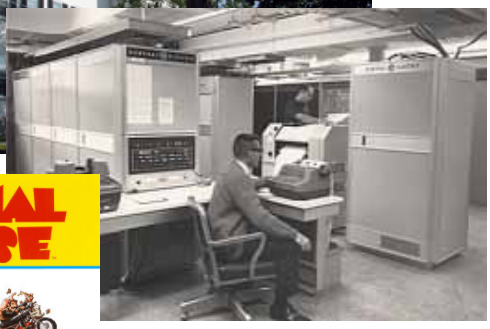
Wireless Network Monitoring Using Coordinated Sampling

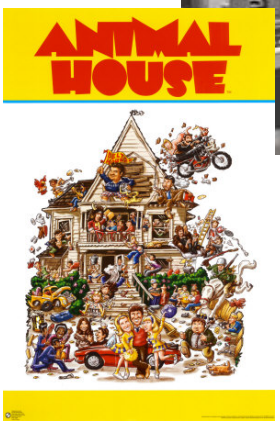
Chris McDonald (The University of Western Australia)
Udayan Deshpande and David Kotz (Dartmouth College)

Effective monitoring of wireless network traffic, using commodity hardware, is a challenging task due to the limitations of the hardware. IEEE 802.11 networks support multiple channels, and a wireless interface can monitor only a single channel at one time. Thus, capturing all frames passing an interface on all channels is an impossible task, and we need strategies to capture the most representative sample.

The competing goals of effective wireless monitoring are to capture as many frames as possible, while minimizing the number of those frames that are captured redundantly by more than one monitoring station. Both goals may be addressed with a sampling strategy that directs neighbouring monitoring stations to different channels during any period.

Dartmouth College (founded 1769)





1940 - the first remote access to a digital computer using phone lines, Dartmouth to Bell Labs New York.

1956 - the term artificial intelligence (AI) was coined by Dartmouth mathematician John McCarthy.

1964 – NSF funds the Dartmouth Time Sharing System and the development of computer language BASIC.

1982 - the College began implementation of X.25 international protocols for network data transmission.

1987 - the file-transfer program named Kermit.

1991 – all students required to own personal comp.

1996 - Intermapper software developed and released,

1997 - Foundation member of Internet-2.

2000 – ISTS – research & education for cybersecurity.

2001 - first Ivy League school to offer wireless Internet access on campus.

2004 – Newsweek - "Hottest for the Tech-Savvy."

2005 - Convergence of all phones, television, and data.

In New Hampshire, this is a tree

The MAP team

Dartmouth College – ISTS
(Institute for Security Technology Studies)

- David Kotz, Professor and PI
- Chris McDonald, adjunct Professor from Western Australia
- Bennet Vance, programmer
- Michael Locasto, postdoc
- Udayan Deshpande, Keren Tan, graduate students

University of Massachusetts Lowell

- Guanling Chen, Professor and Co-PI
- Bo Yan, graduate student

Aruba Networks

- Joshua Wright, Wi-Fi security expert

Supported by award NBCH2050002 from HSARPA, DHS Science and Technology Directorate

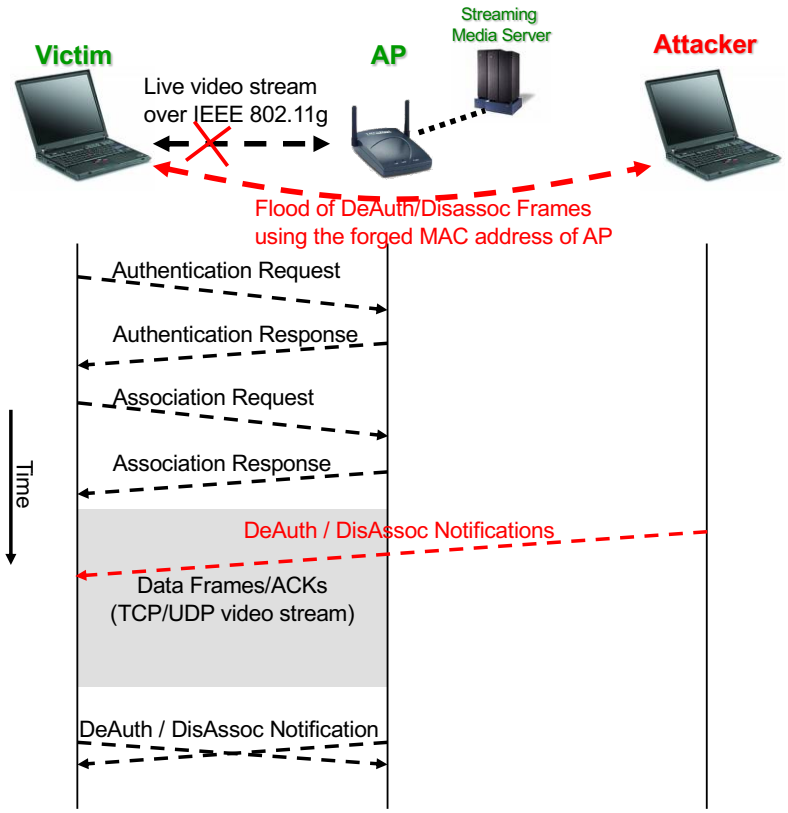
Wi-Fi security needed

- Wireless LANs becoming the dominant transport
  - Mission-critical, voice/video over wireless
    - VoWLAN $15B/yr by 2012 (Juniper07)
  - Fast moving area; new device and packet technologies
    - 802.11i, 802.11n, 802.11e, 802.16
  - presenting many new vulnerabilities

- Growing set of simple but effective attacks
  - Denial of Service (DoS) attacks, Reduction of Quality (RoQ) attacks, consuming excessive bandwidth, disrupting VoIP and video protocols
    - 160 entries in WVE.org database (as of April '08)

- Challenge
  - Capture all "over the air" 802.11 frames and analyze them [NSA guidelines for 802.11 wireless IDS, November 2005]
  - There are no wireless IDS systems capable of doing that today, particularly at the scale of a business, campus, town, or city.

Attack – DeAuth/DisAssoc Flood



- This attack belongs to
  - spoofing attacks
  - Denial-of-Service (DoS) attacks.
- Impact on video quality
  - is different on UDP and TCP based video
- MAP can detect this attack by observing
  - abnormally high rate of DeAuth/DisAssoc frames
  - sequence number gaps (anomalies).

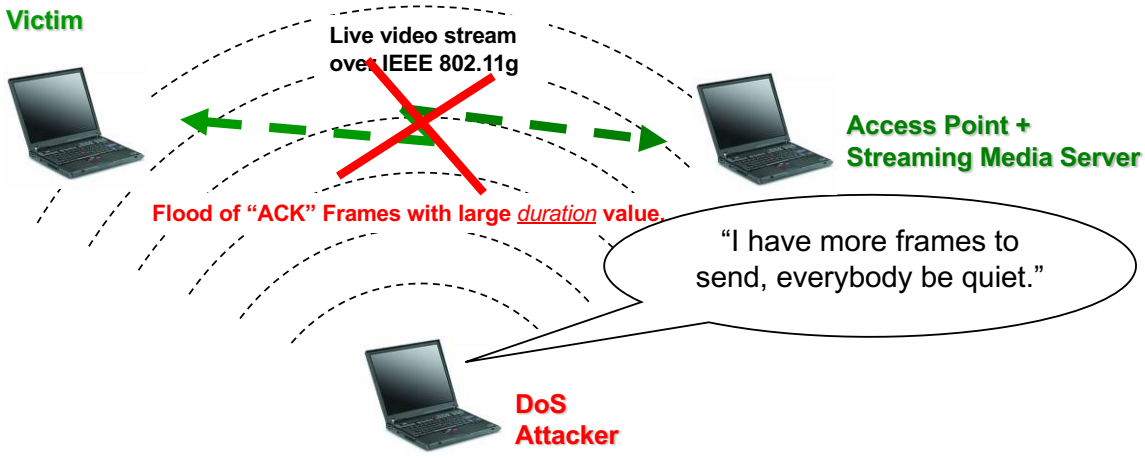
Attack – NAV Flood using ACK Frames

NAV - Network Allocation Vector: a register in each station, of the time periods it should not send frames.

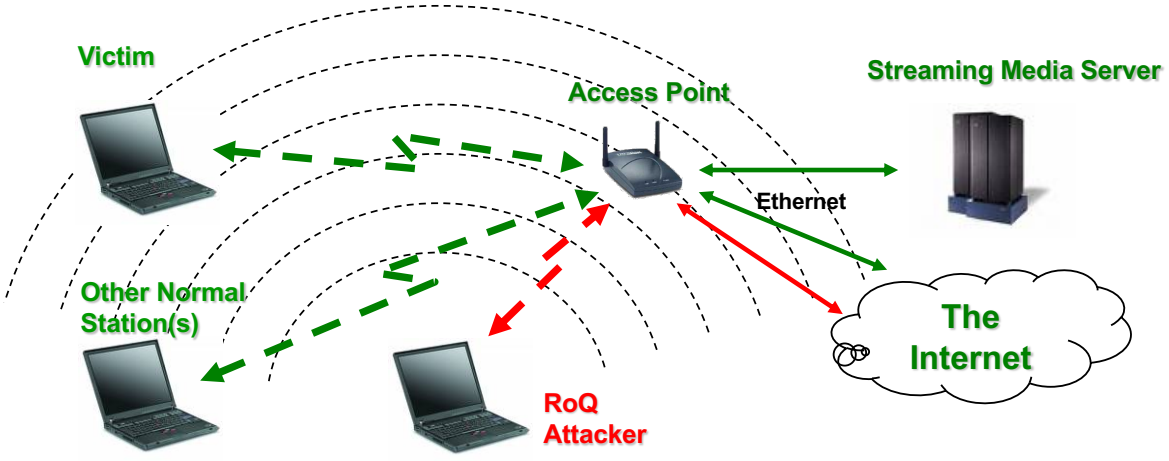
ACK Frame - a type of 802.11 control frames, its *duration* field is used to reserve the wireless medium.

This attack sends flood of “ACK” frames with large *duration* value.

- reserves the wireless medium without using it.

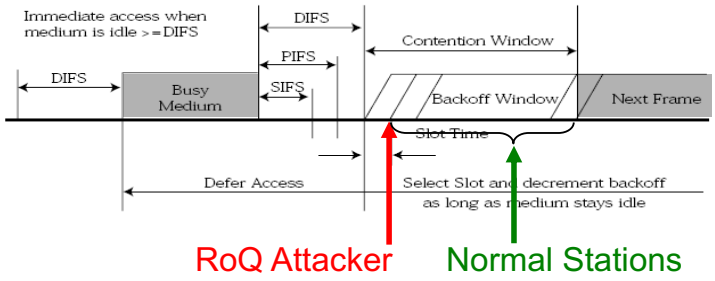


Attack – RoQ (Reduction of Quality)



A subtle attack targets the 802.11 DCF (Distributed Coordination Function), and is difficult to detect.

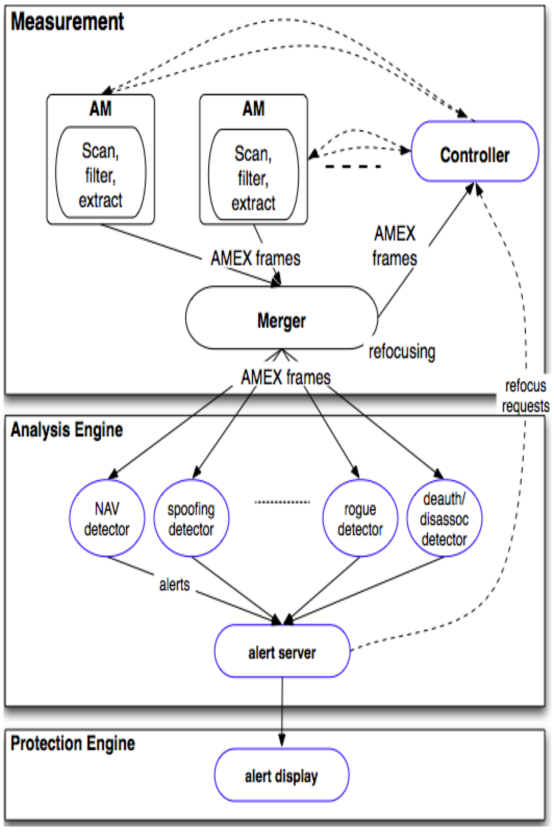
MAP detects this attack by looking at the rate of BEACON frames sent by the AP.

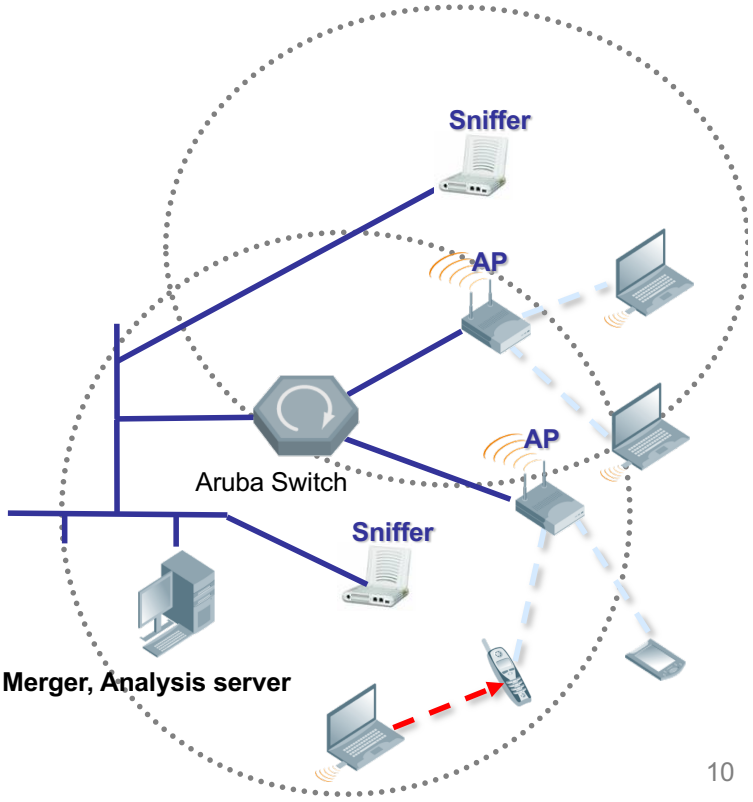


Wi-Fi network management

- "Help desk" support
  - Student reports trouble with connections
  - Need after-the-fact analysis of the network conditions in that location at that time.
- Locating areas of poor coverage
  - Proactively discover coverage problems
  - Examine PHY-layer and MAC-layer behaviour of clients in the region

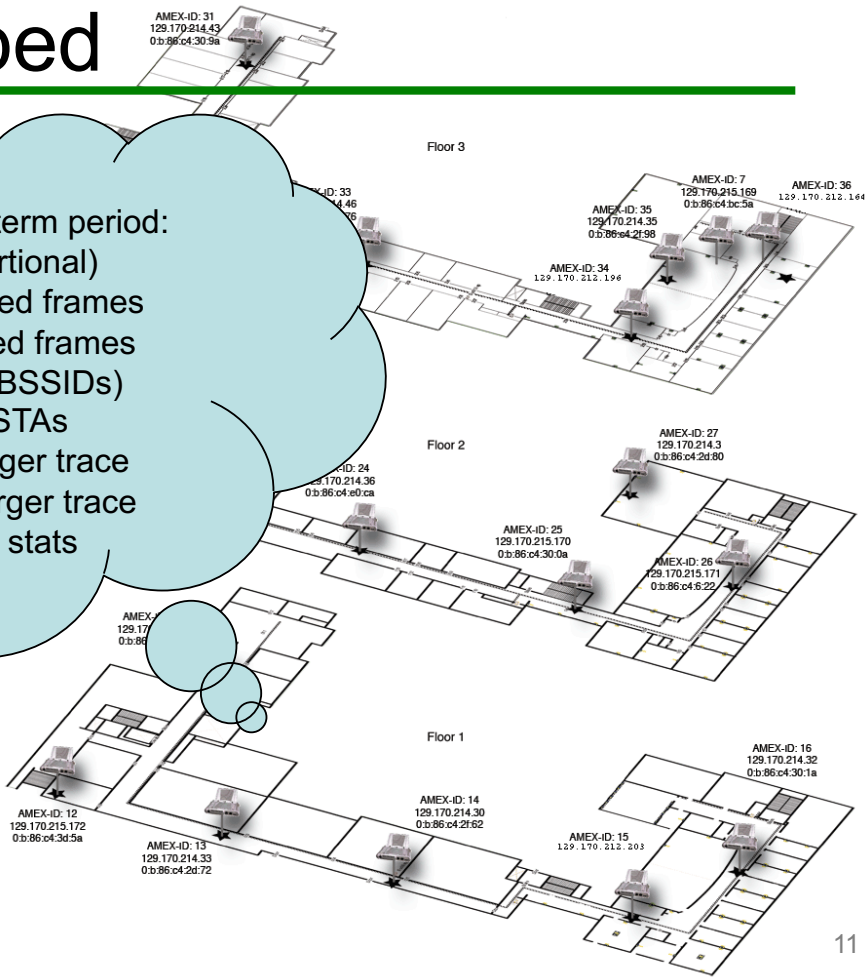
MAP Architecture





MAP testbed

In a typical 24hr in-term period:
(Normal-proportional)
317 million captured frames
161 million merged frames
98 distinct APs (BSSIDs)
696 distinct STAs
37.8 GB pre-merger trace
23.4 GB post-merger trace
approx. 1 GB stats



Sniffer nodes: Aruba AP70s

- Goals for deployment of the Air Monitors (AMs)
  - Coverage of wireless network
  - Must be aesthetically unobtrusive
  - Power over ethernet required
  - *Goals sometimes conflict*
  - *Undergone a detailed security audit*

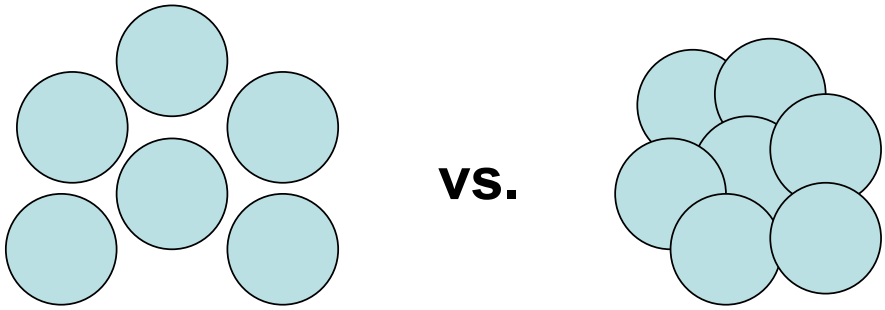
Tool: *MAPmaker*

- Global start/stop of sniffers, merger, etc.
- Independent concurrent instances
- Automates encryption and anonymization
- Systematic experimental record
  - Stores data in designated directory tree
  - Saves configuration snapshot, logs

Secure collection of traffic

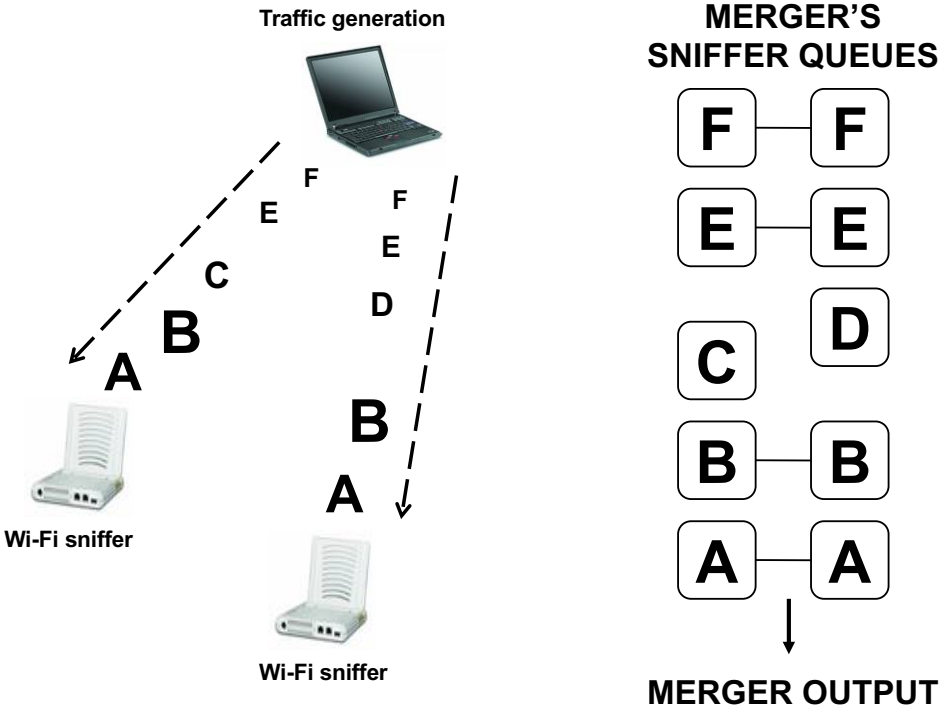
- Encrypt UDP-based traffic crossing the untrusted wired Ethernet between the AMs and server.
  - captured AMEX wireless frames,
  - commands and statistics
- We support
  - the NLSv2 stream cipher,
  - the AES Rijndael block cipher.
  - Additional algorithms may easily be added.

Density dilemma



- Sparse sniffers leave gaps
  - Traffic in gaps will be lost
- Dense sniffers give overlapping coverage
  - Traffic may be heard redundantly
  - Improves overall capture (but requires merging)

Merging Wi-Fi frames



Synchronization challenge

- Sniffer timestamps not reliable
- NTP synchronization inadequate
  - Resolution too coarse
  - Unpredictable discontinuities
- NIC timers accurate but jumpy
- Remedy: merger corrects timestamps
  - Uses common beacons as guideposts
  - Corrections propagate to unify all sniffers

Frame sampling challenges

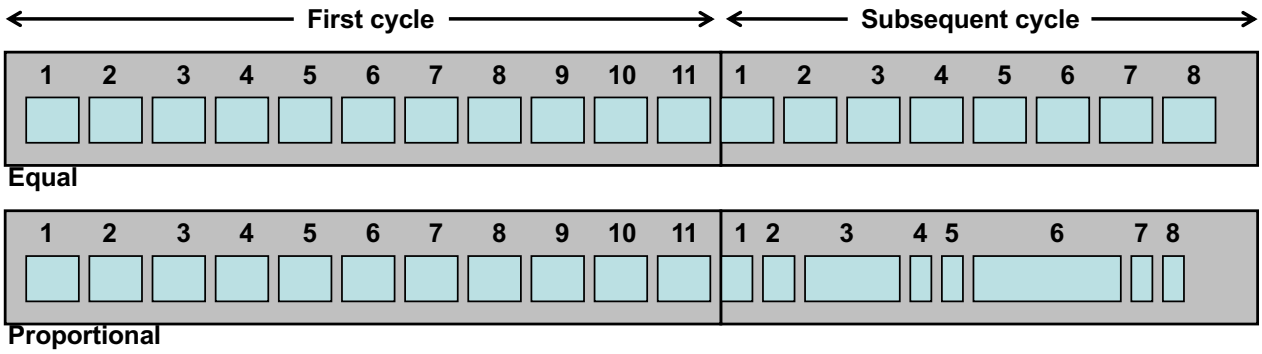
- IEEE 802.11 networks support multiple channels, but a wireless interface can monitor only a single channel at once.
- Changing channels takes (randomly) 5-70msec, during which frames cannot be captured.

Frame sampling strategies

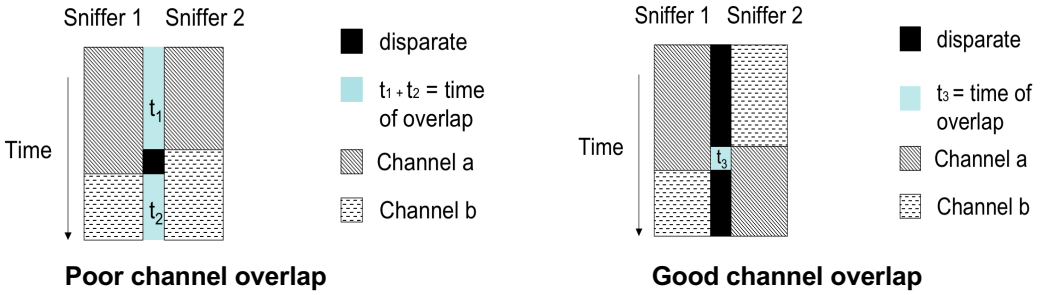
- Goal: capture a *representative* sample.
- A simple taxonomy of *sampling strategies*:
  - Random channel sampling
  - Equal time on each channel
  - Proportional time on each channel
  - Coordinate the activities of each AM so as to maximize the likelihood of hearing desired traffic
- Minimize *redundant* or *unnecessary* effort
- Maximize number of *unique* frames captured

Channel sampling strategies

- Per-sniffer (local) strategies
  - Equal channel sampling
  - Proportional channel sampling



Problems with simplistic sampling



Poor channel overlap

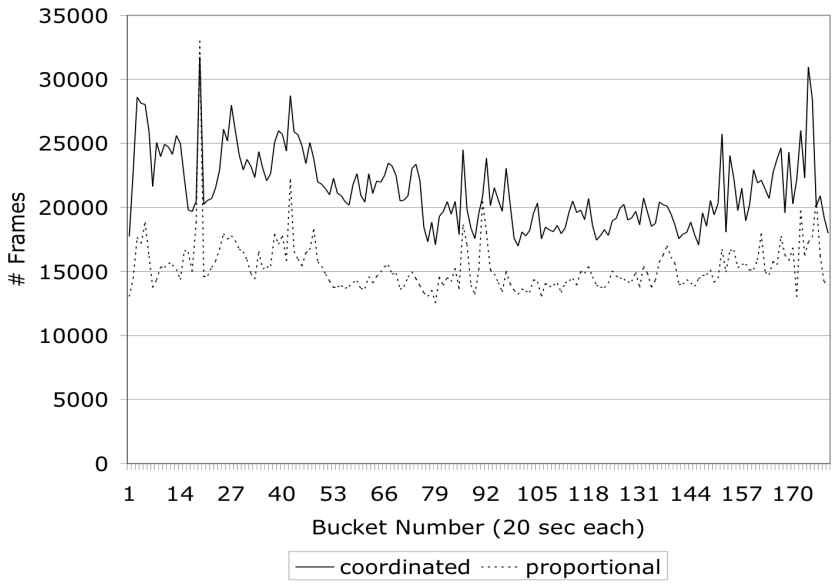
Good channel overlap

Our hypothesis is that scheduling the channels on AMs, such that the coverage includes minimal overlap, should result in even greater unique frame capture.

Coordinated sampling

- Using the *merger's* stream of unique frame information, the *controller* builds a *neighbour graph* recording which sniffers recently saw the same frames.
- The *controller* employs simulated annealing to shuffle sniffer sampling schedules to reduce the overlap.

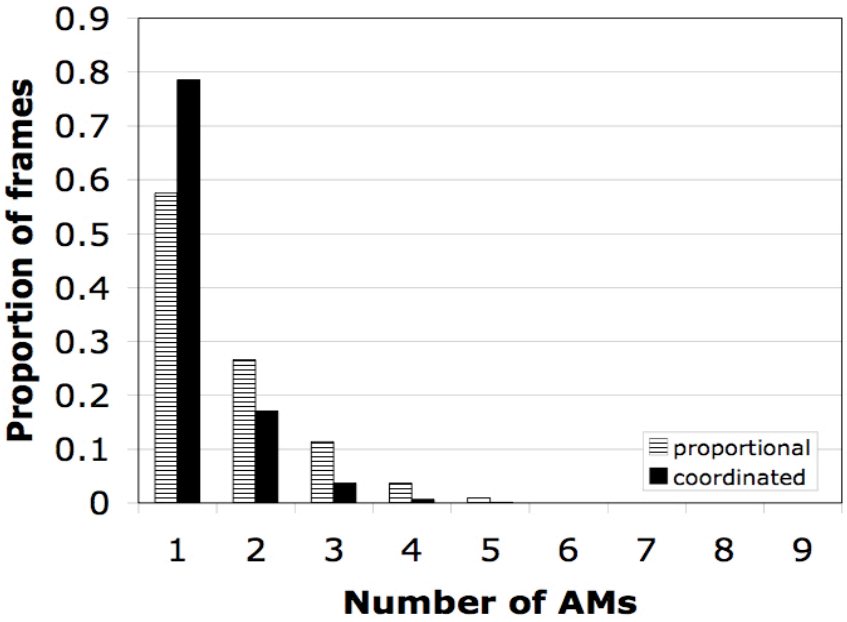
Unique frame capture



"There are people who will commit unspeakable acts for another ten percent"
– John Mashey, founder of MIPS.

ICON 2007: *"Coordinated Sampling to Improve the Efficiency of Wireless Network Monitoring"*

Redundant frame capture



78% of frames captured using coordinated sampling are unique, compared with only 58% of those using proportional sampling

Refocusing on traffic types

"Refocusing" allows the analysis engine – or sysadmin – to ask the measurement system to focus more effort on a particular kind of traffic.

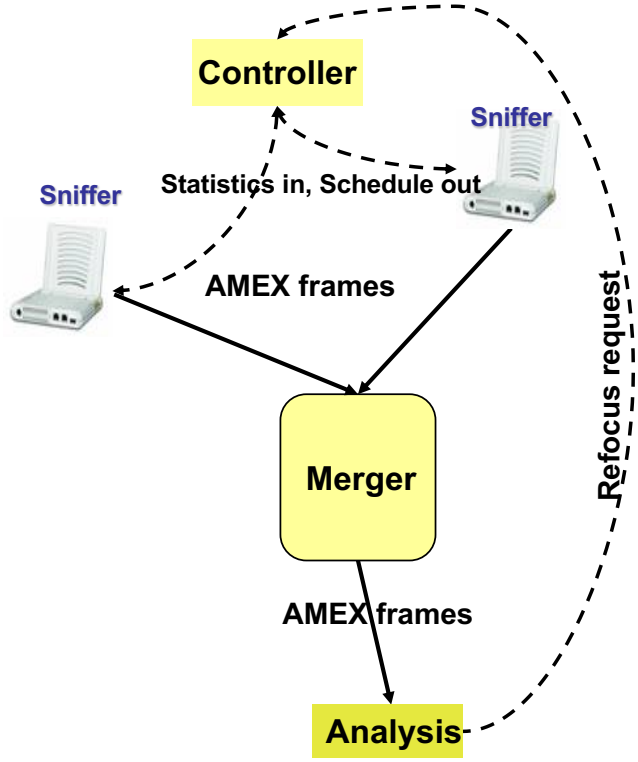
To refocus, we define a predicate, e.g.,

```
"src == 00:16:cb:b7:18:82"
```

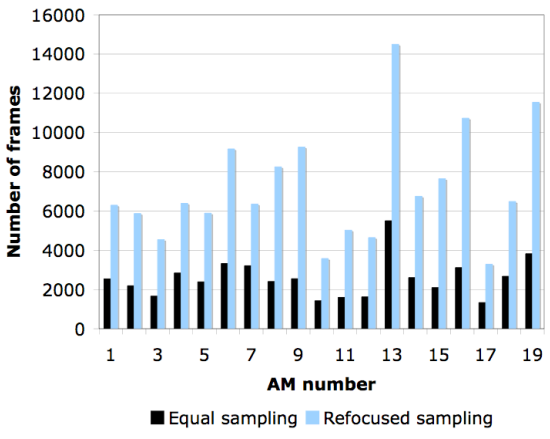
And run the sniffer policy:

```
set cyclelen 3000ms
run -c 1-5 -p npredicate
```

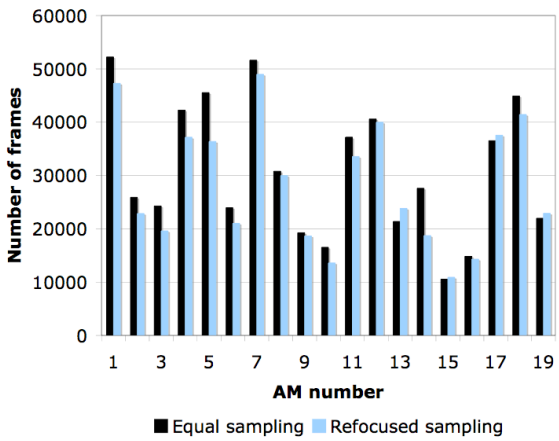
This policy will capture traffic on channels 1 to 5 in proportion to the number of frames matching the predicate.



Refocusing results



Matching frames



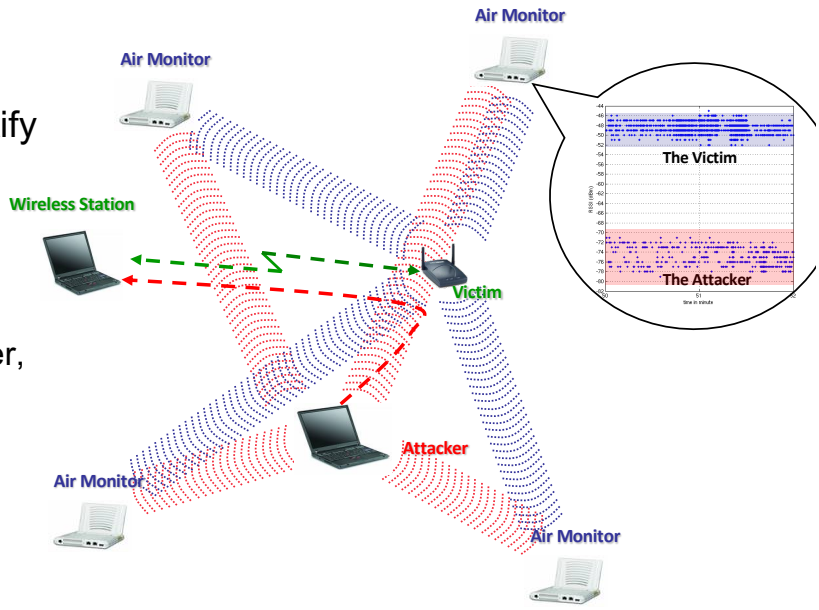
Non-matching frames

Improved focus without losing baseline capture

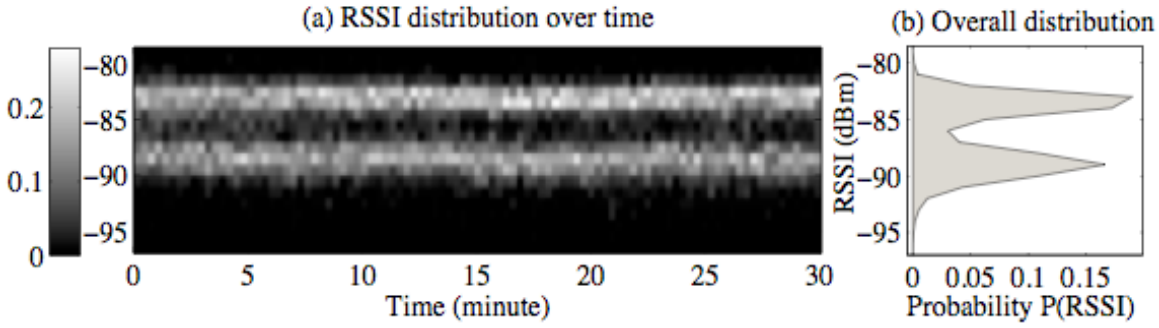
PAM 2008: *"Refocusing in 802.11 Wireless Measurement"*

Spoofing detection using RSSI (Received Signal Strength Indication)

- Spoofing is the foundation of many 802.11 MAC layer attacks
- RSSI can be used to identify "senders"
  - The RSSI patterns are statistically identifiable
  - RSSI is difficult to forge
  - Works with a single sniffer, and even better with multiple sniffers.



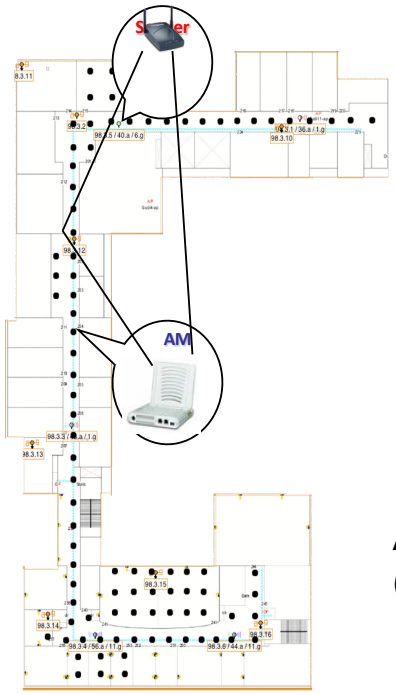
Multi-Modal RSS patterns

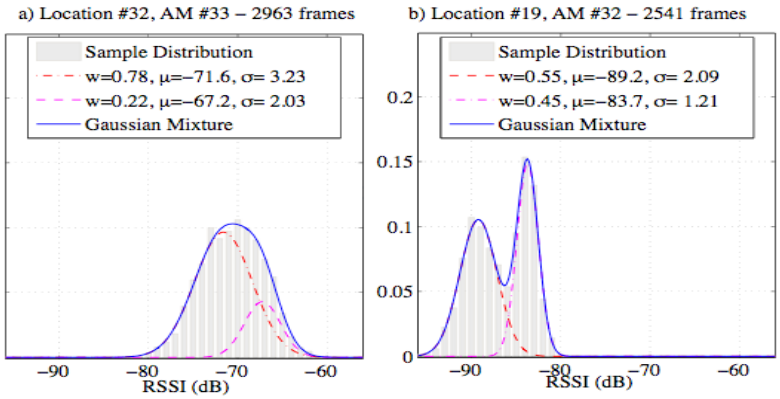


An example of multi-modal RSS distribution from one wireless sender

The above plot suggests the received RSS comes from two active and stable sources, due to "Antenna Diversity" and multi-path propagation.

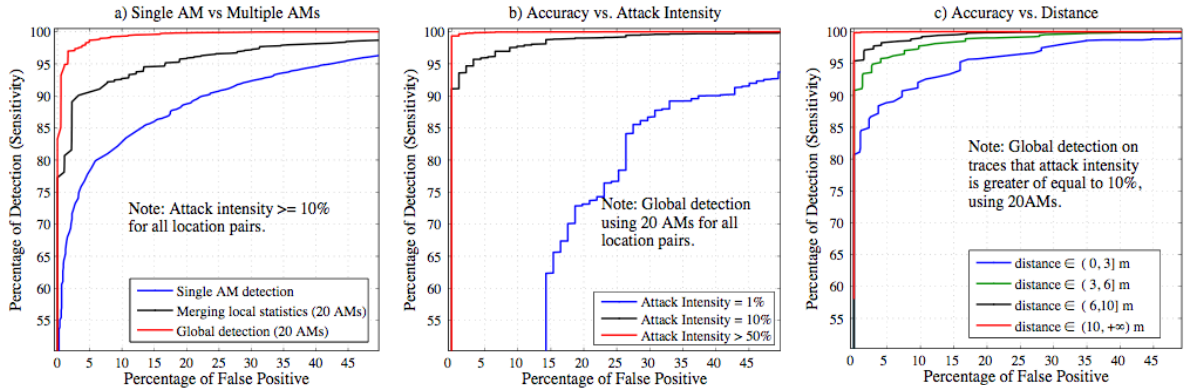
Multi-Modal RSS pattern profiling





An example to show how a Gaussian Mixture Model (GMM) profiles multi-modal RSS patterns well

Detection results

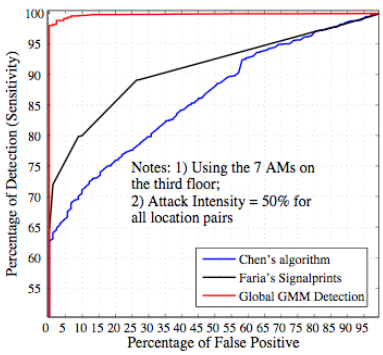


By matching the RSS pattern to pre-obtained profiles, any mis-matches will be reported as “spoofing”.

The best-performing algorithm is global detection based on the frame-by-frame merging from multiple AMs.

Our algorithms out-perform other leading techniques.

Future work: mobile stations.

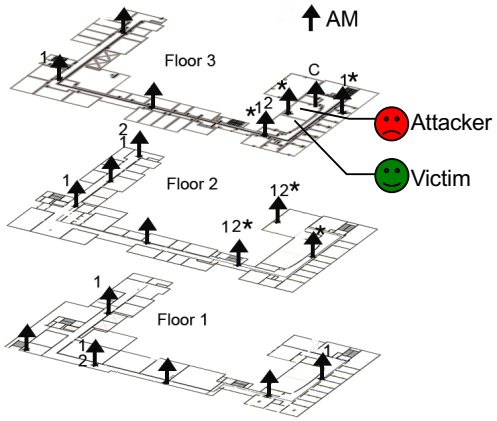


INFOCOM 2008: *"Detecting 802.11 MAC Layer Spoofing Using Received Signal Strength"*

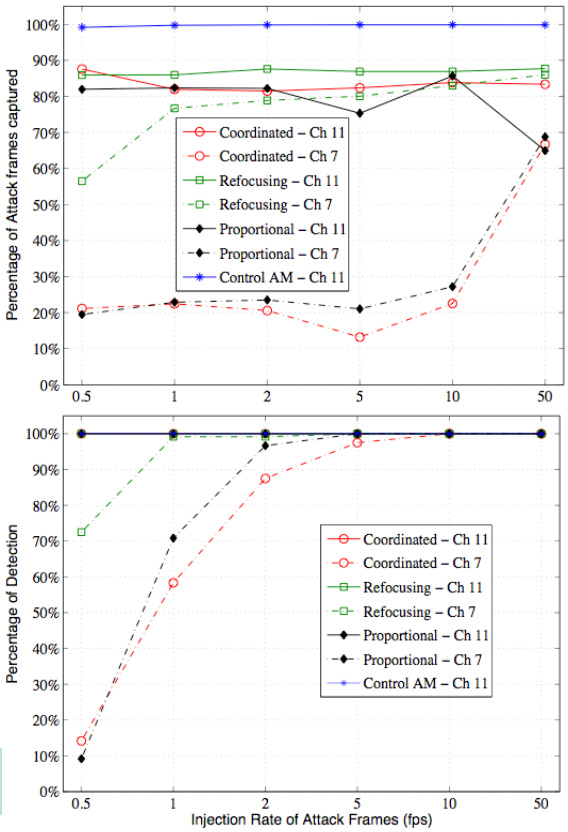
Effectiveness of detection

At each location marked "*", we deployed
- Group 1: 6 AMs, local proportional;
- Group 2: 6 AMs, coordinated proportional;
- Group 3: 6 AMs, coordinated + refocusing;

and
- A controler AM, marked "C", no sampling;
- An attacker injects DeAuth/DisAssoc Frames
  - onto channel 7 (idlest) and 11 (busiest)
  - at various injecting rate 0.5 ~ 50 Fps



Refocusing is an effective strategy to improve detection accuracy.



Layer-3 rogue AP detection

- Existing commercial solutions cannot detect protected layer-3 rogue APs
  - Network Computing survey, June 2006
  - Layer-3 APs = off-the-shelf wireless routers
- New approach based on
  - Wired "verifier" listens on packet streams or transport layer netflow records exported by routers/switches
  - Verifier sends forged test packets
  - Wireless sniffer picks up these special test packets
  - 100% accurate in lab environment
  - Works for all 802.11a/b/g APs
  - Works even if the packets are encrypted by AP

Active fingerprinting

- Inject non-standard or malformed frames to target device
  - FromDS and ToDS bits are expected to be cleared in Probe/ Authentication Request Requests
  - Probe/Authentication Requests are not supposed to be fragmented
  - Responses to Probe Requests with other Frame Control bits set (in particular, More Fragments, Power Management, More Data, and Order bits) differed between APs.
- Classify devices based on response behaviours

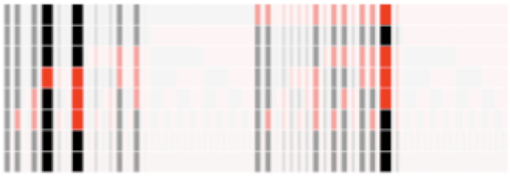


Figure 1: Cisco-Linksys WRT54g *AuthFCTest*

Figure 3: Madwifi-ng soft AP *AuthFCTest*

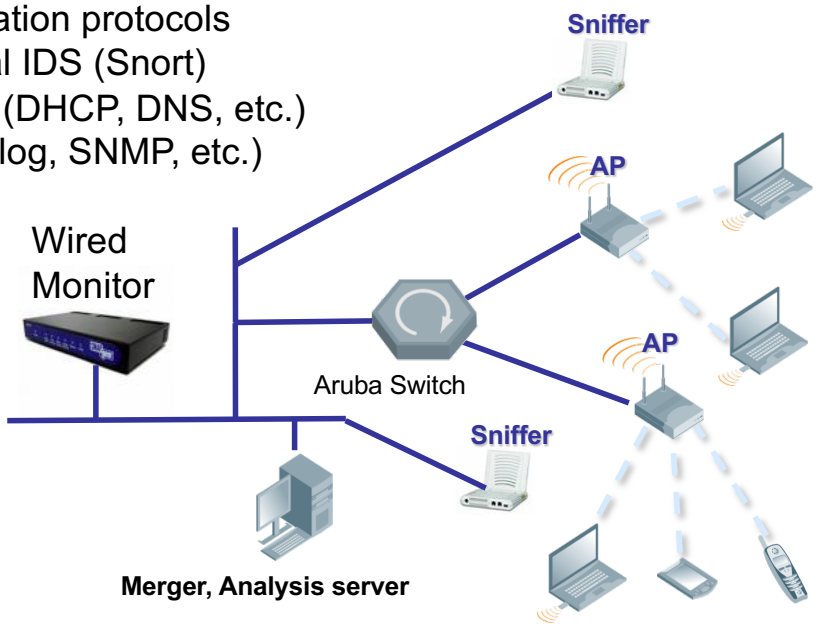
WiSec 2008: *"Active Behavioral Fingerprinting of Wireless Devices"*

Voice over WLAN Diagnosis

- Emerging VoWLAN applications
  - Projected to be a $15B market by 2012 (Juniper Research 2007)
  - Sensitive to QoS and vulnerable to various DoS/RoQ attacks
- Automated diagnosis when QoS degrades
  - Link MAP MAC-layer IDS output with application-level measurements
  - Both malicious attacks and benign faults may degrade QoS
  - Need automatic fault classification and localization
  - Very few existing solutions address this problem
    - VeriWave (offline method, no IDS functionalities), AirMagnet (client-only, limited)
- Our new approach
  - Cross-layer correlation of both wired and wireless network measurements
    - Isolate problems of AP, wireless channel, DHCP, DNS, SIP, authentication, etc.
  - Two-stage analysis
    - Lightweight state-machine analyzer finds anomalous traffic events
    - High-level model-based event correlation engine for auto diagnosis
  - Works directly on APs or as an appliance attached to WLAN switches
- Status
  - Building prototype

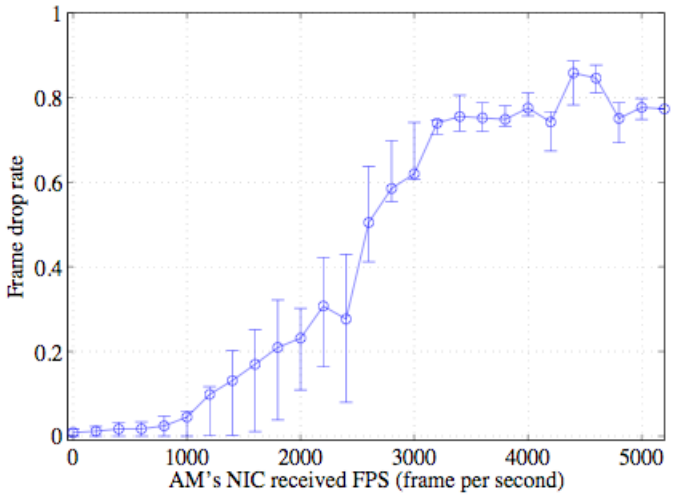
Multi-Source Aggregation

-Higher layers, such as VoIP
-Wireless authentication protocols
-Correlate traditional IDS (Snort)
-Other components (DHCP, DNS, etc.)
-Other sources (syslog, SNMP, etc.)



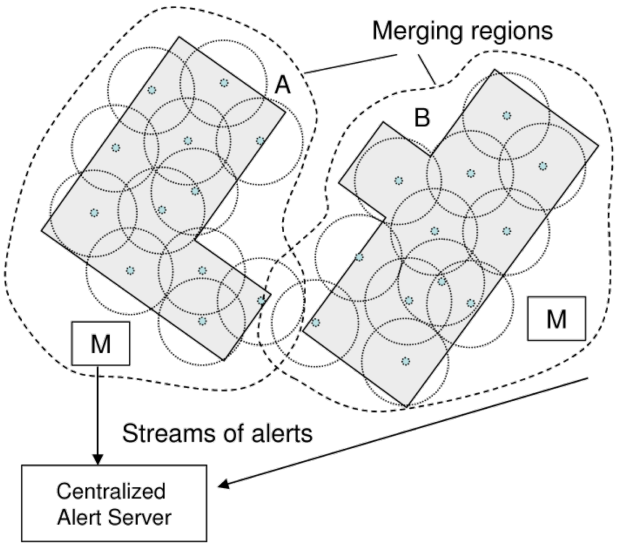
Performance of sniffing

Our AMs (Aruba AP70)
- 266M-MIPS32 CPU
- 32MB DRAM
- 2xAtheros 5212 NIC (a/b/g)
- 2x100MB Ethernet NIC
- OpenWRT
- Madwifi Driver
- Dingo and AMEX libraries



AMs drop frames
- Resource-constrained devices.
- Running complicated analysis software on AMs – unrealistic.

Scaling the MAP architecture



M MAP (including merger and detection engine)

Coverage of a single AM

Although our merger should scale well, not to a whole campus!

Need concept of 'Merging regions'.

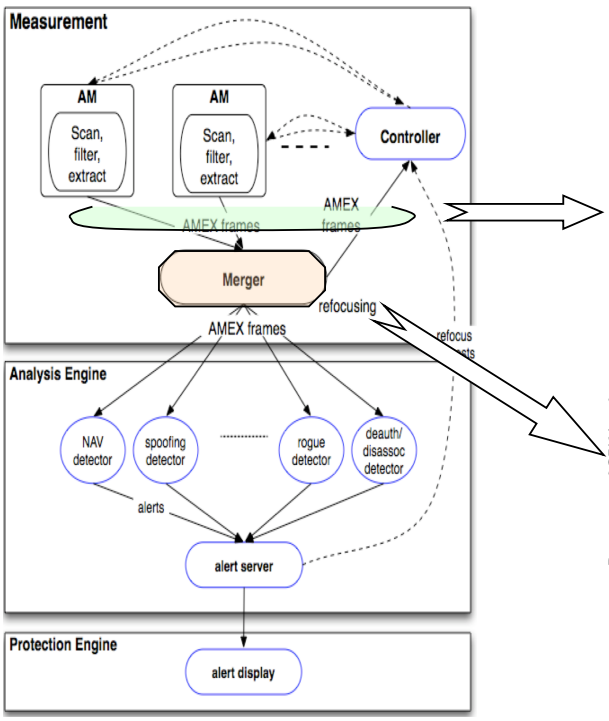
Merging regions may overlap.

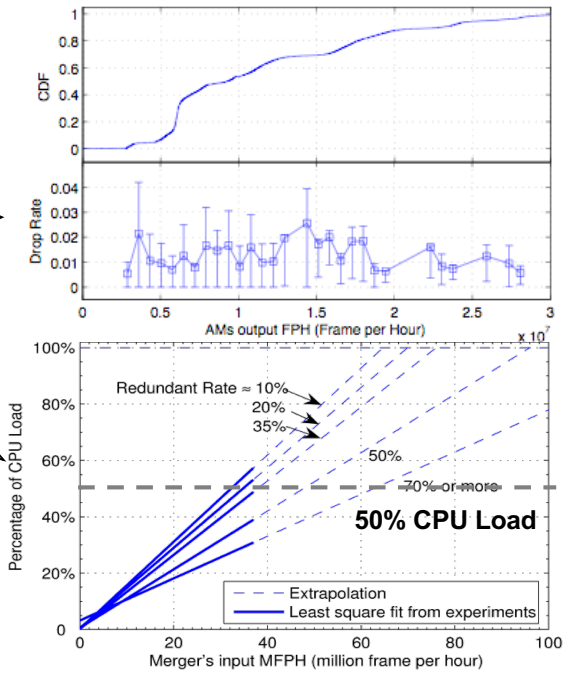
We depend on a centralized alert server to remove duplicated alerts.

The overlapped region can be minimized by careful selection of merging regions and AM deployment.

We'd like to know the maximum region a merger can cover.

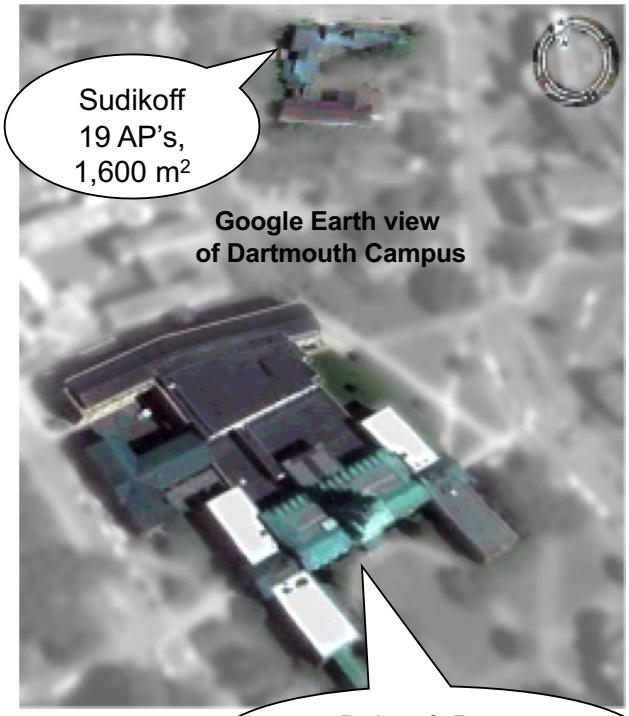
What's the bottleneck?

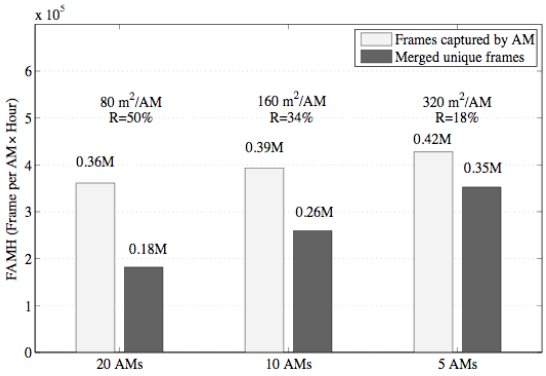




The bandwidth is not the bottleneck. We believe it should be the CPU load for merging frames.

Capacity at scale



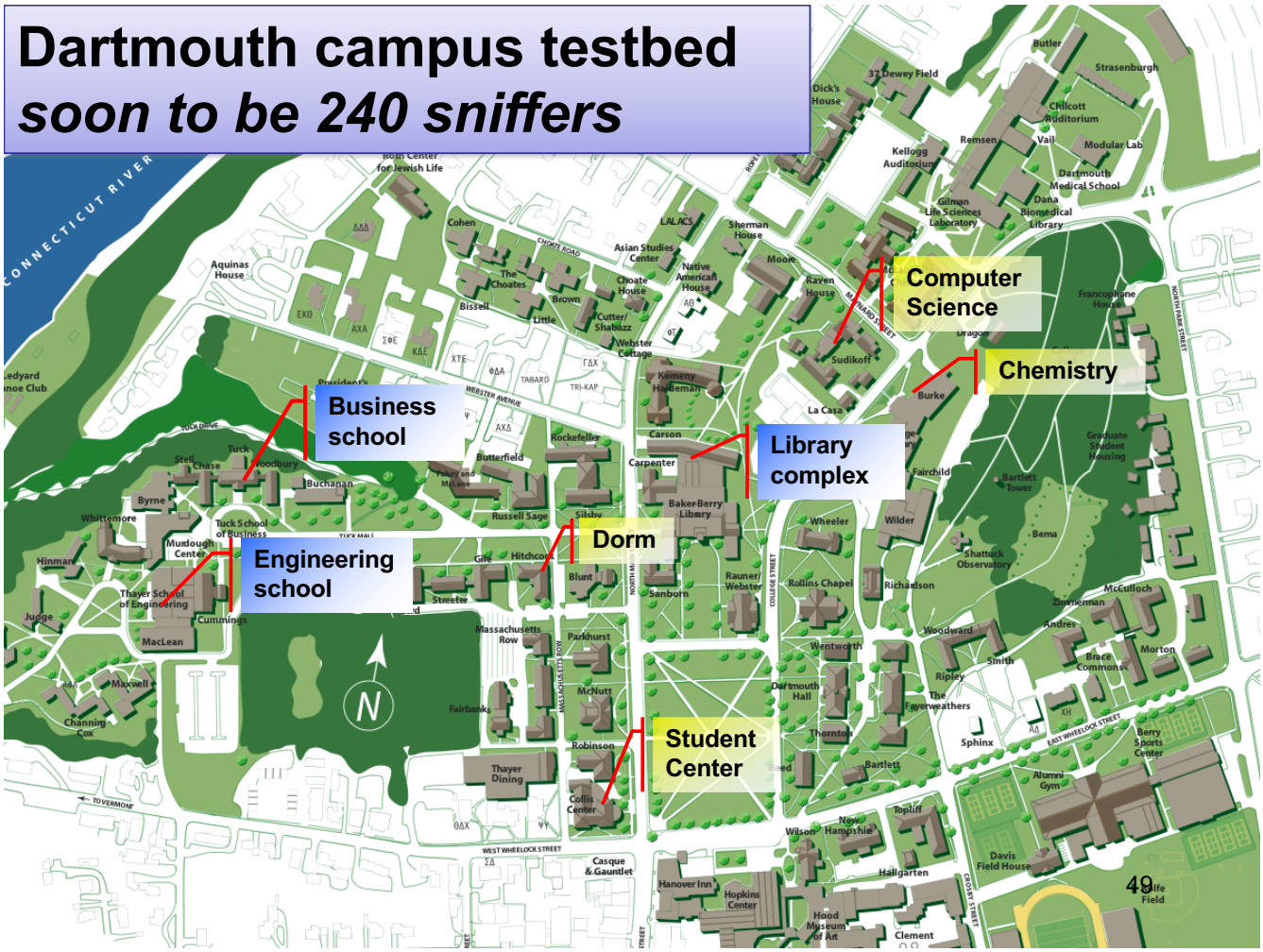


Frame captured by AM under various AM deployment density.

R denotes the rate of redundancy.

Desired AM density (m^2/AM)	80	160	320
AP:AM ratio	1:1	2:1	4:1
Projected max coverage at 50% CPU load			
Max AMs	117	95	80
Max area covered (m^2)	9,360	15,200	25,600

Dartmouth campus testbed
soon to be 240 sniffers



MAP benefits and summary

- Detailed view of wireless traffic
  - *sniffers* capture MAC-layer headers
    - not available at wired monitors
  - *merger* produces a more complete packet stream
    - each sniffer has limited range (view)
    - each sniffer may lose frames
- Flexible controller architecture
  - Different channel sniffing *strategies*
  - Dynamic and manual *refocusing* capability
- Near real-time traffic analysis
- Many spin-off security and monitoring projects