

Redes Inalámbricas y Técnicas de Minería Web Aplicadas a un Modelo de Telemedicina

Marco Javier Suarez Baron¹ and Carlos Andrés Cifuentes García¹,

¹ Escuela Colombiana de Carreras Industriales ECCI- Facultad de Ingeniería, Carrera 19 No. 49-20, Teléfono 57 2715353 Ext. 196, Bogotá D.C, Colombia
marcojaviersuarezbaron@gmail.com

Abstract. This paper presents the integration of two technology tendencies: telemedicine and web mining with the aim of propose a model of real-time diagnosis, along with added services for consultation and monitoring to assist medical staff to use new tools ICT for disease diagnosis. It has defined four stages: (A) Design of hardware, including communications and sensor integration of the patient. (B) Implementation of the Zigbee network for monitoring biomedical sensors. (C) Classification of information Web on medical domain, (D) Development of an algorithm for extraction and information integration. The results are presented as biomedical signal transmission to Internet server in real time, and extracted classified information internet medical domain, in order to generate diagnoses and treatment based on patient monitoring.

Keywords: Zigbee Networks, Web Mining, semantic web, bioinstrumentation.

1 Introduction

La integración de sistemas de monitoreo de la salud con sistemas de telemedicina permitirá la detección temprana de condiciones anormales y la prevención de graves consecuencias [1,2]. El monitoreo continuo como parte del procedimiento de diagnóstico puede beneficiar a muchos pacientes, que se encuentren en observación por condiciones crónicas o durante la supervisión de una recuperación desde un evento agudo hasta el procedimiento quirúrgico [3].

Introducir sistemas de redes inalámbricas para lectura de datos provenientes de sensores biomédicos permite gran flexibilidad tanto para el paciente como para el cuerpo médico. El uso de tecnologías de redes inalámbricas en estos casos, asegura un manejo completamente digital de los datos, con lo cual se mejoraría la gestión de la historia clínica y algunas aplicaciones logísticas. Adicionalmente los datos se pueden almacenar para análisis posteriores y pueden ser consultados en cualquier momento, lo que se puede hacer desde cualquier parte dentro del hospital o de manera remota a través de computadores, PDA's o utilizando telefonía celular. De igual manera el cuerpo médico puede monitorear pacientes dentro de la red sin tener que ir hasta su

posición [4].

El proyecto propuesto integra dos tendencias tecnológicas: telemedicina y minería web, con el objetivo de proponer un modelo de diagnóstico en tiempo real, junto con servicios agregados para consulta y seguimiento. Este diagnóstico se apoya en conocimiento obtenido en un proceso de minería web, tomando como base lecturas de medidas de parámetros biomédicos, siendo adquiridos en línea con el paciente sin alterar su desempeño diario.

2 Metodología

La demanda de pacientes a ser monitoreados por sistemas de salud, está dada por todos los que padecen enfermedades que requieren diagnóstico temprano y tratamientos pertinentes, dentro de esto los más afectados son los que padecen enfermedades cardiovasculares. Entre estas se encuentran las anomalías en el ritmo cardíaco (Arritmias), como las producidas por fibrilación auricular que son encontradas en la clínica práctica, ocurriendo en alrededor del 4% de la población sobre los 60 años, incrementando en un 9% en los mayores a 80 años [5]. Síntomas tempranos de fibrilación auricular incluyen fatiga y palpitaciones, y con frecuencia el paciente tiene que visitar el servicio médico. Las redes inalámbricas y la utilización de sensores biomédicos ofrecen la posibilidad de diagnosticar arritmias cardíacas tempranas en la población mayor. De igual manera permiten monitorear el progreso de las enfermedades y la respuesta a los tratamientos iniciados [6].

Basados en esta problemática este trabajo propone una arquitectura para el monitoreo continuo de pacientes basado en el soporte de la minería web (fig. 1), la cual se compone de tres módulos básicos: (1) Red de área corporal para lectura de parámetros biológicos, está compuesta por una serie de sensores biomédicos que registran diferentes parámetros como: electrocardiografía, pulsioximetría, presión arterial, variables biomecánicas entre otros. (2) Modelo de minería web e integración de información, tiene como propósito extraer la información de conocimiento médica a partir de una colección de documentos web previamente clasificados, estructurando los datos obtenidos en un único formato o estructura RDF/XML [7]. (3) Integración de pacientes a través de una red Zigbee, este componente permite interconectar las lecturas de parámetros biológicos de pacientes al modelo de minería web e integración de información. La conectividad entre la red Zigbee de pacientes y la base de conocimiento en salud se realiza a través de un dispositivo que provee una conexión TCP-IP a internet.

Para el año 2010 las aplicaciones de Web semántica han cobrado un marcado interés, esencialmente en el campo biomédico y de la salud. El W3C (World Wide Web Consortium) ha puesto en marcha un Grupo de Interés con el objetivo de establecer una relación entre la medicina y expertos en Web Semántica en un intento por potenciar la colaboración, la investigación y el desarrollo, y la adopción de innovación en el ámbito de la salud y de las ciencias de la vida. El Grupo de Interés de Web Semántica para la Salud y las Ciencias de la Vida (HCLSIG), el primero de estas características dentro del Consorcio, que aplica estándares basados en Web

Semántica a servicios concretos definidos por una comunidad de usuarios.

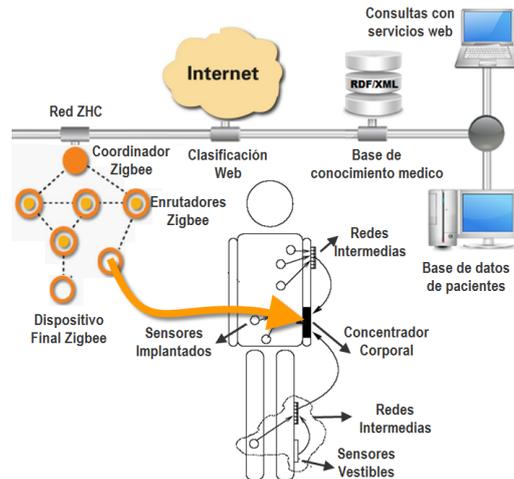


Fig. 1. Arquitectura propuesta para la integración de redes Zigbee de sensores biomédicos y minería web para la aplicación en telemedicina.

Nuestro proyecto permite cumplir claramente las recomendaciones emitidas por el grupo (HCLSIG); La aplicación Semántica aquí generada y sus datos vinculados se han convertido en una opción clara para otros proyectos que tienen un especial interés en compartir la interoperabilidad y datos a través de redes de sensores biomédicos con acceso a la web. Por ende se incluye una participación activa en la Comunidad Europea de Iniciativas Innovadoras Medicina (IMI), en la cual los proyectos que quieran compartir los resultados y la información a través de varios dominios puedan: Incluir acceso al descubrimiento de nuevos medicamentos, acceso a expedientes electrónicos de pacientes junto a los ensayos clínicos con el banco de órganos y tejidos.

3 Materiales y métodos

La integración de redes zigbee de sensores biomédicos y minería web para la mejora de sistemas de telemedicina es un proyecto en desarrollo, donde se han definido cuatro etapas de desarrollo: (A) Diseño de Hardware, comprende la construcción de un dispositivo IEEE 802.15.4/Zigbee, la integración de sensores y el desarrollo de un dispositivo de enlace Zigbee/Internet. (B) Implementación de la Red Zigbee basada en el ZHC (ZigBee Health Care public application profile) para monitoreo de sensores biomédicos. (C) Clasificación de información web del dominio médico, (D) Desarrollo de una algoritmo de extracción e integración de información, generando

una base de conocimiento para realizar análisis fundamentado en datos biomédicos monitoreados en tiempo real.

3.1 Desarrollo de Hardware

Basado en el desarrollo y evaluación de dispositivos IEEE 802.15.4/ZigBee para aplicación en una red de instrumentación biomédica presentado en [8]. El dispositivo permite la comunicación inalámbrica, y cuenta con sensores de aceleración de tres ejes, presión y temperatura, y un puerto para conexión a módulos de señal externos. En la Fig. 2b se observa el dispositivo que permite interconectar un dispositivo Zigbee con internet, para así, transmitir las diferentes señales biomédicas y realizar la conexión TCP/IP a internet por medio de un modem GSM/GPRS GR48 de la compañía Sony Ericsson.

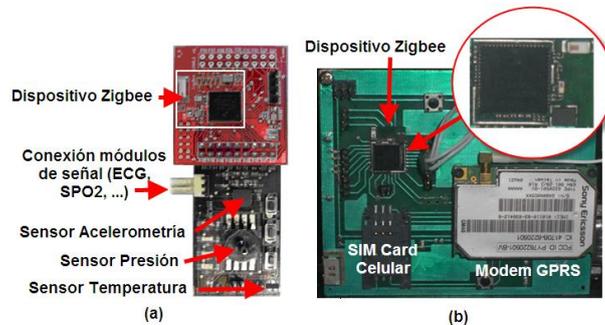


Fig. 2. (a) Módulo de nodo de red biomédico. (b) Dispositivo de enlace Zigbee y GPRS/Internet [8].

3.2 Implementación de una Red Zigbee basada en ZHC

Basado en los dispositivos desarrollados y la integración de sensores, se implementó una red ZHC, donde se desarrolló un dispositivo monitor de ECG (ID 0x1006) y un monitor de pulsioximetría (ID 0x1004) en la categoría de tratamiento de enfermedades, por otro lado, en la categoría de monitoreo del estado físico personal, se implementó un monitor de actividad física (ID 0x102b), el cual está orientado a medidas de variables cinemáticas como aceleración, velocidad, posición y orientación [9].

3.3 Clasificación de información web del dominio medico

La red Zigbee toma las señales biomédicas en tiempo real de los pacientes, estos parámetros son enviados y almacenados en internet, siendo esto una de dos entradas a los sistemas de minería web, la otra entrada es información web del dominio médico

que sea compatible con los parámetros monitoreados, esto se refiere a páginas y documentos en la web que profundicen en la caracterización de las señales transmitidas.

La clasificación web selecciona documentos previamente filtrados y analizados con información no estructurada de dominio médico. Este repositorio de páginas está alimentado por documentos de tipo PDF con información presentada en forma de tablas. A través de un proceso de extracción e integración de información sobre este repositorio se obtiene un conjunto de documentos RDF/XML; que para este trabajo se denomina base de conocimiento RDF/XML, que sirve de fuente para descubrir conocimiento en salud y que soportará la toma de decisiones, para generar información relevante como diagnóstico y tratamientos a pacientes monitoreados en la red Zigbee.

En la Fig. 1, se muestra el modelo general de descubrimiento de conocimiento, que permite interacción con la red de monitoreo a pacientes, con dos objetos finales: La base de datos de pacientes y los servicios web semánticos de consulta y seguimiento.

Los procesos de extracción e integración de información son algoritmos desarrollados y probados a través de un sistema mediador, estos algoritmos fueron implementados a partir de técnicas de web semántica y minería web, y son detallados en las secciones siguientes. Es preciso indicar que la fuente de información es totalmente heterogénea es decir documentos PDF mapeados que para su análisis textual fue necesario convertirlos al tipo HTML.

3.4 Desarrollo de un algoritmo de extracción e integración de información

En esta sección se describe procedimentalmente los modelos de extracción e integración semántica de información [7]. El análisis textual de los documentos PDF de dominio salud se apoya en un proceso de minería web que utiliza el algoritmo de extracción HTML, para ello se desarrolló un sistema mediador que detecta la presencia de tablas en los PDF y apoyado en un conjunto de reglas deductivas convierte este formato a un tipo de archivo textual de tipo HTML [10].

Una vez generado el documento HTML, el algoritmo de extracción procede a analizar las etiquetas obtenidas y selecciona de manera jerárquica la información relevante desde los párrafos hasta los títulos de una tabla.

Una jerarquía sintáctica es el elemento funcional dado para extraer datos e información relevante de la fuente, se inicia la extracción con elementos particulares de información desde las celdas (<td>), desde los textos (<p>), desde las filas (<tr>), hasta llegar a las etiquetas raíz (<caption>), (<colspan>), (<rowspan>) y finalmente la etiqueta (<title>).

El algoritmo de extracción permite identificar y extraer estos datos etiquetados realizando un recorrido secuencial al texto HTML, utilizando la jerarquía sintáctica, como se muestra en la Fig. 3a. La salida de este proceso de extracción corresponde a documentos de tipo XML preliminares (Fig. 3b), que una vez validados deben ser indexados por una ontología en el dominio específico de la señal biomédica, a esta fase se denomina enriquecimiento semántico, aquí se procederá a unificar mediante

un modelo de indexación semántica la información contenida en las estructuras detectadas, que se encuentran en los diferentes documentos web analizados, el formato apropiado para estructurar la información será XML por considerarse un lenguaje de marcado extensible universal.

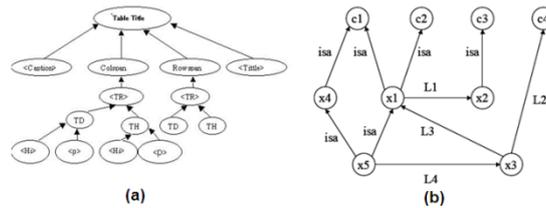


Fig. 3. (a) Árbol sintáctico generado a partir del Documento HTML, (b) jerarquía semántica basada en XML.

4 Resultados

En la Fig. 4 Se observa el montaje de una red ZHC basada en tres dispositivos monitores de actividad física y un monitor ECG, para lo cual se usaron cuatro nodos de red definidos como ZED (Dispositivo Final Zigbee) y un ZC (Coordinador Zigbee), los ZED se distribuyeron en la cadera, rodilla y tobillo para acelerometría, y en el pecho (electrodos/cable ECG) del paciente. El ZC está dentro del dispositivo de enlace Zigbee/Internet, con la función de enviar las señales a un servidor en internet.

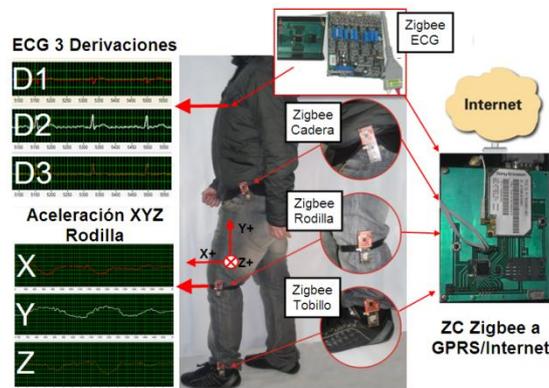


Fig. 4. Montaje de la red ZHC con conexión a internet.

En la Fig. 5 se visualiza la captura y análisis de un documento PDF, definiendo parámetros cardiovasculares, clasificado a través de un sistema mediador. Su análisis y posterior mapeo permiten la conversión a formato HTML.

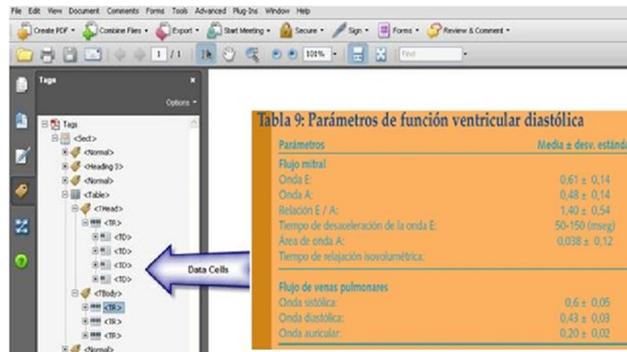


Fig. 5. Sistema crawler para extracción de información. RDF/XML.

Con la ayuda de editores XML como SPY de Altova se sustentara la validez de los documentos obtenidos. A continuación se visualiza el documento RDF/XML final luego del proceso de enriquecimiento semántico aplicado al documento anterior.

```

<table>
  <title>
    <table-title>función ventricular diastólica</table-title>
    <Columns Title> Parámetros</Columns Title>
    <Columns Title> Media +- desv. estándar</Columns Title>
  </title>
  <Number- Columns> 2</ Number- Columns>
  <Content>
    <body site>
      <neoplasm>Flujo mitral</neoplasm>
      <vascular>0.61+-0.14</vascular>
    </body site>
    <body site>
      <neoplasm>Flujo mitral</neoplasm>
      <vascular>0.61+-0.14</vascular>
    </body site>
  </Content>
</table>

```

5 Conclusiones y Trabajo Futuro

Con los dispositivos Zigbee y GPRS/Internet desarrollados se pretende brindar herramientas de bajo costo para monitoreo continuo de pacientes, donde se puede integrar dispositivos portables certificados y tener rápida inclusión en el ambiente clínico.

En el desarrollo informático se ha consolidado en un sistema que extrae información clasificada de internet [11], abriendo un amplio panorama en el desarrollo de algoritmos de específicos; para la obtención de información basada en señales obtenidas en tiempo real en el ambiente del paciente.

Actualmente se está trabajando en un sistema prototipo que genere diagnóstico y tratamientos obtenidos de la web basado en señales de electrocardiografía obtenidas en tiempo real. Por otro lado, basado en medidas de variables biomecánicas para el análisis de posturas y entrenamiento físico. Para esta nueva etapa se pretende trabajar con pacientes de la Clínica Universitaria Teletón.

References

1. Istepanian, R. S. H., Jovanov, E., Zhang, Y. T.: "Guest Editorial Introduction to the Special Section on M-Health: Beyond Seamless Mobility and Global Wireless Health-Care Connectivity", *IEEE Transactions on Information Technology in Biomedicine*, pp. 405-414 (2004)
2. Wearable Technology. *Special Issue of the IEEE Engineering in Medicine and Biology Magazine*, pp 22 (2003)
3. Jovanov, E., Milenkovic, A., Otto, C. P. C. de Groen. (2005, January) "A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation", *Journal of NeuroEngineering and Rehabilitation* [Online], 2(6). pp. 1-3. Available: <http://www.jneuroengrehab.com/content/2/1/6#B1>
4. Hansen, M. S.: "Practical Evaluation of IEEE 802.15.4/ ZigBee Medical Sensor Networks," Norwegian University of Science and Technology, Trondheim, Norway (2006)
5. Go, S., Hylek, E. M., Phillips, K. A., Chang, Y., Henault, L. E., Selby, J. V.: "Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial fibrillation (ATRIA) study". *The Journal of the American Medical Association*, vol. 285, no. 18, pp.2370-2375, (2001)
6. Azis, O., Lo, B., Darzi, A., Yang, G.: *Body Sensor Networks*, London:Springer, ch. 1.
7. M. J. Suárez, 2005. An Approach to Semantic Indexing and Information Retrieval. In Proceedings CIIDET, IEEE México, pp. 34-350 (2007)
8. Cifuentes, C. A., Gentiletti, G. G., Suarez, M. J., Rodríguez, L. E.: Development of a Zigbee platform for bioinstrumentation, 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Buenos Aires, Aug. 31 2010-Sept. 4 2010, pp.390-393 (2010)
9. ZigBee Alliance, ZigBee Health Care™ Profile Specification, ZigBee Profile: 0x0108, Revision 15, Version 1.0.
<http://www.zigbee.org/Markets/ZigBeeHealthCare/download.aspx>, (2010)
10. Lim, S. J.: *Extracting Information from semistructured sources*, pp. 71-80 (2002)
11. Suarez, M. J.: "Modelo inteligente de extracción de información a partir de documentos web". *Revista energía y computación UNIVALLE*, pp 32-40 (2005)
12. <http://www.w3c.es/consorcio/historia>