On Admitting Sensor Fault Tolerance while Achieving Secure Biosignal Data Sharing

Vishwa Goudar and Miodrag Potkonjak
Computer Science Department
University of California, Los Angeles

***Abstract*—Remote health monitoring BASNs promise substantive improvements in the quality of healthcare by providing access to diagnostically rich patient data in real-time. However, adoption is hindered by the threat of compromise of the diagnostic quality of the data by faults. Simultaneously, unresolved issues exist with the secure sharing of the sensitive medical data measured by automated BASNs, stemming from the need to provide the data owner (BASN user / patient) and the data consumers (healthcare providers, insurance companies, medical research facilities) secure control over the medical data as it is shared. We address these issues with a robust watermarking approach constrained to leave primary data semantic metrics unaffected and secondary metrics affected minimally. Further, the approach is coordinated with a fault tolerant sensor partitioning technique to afford high semantic accuracy together with recovery of biosignal semantics in the presence of sensor faults, while preserving the robustness of the watermark so that it is not easily corrupted, recovered or spoofed by malicious data consumers. Based on experimentally collected datasets from a gait-stability monitoring BASN, we show that our watermarking technique can robustly and effectively embed up to 1000 bit watermarks under these constraints.**

***Keywords*-Body Area Networks; Watermarking; Fault Tolerance; Medical Data Sharing; Medical Data Security;**

I. Introduction

Body Area Sensor Networks (BASNs) offer the promise of shortened time-to-treatment and reduced frequency and duration of hospitalization [1] by enabling unfettered and unintrusive collections of patient physiological and biomechanical signals over extended periods. Beyond health monitoring, the diagnostically rich datasets collected by these devices can support several potential killer apps in the healthcare domain. For example, the datasets can be used as supporting documentation in insurance claims, or as measures in assessing coverage and indemnity requirements. Data collected periodically with BASNs over long periods may also be selectively shared in lieu of compensation with medical researchers conducting clinical studies.

In each of these applications, the significance of the datasets lies in their semantics which are derived based on relevant functional/diagnostic metrics. While the ubiquity of semantically rich patient histories will undoubtedly follow the rising adoption of BASNs, two key problem stem from the processing and sharing of these datasets. First, faulty sensors/sensor readings are likely to render the semantic quality of the datasets worthless. Second, sharing medical datasets in the absence of infrastracture that supports secure data exchange is likely to result in patient privacy and security violations on a large scale. Specifically, we focus on three secure sharing use-cases: proof of ownership, wherein the data owner must prove that she/he is the originator of the data; data tracking, wherein the data owner must be able to trace unauthorized sharing of her/his biosignal data; and content authentication, wherein the data owner must prove whether the biosignal data has been maliciously altered as it is shared between consumers.

While cryptography-based systems have been proposed to address these security concerns, their security itself lies in the strength of their key-management protocols. Further, a malicious consumer may come into possession of the decryption keys through legitimate channels only to then illegitimately claim ownership of the data and subsequently share it with other consumers, possibly after altering it. Instead, we offer digital watermarking as an alternative medical data sharing security solution. Watermarking is defined as the practice of imperceptibly altering a Work to embed a message about that Work [2]. Embedding relevant information will provide solutions to each of out security use-cases, while obviating the need for key management, increasing overall security and reducing the transmission and storage overhead necessary to achieve data security.

The need to ensure high semantic quality of the data in the face of faulty sensors can severly constrain the data-sharing infrastructure's ability to support the targeted use-cases with robust watermarking techniques. Signal alterations necessary for the watermark may hinder the quality of the semantics derived from a watermarked dataset, and to address this we shall constrain our watermarking scheme in a manner that limits and/or prevents any such semantic degradation. Further, a majority of watermarking techniques, including the one we propose, derive their robustness to security threats by distributing the embedded message across the entire signal. Yet, a common approach to achieving semantic fault tolerance in sensor networks involves redundantly sampling the spatio-temporal signal and substituting faulty readings with accurate ones while deriving the signal's semantics. In doing so, parts of the watermark may be lost as well

978-1-4799-5701-9/14 $31.00 © 2014 IEEE
DOI 10.1109/ICHI.2014.44

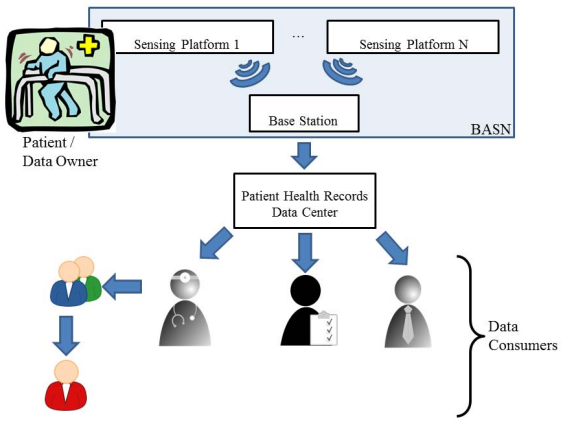


Figure 1: Body Area Sensor Network Architecture.

resulting in the eliminations of its security guarantees. To address this, our watermarking scheme will be a fault-tolerant one that will simultaneously minimize the impact on the dataset's semantic quality from the watermarking and the fault tolerance approaches.

In what follows, we will detail a linear programming based technique to watermark bio-signal data collected by BASNs, for the purposes of secure sharing of the data, such that (i) the watermarks are robust, (ii) the watermarks produce imperceptible semantic alterations to the data (i.e. the watermarks do not affect the diagnostic quality of the data), and (iii) the watermarks admit fault-tolerance induced by redundant sampling. It is necessary that the watermarking process produce robust watermarks so that they are not easily corrupted, erased or spoofed by malicious data consumers, and in this sense the robustness of a watermark is key to the level of security that the watermarking technique will provide. We study the performance of our technique in the context of foot plantar pressure datasets collected from multiple subjects by a wireless gait stability monitoring BASN. In this application context, we evaluate the technique in terms of embedding effectiveness, i.e. the probability of successful watermarking, embedding capacity, i.e. the number of bits that can be successfully embedded, and fidelity, i.e. the extent to which the original and watermarked works differ perceptually. We show that while altering the signal by less than 5%, up to 1000 bits can be embedded robustly and with high effectiveness ($> 99\%$) per stride of the HERMES user, despite several semantic constraints.

II. Related Work

The confidentiality and integrity of BASN collected patient data during exchange and storage within the network, as well as over the course of transmission outside the network for post-processing and storage at the data sink, have been identified as a key security issue for BASN deployments and addressed by a number of studies [3]. Further, secure sharing of patient health records has been studied in the context of data encryption-based solutions. The framework described in [4] provides mechanisms based on Attribute Based Encryption (ABE) for scalable key management, flexible access for multiple classes of data consumers, and efficient user revocation. Watermarking techniques have been thoroughly studied as a means to achieve proof of ownership and transaction tracking [2]. Primarily, these techniques have been applied to image, video and audio data, and include least significant bit alteration methods, quantization methods and signal decomposition based methods. These watermarking techniques have been extended to the domain of medical images, as well as physiological signal measurements including ECG and EEG signals. The authors of [5] compare and contrast three watermarking techniques with regards to their ability to verify EEG signal integrity after noise contamination resulting from communication. To our knowledge, our work is the first to propose robust watermarking of bio-signal data under perceptual and semantic constraints, to address security issues arising from data sharing rather than communication.

Fault tolerance is also a key performance criterion for BASNs. Often, fault-tolerant designs involve redundancy, wherein the coverage "area" is spanned by more sensors than are necessary, which enables the system to reconstruct the spatio-temporal signal profile of the entire coverage area from a functioning subset of sensors when others report faulty measurements. For example, sensor selection approaches have been proposed [6] wherein a sensor subset are constructed based on their ability to jointly predict measurements at all sensors. In this paper we take a similar approach towards fault tolerance with the construction of multiple redundant subsets such that the semantics of the entire spatio-temporal signal may be reconstructed accurately and indepedently from the individual subsets, while the subsets simultaneously admit effective and independent embedding of watermarks in data generated by each of the subsets.

III. Preliminaries

A. System Architecture

Fig. 1 illustrates a BASN architecture in terms of its data flow. A BASN typically consists of several body worn sensors/sensing platforms which continuously measure physiological and/or bio-mechanical signals and wirelessly transmit these measurements over a short range to a base station, such as a smart phone. The data is then forwarded by the base station to a data sink that logs the it to the patient's health records. A common use-case is one where the sensing platforms and the base-station collaboratively apply signal processing techniques to detect "anomalous" health events and alert the patient and/or the healthcare provider(s). Aside from such use-cases involving real-time data consumption, the data may also be consumed offline by its owner, her/his healthcare providers, insurance providers,

family, research organizations, etc. Further, the data may be shared between data consumers as well, after authorization has been provided by the data owner.

We shall study the performance of our watermarking technique in the context of a gait stability monitoring BASN called HERMES [7][8][6][9]. HERMES is a smart-shoe composed of 99 passive-resistive pressure sensors located across the shoe insole, that are used to monitor foot plantar pressure as the subject ambulates. Typically, HEREMES samples each of the plantar pressure sensors at 50Hz as the patient ambulates, and forwards the data to a smart phone via a low-power Bluetooth radio. Features extracted from the measurements of HERMES's sensors can be used to compute several functional gait metrics. For example, the foot contact metric of the GARS-M gait assessment scale [10] evaluates the heel strike angle, which can be estimated by HERMES based on the difference in time between heel strike and forefoot strike. Similarly, the staggering metric from the GARS-M scale determines laterally directed loss of balance, and can be monitored based on the plantar pressure difference between lateral areas of the foot. Other metrics that can be gleaned include the inter foot-strike interval that is relevant to the step symmetry and continuity measures of the Tinetti gait assessment tool [10], and the spatial average of peak plantar pressure that is of interest in monitoring diabetic patients for plantar ulcers. Signal processing for functional gait metric computations based on the data collected by HERMES either occurs at the base station (smart phone), or may be performed offline.

Depending on the needs of the data consumers, some metrics are more relevant than others. For example, while the values of the metrics listed above are crucial to gait assessment, the precise values of other metrics may be less important while they still contribute to a diagnosis. These can include the extent of flat footedness, or an estimate of the subject's weight. We distinguish between metrics along these lines and will constrain the problem to require the former type of metrics (*primary* metrics) to be precisely recoverable from the watermarked signals, but the latter type (*secondary* metrics) to be recovered within a user-specified degree of change.

B. Threat Model

In the three security use-cases that we address, namely proof of ownership, data tracking and content authentication, our goal is to thwart the effort of a malicious data consumer attempting to corrupt or erase the embedded watermark and possibly replace it with his own. By successfully achieving this in the proof of ownership and data tracking use-cases, the malicious data consumer gains the ability to illegally share data that is not owned by him and cover his tracks while doing so. In the case of content authentication, the malicious consumer can alter the data without leaving any evidence of the transformation. For example, this can be used by a disingenuous healthcare provider looking to avoid liability lawsuits, or by a duplicitous insurance provider looking to deny legitimate claims. Therefore it is important to validate the robustness claim of our watermarking procedure.

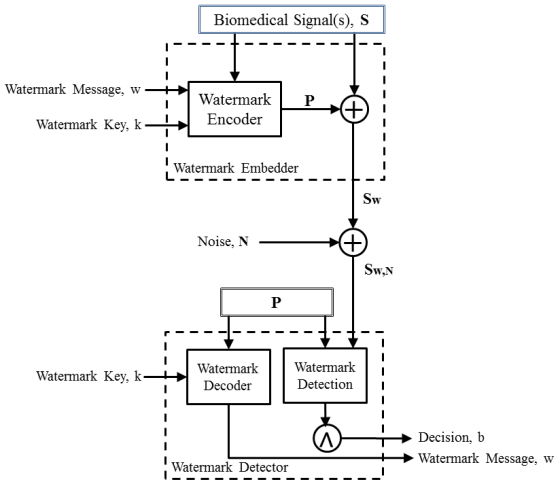


Figure 2: Watermarking System.

Towards this end, we consider three types of threats. First, we consider a malicious data consumer that adds noise to the watermarked signal in an attempt to corrupt the watermark and decrease the likelihood of successfully detecting it. This will hinder the ability of a data owner to prove ownership or track shared data. Seconds, we consider a malicious consumer who attempts to gain access to the watermark encoding, either by random guesses or by reverse engineering the watermarking process. Thereafter, he may remove the watermark and arrive at the original signal, which he is free to share in an untraceable manner. Third, we consider an attack wherein the malicious consumer attempts to prove the false claim that he is the source of the watermarked data, by showing that his watermark, and not the data owner's, is embedded in the watermarked signal. We refer to this as watermark spoofing.

IV. Robust Watermarking of Biomedical Signals

For clarifty, we first describe our watermarking approach in the absence of constraints from the need to provide fault tolerance. The solution is then extended in a subsequent section to address the problem of fault tolerance refine the watermarking solution to support it.

Fig. 2 outlines the watermarking system we will use to address the proof of ownership, data tracking and content authentication requirements for biomedical signal data collected by BASNs. The system consists of a watermark embedder used by the data owner to imperceptibly and robustly embed a watermark, and a watermark detector that is used by an independent authority to test for the existence of the watermark in biomedical signal datasets,

as required. The embedder uses an encoding algorithm to generate an encoding vector, P, corresponding to the watermark message, w, that is most suitable for the biomedical signal vector S. The encoding process is reliant on a secret watermark key k, known only to the true owner of the data. Suitability of the encoding is defined in terms of the robustness and imperceptibility afforded by it once it is embedded into the original signal. The embedder then embeds the encoding vector into the original signal vector, S, to generate an altered biomedical signal, S_w, which now includes the watermark.

As per the first threat in our threat model, over the course of sharing the watermarked signals, a malicious consumer may further alter the signal into $S_{w,N}$ by adding noise to it in an attempt to corrupt the watermark. Once this version of the signal needs to be validated by checking for the existence of the data owner's watermark, the watermark detector accepts the publicly watermarked signal ($S_{w,N}$), and the watermark encoding vector (P) and key (k) from the data owner. Following this, she (i) confirms that the encoding vector exists in the signal under test ($S_{w,N}$) via decision bit b, and (ii) decodes the encoding vector (P) with the help of the watermark key (k) into the data owner's watermark message (w). Here, the decision bit (b) indicates whether the encoded vector exists in the signal under test or not. Note that aside from the watermark key (k), the encoding vector P and the original signal S are also kept secret from the data consumers.

To enable proof of ownership, a watermark identifying the data owner is embedded into the medical signal. Proving ownership is as simple as proving that the data owner's identifier is embedded in the signal when it is tested. As long as the watermarking procedure produces robust watermarks, the watermark detector should indicate the watermark's presence. Data tracking is achieved with the data owner embedding distinct identifiers in each shared copy of her/his biomedical signal data. For example, data shared with one's healthcare provider is embedded with a different watermark message than the one shared with the insurance provider or with one's friends. If the data ends up in the possession of an unauthorized party, the data owner can iterate through the set of watermark messages she/he has used and identify the legitimate but malicious data consumer that released the patient's data without her/his authorization. Finally, content authentication is achieved by embedding the data with a watermark comprised of a cryptographic hash of the original signal. On successful detection of the watermark by the watermark detector, authentication of the data is achieved by removing the watermark from the watermarked signal and comparing it to a cryptographic hash of the signal that remains. The watermark and hash will match only if this is the original signal.

Before we delve into details regarding the construction of such a watermarking procedure, we highlight an important property of most biomedical signals which we leverage in our watermark encoding algorithm. Bio-medical signals are often segmentable in a straight-forward manner since they arise from human physiology. For example, ECG signals are based on heart-beats and have a distinct waveform (e.g. QRS complex in ECG) that is well-studied and the origins of which are well-understood. The same is true of EEG waveforms. In the case of bio-mechanical measurements, human movement and gait usually consists of periodic and repeated sequences of movements which can be used to segment spatio-temporal bio-signals comprised of accelerometer, gyroscope and pressure/force measurements at the joints and limbs. For example, the foot plantar pressure measurements of HERMES can be attributed to the well-studied stance and swing gait phases. Stance is the phase of human gait when the foot is in contact with the ground, while swing in the phase when the foot is in the air. Therefore, the pressure measurements at all sensors can be segmented based on these phases. Consequently, we denote sensor s_i's measurement at epoch j of the segmented signal as sample s_{ij}. Encoding a watermark message (w) into the samples can be equated to perturbing each sample s_{ij} by some amount p_{ij}:

$$s_{w_{ij}} = s_{ij} + p_{ij} \tag{1}$$

Note that signal vector, S, encoding vector P, and watermarked vector S_w are vectorized versions of these matrices. Now, there are three requirements for the watermarking procedure.

First, the watermark should be robustly embedded in the original signal. Recall that this means the watermark is detectable even if the watermarked signal is altered maliciously by the addition of noise. To satisfy this requirement, robust detection of the watermark will be achieved with a linear-correlation based detection procedure. In other words, detection of the watermark depends on how well correlated the watermarked signal and the watermark are. The detection metric z is calculated as:

$$z = S_t \cdot P \tag{2}$$

where S_t is the signal under test for the watermark encoded as P. The linear-correlation based detection technique evaluates the existence of the watermark in the tested signal, based on the value of the detection metric z. Specifically, if the metric value is greater than a threshold τ, positive detection occurs. Otherwise negative detection occurs. However, there is always a chance that the watermark detector will produce incorrect results. Either the detector will indicate the presence of the watermark when it does not exist, which is called a false positive. False positives can occur when the original signal happens to be well correlated with the encoded watermark by chance, driving the detection metric beyond the threshold. Alternatively, the detector may fail

to detect the watermark even though the signal includes the watermark, which is known as a false negative. This occurs when the original signal exhibits a strong negative correlation with the encoded watermark by chance, and reduces the detection metric below the threshold, thereby eluding detection. Therefore, the false positive and false negative rates play and important role in achieving a robust watermarking scheme and must be taken into consideration by the encoding algorithm.

The **second** requirement for a watermarking procedure is that the watermark must be successfully embedded in the original signal. If it cannot be embedded, the procedure must indicate this. Further, the larger the watermark that can be successfully embedded, the stronger the procedure's ability to address the use cases. For example, a larger watermark in the content authentication use-case can support a larger cryptographic hash of the original signal making for a much stronger guarantee regarding malicious alterations to the biosignal data.

Third, the embedding should only result in imperceptible alterations of the original signal. Imperceptible embedding in the context of biomedical signals translates to no loss in the diagnostic value of the signal data which we call *semantic imperceptibility*. This requirement is met when it is ensured that the watermarking procedure does not alter any values of the primary diagnostic metrics. Imperceptibe embedding also translates to a restriction on the changes to the signal waveform, which we shall refer to as *signal waveform imperceptibility*. This helps ensure controlled modifications to the secondary metrics that the data will provide.

A. Assessing the Watermark Corruption Threat

Equations (3a) - (3c) outlines why the value of the detection metric is most likely to be high only when the signal under test embeds the watermark (i.e. $S_t = S_w$). An arbitrary original signal S is likely to correlate less and less with P as the length of these vectors increases. For longer vectors, therefore, $(S \cdot P)$ is very likely to have a small value. It can be shown, for example, that even if two vectors are conservatively categorized as well correlated if the angle between them is less than or equal to $45°$, the probability that two vectors are well correlated is proportional to $\frac{1}{n^2}$ where n is the length of the vector.

Now, since P is fully correlated with itself, the value of $(P \cdot P)$ is large. Hence, even though $(S \cdot P)$ is small in value, the contribution of $(P \cdot P)$ to the detection metric boosts its value up and beyond the threshold τ. Similarly, if S_t does not embed the watermark message w that has been encoded to P (i.e. $S_t \neq S_w$), then z is highly likely to evaluate to a low value since S_t and P will exhibit low correlation.

$$z = S_w \cdot P \tag{3a}$$
$$= (S + P) \cdot P \tag{3b}$$
$$= (S \cdot P) + (P \cdot P) \tag{3c}$$

This procedure also yields robust watermarks for the reasons illustrated in (4a)-(4c). Here, the tested signal S_t is a maliciously altered version of the watermarked signal with the goal of erasing the watermark. Assuming the malicious alteration is in the form of additive gaussian noise, the watermark should still be detectable. The reason is similar to our discussion above - a random noise vector generated as additive gaussian noise is unlikely to correlate well with the encoded watermark P. Therefore, although the addition of such noise might change the detection metric z, the probability of a false negative is low.

$$z = S_{w,N} \cdot P \tag{4a}$$
$$= (S + P + N) \cdot P \tag{4b}$$
$$= (S \cdot P) + (P \cdot P) + (N \cdot P) \tag{4c}$$

B. Assessing the Watermark Reverse Engineering Threat

In the reverse engineering threat, we assume that the attacker may have knowledge regarding the watermark message (w), but he does not know the encoding vector P. He also does not have access to the original bio-signals (S) or to the watermark key (k). Of course, he is assumed to have access to the watermarked signal (S_w). We analyze two approaches an attacker may take to reverse engineer the watermarked signal to arrive at the original bio-signal, S. One, the attacker could randomly generate an encoding vector P' in an attempt to arrive at S (since $S = S_w - P$). Here, we will further assume that the attacker will be satisfied with an estimate of S' of the original bio-signal that is close enough to it but not necessarily exactly equal to it. In other words, the attacker succeeds if $S' - S$ lies within an ϵ-ball in n-dimensional space, where n is the length of the signal and encoding vectors. Under this assumption, it can be shown that the probability of the attacker's success is equal to $\left(\frac{\epsilon}{p_{max}}\right)^n$, where p_{max} is an upper bound on the perturbations p_{ij} composing the encoding vector P.

Another approach an attacker may take to reverse engineering the original bio-signal from the watermarked signal, is to programmatically find the encoding vector (P) subject to the constraints in (5a)-(5b). Given that P is an unknown, the constraint in (5a) is non-linear. Furthermore, it is non-convex making the search for the encoding vector sensitive to initial conditions, and therefore unable to guarantee that it will yield an accurate solution. Consequently, the linear-correlation based detector makes it very difficult for an attacker to successfully reverse engineer the encoding vector P.

$$(S_w - P) \cdot P \leq \tau \tag{5a}$$
$$S_w \cdot P > \tau \tag{5b}$$

C. Assessing the Watermark Spoofing Threat

To successfully spoof the watermark, an attacker must show that a message w' (other than w) of his choosing yields a encoding vector P', which when embedded into a spoofed original bio-signal vector S' (other that S) yields the watermarked signal S_w. A necessary condition for the success of this attack is that the spoofed original bio-signal S' must not show strong correlation with the true data owner's encoding vector P. Given that strong correlation between encoding vectors and original (unwatermarked) signals is an unlikely event, the attacker cannot claim to be the data owner if his "original" bio-signal contains the true data owner's encoding vector. Again, it can be shown that the probability that S' does not to contain P rapidly diminishes as n, the length of the vectors, increases.

V. Watermark Encoding Algorithm

We now describe a linear-programming approach for watermark encoding that addresses the three requirements for a biomedical signal watermarking scheme described in section IV.

A. Constraint: Embedding the Watermark

To address the successful embedding requirement, it is necessary that the watermark message w be correctly encoded by the vector P. This is achieved by the constraint in (6a)-(6b). The watermark message is encoded as a set of perturbations over all sensors and for all epochs of the segmented signal. The constraint in (6a)-(6b) requires that the sign of the perturbation at position m specified by the watermark key (k) be constrained by the bit m of the binary encoding for the watermark message (w). In other words, k acts as a seed to a random number generator (pos) that generates (sensor, epoch) pairs indicating the position of perturbation for each bit in the watermark message. The perturbation at a (sensor, epoch) position must be positive if the corresponding watermark message bit is 1 and negative otherwise. Note that there may be several positions not generated by pos where the perturbations are free to attain positive or negative values subject to the other constraints in the Linear Program (LP). Also note that an attacker will not be privy to the value of k, as it is kept secret by the data owner.

$$p_{pos_{k,m}} > 0 \quad if \quad w_m = 1 \quad \forall m \in [1, log(w)] \tag{6a}$$
$$p_{pos_{k,m}} < 0 \quad if \quad w_m = 0 \quad \forall m \in [1, log(w)] \tag{6b}$$

Aside from computational efficiency, the advantage of specifying the solution as a linear program is that a failure to embed is immediately identified as a failure to meet all constraints.

B. Constraint: Signal Waveform Imperceptibility

We address the requirement for imperceptibility of the changes to the original signal waveform in two ways. First, each perturbation p_{ij} should be within a small percentage of the original sample s_{ij} (see (7)). In other words, the perturbations are bounded by $[100\times(1-\alpha), 100\times(1+\alpha)]\%$ of the original samples.

$$-\alpha s_{ij} \leq p_{ij} \leq \alpha s_{ij} \qquad \forall i, j \tag{7}$$

The second signal waveform imperceptibility constraint applies to signal smoothness after perturbation. Here, a bound is applied to the first derivative of the altered signal (see (8a)). Equations (8b)-(8c) linearize this constraint. If necessary, higher-order smoothness constraints may be applied in a similar manner.

$$|p_{ij-1} - p_{ij}| \leq \Delta \qquad \forall i, j > 1 \tag{8a}$$
$$p_{ij-1} - p_{ij} \leq \Delta \qquad \forall i, j > 1 \tag{8b}$$
$$p_{ij} - p_{ij-1} \leq \Delta \qquad \forall i, j > 1 \tag{8c}$$

Together, these two sets of constraints restrict alterations to the bio-signal waveform by the watermarking process. Restrictions on higher order moments of the signal may be similarly linearized and addressed, if they are deemed relevant to the secondary metrics derived from the data. This approach of addressing secondary metrics by constraining the shape of the watermarked signal rather than the quality of specific secondary metrics enables flexibility in the selection of secondary metrics that are derived from the datasets as they are shared, rather than limiting the choice of secondary metrics available to data consumers.

C. Constraint: Semantic Imperceptibility

The semantics of a bio-signal are strongly tied to the type of bio-signal that is being measured, and the manner in which it is used to infer the status of the patient's health. Therefore, semantic imperceptibility can only be defined in this context. We address semantic imperceptibility in the context of the maximum amplitude and foot contact signal features of plantar pressure signals measured by HERMES. The maximum amplitude feature is computed as the maximum pressure observed at each sensor over the stance phase. The maximum pressure diagnostic metric is then derived as the mean maximum pressure over the set of all sensors. Similarly, the staggering diagnostic metric is derived as the average difference between the maximum amplitude features of lateral areas of the foot sole. To prevent loss of semantic accuracy for these metrics, the maximum amplitude signal feature is computed from the watermarked signal (see (9a)), and the maximum amplitude

and staggering diagnostic metrics derived from these signal features are constrained to the corresponding metric values derived from the original signals (see (9b)-(9c)). Here, A and B correspond to the maximum amplitude and staggering diagnostic metrics based on the original signals, and a_i and b_i are the corresponding coefficients for each sensor i.

$$\max_j(s_{ij} + p_{ij}) = M_i \qquad \forall i \tag{9a}$$

$$\sum_i (a_i M_i) = A \tag{9b}$$

$$\sum_i (b_i M_i) = B \tag{9c}$$

Note that (9a) is non-linear and must be linearized before input to an LP solver (see (10)). However, (10) does not provide the true maximum (M_i) of the perturbed samples. To arrive at the true maximum, m_i must be minimized, and this will be addressed by the objective function, which we will discuss in the following subsection.

$$s_{ij} + p_{ij} \leq m_i \qquad \forall i, j \tag{10}$$

The foot contact signal feature is derived as the sampling epoch for the segmented signal at which foot contact is sensed at a sensor. Note that foot contact at different sensors will be different. For example, contact at the heel occurs much before contact at the toes. Contact is sensed at a sensor when its pressure measurements exceed some threshold tr. To guarantee that the foot contact based diagnostic metrics, namely heel strike angle and inter-footstrike interval, are left unaltered after the watermark has been embedded, we ensure that the foot contact signal feature is itself left unaltered by the watermarking process. To do so at each sensor, the perturbed signal at all epochs prior to the true foot contact epoch is constrained to be less than the threshold tr (see (11a)), while the perturbed signal at the true foot contact epoch is constrained to be greater than or equal to tr (see (11b)). Here, the true foot contact epoch derived from the original signal at sensor i is denoted t^i_{fc}. In this way, the foot contact signal feature derived from the watermarked signal at each sensor will be identical to that derived from the original signal, as will the related diagnostic metrics.

$$s_{ij} + p_{ij} < tr \qquad \forall i, j < t^i_{fc} \tag{11a}$$

$$s_{it^i_{fc}} + p_{it^i_{fc}} \geq tr \qquad \forall i \tag{11b}$$

D. Objective Function: Maximizing Robustness

Finally, the robustness requirement is addressed by the objective function. Recall from our discussion in section IV that minimizing false positives and false negatives translates to ensuring that the linear-correlation based detection metric (z) is greater than tau for a watermarked signal and less that tau for an un-watermarked signal. To minimize the chance that a noise vector (n) introduced into the watermarked signal decreases its z value below τ, we must maximize the distance or difference between $(S \cdot P)$ and $(S_w \cdot P)$. In other words, we must maximize $(P \cdot P)$. In LP form, this is similar to saying that we must maximize $\sum_{i,j} |p_{i,j}|$.

Although thus far, we have discussed the generation of the encoding vector P for one signal segment as if it were independent of the encoding vectors for all other signal segments, tau imposes a unified constraint across signal segments. In other words, $(S \cdot P)$ and $(S_w \cdot P)$ must be separated by τ regardless of the segment under consideration. To retain the ability to solve the LP individually for each segment, we additionally force $(S \cdot P)$ to be as close to some constant X as possible (see (12a)-(12b)). In doing so, $(S \cdot P)$ is tied as close to X as possible, while the value of $\sum_{i,j} |p_{i,j}|$ is simultaneously maximized. This drives $(S_w \cdot P)$ as far away from X as possible.

$$\sum_{i,j} s_{ij} p_{ij} \geq X \tag{12a}$$

$$minimize: \quad \sum_{i,j} s_{ij} p_{ij} - \sum_{i,j} |p_{i,j}| + \sum_i m_i \tag{12b}$$

The best value for X can be arrived at by a line search within bounds generated based on the values of S and α, while seeking out the value of X that yield the lowest embedding failure rate for training set of bio-signal segments. Finally, note that the objective function in (12b) not only aims to limit false positive and false negative rates, but also addresses the maximum amplitude related constraint from the previous subsection.

VI. Watermarking Support for Fault Tolerance

Having addressed the problems of encoding and embedding the watermark across the entire spatio-temporal signal, we now turn to the problem of catering to fault tolerance requirements. Fault tolerance is addressed by a sensor partitioning scheme whose goal is to ensure that the quality of primary and secondary metrics derived from a single partition is maximized. This way, when faults are detected in a sensor belonging to one partition, metric values from the other partition(s) can be relied upon as surrogates. Partitioning schemes are also useful in detecting faults, however this topic lies outside the scope of this paper. While a partitioning scheme allieviates the problem of semantic corruption of datasets by faulty measurements, it introduces a tradeoff by limiting the maximum number of bits that may be successfully embedded (*embedding capacity*). In other words, the more fault tolerant an architecture is, the less secure it can be owing to the limit on the size of the message that is embeddable. However, we posit that embedding independent messages in each partition serves to bolster security and limit this tradeoff. While a smaller watermark

may be easier to attack, there are now as many smaller watermarks as there are partitions. Therefore, the attacker must succeed in attacking the watermark in a majority of the partitions to truly succeed in his attack.

The stated goal of the partitioning scheme is achieved with an objective function that maximize the similarity between the partitioned subset of sensors. We translate this to a graph partitioning problem with the objective of edge cut minimization. Here, the weights of the edges between nodes (representing sensors) is a dis-similarity index between the measurements of the the respective sensors. We chose to use a the inverse of the absolute value of the linear correlation co-efficient, as datasets collected by HERMES indicate strong linear local correlation among its sensors, although other relevant similarity measure would work well too. Therefore, the more correlated two sensors are, the lower their dissimilarity score, and the more likely they are to be assigned to different partitions. However, this graph partitioning is NP-hard. Nevertheless, we address this with the HMETIS heuristic algorithm [11] that has been shown to perform extremely well on a variety of graph families.

VII. Validation

We validate the ability of our linear programming based approach to robustly watermark biomedical signals by applying the technique to three plantar pressure datasets collected with HERMES. The datasets correspond to plantar pressure measurements of male and female subjects with different weights and gaits, collected over a few hundred footsteps as the subjects walked. The LP is applied independently to the plantar pressure signals at each footstep of each subject. Note that the number of embeddable positions (see (6a)-(6b)) is limited by the number of sensors and epochs per foot steps. Generally, at a walking rate of 1Hz, a vast majority of steps last fewer than 50 epochs. Further, most sensors experience plantar pressure for a minority of these epochs, experiencing 0 pressure at most epochs during which no perturbation is possible (see (7)). Consequently, each of the datasets expose between 1500 and 2400 embeddable positions for each signal segmented or foot step. Out first set of validations pertain to a single partition and we present results for multiple partitions at the end of this section.

Fig. 3 shows the watermarked signals from a single sensor over a single footstep from one of the plantar pressure datasets. Here, w is 200 bits long, α is 10%, Δ is $\delta \times max(s_{ij-1}, s_{ij})$, and δ is 0.01. In other words, Δ is not the same for all bio-signal samples, but changes depending on the sample value that is to be perturbed. We will continue with this definition of Δ for the rest of the paper, varying δ when we study the impact of the smoothness constraint on performance.

Several observations can be made regarding the watermarking procedure from Fig. 3. The signal waveforms are very similar before and after watermarking, with the signals

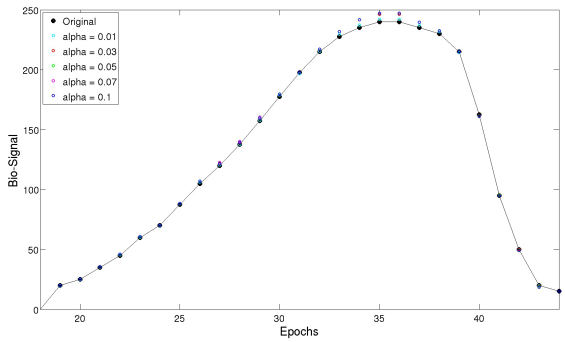


Figure 3: Original vs. Watermarked plantar pressure measurements for 1 sensor over the stance phase of a single footstep from dataset 1. Watermarked signal presented for different values of α ($\delta = 0.01$)

appearing nearly identical for small values of α. The larger α is, the more different the watermarked signal gets, however we will see that this is traded-off for improved robustness. Fig. 3 also shows that the semantic imperceptibility requirement for the foot contact signal feature dictates minimal changes to the waveforms at the initial epochs prior to foot contact. In contrast, the LP solver achieves imperceptibility of the maximum amplitude signal feature in a different way. The maximum amplitude of the watermarked signal for some sensors are greater than that of the original signal (Fig. 3), while the maximum amplitude of the watermarked signal of other sensors are less. However, these difference are balanced out to achieve the overall effect of identical maximum amplitude-related diagnostic metrics before and after watermarking.

Next, we study the robustness of our watermarking procedure in the context of watermarking all signal segments of each dataset, while varying the embedding capacity. While testing with a given embedding capacity, the embedding procedure was repeated 10 times for each foot step of each dataset, as the watermark message w, and therefore its binary encoding, was randomly varied across the repetitions. We set α to 0.05, δ to 0.025 and X to 0. Fig. 4 shows boxplots of the detection metric values before watermarking (i.e. $z = S \cdot P$, in green) and after watermarking (i.e. $z = S_w \cdot P$, in black). Each boxplot provides summary statistics for z over all signal segments of a particular dataset and for a single watermark length. In keeping with (12a)-(12b), the objective function ties $S \cdot P$ to 0 while $P.P$ increases with the embedding capacity. Therefore, selecting a τ value close to 0 provides a whole lot of room ($\cong P.P$) for robustness to noise contamination. The generally increasing trend in the separation of the detection metrics before and after watermarking, are a result of more embedding positions afforded by an increasing capacity.

On the other hand, this increase tapers off at the upper end of the tested range. On a related note, we present, in Fig. 5, the rate of embedding success (*embedding effectiveness*)

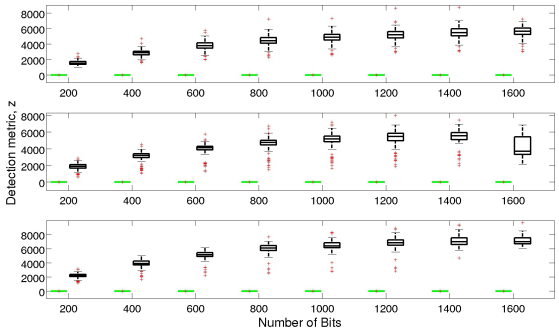


Figure 4: Boxplots of detection metric (z) before (green) and after (black) embedding encoded watermark, as the size of the watermark message is varied. Each panel represents one dataset, ordered from top to bottom.

as the embedding capacity is varied. Note that at 1000 bits, we are at between 42% and 67% of the maximum capacity available (1500-2400 bits). Beyond this capacity, the percent of segments allowing successful embedding begins to fall well below 100%. In other words, between 40%-50% of the maximum capacity can be comfortably supported before embedding effectiveness begins to degrade. The reason for the degradation is that in attempting to embed as many bits, the imperceptibility constraints (particularly, the smoothness and semantic imperceptibility constraints) cannot be easily satisfied. For the same reason, as an attempt is made to embed a large number of bits, the magnitude of the perturbations decrease as the LP solver struggles to meet the smoothness and semantic imperceptibility constraints. This produces the plateau in the separation of the detection metric pre- and post-watermarking. Finally, a line search for the value of X that produces the highest embedding effectiveness showed that substantially poorer results can be achieved at certain values of $X > 0$. However, the best results weren't significantly better that the success rates for $X = 0$ in Fig. 5, particularly at higher embedding capacities. In the set of experiments discussed thus far, the watermark key k was fixed. However, varying the embedding positions by varying the watermark key did not impact the

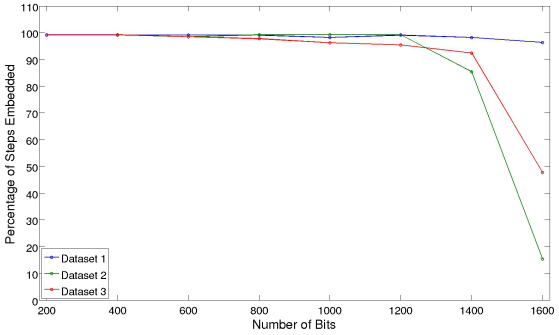


Figure 5: Embedding Effectiveness for each dataset as the size of the watermark message is varied.

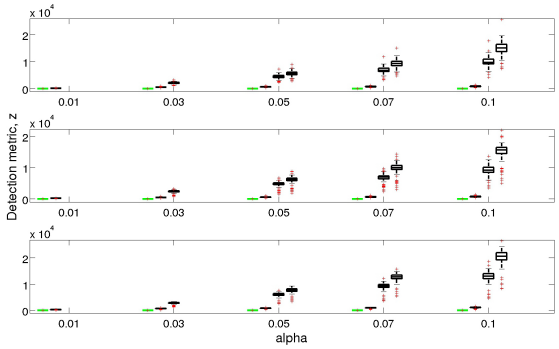
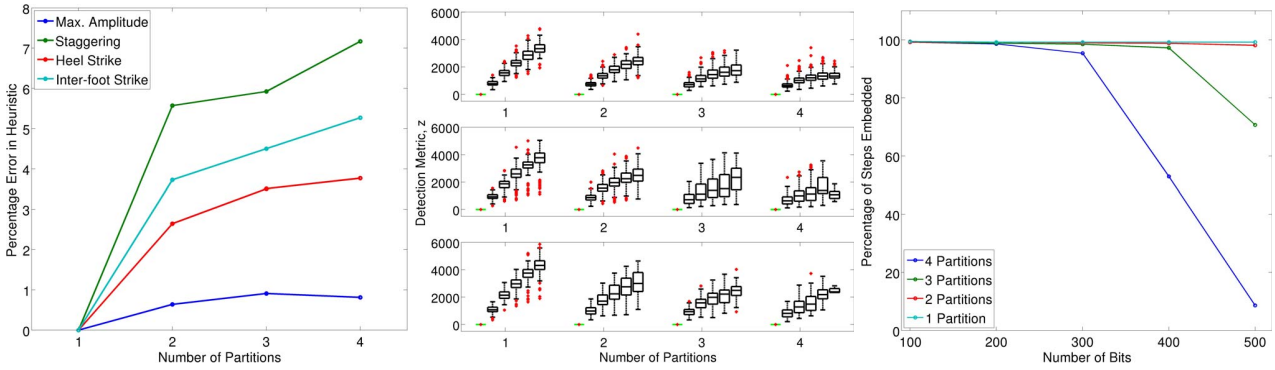


Figure 6: Robustness results for each dataset as the values of α and δ are varied. δ was set to 0.005, 0.025 or 0.04 (left to right, for each value of α). Each panel represents one dataset, ordered from top to bottom.

watermarking robustness or embedding effectiveness at low and high embedding capacities (data not shown for brevity).

We also studied the sensitivity of watermarking robustness and embedding efficacy to the signal waveform imperceptibility parameters α and δ. Fig. 6 presents summary statistics for the detection metric z pre- and post-watermarking for each of the 3 datasets, as α and δ are varied. Each grouping of black box plots for a single dataset and a single value of α represents values of δ set to 0.005, 0.025 or 0.04 (left to right). Parameter pairs where $\alpha < \delta$ are invalid and are omitted from study (e.g., $\alpha = 0.01$ and $\delta = 0.4$). The embedding capacity was set to 800 bits and X to 0. We observed that the separation of z pre- and post watermarking, and hence the robustness, increases with α and with δ. The less stringent the imperceptibility constraints were, the larger the amplitude of P was able to get. Consequently, the watermarking process realized more robust watermarks. We also observed that the watermarking robustness is about as sensitive to changes in δ as it is to changes in α. The reason δ plays such an important role in watermarking robustness, is that even if the magnitude of perturbations are allowed to take on larger values, strong limits on the extent to which consecutive perturbations can differ (δ) often prevent the maximum perturbation magnitudes from being achieved. Again, these set of experimental runs revealed little-to-no impact of α and δ on embedding effectiveness.

Finally, we studied the robustness and embedding effectiveness of our watermarking approach under the constraints of the fault tolerance imposed partitioning scheme. Fig. 7 first presents the impact of the partitioning scheme on the accuracy of the primary semantic metrics considered here. Cross-validation was performed by applying 50% of the data towards the partitioning scheme and towards training linear regression models to estimate aggregate metrics from measurements of individual partitions. The results, averaged across the 3 datasets, show a general degradation of accuracy as the number of partitions are increased. However, with



(a) (b) (c)

Figure 7: Performance results under fault tolerance constraints. The plot on the left shows average error performance for each of the primary metrics as the number of partitions in increased. The plot in the middle (right) shows robustness (embedding effectiveness) results as the number of partitions and the embedding capacity was varied. Boxplots within each group of the robustness results represent increasing embedding capacities from 100 to 500 bits in steps of 100 bits.

up to 4 partitions, the maximum error observed across all metrics was 7%, a testament to the strength of the HMETIS graph partitioning algorithm. Holding α at 10% and δ at 0.01, we measured robustness and embedding effectiveness while varying the number of partitions and the embedding capacity. While the robustness improved with the embedding capacity, this effect reduced as the number of partitions increased. The reason for this becomes clear from the embedding effectiveness results. As expected, increasing the number of partitions constrains the maximum number of bits that can be successfully embedded. Consequently, it becomes more and more difficult for the encoding algorithm to achieve strong robustness.

VIII. Conclusion

We have addressed security concerns of sharing BASN collected biosignal measurements with and between medical data consumers, under constraints imposed by fault tolerance requirements. We use a linear-correlation based detector that is shown to be successful in thwarting watermark corruption, erasure and spoofing threats. We also propose an LP formulation to encode watermarks in a robust manner while adhering to semantic imperceptibility constraints. Further, A BASN partitioning algorithm is applied to satisfy the fault tolerance requirements while minimizing its impact on embedding effectiveness. Validation studies indicate that up to a 1000 bits can be effectively and robustly embedded signals sampled by a gait stability monitoring BASN.

References

[1] P. Pawar, V. Jones, B.-J. F. Van Beijnum, and H. Hermens, "A framework for the comparison of mobile patient monitoring systems," *J. of Biomedical Informatics*, vol. 45, no. 3, pp. 544–556, Jun. 2012.

[2] I. Cox, M. Miller, J. Bloom, J. Fridrich, and T. Kalker, *Digital Watermarking and Steganography*, 2nd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2008.

[3] M. Li, W. Lou, and K. Ren, "Data security and privacy in wireless body area networks," *Wireless Communications, IEEE*, vol. 17, no. 1, pp. 51–58, 2010.

[4] M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, "Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 1, pp. 131–143, 2013.

[5] X. Kong and R. Feng, "Watermarking medical signals for telemedicine," *IEEE Transactions on Information Technology in Biomedicine*, vol. 5, no. 3, pp. 195–201, 2001.

[6] V. Goudar and M. Potkonjak, "Fault-tolerant and low-power sampling schedules for localized basns," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 3, no. 1, pp. 86–95, March 2013.

[7] H. Noshadi, S. Ahmadian, H. Hagopian, J. Woodbridge, F. Dabiri, N. Amini, M. Sarrafzadeh, and N. Terrafranca, "HERMES - mobile balance and instability assessment system," in *Proc. BIOSIGNALS*, 2010, pp. 264–270.

[8] V. Goudar, Z. Ren, P. Brochu, Q. Pei, and M. Potkonjak, "Optimizing the configuration and control of a novel human-powered energy harvesting system," in *23rd International Workshop on Power and Timing Modeling, Optimization and Simulation*, Sept 2013, pp. 75–82.

[9] V. Goudar and M. Potkonjak, "Addressing biosignal data sharing security issues with robust watermarking," in *IEEE International Conference on Sensing, Communication and Networking*, 2014.

[10] A. Wendy and M. Moffa-Trotter, "Functional tools for assessing balance and gait impairments," *Topics in Geriatric Rehabilitation*, vol. 15, no. 1, pp. 66–83, 1999.

[11] G. Karypis and V. Kumar, "A fast and high quality multilevel scheme for partitioning irregular graphs," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 359–392, 1998.