2nd USENIX Workshop on Health Security and Privacy (HealthSec '11)

August 9, 2011
San Francisco, CA

Keynote Address

Joy Pritts, Chief Privacy Officer, Office of the National Coordinator for Health Information Technology

No report is available for this session.

Short Papers

Summarized by Ben Ransford (ransford@cs.umass.edu)

Implantable Medical Device Communication Security: Pattern vs. Signal Encryption

Fei Hu and Qi Hao, University of Alabama; Marcin Lukowiak, Rochester Institute of Technology

Fei Hu discussed his group's "cyber-physical approach" to the security of implantable medical devices (IMDs). His group has built body area networks (BANs) based on sensor motes

and RFID readers. The BANs are structured as peer-to-peer networks whose trust relationships exhibit a ring structure. Hu described an assortment of attacks on IMDs and proposed preliminary defenses. He concluded with a description of "intentional signal entanglement," a mechanism by which an external device could destructively interfere with an IMD's traffic to hide private data from eavesdroppers.

Raj Rajagopalan asked to what degree intentional signal entanglement would depend on the signaling protocol the IMD uses. Hu responded that it was a physical-layer technique that would work independent of higher-level protocols.

Exposing Privacy Concerns in mHealth Data Sharing

Aarathi Prasad, Jacob Sorber, Timothy Stablein, Denise Anthony, and David Kotz, Dartmouth College—Institute for Security, Technology, and Society

Aarathi Prasad presented the preliminary results of a study on patients' privacy concerns with respect to data collected using mobile health (mHealth) devices. They conducted focus groups with patients of all ages in order to learn what people saw as the benefits and drawbacks of electronic health records (EHRs) and mHealth. In the context of a "typical" mHealth architecture, in which patients upload data to their EHRs for sharing with caregivers and family, the authors found that patients wanted the ability to turn mHealth devices on and off and to control the release of the collected data. They found that people perceived a variety of privacy risks, with diet and exercise information considered least sensitive and social interactions the most sensitive. Some patients did not understand why data such as heart rate would be considered sensitive. Patients felt more comfortable sharing data with caregivers than with their friends, families, or insurance companies.

Prasad concluded with several suggestions for mHealth architects. First, privacy controls should have access logs, and changes to privacy settings should be logged. Second, mHealth data should be annotated to aid patient understanding. Third, mHealth devices should have sensible, conservative default privacy settings, because users are unlikely to change them. Fourth, mHealth data can be presented and privacy-controlled in a hierarchical manner that matches patients' mental models.

An audience member asked whether the authors studied the effect of monetary incentives on patients' willingness to divulge data; Prasad said they had not. Raj Rajagopalan noted that privacy decisions can be context-sensitive; Prasad agreed and noted for an example that patients perceive the privacy risks of continuous versus periodic monitoring differently. Another audience member asked whether the authors told the patients how their mHealth devices worked; Prasad said they allowed patients to form their own opinions but were allowed to ask questions. Jim O'Leary asked about the effect of patient age on perceived privacy risks, considering that younger people tend to be better connected on online social networks; Prasad acknowledged that although there were clear differences between age groups, everyone in the study seemed to know about social networks.

Persistent Security, Privacy, and Governance for Healthcare Information

W. Knox Carey, Jarl Nilsson, and Steve Mitchell, Intertrust Technologies

Knox Carey pointed out that healthcare data is not flowing as easily as it should. Despite technological advances and huge investments, healthcare systems lack interoperability standards. Different organizations exhibit mismatching policies that hinder data sharing; enforcing policies across organizations is a nightmare. The current situation, Carey said, incentivizes wrong behavior such as data hoarding.

The authors propose a DRM-like approach to healthcare data, with data being encrypted at the source and persistent policies attached (and governmentally enforced). They propose associating healthcare data with sets of well-defined computations that result in different views of patient data for different interested parties, such as patients, doctors, and insurance companies.

An audience member likened the authors' proposal to a fine-grained informed-consent system, then pointed out that change in circumstances requires patient consent to be revisited; would the proposed system offer backward compatibility? Carey said it would have to. Another audience member asked how to do key management in the DRM context. Carey agreed that that was a problem and cited the additional problem of making a uniform, trustworthy DRM enforcement environment. He suggested that patients should hold their own DRM keys somehow. Carey concluded by summarizing some computations that would produce different views of healthcare data for different observers.

Who Does the Autopsy? Criminal Implications of Implantable Medical Devices

Marc Goodman, Future Crimes Institute

Marc Goodman, who has worked with the LAPD, Interpol, and FBI, offered a law-enforcement view of medical-device-related threats on the horizon. He gave examples of technology being integrated in human bodies and suggested that people might soon receive elective implantable medical devices (IMDs). This integration raises questions for forensics, such as: can medical examiners conduct forensic

analysis of an IMD? The answer, given the current state of medical exams, is no, meaning that a forensic analysis might fail to discover an IMD's role in an event. He further noted the increasing use of consumer-grade computing devices in health care. He invited the audience to consider what kinds of recording and recovery mechanisms IMDs should use to alleviate the problems he mentioned.

Raj Rajagopalan noted the dearth of standards for forensic analysis of mainstream computers and asked what hope there was that the niche of IMDs would be standardized. Goodman pointed out that computer-forensic standards were beginning to appear in Europe and suggested that now was a good time to innovate. Another audience member asked whether there were standards related to default passwords on medical equipment. Goodman agreed that there ought to be standards now in order to set a precedent, since "past is prologue."

Long Papers

Summarized by Shane S. Clark (ssclark@cs.umass.edu)

A Research Roadmap for Healthcare IT Security Inspired by the PCAST Health Information Technology Report

Matthew D. Green and Aviel D. Rubin, Johns Hopkins University

Matt Green presented this work on the recent report by the President's Council of Advisors on Science and Technology (PCAST) titled "Realizing the Full Potential of Health IT," which outlines a vision for the future of electronic medical records (EMRs). Green noted that deployment of EMRs in the US is increasing but that the systems are generally not interoperable and that both sharing and security are at an early stage. He also noted that there is significant existing legislation such as the HIPAA, CCR, and HITECH acts, but that much of the legislation suffers from excessive complexity or underspecification.

Green next listed several open research areas that he feels must be addressed before a vision like that articulated in the PCAST report can be realized. The list included managing user identity, audits and logging, patient interaction with EMRs, cryptographic access controls, de-identification, and security metrics. Green contended that all of these areas require significant new work and that researchers should seek new results in each area before those outside the academic community implement poor technical solutions.

During the Q&A, an audience member asked about the problem of legacy data. Green answered that he does not believe the PCAST report addresses legacy data at all, but that it is definitely an important problem. Another audience member asked about defining security metrics and how one can claim success in solving the problem. Green said that he does not have the answer, but that an important first step is separating apparent security from actual security.

Take Two Software Updates and See Me in the Morning: The Case for Software Security Evaluations of Medical Devices

Steve Hanna, University of California Berkeley; Rolf Rolles, Unaffiliated; Andrés Molina-Markham, University of Massachusetts Amherst; Pongsin Poosankam, University of California Berkeley and Carnegie Mellon University; Kevin Fu, University of Massachusetts Amherst; Dawn Song, University of California Berkeley

Steve Hanna presented this work on software security for software-based medical devices. The researchers chose to examine the security of automated external defibrillators (AEDs) because they are widely deployed (an estimated 1.9 million worldwide in 2009) and make heavy use of software. The researchers reverse-engineered an AED's firmware, as well as the associated update and reporting programs, uncovering a variety of vulnerabilities and successfully deploying a benign worm capable of infecting the tested AED.

The first vulnerability that the researchers identified is a weak firmware verification system that allowed them to create malicious firmware for the device. The second vulnerability is a buffer overflow in the update program that leads to arbitrary code execution on the PC running the software. They also found that the AED's PC software used hard-coded plaintext passwords for data upload and stored other user credentials unprotected on the Windows host. Hanna outlined a scenario for a malicious worm using these vulnerabilities. If an attacker is able to compromise a single AED, he could use the buffer overflow in the update program to gain arbitrary code execution on the host during the next update. The compromised host could then infect other AEDs during the update process.

Finally, Hanna outlined a series of recommendations to improve the state of medical device software security. He suggested that machines used for updates be physically isolated from the network or that updates be run only within fresh virtual machines. He also suggested that device owners carefully monitor physical access to the devices. Hanna closed by saying that the researchers had notified both the FDA and the OEM of the vulnerabilities and advocated continued use of AEDs based on their life-saving potential and the low risk of compromise.

Panel

Do Medical Devices Have Significant Forensic Value?

Moderator: Ben Adida

Panelists: Kevin Fu, University of Massachusetts Amherst; Marc Goodman, FutureCrimes Institute; Nate Paul, University of Tennessee/Oak Ridge National Laboratory; Mark Day, iRhythm Technologies, Inc.

Summarized by Shrirang Mare (shrirang@cs.dartmouth.edu)

Ben Adida started by asking the panelists about their position on the topic. Kevin Fu said that software-controlled medical devices ought to be trustworthy, more particularly for forensics; otherwise, how can one tell whether a failure is accidental or malicious? Mark Day (the industry representative in the panel) made two points: first, that the industry is overwhelmed with regulations, budgets, and various other issues, and among all these issues, it is hard to have state-of-the-art security in medical devices; second, that the raw data from medical devices is very sensitive, and people don't realize that. From raw data from medical devices one can infer many things about the user. Marc Goodman said that today's medical devices are used for health alignments, but increasingly they will be used for enhancements and conveniences. As the number of medical devices increases, he thinks it will be even more important that these devices should have secure logs that will help forensics identify the root cause of a failure. Nate Paul shared his concerns and experiences with medical devices that control therapies (e.g., insulin pumps). He also thinks that it is important to add security to these devices, and ways to do forensics analysis later on, if required.

Ben Adida asked Mark Day to elaborate on what kinds of inferences one can draw from raw data. Mark Day said that from 14 days of heart rhythm data (gathered using a single channel ECG at 200 Hz sampling rate), they could identify different user activities (e.g., brushing teeth, sleeping), respiratory rate, quality of sleep, whether the user was right-handed or left-handed, and much more. An audience member asked what security measures manufacturers have in their devices. Mark Day said that people in industry try to implement what security they can (e.g., encryption, checksums), but they do not have good imaginations for future attack surfaces. The development cycle for medical devices is four to five years, and so by the time their devices are out, things have changed in the real world (i.e., new attacks emerge). He said that remotely programming a device is possible and would help a lot, but it has its own risks. He stressed the point that the people in industry are under enormous pressure—from regulatory bodies, budgets, market—and they cannot change things easily in their devices. Marc Goodman commented that he understands the pressure in the industry and thinks having dialogues between different sectors (e.g., between manufacturers and the FBI) will help manufacturers think about future attack surfaces.

David Kotz (Dartmouth College) asked how difficult it will be to implement forensic techniques in low-power sensors. Kevin Fu said the sensors have tight-resource constraints and that they don't have enough memory for additional code, but he is hopeful that in the future sensors will have more resources to work with. Mark Day pointed out that to bring ideas into reality we must bring economics into the picture. Device manufacturers already have many issues to deal with, but if we can put the security risks and benefits in terms of economics, then manufacturers will start taking security issues seriously. To a follow-up question about whether it might be too much to ask of tiny sensors, Marc Goodman said that everything does not have to be done on the sensor; a few things can be offloaded. But he thinks that because of Moore's law, sensors will have enough resources for security/forensics requirements in the near future.

An audience member raised a concern about the four- to five-year development cycle, and asked if there was any way to add security easily and quickly. Mark Day said there are many reasons why it takes so long: proprietary platforms, need to support legacy systems, long approval process, to name a few. The same person commented that we have done it for automobiles—we have added on-board diagnostic systems and, going forward, we support newer auto models. Mark Day thinks that it is not an option for medical devices. Marc Goodman said that the idea of an industry alliance coming together and forming something like an on-board diagnostic tool is great, but such a tool will also be quickly available to an attacker, increasing the attack surface.

An audience member commented that adding IT to hospitals is hard. Nurses need 50% more time to add data to devices, and it takes away from their time doing actual health care. Kevin Fu said that auditing or data logging can be automated to some extent, and he thinks that an effective and safe system does not mean that the system is going to be unusable. Ben Adida asked, if less usable might be better for patient care, does it mean that less secure might be better for patient care as well? Nate Paul pointed out that for any solution, you have to balance different factors—security, privacy, usability, and cost. Finding the right balance among these factors is the key. Mark Day thinks that people are trained for patient care, not for security, and so they do not take security seriously and they try to circumvent it whenever a system is secure but is a little hard to use.

Carl Gunter (UIUC) asked the panel for their take on regulations. For example, flights are required to have a black box

as a recording device. In medical space we have the FDA deciding what the scope of regulations should be, but the regulation spectrum makes it really unclear where things stand. Marc Goodman pointed out the trend in Europe, where authorities are looking into black-box technologies in automobiles, and he thinks there is no reason not to have them in sensors in future.

Concluding the discussion, all the panelists agreed that people from different sectors need to talk to each other and get a better understanding of perspectives and problems of other groups. Nate Paul mentioned that they had some success in their talks with the FDA. He thinks physicians share the security concerns of medical devices and are interested in helping security researchers. Mark Day emphasized the need to understand real-life problems and the importance (and difficulties) of regulations.

Short Papers

Summarized by Aarathi Prasad (aarathi.prasad@dartmouth.edu)

Providing an Additional Factor for Patient Identification Based on Digital Fingerprint

Guy C. Hembroff and Xinli Wang, Michigan Technological University; Sead Muftic, KTH—Royal Institute of Technology

Guy Hembroff conducted a study which involved 13 hospitals, including critical care households at a rural setting and trauma care facilities associated with a federation. All these hospitals follow the HL7 versions for Health Information Exchange (HIE). They have seen some success with PKI. The hospital security architecture involves patients, medical staff, physicians, roaming physicians, databases, and ID management servers and certificate authorities. Sometimes test results end up with incorrect patient information. Medical staff get additional rights such as search capabilities, which they should not get. Patient-matching algorithms occasionally return duplicate results.

Given the existence of sophisticated fingerprint identification algorithms and improved biometric recognition technology, Hembroff suggests that patient identification can be based on their fingerprints, which can be indexed as a master patient identifier. This identifier becomes the biometric part of the HL7 stream, along with the patient's photo ID. A record locator service can then identify the patient based on their fingerprint and retrieve their health information, based on the security policy associated with the information. If more than one record is retrieved, the photo ID will be used to identify the correct record. Hembroff is concerned about cultural issues regarding the acceptance of fingerprints as a source of identification, and how to convince people that their digital fingerprint is secure and won't be used for other purposes.

An audience member asked about issues regarding legacy systems—what happens when biometric readers change. Hembroff answered that he knows of seven such fingerprint readers; some of them have changed since their origin, but not all of them. Another audience member asked whether it was necessary for the patient to be there every time, to which Hembroff answered that the patient needs to be there the first time her fingerprint is collected. Another question was how to deal with a patient who lost his finger. Hembroff answered that the patient would have to re-enroll in the system, and hence it is better to use multifactor biometrics.

Context-Aware Anomaly Detection for Electronic Medical Record Systems

Xiaowei Li, Yuan Xue, You Chen, and Bradley Malin, Vanderbilt University

Xiaowei Li presented an intrusion detection system for electronic medical records (EMR) using existing knowledge and traces from the clinical environment. Context information—organizational information, user role, etc.—is extracted from traces and applied to the model at runtime. In one clinical workflow, for example, you have a physician who needs to check a patient's lab test results before prescribing medications.

The workflow model he presented works in three tiers. In the first tier, a profile of the user behavior is constructed for each user or role. Next, the session is decomposed into a set of record-oriented clinical workflows. The third tier indicates the treatment guideline for the patient, which involves multiple users and roles.

An audience member asked how an anomaly is usually detected and what features are used for this detection. Li replied that normally action sets, action sequences, or other modeling techniques are used. Another member asked what would happen if the decisions in the workflow do not happen in the correct order, as in the example Li presented. Li replied that such challenges will be handled in the future with some preprocessing of the data.

Role Prediction Using Electronic Medical Record System Audits

Wen Zhang, Vanderbilt University; Carl A. Gunter, University of Illinois at Urbana-Champaign; David Liebovitz, Northwestern University; Jian Tian, Vanderbilt University; Bradley Malin, Vanderbilt University

Wen Zhang talked about role prediction, which uses audit logs to predict automatically whether a user is associated with a role. The group's work attempts to find a synergy

between the two dominant strategies: role-based access control (RBAC), and experience-based access management (EBAM). They used role prediction on the EMR system deployed at Northwestern Memorial Hospital and found 8095 users, 140 roles, 143 reasons to access records, and 43 services provided at 58 locations. The role predictor accurately predicted the job titles of 51.3% (4152) of the users in the system.

For better role prediction, Zhang introduced the concept of role hierarchy. It was observed that prediction accuracy increases as you go higher up in the hierarchy. But at higher levels, the number of roles is small and thus the "separation of duty" will be violated. He also talks about the "role-up" algorithm which tries to find a balance between prediction accuracy and role number. It was found that when the algorithm was biased to accuracy and there were a small number of resulting roles, the accuracy of role prediction was 63%; when it was biased towards specificity and number of roles was high, accuracy was 52%.

One audience member asked how many beds were in the hospital. Zhang said that the study involved 8000 users, though there were not necessarily that many beds. Was "physician" considered a role? The system deployed at Northwestern is Cerner, where physician is not a role. It was also pointed out that roles and privileges are mapped from a physician's nature; when a new physician comes in, it is unclear whether a new role should be assigned.

Audit Mechanisms for Privacy Protection in Healthcare Environments

Jeremiah Blocki, Nicolas Christin, Anupam Datta, and Arunesh Sinha, Carnegie Mellon University

Anupam Datta talked about how audit mechanisms are essential for privacy protection in healthcare environments. However, the cost of inspections by a human auditor would be very high if the auditor were to inspect each access to a patient record to determine whether it was appropriate or not. Their approach, "regret minimizing audits," learns from experience to recommend budget allocations for audits in every cycle to different types of accesses. For example, if in a given audit cycle celebrity record violations caused greater loss to the organization, then the algorithm ensures that there is a higher probability that the next time the auditor performs an audit, accesses to celebrity records will be checked more. The algorithm provably converges to the best fixed strategy (e.g., budget allocation) in hindsight.

He explained that the algorithm doesn't make any assumptions about the adversary's incentives; the learning is based on the loss that is incurred during each cycle. As future work, he wants to consider alternative adversary models and audit mechanisms (which incorporate incentives), test whether experts can be identified from the logs using machine-learning techniques, and conduct experimental evaluation of the approach.

An audience member asked how it is possible to figure out who the celebrities are, to which Datta answered that their records are typically marked as celebrity records and audited separately. Another audience member asked whether logs are perfect and what would happen if all actions are not captured in the logs. Datta replied that the auditing is only as good as the information recorded in the log; he gave an example of how someone might look up information on a record and make a phone call, and not alter the data; this action would not be captured in the logs. Cory Cornelius asked whether attackers would be able to run this algorithm. Datta answered that the guarantees of the algorithm hold even when the attacker runs the algorithm.

Panel

Can We Do Meaningful De-identification of Medical Data?

Panelists: Arvind Narayanan, Stanford University; Lee Tien, Electronic Frontier Foundation; Kelly Edwards, University of Washington; Sean Nolan, Microsoft

Summarized by Aarathi Prasad (aarathi.prasad@dartmouth.edu)

Sean Nolan presented an organizational perspective on the topic. He said that it is fiction that data is anonymized and cannot be re-identified. He stated, however, that there is an increased willingness to disclose identified information to allow research to happen. The question, he pointed out, is how we can maximize people's understanding of doing it and how to maximize the value of doing it. Kelly Edwards presented an ethics perspective. She agreed that her goal was also to protect people while promoting clinical care. She said that we are caught up in the negative sense of privacy. The positive sense is that people have the right to choose and can opt in. She said that people are willing to participate at a high level, if they perceive benefits in doing so. A trustworthy system, in an ethical sense, is based on relationships and accountability. The question, she pointed out, is how to launch a public campaign to educate people about what is happening to their data.

Arvind Narayanan agreed that anonymization is pure fiction. He pointed out that understanding the data flows and who gets access to the data is very complex, so narrowing the process to a set of identifiers is not the right approach. Lee Tien's focus is on privacy, with an emphasis on health

privacy. The big takeaway, according to him, was that no one knows anything about laws in health privacy, health information exchange architectures, etc. So he said it is not right to put the burden on the doctor to inform the patients where their data goes.

An audience member asked whether de-identification is the right way to go. Nolan said the question is what you are doing the research for—treatment for one person vs. analysis of 10,000. If the data is identified, you can reach back to the participants and learn more about them. Edwards said that providers are more nervous than patients, and no regulations say that identification has to be stripped from clinical studies. Narayanan replied that there are a variety of context and threat models. De-identification is useful in case of celebrity records and with an adversary who does not have technical expertise. He suggested differential privacy as a possible option instead of having fully identifiable data and de-identified data.

Another audience member pointed out that de-identification doesn't work as well as people think, especially if there is a threat from an adversary. He asked what is more important—privacy protection with de-identification or having the ability to cure AIDS if we have identifiable data? Nolan said that in the future we might have sufficient opt-in raw material to make public health claims. Edwards replied that in the US people want individual benefits and are willing to be part of something that might benefit them in the future. They are willing to contribute if we ask them. Nolan gave the example of how people donate blood because they know it is safe to do so. Narayanan argued that it was not clear to him if this could be scaled to a large population. He wondered if it was possible to provide incentives, using game theory, so that individuals could see some benefits of providing their data for research. Tien said that it is important for participants to know who is conducting the research. Sean said that a patient might not want his data to be used for research, when he is going to the doctor for treatment; the patient has to trust the system before contributing her data.

Ben Adida pointed out that hospitals were able to find correlations between patients with negative heart rates and a drug. In such cases it might be good to have identifiable data, but where do we draw the line? Tien said that when providing data for research, the patient might not know what utility there is in his data. According to Edwards, no one can decide what counts as a benefit for a diverse population; maybe a educational campaign is the solution.

Another audience member asked whether researchers could write programs, able to be run by the data holders, in such a way that the data collected could not be identified. Nolan replied that this solution, though exciting, may not work, since it is not possible to get aggregated data in all cases. In order to build this program, synthetic data is required, which is difficult to generate. There is not enough incentive for companies to adopt this solution—you will have to charge the patients to run this program—so this solution will need fundamental infrastructural changes.

Carl Gunter asked whether the panel could comment about consent bias—how to measure who opted out or opted in. Nolan said that we are still trying to comprehend consent. Edwards talked about exempt research, where it is possible to do research without requiring the participant's consent. Narayanan wondered whether we could work with self-reported data, but this data might not be useful in all cases, since the fidelity is questionable.

An audience member pointed out that biologists are required to publish their data, so that their results can be verified. Can re-identified data be used for other purposes? Edwards pointed out that biological data is usually de-identified and comes with lots of restrictions. Tien was concerned about how if some (remotely identified) data is released, people might want access to it, and access cannot be denied. Narayanan said that companies also should have a system, similar to IRBs, when conducting studies to collect data, that could audit the research.

The final question was about why data gets disclosed and about differential privacy, which, according to the audience member, has not been verified with studies other than those by the authors. Nolan pointed out that data is usually disclosed so that it can be verified. Maybe there are other ways to verify data. Narayanan said that in cases of differential privacy, anonymization comes, not from privacy protection, but from the noise that is included in the data. This has been verified in academic settings but has not been adopted anywhere.

Long Papers

Summarized by Shrirang Mare (shrirang@cs.dartmouth.edu)

Quickshear Defacing for Neuroimages

Nakeisha Schimke and John Hale, University of Tulsa

Nakeisha Schimke presented her work on de-identification method for neuroimages. The goal of this work is to sufficiently de-identify neuroimages that they cannot be linked back to an individual, and to do this task efficiently compared to existing techniques.

In neuroimages, facial features can be used to identify an individual. There are two existing de-identification meth-

ods used to remove these facial features: skull stripping, a process of segmenting brain and non-brain elements (which include facial features), and MRI defacing, a process of removing only the identifying facial features leaving the brain and surrounding tissues intact. The MRI Defacer algorithm relies on a manually labeled atlas to identify facial features. The skull stripping process is not always accurate, and it is hard to automate. The MRI Defacer process is accurate, but it requires a manually constructed atlas and is also computationally expensive. Quickshear is an effort to make the de-identification process better by making it automatic and computationally inexpensive.

The Quickshear algorithm finds a plane that divides the volume (i.e., the head in the image) into two parts: one containing the facial features and the other containing the entire brain volume. All the voxels of the "face" side are set to zero, which (apparently) is effective to de-identify the face. The key is to find the right plane such that the brain volume is kept intact. The researchers use convex hull algorithms (Andrew's monotone chain) to identify the brain mask, and once the points on the convex hull are identified, the dividing plane is drawn using the points closest to the forehead. The researchers compared their method against MRI Defacer, using 42 images from 21 subjects. They used OpenCV Face Detector to evaluate how well a method has de-identified an image. Out of the 42 de-identified images, OpenCV identified nine MRI Defacer images as faces and about 10 Quickshear images as faces. However, according to the researcher, Quickshear removes fewer brain voxels, and its running time was less than MRI Defacer; thus, Quickshear achieves nearly the same output in terms of preserving the user's privacy but is more efficient.

Arvind Narayanan (Stanford) wondered about the possibility of identifying an image based on geometry of the face; for example, distance between eyes (eye sockets are present in Quickshear images). Schimke agreed that it's a possibility but pointed out that it is hard to measure the precise distance between eyes using the eye sockets in the Quickshear images.

Adaptive Security and Privacy for mHealth Sensing

Shrirang Mare and Jacob Sorber, Institute for Security, Technology, and Society, Dartmouth College; Minho Shin, Myongji University, South Korea; Cory Cornelius and David Kotz, Institute for Security, Technology, and Society, Dartmouth College

Shrirang Mare presented his work on an adaptive protocol for mobile health sensing. People are increasingly using mobile medical sensors to measure their activity and health information, and these sensors transmit data to an aggregator device like a smartphone. Together, the sensors and the aggregator device form a body area network (BAN). BAN has similar security and privacy problems as a WiFi network. The privacy-preserving wireless protocols (proposed for WiFi networks) cannot be used for BAN because of their large overhead. The proposed adaptive method is designed to make these privacy-preserving wireless protocols efficient so that they can be used in the low-power BAN, while preserving the security and privacy properties of those protocols.

The protocol overhead is typically the header and message authentication code (MAC). The larger the overhead, the stronger the security, increasing, for example, the resistance to forgery attacks. Non-adaptive protocols use a fixed long MAC for strong security. Mare argued that a user (a user's BAN, really) is not always in a hostile environment, so always using strong security is inefficient. Instead, he suggests using a small overhead, but increasing the overhead when an adversary attacks, when the adversary is trying to forge a message. The condition on "when" to increase the overhead is critical, and he presented a probabilistic condition to identify an ongoing attack based on the number of corrupted packets (i.e., packets that fail MAC verification).

An audience member asked what happens in the case of a passive attack. Mare said the adaptive protocol does not change any parameter that would make it easier for a passive adversary to learn anything about the payload. For example, changing the MAC size does not help the adversary learn the contents of an encrypted payload. That is, the proposed method does not make the adaptive protocol any more vulnerable to passive attack than the original non-adaptive protocol.

Controlled Dissemination of Electronic Medical Records

Guido van 't Noordende, University of Amsterdam, The Netherlands

Building upon a security analysis of the Dutch electronic patient record system, Guido van 't Noordende presented his ideas on how to share electronic medical records. His approach is decentralized and keeps access to patients' information to a minimum. In this talk, he identified several paths that can be used to share information between different parties, such as patient, physician, family.

Noordende first described the traditional healthcare model: a patient visits a doctor (Doc #1). The doctor keeps all the records of the patient. When the patient visits another doctor (Doc #2), Doc #2 asks for the patient's records from Doc #1 (a pull-based model). Alternatively, Doc #1 can also send the records to Doc #2, if the patient's visit to Doc #2 is planned (a push-based model). Noordende thinks that using a controlled push-based approach with the convenience of a pull-based model is the right approach to sharing patients' records. He

then presented his architecture model, outlining different paths to disclosure.

He described five different paths to disclosure for the patient's information. The idea is that the data stays in one place, but the pointer to the data is moved across different parties in a controlled fashion (i.e., controlled dissemination). The five different paths to disclosure basically describe the medium through which the pointer is shared with the doctors. The five different paths are: professional (secure) push model (e.g., emails), patient's mailbox, USB drive, smartcards, and paper (pointer writing on paper). The idea is that the patient carries the pointer to the data with him, and whoever gets the pointer from the patient gets access to the data. Thus, the patient controls the dissemination of the information.

An audience member wondered if this model can be extended to include insurance companies. Noordende said, for simplicity, he did not include insurance companies in his slides, but it is certainly possible to include them in his model. Another audience member asked how the patient can get all the active references (i.e., pointers) that are floating around. Noordende said the pointers are floating but the data is in one place. One can have access logs at that place,and can tell who has accessed your data, and also keep a log of all the pointers.

Rump Session

Summarized by Aarathi Prasad (aarathi.prasad@dartmouth.edu)

Atif Khan, University of Waterloo

Atif Khan's interests lie in patient consent and consent management. His goal is to understand what a patient wants out of the system. Can the patient choose what her data is used for, whether it is shared with her family physician or with hospitals in another state? Khan uses semantic Web technologies to define information using ontologies. The patient consent rules will be based on these ontologies.

Ben Ransford, University of Massachusetts Amherst

Ben Ransford previewed a SIGCOMM paper he coauthored. It is well known that certain medical devices are vulnerable to passive eavesdropping or the issuance of unauthorized commands. The authors' methods can protect devices that are already implanted and cannot easily be replaced. They developed an auxiliary, wearable device, called an IMD Shield, that acts as a proxy. The Shield has two antennas: one that sends a random jamming signal and another that transmits and receives data. The IMD Shield's jamming reduces the risk of private data loss and active commands by jamming transmissions to and from the IMD. The IMD Shield cancels its own jamming signal so that it is the only device that receives the bidirectional communications in the clear.

Andrés Molina-Markham, University of Massachusetts Amherst

Andrés developed a platform for medical applications, called Moo. It includes an RFID reader that provides energy to power this device, which has no battery. The microcontroller can be programmed in C. Moo has an accelerometer and temperature sensor; external sensors, storage, and harvesters can be added to the device as well.

Kevin Fu, University of Massachusetts Amherst

Kevin Fu talked about the Open Medical Device Research Laboratory, which helps researchers conduct trustworthy computing research on IMDs. MIT and Berkeley have already used IMDs from this library. A student at the University of Pennsylvania opened up the devices to understand the digital logic that goes on inside them. The devices are sterilized so that they are safe for research.

Joseph Ayo Akinyele, Johns Hopkins University

Joseph Ayo Akinyele developed a framework, called Charm, to help cryptographers who want to apply ideas to medical applications and to secure health data in the cloud, in mobile devices, etc. Implementing, measuring, and comparing crypto methods is difficult, especially since it takes a long time to write crypto code. The functional library has math libs at the lowest level and crypto schemes that focus on the algorithms at the higher level. Charm is implemented in Python. The alpha version has already been released and has been used. This version has implementations of attribute-based encryption, key policies, and ID-based encryption.

Matthew Pagano, Johns Hopkins University

Matthew Pagano's work is focused on using attribute-based encryption (ABE) to secure electronic medical records (EMRs) on mobile devices. It is difficult to get access to EMRs and other medical data during catastrophes or network outages. Access policies in healthcare can be complex, and medical systems might not have adequate security policies. With ABE, EMRs can be encrypted with expressive policies that allow the records to be exported outside the trust boundary of a medical institute. This provides self-protecting, offline access control, which is especially vital when network access is unavailable.

This solution allows patients to access their medical records and potentially store them on untrusted storage servers. In this system, the medical institute encrypts a patient's records using ABE with a suitable access-control policy.

The encrypted records are then stored on a Web server, from which patients can download their records onto their mobile devices. After receiving an ABE private key from the medical institute in an out-of-band channel, patients will be able to access their records at any time. Patients can also store their medical data with their PHR providers, either unencrypted, partially encrypted, or fully encrypted.

Mike Rushanan, Johns Hopkins University

Mike Rushanan is working on creating a trusted computing base (TCB) for mobile electronic health records (EHR). Mobile devices could have malware, and it might not be safe to build mobile health applications that can store EHR. His approach involves a Java card with attribute-based encryption (ABE), so that this card will become a trusted ABE service on the phone. The card can be installed in the phone, and it can store the patient's health data on it. They will also develop a communication protocol for the phone to interact with the card. Some processing will have to be done in the cloud, due to the resource limitations on the mobile phone. ABE can be broken up so that processing can be done away from the trusted base.

Michael LeMay, University of Illinois

Michael LeMay's research focuses on providing strong isolation for medical applications on a mobile platform. He presented the idea of a dual persona smartphone, which could be used either by the patient or the physician. However, this phone could have enterprise data or the user's personal data. It is necessary to provide clear isolation of the user's medical information on the phone. Existing software solutions have drawbacks. Protection policies are distributed and access controls are discretionary. He pointed out that errors can compromise protection if they are related to memory management and that VMMs are not enough for isolation. He also said that resource sharing could lead to vulnerabilities such as covert channels.

Raj Rajagopalan, HP

Raj Rajagopalan presented a new general notion of privacy. If you release information, you leak more information than you want. He said it is better to measure the relative release of information. He pointed out that a tradeoff should be drawn between utility (explicit disclosure) and privacy (implicit disclosure);that way you can reveal data with different levels of precision. Data exchanges involve a lot of people and sometimes time is important, so it is better if the data is not deleted. Rajagopalan wants to know whether it is possible to provide positive incentives for data holders to obey the privacy needs of individuals and whether it is possible to establish joint ownership of medical data. He also wants to know if it is possible to have a common interface between security and privacy. An audience member asked whether it is legal to sell data, to which Rajagopalan replied that there is a 4-billion-dollar industry based on selling medical data.