

DETECCIÓN DE CONTAMINANTES MEDIANTE UNA RED INALÁMBRICA DE SENSORES Y RECONOCIMIENTO DE PATRONES

Patricia Arroyo, Jesús Lozano, José Ignacio Suárez, José Luis Herrero, Pablo Carmona
Escuela de Ingenierías Industriales, Universidad de Extremadura, Av. Elvas s/n, Badajoz
email: jesuslozano@unex.es

Resumen

En este trabajo se presenta una red de sensores inalámbrica de bajo consumo y bajo coste para la detección de compuestos contaminantes disueltos en agua. Esta red está compuesta por nodos con varios sensores de gases encargados de recibir la información y de forma inalámbrica, a través del protocolo ZigBee, enviarla a una puerta de enlace conectada a Internet. Una nube de datos se utiliza para almacenar los datos de cada nodo y realizar el pre-procesamiento de los mismos. Para la realización del análisis o procesado de los datos obtenidos, se ha realizado una reducción de la dimensionalidad mediante análisis de componentes principales (PCA) y una clasificación mediante redes neuronales artificiales (RNA) de las clases previamente aprendidas.

Palabras Clave: red inalámbrica de sensores, sensores de gases, Zig Bee, reconocimiento de patrones.

1 INTRODUCCIÓN

La contaminación ambiental es un problema de primer orden tanto en los países desarrollados como en los países emergentes y en vías de desarrollo. Dicha contaminación tiene efectos nocivos para la salud de los seres vivos y recientemente la Agencia Internacional para la Investigación sobre el Cáncer (IARC) ha notificado sobre las consecuencias carcinógenas que tiene la contaminación atmosférica en los seres humanos [17]. Como consecuencia, la preocupación sobre este asunto está aumentando a nivel mundial, dando lugar a que las normativas de calidad ambiental establecidas por los gobiernos de todo el mundo tiendan a ser cada vez más exigentes. En este contexto, la detección y el control de los contaminantes que originan los problemas ambientales, llevada a cabo de un modo fiable, compacto y de bajo costo es aún una meta por conseguir.

El trabajo en el que se enmarca este artículo, se centra, principalmente, en la realización de medidas con sistemas diseñados para la detección de gases objetivo, la caracterización de la respuesta de los sensores, la implementación de técnicas basadas en inteligencia artificial para el procesado de los datos de los sensores, y la realización de redes inalámbricas de sensores (basado en motas y en nodos sensores para dispositivos inteligentes) para la realización de medidas de campo.

En este artículo se presenta una red inalámbrica de sensores de bajo coste, la cual se ha utilizado para la detección en tiempo real de componentes contaminantes disueltos en agua. Los datos medidos por la red de sensores inalámbrica son almacenados en una nube de datos. De esta forma se pueden descargar las mediciones realizadas través de cualquier dispositivo con acceso a internet para su posterior procesado.

Este trabajo se estructura de la siguiente forma: en el capítulo 2, se realiza una descripción del dispositivo desarrollado mostrando sus principales características constructivas y de funcionamiento, y desarrollando el método de clasificación empleado. A continuación, en la sección 3, se muestran las medidas realizadas y los resultados obtenidos y, por último, en el capítulo 4 se detallan las conclusiones obtenidas del trabajo.

1.1 REDES INALÁMBRICA DE SENSORES

Las redes inalámbricas de sensores [1] [24] (WSNs, Wireless Sensor Networks), gracias al avance en la tecnología de Sistemas Micro-Electro-Mecánicos (MEMS) que han hecho más fácil el desarrollo de sensores inteligentes, se han convertido en el foco de atención mundial en los últimos años.

Los nodos sensores inteligentes son dispositivos de bajo consumo equipados con uno o más sensores, un procesador, memoria, una fuente de alimentación, una radio, y un actuador. Gran variedad de sensores mecánicos, térmicos, biológicos, químicos, ópticos y magnéticos pueden estar conectados al nodo sensor para medir las propiedades de entorno. Además, constan de una radio inalámbrica con el fin de enviar

los datos a una estación base. Esto es debido a que los nodos sensores tienen memoria limitada y a que, en numerosas ocasiones, son desarrollados para colocarlos en localizaciones de difícil acceso.

En la actualidad, las WSNs se pueden clasificar en cinco tipos: terrestres [22], consistentes en una cantidad de entre cientos a miles de nodos sensores inalámbricos desplegados en un área determinada; subterráneas [2], basados en nodos sensores enterrados bajo tierra y utilizados para monitorizar las condiciones del subsuelo; subacuáticas [6], que consisten en nodos sensores y vehículos distribuidos bajo el agua; multi-media [1] [9], que se propusieron para habilitar la monitorización y el rastreo de eventos en formas multimedia tales como video, audio e imagen; y móviles, incluidas en monitorización del entorno, rastreo de objetivos, búsqueda y rescate, y monitorización a tiempo real de materiales peligrosos.

Las WSNs son útiles para muchas aplicaciones en escenarios como seguimiento y vigilancia de objetivos militares [21], socorro en desastres naturales [20], monitoreo de salud en biomedicina [11] [19], exploración de ambientes peligrosos y sensores sísmicos [23].

Existen además numerosas aplicaciones de redes de sensores inalámbricas con uso específico de sensores de gases, como pueden ser aplicaciones para la detección de fuga de gases [7], medición de calidad de aire interior [5] o monitorización y prevención de posibles explosiones de gas y carbón [8].

Sin embargo, las WSNs tienen limitaciones en términos de memoria, energía, computación, comunicación y escalabilidad. De forma paralela, la computación en la nube se está convirtiendo en una tecnología prometedora para proporcionar una pila masiva de computación, almacenamiento y servicios de software.

1.2 "SENSOR CLOUD"

Muchas aplicaciones que adquieren y procesan información del mundo físico son posibles gracias a las redes inalámbricas de sensores. De igual forma, la posibilidad de compartir recursos distribuidos se lleva a cabo a través de uso de nubes de datos. La computación en la nube es un término utilizada para describir tanto una plataforma como un tipo de aplicación. Una plataforma de computación en la nube provisiona, configura, reconfigura servidores según la necesidad, lo cual constituye una alternativa a tener servidores locales. Además una nube puede describir una aplicación que se extiende para ser accesible a través de internet. La computación en la nube son servidores desde Internet encargados de

atender las peticiones en cualquier momento. Se puede tener acceso a su información o servicio, mediante una conexión a internet desde cualquier dispositivo móvil o fijo ubicado en cualquier lugar. La extensión del paradigma de las nubes de datos al compartir recursos de sensores en WSNs resulta en una prometedora tecnología llamada "Sensor Cloud" [3].

Aunque se trata de una novedosa infraestructura, ya han surgido aplicaciones en distintos ámbitos. Por ejemplo, se ha diseñado un sistema de monitoreo estructural de edificios con este tipo de tecnología para la mitigación del riesgo sísmico [15]. En el mundo de la salud también han surgido varias aplicaciones, como sistemas para la recogida y el acceso a gran cantidad de datos generados por las redes de sensores médicos [13] o sistemas de monitorización de sueño para estudiar los desórdenes del mismo [12]. También se ha diseñado un sistema de monitorización y almacenamiento de datos referentes a tráfico de vehículos y parámetros particulares de cada vehículo [4].

2 DESCRIPCIÓN DEL DISPOSITIVO.

El principal objetivo de este trabajo es el desarrollo y construcción de una red inalámbrica de sensores para la monitorización de la calidad de aire. El diagrama de bloques del sistema es mostrado en la Figura 1. La red consiste en unos nodos sensores responsables de recibir la información y enviarla, siguiendo el protocolo ZigBee, a la puerta de enlace (Connect Port). Éste se encuentra conectado a internet, vía Ethernet, y se encarga de remitir los datos a la nube "Device Cloud" de Digi donde serán almacenados en formato ".xml".

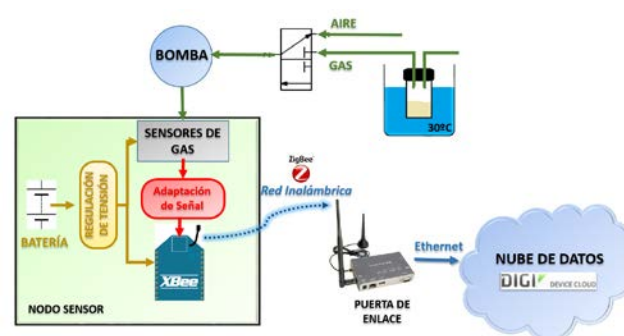


Figura 1: Diagrama de bloques del prototipo desarrollado.

2.1 NODOS

Los nodos sensores son unos dispositivos o sistemas de bajo coste, tamaño reducido y bajo consumo, con la capacidad de recolectar información referente a la

calidad de aire usando diferente tipos de sensores de gas. El nodo diseñado se muestra en la figura 2 y está compuesto por los siguientes dispositivos:

- Módulo XBee y XBee PRO: son los responsables de la comunicación con el nodo coordinador o la puerta de enlace. Se tratan de módulos de radiofrecuencia, fabricados por Maxstream, que trabajan dentro de la banda de 2,4GHz, con un rango de alcance de 100m (XBee) o 1500 m (XBee PRO).
- Circuito de alimentación: está principalmente compuesto por una batería de litio que suministra una tensión de 3.7V y convertidores DC-DC.
- Sensores de Gas: hasta 4 sensores resistivos basados en sistemas microelectromecánicos (MEMS). En concreto los sensores utilizados han sido MiCS-5135, MiCS-5521 y MiCS-5914 SGX (fabricados por Sensortech).
- Circuitos de acondicionamiento de señal: se encargan de adaptar las señales al rango necesario para el correcto funcionamiento de los módulos de comunicación.

Para la realización de las pruebas, los nodos también incluyen una bomba y una electroválvula de tres vías para conducir las muestras de aire a medir o el gas de referencia a la célula de sensores.

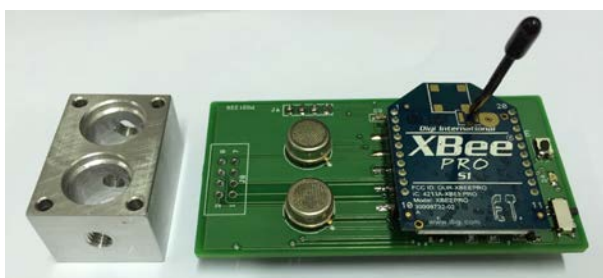


Figura 2: Nodo sensor desarrollado.

2.2 RED INALÁMBRICA

La red inalámbrica responde a las especificaciones para el conjunto de protocolos de comunicación de alto nivel de ZigBee. Éste está basado en la norma IEEE 802.15.4. Se dirige a aplicaciones que requieren comunicaciones seguras con baja tasa de envío de datos y la maximización de la vida útil de sus baterías. ZigBee se diferencia de otras tecnologías de características tales como bajo consumo de energía, su topología de red en malla y la capacidad de integración. En la red creada en este trabajo, se hace uso del modo de 16 bits de direccionamiento debido a la existencia de un bajo número de nodos.

2.3 PUERTA DE ENLACE Y ALMACENAMIENTO EN NUBE DE DATOS

La información recogida por los nodos sensores es enviada a la puerta de enlace: “ConnectPort X4” de Digi. Esta puerta de enlace ofrece una gran variedad de opciones de interfaz LAN/WAN para la creación de redes de extremo a extremo con dispositivos remotos, incluidos dispositivos XBee. Por otra parte, a través de un entorno de desarrollo Python existe la posibilidad de la ejecución local de aplicaciones personalizadas.

En este estudio, se ha desarrollado una aplicación Python que permite separar y pre-procesar la información recibida por los nodos para después enviarla y almacenarla en la nube de datos.

En concreto, la nube utilizada ha sido la suministrada por el fabricante: “Digi Device Cloud”. Ésta se trata de una plataforma máquina a máquina que permite la gestión de los dispositivos a través del acceso y control de las puertas de enlace y enrutadores de la red (Figura 3). Los usuarios pueden controlar los dispositivos a través del envío de comandos o mediante tareas programadas. Por otro lado, los dispositivos pueden cargar datos a la secuencia de datos de la “Device Cloud”.

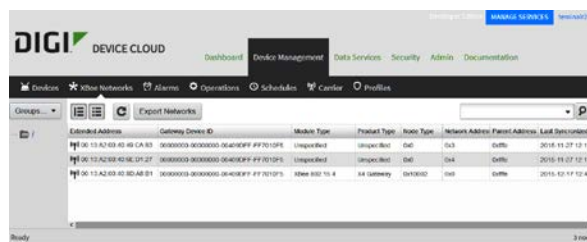


Figura 3: Captura de “Digi Device Cloud”. Pestaña de gestión de dispositivos donde los dispositivos conectados a la nube y sus características pueden ser consultados.

2.4 EXTRACCIÓN DE GASES

Dentro de las distintas técnicas de extracción de gases disponibles [18], se ha hecho uso de la técnica de espacio de cabeza modificada para la extracción de las muestras gaseosas desde viales de vidrio. Esta técnica consiste en aislar la muestra de agua en forma líquida en un recipiente herméticamente cerrado, en nuestro caso un vial de vidrio, de manera que la disolución de contaminantes de agua desprenda los volátiles y queden suspendidos en la zona superior del recipiente. A continuación, a través de unas perforaciones para la entrada y salida del gas portador, se realiza el transporte de los volátiles desprendidos de la muestra hacia los sensores.

2.5 PROCESADO DE DATOS

El procesado de los datos se realiza una vez que las señales de los sensores han sido almacenadas en la nube. En concreto, los datos guardados son el tiempo de la medida, las respuestas de los sensores y otros parámetros como temperatura ambiente, estado de las electroválvulas, etc.

Como se muestra en la figura 4, este proceso puede ser dividido en cuatro etapas: extracción de características y preprocesado de la señal, reducción de la dimensionalidad, clasificación y toma de decisión [10]. El bloque inicial en esta figura representa el conjunto de sensores e instrumentación, cuya salida es el fichero con las medidas temporales realizadas. Los programas del procesado de datos han sido realizados con Matlab y se ha utilizado las toolboxes de Signal Processing y Neural Networks.

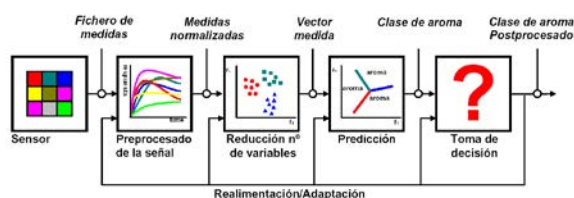


Figura 4: Diferentes etapas/Adaptación en el procesado de la señal obtenida de los sensores.

En la primera etapa se realiza el preprocesado de los datos medidos de los sensores. Éste se efectúa mediante un programa desarrollado en Matlab cuya entrada es el fichero de texto almacenado con los datos temporales y las medidas. A partir de él, se realiza la extracción de característica obteniendo, de esta forma, un único valor representativo de cada muestra medida. Para el cálculo de ese valor característico, se hace uso del algoritmo de resistencia relativa: división del valor en estado estacionario y el valor del gas de referencia.

En la segunda etapa tiene lugar la reducción de la dimensión de los datos extraídos. Para ello se hace uso del análisis de componentes principales (PCA). El PCA es una técnica estadística de síntesis de la información, o reducción de la dimensión (número de variables). Es decir, ante un conjunto de datos con muchas variables, el objetivo será reducirlas a un menor número perdiendo la menor cantidad de información posible. Los nuevos componentes principales o factores [16] serán, en definitiva, una combinación lineal de las variables originales, y además serán independientes entre sí.

La siguiente etapa tiene como objetivo realizar la predicción haciendo uso del vector de baja dimensión resultante de la etapa anterior. Las tareas que se deben desarrollar son agrupamiento (clustering),

regresión o clasificación. En los métodos de agrupamiento o clustering, el objetivo es aprender las relaciones estructurales entre diferentes contaminantes. En tareas de regresión, el objetivo es predecir un conjunto de propiedades (por ejemplo: concentración, calidad, etc.) de un analito. Finalmente, las tareas de clasificación buscan la identificación de una muestra desconocida como una clase dentro de un conjunto de contaminantes aprendidos anteriormente.

En la etapa final del sistema de Reconocimiento de Patrones diseñado en este trabajo, se realiza la clasificación de cada patrón de forma automática. Para ello, la técnica utilizada está basada en el uso de redes neuronales. Este tipo de algoritmos, son conjuntos de elementos de cálculo simples, usualmente adaptativos, interconectados masivamente en paralelo y con una organización jerárquica que le permite interactuar con algún sistema del mismo modo que lo hace el sistema nervioso biológico [14]. Para la implementación de los algoritmos de las tareas de clasificación (se han utilizado red feedforward con aprendizaje backpropagation y Redes con Funciones de Base Radial) se realizó un programa en Matlab. Para modificar los parámetros de las redes neuronales entrenadas, basta con editar los ficheros .m (de Matlab) que contienen la programación de las diferentes redes, en las que dichos parámetros se definen al principio de cada programa como variables.

Para determinar la validez de los modelos de clasificación con redes neuronales es necesario realizar predicciones de medidas de las que se sabe el valor correcto de salida para poder comparar la estimación realizada con el valor real. Este paso final en el reconocimiento de patrones permite la estimación de errores o rendimientos del modelo entrenado mediante técnicas de validación. En este trabajo, debido a la dificultad de obtener una gran cantidad de medidas en algunos de los experimentos y para sacar más rendimiento a las medidas obtenidas, se han utilizado validaciones cruzadas de orden 1 (también denominadas leave-one-out), lo que supone haber generado tantas redes como medidas se disponen. Esta filosofía no está pensada para proponer un modelo concreto para la clasificación de muestras o para la predicción de cada parámetro, sino para probar la validez de la aproximación.

3 RESULTADOS

Con objeto de estimar el rendimiento del sistema desarrollado en la identificación de gases, se han realizado una serie de medidas con unas disoluciones de prueba. La configuración de las medidas incluye

ciclos de 60s de adsorción y 540s de desorción. Las temperaturas de operación de los sensores varían entre 400 y 500°C. Se han tomado al menos 20 medidas de cada una de las muestras preparadas como disoluciones de varios compuestos en el agua: agua (W), acetona (Ac), tolueno (To), amoníaco (AM), formaldehído (FO), peróxido de hidrógeno (HY), etanol (Et), benceno (Be), diclorometano (Dc), ácido acético (AA), xileno (XY) y dimetilacetamida (Da).

Una vez realizadas las medidas y almacenados los datos en la nube, se preprocesan los datos y se realiza la extracción de características, es decir, se calcula la respuesta de cada sensor a partir de un algoritmo de cálculo, en este caso se ha utilizado el de resistencia relativa, en el que la respuesta de cada sensor se corresponde con el cociente entre el valor estable de cada sensor al gas de referencia y el mínimo valor obtenido en la medida de la muestra.

A continuación se realiza el análisis de componentes principales para reducir la dimensión de los datos y de esta forma poder representarlo en un gráfico.

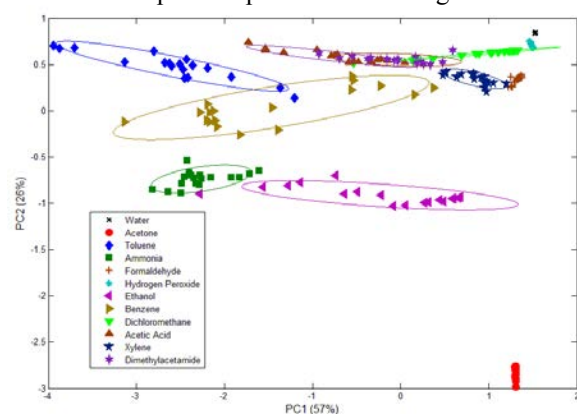


Figura 4: PCA de las medidas realizadas con un nodo de la red.

Se puede observar que los grupos de los diferentes contaminantes están claramente separados aunque aparece cierta superposición parcial entre algunos clusters.

Estos resultados se confirman con la clasificación con varios tipos de redes neuronales artificiales (RNAs). Se han utilizado: Red Neuronal Probabilística con funciones de base radial (RBF) y Red alimentada hacia delante con algoritmo de aprendizaje retropropagación (BP). Se utiliza una validación cruzada de orden 1 para estimar el rendimiento. La tasa de éxito (porcentaje de casos correctamente clasificados en la validación frente al número total de casos) obtenida fue de 94,4% utilizando la red RBF y 91,2% utilizando la red BP. Por otra parte, se ha desarrollado una aplicación basada en la web para el procesamiento de los datos

de forma remota, donde los usuarios pueden introducir los valores de los sensores y solicitar un valor de clasificación. Esta herramienta integra una red neuronal artificial para prestar el servicio de clasificación.

4 CONCLUSIONES

Se ha desarrollado una red inalámbrica de sensores con conexión a una nube de datos para la identificación de contaminantes. Las pruebas realizadas confirman la capacidad del sistema diseñado para discriminar entre las muestras medidas. Análisis de Componentes Principales y Redes Neuronales Artificiales se han utilizado para el procesamiento de los datos, obteniendo valores superiores al 91% de éxito en la clasificación.

Agradecimientos

Este trabajo ha sido realizado gracias al apoyo del Ministerio de Economía y Competitividad a través del proyecto TEMINAIR (TEC2013-48147-C6-5).

Referencias

- [1] Akyildiz, I.F., Melodia, T., Chowdury, K.R., (2007) "Wireless multimedia sensor networks: A survey." *IEEE Wireless Communications*, Vol. 14, Issue 6, 32-39.
- [2] Akyildiz, I.F., Stuntebeck, E.P., (2006) "Wireless underground sensor networks: Research challenges", *Ad Hoc Networks*, Vol. 4 (6), 669-686.
- [3] Alamri, A., Ansari W. S., Hassan, M. M., Hossain, M. S., Alelaiwi, A., Hossain, M. A., (2013), "A Survey on Sensor-Cloud: Architecture, Applications, and Approaches", *International Journal of Distributed Sensor Networks*, 18.
- [4] Alexe, A., Exhilarasie, R., (2011) "Cloud computing based vehicle tracking information systems", *International Journal of Computer Science and Telecommunications*, Vol. 2 (1).
- [5] Al-Haija, Q.A., Al-Qadeeb, H., Al-Lwaimi, A., (2013) "Case Study: Monitoring of AIR Quality in King Faisal University Using a Microcontroller and WSN", *Procedia Computer Science*, Vol. 21, 517-521.
- [6] Ayaz, M., Baig, I., Abdullah, A., Faye, I., (2011) "A survey on routing techniques in underwater wireless sensor networks." *Journal of Network and Computer Applications*, Vol. 34 (6), 1908-1927.
- [7] Brunelli, D., Rossi, M., (2014), "Enhancing lifetime of WSN for natural gas leakages detection." *Microelectronics Journal*, Vol. 45 (12), 1665-1670.

- [8] Deng, M., Chen, Q., (2010) "Coal and gas outburst monitoring system based on WSN." *Procedia Engineering*, Vol. 7, 387–391.
- [9] Gürses, E., Akan, Ö.B., (2005) "Multimedia communication in wireless sensor networks." *Annales des Telecommunications/Annals of Telecommunications*, Vol. 60 (7-8), 872-900..
- [10] Gutierrez-Osuna, R., (2002) "Pattern Analysis for Machine Olfaction: A Review." *IEEE Sensors Journal*, Vol. 2 (3), 189-202.
- [11] Hsu, W-C., Kuo, C-W., Chang, W-W., Chang, J-J., Hou, Y-T., Lan, Y-C., Sung, T-J., Yang, Y-J., (2010) "A WSN smart medication system." *Procedia Engineering*, Vol. 5, 588–591.
- [12] Jit, B., Maniyeri, J., Louis, S., Gopalakrishnan, K., Yap, P., (2009) Design and trial deployment of a practical sleep activity pattern monitoring system", *Proceedings of the International Conference on Smart Homes and Health Telematics*.
- [13] Jit, B., Maniyeri, J., Gopalakrishnan, K., (210) "Processing of wearable sensor data on the cloud – a step forward scaling of continuous monitoring of health and well-being", *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 3860-3863.
- [14] Kohonen, T., (1989) "Self-organization and associative memory". Springer Verlag, New York.
- [15] Kurata, N., Suzuki, M., Saruwatari, S., Morikawa, H., (2008) "Actual application of ubiquitous structural monitoring system using Wireless sensor networks", *Proceedings of the 14th World Conference on Earthquake Engineering*.
- [16] Llobet, E., Brezmes, J., Vilanova, K., Sueiras, J.E., Correig, X., (1997) "Qualitative and quantitative analysis of volatile organic compounds using transient and steady-state responses of a thick-film tin oxide gas sensor array", *Sensors and Actuators B*, 41 (1-3), 13-21.
- [17] Loomis, D., Grosse, Y., Lauby-Secretan, B., El Ghissassi, F., Bouvard, V., Benbrahim-Tallaa, L., Guha, N., Baan, R., Mattock, H., Straif, K., (2013) "The carcinogenicity of outdoor air pollution". *Lancet Oncol*, 14 (13):1262-3.
- [18] Lozano, J., Santos, J.P., Gutiérrez, J., Horrillo, M.C, (2007) "Comparative study of sampling systems combined with gas sensors for wine discrimination", *Sensors and Actuators B: Chemical*, Vol 126, 616-623.
- [19] Preve, N., (2011), "SEGEDMA: Sensor grid enhancement data management system for Health Care computing", *Expert Systems with Applications*, Vol. 38 (3), 2371–2380.
- [20] Ram G. Lakshmi Narayanan, Oliver C. Ibe, (2012), "A joint network for disaster recovery and search and rescue operations.", *Computer Networks*, Vol. 56 (14), 3347–3373.
- [21] Simon, G., Maroti, M., Ledeczki, A., Balogh, G., Kusy, B., Nadas, A., Pap, G., Sallai, J., Frampton, K., (2004), "Sensor network-based countersniper system.", *Proceedings of the Second International Conference on Embedded Networked Sensor Systems (Sensys)*, Baltimore, MD.
- [22] Toumpis, S., Tassiulas, T., (2006) "Optimal deployment of large wireless sensor networks.", *IEEE Transactions on Information Theory*, 52, 2935–2953.
- [23] Wener-Allen, G., Lorincz, K., Ruiz, M., Marcillo, O., Johnson, J., Lees, J., Walsh, M., (2006) "Deploying a wireless sensor network on an active volcano. Data-Driven Applications in Sensor Networks (Special Issue)", *IEEE Internet Computing*.
- [24] Yick, J., Mukherjee, B., Ghosal, D., (2008) "Wireless sensor network survey.", *Computer Networks*, Vol. 52 (12), 2292–2330.