A Dynamic Distributed Architecture for Preserving Privacy of Medical IoT Monitoring Measurements

Salaheddin Darwish[1], Ilia Nouretdinov[1], and Stephen Wolthusen[1 2]

[1] School of Mathematics and Information Security, Royal Holloway, University of London, Egham, United Kingdom
[2] Department of Information Security and Communication Technology, Norwegian University of Science and Technology, Trondheim, Norway
{salaheddin.darwish,i.r.nouretdinov,stephen.wolthusen}@rhul.ac.uk

Abstract. Medical and general health-related measurements can increasingly be performed via IoT components and protocols, whilst inexpensive sensors allow the capturing of a wider range of parameters in clinical, care, and general health monitoring domains. Measurements must typically be combined to allow e.g. differential diagnosis, and in many cases it is highly desirable to track progression over time or to detect anomalies in care and general monitoring contexts. However, the sensitive nature of such data requires safeguarding, particularly where data is retained by different third parties such as medical device manufacturers for extended periods. This appears to be very challenging especially when standards-based interoperability (i.e using IoT standards like HyperCAT or Web of Things-WoT) is to be achieved. This is because open meta-data of those standards can facilitate inference and source linkage if compiled or analysed by adversaries. Therefore, we propose an architecture of pseudonimyised distributed storage including a dynamic query analyser to protect the privacy of information being released.

Keywords: medical IoT, differential privacy, pseudonymisation, meta-data, anonymisation

1 Introduction

Privacy has been identified as a major concern in the Internet of Things (IoT) [27], but earlier it was mostly concentrated on such aspects as identification of individuals and interactions, localisation and tracking, without paying much attention to the profiling of individuals and their behaviour based on data sources ranging from radio-frequency identification (RFID) tags via surveillance devices to wearable components. Nevertheless, there are few attempts to address this particular issue in IoT like in RFC-7744 [24]. However, despite the IoT potential for improved outcomes as well as cost savings identified in various domains including the health sector [7], individuals are subject to monitoring by diverse sensors over extendable time-periods, resulting in **linkage** of such different sources as a major risk for **re-identification** [10]. Measurements and observations not only limited to IoT environments may be linked together eventually as individual data sources become interchangeable and are no longer restricted to vertical application domains in which anonymisation can take place as required.

Current practice frequently relies on information being de-identified in a particular context, but without considering how such information may be linked with other sources or over longer time-periods as may become feasible for health monitoring and care where symptoms may be analysed algorithmically or are desirable for research purposes. It is, however, well known that merely anonymising a *pre-defined* subset of attributes will not prevent re-identification when combined with other attributes [17]. This, however, has severe implications for how such information arising in a medical context may be processed, stored, and presented. In the United Kingdom, common law and a number of laws including the UK Data Protection Act (1998), the NHS Act (2006), Social Care Act (2012), Human Rights Act (1998), and Data Protection (Processing of Sensitive Personal Data) Order (2000) impose bounds on handling of sensitive information, whilst EU General Data Protection Regulation (2016/679) coming into effect in 2018 imposes further constraints. Also, similar (less prescriptive) considerations apply in other jurisdictions where e.g. the U.S. Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule defines health information as individually identifiable if (1) it identifies the individual; or (2) there is a reasonable basis to believe the information can be used to identify the individual.

In order to allow the flexible effective aggregation of diverse IoT data sets and measurements in a privacy-preserving and lawful manner, it is insufficient just to aggregate even de-identified sources, but rather must provide further control over processing and queries, *particularly where data sources may be aggregated over time or new sources added.* In this work, we specifically address the impact of involving *external sources (prior knowledge)* on the control of privacy protection via using e.g. a **meta-data** element in medical IoT (referring to HyperCAT [4] and W3C Web of Things-WoT [2] standards) which is typically used for describing IoT asset and attributes semantics for interoperability and discovery purposes. For this reason, we propose an architecture of pseudonimyised distributed storage incorporating dynamic query-based privacy protection relying on differential privacy models to ensure that constraints are honoured.

Sec.2 reviews related work. Sec.3 describes meta-data potentials in privacy. Sec.4 addresses the proposed distributed storage model for protecting privacy. Sec.5 analyses the problem of aggregation and proposes a query-based approach for selective release. Conclusions and future work are in Sec.6.

2 Background

Emerging of various IoT-based and mobile health applications has offered an open and seamless way of tracking health of population such as HealthKit [1], Medical IoT monitoring [26,20,21,23], etc. However, using such complex systems poses more privacy concerns about how the sensitive information is being handled. Also, most IoT-related security works are focusing on securing communication rather than data at rest. Therefore, it is very demanding to expand the vision of how to effectively protect privacy of data in this system.

In order to protect the privacy of data records, some transformations need to be applied. They may involve reversible operations (such as encryption and pseudonymisation) and irreversible ones (anonymisation, deleting, obscuring the data [5]). According to [19,18], pseudonymisation is developed as an alternative for encryption as a privacy-enhancing technique in which identification data must be replaced with cryptographically generated pseudonyms to keep some form of secret association with original data. In this particular approach, original data or measurements are presumed to be held separately and securely from processed data. However, this technique appears to be insufficient alone for maintaining the privacy, because an attacker can make a data analysis of the open measurements when data from different IoT devices for the same patients are collected together.

As presented in [5], there are different types of obscuring anonymising strategies. They include: replacing data with synthetic one [8], data swapping [6], imputation of gaps [22] or noise [9], rounding, binning and suppression [12,13]. In our context, rounding (binning) is considered as a simple and 'fair' deterministic approach of data generalisation (i.e. refer to numerical discrete and continuous data type) that does not require imputing any wrong information into the data, such as artificial noise. These strategies have to be measured with some quantitative criterion to ensure that the privacy defence is kept at some satisfactory level. Two the most well-known criteria are *k-anonimity* [25] and *differential privacy* [9]. An example of privacy defence for time series MIoT data is shown in [16] based on differential privacy framework. In this certain work, data are collected from one sensor and the goal is to prevent identifying the small time changes by the attacker. In addition, we need to involve the prior knowledge about feature dependence into differential privacy as some different sensor measurements may be correlated. Necessity to modify the differential privacy approach [9] for dependent features is stated and partially addressed in [14].

On the other side, the meta-data aspect becomes very critical since the meta-data may possibly contain some information which can be linked to an individual or group resulting in privacy violation. Madaan et al. [15] discuss the impact of meta-data on the privacy goals. This work demonstrates the potential role of meta-data which can play in constructing prior information threatening privacy, and how to mitigate this risk by adopting a differential privacy framework.

3 Meta-Data

For interoperability and integration purposes, most well known IoT standards incorporate meta-data to describe IoT devices and its interactions, for example, HyperCAT [4] and Web of Thing (WoT) - Thing Description (TD) [2]. Meta-data involved in the medical IoT system without mindful consideration appears to be intuitively harmless to data privacy. However, nowadays because of new technological innovations, meta-data can be easily compiled and analysed, leading to disclosure of sensitive information: device attributes, patient location, etc. Analysed meta-data of a given medical IoT system enables adversaries to establish a comprehensive profile of a patient's location, medical conditions, medical devices in use, etc. Also, some meta-data may contain some explicit and implicit potentials (e.g. some semantics) in drawing a picture about patients patterns of behaviour, interactions, and associations, exposing even more about that patient than the content of his\her medical conditions. Therefore, meta-data becomes a significant source of knowledge which can be encoded and exploited by adversaries to threaten privacy and this element needs to be taken into account seriously when tackling privacy issues in the medical IoT system.

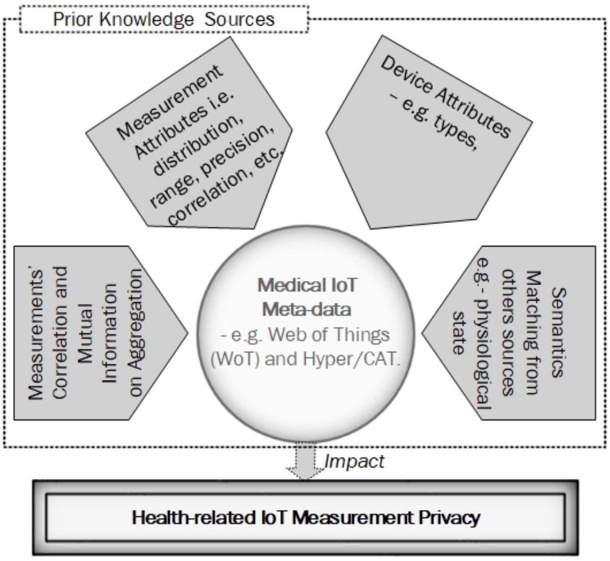
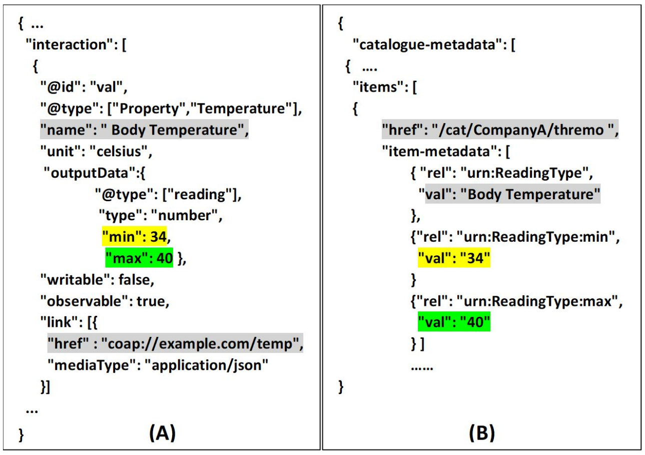

Prior Knowledge Sources
Measurement Attributes i.e. distribution, range, precision, correlation, etc,
Device Attributes - e.g. types,
Measurements' Correlation and Mutual Information on Aggregation
Medical IoT Meta-data - e.g. Web of Things (WoT) and Hyper/CAT.
Semantics Matching from others sources e.g. - physiological state
Impact
Health-related IoT Measurement Privacy

Fig. 1. Meta-data capabilities on privacy

To realise the significance of meta-data privacy leakage, we rely in this work on two well-known IoT standards, HyperCAT [4] and W3C Web of Things (WoT) [2]. The HyperCAT standard introduces an extensible, lightweight JSON-based catalogue which facilitates description and discovery of IoT resources over the Web using REST APIs, meta-data, semantic annotations, and special URI conventions. This standard is proposed in order to enable distributed data sources (i.e. hubs) to be utilised collaboratively by applications in a uniform machine-readable format. While, Web of Things presents a versatile IoT standard refining the Internet of Things by integrating smart things not only into the Internet (the network) but into the Web (the application layer). This standard leverages platform-independent APIs for web developers and offers a means for different platforms to discover and inter-operate with each other. This depends on rich JSON-based meta-data, Thing Description (TD), to define the data and interaction models for applications, and the communications and security requirements for platforms to communicate seamlessly. Also, it is important to address that Web of Things is still in the early stage of fully incorporating semantics of things and the domain constraints associated with this semantics, seeking for building W3C's extensible RDF and Linked Data. Therefore, we propose an exhaustive set of related-measurement and device information which can be extracted from accessible meta-data and may be linked, causing a privacy risk to a given patient as shown in Fig. 1. To perceive meta-data capabilities, we use meta-data from HyperCAT or W3C Web of Things (WoT) as an illustrative example to support our claims. we identify four main properties: (1) device attributes, (2) measurement attributes, (3) measurement correlation / mutual information and (4) semantics matching from others sources.



```
{ ...
 "interaction": [
  {
   "@id": "val",
   "@type": ["Property","Temperature"],
   "name": " Body Temperature",
   "unit": "celsius",
    "outputData":{
            "@type": ["reading"],
             "type": "number",
             "min": 34,
              "max": 40 },
   "writable": false,
   "observable": true,
   "link": [{
    "href" : "coap://example.com/temp",
    "mediaType": "application/json"
   }]
 ...
}
```

(A)

```
{
   "catalogue-metadata": [
  {  ....
   "items": [
   {
         "href": "/cat/CompanyA/thremo ",
         "item-metadata": [
             { "rel": "urn:ReadingType",
               "val": "Body Temperature"
             },
             {"rel": "urn:ReadingType:min",
               "val": "34"
             }
             {"rel": "urn:ReadingType:max",
               "val": "40"
             } ]
             ......
}
```

(B)

Fig. 2. Two simple meta-data samples for (A) TD-WoT and (B) HyperCAT

Measurement attributes: It is noticeable that meta-data used in IoT context seems to enclose some useful information about attributes of individual measurement such as measurement prior distribution, ranges, precision (error rate) and correlation. However, these extracted attributes can be exploited and linked by adversaries resulting in a privacy violation of a patient associated with these measurement attributes. For example, learning about a range and probability density function of an individual measurement from meta-data along with incorporating a sample of measurement data reveals sensitive information about a patient medical status. Some types of measurements have the capability to show dependence which can be linked to the individual, for example, opening door, motion, etc. Fig. 2 presents two simple examples of meta-data representations of thermometer sensor using Web of Things and HyperCAT.

Device attributes: Meta-data may include device attributes which are unassociated with the measurement being generated, but likely to leak some sensitive information about the patient (e.g. location, portability, etc). Some attributes of a device like being wearable or using some communication medium and protocols (e.g. BLE, Zigbee, etc.) may indicate to the patient localisation.

Measurement Correlation / Mutual Information on Aggregation: In some cases in the system, measurements of different medical devices or sensors are usually aggregated or grouped in a gateway before sending them to the back-end storage. Knowing about this aggregation and some features of associated devices from meta-data, some information sensitive can be easily obtained or inferred (i.e. patient location, health condition). Also, some information about a possible influence of the physiological measurements on each other (e.g. having a certain blood pressure and heartbeat measurements can be connected with heart problems) can be a threat to the privacy see Sec. 5. For example, the W3C WoT standard, unlike the HyperCat standard, offers a structured and visible meta-data (i.e. TD of Gateway Servient) indicating where and how data is combined. Therefore, this particular information along with some extra information about some device attributes (e.g. TD for Bluetooth devices communicating to the same gateway) may reveal sensitive location properties. Being aware of aggregation of some certain types of medical measurements (e.g. ECG, body pressure, GPS, motion, etc.) may often expose some medical status or health problem.

Semantics Matching from others sources: most meta-data models are typically enriched with semantics (using common vocabularies) for more machine and human readability and involve some semantic annotations for facilitating resource discovery and knowledge reasoning. This information of semantics may enable adversaries to find and learn a lot of extra related-measurement and -device information either manually or automatically and this particular prior knowledge can be exploited to breach privacy. However, the current HyperCAT and Web of Things standards only rely on statics approaches incorporating the semantics by using semantics annotation for some properties in the meta-data (e.g. using JSON-LD in Web of Things whereas using RDF-based Uniform Resource Identifiers (URIs) which typically identify data sources in HyperCAT). For example, some particular semantic queries for the meta-data in these standards can be requested in order to discover more details about medical devices in the system including related measurement attributes and constraints.

4 Pseudonymisation

The diverse and pervasive nature of the measurement data in a medical IoT system incites several privacy threats (e.g. identification and attribute disclosures including information linking [27,11]) as a result of subject (i.e. patients and devices) asset association, intermediate data aggregation and system meta-data being incorporated. Also, for the purpose of system control and data utilisation, individual measurements in this particular system unlike a typical health systems demonstrate a strong link with the device and patient information (e.g. patient and device records IDs) whether directly or indirectly as most IoT devices and sensors is normally bound to a certain patient or vicinity. Therefore, pseudonymi-

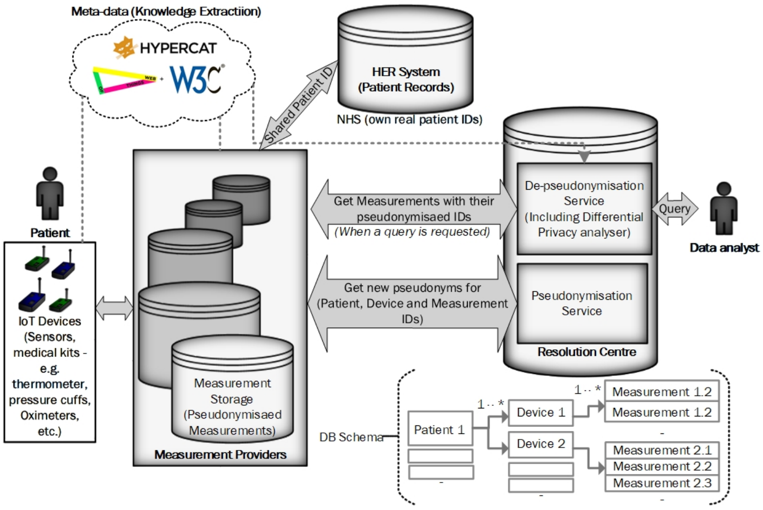


Fig. 3. The proposed distributed architecture: a pseudonymising system

sation offers a distinct approach to preserve privacy by anonymising data with the advantage of reversing this anonymisation if required. Pseudonymisation is a technique where all data identifiers should be changed by one or more artificial identifiers or pseudonyms [11]. Various approaches for patient pseudonymisation are proposed as shown in [3]. In this work, we adopt the simple approach of substituting real indentifiers with pseudorandom identifiers as some particular approaches in [3] appear to suffer from performance and management overheads because of heavily using crypto methods and intricate interactions. We propose a model of pseudonymised distributed storage offering an effective means to protect measurement privacy with keeping utility of measurement data competently, in medical IoT systems as shown in Fig. 3. This model mainly assumes there are multiple measurement providers (e.g. Third-party or IoT vendors) which typically store medical or health-related measurements collected from IoT sensors or gateways. Each measurement provider should keep measurements along with their own associated IoT devices' and patients' information in a pseudonymised form via using the pseudonymisation service provided by a resolution storage centre. In addition, all providers are assumed to share different pseudo-random identifiers with the resolution centre not the original identifiers associated with their patients for protection purposes in case a resolution centre is compromised. In other words, each provider must have a a table for mapping between its real IDs and their random IDs generated for sharing with the centre. The resolution centre has twofold roles. The first one is to generate and store random pseudonyms for the different identifiers (i.e patient, device and measurement IDs) whereas the second role is to control and resolve analyst queries with preserving privacy. The resolution centre relies on a set of master tables for mapping between different generated pseudonyms with their corresponding provider identifiers. Important to stress that adopting this specific pseudonymisation approach is to impede any privacy breaches coming from device and patient levels. On the other side, the de-pseudonymised service in the resolution centre as a second level of privacy protection involves a query analyser to anonymise results of a query sufficiently before posting to the data analyst. Also, the query analyser is used to restrict some queries if releasing those queries may lead to re-identification by exploiting some prior knowledge extracted from meta-data and query history, the details will be discussed in the next section.

Finally, we argue that our distributed storage model with providing purely randomised (not derived) pseudonyms for patient, device and measurement identifiers does not guarantee unlinkability, but it makes the process of linking a pseudonym to an individual very cumbersome and demand a lot of effort and resources. In our architecture, the resolution data centre is only a map between different measurement data repositories and patient and device repositories, so any compromise which may occur to the resolution centre or other repositories, will arguably not compromise the whole system. In other words, the key privacy advantage of a distributed system is avoiding a central point of data aggregation. Important to mention that the proposed architecture is assumed to have a standard access control managing access of front-end and back-end parties (e.g. RBAC) and also establish a secure communication between different endpoints.

5 Differentially Private Query De-Pseudonymisation

Even if the data records are pseudonymised, there is still a chance for adversaries to identify a patient from data analysis, i.e. analysis of the measurements generated by different sensors for a certain patient (see DB Schema diagram at Fig.3) collected at the resolution centre storage.

According to the scheme of data processing, we assume that the *Measurement Storage* (see Fig.3) collects together only the records and strictly pseudonymised references to patients. They should not include any other information related to the patients such as their meta-data or history of illness.

We also expect some input in the form of *queries* from a user (i.e. *data analyst*). The system contains a *query analyser* block (*Differential Privacy analyser*) which uses *prior knowledge* for decisions.

We assume that the data record is collected for a patient in the form of $(t_1, s_1, d_1), \ldots, (t_m, s_m, d_m)$ where t_i is a time stamp, s_i is a reference to a sensor, d_i is the numerical value of a *measurement.*

The prior knowledge of the data comes from knowledge resources such as meta-data (obtained by HyperCat and Web Of Things), and domain-specific expertise. We assume that it comes in the form of restrictions on possible *joint* distribution of the observations coming from different sensors.

5.1 Privacy Constraint

The principal way of keeping the privacy is an *anomymising strategy* $\mathcal{A}$ transforming the original measurement sequence to a form observable to the user. We prefer this strategy to be deterministic, but assume it applied to the measurement values only, while the schedule is open.

We consider the following version of differential privacy constraint for deterministic strategies, under prior knowledge. Let $D \in R^m$ mean an individual data record, $\mathcal{A}$ be a strategy, $\mathcal{P}$ be the set of possible joint probability density functions P on R^m. The $(\varepsilon, \mathcal{P})$-*differential privacy constraint* for $\mathcal{A}$ is

$$\forall P \in \mathcal{P} : Prob_P \left\{ \mathcal{A}(D) = \mathcal{A}(D') \right\} = \int_{D \in R^m s.t. \mathcal{A}(D) = \mathcal{A}(D')} P(dD) \geq e^{-\varepsilon}$$

where P generates m-dimensional data records D and D' independently of each other, $\mathcal{P}$ is the class of possible density functions on R^m according to the prior knowledge and the parameter ε (known as a *privacy budget*) quantifies strength of the constraint. This requirement means that the data record produced for a patient should be with high probability indistinguishable from another record generated for another patient with same (or similar) schedule of measurements.

The useful property of $(\varepsilon, \mathcal{P})$-differential privacy constraint is its decomposability, related to sharing the privacy budget over the queries. Assume that $q \leq m$ is the overall number of queries, $k_1, \ldots, k_q$ are the number of measurements addressed by the queries and

$$0 = \varepsilon_0 \leq \varepsilon_1 \leq \cdots \leq \varepsilon_{k-1} \leq \varepsilon_k \leq \cdots \leq \varepsilon_q \leq \varepsilon$$

where ε_k is a measure of the volume of information available for disclosure after first k queries. A way to satisfy the privacy constraint is dividing it into steps:

$$Prob_P\{\mathcal{A}(d_{q_1}, \ldots, d_{q_k}) = \mathcal{A}(d'_{q_1}, \ldots, d'_{q_k})$$
$$|\mathcal{A}(d_{q_1}, \ldots, d_{q_{k-1}}) = \mathcal{A}(d'_{q_1}, \ldots, d'_{q_{k-1}})\} \geq e^{-(\varepsilon_k - \varepsilon_{k-1})}.$$

5.2 Scheme of Differential Privacy Analyser

The central system has to include Differential Privacy analyser block for the queries in its Resolution Centre (see Fig.3). We assume that this block contains the prior information available for the work. In order to give a safe answer to the next query, it stores the history of preceding queries (Fig.4).

We introduce a anonymisation strategy of sequential binning (rounding) for the measurements in this architecture as most data are numerical:

$$\mathcal{A}_k : D = (d_{q_1}, \ldots, d_{q_k}) \to (b_1, \ldots, b_k) = B$$

where $b_j = \left[\frac{d_{q_j}}{r_j} \right] \times r_j$. Here square brackets mean eplacing a number with the closest integer. The values r_j is the *resolution level* (precision) for the j-th query.

Calculation of the resolution coefficients is linked to the order of queries: r_k is a function of the previously observed feature values $(d_{q_1}, \ldots, d_{q_k})$ but not of the later ones. We can also assume that, up to fixed prior knowledge, r_j is a function of $b_1, \ldots, b_{k-1}$ and d_{q_k} only, as shown by the arrows on the Fig.4. This way of

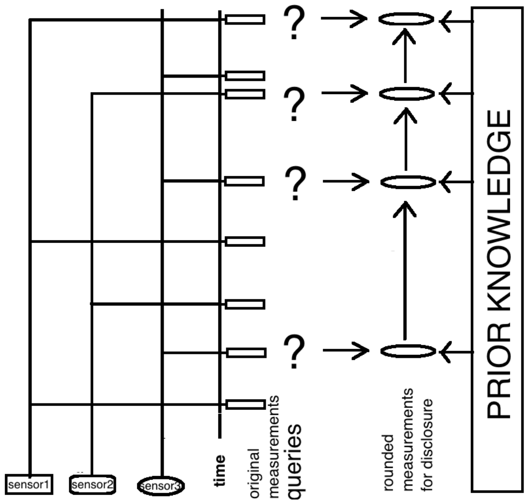


Fig. 4. Differential Privacy analyser for De-pseudonymisation service.

binning means that classes of indistinguishable records are represented in the vector space of possible records as multi-dimensional parallelepipeds ('bricks') of possibly different sizes.

The exact way of calculation depends on the form of the prior knowledge $\mathcal{P}$. In the sequential setting done above, the estimation of r_i becomes a relatively easy task for empirical estimation in the most typical cases. Especially, if $\mathcal{P}$ consists of only one or several distributions P one can make simulation of the conditional distribution of d_{q_k} given $b_1, \ldots, b_{k-1}$, and decrease r_i as far as the conditional privacy constraint is not empirically broken for any P. If $\mathcal{P}$ is a parametric distribution with some range of parameters, then it can be reasonably approximated by scanning over a grid within the allowed parameter range. We recommend users to give a desirable resolution level $\hat{r}_i$ which is sufficient, so that attempts for further decreasing r_i can be stopped when $r_i = \hat{r}_i$ is reached.

It is also required to select a strategy of sharing privacy budgets. Possible examples may be as follows. *Equal share:* for an initially fixed positive number q, $\min\{\varepsilon/q, \varepsilon_r\}$ is considered as *upper bound* for $(\varepsilon_k - \varepsilon_{k-1})$, that is either spent totally at a step, or decreased if the reachable $\hat{r}_k$ is smaller than r_k required by the user. *Share in geometric progression:* for an initially fixed $h < 1$, $h\varepsilon_r$ is considered as *upper bound* for $(\varepsilon_k - \varepsilon_{k-1})$; all the rest is done the same way as above. Those strategies can be modified in various ways, e.g. higher weights may be given to more important sensors.

6 Conclusions

We propose a prototypical privacy architecture integrating both pseudonymisation and anonymisation techniques in a Medical IoT system for a sake of protecting data privacy and maximising utility. A distributed pseudonymisated storage using pseudo-random identifier generators is developed to suite the distributed MIoT system as different sensors or IoT devices may be provided by different MIoT providers. On the other side, apparently, medical IoT meta-data like HyperCAT, Web of Things, etc. may become a key enabler to directly or indirectly infer about patient measurements leading to more privacy breaches in such a system. Therefore, we design a query analyser to perform the anonymisation stage and this particular analyser with considering prior knowledge (from meta-data, domain experts, measurement dependence, etc.) must control releasing queries requested by data analysts. In addition, the query analyser may use a quantitative method of disclosing information in reply to the queries, based on a limited privacy budget for a differential privacy model.

One direction of the future work is to tackle different types of measurements, for example, textual or categorical. The differential privacy model can be developed further e.g. addressing the leakage of information through the schedule. The may involve elements of inter-feature binning suggested in [16].

Acknowledgments

This work was supported by Technology Integrated Health Management (TIHM) project awarded to the School of Mathematics and Information Security at Royal Holloway as part of an initiative by NHS England supported by InnovateUK.

References

1. HealthKit — Apple Developer Documentation, `https://developer.apple.com/documentation/healthkit`
2. W3C Web of Things Architecture, `https://w3c.github.io/wot-architecture/#sec-building-blocks-thing-description`
3. Aamot, H., Kohl, C.D., Richter, D., Knaup-Gregori, P.: Pseudonymization of patient identifiers for translational research. BMC Medical Informatics and Decision Making 13(1), 75 (Jul 2013)
4. Beart, P., Jaffey, T., Davies, J.: Hypercat 3.00 Specification (2016), `http://www.hypercat.io/standard.html`
5. C.M.O'Keefe: Protecting confidentiality while making data available for research and policy analysis, `http://www.bioss.ac.uk/rsse/2016/15Nov2016RSS_Protecting.pdf`
6. Dalenius, T., Reiss, S.P.: Data-swapping: A technique for disclosure control. 6(1), 73–85 (1982)
7. Dimitrov, D.V.: Medical Internet of Things and Big Data in Healthcare. Healthcare Informatics Research 22(3), 156–163 (Jul 2016)
8. Duncan, G.: Statistical confidentiality: Is synthetic data the answer? (2006), `http://slideplayer.com/slide/9374068/`, in UCLA IDRE:UCLA
9. Dwork, C., Roth, A.: The Algorithmic Foundations of Differential Privacy. Foundations and Trends in Theoretical Computer Science 9(3/4), 211–407 (Aug 2014)

10. El Emam, K., Jonker, E., Arbuckle, L., Malin, B.: A Systematic Review of Re-Identification Attacks on Health Data. PLOS One 6(12), 1–12 (Dec 2011), Correction published in PLOS ONE 10(4)e0126772
11. Garfinkel, S.L.: NISTIR 8053. de-identification of personal information. Tech. rep., Technical report, National Institute of Standards and Technology (NIST) , Gaithersburg, MD, USA (2015)
12. HESA: Rounding and suppression to anonymise statistics, `https://www.hesa.ac.uk/about/regulation/data-protection/rounding-and-suppression-anonymise-statistics`
13. Lin, Z., Hewett, M., Altman, R.B.: Using binning to maintain confidentiality of medical data. Proceedings. AMIA Symp. pp. 454–8 (2002)
14. Liu, C., Chakraborty, S., Mittal, P.: Dependence Makes You Vulnerable: Differential Privacy Under Dependent Tuples. In: Network and Distributed System Security Symposium (2016)
15. Madaan, N., Ahad, M.A., Sastry, S.M.: Data integration in iot ecosystem: Information linkage as a privacy threat. Computer Law & Security Review (2017)
16. M.Hadian, X.Liang, T.Altuwaiyan, M.M.E.A.Mahmoud: Privacy-Preserving mHealth Data Release with Pattern Consistency. IEEE Global Communications Conference pp. 1–6 (2016)
17. Narayanan, A., Shmatikov, V.: Myths and Fallacies of "Personally Identifiable Information". Communications of the ACM 53(6), 24–26 (Jun 2010)
18. Neubauer, T., Kolb, M.: An evaluation of technologies for the pseudonymization of medical data. In: Stud. Comput. Intell. vol. 208, pp. 47–60 (2009)
19. NOMINET: Privacy guidelines for IoT: what you need to know, `https://www.nominet.uk/researchblog/privacy-guidelines-iot-need-know-infographic/`
20. Par, G., Moqadem, K., Pineau, G., St-Hilaire, C.: Clinical Effects of Home Telemonitoring in the Context of Diabetes, Asthma, Heart Failure and Hypertension: A Systematic Review. J. Med. Internet Res. 12(2), e21 (2010)
21. Rahmani, A.M., Gia, T.N., Negash, B., Anzanpour, A., Azimi, I., Jiang, M., Liljeberg, P.: Exploiting Smart e-Health Gateways at the Edge of Healthcare Internet-of-Things: A Fog Computing Approach. Future Generation Computer Systems 78(2), 641–658 (Jan 2018)
22. Reiter, J.: Simultaneous Use of Multiple Imputation for Missing Data and Disclosure Limitation. Survey Methodology (2004)
23. Riazul Islam, S.M., Daehan Kwak, Humaun Kabir, M., Hossain, M., Kyung-Sup Kwak: The Internet of Things for Health Care: A Comprehensive Survey. IEEE Access 3, 678–708 (2015)
24. Selander, G., Mani, M., Kumar, S.: RFC 7744 - Use Cases for Authentication and Authorization in Constrained Environments. Tech. rep., Internet Engineering Task Force (IETF) (May 2016), `https://tools.ietf.org/html/rfc7744`
25. Sweeney, L., P.Samarati, Pierangela: Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression. Harvard Data Privacy Lab (1998)
26. Tarouco, L.M.R., Bertholdo, L.M., Granville, L.Z., Arbiza, L.M.R., Carbone, F., Marotta, M., de Santanna, J.J.C.: Internet of Things in healthcare: Interoperability and security issues. In: 2012 IEEE Int. Conf. Commun. pp. 6121–6125. IEEE (Jun 2012)
27. Ziegeldorf, J.H., Morchon, O.G., Wehrle, K.: Privacy in the Internet of Things: Threats and Challenges. Security and Communication Networks 7(12), 2728–2742 (Dec 2014)