

Generación de Datos Sintéticos para Arquitecturas de Procesamiento de Datos del Internet de las Cosas

¹Alfonso Garcia-de-Prado, ¹Guadalupe Ortiz, ²Juan Hernández, ²Enrique Moguel

¹Grupo de Ingeniería del Software UCASE
Escuela Superior de Ingeniería, Universidad de Cádiz
Avda. de la Universidad de Cádiz, 10. 11519 Puerto Real, Cádiz, España
{alfonso.garciadeprado, guadalupe.ortiz}@uca.es

²Grupo Quercus de Ingeniería del Software
INTIA, Universidad de Extremadura
Av. de la Universidad, S/N, 10003 Cáceres, España
{juanher, enrique}@unex.es

Resumen. La vertiginosa evolución del Internet de las Cosas, sumada a las grandes cantidades de datos heterogéneos que fluyen por los sistemas de información, han dado lugar a diversas plataformas software que analizan dichos datos con el objetivo de mejorar la toma de decisiones. Estas plataformas requieren de una prueba en materia de eficacia y eficiencia antes de su puesta en producción; para ello requieren de grandes cantidades de datos del dominio tecnológico y de aplicación en cuestión. Con este fin se implementa nTROGEN: un generador de datos sintéticos para el IoT que cubre las necesidades de estos sistemas.

Palabras clave: Arquitecturas Orientadas a Servicios, Internet de las Cosas, Consciencia del Contexto, Generador de Datos Sintéticos

1 Motivación

El término Internet de las Cosas (*Internet of Things* (IoT)) fue acuñado por Kevin Ashton en 2009 en el ámbito de la cadena de suministros [1]; desde entonces se ha aplicado en muchos otros dominios y en un sentido amplio, refiriéndose a dispositivos conectados a la red que pueden obtener y compartir información a través de diversas plataformas proporcionando valor añadido a las aplicaciones. La rauda evolución del IoT, junto con una mayor implementación de sistemas de información, ha favorecido la generación de grandes cantidades de datos. Se acuña así el término *Big Data*, refiriéndose a grandes cantidades de datos heterogéneos que fluyen por los sistemas y que se analizan con el objetivo de mejorar la toma de decisiones [2]. Es más, la información contextual es crucial para la toma de decisiones [3]; en el ámbito de IoT, donde se obtienen datos de diversas fuentes en el contexto del dominio de aplicación, la consciencia del contexto es fundamental, y así lo identifica la Unión Europea en sus retos de Horizonte 2020 [4].

Por otra parte, el IoT ha evolucionado en un entorno donde la integración de servicios juega un papel muy relevante [5]. Las arquitecturas orientadas a servicios (*Service Oriented Architecture* (SOA)) suponen un soporte clave para el IoT [6–8], facilitando

la inclusión de datos que provienen de diversos dispositivos y facilitando el intercambio de datos entre los componentes del sistema. Para probar la eficiencia y eficacia de dichas arquitecturas previo a su puesta en producción, es necesario un simulador capaz de generar datos de acuerdo a la heterogeneidad de datos y protocolos del dominio y arquitectura en cuestión, así como a la rapidez en función de las necesidades de tiempo real del sistema. La mayoría de las herramientas en este ámbito se centran solo en pruebas funcionales [9]; para suplir esta falta se propone e implementa nITROGEN (*Internet of Things RandOm GENerator*): un generador de datos sintéticos para el IoT.

2 Descripción de nITROGEN

nITROGEN se puede usar por línea de comandos o con su interfaz gráfica. Definiremos uno o más ficheros de configuración que posteriormente lanzamos para la generación de datos. En la configuración especificamos una serie de *grupos* que marcan la frecuencia con la que se enviarán los datos a cada uno de los conectores de este grupo. Actualmente se implementan los siguientes conectores, en constante ampliación y mejora: fichero, plataforma ThingSpeak, protocolos AMQP, MQTT y TCP y *TestSpeed* (prueba la velocidad a la que se generan los datos, sin considerar el tiempo de transmisión).

En relación a la periodicidad, se puede expresar en tics por segundo, tics por minuto, milisegundos por tics, o establecer que genere los datos a la máxima velocidad soportada por el equipo. Además, se puede indicar si la generación de datos es secuencial o usar uno o varios niveles de paralelismo (por grupo, por conector o por dato generado).

Los datos a generar se especifican como una cadena de texto seguida de una serie de valores generados conforme a la definición del fichero de configuración, donde pueden definirse distintas distribuciones de probabilidad. Los datos generados pueden ser locales —específicos de un conector y no compartidos—, de grupo —específicos de un grupo y por tanto compartidos entre todos los conectores del grupo—, globales —se generan al comienzo de la simulación y son compartidos por todos los grupos de la configuración— y de repetición —permite definir un solo conector actuando como si se hubiesen definido varios del mismo tipo—. El tipo de datos a generar pueden ser:

- Numérico: se pueden generar de forma aleatoria, secuencial o especificando tendencias, permitiendo cada método diversas opciones para facilitar la mayor correspondencia con el mundo real o el escenario que queramos simular.
- Cadenas aleatorias: de longitud fija o variable, y porcentaje de mayúsculas, minúsculas, números y símbolos.
- Elementos de una serie: dada una serie definida por el usuario se generan elementos de la serie susceptibles de configuración (elemento inicial, final, paso, orden, etc.)
- Rangos temporales: permite generar fechas o periodos de tiempo en diversos formatos, pudiéndose especificar rangos de fechas u horas, incrementos o decrementos, con paso fijo o variable, con inicio fijo o en el instante de generación, etc.

En relación a la velocidad de generación; con un i3-2 de 8GHz, 8GB RAM y Windows 10, genera una media de 55.000 datos por segundo. A lo que habría que descontar los tiempos consumidos en conexiones y transferencias en función del protocolo y la red.

3 Conclusiones

Hemos presentado nITROGEN, un generador de datos sintéticos para el IoT, que permite probar la eficiencia y eficacia de las arquitecturas que tratan con dichos datos. nITROGEN se encuentra en continua mejora con el fin de dar servicio a la comunidad* para la prueba de sus arquitecturas; en sus versiones preliminares facilitó con éxito de las pruebas de rendimiento y estrés de CARED-SOA [7] y COLLECT [8]. En nuestro trabajo futuro más próximo estamos abordando que varias instancias de nITROGEN trabajen de forma colaborativa y ofrecer un soporte de ayuda contextual y/o en línea.

Agradecimientos

Trabajo financiado por MINECO/FEDER (TIN2015-65845-C3-3-R, TIN2016-81978-REDT, TIN2015-69957-R), por el programa POCTEP 2014-2020 (0045-4IE-4-P) y por el Programa de Fomento e Impulso de la Actividad Investigadora de la Univ. de Cádiz.

Referencias

1. Ashton, K.: That “Internet of Things” Thing. *RFID J.* 22, 97–114 (2009).
2. Bumblauskas, D., Gemmill, D., Igou, A., Anzengruber, J.: Smart Maintenance Decision Support Systems (SMDSS) based on corporate big data analytics. *Expert Syst. Appl.* 90, 303–317 (2017).
3. Cheng, B., Wang, M., Zhao, S., Zhai, Z., Zhu, D., Chen, J.: Situation-Aware Dynamic Service Coordination in an IoT Environment. *IEEEACM Trans. Netw.* 25, 2082–2095 (2017).
4. The Alliance for Internet of Things Innovation: Internet of Things Applications, <https://ec.europa.eu/digital-single-market/en/news/aioti-recommendations-future-collaborative-work-context-internet-things-focus-area-horizon-2020>, (2015).
5. Khodadadi, F., Dastjerdi, A.V., Buyya, R.: Internet of Things: an overview. In: *Internet of Things*. pp. 3–27. Elsevier (2016).
6. Issarny, V., Bouloukakakis, G., Georgantas, N., Billet, B.: Revisiting Service-Oriented Architecture for the IoT: A Middleware Perspective. In: Sheng, Q.Z., Stroulia, E., Tata, S., and Bhiri, S. (eds.) *Service-Oriented Computing*. pp. 3–17. Springer International Publishing, Cham (2016).
7. Garcia De Prado, A., Ortiz, G., Boubeta-Puig, J.: CARED-SOA: A Context-Aware Event-Driven Service-Oriented Architecture. *IEEE Access.* 5, 4646–4663 (2017).
8. Garcia-de-Prado, A., Ortiz, G., Boubeta-Puig, J.: COLLECT: COLLaborativE Context-aware service oriented architecture for intelligent decision-making in the Internet of Things. *Expert Syst. Appl.* 85, 231–248 (2017).
9. Völker, M., Mandal, S., Hewelt, M.: Testing Event-driven Applications with Automatically Generated Events. In: *BPM (Demos)* (2017).

* Desde <https://ucase.uca.es/nITROGEN/> puede obtenerse la herramienta así como tutoriales, ejemplos de uso y pruebas realizadas.