Persistent Security, Privacy, and Governance for Healthcare Information

W. Knox Carey, Jarl Nilsson, and Steve Mitchell
Intertrust Technologies, 955 Stewart Drive, Sunnyvale, CA 94085

***Abstract*—A fundamental tension between accessibility and governance exists in the design of healthcare information systems. In order to be useful in practice health information must be distributed, but as the information moves between systems — and different information governance policies — the risk of privacy and security violations increases. The lack of a persistent policy enforcement mechanism thus inhibits the dissemination of health information, making it less useful for research and treatment. In this paper, we argue that trusted computing and policy management technologies are required to allow for broad distribution of health information while preserving security and privacy. We also introduce the concept of *derived resources*, which helps to address many of the unique challenges in the governance of health information.**

I. Introduction

Medical information is everywhere, but only rarely where it needs to be to make a difference. A public health researcher studying the propagation of a new disease needs timely, comprehensive information drawn from front-line healthcare institutions across the country. A family doctor seeing a new patient might avoid repeating expensive tests if he could only access the patient's lab results from last year. Health data recorded by home monitoring systems sit locked in a personal computer, when they should be informing diagnosis and treatment in the hands of a clinical practitioner. All of these examples depend on the frictionless flow of medical information through a heterogeneous network of devices and systems. Unfortunately, this information flow is not happening.

In practice, sharing data across systems involves significant risks to the security and privacy of patient data. When patients and institutions *do* release information, they have little assurance that it will be governed in a manner that is consistent with their policies.

In this position paper, we describe some of the elements necessary for a solution to the problems of governed information sharing, and discuss how new developments in trusted computing enable new applications for patient privacy, data security, and medical research.

II. Elements of the Solution

Many of the technologies required to provide persistent governance of medical information have been in use for many years in other fields. Applying these technologies to healthcare will require a rethinking of basic assumptions about security and governance for healthcare information systems.

A. Persistent Governance of Medical Information

Providing for governance of medical information across heterogeneous systems requires an expanded perspective on the nature of data security, one that takes into account not only *access* to governed information, but also the *managed use* of that information.

Since the earliest days of computing, sensitive information has been secured primarily by keeping it isolated within a carefully guarded perimeter that admits only authorized individuals [CSTPS]. Unfortunately, most access control models share the property that the use of sensitive information — once access is granted — is relatively unrestricted.

In the 1990s, the first digital rights management (DRM) systems [DBOX] introduced the notion of *persistent governance*. Not only were data protected cryptographically, but use of those data were subject to certain rules that were securely associated with the data and enforced consistently wherever the data traveled. By contrast, the policies governing information in older systems depended as much on the location of the data as they did on the intention of the data originator.

The ability to persistently govern information across systems enables new possibilities for the dissemination of sensitive information, possibilities that are not realizable with more traditional forms of access control:

- Data can transmitted through a heterogenous network with no degradation in security.
- Data can be consumed offline — rules are evaluated locally at the point of access.
- Data and the rules governing it can be distributed separately. Data may be distributed in advance of rules, and new rules may be associated with the data at any time.
- Data can be packaged with rules that enforce very fine-grained usage policies. For example, access to data can depend upon time, the accessing principal, the membership of the accessing device in a group, and so forth.

B. Consistent Trust Management

Advisory bodies to the US federal government have begun to address requirements in trust management [FHITS]. The context for these recommendations however, has tended to focus on securing communications between endpoints rather than providing assertions of compliance to certain policies. Trust management systems for governed medical information will need to ensure that originators of medical data — the entities that associate policy with the data — can rely upon the policy enforcement performed by any credentialed recipient without the necessity of establishing a relationship in advance.

C. Derived Resources

Applying trusted computing technologies to the governance of medical information presents several unique requirements:

- Various stakeholders require different views of the same data. For example, a patient may be interested in every data point in a series of health metrics, whereas his physician is only interested in a summary of the trends. Epidemiology researchers might see the data, but they should see it at perhaps a lower resolution, and certainly in an anonymized form.
- Different aspects of the governed data may be important over time. Initially, for example, a doctor may be interested only in the overall trend for a particular health metric. When an anomalous circumstance is discovered, however, the full data series may become important. This phenomenon is also important in research, where revisiting data years later can shed new light on older studies [CURRY]. It should not be necessary to repackage older information to enable this feature.
- To maximize utility, the data need to be distributed as broadly as possible, but the data should not be distributed in an ungoverned manner — it should always be possible for the owner or originator of the information to control data access policies and to audit actual usage.

To meet these requirements, we introduce the concept of a *derived resource*. In most governance systems, granting access to a resource consists of applying a set of conditions to determine if access is possible and then producing the key that allows the consuming system to decrypt the static resource, which is uniform for all users. As the requirements above indicate, this uniformity is not sufficient in healthcare applications — different stakeholders have different interests in the data that may change over time.

Derived resources address this problem by securely associating a set of specified computations — performed on the resource itself — with the packaged resource. In this scheme, a resource to be protected is associated with a set of rules governing access, keys to allow decryption, and computations to be applied to the original resource before returning it to the requesting system. The computations may depend on several factors, including the identity of the principal that will access the data, environmental considerations at the point of evaluation of the computation, or state information maintained by the system at the client accessing the derived resource.

- The precise view required of a set of raw information need not be computed in advance; the packager simply associates a computation that produces the derived resource. These computations can be reusable for different data sets, e.g. produce a five-day trailing average over the enclosed data series.
- Since computations may depend on conditions such as the accessing principal, different stakeholders in the resource may see different views; different computations are associated with each principal.
- New views of the resource can be provided after the fact. If a new type of derived resource becomes important later, the packager can simply provide a new set of associated computations rather than recomputing and repackaging the entire data set. New computations can be generated by the original packager or proposed by the users of the data and selectively authorized by the owner of the resource.
- Creating a derived resource can be lossless. Using derived resources, the original data need not be repeatedly filtered and repackaged, so no information is lost that may be of use in the future.

Adding derived resources to existing trust computing models enables new uses that cannot be realized with older technologies. For example, consider a diabetic patient who is recording blood glucose levels at home using an non-invasive glucose monitor. The data are synchronized with an online service that (a) allows the patient to track his blood glucose over time and (b) actively packages and forwards the information to the patient's physician with rules that grant access only to a set of authorized principals. The computations can be specified such that the physician herself has access to the full set of data, whereas her colleagues may see the data only in a partially anonymized form.

A physician may be interested only in the peak blood glucose over a week rather than the hour-by-hour data points. Default computations associated with the resource can therefore produce just the desired indicators. On the other hand, when the situation merits a more detailed investigation — e.g. the blood glucose peaks predictably at certain times — the physician can propose and apply alternative computations that produce higher-resolution data.

The same technologies may be used to provide selectively anonymized data for epidemiology and medical research purposes, as well as enabling researchers to publish data in a way that allows for third parties to validate the computations that were performed to draw a scientific conclusion.

References

[CSTPS] J. P. Anderson *Computer Security Technology Planning Study, ESD-TR-73-51, ESD/AFSC*, Hanscom AFB, Bedford, MA, Oct. 1972, Volume 1.

[MTMIS] Blaze, Feigenbaum, and Lacy. *Managing Trust in Medical Information Systems*. AT&T Internal Research Paper. http://www.eyetap.org/~maali/trust-papers/blaze96managing.pdf

[DBOX] Sibert, Bernstein, and Van Wie. *A Self-Protecting Container for Information Commerce*. Proceedings of the First USENIX Workshop on Electronic Commerce, New York, New York, July 1995. http://www.usenix.org/publications/library/proceedings/ec95/full_papers/sibert.txt

[FHITS] Office of the National Coordinator. *Federal Health Information Technology Strategic Plan, 2011-2015*. http://healthit.hhs.gov/portal/server.pt/community/federal_health_it_strategic_plan_-_overview/1211

[CURRY] Andrew Curry. *Rescue of Old Data Offers Lesson for Particle Physicists*. Science, vol. 331, pp. 694–695, 11 Feb. 2011.