

Generación automática de catálogos de activos mediante Inteligencia Artificial para aplicar Confiabilidad Operacional

Dayrien Corrales Díaz

Centro de Informática Industrial, Universidad de las Ciencias Informáticas, La Havana, Cuba, dcorrales@uci.cu

Luanner Kerton Martínez

Centro de Informática Industrial, Universidad de las Ciencias Informáticas, La Havana, Cuba, luanner@uci.cu

Yoel Martín Pardo

Centro de Informática Industrial, Universidad de las Ciencias Informáticas, La Havana, Cuba, ympardo@uci.cu

ABSTRACT

The Integral Reliability for Assets System (SCIA) makes a criticality analysis, maintenance plan and Cost-Risk-Benefit for helping to decide what maintenance tasks will be applied in PDVSA industries. However, to make operational reliability process to the assets, users have to input all its data manually, which set a limitation to the usage of the system only for objects classified as prioritized.

An excellent assets data source could be the databases of several departments in the industry. This research propose an extraction information method about the equipment before submit them to a reliability cycle, creating a dynamical catalog through the Kohonen's self-organizing maps.

The solution allows to get a list of the identified assets in the Configuration SCADA Guardián del ALBA subsystem, installed in some productive plants, been an alternate way to get a technological independence and remove importations. Others empirical test were made, looking for an effectively prove of the correct execution of the algorithm in different settings.

Keywords: SCIA, PDVSA, Operational Reliability, PDVSA, Artificial Intelligence, SCADA GALBA.

RESUMEN

El Sistema de Confiabilidad Integral de Activos (SCIA) realiza análisis de la criticidad, planificación de mantenimiento y optimización basada en Costo-Riesgo-Beneficio que ayuda a decidir la realización de las tareas de mantención en instalaciones de la industria PDVSA. Actualmente, la aplicación de la confiabilidad operacional sobre los activos se realiza de forma manual, lo cual limita el uso de este sistema solo a los equipos clasificados como prioridad.

Una fuente idónea de datos de los activos la constituyen las bases de datos de los diversos departamentos de la industria. Esta investigación propone un método de extracción de la información referente a los activos para someterlos a un clico de confiabilidad, generando dinámicamente un catálogo mediante el uso de mapas auto-organizados de Kohonen.

El resultado alcanzado permite obtener un catálogo de los activos identificados en el módulo de Configuración del SCADA Guardián del ALBA (GALBA), implantado en distintas instalaciones de producción como vía a la soberanía tecnológica y sustitución de importaciones. También se realizaron pruebas experimentales para comprobar su efectividad de ejecución sobre otros escenarios.

Palabras claves: SCIA, Confiabilidad Operacional, PDVSA, Inteligencia Artificial, SCADA GALBA.

1. INTRODUCCION

Las grandes empresas siempre abogan por incrementar al nivel de productividad. Para ello un fundamental es la aplicación de mantenimiento a los activos para alargar su vida útil y así evitar paradas en los procesos de producción debido a fallas irreparables y cambio de equipamiento.

Un equipo multidisciplinario formado por especialistas del Centro de Informática Industrial (CEDIN) de Universidad de las Ciencias Informáticas (UCI) en La Habana, Cuba; y el departamento de Automática, Informática y Telecomunicaciones (AIT) de la empresa PDVSA, Venezuela; desempeñan el desarrollo del Sistema de Confiabilidad Integral de Activos para los pueblos del ALBA (SCIA). Esta herramienta permite aplicar un ciclo de confiabilidad operacional a los activos para realizar la planificación de mantenimiento.

Actualmente el proceso de captura de los datos de los activos se realiza de forma manual. Esto provoca que se limite el uso de la herramienta de confiabilidad solo a los activos identificados con prioridad por parte de los ingenieros o como resultado de una decisión gerencial. En cualquiera de los casos se emplea tiempo y esfuerzo debido a la incapacidad que tiene el SCIA de gestionar automáticamente la información de los activos contenido en los diferentes medios de almacenamiento.

En este punto se define: La necesidad de identificar cada activo a partir de la información almacenada en las bases de datos, teniendo en cuenta que la nomenclatura de los datos varía en cada instalación, como problema a resolver.

2. DESARROLLO

2.1 DATOS, INFORMACIÓN Y CONOCIMIENTO

Los sistemas operacionales centran su funcionamiento en el registro de los datos de las principales actividades que realizan. Estos están estrechamente vinculados con las Base de Datos Relacionales donde persisten los resultados de las operaciones. Un dato por sí mismo tiene poca utilidad, solo representa un valor, pero al relacionarlos entre sí ya toma otro significado.

Un conjunto de datos analizados sobre un tema en común toma cierta relevancia que correctamente interpretados conciben una información, siendo esta, a su vez, quien origina un conocimiento (Maneiro, 2008).



Figura 1. Relación entre los Datos, la Información y el Conocimiento.

La obtención del conocimiento es aplicado a la solución de problemas, en caso contrario se desperdicia. Partiendo de esto es posible conocer los activos de una instalación si al menos se realiza un procesamiento transaccional en línea (OLTP) que tenga en cuenta los datos correspondientes a los mismos.

El descubrimiento del conocimiento basa su fortaleza en el uso de mecanismos de inteligencia artificial, capaces de identificar y extraer patrones de grandes cantidades de datos crudos, a partir de argumentos que representan el resultado deseado según las metas de los usuarios. En este caso el proceso está centrado en la clasificación de los datos.

2.2 TÉCNICA INTELIGENTE

El algoritmo empleado en la solución propuesta es el de mapas auto organizados de Kohonen. Es una red neuronal presentada por T. Kohonen en 1982, denominada originalmente SOM. Es un mecanismo de la inteligencia artificial que posee un aprendizaje no supervisado y competitivo (Ultsch, 1990).

Basándose en el descubrimiento de rasgos comunes, regularidades, correlaciones en los datos de entrada; se auto organiza en función de parámetros recibidos. Guarda similitud con las observaciones realizadas al córtex de los animales superiores en las que aparecen zonas donde las neuronas detectoras de rasgos se encuentran ordenadas topológicamente (Yin, 2007).

Este mecanismo inteligente cuenta con las siguientes ventajas:

- Clasificar muestras mediante observaciones. Usa métodos matemáticos y a través de un comportamiento competitivo crea las propias categorías a partir de características presentes en los datos.
- Representa la información en forma de mapa. Contribuye a la clasificación directa de los activos de entrada, permitiendo la configuración de condiciones de los objetivos.

Funcionamiento del algoritmo.

Este modelo está compuesto por dos capas de neuronas. La capa de entrada tiene N neuronas que reciben la información y la transmiten a la capa de salida. En la capa de salida se tiene M neuronas que procesan la información y forma los mapas de rasgos. Cada neurona de la capa de entrada está conectada a todas las neuronas de la capa de salida mediante un peso. De esta forma, cada una de las neuronas de salida tiene un peso conocido como vector de referencia (o codebook).

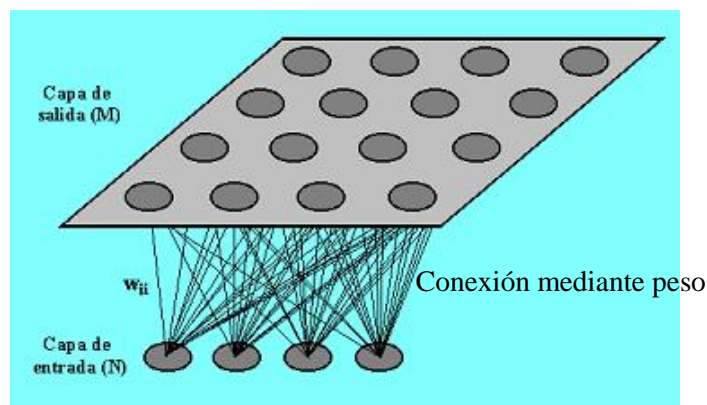


Figura 2. Representación de las capas de los mapas auto-organizados de Kohonen.

Puede decirse que entre las neuronas de salida existen conexiones laterales de excitación, pues cada una influye sobre sus vecinas, siendo este el elemento que rige el proceso de competición y sumado a la aplicación de una función de vecindad, que produce una topología en el mapa.

Es resultado del algoritmo es fácilmente representable, pues se obtiene un modelo bidimensional agrupado. Ejemplo de ello son las imágenes siguientes, tomadas de un prototipo que prueba el funcionamiento de esta red neuronal; a la izquierda la entrada, a la derecha el resultado del algoritmo.

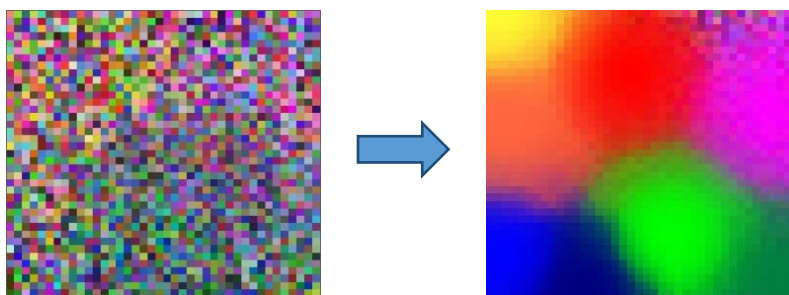


Figura 3. Ejemplo del funcionamiento de los mapas de Kohonen.

Se tienen pruebas de la utilidad de este método el área empresarial, con diversos propósitos como la determinación de criticidad de activos para procesos de mantenimiento, ordenamiento de información en sistemas referenciados, ubicación geoespacial, desarrollo de catálogos con fines estratégicos empresariales, auditoría de sistemas informáticos, clasificación de señales emitidas por equipos industriales y determinación de cuotas económicas.

En el sector de la salud se aprecia su aplicación en identificación de tumores cerebrales a partir de imágenes cerebrales. En el área educativa se empleó para la selección y clasificación de estudiantes, a partir de la definición de un grupo de variables que resumen el potencial mostrado durante el ciclo de estudio; con el objetivo de asignarlos a diferentes procesos del ciclo de formación (Rizo, 2012).

2.3 MATERIALES Y MÉTODOS

La solución propuesta es un subsistema para el SCIA capaz de generar automáticamente un catálogo de activos. Está desarrollada usando varias tecnologías y lenguajes que aportan la capacidad de operar independientemente del Modelo Entidad-Relación de las bases de datos.

Se emplea Python como lenguaje principal para el desarrollo de la herramienta, el cual aporta una gran productividad en el desarrollo. La arquitectura principal cuenta con un modelo de 3 capas: Interfaz de Usuario (GUI), Lógica de Negocio (Business Logic) y Acceso a Datos (Data Access). La capa de interfaz de usuario está desarrollada con una arquitectura interna MVP (Model-View-Presentation), y usa el framework QT versión 5.0, mediante el binding PyQt. En la capa de lógica de negocio se usa de la biblioteca Kohonen en su versión 0.1.4 escrita en C++. El empleo de la biblioteca XLWT de Python permite exportar el resultado a archivos Excel.

El nivel de acceso a datos cuenta con el Object Relational Model (ORM) SQLAlchemy, versión 0.9. El uso del ORM permite extender la solución a distintos tipos de base de datos sin necesidad de realizar cambios en el código. Facilita el mapeo dinámico de las tablas, la navegación por las mismas y el manejo de las consultas necesarias para obtener el conjunto de datos requeridos.

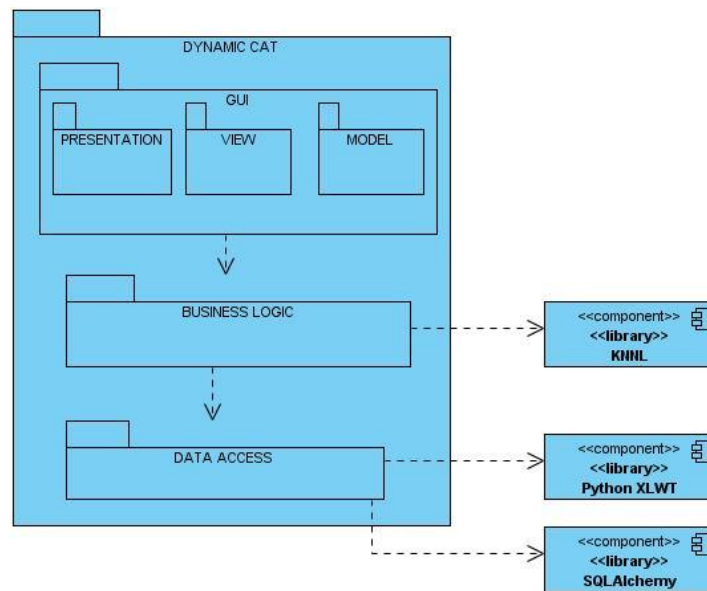


Figura 4. Arquitectura de la solución propuesta.

Esta aplicación tiene como punto de partida la definición de un diccionario (estructura de datos definida en el lenguaje Python) del tipo **{etiqueta: valor}** el cual se usa como argumento en los criterios de búsqueda en la base de datos, y como capa de entrada en la red neuronal para proveer el vector de referencia. Las consultas generan un conjunto de datos que forman parte de la capa de salida del algoritmo. Finalmente se define una estructura de árbol para ajustar el resultado de los mapas auto-organizados de Kohonen.

3. RESULTADO Y DISCUSIÓN

La producción de petróleo y sus derivados (Gasolina, Gas natural, Azufre, Coke, entre otros) es el principal proceso en la industria de PDVSA. Los sistemas de Supervisión, Control y Adquisición de Datos (SCADA: por sus siglas en inglés) son las herramientas que permiten observar constantemente qué está ocurriendo con cada dispositivo en el campo. Para llevar a cabo estas tareas primero debe configurarse cada dispositivo o sensor y variables asociadas a ellos, punto en el cual se dan una serie de especificaciones, dentro de las cuales se nombra el equipo al que está integrado.

Toda la información inicial de un SCADA se almacena en el módulo de Configuración, por lo cual este sería nuestra fuente de datos ideal para la generación automática del catálogo de activos.

Luego de seleccionar una muestra de 2 instalaciones en las cuales se emplea el SCADA Guardián del ALBA (GALBA), cada una con una norma distinta definida, se definieron diccionarios en la herramienta para la generación automática de catálogos. El módulo de Configuración del GALBA usa PostgreSQL como sistema gestor de base de datos. A continuación se muestran para cada entrada de datos la salida obtenida.

Diccionario número 1 para equipos industriales de PDVSA.

{PT: TANQUE, PV: VALVULA, PB: BOMBA, PD: DEPURADOR, PS: SEPARADOR}

Fragmento del archivo XLS generado por la solución a partir de la base de datos de Configuración según una norma ficticia:

	A	B	C	D
1	ETIQUETA	VALOR		
2	PT-1200A	TANQUE DE RESERVA 1200 T		
3	PT-1200B	TANQUE DE ALMACENAMIENTO 1200 T		
4	PT-500X	TANQUE DE PRUEBA DE AGUA		
5				

Figura 5. Fragmento del archivo Excel generado para el diccionario 1.

Diccionario número 2 para equipos industriales de PDVSA.

{PT: TANQUE, PV: VALVULA, PB: BOMBA, PD: DEPURADOR, PS: SEPARADOR}

Fragmento del archivo XLS generado por la solución a partir de una colección de puntos seleccionados de la base de datos de Configuración de una instalación del oriente de Venezuela:

	A	B	C	D	E	F
1	ETIQUETA	VALOR				
2	PV_550A	VALVULA 550 DE CONTROL DE BOMBEO A PATIO DE TANQUE				
3	PV_650A	VALVULA 550 DE CONTROL DE BOMBEO A PATIO DE TANQUE				
4	PV_1000V	VALVULA 1000 A VENTEO 1				
5	PV_2000V	VALVULA 2000 A VENTEO 1				
6	PV_3000V	VALVULA 3000 A VENTEO 1				

Figura 7. Fragmento del archivo Excel generado para el diccionario 3.

Las listas de activos resultantes demuestran la posibilidad de obtener los equipos inmersos en el proceso de producción a partir de la base de datos de módulo de Configuración del SCADA GALBA, lo cual soluciona el problema de gestionar la información requerida para cada activo. Los catálogos generados se insertan al Repositorio de Información Unificada (base de datos de activos del SCIA) teniendo en cuenta la preferencia de ubicación del usuario.

3.1 RESULTADOS EXPERIMENTALES

A la solución también se le realizó una prueba experimental en la que se tomó otra muestra de bases de datos de empresas ajenas. Igualmente se aplicó el método de generación automática de catálogo de activos y el resultado obtenido fue similar al anterior. Sendas empresas usan Microsoft SQL Server como sistema gestor de bases de datos.

Gracias a la flexibilidad que aporta la definición de diccionarios para establecer criterios personalizados y el uso del ORM para abstraer el resto del sistema del acceso a los sistemas de almacenamiento, la solución propuesta es aplicable en varios escenarios.

Diccionario número 3 sobre vehículos de una empresa de transportación de la provincia de La Habana:

{B004: AUTO, B007: AUTO, B003: OMNIBUS}

Fragmento del archivo XLS generado por la solución a partir de la base de datos de la empresa de transportación. De la misma se toma como norma el estándar definido para identificar vehículos en el país:

	A	B	C	D	E	F	G
1	ETIQUETA	VALOR					
2	B0030489	OMNIBUS RUTERO 3909, RUTA 490A, PARADERO LA LISA					
3	B0039021	OMNIBUS RUTERO 0540, RUTA P5, PARADERO SAN AGUSTIN					
4	B0030037	OMNIBUS RUTERO 1304, RUTA P13, PARADERO SANTIAGO DE LAS VEGAS					
5	B0033094	OMNIBUS ESCOLARES 3404, PARADERO SAN ANTONIO					
6	B0034408	OMNIBUS ESCOLARES 2990, PARADERO SAN ANTONIO					

Figura 7. Fragmento del archivo Excel generado para el diccionario 3.

4. CONCLUSIONES

Se mejora el proceso de captura de información de los activos para la realización del ciclo de confiabilidad operacional.

Se crea una abstracción para la extracción de los datos requeridos, extendiendo la portabilidad de los mismos a distintos sistemas gestores de bases de datos como son: PostgreSQL, MySQL, Microsoft SQL Server, SQLite.

El uso de los diccionarios permite personalizar tanto las consultas como la técnica de clasificación, lo cual hace extensible el uso de la solución.

5. REFERENCIAS

- Acosta Gonzalez, Riolvi (2010). "Herramienta de simulación de Montecarlo para el Sistema de Confiabilidad Integral de Activos para los pueblos del ALBA". La Habana: Universidad de las Ciencias Informáticas, Cuba 2010.
- Bermejo, Sergi (2000). Curso de Redes Neuronales Artificiales. Curso Escolar 2000-2001.
- Ceballo Gastell, Daimerys. De Diego Ceruto, Yanet del Carmen (2008). "Categorización de texto usando Redes Neuronales Artificiales". La Habana: Universidad de las Ciencias Informáticas. Cuba 2008.
- Corrales Estrada, Elisandra (2012). "Minería de datos aplicada al sistema AiresProxy". La Habana: Universidad de las Ciencias Informáticas, Cuba 2012.
- Cross, Mark (2001). "Decision support systems: using technology for succesful management". s.l.: Hamilton, p. 48. Vol. 75.
- Kerton, Luanner (2012). "Sistema para la toma de decisiones sobre datos parametrizados". La Habana: Universidad de las Ciencias Informáticas, Cuba 2012.
- Maneiro, Mariela Yanina (2008). "Minería de Datos". Facultad de Ciencias Exactas, Naturales y Agrimensura. Universidad Nacional del Nordeste.
- Martín Prado, Yoel (2010). "Desarrollo de la herramienta para la Política de Cuidado Integral de Activos del Sistema de Confiabilidad Integral de Activos". La Habana: Universidad de las Ciencias Informáticas, Cuba 2010.
- Nuevo León, Heidy Alina (2009). "Desarrollo del flujo de requisitos para la Metodología Análisis de Criticidad Integral de Activos para el producto SCIA". La Habana: Universidad de las Ciencias Informáticas. Cuba 2009.
- Rizo Constanten, Rosa María (2012). "Método para la selección y clasificación de estudiantes basado en reglas inteligentes". XI Jornada Científica Estudiantil. Universidad de las Ciencias Informáticas. Cuba, 2012.
- Ultsch, Alfred. (1990). "Kohonen's Self Organizing Feature Maps for Exploratory Data Analysis". Paris: Proceedings of the International Neural Network Conference, 1990. ISBN 978-0-7923-0831-7 (0-7923-0831-X).
- Yin, Hujun (2007). "Learning Nonlinear Principal Manifolds by Self-Organising Maps". Berlín: Springer, 2007. ISBN 978-3-540-73749-0.

Authorization and Disclaimer

Authors authorize LACCEI to publish the paper in the conference proceedings. Neither LACCEI nor the editors are responsible either for the content or for the implications of what is expressed in the paper.