Securing Data Provenance in Body Area Networks using Lightweight Wireless Link Fingerprints

Syed Taha Ali,
Vijay Sivaraman
University of New South Wales
Sydney, Australia
{taha,
vijay}@unsw.edu.au

Diethelm Ostry
CSIRO Computational Informatics
Sydney, Australia
diet.ostry@csiro.au

Sanjay Jha
University of New South Wales
Sydney, Australia
sanjay.jha@unsw.edu.au

ABSTRACT

Wireless bodyworn sensing devices are becoming popular for fitness, sports training and personalized healthcare applications. Securing the data generated by these devices is essential if they are to be integrated into the current health infrastructure and employed in medical applications. In this paper, we propose a mechanism to secure data provenance for these devices by exploiting symmetric spatio-temporal characteristics of the wireless link between two communicating parties. Our solution enables both parties to generate closely matching 'link' fingerprints which uniquely associate a data session with a wireless link such that a third party, at a later date, can verify the links the data was communicated on. These fingerprints are very hard for an eavesdropper to forge, lightweight compared to traditional provenance mechanisms, and allow for interesting security properties such as accountability and non-repudiation. We validate our solution with experiments using bodyworn devices in scenarios approximating actual device deployment, and we present optimization mechanisms. We believe this is a promising first step towards using wireless-link characteristics for data provenance in body area networks.

Categories and Subject Descriptors

E.0 [**Data**]: General

Keywords

Body Area Networks; Data Provenance

1. INTRODUCTION

Body area networks is an emerging technology paradigm that is anticipated to revolutionize the healthcare domain and significantly reduce soaring national health expenditures. Miniaturized, unobtrusive sensors worn on the body allow for mobility, remote monitoring, and reduce the burden on hospital and professional staff. This technology has already found popular application in sports and fitness training, and in lifestyle monitoring. Examples include the Nike+ FuelBand [1], Fitbit Flex [2], the Toumaz Sensium Digital Plaster [3], and the Natalia Project [4]. Several companies, including Apple [5], are reportedly innovating in this field, and ABI Research [6] predicts that shipments of disposable wireless sensors are expected to reach 5 million by 2018.

A typical body area networks topology consists of bodyworn sensors communicating wirelessly with a handheld device or an off-body basestation which forwards the data to an online database to be accessed and analyzed by professionals. Miniaturized sensor devices are severely resource-constrained, unable to deploy traditional cryptographic solutions because of the high overheads, and are the weakest link in this architecture. However, ironclad security is needed because these devices deal in personal medical data, wrongful disclosure of which can result in serious ethical and legal implications. Developing lightweight security solutions for these devices is therefore a popular research area.

Thus far there has been significant research in maintaining confidentiality and authenticating sensor data. However, for these devices to integrate successfully into the healthcare infrastructure and for patients and medical professionals to trust this data, there are other guarantees that must be provided, such as information about the data, sensor-patient association, sensor-device association, which parties handled the data, etc. This metadata relating to how the sensor data is generated, manipulated and communicated falls in the purview of *data provenance*.

Consider the case of a user, Alice, who has had a heart attack and is informed by her insurance provider that they will cut her insurance rates if she gives up smoking. To ensure she complies, they hand her a wearable sensor device to monitor her for a trial period. This device periodically sends readings to her smartphone which forwards them to an online database. Thwarting this mechanism to secure benefits is easy: Alice could easily hack into the phone and forge her readings. Or she could record some readings and replay those into the network to cover up smoking episodes. Or, in the identity transference attack, she could even affix the sensor to a non-smoker friend for the duration of the trial without anyone finding out.

In this paper we develop a novel provenance solution for bodyworn devices. Provenance allows for an assessment of the trustworthiness of the data, which can prove critical especially in data forensics. In Alice's case, it would be useful to reliably know certain information about the data such

TrustED'13, November 4, 2013, Berlin, Germany.

http://dx.doi.org/10.1145/2517300.2517303.

as, for example, the common sensor data off-load points: it would enhance trustworthiness of the medical readings to discover that Alice's sensor's first point of contact is usually her personal mobile phone, her home WiFi access point, a basestation device in her office, or even her regular gym. In the event of an incident, investigators should be able to return to the archives, check data associations, reconstruct scenarios, identify faults and assign liability.

Existing work in provenance for sensor devices, reviewed in the next section, generally relies upon frequent use of cryptographic mechanisms which is impractical for resource-constrained sensor devices [7]. In contrast, we suggest an information-theoretic approach: we propose that wireless channel characteristics between sensor device and basestation be used to generate 'link fingerprints'. Characteristics of the wireless link, such as radio signal strength or signal phase, are unique to the two communicating parties, very hard for an eavesdropper to forge, and can be leveraged to provide a shared and provable record of data sessions between two devices.

If both parties were to digitally sign the data they exchange and their corresponding link fingerprints, this would authenticate the session, such that forensics, at a later date, could verify the sensor-basestation association for the transaction, effectively confirming that said data was transmitted over that particular link. This process is **secure** in that the fingerprints cannot be forged, and it is **lightweight** compared to cryptographic alternatives. Unlike existing provenance solutions, our scheme also provides **accountability**, i.e. the user of the sensor device can verify that provenance information has not been tampered with or hacked en route in the network. This is especially applicable where data flow is from basestation to sensor device, i.e. in certain scenarios or for bodyworn actuator devices (such as insulin pumps), Alice (or her doctor) may choose to reprogram her bodyworn device, and, the signed link fingerprint would be **non-repudiable** evidence of that operation.

In this paper, we take initial steps towards developing such a protocol solution and demonstrate a proof of concept. Our specific contributions are:

1. We present a data provenance protocol using wireless channel characteristics to generate link fingerprints.
2. We present experimental results confirming that this solution can generate unique and near-perfect matching link fingerprints for typical data exchanges between bodyworn sensor device and off-body basestation.
3. We present optimization mechanisms that significantly reduce the memory and transmission overheads in handling link fingerprints, making this scheme feasible for resource-constrained devices.

Our results indicate that, in a typical usage environment, two parties can generate a usable 128-bit link fingerprint approximately every $10 \sim 15$ minutes which eavesdroppers are unable to replicate. We believe this is a promising first step in using wireless link characteristics to enable secure data provenance.

This paper is organized as follows: Section 2 covers prior work in this domain. Our link fingerprint protocol is described in Section 3. The fingerprint generation technique is experimentally validated in Section 4, and we suggest optimization mechanisms in Section 5. We conclude in Section 6.

2. PRIOR WORK

Provenance may be defined as a record of the origin and evolution of data within a system, and has application in terms of contextualizing the data, and evaluating its trustworthiness in an objective manner. It is key for digital forensics: with the surge in computer crime, provenance is critical in reconstructing incidents and assigning liability. This also motivates the need for *securing* provenance (distinct from generating it), as discussed in [8].

The *granularity* of provenance varies as per application requirements and device capability: [9] makes the case for 'high-fidelity' provenance, compiled at the kernel level, enabling very detailed forensics analysis. On enterprise networks, administrators can log file and system operations in detail. However, on resource-constrained mobile devices, all that might be possible is a digital signature or a timestamp association. Additionally, provenance need not be binary: especially in distributed environments, such as large multi-hop sensor networks, it may be more practical to express confidence in sensor data using a probability value [10] or a trust score [11] [12].

In body area networks, provenance has mostly been limited to verifying data-sensor and data-patient associations. Chowdhury et al. [13] survey existing research on associating sensor data with the human subject, and they consider different authentication solutions, typically relying on frequent use of costly mechanisms such as expensive cryptographic protocols, trusted third parties, and additional sensor node capability (e.g. biometrics readers). The authors in [14] amortize digital signatures for bodyworn devices, enabling a secure and irrevocable binding between patient data and the device from which it originated. [15] proposes binding patient data to the subject's own unique vital signs readings (real-time ECG and accelerometer data), enabling user authentication in a continuous manner. However, none of these schemes explicitly address the data path or link association between two parties.

Radio fingerprinting techniques [16] may be used to identify a transmitting party (with up to 70% probability) by examining its radio signal, but require a strictly stationary deployment, specialised sampling hardware and a database to train the system. This approach is not scalable to multi-hop networks, and a sophisticated attacker may even forge legitimate radio fingerprints.

Identifying the links the data traverses in a wireless network is examined in [17], a provenance mechanism in which intermediate nodes in a multi-hop sensor network use Bloom filters to imprint path information on transit packets such that the basestation is able to verify the path each packet takes in the network. This concept is similar to that of secure routing protocols: for instance, in the Secure Ad-hoc On-demand Distance Vector (SAODV) [18], routers in an ad hoc network append digital signatures to all routing information packets so that other routers in the network are able to verify the route the packet has taken.

Our mechanism differs primarily in that we generate provenance concerning the wireless link association between two parties, on a *per session* basis. This makes minimal use of cryptographic operations and is ideal for resource-constrained devices. In much of the earlier work, digital signatures and/or encryption is used on a per packet basis, whereas in our case, these operations are performed only after a usable link fingerprint is derived, approximately every fifteen

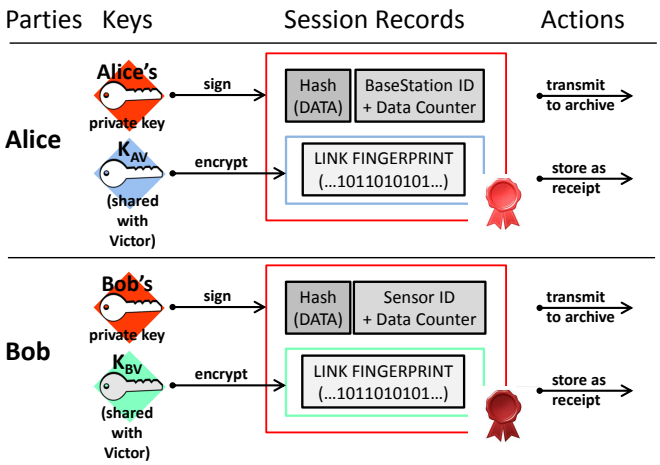


Figure 1: Protocol for Alice and Bob

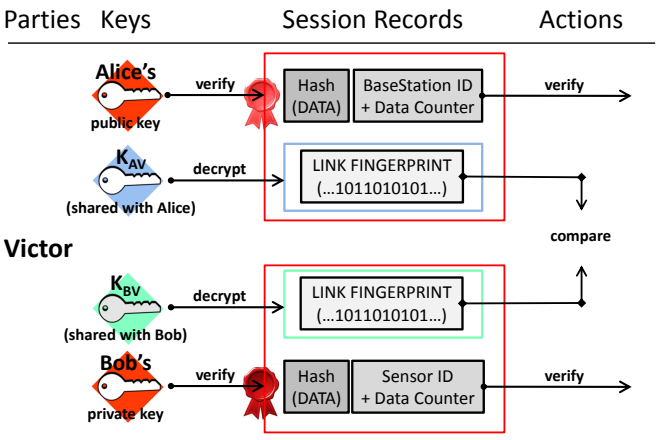


Figure 2: Protocol for Victor

minutes (every 900 data packets). The fingerprint is generated as a function of every packet reliably exchanged over the link, by exploiting wireless channel characteristics.

3. PROTOCOL FOR LINK FINGERPRINTS

We used the example of Alice earlier to motivate the need for associating a sensor device with the basestation for the duration of a data transaction. Theoretically, provenance protocols can be built using cryptographic primitives, but, as we noted earlier, these techniques are not practical for resource-constrained devices and can only be used sparingly. These devices are also lightweight with very small form factor to easily fit on the body (the Sensium device weighs less than 10 grams) and it is not possible to attach locationing technology, such as GPS receivers, onto them. A lightweight solution is needed that does not require extra hardware and can still provide strong security guarantees.

Recently there has been considerable interest in using the wireless physical layer between two devices to construct security primitives. The essential idea here is that wireless channel characteristics are symmetric for two parties, say Alice and Bob (and preclude an eavesdropper, Eve), and highly sensitive to spatial-temporal changes. If Alice and Bob sample the variation in channel characteristics over a period of time, they can use the measurements as a shared 'channel signature' for a range of functions including secret-key agreement [19], weak authentication, detecting certain attacks, masking channel state [20], intrusion detection [21], and location distinction [22]. These wireless channel-based techniques have also been successfully applied to body area networks, specifically to generate low-cost shared secret keys [23] and on-body validation of device-patient association [24]. An early work in this domain [20] has argued that these lower layer techniques complement cryptographic mechanisms, and can be employed to augment system security.

We use the 'channel signature' to uniquely fingerprint the Alice-Bob link and associate it with the data they exchange, such that a verifying party, Victor, is able to confirm wireless data-to-link associations in the network, potentially at a later date. The fingerprint is a unique bitstring of pre-configured length, generated individually by Alice and Bob sampling their common wireless link, and is therefore highly correlated between the two. Details of how both parties generate the fingerprints are provided in the following sections. Here we discuss how the fingerprints can be used as a building block in a protocol for data provenance.

The protocol for the bodyworn sensor device and basestation, denoted as Alice and Bob respectively, is depicted in Fig. 1. It is essential that Alice and Bob encrypt their fingerprints to keep them private from unauthorized parties. If the signature were transmitted in plain sight, an attacker could possibly copy it and claim the link association with itself, raising confusion as to which was the actual data offload point. Alice and Bob should also not be able to view each other's fingerprint: if any of them were maliciously inclined, they might share the fingerprint with an attacker or use the fingerprint and its signature binding to try and mount a replay attack. An easy way to protect the fingerprint is for each party to encrypt their own with a key they share only with Victor, in this case, K_{AV} for Alice and K_{BV} for Bob.

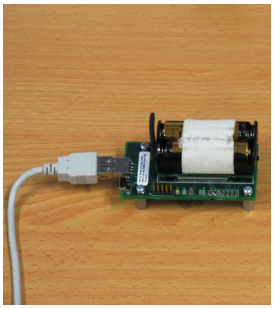
Once the fingerprint is encrypted, it is bundled with a hash digest of the data and session identifiers (timestamp or counter value, identity of the devices, etc.), into a *session record* which is digitally signed, and transmitted to the database. The signature ensures both parties commit to the data and the link association in a non-repudiable way.

Furthermore, each party can retain a copy of this signed session record as a *receipt* for the transaction, enabling system-wide **accountability**. In existing provenance solutions, sensor devices usually offload trust on to the network or third parties (as in [11] and [10]), and if trusted elements collude for malicious purpose, the sensor device has no way to prove it. However, the symmetric property of the link fingerprint and the digital signature binding enables the sensor to verify at a later date if a record has been tampered with.

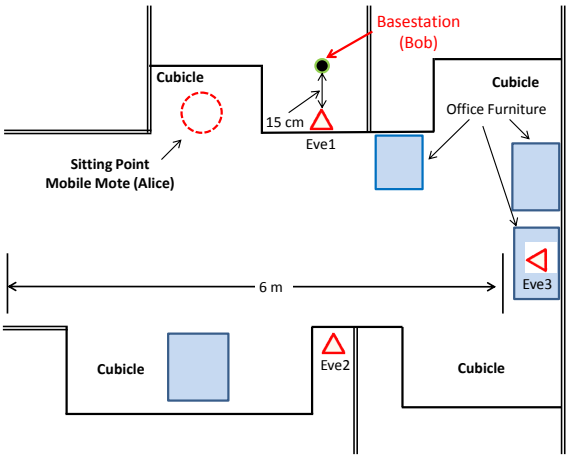
Fig. 2 depicts the verification process for Victor: as part of a forensics investigation, Victor could revisit archived session records, confirm the digital signatures and session identifiers for the data items of interest, decrypt Alice and Bob's link fingerprints using the individual symmetric keys he shares with them to check their agreement. If there is a very high correlation for the two fingerprints, he can be certain that Alice and Bob used the wireless link between them to communicate that particular data item.

Furthermore this scheme also identifies man-in-the-middle attacks. If Eve were to insert herself as a relay point between

(a) Mobile mote (Alice)



(b) base-station (Bob)



(c) Layout of experimental setup

Figure 3: Mobile node, base-station and experimental layout for indoor environment

Alice and Bob without their knowledge, the legitimate Alice-Bob communication would span two different wireless links (Alice-Eve and Eve-Bob) and the resulting Alice-Bob fingerprints would therefore disagree with very high probability.

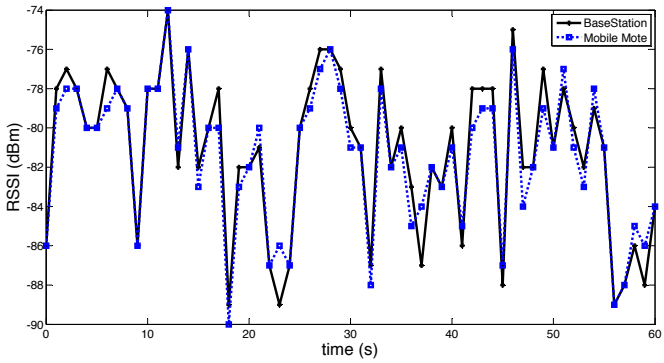
We note here that this is not a comprehensive data provenance solution but a first step in that direction. Our aim is to demonstrate how this technique may be used as an important building block in actual solutions, and highlight its security properties. In the next section, we perform experiments to validate that usable link fingerprints can be generated using wireless channel characteristics.

4. EXPERIMENTAL VALIDATION

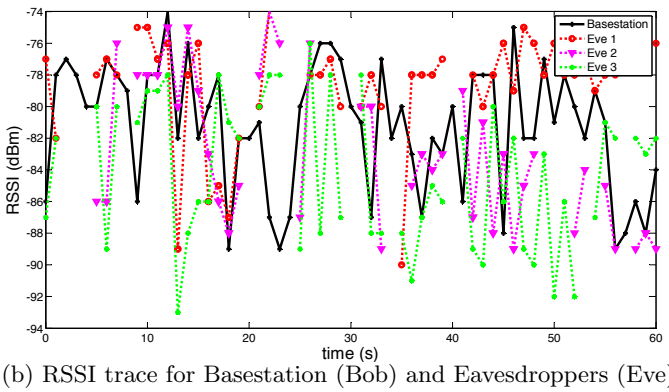
We used MicaZ motes, running TinyOS, and operating in the 2.4 GHz band. These radios provide received signal strength indicator (RSSI) values, which is a measure of the signal power in logarithmic units, and suffices for generating link fingerprints. To recreate an actual bodyworn sensor deployment, we mounted the device on a human subject Alice (on the upper right arm as shown in Fig. 3(a)), that communicates with an off-body basestation Bob (pictured in Fig. 3(b)). Our indoor environment is an office space with multiple cubicles, furniture and people. The layout is depicted in Fig. 3(c), marking out the locations of the basestation (Bob) and three eavesdroppers (Eve1, Eve2, Eve3) at a distance of greater than one wavelength away from the legitimate parties. The bodyworn device transmits packets at the rate of 1 packet/second, typical for healthcare devices transmitting physiological information such as heart-rate, ECG, etc. The basestation responds within $10 \sim 20$ milliseconds with an acknowledgement to every message. This routine device communication enables both parties to sample the wireless link in succession and record the RSSI values.

Our experiments consist of two activity modes: *High Activity* where the subject, Alice, walks around the office space to different cubicles, and interacts with other people in the room, and *Low Activity* where she is mostly seated at her cubicle, occasionally getting up to fetch items from other cubicles. For each activity, we collect RSSI trace readings for the bodyworn sensor device (Alice), basestation (Bob), and eavesdroppers (Eve1-3), spanning approximately 40 minute periods, which we analyze offline with Matlab.

We provide in Fig. 4 a one-minute sample of the RSSI trace for the *High Activity* experiment. It is observed in Fig. 4(a) that the bodyworn sensor device, Alice, and the basestation, Bob, channel measurements are highly correlated. Eavesdroppers, however, experience a different channel and are unable to replicate the RSSI profile, as shown in Fig. 4(b). This follows from Jakes uniform scattering model [25], which states that there is rapid signal decorrelation at distances of over half a wavelength, and independent signals can be assumed for a separation of one or two wavelengths and more. In the case of the 2.4 GHz band, this indicates that if Eve is at a distance greater than 13 cm of Alice or Bob, she will experience different fading characteristics than the legitimate Alice-Bob channel and not be able to deduce the channel profile. For this reason, research solutions based on wireless link characterization stipulate as part of the threat model that eavesdroppers be situated at least two wavelengths away from the legitimate parties.



(a) RSSI trace for Basestation (Bob) and Mobile Mote (Alice)



(b) RSSI trace for Basestation (Bob) and Eavesdroppers (Eve)

Figure 4: Sample of RSSI trace for *High Activity*

Table 1: Correlation coefficient (r) of RSSI measurements observed by various parties

Experiment	Alice-Bob (r)	Alice-Eve1	Alice-Eve2	Alice-Eve3
High Activity	0.974	0.197	0.088	0.038
Low Activity	0.950	0.129	0.102	0.158
High Activity (filtered)	0.986	0.281	0.118	0.065
Low Activity (filtered)	0.976	0.205	0.152	0.224

To quantify the correlation for channel measurements for the different parties, we compute the Pearson correlation coefficient r:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}.\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

where X_i and Y_i are the RSSI values of the ith packet of each party and $\bar{X}$ and $\bar{Y}$ are the respective mean RSSI values of a sequence of n packets. The correlation coefficient r returns a value in $[-1, 1]$ where 1 indicates perfect correlation, 0 indicates no correlation, and -1 indicates anti-correlation. This metric is ideal for channel profile characterization in that it measures variations and not absolute values, and is therefore unaffected by offsets in RSSI measurements arising from differences in receiver sensitivities or transmit powers.

The results are presented in Table 1. Again, it is observed that there is strong correlation (well in excess of 0.9) between the legitimate parties (Alice and Bob), whereas it is poor for the eavesdroppers (typically below 0.2). We also present results for filtered versions of the channel profile. Filtering is useful because it reduces nonsymmetric discrepancies between the two parties (due to elements such as random noise, thermal effects, etc.) and has been recommended in the literature [26] [27]. For our purposes we use the Savitzky-Golay filter (other filters such as moving average techniques can also be used) and correlation is seen to improve slightly, giving higher confidence (> 0.95 correlation) in the shared fingerprint between Alice and Bob, while still keeping correlation low (< 0.3) for eavesdroppers.

These results are as expected and in good agreement with prior work. We do not provide a thorough characterization of the wireless channel in this paper as it has been extensively documented in the literature: the interested reader is referred to a detailed study in [19] and specifically [23] for the off-body channel.

Considering the strong correlation between the bodyworn sensor device and the basestation, technically the RSSI measurements themselves could be the link fingerprint. Both parties, Alice and Bob, could simply encrypt and sign their RSSI measurements, and a third party, Victor could compare the RSSI trace results, measure the correlation coefficient (much as we have done), and if he calculates a value $r > 0.9$, he can be certain that the fingerprint is valid.

However, there are issues with using raw RSSI values as link fingerprints: for one, both parties will have to store every RSSI value for every transaction in memory which may not be feasible for memory-constrained sensor devices. For example, the MicaZ motes record RSSI in single byte-sized values and, at a sampling rate of 1 packet/second, would exhaust their 4 KB of RAM in little over an hour. Second, radio usage is a very expensive operation for these devices [28], and these RSSI measurements have to be offloaded from the sensor device as part of the session record, resulting in extra transmission overheads. In the next section we show how the storage and communication overheads can be dramatically reduced by leveraging known quantisation techniques for compressing the RSSI-based link fingerprint.

5. OPTIMIZATION AND DISCUSSION

Quantization is a signal processing technique that can efficiently distil the raw RSSI data to a much smaller and manageable size. Another advantage is that it has been well-studied in the literature (especially in the context of wireless channel-based secret-key generation [19]) and can be designed to further reduce nonsymmetric noise components in the signals observed by Alice and Bob.

In generating link fingerprints, legitimate communicating parties Alice and Bob sample the wireless channel over a period of time to gather sufficient channel variation (or entropy) which is then quantized to yield a common bitstring. Quantization mechanisms typically consist of level crossing or ranking techniques and the operator of the scheme can choose one depending on application requirements. We describe an example of each approach here, validate them with our experimental RSSI traces, and compare their properties.

5.1 Level Crossing Quantization

Fig. 5 depicts a basic level-crossing quantizer (defined in prior work [19]). The bodyworn sensor device and the basestation define an adaptive moving window of size W_Q, within which consecutive (filtered) RSSI readings are processed. For each window, two threshold values are calculated:

$$q+ = \mu + \alpha.\sigma$$
$$q- = \mu - \alpha.\sigma$$

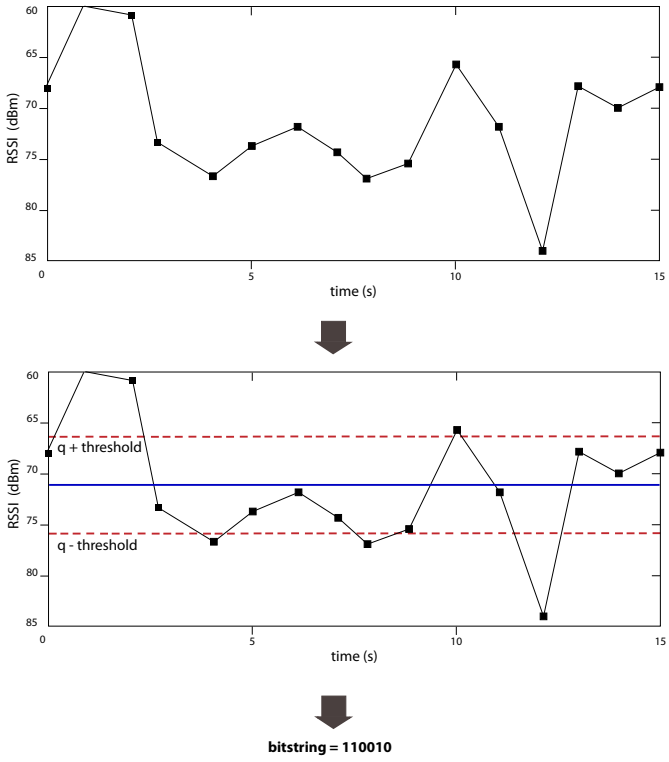


Figure 5: Level crossing quantization technique

where μ is the mean, σ is the standard deviation, and $\alpha \geq 0$ is an adjustable parameter. If an RSSI reading within a window exceeds $q+$, it is encoded as 1, and if less than $q-$, as 0. The thresholds define an exclusion zone and values lying in between them are discarded. This helps to further remove small scale discrepancies between the two endpoints, whereas there is usually very good agreement for excursions larger than the standard deviation. The α parameter allows the operator to adjust quantizer performance to balance between bit generation rate and bit agreement. For our purposes, we use a window size of $W_Q = 5$ and $\alpha = 1$, consistent with prior work.

5.2 Ranking Quantization

A multi-bit ranking quantizer is depicted in Fig. 6. The algorithm sorts the RSSI values in order to divide them into n equal-sized 'buckets' ($n = 4$ in this case). Each RSSI value in the original channel profile can then be encoded with $\log_2 n$ bits . Gray coding is used to number the buckets instead of binary coding, because successive values in Gray coding differ in only one bit, and will therefore limit small RSSI disagreements between Alice and Bob to a single bit per discrepancy.

5.3 Results

We perform the quantization process to generate fingerprints for the two activity modes for all parties using level crossing and ranking technique, and present results in Table 1. We briefly discuss here our findings and the metrics we use to evaluate our solution:

1. **Bit Agreement**: is the percentage of bits in the fingerprint that are matching between the bodyworn device and the basestation. As can be seen, this is 90% or greater for the legitimate parties and can be used to conclusively validate the link between them. Bit agreement is better in general for the level crossing quantizer because, unlike for the case of ranking where every RSSI value is quantized, the level crossing algorithm discards those values within the exclusion zone that are likely to cause disagreement.

2. **Bit Rate**: is the average number of bits that can be extracted from the channel per unit time. The ranking quantizer performs at a constant rate of 2 bits/s since all of the raw RSSI values are encoded. The level crossing technique exhibits a much lower rate because a single RSSI value can only be encoded to a single bit and several RSSI values are discarded. There is a tradeoff between bit agreement and rate.

3. **Minimum Session Length**: is the fingerprint length divided by the bit rate. The operator of the scheme can choose the length of the desired fingerprint. For level crossing, depending on the activity mode, it takes approximately 11 to 16 minutes to generate a 128 bit link fingerprint. For the ranking quantizer, which has a much faster rate, a fingerprint can be generated in approximately 2 minutes.

4. **Eavesdropper Agreement**: Fingerprints generated by the eavesdroppers should ideally match with the legitimate parties for 50% of the bits, which we see in our results. This translates to their knowledge of the legitimate fingerprint gained by eavesdropping as being no more useful than a tossing a coin.

5. **Entropy**: is a measure of the inherent randomness or uncertainty in the key. For a random variable X, over the set of n symbols $x_1, x_2, ..., x_n$, entropy is typically measured as follows:

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i)$$

where $p(x_i)$ is the probability of occurrence of symbol x_i. For binary symbols, a value close to 1 indicates high entropy, which we achieve in our results. Furthermore, the fingerprints we generate clear the entropy tests in NIST test suite [29], a battery of tests with a pass/fail result, typically used in the literature to confirm randomness of wireless channel-based secrets.

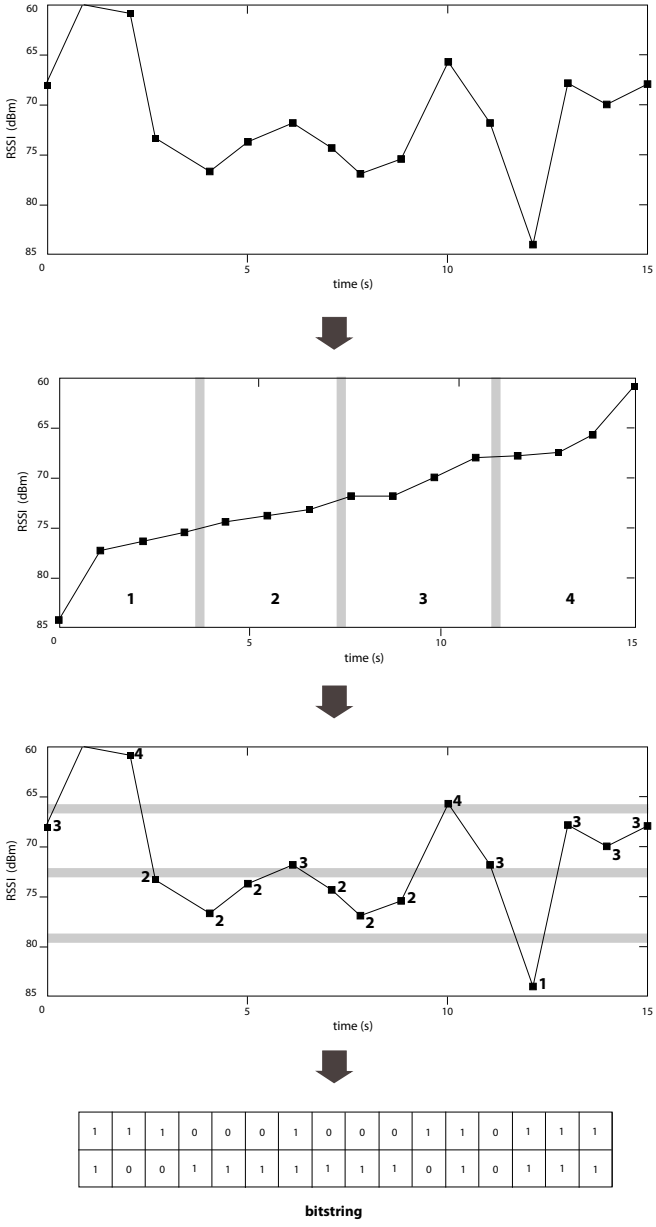


Figure 6: Ranking quantization technique

Activity (Quantization)	Fingerprint Agreement	Bit (bit/s)	Min. Session Length (mins)	Eve1 Agreement	Eve2 Agreement	Eve3 Agreement	Entropy
High Activity (Level Crossing)	98.40%	0.205	10.41	47.11%	46.48%	47.34%	0.997
Low Activity (Level Crossing)	95.53%	0.139	15.35	46.26%	46.80%	47.60%	0.997
High Activity (Ranking)	93.60%	2	2.13	44.39%	46.92%	48.74%	1
Low Activity (Ranking)	96.08%	2	2.13	50.54%	50.41%	52.92%	1

Table 2: Link fingerprint performance for experimental scenarios

We believe these are encouraging results which validate our proposed approach, and lay the grounds for future work. The tradeoffs between the different quantizers are also highlighted: ranking can be used to build lossless multi-bit quantizers with high bit generation rate, more suited to applications where the average session time between sensor device and basestation is low. A level crossing technique could be used for longer session times, and yields a fingerprint with higher agreement between the two ends. Quantizers could even be further customised to prioritize a desired metric as per application requirements.

In some settings, such as hospitals or gyms, where there are multiple basestations and possibilities for roaming, a bodyworn sensor device may form link associations with different basestations over a period of time. The sensor device worn by a hospital patient may communicate for the most part with the basestation in the ward, except when the patient visits the hospital cafeteria where it associates with another basestation. Associations may be very brief and frequently disrupted. In this case, the sensor device and individual basestations could maintain running counter values and incremental fingerprints for communications, such that a complete fingerprint may be generated and signed over multiple sessions between the sensor and basestation pair, only when sufficient data has been communicated between the two.

Furthermore, our protocol solution should be considered a starting step, and several variations are possible on this basic design: for instance, if session times are very short, and frequent use of a digital signature is too resource-intensive, it is possible to incorporate signature amortization techniques [14] to distribute a digital signature over several sessions.

We also note that it is possible to extend this concept to networks with multiple hops (such as mobile sensor networks, delay tolerant networks, etc.) and document the entire wireless path for a data item. If the sensor device transmits data to the basestation which in turn forwards it to another device using the wireless channel, each party in the path could generate the associated signature records and maintain receipts. The verifying party, Victor, could map out the entire wireless path by performing the fingerprint verification process for every link, and, in a loose sense, may even be able to track the mobile parties. Mechanisms could be developed to ensure that malicious parties in the path do not collude with each other.

We intend to explore all of these ideas in future work. We are also working on prototyping our solution to quantify the energy requirements, and study performance across a wider range of environments and activities.

6. CONCLUSION

In this paper we have proposed a data provenance protocol for bodyworn devices that exploits symmetric spatio-temporal characteristics of the wireless channel. Our solution generates unique link fingerprints that we use to form data to wireless link associations. In contrast to existing provenance mechanisms which operate on a per packet basis, this solution generates provenance on a per session basis, which minimizes the use of cryptographic techniques and associated overheads. The link fingerprints can be built using routine data transmissions, they are unique to the two communicating parties and cannot be deduced in detail by an eavesdropper situated at a distance. Our provenance solution also provides system-wide accountability and non-repudiation.

We performed experiments using bodyworn devices in an indoor office space to demonstrate the high correlation in channel measurements between the two endpoints. We suggest two optimization techniques, level crossing and ranking, to quantize the raw RSSI values to a manageable size, and we discuss possibilities for adapting the fingerprinting process for different application requirements. We believe this is a promising first step in using wireless link-based techniques to secure data provenance.

7. REFERENCES

[1] Nike+ FuelBand. Retrieved on 21 July, 2013. `http://www.nike.com/us/en_us/c/nikeplus-fuelband`.

[2] Fitbit Flex. Retrieved on 21 July, 2013. `http://allthingsd.com/20130715/fitbit-flex-vs-jawbone-up-and-more-a-wearables-comparison/`.

[3] Toumaz Technology Ltd. *Sensium Life Platform.* `http://www.toumaz.com/page.php?page=sensium_intro`.

[4] David Szondy. Bracelet Uses Social Network to Protect Civil Rights Activists. 7 April, 2013. `http://goo.gl/MIEk2`.

[5] Waiting for Apple's iWatch. 22 March, 2013. `http://goo.gl/xTsQO`.

[6] Disposable Wireless Sensor Market Shows Signs of Life: Healthcare Shipments to Reach 5 Million in 2018. 3 March, 2013. `http://goo.gl/V9QoP0`.

[7] A. S. Wander, N. Gura, H. Eberle, V. Gupta, and S. C. Shantz. Energy Analysis of Public-key Cryptography for Wireless Sensor Networks. In *3rd IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 324–328, 2005.

[8] U. Braun, A. Shinnar, and M. I. Seltzer. Securing Provenance. In *USENIX Summit on Hot Topics in Security (HotSec)*, 2008.

[9] D. J. Pohly, S. McLaughlin, P. McDaniel, and K. Butler. Hi-Fi: Collecting High-fidelity Whole-system Provenance. In *Proceedings of the 28th ACM Annual Computer Security Applications Conference*, pages 259–268, 2012.

[10] B. Przydatek, D. Song, and A. Perrig. SIA: Secure Information Aggregation in Sensor Networks. In *Proceedings of the 1st ACM International Conference on Embedded Networked Sensor Systems*, pages 255–265, 2003.

[11] H. Lim, Y. Moon, and E. Bertino. Provenance-based Trustworthiness Assessment in Sensor Networks. In *Proceedings of the Seventh ACM International Workshop on Data Management for Sensor Networks*, pages 2–7, 2010.

[12] H. Lim, G. Ghinita, E. Bertino, and M. Kantarcioglu. A Game-Theoretic Approach for High-Assurance of Data Trustworthiness in Sensor Networks. In *28th IEEE International Conference on Data Engineering (ICDE)*, pages 1192–1203, 2012.

[13] M. A. Chowdhury, W. Mciver, and J. Light. Data Association in Remote Health Monitoring Systems. *IEEE Communications Magazine*, 50(6):144–149, 2012.

[14] S. T. Ali, V. Sivaraman, and D. Ostry. Authentication of Lossy Data in Body-sensor Networks for Healthcare Monitoring. In *9th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, pages 470–478, 2012.

[15] J. C. Sriram, M. Shin, T. Choudhury, and D. Kotz. Activity-aware ECG-based Patient Authentication for Remote Health Monitoring. In *Proceedings of the ACM International Conference on Multimodal Interfaces*, pages 297–304, 2009.

[16] K. B. Rasmussen and S. Capkun. Implications of Radio Fingerprinting on the Security of Sensor Networks. In *Third International Conference on Security and Privacy in Communications Networks, SecureComm*, pages 331–340. IEEE, 2007.

[17] B. Shebaro, S. Sultana, S. R. Gopavaram, and E. Bertino. Demonstrating a Lightweight Data Provenance for Sensor Networks. In *Proceedings of the ACM Conference on Computer and Communications Security*, pages 1022–1024, 2012.

[18] M. G. Zapata. Secure Ad hoc On-demand Distance Vector Routing. *ACM SIGMOBILE Mob. Comput. Commun. Rev.*, 6(3):106–107, 2002.

[19] S. Jana, S. N. Premnath, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy. On the Effectiveness of Secret Key Extraction from Wireless Signal Strength in Real Environments. In *Proceedings of the 15th ACM Annual International Conference on Mobile Computing and Networking*, pages 321–332, 2009.

[20] Z. Li, W. Xu, R. Miller, and W. Trappe. Securing Wireless Systems via Lower Layer Enforcements. In *Proceedings of the 5th ACM Workshop on Wireless Security*, pages 33–42, 2006.

[21] J. Tang and P. Fan. A RSSI-based Cooperative Anomaly Detection Scheme for Wireless Sensor Networks. In *International Conference on Wireless Communications, WiCom*, pages 2783–2786. IEEE, 2007.

[22] N. Patwari and S. K. Kasera. Robust Location Distinction using Temporal Link Signatures. In *Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking*, pages 111–122, 2007.

[23] S. T. Ali, V. Sivaraman, and D. Ostry. Zero Reconciliation Secret Key Generation for Body-worn Health Monitoring Devices. In *Proceedings of the 5th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, pages 39–50, 2012.

[24] L. Shi, M. Li, S. Yu, and J. Yuan. Bana: Body Area Network Authentication Exploiting Channel Characteristics. In *Proceedings of the 5th ACM conference on Security and Privacy in Wireless and Mobile Networks*, pages 27–38, 2012.

[25] W. C. Jakes. *Microwave Mobile Communications*. Wiley, 1974.

[26] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik. Radio-telepathy: Extracting a Secret Key from an Unauthenticated Wireless Channel. In *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking*, pages 128–139, 2008.

[27] L. Yao, S. T. Ali, V. Sivaraman, and D. Ostry. Improving Secret Key Generation Performance for On-body Devices. In *Proceedings of the 6th International Conference on Body Area Networks*, pages 19–22. ICST, 2011.

[28] D. Culler, D. Estrin, and M. Srivastava. Guest editors' introduction: Overview of sensor networks. *Computer*, 37:41–49, 2004.

[29] NIST. *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*, 2001.