Detección de anomalías en redes de sensores inalámbricos

Detection of anomalies in wireless sensor networks

VADILLO-MEJÍA, C. †, MOO-MENA, F*. y GÓMEZ-MONTALVO, J.

Universidad Autónoma de Yucatán, Anillo Periférico Norte, T.C. 13615, Chuburná Hidalgo Inn, Mérida, Yucatán, México.

ID 1er Autor: *C. Vadillo-Mejía* / **ORC ID**: 0000-0001-5904-8517, **CVU CONACYT ID:** 922808

ID 1er Coautor: *F. Moo-Mena* / **ORC ID**: 0000-0002-8812-2525, **CVU CONACYT ID:** 45837

ID 2do Coautor: *J. Gómez-Montalvo* / **ORC ID**: 0000-0002-5606-7517, **CVU CONACYT ID:** 56793

DOI: 10.35429/JTD.2019.11.3.22.37 Recibido: 18 de Julio, 2018; Aceptado 05 de Septiembre, 2018

Resumen

Con el tiempo, las redes de sensores inalámbricas (WSN) se han utilizado para una variedad de aplicaciones. Se ha dedicado un amplio trabajo a diversas aplicaciones de WSN. Es importante resaltar que, debido a sus limitaciones físicas, los sensores son propensos a varios tipos de fallas. Estas restricciones pueden plantear graves problemas en las aplicaciones de detección de eventos. Sobre todo, si las WSN son desplegados en entornos hostiles, como el sector industrial o ambiental. La detección de anomalías ha atraído recientemente la atención de la comunidad científica, debido a su relevancia en aplicaciones del mundo real. Las soluciones propuestas dependen en gran medida en la supervisión y en la comunicación, utilizando técnicas basadas en herramientas tales como Aprendizaje Automático y Redes Neuronales. En este contexto, realizamos una introducción a las técnicas de detección de anomalías más utilizadas en WSN. Recopilando y comparando los principales métodos aplicados en escenarios específicos, analizamos las ventajas y conveniencias de usar alguno de ellos.

Redes de Sensores, Detección de Fallas, Aprendizaje Automático

Abstract

Over time, wireless sensor networks (WSN) have been used for a variety of applications. Extensive work has been dedicated to various WSN applications. It is important to note that, due to their physical limitations, the sensors are prone to several types of faults. These restrictions can pose serious problems in event detection applications. Especially if the WSNs are deployed in hostile environments, such as the industrial or environmental sector. The detection of anomalies has recently attracted the attention of the scientific community, due to its relevance in real-world applications. The proposed solutions depend to a large extent on supervision and communication, using techniques based on tools such as Machine Learning and Neural Networks. In this context, we introduce the most commonly used anomaly detection techniques in WSN. Compiling and comparing the main methods applied in specific scenarios, we analyze the advantages and conveniences of using any of them.

Sensor Networks, Anomalies Detection, Machine Learnin

Citación: VADILLO-MEJÍA, C., MOO-MENA, F. y GÓMEZ-MONTALVO, J. Detección de anomalías en redes de sensores inalámbricos. Revista del Desarrollo Tecnológico. 2019 3-11: 22-37

* Correspondencia del Autor (Correo electrónico: a18016374@alumnos.uady.mx)
† Investigador contribuyendo como primer autor.

Introducción

Los recientes avances tecnológicos en hardware han hecho posible implementar pequeños nodos de sensores inalámbricos, de baja potencia, con poco ancho de banda y multifuncionales para monitorear e informar las condiciones y eventos en sus entornos locales (Liu D. a., 2007). Un gran conjunto de estos nodos sensores pueden formar redes de sensores inalámbricos de una manera ad hoc, creando un nuevo tipo de sistemas de información (Shahid, Naqvi, & Qaisar, 2014). Entre sus usos están el monitoreo ambiental, manejo de desastres, monitoreo médico, vigilancia inteligente, y recientemente en la creación de vehículos autónomos (Ogundile & Alfa, 2017). Las oportunidades y desafíos de estas redes de nodos pequeños, han atraído a una gran comunidad de investigadores y desarrolladores. En específico, en temas de identificación de anomalías en WSN. La implementación de una WSN en entornos extremos, tales como sistemas industriales, produce un conjunto propio de retos que pueden llevar a los sistemas a entrar en estados de falla (Gaura, 2010). Por ejemplo, la aparición de datos atípicos, ya sea por falta de energía, daño físico o interferencia ambiental (Akyildiz, Su, Sankarasubramaniam, & Cayirci, 2002). En la mayoría de las áreas de aplicación, los métodos tradicionales para el diagnóstico de anomalías en WSN dependen profundamente de la pericia de los técnicos. Este método es ineficiente y costoso en lugares donde es crítico tener un buen control de prevención de fallas, como es el caso en grandes sistemas industriales.

De manera general, la detección de anomalías se ha estudiado en múltiples contextos, desde entornos extremos como sistemas industriales (Ramotsoela, Abu-Mahfouz, & Hancke, 2018; Martí, Sanchez-Pi, Molina, & Garcia, 2015; Yi, y otros, 2015; Liu, Liu, Zhang, & Peng, 2016; Rabatel, Bringay, & Poncelet, 2011; Vries, Van Den Akker, Vonk, De Jong, & Van Summeren, 2016), redes de telecomunicaciones, monitoreo ambiental (Aslan, Korpeoglu, & Ulusoy, 2012; Ul Islam, Hossain, & Andersson, 2018; Mainwaring, Polastre, Szewczyk, & Culler, 2002; Rajasegarar, Leckie, Bezdek, & Palaniswami, 2010; Rassam, Maarof, & Zainal, 2014; Magán-Carrión, Camacho, & Garciá-Teodoro, 2015; Conde, 2011) , sistemas mecánicos o monitoreo médico (Ayadi, Ghorbel, Obeid, & Abid, 2017) (Smarsly & Law, 2014; Cowton, Kyriazakis, Plötz, & Bacardit, 2018; Haque, Rahman, & Aziz, 2015; Salem, Liu, & Mehaoua, 2013; Alemdar & Ersoy, 2010).

En años recientes han surgido métodos más eficaces que hacen frente a las carencias que, como humanos, podemos tener. Entre los que destacan métodos no paramétricos, como técnicas de Inteligencia Artificial o Aprendizaje Automático (ML, por sus siglas en ingles). Hay varios artículos de revisión dedicados a estudiar extensamente las técnicas principales para la detección de anomalías, donde se destaca (Chandola, Banerjee, & Kumar, 2009) por su aportación que nos permite comprender mejor las técnicas existentes de acuerdo al dominio de la investigación. Para redes de sensores encontramos artículos de revisión recientes como (Ayadi, Ghorbel, Obeid, & Abid, 2017) , y otros no tan recientes, pero con gran aportación al tema (Xie, Han, Tian, & Parvin, 2011; S. Rajasegarar & Palaniswami, 2008; Shahid, Naqvi, & Qaisar, 2012; Savage, Zhang, Yu, Chou, & Wang, 2014).

En este documento se retoman las técnicas más destacadas de detección de anomalías en WSN mencionadas en estos artículos de revisión. El propósito es actualizar las técnicas con trabajos más recientes y hacer una introducción a las más utilizadas para detección de anomalías en una WSN.

El documento se encuentra estructurado de la siguiente forma: en la Sección 2 se introducirá el tema de anomalías en las redes de sensores y qué tipos existen. En la Sección 3 se hablará brevemente sobre los métodos más destacados utilizados en la detección de anomalías en redes de sensores. Para finalizar, se presentará un análisis de las técnicas en la Sección 4.

Detección de anomalías

Una anomalía o un valor atípico, se define como una observación que es inconsistente con el resto de las muestras observadas (S. Rajasegarar & Palaniswami, 2008). Por otro lado, el concepto de detección de anomalías puede variar dependiendo del contexto en que se emplea. En (Kumarage, Khalil, Tari, & Zomaya, 2013) se define como una rama de la detección de intrusiones, que identifica un comportamiento anormal sin conocimiento previo de la naturaleza de ese comportamiento. Por otro lado, en (Al-Thani, 2018) se define como un campo cuyo objetivo es encontrar patrones o datos que no encajen con el comportamiento esperado dentro de un conjunto general de datos.

Mientras que en (Behravan, y otros, 2017), se define como el área de investigación que busca automatizar el proceso de detección de comportamientos anormales en sistemas físicos y diagnosticar las causas. En WSN, la detección de anomalías se puede definir como un proceso de identificación de comportamientos anómalos, de una forma precisa y utilizando una cantidad mínima de recursos disponibles en la WSN (Maleh & Ezzati, 2015).

Tipos de anomalías en WSN

Por la naturaleza de los sensores, existen varios factores que hacen a las WSN ser propensas a la aparición de datos atípicos. En (Conde, 2011) se describen algunas, en las que se encuentran: errores de calibración o instalación, falta de mantenimiento, daño físico, cambios en el ambiente del sensor.

– Errores de calibración o instalación: este tipo de errores suelen producir ruido en las mediciones. Es posible detectarlos por la extrema diferencia que existe con los datos normales (Ayadi, Ghorbel, Obeid, & Abid, 2017).
– Falta de mantenimiento: los sensores, como todo instrumento físico tienen un lapso de vida limitado. La mayoría de ellos funcionan con baterías y con el tiempo su rendimiento tiende a deteriorarse, produciendo mediciones erróneas o atípicas (Ramotsoela, Abu-Mahfouz, & Hancke, 2018).
– Cambios en el ambiente: Los sensores suelen usarse en escenarios extremos donde su integridad física suele exponerse a altas temperaturas, radiación, o zonas químicamente corrosivas por periodos prolongados. Este tipo de estrés puede producir desviaciones significativas en los patrones normales de los datos (Shahid, Naqvi, & Qaisar, 2012).

Un aspecto importante de la detección de anomalías es la naturaleza de la misma. En (Ahmed, Naser Mahmood, & Hu, 2016) se menciona la siguiente clasificación:

– Anomalía puntual: Cuando una instancia de datos particular se desvía del patrón normal del conjunto de datos.
– Anomalía contextual: Cuando una instancia de datos se comporta de manera anómala en un contexto particular.
– Anomalía colectiva: Cuando una colección de instancias de datos similares se comporta de forma anómala con respecto a todo el conjunto de datos.

Métodos de detección de anomalías en WSN

Un método de detección de anomalías en redes de sensores debe tener las siguientes propiedades. Primero, debe poder identificar con precisión todos los tipos de anomalías, así como el comportamiento normal (Ayadi, Ghorbel, Obeid, & Abid, 2017). Segundo, debe ser robusto, es decir, debe poder manejar los cambios de patrones en los conjuntos de datos (Ramotsoela, Abu-Mahfouz, & Hancke, 2018). Tercero, por las limitaciones de recursos en los sistemas de sensores, este debe ser eficiente (Chandola, Banerjee, & Kumar, 2009). Por último, también es deseable que un algoritmo de detección pueda descubrir anomalías en tiempo real o casi en tiempo real (Yao, Sharma, Golubchik, & Govindan, 2010).

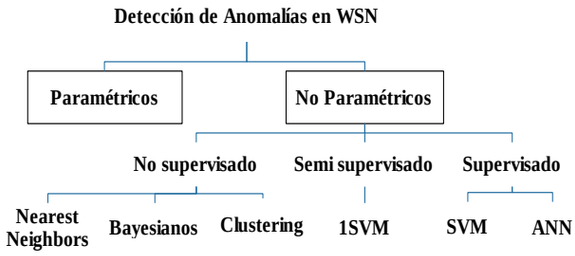

Figura 1 Clasificación de técnicas para detección de anomalías en WSN

En la Figura 1 se desglosa las técnicas analizadas en este documento. La elección de algunas de estas técnicas está en función de las características del problema. Existen varios indicadores como la naturaleza de los datos, la disponibilidad de datos etiquetados, el tipo de anomalía que se va a detectar y la arquitectura de comunicación de la WSN que ayudan a definir el problema (Chandola, Banerjee, & Kumar, 2009). Por ejemplo, dependiendo de la forma de comunicación y computación entre los nodos de la WSN, se distinguen dos esquemas para la detección de anomalías, centralizado y distribuido. En un enfoque centralizado, todos los datos recibidos en nodos individuales se transmiten a un nodo central.

El nodo central es responsable de procesar todos los datos recibidos de la red y determinar los valores o eventos atípicos (Ayadi, Ghorbel, Obeid, & Abid, 2017). Estos enfoques requieren que se comuniquen grandes cantidades de mediciones sin procesar a un nodo central para su procesamiento (S. Rajasegarar & Palaniswami, 2008). Esta carga genera un aumento en la energía de la WSN, reduciendo la vida útil de la red (Miao, Liu, Zhao, & Li, 2018). Además, dado que el nodo central es una pieza importante de un enfoque centralizado, si este nodo falla, todo el proceso de detección de anomalías se ve afectado de igual forma.

En un enfoque distribuido, las tareas se ejecutan en cada nodo localmente (Miao, Liu, Zhao, & Li, 2018). Se usan los datos locales y la información recibida de sus vecinos de un salto (Ayadi, Ghorbel, Obeid, & Abid, 2017). Se realiza algún procesamiento que permita determinar las estadísticas de los datos recopilados. Posteriormente, cada nodo transmite sus datos a un jefe de clúster en la red. El jefe es responsable de procesar las estadísticas recibidas de todos los nodos y determinar las anomalías (S. Rajasegarar & Palaniswami, 2008). Este enfoque tiene muchas ventajas en comparación con el procesamiento centralizado. Por ejemplo, dado que el procesamiento distribuido no necesita un nodo central, solo intercambia mensajes entre vecinos a un solo salto. Además, es escalable, robusto y la sobrecarga de comunicación por nodo se puede mantener a un nivel considerable (Miao, Liu, Zhao, & Li, 2018).

A partir del tipo de esquema de comunicación, centralizado o distribuido, se pueden identificar subclasificaciones de métodos de detección de anomalías en WSN, por ejemplo, en (Dunning & Friedman, 2012) se catalogan con base a dos posibles escenarios. Primero, podríamos no tener un conocimiento a priori de lo que se está buscando. Segundo, podríamos tener información etiquetada de las anomalías. En el primer enfoque se detectan anomalías sin un conocimiento previo de los datos. Técnicas de aprendizaje no supervisado o clustering destacan en este enfoque. En el segundo caso, se cuenta con datos etiquetados, es decir, los datos están marcados como normales o anormales. Se destacan técnicas de aprendizaje supervisado, donde se entrena el clasificador a partir de los datos etiquetados. Luego, el clasificador entrenado se usa para clasificar datos nuevos.

Los autores de (S. Rajasegarar & Palaniswami, 2008), dan un paso más y consideran una tercera categoría a las dos descritas previamente. Este tercero es análogo a un enfoque semi supervisado, donde un clasificador aprende una generalización breve de un conjunto de datos dado, que luego puede usarse para reconocer anomalías. Enseguida, este puede aprender de manera incremental el modelo normal a medida que los datos estén disponibles.

En (Ahmed, Naser Mahmood, & Hu, 2016), se agrupan las técnicas en cuatro categorías: Clasificación, Estadísticos, Clustering y Teóricas de información. En (Ramotsoela, Abu-Mahfouz, & Hancke, 2018) y (S. Rajasegarar & Palaniswami, 2008) los autores utilizan dos categorías, paramétricos (estadísticos) y no paramétricos (asumen no tener conocimiento previo sobre la distribución de los datos). En los métodos paramétricos encontramos técnicas basadas en métodos estadísticos, que asumen o estiman un modelo de la distribución de los datos y evalúan las instancias de datos con respecto a qué tan bien se ajustan al modelo (Shahid, Naqvi, & Qaisar, 2012). En los no paramétricos encontramos técnicas de clasificación enfocados en métodos de minería de datos e inteligencia artificial. Estos se pueden agrupar en tres subcategorías: supervisado, semi supervisado y no supervisado.

Las Tablas 1 y 2, resumen las características de estas dos categorías.

Categoría	Características
Paramétrico	• Distribución estadística. • Requiere conocimiento a priori del modelo. • Entorno estático. • Detección rápida.
No paramétrico	• Datos etiquetados. • No requiere conocimiento a priori. • Más flexibles. • Entorno dinámico. • Detección moderada/lento.

tsoela, Abu-Mahfouz, & Hancke,

Tabla 1 Métodos paramétricos y no paramétricos

Métodos paramétricos

Las técnicas paramétricas ajustan un modelo estadístico asumiendo el conocimiento de la distribución de densidad (Chandola, Banerjee, & Kumar, 2009).

Se detecta una anomalía cuando se observan datos que tienen una probabilidad pequeña de ocurrir (Ramotsoela, Abu-Mahfouz, & Hancke, 2018). Los métodos paramétricos pueden ser univariados o multivariados, siendo esta última más adecuada para WSN (Ramotsoela, Abu-Mahfouz, & Hancke, 2018). El primero solo trata con una variable aleatoria a la vez, mientras que el último permite que más de una variable aleatoria se modele utilizando la misma función de distribución.

Métodos no paramétricos

En las técnicas no paramétricas la estructura del modelo no está definida a prioiri, sino que se determina a partir de los datos dados (Chandola, Banerjee, & Kumar, 2009). Estas técnicas suelen hacer menos suposiciones con respecto a los datos, en comparación con las técnicas paramétricas. Además, realizan el de patrones a través del aprendizaje automático en el que se utiliza un conjunto de datos conocido para encontrar la relación entrada/salida del sistema. Implementados correctamente son eficientes con los recursos de las WSN (Ramotsoela, Abu-Mahfouz, & Hancke, 2018).

Método	Conocimiento previo	Ventajas	Desventajas
Supervisado	Si	-Detección rápida de anomalías.	-Requiere datos etiquetados. -Entornos estáticos. -Propenso a sobre entrenamiento.
No supervisado	No	-No requiere datos reales de entrenamiento. -Entorno dinámico.	-Detección moderada/lenta.
Semi supervisado	Parcial	-Puede aprender progresivamente a medida que los datos estén disponibles. -Flexible a cambios en los datos. -Detección rápida/moderada.	-Si el aprendizaje falla, los errores pueden reforzarse a sí mismos.

Tabla 2 Técnicas no paramétricas
Fuente: (Ramotsoela, Abu-Mahfouz, & Hancke, 2018)

En función del grado de disponibilidad de las etiquetas de los datos, estás técnicas pueden funcionar en tres modos: supervisado, semi supervisado y no supervisado (Chandola, Banerjee, & Kumar, 2009).

Supervisado

Los enfoques supervisados necesitan aprender modelos de normalidad y anormalidad utilizando datos preestablecidos (Han, 2012). Por lo tanto, la identificación de un nuevo punto de datos como normal o atípico depende del modelo que encaje en el punto de datos (Ayadi, Ghorbel, Obeid, & Abid, 2017). Los ejemplos incluyen redes neuronales artificiales (ANN), Support Vector Machines (SVM), etc.

También conocidas como técnicas de clasificación supervisada (Han, 2012), producen una función f, el *clasificador*, capaz de asociar algunos datos de entrada, normalmente un vector x de atributos numéricos x_i , llamados *características*, a un valor de salida y, la etiqueta de clase, tomada de una lista Y de posibles datos.

Para construir esta función de mapeo, el algoritmo de clasificación supervisada necesita datos de ejemplos ya etiquetados. En otras palabras, un conjunto de parejas (x, y), también llamado conjunto de entrenamiento (Schatz, Hoßfeld, Janowski, & Egger, 2013) .

Uno de los principales requisitos de estos métodos es la calidad de los datos. En aplicaciones reales, no es tan sencillo obtener datos de una alta calidad para el entrenamiento, ni mucho menos datos sobre anomalías (Blomquist & Möller, 2015).

Este enfoque no se puede utilizar para una clasificación en línea, donde el clasificador aprende el modelo de clasificación con la llegada de nuevas muestras de datos y luego clasificar las siguientes muestras de datos según el modelo aprendido. Técnicas basadas en estadísticas son en su mayoría supervisadas (Shahid, Naqvi, & Qaisar, 2012).

Redes neuronales artificiales

Una técnica básica de detección de anomalías que utiliza redes neuronales funciona en dos pasos. Primero, una red neuronal se entrena con datos de estados normales. Segundo, cada instancia de prueba se proporciona como una entrada a la red neuronal. Si la red acepta la entrada de prueba, es normal y si la red rechaza una entrada de prueba, es una anomalía (Chandola, Banerjee, & Kumar, 2009).

La elección del algoritmo de entrenamiento, la arquitectura de la red, la representación de la señal de entrada y el conjunto de entrenamiento, juegan un papel importante para el entrenamiento de estas redes (Azimisadjadi, Poole, Sheedvash, Sherbondy, & Stricker, 1992). Por ejemplo, se ha demostrado que una red neuronal de tres capas, con no linealidad sigmoidal en los nodos, puede aproximar cualquier función no lineal arbitraria y generar cualquier región de decisión compleja necesaria para las tareas de detección y clasificación (Hecht-Nielsen, 1988). La elección del algoritmo de entrenamiento, por otro lado, determina la tasa de convergencia hacia una solución, el tiempo requerido para alcanzar una solución y la optimización de la misma. Si se utilizan suficientes muestras de entrenamiento y parámetros internos, la transformación entrada-salida puede definirse con una precisión arbitraria (Hecht-Nielsen, 1988).

Las redes neuronales se han usado en varios dominios como procesamiento de imágenes, sin embargo, suelen tener un alto requerimiento computacional. Para detección de anomalías en redes de sensores, se han mezclado técnicas de redes neuronales con otras técnicas, como los métodos estadísticos. En (Hawkins, He, Williams, & Baxter, 2002), se utiliza una Red Neural Replicante (RNN) para proporcionar un factor incontable para el tráfico anómalo de la red. Se trata de una percepción multicapa de retroalimentación con tres capas ocultas situadas entre las capas de entrada y salida.

Su objetivo es reproducir el patrón de datos de entrada en la capa de salida con un error minimizado a través de la formación. En (Ma, Wang, Cheng, Yu, & Chen, 2016) se propone un enfoque novedoso llamado SCDNN, que combina algoritmos de agrupamiento espectral (SC, por sus siglas en ingles) y de red neuronal profunda (DNN, por sus siglas en ingles).

Los resultados experimentales indican que el clasificador SCDNN no solo funciona mejor que una red neuronal de propagación hacia atrás (BPNN, por sus siglas en ingles), Support Vector Machines (SVM), Random Forests (RF) y modelos de redes Bayesianas en la precisión de detección y los tipos de ataques anormales encontrados, también proporciona una herramienta eficaz de estudio y análisis de detección de intrusos en redes de sensores.

En (Subba, Biswas, & Karmakar, 2018) se utiliza una combinación de reglas de especificación y un módulo ligero de detección de anomalías basado en redes neuronales para identificar nodos de sensores maliciosos. Los resultados de la simulación muestran que el framework propuesto logra una mayor precisión y tasa de detección en una amplia gama de ataques, mientras que al mismo tiempo minimiza el consumo total de energía y el volumen de tráfico en la WSN.

Support Vector Machines

Las técnicas basadas en Support Vector Machine, utilizan algoritmos robustos de aprendizaje supervisado que se basan en el principio de minimización del riesgo estructural de la teoría del aprendizaje estadístico (Vapnik, 1998). Su objetivo es encontrar un hiperplano lineal que separe un conjunto de muestras positivas de un conjunto de muestras negativas con un margen máximo.

Este margen se define por la distancia del hiperplano al punto más cercano de las muestras positivas y negativas, también llamados vectores de soporte (Yélamos, Escudero, Graells, & Puigjaner, 2009). Sin embargo, las técnicas estándares de SVM no tienen un buen desempeño en casos donde hay escasez de datos atípicos. En estos casos se emplean modelos híbridos (Erfani, Rajasegarar, Karunasekera, & Leckie, 2016; Shahid, Naqvi, & Qaisar, 2014; Saeedi Emadi & Mazinani, 2018; Maleh & Ezzati, 2015; Raghuvanshi, Rajeev, & Sudarshan, 2000).

Otro problema común es que, en la mayoría de los casos, los datos no son linealmente separables y es necesario utilizar estrategias como proyectarlos a otras dimensiones (Raghuvanshi, Rajeev, & Sudarshan, 2000). En este caso, las funciones del kernel se utilizan para transformar el espacio multidimensional original en otro, donde las clases son lineales y separables (Maleh & Ezzati, 2015).

La idea clave es mapear, mediante una función de kernel, los puntos de entrenamiento en un espacio recién transformado, generalmente de mayor o incluso de infinita dimensionalidad, donde los puntos pueden ser separados eficientemente con un hiperplano (Schatz, Hoßfeld, Janowski, & Egger, 2013).

En la práctica, las SVM son entrenadas usando diferentes kernels para seleccionar el que tenga el mejor rendimiento para el problema planteado (Morales, Cebrián, Fernandez-Blanco, & Sierra, 2016).

Semi supervisado

En la mayoría de las aplicaciones, las muestras anómalas son generalmente insuficientes e inexactas, lo que complica el uso de métodos supervisados. Para abordar esto, se aplican técnicas semi supervisadas para modelar los registros normales, y solo los registros que no cumplen con el modelo generado se etiquetan como anómalos (Pang, 2018). Estás técnicas se caracterizan por primero enseñarles la clase normal/anormal, pero posteriormente el algoritmo aprende a reconocer a la otra clase deseada (Hodge & Austin, 2004). Es adecuado para datos dinámicos, ya que sólo aprende una clase que proporciona el modelo de normalidad o anormalidad. Puede aprender el modelo gradualmente a medida que llegan nuevos datos, ajustando el modelo para mejorar los resultados, a medida que cada nuevo ejemplar esté disponible.

Técnicas basadas en derivaciones de SVM, como One-Class SVM (1SVM), se han vuelto una opción recurrente en WSN (Bahrepour, Meratnia, Poel, Taghikhaki, & Havinga, 2010; Rajasegarar, Leckie, Bezdek, & Palaniswami, 2010; Zhang, Meratnia, & Havinga, 2009; Sánchez, 2003).

1SVM

Recientemente varios enfoques 1SVM han sido propuestos para la detección de anomalías. El funcionamiento general de 1SVM es mapear primero los vectores de datos (mediciones) desde el espacio de entrada al espacio de la característica, mediante una función no lineal (Rajasegarar, Leckie, Bezdek, & Palaniswami, 2010). Los vectores mapeados en el espacio de la característica son vectores de imagen. Luego, se encuentra una superficie suave o un límite en el espacio de características que separe los vectores de imagen en mediciones normales y anómalas (Rajasegarar, Leckie, Bezdek, & Palaniswami, 2010). En otras palabras, aprenden el límite alrededor de las instancias normales durante el entrenamiento, mientras que ignoran alguna instancia anómala en los datos.

Es decir, cualquier instancia nueva que se encuentre fuera de este límite como un valor atípico (Shahid, Naqvi, & Qaisar, 2012).

Al usar una función kernel para mapear implícitamente el espacio de entrada a un espacio de características de mayor dimensión, estos métodos pueden modelar patrones altamente no lineales de comportamiento normal de una manera flexible (Maleh & Ezzati, 2015).

No supervisado

Un reto en la detección de anomalías no paramétricas es obtener datos etiquetados para entrenar un clasificador. La obtención de estos datos limpios y etiquetados suele ser costoso o una tarea manualmente intensiva (Rajasegarar, Leckie, Bezdek, & Palaniswami, 2010). Además, en el caso de las redes de sensores, el entrenamiento debe realizarse con frecuencia para adaptarse a los cambios en el comportamiento normal a lo largo del tiempo, ya sea periódicamente o en línea, sin que esté disponible ninguna de estas etiquetas.

Las técnicas no supervisadas, son unas opciones validas que no requieren de un proceso de entrenamiento. Pueden identificar valores atípicos basados en modelos estándar de distribución estadística o en la distancia total entre un punto y sus vecindarios (Ayadi, Ghorbel, Obeid, & Abid, 2017). Suponen implícitamente que los casos normales son mucho más frecuentes que las anomalías en los datos (Portnoy, Eskin, & Stolfo, 2001).

Cuando esta suposición no es cierta, estas técnicas sufren de altas tasas de falsas alarmas. Muchas técnicas semi supervisadas pueden ser adaptadas para funcionar en modo no supervisado utilizando una muestra del conjunto de datos sin etiquetar como datos de entrenamiento (Chandola, Banerjee, & Kumar, 2009). Esta adaptación supone que los datos de las pruebas contienen muy pocas anomalías y que el modelo aprendido durante el entrenamiento es los suficientemente robusto frente a estas pocas anomalías. Esta es la principal ventaja sobre los métodos supervisados, y es que no requieren los datos reales de entrenamiento, lo cual es muy importante, especialmente en la práctica (Li & Teng, 2006). Las técnicas que más destacan en esta categoría son las basadas en clustering y Nearest Neighbors.

Clustering

Clustering es una técnica popular dentro de la comunidad de minería de datos para agrupar instancias de datos con comportamientos similares en clústeres (Ayadi, Ghorbel, Obeid, & Abid, 2017). En WSN, las mediciones o los datos recopilados por los nodos sensores pueden agruparse mediante la identificación de grupos con mediciones similares en los datos (Park, 2018). Aquí la similitud significa la proximidad de los vectores de datos entre sí. La detección de anomalías se hace con base a las agrupaciones formadas. Sí una instancia de dato nueva no pertenece a una agrupación, o si un grupo de datos forman agrupaciones pequeñas en comparación con otras agrupaciones, entonces se consideran datos anómalos (Ayadi, Ghorbel, Obeid, & Abid, 2017). Para la comunicación, los datos en cada nodo se agrupan mediante clústeres hiperesféricos.

En la literatura, existen variantes de algoritmos basados en clustering para extracción de datos con valores atípicos, como (Ma, Wang, Cheng, Yu, & Chen, 2016; Moshtaghi, 2011; Rajasegarar, Leckie, & Palaniswami, 2014; Loo, 2006)

Nearest Neighbors

Estos enfoques utilizan varias nociones de distancias (medida de similitud) entre dos instancias de datos (Rajasegarar, Leckie, & Palaniswami, 2013). Ejemplos de medidas incluyen la distancia al vector de datos vecino más cercano (NN), la distancia al kth vector de datos vecino más cercano (kNNM), y la distancia al promedio de los k vectores de datos más cercanos (kNN promedio) (Rajasegarar, Leckie, & Palaniswami, 2014).

K es un parámetro definido por el usuario. Estas medidas de similitud se utilizan para ordenar los vectores de datos y clasificarlos como normales o anómalos. Por ejemplo, una instancia de datos se declara como un valor atípico si se encuentra lejos de sus vecinos (Rajasegarar, Leckie, & Palaniswami, 2013).

K-Nearest Neighbors (KNN) es una de las técnicas ya establecidas y usadas en detección de anomalías (Zhang, y otros, 2010; Yihua Liao, 2002).

Este enfoque se ha utilizado para diversos propósitos, como clasificación, clustering y detección de valores atípicos (Ayadi, Ghorbel, Obeid, & Abid, 2017). Sin embargo, es difícil de implementar en WSN debido a la complejidad de los cálculos y recurrencia de los mismos, que demandan un alto consumo de energía (Haque, Rahman, & Aziz, 2015). Algunos trabajos basados en estas técnicas son (Zhang K. a., 2007; Branch, 2013; Zhuang, 2006; Magán-Carrión, Camacho, & Garciá-Teodoro, 2015; Xie, Hu, Han, & Chen, 2013)

Bayesianos

Los métodos Bayesianos se caracterizan por encontrar que un sensor sea defectuoso a través de la probabilidad según el teorema de Bayes (Muhammed, 2017). Las redes bayesianas (BN) son una de las técnicas usadas en WSN que se basan en métodos Bayesianos. Una BN usa un modelo gráfico probabilístico que es aprendido de un conjunto de datos de entrenamiento y estima un valor de sensor calculando una probabilidad condicional (Zhang H. a., 2018). Además, pueden agregar lecturas de diferentes sensores en diferentes momentos para proporcionar una mejor precisión de estimación (Zhang H. a., 2018).

En una red bayesiana, existe una relación padre-hijo entre los nodos que indica que una variable representada por un nodo hijo depende de aquellos representados por los nodos padres.

Las BN se pueden utilizar en un esquema de clasificación de eventos, haciéndolas aplicables para la detección de anomalías (Ahmed, Naser Mahmood, & Hu, 2016). El uso de BN para la detección de valores atípicos permite tener en cuenta la dependencia probabilística entre las variables aleatorias (De Paola, 2015).

El aprendizaje u otras formas de razonamiento se realizan mediante las reglas de probabilidad. Estas consisten en encontrar la categoría que hace que la probabilidad posterior sea mayor si se proporciona un conjunto de datos. Cada nodo de la WSN implementa solo una parte de una BN. Cada vez que un nodo sensor coopera con otros nodos, su parte de BN se conecta con aquellos que residen en otros lugares (De Paola, 2015).

ISSN 2531-2960

En la literatura se han encontrado modelos exitosos para su implementación, tales como Näive Bayesian, Bayesian Networks y Bayesian Neuronal Networks (Liu, Qi, Hou, & Chang, 2008). El uso de redes Bayesianas como medio para el aprendizaje no supervisado y la detección de anomalías en redes de sensores de monitoreo de gas para minas de carbón subterráneas se describe en (Hutchison, 2003).

Los autores demostraron que el modelo de red Bayesiano puede aprender líneas de base cíclicas para concentraciones de gas, reduciendo así las falsas alarmas causadas generalmente por umbrales de línea plana. Su solución ha demostrado ser eficaz tanto en el enfoque distribuido como en el centralizado.

Análisis de técnicas

Muchos de los métodos que acabamos de describir, específicamente los no paramétricos, necesitan de modificaciones y optimizaciones al algoritmo estándar para funcionar eficientemente en WSN. En los métodos paramétricos el modelo crece solo con la complejidad del modelo, no con el tamaño de los datos (Ramotsoela, Abu-Mahfouz, & Hancke, 2018).

Por lo tanto, permiten que el modelo se evalúe rápidamente en nuevas instancias, lo que los hace adecuados para grandes conjuntos de datos. Si se sabe que los datos se ajustan a dicho modelo de distribución y que esta distribución no cambiará durante el tiempo de vida de la WSN, estos enfoques son una buena opción para la detección de anomalías (Hodge & Austin, 2004). Sin embargo, debido a que su aplicabilidad depende del conocimiento a priori del modelo, que a menudo no está disponible o es costoso de obtener, no son la mejor opción para nuestro sistema de interés. Además, como sistema complejo la distribución de datos evoluciona a lo largo de la vida útil de la WSN, dificultando más la aplicabilidad de estos métodos en el sistema de interés (S. Rajasegarar & Palaniswami, 2008).

Por otro lado, las técnicas no paramétricas no requieren tener conocimiento previo sobre la distribución de los datos. Estas técnicas son adecuadas para redes de sensores con recursos limitados donde la distribución de datos puede cambiar con frecuencia.

Por ejemplo, estos cambios pueden ser causados por el agotamiento de la energía de los sensores a lo largo de la vida útil de la red, lo que afecta la estabilidad de la topología de enrutamiento y, por lo tanto, puede afectar la detección de intrusiones basada en anomalías. Los cambios sobre el tipo de ambiente monitoreado también pueden afectar la distribución de los valores de medición. Estas características descritas encajan en el caso de estudio, haciendo viable el uso de técnicas no paramétricas. Sin embargo, la complejidad computacional es mayor a las paramétricas. Por ejemplo, las basadas en SVM requieren una solución de optimización cuadrática o lineal en cada instante. Por su parte, las basadas en Nearest Neighbors o clustering, tienen la mayor complejidad computacional, ya que requieren el cálculo de la distancia euclidiana multivariada entre cada par de muestras de datos. A pesar de esto, la complejidad de la comunicación es comparable a los métodos paramétricos, ya que sólo se necesita transmitir unos pocos parámetros entre los distintos nodos de la red, reduciendo el consumo de energía de la WSN y alargando su ciclo de vida.

Entre los métodos no paramétricos, los supervisados tienen la desventaja de requerir tanto datos normales como anormales para el entrenamiento del modelo. Estos datos pueden ser reales o simulados. Además, son susceptibles al sobre entrenamiento cuando no se generalizan bien en casos completamente nuevos. Sin embargo, proporcionan un alto grado de detección de datos anómalos, siempre que sean implementados correctamente.

Los SVM son modelos de aprendizaje supervisado no paramétricos, cuya complejidad crece de forma cuadrática con el número de muestras. Se adaptan mejor a conjuntos de datos pequeños con muchas características, pero también a entrenamiento a gran escala con datos en altas dimensiones a través del uso de kernels. Sin embargo, SVM requiere mucho tiempo y memoria para el entrenamiento (Erfani, Rajasegarar, Karunasekera, & Leckie, 2016).

Las redes neuronales ofrecen soluciones robustas y adaptables para detectar y clasificar objetivos en un dominio muy desordenado. Cuando se agregan nuevos datos o reglas al sistema, no es necesario volver a entrenar el sistema, principalmente solo agregando nuevas reglas.

Tienen la capacidad de generalizar a partir de datos limitados, con mucho ruido e incompletos. Dado a que tienen altos requisitos computacionales, no son recomendables para la detección de anomalías en un WSN. Pero en conjunto con otras técnicas, como un método estadístico o variantes de la misma, resultan ser bastante poderosas.

En el caso de no contar con suficientes datos anómalos, los métodos semi supervisados y no supervisados nos dan opciones aplicables.

El primero es adecuado para datos dinámicos, ya que sólo aprende una clase que proporciona el modelo de normalidad o anormalidad. Posteriormente, este va aprendiendo progresivamente el modelo a medida que llegan nuevos datos, ajustándolo con cada nueva instancia disponible.

Por otro lado, las técnicas no supervisadas, son unas opciones validas que no requieren datos etiquetados.

Pueden identificar valores atípicos basados en modelos estándar de distribución estadística, por ejemplo, redes bayesianas, o en la distancia total entre un punto y sus vecindarios, como los métodos de Clustering.

Sin embargo, tienen una alta complejidad computacional y cuando no se cumplen con los supuestos definidos, estas técnicas sufren de altas tasas de falsas alarmas.

Los enfoques basados en redes bayesianas utilizan un modelo gráfico probabilístico de un conjunto de variables y sus dependencias probabilísticas. Agregan datos de diferentes instancias y proporcionan una estimación de que un evento pertenezca a la clase aprendida.

La Tabla 3 resume las características de cada de una de las técnicas no paramétricas contempladas para WSN.

Además, se anexan algunos trabajos de investigación donde se han aplicado estas técnicas de detección de anomalías en WSN.

Categoría	Técnica	Trabajos	Características
Supervisado	SVM	(Hu, Granderson, Auslander, & Agogino, 2019; Martí, Sanchez-Pi, Molina, & Garcia, 2015; Saeedi Emadi & Mazinani, 2018; Erfani, Rajasegarar, Karunasekera, & Leckie, 2016; Shahid, Naqvi, & Qaisar, 2014; Ma, Wang, Cheng, Yu, & Chen, 2016; Rajasegarar, Leckie, Bezdek, & Palaniswami, 2010; Raghuvanshi, Rajeev, & Sudarshan, 2000; Maleh & Ezzati, 2015) (Granjal, Silva, & Lourenço, 2018) (Miao, Liu, Zhao, & Li, 2018) (Feng, Fu, Du, Li, & Sun, 2017)	– Requiere datos etiquetados. – Alta complejidad computacional. – Buen desempeño en conjuntos de datos pequeños. – Usa kernels para lidiar con altas dimensiones.
	ANN	(Díaz, Carta, & Matías, 2018; Tanprasert, Saiprasert, & Thajchayapong, 2017; Ma, Wang, Cheng, Yu, & Chen, 2016; Azimisadjadi, Poole, Sheedvash, Sherbondy, & Stricker, 1992; Reddy, Sarkar, Venugopalan, & Giering, 2016; Subba, Biswas, & Karmakar, 2018; Curiac & Volosencu, 2012; Cowton, Kyriazakis, Plötz, & Bacardit, 2018; Conde, 2011)	– Requiere datos etiquetados. – Robustos y adaptables. – capacidad de aprender y modelar relaciones no lineales y complejas
Semi supervisado	1SVM	(Ma, Wang, Cheng, Yu, & Chen, 2016; Moshtaghi, 2011; Rajasegarar, Leckie, & Palaniswami; Loo, 2006)	– Requiere mínimos datos etiquetados. – Útil cuando se tiene poca información de las anomalías.
No supervisado	Clustering	(Ma, Wang, Cheng, Yu, & Chen, 2016; Moshtaghi, 2011; Rajasegarar, Leckie, & Palaniswami, 2014; Loo, 2006)	– No requiere datos etiquetados – Consumo elevado de energía – Fácil de implementar. – Crea agrupaciones de elementos con las mismas características
	Nearest Neighbors	(Rajasegarar, Leckie, & Palaniswami, 2014; Janeja, Adam, Atluri, & Vaidya, 2010; Xie, Hu, Han, & Chen, 2013; Zhu, Feng, & Huang, 2016; Rajasegarar, Leckie, & Palaniswami, 2013; Liu & Deng, 2013; Bosman, Iacca, Tejada, Wörtche, & Liotta, 2017) (Zhang K. a., 2007; Branch, 2013; Zhuang, 2006; Magán-Carrión, Camacho, & Garcíá-Teodoro, 2015; Xie, Hu, Han, & Chen, 2013)	– No requiere datos etiquetados – Consumo elevado de energía – Intuitivo y simple – No requiere entrenamiento – Utiliza la distancia como medida de similitud entre los datos
	Bayesiano	(Wang, Lizier, Obst, Prokopenko, & Wang, 2008; Krishnamachari & Iyengar, 2004; Janakiram, Reddy, & Kumar, 2006; Hill, Minsker, & Amir, 2007)	– No requiere datos etiquetados – Detectan anomalías a través de probabilidades

Tabla 3 Técnicas analizadas para detección de anomalías en WSN

Conclusiones

De manera general, la detección de anomalías consiste en identificar mediciones que se desvían significativamente de un perfil establecido, para el comportamiento normal dentro de un dominio particular.

La capacidad de detectar fallas con alta precisión en redes de sensores se ha vuelto cada vez más importante, especialmente en dominios como la industria y la salud. En redes de sensores, es altamente deseable que la detección de anomalías se realice de manera distribuida para prolongar la vida útil de la red.

En este documento se han descrito dos formas generales para clasificar las técnicas para detección de fallas en WSN, paramétricos y no paramétricos. Los métodos paramétricos son adecuados en entornos estables donde la distribución de datos es bien conocida y es poco probable que cambie con frecuencia. Por otro lado, los métodos no paramétricos pueden utilizarse en entornos dinámicos, en los que la distribución estadística es desconocida, pero inferida a través de los datos. Además, la mayoría de los métodos no paramétricos son ideales para dispositivos con restricción de recursos y propensos a cambios, como los nodos de sensores. Los métodos no paramétricos se pueden subclasificar en otros tres, supervisado, semi supervisado y no supervisado.

Los métodos supervisados requieren tener datos de estados normales y anormales. Además, que el clasificador se vuelva a entrenar si las características de los datos cambian. Los semi supervisados aprenden una generalización breve de un conjunto de datos dado. Posteriormente, se retroalimentan de manera incremental a medida que los datos estén disponibles. De modo que pueda adaptarse a los cambios en la distribución de los datos. Por último, los métodos no supervisados, como Clustering y basados en Nearest Neighbors, suponen que los valores atípicos están bien separados de los puntos de datos que son normales.

Cada una tiene sus ventajas y desventajas, pero es de destacar que SVM y todas sus modificaciones se han vuelto populares en años recientes, dado por su eficiencia y por su alto nivel de detección. La elección de una técnica dependerá del problema y el contexto del mismo. Para nuestro caso de estudio, por las características descritas previamente, los métodos no paramétricos suponen una opción aplicable para detectar comportamientos anormales en la WSN. Además, se cuenta con datos históricos del funcionamiento de los procesos, teniendo a disposición datos etiquetados.

En este caso, los enfoques supervisados y semi supervisados son una opción a considerar. Aunque por lo descrito anteriormente se consideran los adecuados, no debemos descartar totalmente el uso de otros métodos.

Agradecimientos

Los autores agradecen al CONACYT por el financiamiento recibido para la realización de este trabajo.

Referencias

Ahmed, M., Naser Mahmood, A., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19-31. doi:10.1016/j.jnca.2015.11.016

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). A Survey on Sensor Networks. 102-114. doi:10.1109/MWC.2010.5416354

Alemdar, H., & Ersoy, C. (2010). Wireless sensor networks for healthcare: A survey. Computer Networks, 54, 2688-2710. doi:10.1016/j.comnet.2010.05.003

Al-Thani, H. a.-M. (2018). Unsupervised Technique for Anomaly Detection in Qatar Stock Market. 2018 International Conference on Computer and Applications (ICCA). doi:10.1109/COMAPP.2018.8460282

Aslan, Y. E., Korpeoglu, I., & Ulusoy, ö. (2012). A framework for use of wireless sensor networks in forest fire detection and monitoring. Computers, Environment and Urban Systems, 36, 614-625. doi:10.1016/j.compenvurbsys.2012.03.002

Ayadi, A., Ghorbel, O., Obeid, A. M., & Abid, M. (2017). Outlier detection approaches for wireless sensor networks: A survey. Computer Networks, 129, 319-333. doi:10.1016/j.comnet.2017.10.007

Azimisadjadi, M. R., Poole, E. E., Sheedvash, S., Sherbondy, K. D., & Stricker, S. A. (1992). Detection and Classification of Buried Dielectric Anomalies Using a Separated Aperture Sensor and a Neural Network Discriminator. Ieee Transactions on Instrumentation and Measurement, 41, 137-143. doi:10.1109/19.126648

Bahrepour, M., Meratnia, N., Poel, M., Taghikhaki, Z., & Havinga, P. J. (2010). Distributed event detection in wireless sensor networks for disaster management. Proceedings - 2nd International Conference on Intelligent Networking and Collaborative Systems, INCOS 2010, 507-512. doi:10.1109/INCOS.2010.24

Behravan, A., Obermaisser, R., Hanike, D., Mallak, A., Weber, C., & Fathi, M. (2017). Fault Injection Framework for Fault Diagnosis based on Machine Learning in Heating and Demand-Controlled Ventilaion Systems. 273-279.

Blomquist, H., & Möller, J. (2015). Anomaly detection with Machine learning Quality assurance of statistical data in the Aid community. Retrieved from https://pdfs.semanticscholar.org/eda8/b4887feae76f2a64181887bc4fb0a45d7ec4.pdf

Bosman, H. H., Iacca, G., Tejada, A., Wörtche, H. J., & Liotta, A. (2017). Spatial anomaly detection in sensor networks using neighborhood information. Information Fusion, 33, 41-56. doi:10.1016/j.inffus.2016.04.007

Branch, J. W. (2013). In-network outlier detection in wireless sensor networks. Knowledge and Information Systems, 34(1), 23--54.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly Detection: A Survey. ACM Computing Surveys, 41, 1-58. doi:10.1145/1541880.1541882

Conde, E. F. (2011). ENVIRONMENTAL SENSOR ANOMALY DETECTION by Erick F . Conde A thesis submitted in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE in Civil and Environmental Engineering Approved : Dr . Mac McKee .

Cowton, J., Kyriazakis, I., Plötz, T., & Bacardit, J. (2018). A combined deep learning GRU-autoencoder for the early detection of respiratory disease in pigs using multiple environmental sensors. Sensors (Switzerland), 18. doi:10.3390/s18082521

Curiac, D. I., & Volosencu, C. (2012). Ensemble based sensing anomaly detection in wireless sensor networks. Expert Systems with Applications, 39, 9087-9096. doi:10.1016/j.eswa.2012.02.036

De Paola, A. a. (2015). Adaptive distributed outlier detection for WSNs. IEEE Transactions on Cybernetics, 45(5), 888--899.

Díaz, S., Carta, J. A., & Matías, J. M. (2018). Performance assessment of five MCP models proposed for the estimation of long-term wind turbine power outputs at a target site using three machine learning techniques. Applied Energy, 209, 455-477. doi:10.1016/j.apenergy.2017.11.007

Dunning, T., & Friedman, E. (2012). Practical Machine Learning A New Look at Anomaly Detection. O'Reilly Media, Inc.

Erfani, S. M., Rajasegarar, S., Karunasekera, S., & Leckie, C. (2016). High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. Pattern Recognition, 58, 121-134. doi:10.1016/j.patcog.2016.03.028

Feng, Z., Fu, J., Du, D., Li, F., & Sun, S. (2017). A new approach of anomaly detection in wireless sensor networks using support vector data description. International Journal of Distributed Sensor Networks, 13. doi:10.1177/1550147716686161

Gaura, E. a. (2010). Wireless sensor networks: Deployments and design frameworks. Springer New York Dordrecht Heidelberg London. doi:10.1007/978-1-4419-5834-1

Granjal, J., Silva, J. M., & Lourenço, N. (2018). Intrusion detection and prevention in CoAP wireless sensor networks using anomaly detection. Sensors (Switzerland), 18. doi:10.3390/s18082445

Han, J. a. (2012). Data Mining: Concepts and Techniques. Morgan Kaufmann.

Haque, S. A., Rahman, M., & Aziz, S. M. (2015). Sensor anomaly detection in wireless sensor networks for healthcare. Sensors (Switzerland), 15, 8764-8786. doi:10.3390/s150408764

Hawkins, S., He, H., Williams, G., & Baxter, R. (2002). Outlier Detection Using Replicator Neural Networks. (Y. Kambayashi, W. Winiwarter, & M. Arikawa, Eds.) 170-180.

Hecht-Nielsen, R. (1988). Theory of the backpropagation neural network. International 1989 Joint Conference on Neural Networks, 593-605 vol.1.

Hill, D. J., Minsker, B. S., & Amir, E. (2007). Real-Time Bayesian Anomaly Detection for Environmental Sensor Data. Proceedings of the Congress-International Association for Hydraulic Research, 503.

Hodge, V. J., & Austin, J. I. (2004). A Survey of Outlier Detection Methodologies. 85-126. doi:10.4324/9781315744988

Hu, R. L., Granderson, J., Auslander, D. M., & Agogino, A. (2019). Design of machine learning models with domain experts for automated sensor selection for energy fault detection. Applied Energy, 235, 117-128. doi:10.1016/j.apenergy.2018.10.107

Hutchison, D. a. (2003). EWSN 2008: Wireless Sensor Networks. Retrieved from http://www.math.tau.ac.il/ mansour/coursegames/nash-load.pdf

Janakiram, D., Reddy, V. A., & Kumar, A. V. (2006). Outlier Detection in Wireless Sensor Networks using Bayesian Belief Networks. 2006 1st International Conference on Communication Systems Software & Middleware, 1-6. doi:10.1109/COMSWA.2006.1665221

Janeja, V. P., Adam, N. R., Atluri, V., & Vaidya, J. (2010). Spatial neighborhood based anomaly detection in sensor datasets. Data Mining and Knowledge Discovery, 20, 221-258. doi:10.1007/s10618-009-0147-0

Krishnamachari, B., & Iyengar, S. (2004). Distributed Bayesian Algorithms for Fault-Tolerant Event Region Detection in Wireless Sensor Networks. IEEE Transactions on Computers, 53, 241-250.

Kumarage, H., Khalil, I., Tari, Z., & Zomaya, A. (2013). Distributed anomaly detection for industrial wireless sensor networks based on fuzzy data modelling. Journal of Parallel and Distributed Computing, 73, 790-806. doi:10.1016/j.jpdc.2013.02.004

Li, K., & Teng, G. (2006). Unsupervised SVM Based on p-kernels for Anomaly Detection. First International Conference on Innovative Computing, Information and Control - Volume I (ICICIC'06), 2, 59-62. doi:10.1109/ICICIC.2006.371

Liu, D. a. (2007). Security for Wireless Sensor Networks. Springer Science+Business Media, LLC.

Liu, J., & Deng, H. (2013). Outlier detection on uncertain data based on local information. Knowledge-Based Systems, 51, 60-71. doi:10.1016/j.knosys.2013.07.005

Liu, L., Liu, D., Zhang, Y., & Peng, Y. (2016). Effective sensor selection and data anomaly detection for condition monitoring of aircraft engines. Sensors (Switzerland), 16. doi:10.3390/s16050623

Liu, T., Qi, A., Hou, Y., & Chang, X. (2008). Method for network anomaly detection based on Bayesian statistical model with time slicing. Proceedings of the World Congress on Intelligent Control and Automation (WCICA), 3359-3362. doi:10.1109/WCICA.2008.4593458

Loo, C. E. (2006). Intrusion detection for routing attacks in sensor networks. International Journal of Distributed Sensor Networks, 2(4), 313--332.

Ma, T., Wang, F., Cheng, J., Yu, Y., & Chen, X. (2016). A Hybrid Spectral Clustering and Deep Neural Network Ensemble Algorithm for Intrusion Detection in Sensor Networks. Sensors, 16, 1701. doi:10.3390/s16101701

Magán-Carrión, R., Camacho, J., & Garciá-Teodoro, P. (2015). Multivariate statistical approach for anomaly detection and lost data recovery in wireless sensor networks. International Journal of Distributed Sensor Networks, 2015. doi:10.1155/2015/672124

Mainwaring, A., Polastre, J., Szewczyk, R., & Culler, D. (2002). Wireless Sensor Network fot Habitat Monitoring. doi:10.1145/570738.570751

Maleh, Y., & Ezzati, A. (2015). Lightweight intrusion detection scheme for wireless sensor networks. IAENG International Journal of Computer Science, 42, 347-354. doi:10.1155/2015/653232

Martí, L., Sanchez-Pi, N., Molina, J. M., & Garcia, A. C. (2015). Anomaly detection based on sensor data in petroleum industry applications. Sensors (Switzerland), 15, 2774-2797. doi:10.3390/s150202774

Miao, X., Liu, Y., Zhao, H., & Li, C. (2018). Distributed Online One-Class Support Vector Machine for Anomaly Detection Over Networks. IEEE Transactions on Cybernetics, PP, 1-14. doi:10.1109/TCYB.2018.2804940

Morales, I. R., Cebrián, D. R., Fernandez-Blanco, E., & Sierra, A. P. (2016). Early warning in egg production curves from commercial hens: A SVM approach. Computers and Electronics in Agriculture, 121, 169-179. doi:10.1016/j.compag.2015.12.009

Moshtaghi, M. a. (2011). Clustering ellipses for anomaly detection. Pattern Recognition, 44(1), 55--69.

Muhammed, T. a. (2017). An analysis of fault detection strategies in wireless sensor networks. Journal of Network and Computer Applications, 78(October 2016), 267--287.

Ogundile, O. O., & Alfa, A. S. (2017). A survey on an energy-efficient and energy-balanced routing protocol for wireless sensor networks. Sensors (Switzerland), 17, 1-52. doi:10.3390/s17051084

Pang, J. a. (2018). Optimize the coverage probability of prediction interval for anomaly detection of sensor-based monitoring series. Sensors (Switzerland), 18(4). doi:10.3390/s18040967

Park, S. a. (2018). Unsupervised and non-parametric learning-based anomaly detection system using vibration sensor data. Multimedia Tools and Applications.

Portnoy, L., Eskin, E., & Stolfo, S. (2001). Intrusion Detection with Unlabeled Data Using Clustering. In Proceedings of ACM CSS Workshop on Data Mining Applied to Security (DMSA-2001, (pp. 5-8).

Rabatel, J., Bringay, S., & Poncelet, P. (2011). Anomaly detection in monitoring sensor data for preventive maintenance. Expert Systems with Applications, 38, 7003-7015. doi:10.1016/j.eswa.2010.12.014

Raghuvanshi, A. S., Rajeev, T., & Sudarshan, T. (2000). MACHINE LEARNING APPROACH FOR ANOMALY DETECTION IN WIRELESS SENSOR DATA. Stress: The International Journal on the Biology of Stress, 1, 76-99.

Rajasegarar, S., Leckie, C., & Palaniswami, M. (2013). DISTRIBUTED ANOMALY DETECTION IN WIRELESS SENSOR NETWORKS. Conference Record - Asilomar Conference on Signals, Systems and Computers, 428-432. doi:10.1109/ACSSC.2013.6810312

Rajasegarar, S., Leckie, C., & Palaniswami, M. (2014). Hyperspherical cluster based distributed anomaly detection in wireless sensor networks. Journal of Parallel and Distributed Computing, 74, 1833-1847. doi:10.1016/j.jpdc.2013.09.005

Rajasegarar, S., Leckie, C., Bezdek, J. C., & Palaniswami, M. (2010). Centered hyperspherical and hyperellipsoidal one-class support vector machines for anomaly detection in sensor networks. IEEE Transactions on Information Forensics and Security, 5, 518-533. doi:10.1109/TIFS.2010.2051543

Ramotsoela, D., Abu-Mahfouz, A., & Hancke, G. (2018). A survey of anomaly detection in industrial wireless sensor networks with critical water system infrastructure as a case study. Sensors (Switzerland), 18, 1-25. doi:10.3390/s18082491

Rassam, M. A., Maarof, M. A., & Zainal, A. (2014). Adaptive and online data anomaly detection for wireless sensor systems. Knowledge-Based Systems, 60, 44-57. doi:10.1016/j.knosys.2014.01.003

Reddy, K. K., Sarkar, S., Venugopalan, V., & Giering, M. (2016). Anomaly Detection and Fault Disambiguation in Large Flight Data: A Multi-modal Deep Auto-encoder Approach. Phm, 1-8. doi:10.1039/c0Ob00047g

S. Rajasegarar, C. L., & Palaniswami, M. (2008). Anomaly Detection in Wireless Sensor Networks. Ieee Wireless Communications, 34-40.

Saeedi Emadi, H., & Mazinani, S. M. (2018). A novel anomaly detection algorithm using DBSCAN and SVM in wireless sensor networks. Wireless Personal Communications, 98, 2025-2035. doi:10.1007/s11277-017-4961-1

Salem, O., Liu, Y., & Mehaoua, A. (2013). Anomaly detection in medical wireless sensor networks. Journal of Computing Science and Engineering, 7, 272-284. doi:10.5626/JCSE.2013.7.4.272

Sánchez, V. D. (2003). Advanced support vector machines and kernel methods. Neurocomputing, 55, 5-20. doi:10.1016/S0925-2312(03)00373-4

Savage, D., Zhang, X., Yu, X., Chou, P., & Wang, Q. (2014). Anomaly detection in online social networks. Social Networks, 39, 62-70. doi:10.1016/j.socnet.2014.05.002

Schatz, R., Hoßfeld, T., Janowski, L., & Egger, S. (2013). Data Traffic Monitoring and Analysis: From Measurement, Classification, and Anomaly Detection to Quality of Experience. doi:10.1007/978-3-642-36784-7

Shahid, N., Naqvi, I. H., & Qaisar, S. B. (2012). Characteristics and classification of outlier detection techniques for wireless sensor networks in harsh environments: a survey. Artificial Intelligence Review, 43, 193-228. doi:10.1007/s10462-012-9370-y

Shahid, N., Naqvi, I. H., & Qaisar, S. B. (2014). SVM based event detection and identification: Exploiting temporal attribute correlations using sensgru. Mathematical Problems in Engineering, 2014. doi:10.1155/2014/259508

Smarsly, K., & Law, K. H. (2014). Decentralized fault detection and isolation in wireless structural health monitoring systems using analytical redundancy. Advances in Engineering Software, 73, 1-10. doi:https://doi.org/10.1016/j.advengsoft.2014.02.005

Subba, B., Biswas, S., & Karmakar, S. (2018). A Game Theory Based Multi Layered Intrusion Detection Framework for Wireless Sensor Networks. International Journal of Wireless Information Networks, 25, 1-23. doi:10.1007/s10776-018-0403-6

Tanprasert, T., Saiprasert, C., & Thajchayapong, S. (2017). Combining Unsupervised Anomaly Detection and Neural Networks for Driver Identification. Journal of Advanced Transportation, 2017. doi:10.1155/2017/6057830

Ul Islam, R., Hossain, M. S., & Andersson, K. (2018). A novel anomaly detection algorithm for sensor data under uncertainty. Soft Computing, 22, 1623-1639. doi:10.1007/s00500-016-2425-2 Vapnik, V. N. (1998). Statistical Learning Theory.

Vries, D., Van Den Akker, B., Vonk, E., De Jong, W., & Van Summeren, J. (2016). Application of machine learning techniques to predict anomalies in water supply networks. Water Science and Technology: Water Supply, 16, 1528-1535. doi:10.2166/ws.2016.062

Wang, X. R., Lizier, J. T., Obst, O., Prokopenko, M., & Wang, P. (2008). Spatiotemporal Anomaly Detection in Gas Monitoring Sensor Networks. In R. Verdone (Ed.), Wireless Sensor Networks (pp. 90-105). Berlin: Springer Berlin Heidelberg.

Xie, M., Han, S., Tian, B., & Parvin, S. (2011). Anomaly detection in wireless sensor networks: A survey. Journal of Network and Computer Applications, 34, 1302-1325. doi:10.1016/j.jnca.2011.03.004

Xie, M., Hu, J., Han, S., & Chen, H. H. (2013). Scalable hypergrid k-NN-based online anomaly detection in wireless sensor networks. IEEE Transactions on Parallel and Distributed Systems, 24, 1661-1670. doi:10.1109/TPDS.2012.261

Yao, Y., Sharma, A., Golubchik, L., & Govindan, R. (2010). Online anomaly detection for sensor systems: A simple and efficient approach. Performance Evaluation, 67, 1059-1075. doi:10.1016/j.peva.2010.08.018

Yélamos, I., Escudero, G., Graells, M., & Puigjaner, L. (2009). Performance assessment of a novel fault diagnosis system based on support vector machines. Computers and Chemical Engineering, 244-255. doi:10.1016/j.compchemeng.2008.08.008

Yi, W. Y., Lo, K. M., Mak, T., Leung, K. S., Leung, Y., & Meng, M. L. (2015). A survey of wireless sensor network based air pollution monitoring systems (Vol. 15). doi:10.3390/s151229859

Yihua Liao, V. R. (2002). Use of K-Nearest Neighbor classifier for intrusion detection. 21, 439-448.

Zhang, H. a. (2018). A Bayesian network model for data losses and faults in medical body sensor networks. Computer Networks, 143, 166--175.

Zhang, K. a. (2007). Unsupervised Outlier Detection in Sensor Networks Using Aggregation Tree. 158--169.

Zhang, Y., Meratnia, N., & Havinga, P. (2009). Adaptive and online one-class support vector machine-based outlier detection techniques for wireless sensor networks. Proceedings - International Conference on Advanced Information Networking and Applications, AINA, 990-995. doi:10.1109/WAINA.2009.200

Zhang, Y.-Y., Chao, H.-C., Chen, M., Shu, L., Park, C.-H., & Park, M.-S. (2010). Outlier detection and countermeasure for hierarchical wireless sensor networks. IET Information Security, 4, 361. doi:10.1049/iet-ifs.2009.0192

Zhu, Q., Feng, J., & Huang, J. (2016). Weighted natural neighborhood graph: an adaptive structure for clustering and outlier detection with no neighborhood parameter. Cluster Computing, 19, 1385-1397. doi:10.1007/s10586-016-0598-1

Zhuang, Y. a. (2006). In-network Outlier Cleaning for Data Collection in Sensor Networks. Workshop in VLDB.