Visual

- How well can one extract PIN numbers via post-use infra-red imagery? (or anything else fun w/IR)
- Privacy issues related to GoogleEarth or StreetView data? (e.g., finding building violations, identify when people present)
- Automated extraction of latent fingerprints (e.g., via UV photography)
- Video-conferencing filter that obscures identity but looks natural? (autotune for faces)
- Identifying individuals in crowd scenes via non-traditional cues (e.g. bad of the head recognition)
- Watermarking 3D models
- Meaningful visualization of security data (e.g., spam, net, etc)
- Seeing through privacy glass or 3M privacy screens?
- Fingerprint capture at a distance via live video
- Automated mapping of spy satellites

Privacy

- How well can you reconstruct Web logs from Netflow data?
- Study of exploitable PII exposure on flickr, myspace, etc (e.g. credit cards, check routing numbers, SSNs, name/address)
- Javascript code to track fine-grain user interaction with a page (mouse movement, key timing, etc)
- Check which sites you’ve visited recently via DNS lookup timing (was entry in cache?)
- Build a Flash “scrubber” that integrates with Tor to prevent flash-based information leakage
- Privacy analysis of chatroulette (i.e., can you tell who these people are?)
- Tie FAA flight database with network log data to infer which users are travelling and where they came from
- What does your TiVO box know about you?

Malware

- Malware phylogeny
  - Fingerprint malware "family trees" based on the data structures they use, i.e., to determine which variants derive from which earlier specimens
- Build system to identify the kinds of information being targeted by different kinds of malware
- Evaluate malware delivery vectors
  - P2P malware vs web sites vs attachments, etc... are they all carrying the same malware or different?
- Evaluate time-to-detect for commercial malware
- Build IDA plug-in to locate particular “kind” of code in binary (e.g., AES code, CRC code, etc) [could also be a machine-learning project]

Network security

- Differential network hygiene (who fixes infected machines well, who doesn’t?)
- Infer AV updates/patch use from netflow data and see if it correlates with infection outcomes
- DDoS response (use UCSD telescope data to infer how long before ISP “blackhole”’s site)
- Explore ways of inferring the number of visitors to a remote site (e.g. ip_id trick, DNS caching, etc)
Spam/E-Crime

- How much higher is the delivery rate for “laundered” spam (e.g., via Web-mail)?
- How serious is the problem of blog spam? What might be done to detect/characterize it?
- To what degree can you assess the accuracy of blacklist feeds, how effective they are and how redundant?
- We have lots of spam feeds. How skewed are they and how much do they overlap?
- Automatically extract pricing data from UCSD Trajectory database to track price of spam-advertised goods
- Use NLP to track good/service pricing on underground forums/IRC
- Relate use of domain names in various scams to price of domain offered by registrar

Machine learning

- Predict which code changes will produce software vulnerabilities
- Repeat Ma’s on-line URL classification study using Web page content
- Build classifier to predict machine compromise based on what sites you visit
- Apply receiver-reputation idea to Web visits (reputation of sites depends on who visits them)
- Can we predict which Torrents are “bad”?
- Classify spam senders (which program used)
- Clustering of malware families based on behavioral features

More random ideas

- Detection of Bots in MMPORGs
- Automation for “attack surface” estimation
- Analysis of Taser authentication
- Fuzz testing of embedded devices (e.g. Cameras)
- Location verification via “audio-print” (indeed, any way of proving location)
- Analysis of on-line poker (fair deal or not?)
- AppEngine Cartography (repeat our Amazon study on AppEngine)
- Automated safe cracker
- Use new Intel security features for something interesting (e.g., NotABot)
- Hardware support for self-destructing data
- Hardware support for information flow tracking
- Detecting pirated hardware IP (e.g., mp3 or PCI blocks) via unique side-effects
- How does a cloud provider prove that they’ve replicated your data?
- How hard to reliably erase flash storage?
- Security analysis of campus power grid
- Security vulnerabilities the Kindle?
- Repeat Ozment/Schechter’s Milk/Wine study on vulnerability generation w/another system
- Evaluate how people respond to different kinds of warning messages
- Explore use of ABE for real systems problem
- Difficulty in spoofing consumer GPS
- Are there vulnerabilities in Digital FM radio?
- Identify “anomalous” file contents to mitigate file format vulnerabilities (esp Flash, QuickTime and PDF)
- Attacks against smart batteries (drain beyond ability to recharge or make explode)
- Driver detection (infer identity of driver via driving behavior)
• Build a javascript or Flash reference monitor
• Explore use of differential privacy to protect data for interesting network or security trace analysis problem (e.g., pick any of George Varghese's recent papers and see if it can be done with DP)
• Build an interactive biometric system (e.g., proof of presence via eye-tracking) to prevent simple replay attacks
• Design a CAPTCHA that is difficult to outsource to low minimum-wage solvers
• Build a system to fingerprint physical luxury goods (e.g., leather, etc) ala the Princeton work on paper
• SMART disks will move data from failing sectors to spare sectors. Consequently the data on these failing sectors may not be erased when the associated data is erased. Explore if this actually happens and the correct way to erase a disk.