Verification of Physiological Data Collection

Colin Elkin, MS and Vijay Devabhaktuni, PhD
Department of Electrical Engineering and Computer Science, The University of Toledo, Toledo OH USA

Introduction

Many emerging facets of data science are based on the collection and analysis of physiological data, some from commercial wearable sensors (e.g. Fit Bit and Microsoft Band) and others from elaborate methods of obtaining sensor data (e.g. wireless body area sensor networks, or WBASNs [1]). In the realm of wearable technology, security concerns are on par with those of typical mobile devices. For wearables, security is simply whether a user consents to having his or her data monitored [2], but as software for commercialized sensors become more flexible, this security policy will require significant expansion.

Wireless sensor networks, including WBASNs, are a recurring research pursuit in regards to security and thus follow many conventional security principles. At some instances, the concept of localization becomes important, as the overall function and value of some wireless sensor networks (or WSNs) can be dependent upon the accuracy and speed of a localization technique [3]. Consequently, availability and integrity become the security principles of utmost importance. A body area sensor network is a WSN that collects physiological data for which reliability is the most crucial security concern. This is because (i) the system is wireless and (ii) the confidentiality-dependent field of healthcare is a frequent application for this type of network [1].

The goal of this research study is to identify the security and verification challenges associated with data assurance in these varied techniques. While the scope of this research focuses primarily on WBASNs, the solution is designed with enough flexibility that wearable sensors can incorporate an identical security policy.

Challenges

In general, security is already a primary concern and can be derived from four major factors: authentication, availability, integrity, and secrecy [4]. This research commences by identifying the challenges and involved in each of the following four security factors, which is also summarized in Table 1.

1. Authentication is the process of identifying the appropriate roles for the appropriate user, i.e. privileged user access [5]. Thus, an overarching challenge is to clearly define the roles and responsibilities for each user. An extension of this challenge is to acquire as much information from each user as possible in order to accurately formulate each role.

2. Availability is the ability of the data to be reliably accessed at any appropriate time. This security policy is mostly prevalent on the hardware side of the data acquisition. Thus, the greatest challenge is preserving the functionality of a sensor network in the event of a failing node [3], especially an anchor node, which is critical for data processing [4].

3. Confidentiality is the practice of concealing data to only the users who are explicitly authorized to view, modify, or manipulate. Challenges thereof are similar to those of authentication and integrity.

4. Integrity refers to the reliability and trustworthiness of the data. This factor has perhaps the greatest potential for challenges, including malicious nodes, injection of false data, and unintentional harm due to noise.

Required Components

In order to formulate a secure and comprehensive security mechanism for this application, the following questions must be posed:

- How should the data be stored?
- Where should the data be stored?
- Who gets to read the data, write the data, and/or execute the data?
- How can the data be preserved and processed accurately and securely?

Results and Discussion

When typical input and output of data are considered, simple textual file formats, such as comma separated value (CSV) are typically desired. This research selects CSV as the gold standard for file input and output but in the process converts each data file to and from the encrypted PEM format. For the purposes of both redundancy and flexibility, the location of the data will be stored of up to three simultaneous locations: the anchor nodes of a WBASN, an embedded computing device known as the data fusion center [4], and cloud storage to be accessed by the healthcare workers or data scientists involved in executing the data. The role of the latter location will vary based on three proposed levels of patient monitoring needs, as given in Chart 1.

Determination of who is authorized to read, write, or use the data can be simplified to three distinct user classes, as given in Chart 2. This is implemented with a Unix-style separation principle of read-write-execute [5]. As the chart indicates, all authorized user groups have access to reading the data, but only the patient creating the data has write access. The other two groups are able to manipulate and execute the data but in different capacities. A healthcare worker is expected to run the data through pre-programmed applications (e.g. to determine risk of heart disease or to classify blood pressure), while a data scientist can experiment with the data under a greater amount of creativity, creating novel and meaningful algorithms to engage in cutting edge human effectiveness research. The final component of ensuring accurate data mostly lies beyond the scope of this research study but contains guidelines and recommendations for data scientists (e.g. avoiding premature convergence through careful selection of machine learning techniques) and network engineers (e.g. minimizing noise in a WBASN).

Chart 1. Depiction of monitoring levels and corresponding data storage locations.

Intentional	Unintentional
Unauthorized users	Noise in sensor network
Failed/malicious nodes	Machine learning errors
Injection of false data	

Table 1. Types of potential security threats to physiological data.

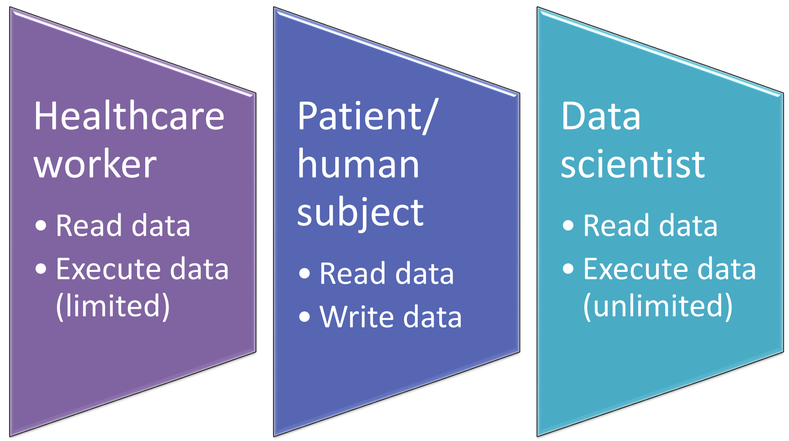

Chart 2. Classes of users and access rights thereof.

Conclusion

A flexible security solution for the preservation and manipulation of physiological data is presented. Current studies aim to bridge the two research fields of wireless sensor networks and human effectiveness by utilizing the realm of information security as well as the application of wireless body area sensor networks. Also anticipated is the potential to improve medical and psychological applications in which human-centric data analysis is needed.

Acknowledgements

This research is supported by the Ohio Federal Research Network project entitled “Sliding-Scale Autonomy through Physiological Rhythm Evaluations (SAPHYRE)” and by the AFRL/DAGSI Ohio Student-Faculty Research Fellowship Program.

Contact

Colin Elkin
The University of Toledo
Colin.Elkin@rockets.utoledo.edu
(419) 530-8140

References

1. Miao, F., Bao, S., and Li, Y. (2012). Physiological Signal Based Biometrics for Securing Body Sensor Network. In Jucheng Yang (Ed.), New trends and Developments in Biometrics. DOI: 10.5772/51856.
2. Oliveira, E. A., Kirley, M., Fonseca, J. C. B., and Gama, K. (2015). Device Nimbus: An Intelligent Middleware for Smarter Services for Health and Fitness. International Journal of Distributed Sensor Networks. DOI: 10.1155/2015/454626.
3. Elkin, C. (2015). Development of Novel Computational Algorithms for Localization in Wireless Sensor Networks through Incorporation of Dempster-Shafer Evidence Theory (Master’s thesis). Retrieved from https://etd.ohiolink.edu/.
4. Padmavathi, G. and Shanmugapriya, D. (2009). A Survey of Attacks, Security Mechanisms and Challenges in Wireless Sensor Networks. International Journal of Computer Science and Information Security. DOI: 0909.0576
5. Anderson, R. (2008). Security Engineering, 2nd Edition. Jersey City, NJ: Wiley.