Security Defenses for Vulnerable Medical Sensor Network

Renchi Yan, Vaibhav Chetan Shah, Teng Xu and Miodrag Potkonjak
Computer Science Department
University of California, Los Angeles
{renchi.yan, vaibhav289, xuteng, miodrag}@cs.ucla.edu

***Abstract*—Medical sensor networks have facilitated a wide range of applications in healthcare. However, these systems are in particular vulnerable to security attacks due to the fact that they are often not physically secured and are used in potentially hostile environments. We have proposed a theoretical and statistical framework for creating attacks and also the corresponding security defenses that include attack detection, diagnosis, and impact removal. We use medical shoes to collect data and demonstrate that low energy and low cost of medical sensor networks increase the probabilities of successful attacks. Our approach maps a semantic attack to an instance of an optimization problem where medical damage is maximized under the constraints of the probability of detection and root cause tracing which may consequence in incorrect medical diagnosis and treatment. Our results show that it is easy to attack several essential medical metrics and to alter corresponding medical diagnosis. Finally, we have developed several low energy and low overhead defense procedures for detecting and analyzing semantic security attacks.**

I. Introduction

Healthcare applications are increasingly using Wireless sensor networks(WSN) to extract important health information from them. Through its remote sensing ability, WSN in healthcare has helped in creating numerous wireless health applications and has also extended the scope of medical diagnosis. For example, medical experts can examine the day-to-day activities of a patient or analyze the patients' physiological data collected from sensors on a regular basis. This also eliminates the need to solely rely on in-person medical check-ups while improving the quality of care.

Embedded medical sensor networks are becoming popular as they are low-cost solutions to a variety of medical challenges. Having accurate information is critical in dealing with these medical sensor networks. For example the medical expert can carry out the right treatment only if he/she has accurate information about the condition of the patient. Inaccurate medical readings can produce irreversible damage which can even lead to the death of a patient.

However, embedded medical sensor network enforces strict constraints on energy and power consumption. Both of these act as major obstacles in the application of traditional security techniques to medical sensors. Security is extremely critical in medical devices as even a small error due to tampering can cause wrong diagnosis and can potentially endanger the life of a patient. Moreover, in making medical devices energy-efficient by reducing the number of sensors used; makes them more susceptible to semantic attacks. Traditional cryptography can only guarantee the data to be safe through the wireless channels. However, it cannot prevent malicious manufacturers because of their special privileges in manufacturing. Cryptographic techniques are not suitable for semantic attacks as the medical devices are sometimes unattended which can lead to their tampering and subsequent malfunctioning. Our focus is that the medical expert should be able to detect the attacks in real time even though the data has been tampered.

Many papers have been written to address communication based security issues of wireless sensor networks in healthcare applications [1] [2]. However, we focus on semantic attacks. A semantic attack is the one in which the attacker modifies information in such a way that the result is incorrect, but looks correct to the casual or perhaps even the attentive viewer. We demonstrate that by implementing simple attacks; the attacker can hamper the result drastically. In this experiment we prove that by attacking several essential medical metrics, the diagnosis can be changed considerably yet generating acceptable results. We also propose defenses against these semantic security attacks that can detect the respective attacks. We evaluate both attacks and defenses under two different scenarios with respect to the number of sensors used to generate the results for comparing and analyzing our techniques.

Our proposed semantic attacks are possible in real scenarios. For example, if a malicious attacker is able to break into the computer of a medical expert or a malicious party can even be the manufacturer of medical shoes who can create back doors into the embedded sensors such that he/she can easily access and tamper the sensor data. However, from the perspective of attacker, he/she cannot be too aggressive as too much deviation from the original data will easily make the medical expert suspicious about the data being tampered.

We evaluate our attacks and defenses on the Hermes shoe platform, which is designed to assess balance and instability in patients [3].It consists of 99 pressure sensors distributed in each insole and integrated with a common computing platform. The special features of a person's gait which are highly correlated to his/her risk of falling as shown by Maki [4] are used for the attack and defense evaluation.

This paper presents the following research contributions: 1) One of the first to analyze the impact of semantic attacks on the security of wireless medical devices; 2) Propose two

novel semantic attacks that can impact the result in a great manner and also corresponding defenses to identify the same; 3) Detailed analysis showing that using more sensors are potentially more resilient to semantic attacks.

The remainder of this paper is organized as follows: In Section II and III we present the related work and the preliminary knowledge required for the experiment. Section IV gives a short overview of the metrics considered and our formulation of the result. Sections V and VI provide the implementation details of our proposed attacks and defenses on the specified metrics and their evaluation. Finally we conclude the paper with Section VII summarizing our findings and stating our conclusions.

II. Related Work

In the last decade, wireless medical devices and corresponding techniques have attracted a great deal of research and development interest. More recently a significant emphasis has been made on security issues. For example several research groups at different universities including University of Michigan, University of Massachusetts Amherst, Massachusetts Institute of Technology, Rice and Princeton have reported techniques that enabled security compromise of so popular and important devices such as pacemaker [5], implantable cardiac defibrillator[6] and insulin pumps [7] [8].

In addition, actual medical community has been rapidly becoming more aware of power and energy limitations of security techniques used for medical devices [9] [10] . In addition to security, other issues such as privacy [11] and trust [12] [13] have received significant attention. As a matter of fact many other aspects related to security in medical applications such as systems that integrate wireless devices and cloud computing have been addressed [14] [15]. We conclude out brief survey in wireless medical devices by pointing to two comprehensive reviews in this research field [16] [17]. Meanwhile, hardware based technology has been proposed to secure the sensor network [18][19]. Protocols to protect the integrity of sensor data are proposed in [20].

While all previous efforts in this field emphasized vulnerabilities of used wireless security protocols and their potential fixes, we focus on actual alteration of collected sensor data in such a way that semantic conclusion of medical experts is altered. This alteration leads to incorrect treatments which might compromise the medical well being of a subject.

III. Preliminaries

A. Medical Shoe

Medical sensor networks are inherently semantics-driven systems. The medical expert is generally not concerned with the actual sensor readings but rather more concerned with the semantic information like the gait characteristics mentioned below. Hence we try to attack this semantic information so that maximum harm can be caused.

We evaluate our attacks and defenses on our medical shoe which consists of 99 sensors distributed about the sole of the foot, a processing unit, flash memory, a radio, and an ADC. The sensor placement is according to the Pedar plantar pressure mapping system [21] as shown in Figure 1. The numbering of the sensors according to their position is also depicted in Figure 1.

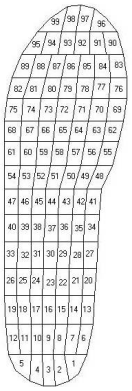


Fig. 1: The 99 sensor mapping of Hermes shoe platform.

B. Data Set

Our dataset includes pressure readings over all the 99 sensors for each shoe sampled at 50hz for four persons respectively using a 16-bit analog-to-digital converter. The resulting time-dependent pressure mappings are used to calculate the below mentioned gait characteristics from the full dataset. The dataset includes hundreds of steps of all the subjects which also incorporates data for each foot of a person separately.

The pressure readings are collected for seven different scenarios namely walk, jump, lean, run, stand, limp, slow-walk. In our paper we focus on the walk readings and calculate the impact of our attacks and defense mechanisms.

IV. Metrics and Formulation

Maki [4] has observed that stride-to-stride variability in speed has a strong correlation with the risk of falling. The spatial gait parameters related to this variability like stride period, double support, stride length, stride width and stride velocity help in predicting the danger of falling. We take into consideration the two important metrics namely stride period and double support to simulate our attacks and defenses respectively.

Stride-to-stride consistency is taken into account by computing the average of differences between two consecutive values of a specific feature. Moreover if the patient is not able to walk with consistency; they are normally at higher risk of falling. Hence higher the variation; higher is the instability. Thus in general, variation in a metric signifies increased instability.

In order to calculate the stride period with the help of our dataset, we generated the walk waveform of each subject's left and right foot separately. The left foot walk waveform of the first subject is shown in Figure 2. The peaks with the highest pressure in this waveform represent the moment of highest pressure and hence it symbolizes the contact of foot with the ground. We measure the difference between two successive peaks to determine the required stride period. Then variation is calculated by taking the average of all the absolute differences

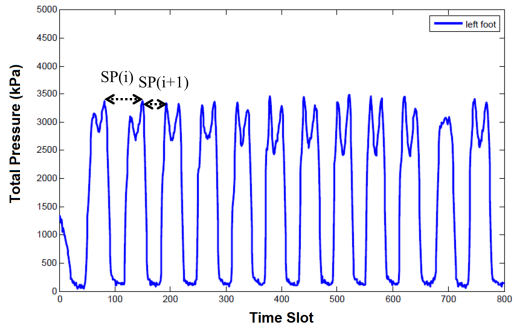


Fig. 2: The waveform of the stride period metric.

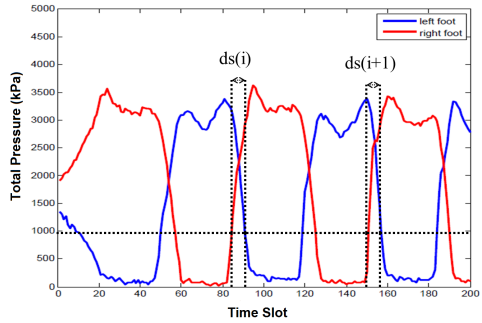


Fig. 3: The waveform of the double support metric.

of stride periods. *SP(i)* represents stride period at i^{th} step and *n* represents the total number of steps.

Similarly to calculate double support, we superimpose the walk waveforms of each person's left and right foot as shown in Figure 3. We measure the overlap time between peaks of the two waveforms at the same time slot giving us the duration for which both the feet are in contact with the ground simultaneously. Variation is determined by taking the average of these absolute differences. *ds(i)* represents the double support time at i^{th} step in both feet and *min(SP(i))* represents the minimum of left and right stride period at i^{th} step. The dotted horizontal line passing through the waveform in Figure 3 indicates the level above which we assume that both the feet of the person are on the ground simultaneously.

$$Var_{SP} = 1/n \sum_{i=1}^{n} |SP(i+1) - SP(i)| \quad (1)$$

$$Var_{DS} = 1/n \sum_{i=1}^{n} |\frac{ds(i+1)}{min(SP(i+1))} - \frac{ds(i)}{min(SP(i))}| \quad (2)$$

Thus instability can be calculated from these variations based on equation 1 and 2. The coefficients γ_{SP} and γ_{DS} indicate the significance of a particular metric. The coefficients can be adjusted by the medical specialists.

$$Instability = \gamma_{SP} Var_{SP} + \gamma_{DS} Var_{DS} \quad (3)$$

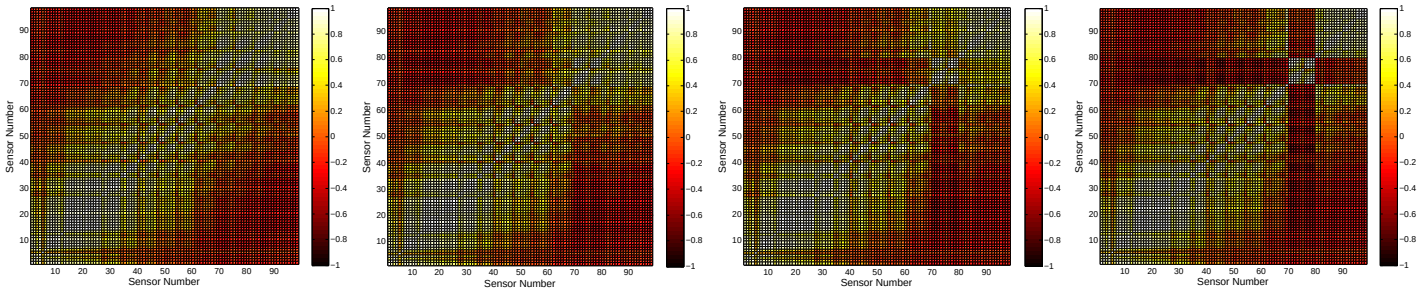
We attack these two important metrics under two different scenarios namely using the summation of 99 sensors and the summation of just twenty fixed sensors for medical diagnosis. Furthermore, we develop two defense procedures against these semantic attacks under the same scenarios.

V. Security Attacks

A. Goals and Challenges

Various types of attacks can be introduced in wireless sensor networks. The starting point in our creation of semantic attacks is to mislead the medical diagnosis as far as possible which can be done by altering the original Var_{SP} and Var_{DS} to some large extent. However, the main challenge for the attacker is to maintain a balance between the outcome of the attack and the risk of the attack being detected. For example, suppose in one attack, the attacker changes the pressure value of many sensors or changes some amount of sensors by a large amount of pressure then that attack will be easily detected if the medical expert looks into the statistical properties of the data and compares it with the historical data. As a result, the criteria in designing semantic attacks is to assume limited access to the sensors for the attacker and also of not being suspicious at the same time. In other words, the attacker must follow the constraints.

In principle, we propose two attacks on the pressure data. In the first type of attack, we assume that the attacker can change n sensors by k percent. In the second type of attack, the attacker postpones the pressure data of some sensors for certain time slots. And we claim that a type of attack is "good" when the attacker only changes a small number of sensors but changes the variation to a certain degree and yet remains unsuspicious. The intuition behind the calibration attack is that when some sensors with high pressure are being changed, the waveform of the pressure is influenced significantly and thus the variation is modified. This type of attack only requires the attacker to change the calibration of the sensor value which can be achieved at either software or hardware level and thus it is easy and feasible. However, if some sensor shows extraordinary high or low value, it can be easily detected by the medical expert. Hence we put constrains on the percentage upto which the pressure of each sensor can be altered. Our second type of attack is from the perspective of timing. The attacker intentionally introduces delay on some sensor readings so that part of the original waveform of pressure is shifted by some time slots. This type of attack is simple, low cost and the only need is to introduce delay to some sensors. But the main problem here is to decide the number of time slots the attacker should delay while still being unsuspicious to the medical experts. One of the simplest defense technique is to look into the sensor correlations and thus to get an idea of it we analyzed the correlation between each pair of the 99 sensors when ten sensors are postponed to respectively five, ten, and twenty time slots. The results are shown in Figure 4a, 4b, 4c, and 4d. It is obviously seen that when the number of postponed time slots increase, the correlation between the sensors changes dramatically which is easily detectable. Therefore, the attacker needs to seek for a balance between the number of time slots to



(a) (b) (c) (d)

Fig. 4: Correlation between each pair of the 99 sensors, tested on P1. (a) Original Correlation without attack. (b) Correlation when postponing 10 sensors by 5 time slots. (c) Correlation when postponing 10 sensors by 10 time slots. (d) Correlation when postponing 10 sensors by 20 time slots.

postpone and the effect of the attack. Similarly in calibration attack, the attacker needs to vary the sensor readings by an appropriate percentage to generate acceptable results.

B. Scenarios

We consider two scenarios which the medical expert uses to diagnose patients. In the first scenario, the medical expert uses the summation of all the 99 sensors to generate the waveform of each metric and then calculates the corresponding variation. However, when we look at the pressure data closely, many sensors have very low value across all the time slots. Most probably these sensors are ones in the middle of the shoe who have high potential of not having any influence on the diagnosis. Keeping the power and the cost of medical shoe in consideration, many designs [10] use much fewer number of sensors rather than all the 99 sensors. Therefore, in the second scenario, the medical expert only uses twenty important sensors to fetch the data and further makes the diagnosis based on that.

C. Attack Modeling

The modeling of attacks revolves around tuning and setting the parameters of the attacks as random attacks will be easily detectable. Two types of attacks and their modeling are proposed in this part.

1) Calibration Attack: In this attack, the attacker has the access to change n sensors by k percent. We convert the attack problem into an optimization problem. The following is the objective function and the constrains which work for both metrics. The goal of optimization is to change the original variation as much as possible under the following three constraints, the first one is to change only a limited number of sensors and the second one is to change the pressure values by only a limited percentage. The third one is that the number of steps of an individual patient cannot be changed beyond a certain percentage after an attack. As we explained, the first constraint is based on the assumption that the attacker has limited access to the pressure data he/she can change. While the second and the third constraints are applied to prevent the attacks from appearing suspicious to the medical expert.

$$\begin{aligned} &\text{Maximize } |Var - Var_0| \\ &\text{Subject to} \\ &\qquad K \leq k \\ &\qquad N \leq n \\ &\qquad |S - S_0|/S_0 \leq \sigma \\ &\text{where} \end{aligned} \tag{4}$$

- Var is the variation after attack.
- Var_0 is the variation before attack.
- K is the percentage of pressure variation.
- N is the number of sensors attacked.
- S is the number of steps after attack.
- S_0 is the number of steps before attack.
- n, k, and σ are constants.

2) Timing Synchronization Attack: In this type of attack, the attacker has the access to some sensors and causes delay between their pressure measurement and pressure reading. Therefore, some sensors would be postponed for some time slots before they are read. For example, when the attacker decides to postpone the pressure value of sensor 1 by three time slots, then every time when the medical expert receives the pressure of sensor 1 in time slot T, it is actually the pressure value in time slot $T-3$. The following is the objective function and the constrains.

$$\begin{aligned} &\text{Maximize } |Var - Var_0| \\ &\text{Subject to} \\ &\qquad T \leq t \\ &\qquad N \leq n \\ &\qquad |S - S_0|/S_0 \leq \sigma \\ &\text{where} \end{aligned} \tag{5}$$

- Var is the variation after attack.
- Var_0 is the variation before attack.
- T is the number of time slots to postpone.
- N is the number of sensors attacked.
- S is the number of steps after attack.
- S_0 is the number of steps before attack.
- n, t, and σ are constants.

3) Algorithm for Attacks: We use dynamic programming (DP.) to implement the optimization problem. The pseudocode is shown in Algorithm 1. Theoretically, it is possible to try all the combination of N sensors out of 99 sensors such that the change in pressure of each sensor is within K percentage. However, the exponential search space of this problem combined with the large number of sensor-samples makes it impossible to solve this problem in a reasonable amount of time, especially in the real-time wireless network scenario. In order to reduce the running time, we do the N-sensor selection step by step. As described in Algorithm 1, the first step is to iteratively choose each sensor to attack which is followed by varying its pressure in every time slot by K percent. Then the variation is calculated after attack and the sensor is put into the set of attacked sensors to create a new attack situation. Then we sort the situations according to the difference in variation before and after the attack. We take the top M best attack situations from all the possibilities and use those situations for next iteration. We repeat the above procedure for another $N-1$ steps, thus to choose N sensors which cause the maximum damage. In this way, we reduce the exponential time complexity to $O(|sensors|MN)$, where $|sensors|$ is the number of sensors in the system.

Algorithm 1 Dynamic Programming for Sensor Selection

Input: P - original sensor pressure at each time slot.
Input: K - percentage of the pressure change in attack.
Input: N - number of sensors to attack.
Input: σ - error rate of the number of steps.
Input: M - number of optimal values to preserve in previous DP. step.
vec is a vector that contains the attacked sensors set, their corresponding data, and the difference of variation.

```
vec.append(<attacksensor=∅, P, diff. = 0>)
for 1 <= i <= N do
  for all sensors s_i do
    for 0 <= t <= M do
      if s_i is not in vec[t].P then
        P'= Attack(vec[t].P, K, σ, s_i)
        vec.append(<vec[t].attacksensor+s_i, P'
                   ,abs(Var.(P)-Var.(P'))>)
      end if
    end for
  end for
  vec = SortByDifferenceTopM(vec)
end for
Output: vec
```

D. Experimental Results

Table I to Table VIII show the effect of calibration attack and timing synchronization attack in different scenarios on different metrics. We assume that the attack can happen on left foot or right foot for stride period metric, and left foot or both feet for double support metric. The readings in the tables indicate the percentage of pressure variation after attack compared to the original variation. We test each attack on both the metrics for the four persons P1, P2, P3, and P4 respectively. The average change in the percentage of new variation is then calculated across the 4 persons as shown in the last row of each table. In order not to change the correlation between the sensors drastically, we change the pressure values to a maximum of 20% in calibration attack and postpone the readings to a maximum of 7 time slots in timing synchronization attack.

E. Evaluation

We plot the the average percentage of new variation across the four persons for each attack according to Table I to Table VIII in Figure 5 and Figure 6 respectively. Through these tables and figures, we can see that more the change in pressure, the more effective is the attack. However, in Table I, it can be seen that the average change in the percentage of variation when K is 10% is higher than both when K is 15% and 20% respectively. This abnormal phenomenon occurs due to two reasons. One is that the metric formulations that transform the raw pressure data to variation is not a linear function. Although changing by more percentage could easily disturb the original pressure waveform but it does not mean that higher the disturbance; higher the variation. The third constraint that the number of steps of an individual patient cannot be changed beyond a certain percentage after an attack is the second reason for this phenomenon. Changing the pressure readings by 20 percent might have disturbed the pressure waveform too much such that the number of steps changed beyond σ. So in order to satisfy the third constraint, it will choose some less effective sensors to attack which will lead to lower change in the percentage of variation.

In general, a few conclusions can be drawn from the results. (i) Both attacks can dramatically change the variation for both stride period metric and double support metric. (ii) Within certain scope, when the pressure of the sensors change by more percentage, the attack is more effective. (iii) Within certain scope, when the pressure of the sensors is postponed for more time slots, the attack is more effective. However, note that in our test, we assume that the number of time slots to postpone is fixed. As part of the future work, one possible improvement is to postpone the pressure of some sensors by a random number of time slots. (iv) the scenario with 99 sensors is more resilient against attacks compared to the scenario with 20 sensors.

VI. Defense

We propose some corresponding technology to possibly detect and analyze semantic security attacks in this section. We first explain the technique in detail and then study their effect when an attack happens through experimental results. Finally we evaluate the different technologies.

A. Goals and Challenges

Two essential goals are addressed for the defense technology. The first one is to detect, diagnose, and also to remove

Test Cases	Original Var.	K=5%	K=10%	K=15%	K=20%	T=1	T=3	T=5	T=7
P1	5.619	-0.23%	+55.37%	+10.45%	+48.02%	+138.81%	+147.27%	+171.62%	+172.78%
P2	10.432	-6.48%	+13.71%	+8.98%	+16.08%	+47.02%	-62.22%	-66.24%	-66.38%
P3	6.067	-31.72%	-35.31%	-35.42%	-35.87%	+107.74%	+77.54%	-42.18%	+81.06%
P4	6.447	-17.16%	-17.16%	-17.39%	-17.62%	+81.44%	+70.14 %	+115.70 %	+86.74%
Average change % (abs)		13.90%	30.39%	18.06%	29.4%	93.75%	89.29%	98.94%	101.74%

TABLE I: Effect of the attacks on stride period metric, left foot, 99 sensors scenario. P1, P2, P3, and P4 are 4 tested persons. The change of k corresponds to calibration attack, and the change of T corresponds to Timing synchronization attack. The last row is the average change percentage of variation across the 4 persons tested.

Test Cases	Original Var.	K=5%	K=10%	K=15%	K=20%	T=1	T=3	T=5	T=7
P1	7.328	-26.15%	-28.32%	-29.16%	+32.61%	+139.04%	+106.99%	+187.30%	+263.99%
P2	24.709	-59.61%	-59.78%	-59.44%	-61.55%	-70.79%	-61.66%	-63.48%	-66.74%
P3	7.05223	-13.13%	-13.13%	-13.15%	+35.41%	+115.00%	+185.86%	+194.16%	+201.18%
P4	10.939	-48.55%	-48.83%	-48.96%	-51.87%	-44.39%	+47.48%	+104.33%	+104.95%
Average change % (abs)		36.86%	37.52%	37.68%	45.36%	92.30%	100.50%	137.32%	159.22%

TABLE II: Effect of the attacks on stride period metric, left foot, 20 sensors scenario.

Test Cases	Original Var.	K=5%	K=10%	K=15%	K=20%	T=1	T=3	T=5	T=7
P1	7.921	-32.94%	-32.93%	-33.72%	-38.37%	+82.33%	+91.50%	+90.49%	+126.32%
P2	10.971	-20.07%	-22.74%	-22.37%	-23.62%	+51.19%	-66.06%	-66.08%	-71.97%
P3	7.015	-14.26%	-14.26%	-15.34%	+22.81%	+77.47%	+52.42%	+47.82%	+52.39%
P4	8.887	-30.12%	-30.29%	-31.98%	-36.96%	-43.99%	-46.21%	-45.79%	-51.41%
Average change % (abs)		24.35%	25.06%	25.85%	30.44%	63.50%	64.05%	62.55%	75.52%

TABLE III: Effect of the attacks on stride period metric, right foot, 99 sensors scenario.

Test Cases	Original Var.	K=5%	K=10%	K=15%	K=20%	T=1	T=3	T=5	T=7
P1	5.508	-45.53%	-54.61%	-55.74%	-53.20%	+152.18%	+160.89%	+185.00%	+228.76%
P2	15.958	-49.02%	-52.21%	-56.27%	-59.49%	-51.98%	-72.544%	-56.94%	-73.75%
P3	10.724	-65.15%	-65.71%	-66.13%	-66.13%	-59.42%	+98.05%	+129.12%	+132.70%
P4	10.527	-41.79%	-52.58%	-53.07%	-54.09%	-43.00%	+92.18%	+99.56%	+118.65%
Average change % (abs)		50.37%	56.28%	57.8%	58.23%	76.65%	105.92%	117.66%	138.47%

TABLE IV: Effect of the attacks on stride period metric, right foot, 20 sensors scenario.

Test Cases	Original Var.	K=5%	K=10%	K=15%	K=20%	T=1	T=3	T=5	T=7
P1	0.832	+1.92%	+1.92%	+5.77%	+33.65%	+46.51%	+65.02%	+58.05%	+65.75%
P2	3.280	+12.53%	+13.45%	+22.62%	+25.37%	-52.59%	-53.08%	-55.24%	-41.16%
P3	2.168	+1.38%	+1.38%	+1.38%	-14.11%	-32.24%	-63.88%	-71.54%	-75.88%
P4	1.254	-0.08%	-2.47%	+1.20%	-6.20%	-16.51%	-20.89%	-20.89%	-20.26%
Average change % (abs)		3.98%	4.81%	7.74%	19.83%	36.96%	50.71%	51.43%	50.76%

TABLE V: Effect of the attacks on double support metric, left foot, 99 sensors scenario.

Test Cases	Original Var.	K=5%	K=10%	K=15%	K=20%	T=1	T=3	T=5	T=7
P1	1.024	-5.57%	-18.16%	-16.60%	+34.96%	+84.92%	+46.43%	+89.54%	+172.18%
P2	4.345	-14.88%	-15.92%	-16.21%	-18.91%	-26.32%	-38.56%	-24.50%	+31.77%
P3	1.617	-2.10%	+34.94%	+31.91%	+142.24%	+114.68%	+245.23%	+276.42%	+221.76%
P4	1.170	+9.49%	+12.22%	+13.33%	+14.78%	+19.06%	+39.67%	+88.81%	+129.31%
Average change % (abs)		8.01%	20.31%	19.51%	52.72%	61.25%	92.47%	119.82%	138.76%

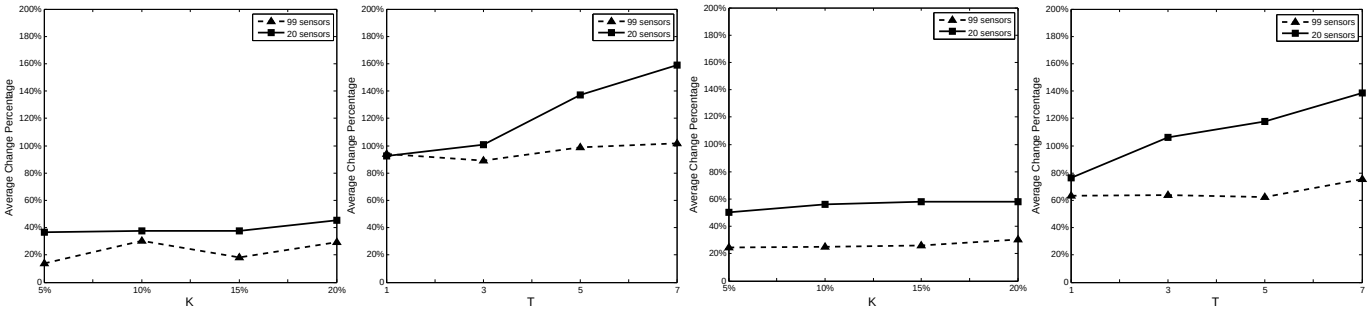
TABLE VI: Effect of the attacks on double support metric, left foot, 20 sensors scenario.

Test Cases	Original Var.	K=5%	K=10%	K=15%	K=20%	T=1	T=3	T=5	T=7
P1	0.832	+130.77%	+130.77%	+135.58%	+140.38%	+158.26%	+170.93%	+178.81%	+232.01%
P2	3.280	+18.96%	+28.54%	+23.54%	+34.09%	+46.95%	-81.52%	-83.80%	-86.59%
P3	2.168	+1.38%	+7.15%	+15.68%	-27.81%	-52.80%	-81.62%	-82.20%	-80.90%
P4	1.254	-45.37%	-45.37%	-45.37%	-50.00%	-58.25%	-61.05%	-61.36%	-61.05%
Average change % (abs)		49.12%	52.96%	55.04%	63.07%	79.07%	98.78%	101.54%	115.14%

TABLE VII: Effect of the attacks on double support metric, both feet, 99 sensors scenario.

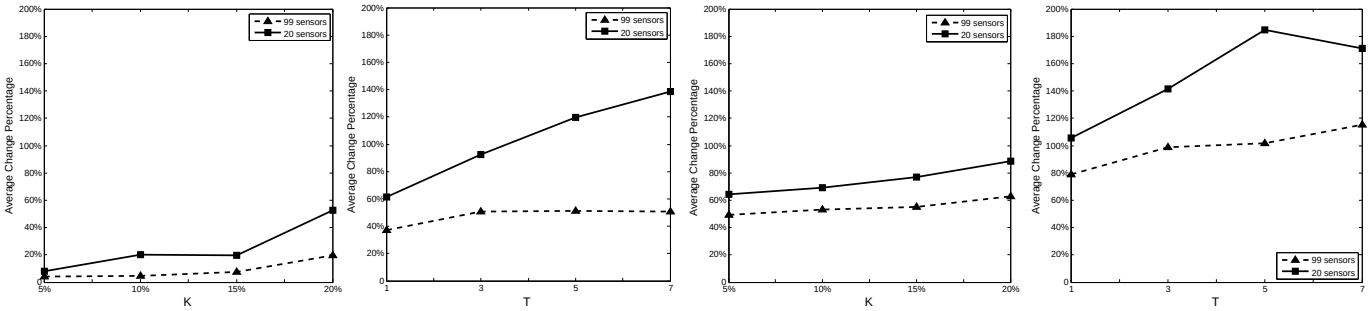
Test Cases	Original Var.	K=5%	K=10%	K=15%	K=20%	T=1	T=3	T=5	T=7
P1	1.024	+123.44%	+125.00%	+125.78%	+131.84%	+134.29%	+175.78%	+229.56%	+228.35%
P2	4.345	-42.28%	-57.38%	-58.92%	-61.15%	-76.40%	-83.80%	-82.08%	-85.25%
P3	1.617	+46.07%	+46.26%	-75.26%	+108.23%	+148.56%	+243.52%	+315.97%	+261.71%
P4	1.170	-44.62%	-47.26%	-48.63%	-52.65%	-62.71%	-62.80%	+110.78%	+110.07%
Average change % (abs)		64.10%	68.98%	77.15%	88.47%	105.49%	141.48%	184.60%	171.35%

TABLE VIII: Effect of the attacks on double support metric, both feet, 20 sensors scenario.



(a) (b) (c) (d)

Fig. 5: Average change percentage in the following conditions. (a) Calibration attack on stride period metric, left foot. (b) Timing synchronization attack on stride period metric, left foot. (c) Calibration attack on stride period metric, right foot. (d) Timing synchronization attack on stride period metric, right foot.



(a) (b) (c) (d)

Fig. 6: Average change percentage in the following conditions. (a) Calibration attack on double support metric, left foot. (b) Timing synchronization attack on double support metric, left foot. (c) Calibration attack on double support metric, both feet. (d) Timing synchronization attack on double support metric, both feet.

the impact of attacks as much as possible. The second one is to design the defense to be low-cost, low-energy and real-time which is due to the unique properties of the wireless sensor networks, especially for medical devices. Most wireless medical devices have high limitations regarding cost and energy consumption.

The main challenge in defense technology is to find out whether the data abnormality is because of an attack or due to the sickness of the person. Suppose the medical expert simply checks the pressure distribution of sensors at each time slot. When an attack happens, it is possible that this pressure distribution might be abnormal. However, if a person is sick, this particular distribution also has high likelihood to show similar abnormality. In this case, the data abnormality found by the medical expert cannot be used to successfully detect and distinguish attacks. The above challenge is taken into consideration in our suggested defenses.

We propose two technologies for defense. The first technology targets at using several subgroups of the sensors to verify the difference in variation. It is simple as the medical expert only needs to repeat the calculation of variation on different groups of sensors. The second one is a new concept but it is even more straightforward, the basic idea is to attack again on the received data to detect previous attacks, thus, only re-attacks are required. We explain their technical details in the following section.

B. Defense Procedure

We demonstrate two types of defense and their modeling in this part. We assume that the defense technology is applied after the attack, which means that the goal of defense is to detect the attack rather than preventing it.

1) Multi-group Verification: The basic idea of this defense is to divide the original sensors into several subgroups. We claim that the variation calculated for each subgroup should be close enough if the subgroups are properly divided. Since the sensors with numbers next to each other are usually present near each other as shown in Figure 1, the pressure values of the two neighboring sensors are highly correlated. Therefore, if the sensors are divided into two randomly selected groups, e.g., odd group and even group according to their sensor number, the two groups should have similar waveforms, which will lead to similar variation. Note that in this defense, even though a person is sick, because of the neighboring correlation, his/her variation between odd/even group will be similar. However, when an attack happens, due to the fact that the effect of an attack can be very different in each individual subgroup, the variation for each of them may vary a lot. In the following experiment, we divide the original sensors into odd and even groups and then we compare α as defined below. If α becomes extraordinary large as compared to the α of history data or some other threshold, the medical expert can conclude that the sensors were attacked. Algorithm 2 describes the multi-group verification defense algorithm.

$$\alpha = \frac{|Var_{odd} - Var_{even}|}{min(Var_{odd}, Var_{even})} \tag{6}$$

Algorithm 2 multi-group verification defense

Input: P - original sensor pressure at each time slot.

```
(P_odd, P_even) = SplitOddEven(P)
Var_odd = Var.(P_odd)
Var_even = Var.(P_even)
α = |Var_odd − Var_even| / min(Var_odd, Var_even)
if α>threshold then
  Attack Detected
else
  Attack Not Detected
end if
```

2) Repeated Attack based Defense: This defense targets at reversing the original attacks. When the attacker performs either attack as mentioned in section V, the change of variation can be in two directions, either increasing or decreasing. We suppose that the variation of the data before attack is Var_0 and the variation after attack is Var_1. Our proposed way to defend is to perform the same attack again on the data, but for two times. For the first time, we perform the attack only in the direction to increase Var_1 (the medical expert only has the data after attack, which has the variation of Var_1) and for the second time, we perform the attack only in the direction to decrease Var_1. So our intuition is that only one of the directions (because the 2 directions are opposite) will reverse Var_1 back to Var_0. However, it would be much harder to change Var_1 in the opposite direction of Var_0 due to the difficulty in changing the variation in one direction continuously. But for a piece of data that has not been attacked before, both directions can cause similar effects (the absolute value of the change) to the original variation. We further use β defined below to compare the effect of performing attack in two opposite directions. If the data has not been attacked before, β will be approaching 1, otherwise it would be a large value. Therefore, the medical expert will be able to deduce whether the data has been attacked or not and also that in which direction the variation has been modified.

$$\beta = \frac{min(|Var_{dir1} - Var_1|, |Var_{dir2} - Var_1|)}{max(|Var_{dir1} - Var_1|, |Var_{dir2} - Var_1|)} \tag{7}$$

C. Experimental Results

Table IX and Table X show the effect of multi-group verification on calibration and timing synchronization attack. In each table, we consider 2 metric and 2 scenarios. We further assume that the attacks happen on left foot. The readings in the tables indicate the value of α (defined in Eq. 6) before and after the attack. Both tables show the average test results on P1, P2, P3, and P4 which are used to illustrate the average performance of the defense methodology. Similarly, Table XI and Table XII demonstrate the effect of repeated attack based defense and use value of β (defined in Eq. 7) to make the decision.

Test Situation	Original α	K=5%	K=10%	K=15%	K=20%
Stride Period, left foot, 99 sensors	0.0589	0.2466	0.2568	0.2796	0.2901
Stride Period, left foot, 20 sensors	0.2032	0.3992	0.3653	0.4695	0.7624
Double Support, left foot, 99 sensors	0.2695	0.2707	0.2996	0.4194	0.3299
Double Support, left foot, 20 sensors	0.2971	0.2794	0.5888	0.4381	0.4572

TABLE IX: Multi-group verification to calibration attack, tested on P1, P2, P3, P4. The value in the table is the average α across the 4 tested persons under different conditions.

Test Situation	Original α	T=1	T=3	T=5	T=7
Stride Period, left foot, 99 sensors	0.0589	0.2636	0.4487	0.3588	0.4526
Stride Period, left foot, 20 sensors	0.2032	0.5907	0.4134	0.5753	0.4764
Double Support, left foot, 99 sensors	0.2695	0.4340	0.3583	0.4621	0.6184
Double Support, left foot, 20 sensors	0.2971	0.5659	0.3331	0.6842	0.4351

TABLE X: Multi-group verification to timing synchronization attack.

Test Situation	Original β	K=5%	K=10%	K=15%	K=20%
Stride Period, left foot, 99 sensors	2.21	21.87	37.78	20.78	29.91
Stride Period, left foot, 20 sensors	2.62	33.90	30.99	8.21	6.85
Double Support, left foot, 99 sensors	2.43	3.99	5.20	27.4	37.66
Double Support, left foot, 20 sensors	2.36	16.03	13.79	8.34	22.39

TABLE XI: Repeated attack based defense to calibration attack, tested on P1, P2, P3, P4. The value in the table is the average β across the 4 tested persons under different conditions.

Test Situation	Original β	T=1	T=3	T=5	T=7
Stride Period, left foot, 99 sensors	2.32	6.69	38.07	39.07	35.03
Stride Period, left foot, 20 sensors	3.28	23.12	13.45	15.29	4.47
Double Support, left foot, 99 sensors	2.50	5.52	7.16	21.62	13.65
Double Support, left foot, 20 sensors	6.77	6.46	10.42	6.35	17.98

TABLE XII: Repeated attack based defense to timing synchronization attack.

Algorithm 3 Repeated Attack based Defense

Input: P - original sensor pressure at each time slot.
Input: K - percentage of the pressure change in attack.
Input: N - number of sensors to attack.
Input: σ - error rate of the number of steps.
Input: M - number of optimal values to preserve in previous DP. step.
AttackToIncrease() and AttackToDecrease() are the DP. algorithm in Algorithm 1.

1: Var_1 = Var.(P)
2: Var_{dir1} = Var.(AttackToIncrease(P, K, N, σ, M))
3: Var_{dir2} = Var.(AttackToDecrease(P, K, N, σ, M))
4: $\beta = \frac{min(|Vardir1 - Var_1|, |Var_{dir2} - Var_1|)}{max(|Var_{dir1} - Var_1|, |Var_{dir2} - Var_1|)}$
5: **if** β>$threshold$ **then**
6: Attack Detected
7: **else**
8: Attack Not Detected
9: **end if**

D. Evaluation

According to the Multi-group verification defense results in Table IX and Table X, although the original α varies from person to person, in most cases, it is relatively small. For both calibration and timing synchronization attack, the α value becomes extraordinarily large as compared to the original α. Similarly, Repeated attack based defense results in Table XI and Table XII demonstrate that the β value increases significantly as compared to original β. However, in some cases the defenses are not fully dependable, as observed in Table XII with metric double support and 20 sensors scenario, when $T = 5$, the value of β drops from 6.77 to 6.35 after timing synchronization attack which is adversary to the purpose of the defense. Besides, there is no significant linear relationship between α and different values of K or T, so as β. This is due to the metric formulations that transform the raw pressure data to variation which is not a linear function. We cannot conclude that higher K or T will result in higher α and β.

For the practical application of these defenses, first we could use history data to get the original α and β. And then we can set a threshold for original α and β based on a particular individual. Once the new α or β is higher than the threshold, the medical expert can conclude that the sensors are attacked. In our experiment, we set the α threshold as double of original

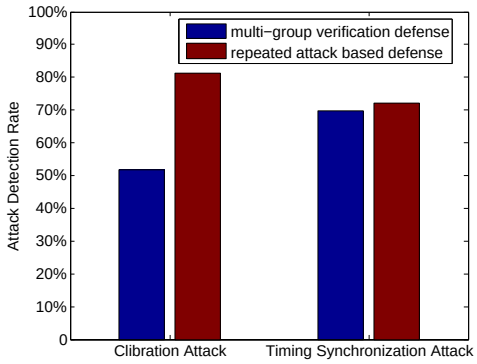


Fig. 7: Attack detection rate of the two defenses.

α and set the β threshold as double of original β. This set threshold is significant enough to detect the change in both α and β and we get the following results. As shown in Figure 7, the multi-group verification defense could detect 51.78% of calibration attack and 69.64% of timing synchronization attack. The repeated attack based defense could detect 81.25% of calibration attack and 72.22% of timing synchronization attack.

In general, a few conclusions can be drawn from the results. (i) The defense is subject to each individual separately. Hence setting a universal threshold is not possible. (ii) Both defenses are effective against the two types of attack. (iii) Considering our assumed threshold, the repeated attack based defense is more effective than the multi-group verification defense.

VII. Conclusion

We have discovered a new area of research by analyzing the impact of semantic attacks on the security of wireless medical devices. We explored the attacks on embedded medical sensor network by way of optimizing the damage yet being under the constraints of producing unsuspicious results. In our experiment, we simulated two different types of attack under two different scenarios to prove the susceptibility of medical devices against these simple attacks. We also developed two real-time defenses in order to detect these semantic attacks. By virtue of comparing the results with respect to the two scenarios, we concluded that energy efficient medical devices are easier to attack as the redundant sensors are not available for defending against these attacks. By attacking the medical device using the calibration and timing synchronization attack; on an average we are able to change the variation up to 88.47% and 184.6% respectively. Variation is the change caused in the medical metrics and hence in the medical diagnosis because of these attacks. After the application of our proposed defenses using our assumed threshold; the multi-group verification defense can detect upto 69.64% of attacks and repeated attack based defense can detect upto 81.25% of attacks.

VIII. Acknowledgement

This work was supported in part by the NSF under Award CNS-0958369, Award CNS-1059435, and Award CCF-0926127, and in part by the Air Force Award FA8750-12-2-0014.

References

[1] H. S. Ng, M. L. Sim, and C. M. Tan, "Security issues of wireless sensor networks in healthcare applications," *BT Technology Journal*, vol. 24, no. 2, pp. 138-144, 2006.

[2] M. A, Ameen, J. Liu, and K. Kwak, "Security and privacy issues in wireless sensor networks for healthcare applications," *Journal of medical systems* vol. 36, no. 1, pp. 93-101, 2012.06.

[3] H. Noshadi, S. Ahmadian, H. Hagopian, J. Woodbridge, N. Amini, F. Dabiri, and M. Sarrafzadeh, "Hermes-Mobile Balance and Instability Assessment System," *BIOSIGNALS*, 2010.

[4] B. E. Maki, "Gait changes in older adults: predictors of falls or indicators of fear." *Journal of the American geriatrics society*, vol. 45, no. 3, pp. 313-320, 1997.

[5] D. Halperin, T.S. Heydt-Benjamin, B. Ransford, S.S. Clark, B. Defend, W. Morgan, K. Fu, T. Kohno and W.H. Maisel, "Pacemakers and implantable cardiac defibrillators: Software radio attacks and zero-power defenses," *In IEEE Symposium on Security and Privacy*, pp. 129-142, 2008.

[6] B. Haran, and D. Senouf, "Remote monitoring and follow-up of pacemakers and implantable cardioverter defibrillators," *Europace*, vol. 11, no. 6, pp. 701-709, 2009.

[7] D. Halperin, T. S. Heydt-Benjamin, K. Fu, T. Kohno, and W. H. Maisel, "Security and privacy for implantable medical devices," *Pervasive Computing, IEEE*, vol. 7, no. 1, pp. 30-39, 2008.

[8] S. Gollakota, H. Hassanieh, B. Ransford, D. Katabi, K. Fu, "They can hear your heartbeats: non-invasive security for implantable medical devices," *ACM SIGCOMM Computer Communication* , vol. 41, no. 4, pp. 2-13, 2011.

[9] W. H. Maisel and T. Kohno, "Improving the security and privacy of implantable medical devices," *New England journal of medicine*, vol. 362, no. 13, pp. 1164, 2010.

[10] J. B. Wendt and M. Potkonjak, "Medical Diagnostic-Based Sensor Selection," *IEEE Sensors*, pp. 1507-1510, October 2011

[11] X. Lin , R. Lu , X. Shen , Y. Nemoto and N. Kato, "SAGE: a strong privacy-preserving scheme against global eavesdropping for ehealth systems," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 4, pp. 365-378, 2009.

[12] M. Potkonjak, S. Meguerdichian, J.L. Wong, "Trusted Sensors and Remote Sensing," *IEEE Sensors*, pp. 1104-1107, 2010.

[13] J. B. Wendt, M. Potkonjak, "Nanotechnology-Based Trusted Remote Sensing," *IEEE Sensors*, pp. 1213-1216, October 2011.

[14] M. Barua, X. Liang, R. Lu, and X. Shen, "ESPAC: Enabling Security and Patient-centric Access Control for eHealth in cloud computing," *International Journal of Security and Networks*, vol. 6, no. 2, pp. 67-76, 2011.

[15] L. Fan, W. Buchanan, C. Thummler, O. Lo, A. Khedim, O. Uthmani, A. Lawson, and D. Bell, "Dacar platform for ehealth services cloud," *In IEEE International Conference on Cloud Computing (CLOUD)*, pp. 219-226, 2011.

[16] P. Kumar and H.J. Lee, "Security issues in healthcare applications using wireless medical sensor networks: A survey," *Sensors*, vol. 12, no. 1, pp. 55-91, 2011.

[17] B. Wayne, S. S. Clark, B. Ransford, and K. Fu, "Design challenges for secure implantable medical devices,"*In Proceedings of the 49th Annual Design Automation Conference ACM*, pp. 12-17, 2012.

[18] T. Xu, J. B. Wendt, M. Potkonjak, "Digital Bimodal Function: An Ultra-Low Energy Security Primitive," *ISLPED*, 2013.

[19] T. Xu, M. Potkonjak, "Robust and Flexible FPGA-based Digital PUF," to appear in *FPL*, 2014.

[20] T. Xu, M. Potkonjak, "Lightweight digital hardware random number generators," *IEEE SENSORS*, pp. 1-4, 2013.

[21] Novel.de, Pedar, 2007, *http://www.novel.de/*