lambda, T)$ cells.
- Normalizing each update operation with a cost of 1 and each search operation with a cost of $\sigma * \min(1/\lambda, T)$, the cost of time-based management per unit time is:

$$C = \lambda * \sigma * \min(1/\lambda, T) + 1/T$$

- When $1/\lambda > T$, $C = \lambda * \sigma * T + 1/T$. Take $dC/dT=0$, $T_{opt}=1/\text{sqrt}(\sigma\lambda)$, implying that when either σ or λ increases, T_{opt} decreases

16

LeZi Update

- Based on a compression algorithm by Ziv and Lempel
- LeZi is a path-based update algorithm by which the movement history, not just the current location, is sent in an update message
  - The history has a list of zone (LA or cell) ID’s the mobile terminal has crossed
  - The location database keeps the history in compact form by means of a search tree structure
    - Can be part of the user’s profile
  - On a call arrival, selective paging is performed

17

Per-User Location Caching

- Lazy Cache Maintenance (Cached information is not updated on location update) vs. Eager Cache Maintenance
- Lazy Cache Maintenance: When is it beneficial to cache a callee’s location in a caller’s cell or a registration area?
  - Let CMR = λ/σ, representing the number of search operations between two consecutive update operations.
  - Let P_h be the cache hit ratio.
  - First search after an update operation will result in a cache miss, after which the remaining (CMR-1) search operations will result in cache hits. Thus P_h =(CMR-1)/CMR.
  - Let A_l be the local access cost and A_r be the remote access cost. Then $A_l + (1- P_h) A_r < A_r$, i.e., $P_h > A_l / A_r$, or, equivalently, CMR > $A_r /(A_r - A_l)$, meaning that caching is beneficial if user’s call-to-mobility-ratio > $A_r /(A_r - A_l)$.

18

Per-User Location Caching

- Eager Cache Maintenance: When is it beneficial to cache a callee's location in a caller's cell or a registration area?
  - A list of registrars is used to cache a user's location information
  - Let T be an observation period
  - Let λi be the average number of calls in cell i for the mobile user during T
  - Let σ be the number of location updates by the mobile user during T
  - Let α be the cost savings when a local lookup succeeds
  - Let β be the cost of updating a cache copy.
  - The location of the mobile user is cached at cell i only if the cost of savings outweighs the cost of location update, i.e.,

$$\alpha * \lambda i > \beta * \sigma$$

19

Eager Caching

- Working Set under eager caching: the working set of a mobile user *m* is the set of registrars (e.g., cells or RAs) that maintain location information about *m*
- A sliding window of length T is maintained by the system to estimate λi and σ for mobile user *m*
- When a new operation occurs, the working set membership is dynamically maintained as follows:
  - The operation is a search operation from registrar i: If registrar i is not in the working set and if $\alpha * \lambda i > \beta * \sigma$ is true, then add registrar i to the working set
  - The operation is an update operation: All the registrars in the working set are evaluated and if $\alpha * \lambda i > \beta * \sigma$ is not met, then delete registrar i from the working set.

20

Replicating Location Information

- A mobile user's location information may be replicated to a number of registrars for fault tolerance
- Two different organizations:
  - Flat: No structure exists among the registrars
  - Hierarchical: A multiple-level tree structure exists to organize location registrars

21

Replicating Location Information based on Flat Organization

- Consider using k replicas:
  - Placing k replicas at registrars i, (i+s) mod n, (i+2s) mod n, …, [i+(k-1)s] mod n where n is the total number of registrars and s = n/k.
  - What is the best value of k?
  - The update cost is k location registrars per update
  - The search cost is n/k location registrars accesses per search
  - The normalized overall cost per time unit is $C=k\sigma + (n/k)\lambda$, which is minimized when k_{opt}=sqrt(n*CMR)
  - o As CMR (i.e., λ/σ) increases, k_{opt} increases
  - o Search and update costs are proportional to sqrt(n) at k_{opt}

22

Replicating Location Information based on Hierarchical Organization

- A tree of location registrars:
  - A registrar that is a leaf node in the tree has information on all the mobile users in the associated RA
  - A non-leaf registrar replicates location information in all the location registrars in the subtree rooted to it.
  - The root registrar in the tree stores information on all the mobile users in the systems.

23

Replicating Location Information based on Hierarchical Organization

- Search: Let the callers be in RA_i and the callee be in RA_j. Let LCA(i,j) be the registrar that is the least common ancestor of LR_i and LR_j. The registrars along the path from the leaf registrar LR_i to LCA(i,j) will be searched until the callee information is found.
- Update: If a mobile user moves from RA_i to RA_j, then location information is deleted in all the registrars along the path from RA_i to LCA(i,j) (except LCA(i,j)), and the location information is updated in all the registrars along the path from root to RA_j.
- The cost of both the search and update is O(log n) where n is total number of registrars in the tree

24

Mobile Terminal Paging

- A process by which the network determines the exact location of a particular mobile terminal
- Polling cycle or search iteration:
  - Polling signals sent over a downlink control channel where the mobile terminal is likely to be
  - If a reply is received before a timeout, the polling ends; otherwise, a new group of cells is chosen
  - A call is dropped when the mobile terminal is not located within an allowable time constraint
  - "Maximum paging delay" is the maximum number of polling cycles allowed to locate a mobile terminal
- The Paging cost is proportional to the number of polling cycles as well as the number of cells polled in each cycle

25

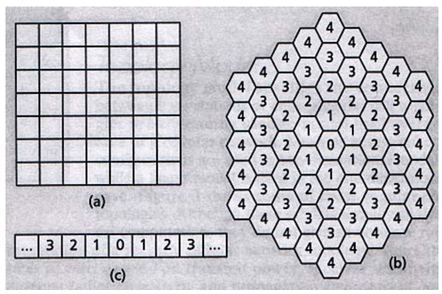
Terminal Paging

- Blanket Polling
  - All cells within an LA are polled at once when a call arrives.
  - The mobile terminal is located in 1 polling cycle
  - Currently deployed on top of LA-based update schemes in existing PCN's
  - Paging cost is high due to a large number of cells in an LA

26

Terminal Paging

- Shortest-Distance-First (Expanding ring search)
  - Starts at last known mobile terminal location
  - Moves outward in a shortest-distance first order.



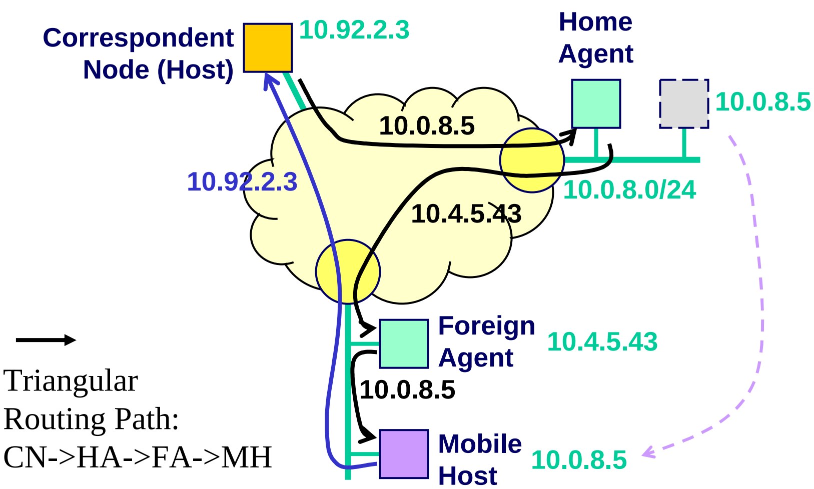
27

Terminal Paging

- Sequential Paging Based on a User's Location Probability
  - Current location is predicted based on its location probability distribution
  - A uniform location distribution gives the highest paging cost and delay
  - Groups of cells can be polled by selecting them with dynamic programming, when using a maximum paging delay constraint

28

Basic Mobile IP: Home Agent and Foreign Agent



Triangular
Routing Path:
CN->HA->FA->MH

29

Mobile IP Operation

- Mobile (foreign and home) agents advertise their availability using agent-advertisement messages
  - A mobile host may optionally solicit an agent-advertisement message
- A mobile host receives agent-advertisement message and decides if it is on a foreign or home network
- If the mobile host is on a foreign network, it obtains a care-of address on the foreign network
  - Foreign agent's IP address
  - Colocated care-of address statically or dynamically through DHCP (this is the only way in Mobile IPv6)

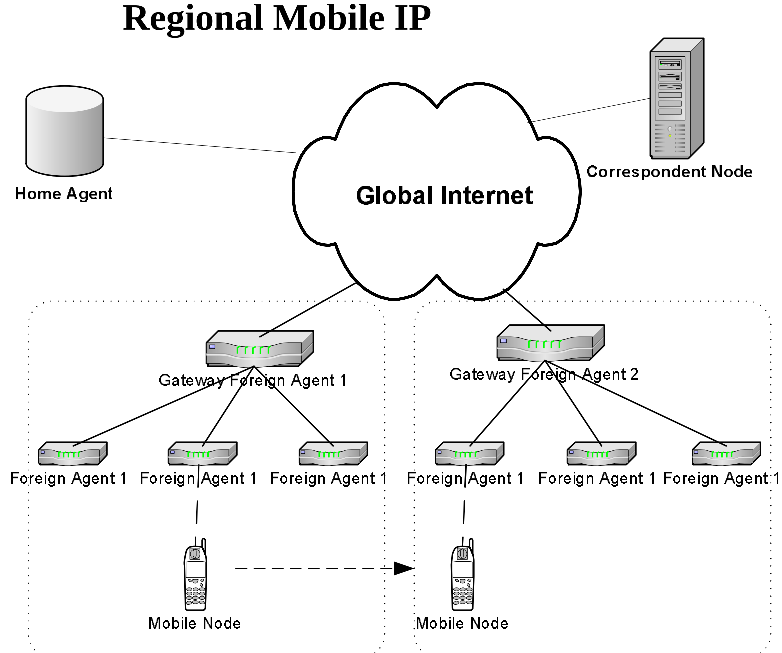
30

Mobile IP Operation

- Mobile host registers the new care-of address with home agent, possibly via a foreign agent
  - Registration request
  - Registration reply
- Home agent intercepts datagrams sent to the mobile node's home address and tunnels datagrams to the registered care-of address
- Tunneled datagrams could be received:
  - At foreign agent and delivered to the mobile host, or
  - Directly at the mobile node (colocated)
- Mobile host can send datagrams directly back to the correspondent node
- In Mobile IPv6, MH also registers new COA with CN which can communicate directly with MH without the overhead of triangular routing (called route optimization).

31

Regional Mobile IP



32

Summary

- Mobility management includes location management and handoff management
- Location management must explore the tradeoff between search and update costs based on each user's call (service) and mobility characteristics:
  - Time-based
  - Movement-based (forwarding is a variation of it)
  - Distance-based (local anchor is a variation of it)
- Caching and replication techniques can be used to provide search efficiency and fault tolerance but must be used with care not to dramatically increase update costs. In many cases, caching and replication must also base on each user's call (service) and mobility characteristics.
- Mobile IP supports mobility management in IP networks – dynamic location updates used for PCN can be applied (e.g., Regional Mobile IP vs. Local Anchor).

33

References

1. Chapter 2, F. Adelstein, S.K.S. Gupta, G.G. Richard III and L. Schwiebert, *Fundamentals of Mobile and Pervasive Computing*, McGraw Hill, 2005, ISBN: 0-07-141237-9.
2. I.R. Chen and B. Gu, "A comparative cost analysis of degradable location management algorithms in wireless networks," *The Computer Journal*, Vol. 45, No. 3, 2002, pp. 304-319.
3. V.W.S. Wong and V.C.M. Leung, "Location management for next-generation personal communications networks," *IEEE Network*, Vol. 14, No. 5, 2000, pp. 18 -24.
4. Lecture slides on Mobile IP, ECE/CS 4570, Virginia Tech, http://www.irean.vt.edu/courses/ececs4570/content2004/lecture10_mobileip.pdf.
5. Mobility Support in IPv6, RFC 3775, 2004, http://www.ietf.org/rfc/rfc3775.txt.

34