Towards a Framework for Safety Analysis of Body Sensor Networks

Philip Asare[1,2], John Lach[1],
John A. Stankovic[2],
[1]Charles L. Brown Department of Electrical and Computer Engineering
[2]Department of Computer Science
University of Virginia, Charlottesville, VA USA 22904
{asare, jlach, stankovic}@virginia.edu

Yi Zhang, Paul L. Jones, Sandy Weininger
U.S. Food and Drug Administration
Silver Spring, MD USA 20993
{Yi.Zhang2, Paul.Jones, Sandy.Weininger}@fda.hhs.gov

ABSTRACT

Body sensor networks (BSNs) are an emerging class of medical cyber-physical systems which have the potential to change the healthcare paradigm. However, they present many new challenges, chief of which (like any medical device) is assuring patient safety. This requires not only a precise definition of safety, but also techniques for assessing the safety of BSN designs. Although solutions are possible and important for specific BSNs used in specific applications, addressing this issue on a case-by-case basis usually results in an ad-hoc process, and more importantly, makes the translation of experiences and solutions between different applications more difficult.

A generic and conceptual framework for guiding the safety analysis process would provide all stakeholders a common basis for communicating, discussing, and examining the safety of BSN designs, and provide manufacturers with an exemplary process that they can follow to improve and gain confidence in the safety of their devices. This paper presents our current efforts in developing such a framework. In particular, we present a theoretical foundation for modeling and analyzing BSNs, and identify the general class of hazards based on this foundation. These efforts explore critical issues that deserve attention in designing safe BSN systems, and more importantly, can help advance the understanding of BSNs and their safety.

Categories and Subject Descriptors

J.3 [**Life and Medical Sciences]:** – *health, medical information systems*; H.1.1 [**Models and Principles]:** Systems and Information Theory – *general systems theory, value of information*

General Terms

Design, Verification

Keywords

Body sensor networks, safety analysis, model-driven design

1. INTRODUCTION

Body sensor networks (BSNs) present a unique opportunity for improving the quality and mobility of healthcare. Such systems enable patients to continue their normal daily lives and 'invisibly' collect patient information under dynamically changing environments. The collected information enables healthcare practitioners to access otherwise-unobtainable information to assist and improve medical decision-making, and to gain better understanding of how the human body functions in various environments. Realization of the BSN vision will significantly influence both medical research and practice, as evidenced by a number of preliminary studies highlighted in [8].

The ultimate challenge is to assure the safety of patients who use BSNs[1]. In this paper, we focus on BSNs that are sense-only systems. Even though BSNs do not directly deliver treatment or medication to patients, they do collect and supply information that, when used in medical decision-making processes, have significant impact on the correctness of decisions made, and hence on the patient's safety.

Assuring patient safety is especially challenging in BSNs because of their differences from their in-clinic counterparts. BSNs are typically governed by more stringent constraints on their resource consumption (e.g., energy and computational resources), as well as mobility and device size constraints (see [3] for discussion of such issues). More importantly, BSNs are operated in scenarios typically outside a clinical environment. Therefore, access to medical practitioners and service technicians is usually limited.

To improve the safety of BSNs, we propose to develop a generic model-driven design process for BSNs, with the following long-term goals:

1. Establish a model-driven design process for BSNs that manufacturers can follow to improve on and gain confidence in the safety of their products.
2. Provide all stakeholders a common basis, in the form of a generic model, for communicating, discussing, and examining the safety properties of BSN designs.
3. Provide mechanisms for formal verification of the generic model. Once the safety properties of this generic model are established (e.g., using formal verification), manufacturers can then use it as a safety reference model to challenge the behavior of their own products, or as a starting point to develop their specific designs.

In this paper, we present the theoretical foundation for achieving these goals. In particular, we

[1] Unlike other more traditional clinically based medical devices supported by trained personnel, BSNs usually interact with patients directly, placing the responsibility of operation on the patient.

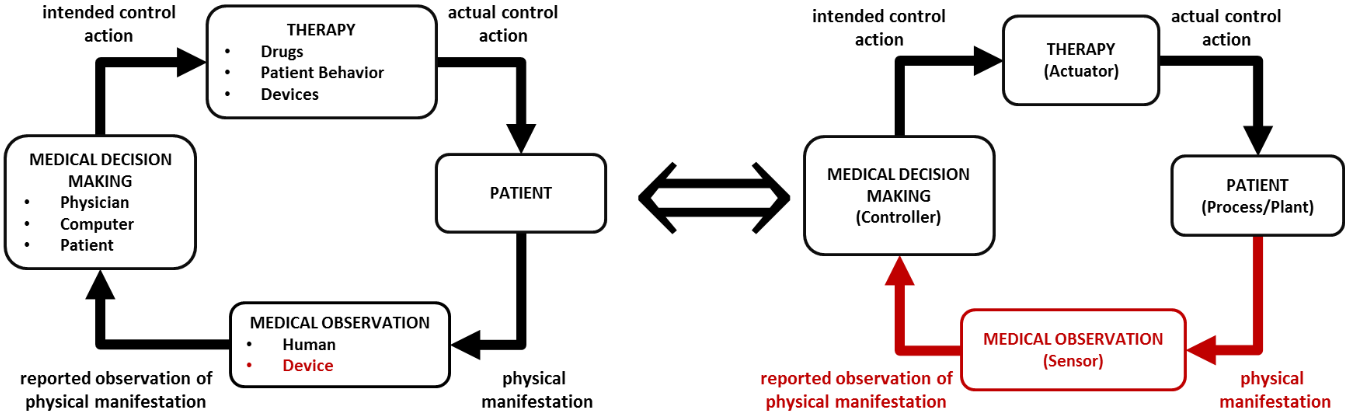

Figure 1: Personal health management as a control loop

- Discuss conceptually the role of BSNs in the health management system loop.
- Provide a formal definition of system scope for BSNs.
- Provide a formal definition of patient safety for BSNs.
- Identify the general class of hazards for BSNs based on the system scope and definition of patient safety.

These contributions set the stage for developing analysis techniques for checking these safety properties, which is part of our on-going work.

2. BSN PATIENT SAFETY

Recent advances in system safety engineering define the notion of safety from the perspective of control loops [7], where hazards result from ineffective control mechanisms and can be mitigated by designing more effective controls into the system. Following the same idea, personal health management can indeed be viewed, at a higher level, as a complex control problem, in which BSNs can play an important and promising role. Figure 1 shows various aspects of health management (the left half of the figure) related to (simplified) control loop concepts (the right half of the figure).

The person whose health is being managed (i.e., the patient) constitutes the (controlled) process or plant. The activities essential to health management can be related to the control loop as follows:

1. **Sensor (medical observation):** devising means of 'observing' the physical manifestations of the patient's state either using human observation or a device (like a BSN).
2. **Controller (medical decision making):** developing techniques for 'estimating' the patient's health state from the observations, and strategies (based on the current understanding of human bodily function) for using results of this inference to make decisions in order to influence the patient's state if necessary.
3. **Actuator (medical therapies):** developing therapies and behaviors that can be delivered or recommended to the patient to control their health state when necessary.

This view of health management assembles all involved elements around the patient, and offers the required viewpoint for considering patient safety: a health management system is unsafe if any undesired behaviors in any elements of the control loop put the patient in an undesired state, harming rather than helping the patient.

To put the above in concrete terms, we consider an example BSN. A BSN capable of electrocardiography and activity sensing (e.g. [10]) can be used to determine if a patient has arrhythmias (irregular heartbeats) and the activity factors surrounding those arrhythmias, so that appropriate therapy can be prescribed. The doctor has some medical knowledge, which based on the information provided by these sensors, will allow him or her to make an appropriate therapy decision. The therapies recommended by the doctor may include drugs, diets, pacemakers, or particular physical therapy to help the patient's heart condition improve.

A number of things can go wrong with the BSN in this example: the sensors may provide misleading information, the doctor may misinterpret the sensor information, the sensors may heat up and burn the patient, and many other scenarios. Any of these would cause the control loop to harm the patient and hence make the control loop unsafe.

In a typical safety analysis (as discussed above), safety is defined as a property of the whole system, especially on its control aspect and interaction with the operational environment. Thus, a meaningful safety analysis of a system requires a comprehensive knowledge of the whole system and an accurate assumption of its environment.

In the BSN case, however, it is not always practicable to acquire all information for the controller and the plant in the loop. It is also difficult to make accurate assumptions on the dynamic environment. Fortunately, complete knowledge of the controller or the plant may not be necessary for deciding if a BSN is 'appropriate' for being used in a health management control loop.

We notice that the knowledge of what information the controller expects from the sensor (BSN) is usually available. Hence, it is possible to abstract away technical details of the controller and its interaction with the patient (process/plant), and derive expectations on the output behavior of the sensor (with respect to the controller). Also known at the design time is how the patient responds to output from the BSN, from which expectations on the specification of the BSN's output with respect to the patient can be derived.

Thus, we consider a BSN as 'appropriate' to be used in a health management control loop, if it meets all output behavior expectations, with respect to both the patient and their environment, and the controller. Normally, such expectations are not absolute. Instead, they must tolerate deviations from the ideal within certain bounds. These bounds essentially define the 'robustness' of the rest of the control loop with respect to the

BSN. The ‘appropriateness’ of deviations in a particular loop is our notion of safety for the BSN.

3. FORMAL BSN PATIENT SAFETY

In this section, we define formally the BSN system scope, provide a general definition of safety based on this system scope, and provide refinements to the relevant aspects of the system scope in order to set the stage for discussing the general BSN hazards.

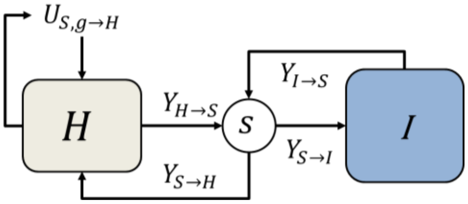

Figure 2: Abstract view of BSN operational scenario

3.1 BSN System Scope

Figure 2 shows an abstract view of the BSN operational scenario described earlier. Part of the BSN (the sensing process S) must be coupled to the patient (the patient and his or her environment H) through the ‘input’ $U_{S,g\to H}$ (“g” indicates that the input is a configuration), so that the BSN can sense the outputs from the patient ($Y_{H\to S}$). For example, in electrocardiography, electrodes are usually placed on the patient’s limbs and chest (the coupling $U_{S,g\to H}$) in order to sense the appropriate voltages (the output $Y_{H\to S}$) from the electrical activity of the heart (the process in H of interest).

Note that the BSN may not be ‘interested’ in all outputs in $Y_{H\to S}$, and some of these outputs may hinder its ability to achieve its goals. For example, electrical interference from other sources (remember that the patient environment is part of H) may show up in an ECG and distort the intended signal. Such inputs to the BSN are termed as *interfering inputs* if they resemble the signal the BSN is trying to actively sense; or *modifying inputs* if they are signals or energies that the BSN is not actively sensing, but is sensitive to, and hence can affect its operation [9] (e.g. heat affecting electrical circuitry in the BSN). The signals that the BSN is ‘interested’ in are termed as *desired inputs*.

The BSN may produce outputs ($Y_{S\to H}$) that the patient is sensitive to: some of these outputs may be intentional energy exposed to the patient to aid in sensing (e.g. light energy used in pulse oximetry), while others may be produced due to the physical nature of the BSN.

The goal of the BSN is to produce medically-relevant information ($Y_{S\to I}$) for the clinician (the inference process *I*). The inference process (I) abstracts the rest of the loop (*i.e.* it accounts for the medical decision-making and therapy parts of the loop, as well as their effect on the patient), and it can provide information ($Y_{I\to S}$) to the BSN to aid in its operation. If not otherwise indicated, all signals and processes discussed in the paper evolve over time.

3.2 Formal Definition of Patient Safety

The $Y_{S\to I}$ and $Y_{S\to H}$ interfaces are the interfaces on which the safety expectations must be placed, since it is through these interfaces that the BSN can either directly harm the patient or drive the rest of the loop to harm the patient. As mentioned previously, safety (or ‘appropriateness’ for a particular loop) can be expressed in terms of deviations from the ideal. That is, if we assume that the expected behavior of the BSN can be characterized by an ideal BSN (an ‘Oracle’), then whether or not a designed BSN is ‘appropriate’ for use can be defined as whether or not the output of this BSN is consistent with that of the ideal BSN under all circumstances.

More formally, a BSN S is appropriate for use, with respect to the ideal BSN S_0, if it satisfies safety constraints formulated in Equation (1) for all safety bounds $b_{S\to(*)|z} \in \mathcal{B}_z$:

$$f_z\left(Y_{S_0\to(*)}, Y_{S\to(*)}\right) \leq b_{S\to(*)|z} \qquad (1)$$

where $Y_{S_0\to(*)}$ and $Y_{S\to(*)}$ denote the outputs to component $(*)$ produced by S_0 and S, respectively; $(*)$ can be either the inference process (I) or the human process (H); $\mathcal{B}_z$ is the set of safety bounds; f_z is a comparison function designated for property z, and z can be timing, accuracy, or output level properties.

The assumption for the above definition is that the ideal BSN is available to the designer. Moreover, recent work [1, 11] shows that safety bounds $b_{S\to(*)|z}$ can be determined on the values of outputs of a medical instrument and used to inform system design. Though the results are preliminary, it is reasonable for our purposes to assume that such bounds exist and can be determined.

3.3 BSN Output Interfaces

The nature of the BSN output interfaces determines the property z identified above.

3.3.1 The Energy Output Interface ($Y_{S\to H}$)

The $Y_{S\to H}$ interface consists of a number of continuous time outputs $y^i_{S\to H}(t) \in Y_{S\to H}$ from the BSN, whose values are energy quantities (e.g., light, heat, electromagnetic radiation, chemical concentrations). These values may be constant or changing over time. As mentioned previously, these energies may be generated intentionally to affect the patient to aid in sensing, or generated intentionally for other purposes (e.g. electromagnetic radiation for wireless communication), or inevitably produced as part of the BSN operation (e.g., heat from circuits).

3.3.2 The Information Output Interface ($Y_{S\to I}$)

The BSN is a sensing process that takes in analog inputs from the patient and produces digital outputs for the clinician or patient/user based on these inputs and its assumptions about its operational environment. We first present the general forms in which the digital outputs can exist with a number of examples, and then present an abstract model of BSN information output that can be used to represent all these forms of information output.

General Forms of BSN Information Output

The simplest general form of information output from the BSN is what we term as *sets of values*. These are essentially snapshots of information from the patient which are not ordered. For example, a clinician may be interested in the blood pressure readings for a patient over the course of a day in order to establish a minimum, an average and a maximum, but may not necessarily be interested in the time at which these readings are taken or when each reading occurred relative to the others. Another example may be that the clinician may be interested in the lengths of walks that a patient took over a day or a week, but again not in when these occurred or how they are related in time.

The next form of information output from the BSN is what we term as an *ordered sequence of values*. Here, the values are ordered by time, even though their actual occurrence time or the relative time between values are not indicated. A patient may use a device to measure blood pressure in the morning after waking up, sometime in the afternoon, and in the evening. The actual times may not be known, but the order in which blood pressure readings were taken would.

In some cases, timing information about the data is important and hence the BSN may produce *time-stamped sequence of values*. These time stamps are usually in ‘wall clock’ time, and are typically used when multiple streams of information from the BSN need to be correlated. In the examples given previously, the wall clock times of all the readings can be logged by the system. When the time stamps are not in wall clock time but in some other time reference like a system clock, or when the time between values is known but their wall clock times are not, we call this a *sequence of relatively-timed values*. An ECG strip could be an example of such a sequence since it is sampled at a constant rate; the relative times between samples are known even if the actual wall clock times of the samples are not.

In some cases where discrete events are being monitored or summary information is being presented, the information may reflect the value over a period of time. We term this type of information as a *time-ranged sequence of values*. For example, if the BSN reports that a patient was walking or jogging over a particular period of time, this particular activity becomes the value for that range of time.

A real information stream could be a hierarchical combination of these forms of information. For example, a BSN may report the sets of ECG samples from detected arrhythmias over a day. This would be the highest level in the hierarchy of the information stream. Each ECG sample is a relatively- timed sequence of values since the information is a waveform sampled at a particular frequency. This comprises the next level of the information stream.

Timing of BSN Information Output

A BSN is typically a software-regulated system and usually has communication networks for coordinating operations between its various components and for communicating with the clinician. Thus, it inevitably exhibits particular temporal behaviors.

Assume that the BSN S in Figure 2 is tracking a continuous signal produced by the patient H, and that it needs a segment of this signal of time size t_{seg} in order to compute the ‘value’ of the information. For example, the BSN could be a pulse oximeter that needs to monitor a few milliseconds of the photoplethysmograph in order to compute the heart rate. At time t_1, S obtains a segment of the signal comprising data points between $t_0 = t_1 - t_{seg}$ and t_1. Since it takes some time for S to compute the ‘value’ of the segment, S may have a clock to log its information and may log the time when it obtains the last point of the segment as t_2, where $t_2 - t_1 \geq 0$.

Let us assume that S transmits its value to a remote point where the clinician can obtain this information as soon as the information is available at the remote point. The information from S will be available at this point at some time t_3, where $t_3 > t_1$. When the information is reported, S will assign t_2 as the time for that particular value. We call this time the *reported observation time* of the particular value. The inference process I is first able to obtain this value at t_3. We call this time the *received time* for that particular value. We term t_1 (or t_0 if the system logs the time of all points in the segment) as the *actual time*. These temporal phenomena in the BSN are illustrated in Figure 3.

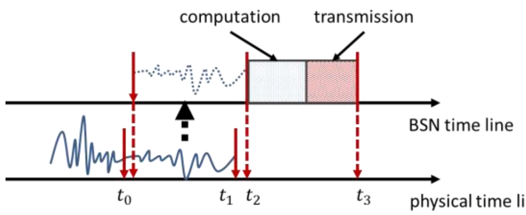

Figure 3: Temporal phenomena in BSN

Abstract Information Output Model

From the inference process’s perspective, the output of the BSN $Y_{S\rightarrow I}$ can be viewed as an output stream, which could consist of multiple substreams $Y_{S\rightarrow I}^i$. The requirement is that each of substreams be homogeneous, in the sense that they must contain the same type of information. Hence, a system that reports activity data correlated with heart rate data may have activity as one substream and the heart rate information as another substream. An information (sub)stream contains a number of data points. For a substream $Y_{S\rightarrow I}^i$, each of its data point, y_j^i, should be a tuple $y_j^i = \left(t(y_j^i), v(y_j^i)\ , t_r(y_j^i)\right)$, where $t(y_j^i)$ is the reported observation time, $v(y_j^i)$ is the value, and $t_r(y_j^i)$ is the received time of y_j^i. Note that $v(y_j^i)$ could itself be a set of data points $y_{j,k}^i$, which have the same tuple form as y_j^i (i.e. they possess the same properties as a data point in a stream). The recursive nature of $v(y_j^i)$ allows us to model hierarchical streams.

It should be noted that we use the term “value” loosely as a value could be as complicated as an image taken by the BSN (in case of a mobile ultrasound, for example). The form of the reported observation time $t(y_j^i)$ determines the stream ‘type’. Table 1 defines different forms of $t(y_j^i)$ and the associate stream types.

Table 1: Association of reported observation time forms and stream types

Form (Notation)	*Meaning*	*Stream Type*
t_j^i	time stamp	time stamped
	relative time	relatively-timed
$[t_{j,0}^i, t_{j,f}^i]$	time range	time-ranged
k_j^i	ordering index	ordered
None or tag	no reported observation time	set

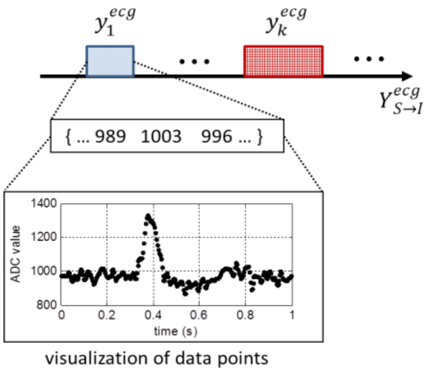

Figure 4: Information output example

Figure 4 shows an example output stream. This stream, $Y_{S\rightarrow I}^{ecg}$, contains sets of data points $\{y_1^{ecg} \dots y_n^{ecg}\}$, each of which is an ECG strip. Each ECG strip contains a sequence of relatively-timed data points (voltage values sampled at 250Hz). Formally, $y_i^{ecg} = (i, v(y_i^{ecg}), t_r(y_i^{ecg}))$, where i is the tag for the strip, and $v(y_i^{ecg})$ is also a sequence of relatively-timed data points $\{y_{i,1}^{ecg} \dots y_{i,n}^{ecg}\}$, such that: 1) $y_{i,j}^{ecg} = (t_j, v(y_{i,j}^{ecg}), t_r(y_{i,j}^{ecg}))$; 2) $t_{j+1} - t_j \approx 4ms$; and 3) $v(y_{i,j}^{ecg})$ is a digital value representing the electrical potential measured at the patients skin surface. If the BSN is assumed to be streaming the ECG data, then the received time of each sample can be assumed to be a small delay from the actual time when the value was measured (i.e., $t_r(y_{i,j}^{ecg}) - t_0(y_{i,j}^{ecg}) < \varepsilon$).

4. GENERAL BSN HAZARDS AND POTENTIAL CAUSAL FACTORS

With the understanding of $Y_{S\to I}$ and $Y_{S\to H}$, we can now define the hazards related to these interfaces. We assume that the ideal BSN (S_0) always produces safe levels of energy output, is infinitely fast in computation, has perfect event detection capabilities, and exhibits no communication delay. Hence, for a data point y in a stream in the ideal BSN that occurs at time $t_0(y)$, its reported time $t(y)$ and received time $t_r(y)$ are both equal to its occurrence time (i.e. $t_r(y) = t(y) = t_0(y)$). In addition, the value $v(y)$ is the actual digital value (i.e. the ideal BSN is immune to interfering and modifying inputs). Lastly, assume that the inference process I is immediately aware of information from the BSN once this information is made available to it and can act on it from that point in time.

In order to discuss the potential causal factors of the hazards we identify below, one needs an understanding of the internal structure of the BSN. Figure 5 shows a refinement of the black box model in Figure 2 that reveals the general structure of S.

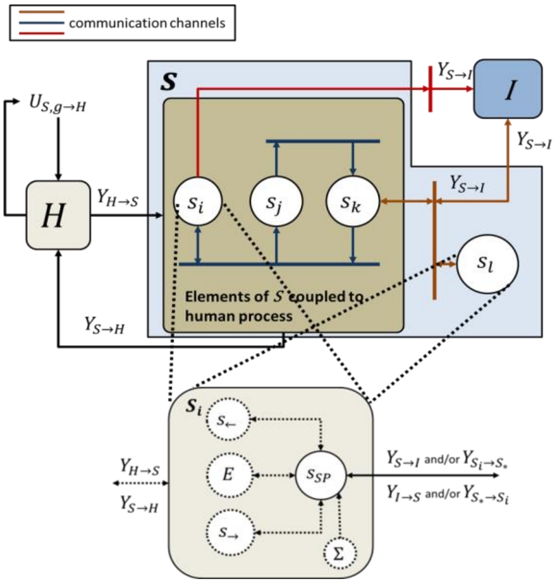

Figure 5: Refined view of BSN operational scenario. Dotted lines indicate context-dependent signals or elements

A BSN is a collection of sensing sub-processes (such as S_i, S_j, S_k, and S_l in Figure 5). Some of these sub-processes may be located on or close to the body and others located away from the body. These sub-processes interact with each other and I through communication channels (either traditional communication network structures or storage media).

Each sensing sub-process $S_{(*)}$ may contain a number of components, and must at least contain some form of computational element (S_{SP}), which is connected to a communication channel to interact with other sub-processes or I. Sub-processes that must interface directly with the human process H (like S_i, S_j, and S_k in Figure 5) would require a transducer ($S_\leftarrow$) for converting physical quantities produced by H to voltages and, if necessary, an "actuator" ($S_\rightarrow$) for producing energy output that affects H to aid in sensing. A sub-process may also have a (potentially) limited energy source (E) if it is mobile. A sub-process may exist within a computational environment (Σ) if it shares computational resources with other processes that are not part of S.

H1 Energy Exposure (EE) Hazard

This hazard represents direct harm to the patient by the BSN.

$$\int_{t_0}^{t} \max\{0, y_{S\to H}^i(t) - y_{S_0\to H}^i(t)\}\, dt < b_{S\to H} \qquad (2)$$

Formally, an energy hazard occurs when constraint (2) is violated for some $y_{S\to H}^i(t) \in Y_{S\to H}$ and$y_{S_0\to H}^i(t) \in Y_{S_0\to H}$. Constraint (2) is defined to account for quick, but intense (very high), as well as slow, but less intense (relatively low), energy output above an accepted threshold. Energy under-exposure is not considered as a hazard because the BSN is not therapeutic and under-exposure would at worst affect the desired input to the BSN, which would most probably cause other hazards.

EE hazards can occur if the BSN intentionally produces energies intended to aid in sensing that are dangerous. For example, a mobile X-ray could produce unsafe levels of radiation. Also, the BSN can potentially produce physical outputs such as heat and electromagnetic radiation (for wireless communication for example) at unsafe levels even when the BSN is coupled to the person in the intended way.

H2 Information Quantity (IQ) Hazard

Such hazards arise when the amount of data points produced by the sensing process deviates from the expectation (as captured by S_0) by a pre-specified threshold. This can be formalized as follows.

Let $N(Y_T)$ denote the number of data points produced by some process T, then IQ hazards occur when the constraint

$$\left|N(Y_{S_0\to I}) - N(Y_{S\to I})\right| \le b_{S\to I|N} \qquad (3)$$

is violated, where $b_{S\to I|N}$ is the threshold for acceptable missing/extra data points. If the thresholds for missing and extra data points are different, then separate bounds can be specified for each.

A primary cause for IQ hazards is the inaccuracy of event detection algorithms used by the sensing process, an example of which is the algorithm(s) to detect activities of the patient. The inaccuracy in such algorithms may produce false events and suppress real events. As a result, the BSN would supply the inference process with spurious or missing information that prevent correct or appropriate medical decisions from being made.

In addition, the instability or corruption of communication channels in BSNs could also cause missing information in the sensing data. For example, the sensed data may get lost during the transmission if the communication buffer is flushed and overwritten by excessive data.

H3 Reported Observation Time (ROT) Hazard

This occurs when the reported observation time of a data point differs unreasonably from its actual occurrence time (as captured by S_0). Here unreasonably implies that for this difference the rest of the loop will probably harm the patient, whereas this would not be the case in S_0.

Formally, let $Y_{S\to I}$ and $Y_{S_0\to I}$ be timed streams (timed-stamped, time-ranged, or relatively timed) sorted in an increasing order of reported observation time (i.e. for data points $y_i(\,i\; = \;1 \ldots n)$ in either $Y_{S\to I}$ or $Y_{S_0\to I}$, $t(y_{i+1})\; > \;t(y_i)$) and with no IQ hazards. Then an ROT hazard occurs as a result of the violation of Constraint (4) for some $y_{i,S\to I} \in Y_{S\to I}$ and $y_{i,S_0\to I} \in Y_{S_0\to I}$.

$$\left|t\left(y_{S_0\to I}\right) - \;t\left(y_{S\to I}\right)\right| \le b_{S\to I|t} \qquad (4)$$

Similar to IQ hazards, separate thresholds can be specified for an observation time unreasonably earlier or later than expected.

The primary cause of ROT hazards are potential flaws or discrepancies in time-stamping and clock synchronization. An out-of-sync component may produce data points with false timestamps. Consider this example: one sensing sub-process is responsible for collecting the ECG waveform, which it then relays to another physically separated sub-process for further processing to identify arrhythmias. If the sensing sub-process does not timestamp data points as they are collected, but leaves it to the arrhythmia detection sub-process, then the claimed time of these data points will be the time at which the arrhythmia detection sub-process first sees them, which might be somehow delayed from the actual collection time (e.g., due to communication delays). Incorrect claimed time, either for all data points or for the first time-stamped one in a processing 'chain', might cause inaccuracy in correlating these data points with the activity events that they are claimed to correspond with.

H4 Value Hazard

These hazards occur when the value of a data point measured and reported by the designed sensing process deviates from its expected value (as captured by S_0) by an intolerable amount. Remember that the word 'value' may refer to information with complicated structures. Also, the deviation could refer to complex semantic features of the reported information like those explored in [1, 11].

Formally, let $Y_{S\to I}$ and $Y_{S_0\to I}$ be BSN output streams sorted in an increasing order of reported observation time, and no ROT or IQ hazards exist. Then a value hazard occurs as a result of the violation of the constraint

$$\left|v\left(y_{i,S_0\to I}\right) - \;v\left(y_{i,S\to I}\right)\right| \le b_{S\to I|v} \qquad (5)$$

for the pair of points $y_{i,S\to I} \in Y_{S\to I}$ and $y_{i,S_0\to I} \in Y_{S_0\to I}$. Similar to the previous hazards, separate bounds can be specified for 'above' and 'below' the expected.

Value hazards are often associated with situations where a patient event or phenomenon does not occur but the BSN claims that it does. The above definition can be extended to cover situations where symbolic values (like those representing events) are used, if some semantics can be established to replace symbolic values with numeric values in a meaningful way.

Value hazards are usually caused by the inability of the sensing process to: 1) compensate for or tolerate modifying and interfering inputs from the human process, or 2) detect that the coupling between itself and the human process is changed to an illegal configuration (e.g., the patient could place an inertial measurement unit (IMU) at a wrong location on the body which, if not detected, could produce erroneous values.). In our example, activities such as running can cause motion artifacts in the ECG. The resulting sweat could either change skin resistance to affect ECG measurements or introduce a level of moisture to affect the readings of other sensors. In addition, sensing sub-processes can potentially produce modifying and interfering inputs for other sub-processes, which might cause value hazards to these sub-processes.

Value hazards can also be caused by limited computational capabilities of the sensing sub-processes. For example, it is technically unfeasible (depending on the configuration) to differentiate between sitting and standing activities from IMU data. If a sub-process with insufficient or limited computational capabilities is used for event detection, even minimal modifying and interfering inputs to a particular signal might cause the sub-process to attribute it to a wrong event by mistake.

H5 Received Time (RT) Hazard

A RT hazard reflects the risk of some data point being reported too late. More formally, a RT hazard arises when there is an unreasonable delay between the time when a data point occurs and the time when it is reported. Let $Y_{S\to I}$ and $Y_{S_0\to I}$ be the actual and ideal output streams, respectively. Assume that data points in both streams are sorted in an increasing order of reporting time (i.e., $y_{k,S\to I}$ is reported earlier than $y_{j,S\to I}$ if $k < j$), and that no IQ hazards exist. Then, a real-time RT hazard arises if $Y_{S\to I}$ is a real-time stream and there exist two data points $y_{i,S\to I} \in Y_{S\to I}$ and $y_{i,S_0\to I} \in Y_{S_0\to I}$ violating the constraint

$$t_r\left(y_{i,S\to I}\right) - \;t_r\left(y_{i,S_0\to I}\right) \le b_{S\to I|t_r} \qquad (6)$$

where $b_{S\to I|t}$ is a pre-specified threshold. If $Y_{S\to I}$ is not expected to be real time, then a RT hazard arises only if Constraint (6) is violated by the last pair of data points $\left(y_{N,S\to I}, y_{N,S_0\to I}\right)$, where N is the number of data points in $Y_{S\to I}$ and $Y_{S_0\to I}$. This is because, in a non-real-time situation, the last received time for a chunk of output data is specified, and only the last pair of data points can possibly violate Constraint (6) due to the fact that all other pairs have been reported before.

The main cause of RT hazards is delays due to computation or communication. An event detection algorithm may require a number of data points to detect an event (with acceptable confidence). However, the overall time cost of collecting and replaying and detecting data points might exceed the pre-specified reporting time bound. Another potential cause of RT hazards is that a sub-process is expected to provide adequate data at a certain rate, but fails to do so.

In our example, the activity detection sub-process may require data from multiple IMUs and may require a number of sample windows in order to make a confident event 'prediction', both of which require excessive amounts of information. As a result, the computational process of activity detection can be fairly slow, and hence cause RT hazards. On the other hand, the IMUs may also delay in presenting their output, especially when such output requires pre-processing.

RT hazards can even occur in non-real-time scenarios. Consider this situation: the amount of information to be processed is large; while the information can only be processed after the last data point is available. Thus, if the time between this last data point is available and the final output must be presented is short, then RT hazards may likely occur.

5. DISCUSSION

Manufacturers must account for a number of variables in order to analyze BSNs properly for their safety properties and must develop their systems to mitigate the potential hazards that can arise from interactions between various components.

There are many ways to use the hazards identified in this paper, and we discuss two of them here. The first is in specification; for particular applications (loops the BSN must operate in), the hazards can be used to specify the requirements that a BSN must satisfy in order to be deemed safe for the particular applications. The second is in communication; manufacturers can use the safety bounds to communicate the capabilities of their BSNs within a particular loop (patient population, intended inference, and control by the clinician). This allows clinicians or technical personnel in the health industry to select BSNs that meet their specification for the particular application.

The bounds can also be used on specific BSN components (sub-processes), when these components are designed by different manufacturers. In this case, each component treats the component that it must report information to as the inference process *I*, and the component receiving the information can specify its expectations using the constraints presented in this paper.

6. RELATED WORK

Previous research on the safety of (non-therapeutic) BSNs mostly focuses on specific safety properties. For example, Banerjee et al. applied formal verification to assess thermal safety associated with the interaction between the BSN and the patient [2]. De Santis et al. used modeling techniques to study the potential risks of ultra-wide band (UWB) radios for the patients [4]. Armenti et al. explored issues in BSNs that can potentially mislead clinicians to make wrong medical decisions [1], but considered such issues more from the data quality perspective. The comparison-based approach in [1] inspired the definition of BSN safety in this paper, and its results provide us the basis for exploring safety issues associated with the interface between the BSN and the clinician.

This paper, on the other hand, generalizes previous research, in the sense that it considers safety of BSNs by understanding the general dynamics of non-therapeutic BSNs and their interaction with other related parties. The generic framework that this work provides is important to regulators (as evidenced by efforts like [5, 6]), especially considering the great possibility of designing arbitrary BSNs.

7. CONCLUSIONS AND FUTURE WORK

Even though the general hazards are few, the potential causal factors are many and may depend on the particular system design. In future work, we will extend the generic model to capture common internal structure and component interaction in BSNs, and undertake a more comprehensive hazard analysis on it.

To aid manufacturers in checking a BSN for the identified hazards, we are developing analysis techniques to evaluate BSNs for the safety properties identified here. Our hope is that, through simulation and further exploration of our generic models, we can gain insight in how to appropriately abstract BSN designs, so that formal verification (such as model checking) techniques can be applied.

We plan to evaluate our framework and analysis techniques with realistic BSN systems, in which safety issues are either well-known or critical to the public health. One example is a diabetes management system.

Our work restricts the BSN's under consideration to non-therapeutic devices. However, we recognize that therapeutic capabilities will be introduced into BSNs or, more generally, body area networks (BANs). In this case, the patient's safety is not only depending on the quality of medical observation ("sensing"), but also on the safety and effectiveness of the medical decision-making and therapy delivery mechanism involved. We believe that our framework can be extended to characterize and analyze therapeutic BSNs.

Two other primary issues of concern for BSNs are security and privacy. We hope to extend our framework to address such issues in the future. Currently, our thoughts are that BSN adversaries could be modeled as sub-processes. Privacy issues could be addressed by preventing these processes from gaining unauthorized access to patient information. Security could be addressed by ensuring that such processes do not disrupt intended functioning of the whole system.

8. ACKNOWLEDGMENTS

This work was funded, in part, by NSF Grants ECCS-0901686, CNS-1035303, and CNS-1240454.

9. REFERENCES

[1] I. Armenti, P. Asare, J. Su, and J. Lach. A methodology for developing quality of information metrics for body sensor design. In Wireless Health, 2012.

[2] A. Banerjee, S. Kandula, T. Mukherjee, and S. K. S. Gupta. Band-aide: A tool for cyber-physical oriented analysis and design of body area networks and devices. ACM Trans. Embed. Comput. Syst., 11(S2):49:1–49:29, Aug 2012.

[3] B. Calhoun, J. Lach, J. Stankovic, D. Wentzloff, K. Whitehouse, A. Barth, J. Brown, Q. Li, S. Oh, N. Roberts, and Y. Zhang. Body sensor networks: A holistic approach from silicon to users. Proceedings of the IEEE, 100(1):91 –106, jan. 2012.

[4] V. De Santis, M. Feliziani, and F. Maradei. Safety assessment of uwb radio systems for body area network by the method. Magnetics, IEEE Transactions on, 46(8):3245–3248, 2010.

[5] J. Hatcliff, A. King, I. Lee, A. Macdonald, A. Fernando, M. Robkin, E. Vasserman, S. Weininger, and J. Goldman. Rationale and architecture principles for medical application platforms. In Cyber-Physical Systems (ICCPS), 2012 IEEE/ACM Third International Conference on, pages 3–12, 2012.

[6] B. Kim, A. Ayoub, O. Sokolsky, I. Lee, P. Jones, Y. Zhang, and R. Jetley. Safety-assured development of the gpca infusion pump software. In proceedings of EMSOFT: 2011 international conference on embedded software, pages 155-164, Oct. 2011.

[7] N. Leveson. Engineering A Safer World: Systems Thinking Applied to Safety. MIT Press, Cambridge, MA, 2011.

[8] A. Pantelopoulos and N. Bourbakis. A survey on wearable sensor-based systems for health monitoring and prognosis. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 40(1):1 –12, Jan. 2010.

[9] J. G. Webster, editor. Medical Instrumentation: Application and Design. John Wiley & Sons, Inc, Hoboken, NJ, 4e, 2010.

[10] W. Wu, M. Batalin, L. Au, A. Bui, and W. Kaiser. Context-aware sensing of physiological signals. In Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE, pages 5271–5275, 2007.

[11] Y. Zigel, A. Cohen, and A. Katz. The weighted diagnostic distortion (WDD) measure for ECG signal compression. Biomedical Engineering, IEEE Transactions on, 47(11):1422–1430, Nov. 2000.