Enabling Privacy Policies for mHealth Studies

Brian Wang
UCLA CS

Mani B. Srivastava
UCLA ECE

Motivation

- mHealth is an active area of research, where health studies are designed to capture sensor data from participants using a variety of mobile and wireless devices.

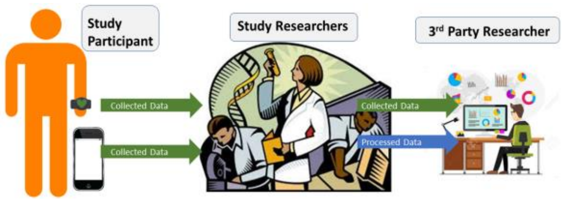


Figure 1: However, data may be processed and shared with many entities, which incurs a privacy risk to participants. They may want to limit access to data with respect to time, location, or other factors. We call this notion **context awareness**. In addition, participants may want to limit with whom and how their data is shared, a notion called **data handling**.

Problem

- Popular access control policy languages such as XACML are able to express data handling properties. On the other hand, system specific languages, such as IpShield for Android devices, enable control of data over contexts (i.e. location). However, neither group of languages are able to provide both data handling and context awareness.
- In addition to sharing collected data, there is sharing of processed data, resulting from data fusion and inference in mHealth studies. This processing creates new data byproducts which do not have privacy policies. mPolicy must also account for these cases.

mPolicy

To model context awareness and data handling, we introduce a policy language based on four constraints:

- **Data window**
  - Used to determine contextual conditions for a privacy policy
- **Operations**
  - Used to determine types of actions that can be performed on the participant data
- **Intents**
  - Purposes that the data may be used for
- **Data stream**
  - Modifications to apply to the data values themselves, such as reducing the precision of values, rounding, etc. An example is reducing the granularity of GPS values from kilometer level accuracy to meter level accuracy

```
ENTITY: MayoClinic
ENTITY-TYPE: Hospital
DATA-STREAM-TYPE: GPS
   DATA-WINDOW:
      ALL            # All of the conditions must be satisfied for this policy to be relevant to an access
           NOT TimeRange(“9pm”, “9am”),
           NOT LocationAt(“34.069455”, “-118.444515”),
           NOT TimeGreaterThan(“December 31st, 2019”),
   DATA-METHODS:
      ALL            # All methods must be applied before downstream dissemination
           CoarseGPSFilter()
   ALLOWED-OPERATIONS-INTENTS:
      ANY            # Any of the conditions must be satisfied for this policy to approve an access
           ( “Clustering Data”, “Recommending Points of Interest”),
           ( “Statistical Analysis”, “Finding Commonly Frequented Locations”)
```

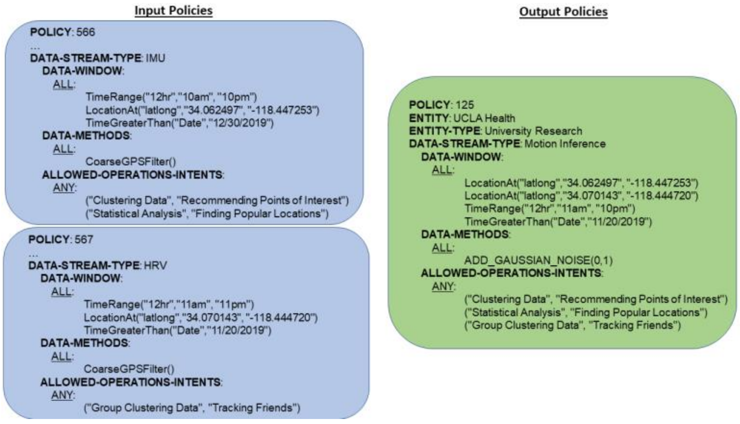
Figure 2: This is an mPolicy that is meant for any data received by MayoClinic, which is a hospital entity. For this participant’s GPS data, this policy applies constraints over the times of 9am to 9pm, with access ending on December 31st, 2019. This policy also states that all GPS values with the data-window constraints must have reduced granularity. Lastly, only certain operations and purposes are allowed on the data. In this case, data clustering for recommending POI is an allowed operation and intent.

Downstream Policies

Generating privacy policies for data byproducts requires the creation of a new byproduct policy, parameterized by the entities, operations, and data involved in future interactions.

$$ETM = w_{entity} * t_{entity} + w_{entity-class} * t_{entity-class}$$
$$+ w_{operation} * t_{operation} + w_{stream-type} * t_{stream-type}$$

Figure 3 (above): Using the entity trust measure (ETM), we alter the data methods for a policy. Each parameter of the equation is manually determined. If the ETM is above a threshold α, we add Gaussian noise. The resulting policy is shown in **Figure 4 (right)**.



Input Policies

```
POLICY: 566
...
DATA-STREAM-TYPE: IMU
   DATA-WINDOW:
      ALL:
           TimeRange("12hr","10am","10pm")
           LocationAt("latlong","34.062497","-118.447253")
           TimeGreaterThan("Date","12/30/2019")
   DATA-METHODS:
      ALL:
           CoarseGPSFilter()
   ALLOWED-OPERATIONS-INTENTS:
      ANY:
           ("Clustering Data", "Recommending Points of Interest")
           ("Statistical Analysis", "Finding Popular Locations")
```

```
POLICY: 567
...
DATA-STREAM-TYPE: HRV
   DATA-WINDOW:
      ALL:
           TimeRange("12hr","11am","11pm")
           LocationAt("latlong","34.070143","-118.444720")
           TimeGreaterThan("Date","11/20/2019")
   DATA-METHODS:
      ALL:
           CoarseGPSFilter()
   ALLOWED-OPERATIONS-INTENTS:
      ANY:
           ("Group Clustering Data", "Tracking Friends")
```

Output Policies

```
POLICY: 125
ENTITY: UCLA Health
ENTITY-TYPE: University Research
DATA-STREAM-TYPE: Motion Inference
   DATA-WINDOW:
      ALL:
           LocationAt("latlong","34.062497","-118.447253")
           LocationAt("latlong","34.070143","-118.444720")
           TimeRange("12hr","11am","10pm")
           TimeGreaterThan("Date","11/20/2019")
   DATA-METHODS:
      ALL:
           ADD_GAUSSIAN_NOISE(0,1)
   ALLOWED-OPERATIONS-INTENTS:
      ANY:
           ("Clustering Data", "Recommending Points of Interest")
           ("Statistical Analysis", "Finding Popular Locations")
           ("Group Clustering Data", "Tracking Friends")
```

Conclusion

- mPolicy is designed to express the privacy needs of mHealth study participants with context awareness and data handling, while also addressing the need to create new byproduct policies from data byproducts.

References

eXtensible Access Control Markup Language (XACML)Version 3.0.” [Online]. Available: http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.html

S. Chakraborty, C. Shen, K. R. Raghavan, Y. Shoukry,M. Millar, and M. Srivastava, “ipShield: A Framework ForEnforcing Context-Aware Privacy,” 2014, pp. 143–156. [Online].Available: https://www.usenix.org/node/179736