

Estimación de KQIs basada en trazas móviles de la capa de aplicación

Alejandro Moreno-Sancho⁽¹⁾, Sergio Fortes⁽¹⁾, Eduardo Baena⁽¹⁾, Francisco Pareja⁽²⁾, Raquel Barco⁽¹⁾

⁽¹⁾Instituto de Telecomunicación (TELMA), Universidad de Málaga, CEI Andalucía TECH

E.T.S. Ingeniería de Telecomunicación, Bulevar Louis Pasteur 35, 29010 Málaga (España)

{aams, sfr, ebm, rbm}@ic.uma.es

⁽²⁾Metricell Limited

Francisco.Pareja@metricell.com

Resumen—The increasing complexity of mobile networks, even more with the development of 5G, makes it no longer possible to relate the traditional low-layer radio metrics used until now, with the real quality of the user. This measures known as KPI (Key Performance Indicator) are easy to obtain in comparison with KQI (Key Quality Indicator), but it is the latter, the ones which give a real approach to the quality of the user. The problem with KQI is the difficulty of obtaining them, they are time and battery consuming to be used in mobile devices, as most of them require downloading and uploading files, HTTP requests, video streaming and others. Then, it is proposed and evaluated in a real cellular network, a complete framework to estimate KQI using KPIs as inputs.

I. INTRODUCCIÓN

Con la comercialización de LTE y el temprano despliegue de las tecnologías 5G, junto con la explosión de los servicios “Over-The-Top” (OTT), encontrar formas de gestionar y optimizar las redes es cada vez más complejo, y las técnicas tradicionales, ya anticuadas, son insuficientes para la tarea. Hasta ahora, los ingenieros han tomado medidas, normalmente mediante “Drive Test” (toma de mediciones en un vehículo pilotado), basadas en las quejas de los clientes. Este enfoque es ineficaz dada la rapidez de los cambios en la red. No solo es un problema su ineficacia, además, las mediciones que se realizan, normalmente de capas bajas, como la potencia o la calidad de la señal, no son representativas de la calidad real del usuario debido a la complejidad de las redes actuales y los procesos implicados.

El uso generalizado de los “smartphones” y el número de sensores que contienen permite que estas mediciones se realicen constantemente mediante aplicaciones instaladas en el dispositivo, obteniendo una información que, a diferencia de los Drive Test, no es puntual y se basa en los problemas detectados por el usuario. Sin embargo, aunque se ha mejorado el problema de los Drive Test, estas mediciones, normalmente de capas bajas y parámetros de configuración de red, junto con información del propio dispositivo como el sistema operativo, la ubicación, la batería u otros, siguen sin ser representativas de la calidad real del usuario. Estos indicadores, normalmente denominados KPIs, son fáciles de recoger y no requieren un alto consumo de batería, mientras que los KQIs, que se aproximan a la calidad real del usuario, requieren la puesta en marcha de experimentos normalmente relacionados con la descarga y subida de archivos, peticiones HTTP o streaming de vídeo, que consumen mucha batería y tiempo. Por lo tanto, una solución que utilice la Inteligencia Artificial capaz de

mapear el KQI a partir de los KPIs permitiría obtener las ventajas de ambos indicadores.

El uso de Inteligencia Artificial para optimizar y organizar redes móviles está en auge y se ha utilizado ampliamente en temas relacionados con la detección de interiores [1] [2], la localización [3] o la detección de agujeros de cobertura [4].

En la literatura se encuentran ejemplos de este mapeo de indicadores [5] [6], sin embargo, siempre se realiza sobre mediciones recogidas a través de simulaciones o en entornos reales pero controlados en los que normalmente se utiliza un solo dispositivo para realizar las mediciones. Aunque hay ejemplos en los que las mediciones se realizan mediante diferentes dispositivos y en un amplio rango espacial, normalmente no se trata de predecir el KQI como un valor continuo, se suele discretizar entre diferentes rangos o incluso si el valor es “bueno” o “malo” como se hace en [6]. Además, estos estudios suelen centrarse en indicadores de calidad específicos (por ejemplo, ping, HTTP).

Hasta el actual conocimiento de los autores, aún no se ha propuesto un sistema completo de análisis, preprocesamiento de datos y predicción para diversos indicadores de calidad (por ejemplo, tiempo de descarga de archivos, tiempo de ping, tiempo de solicitud HTTP) a partir de los KPI.

Este sistema tiene valor por sí mismo debido al estudio estadístico de los datos que se ha realizado, pero también tendrá implicaciones para la renovación de las anticuadas estrategias de organización y optimización de la red. En cuanto a la detección de agujeros de cobertura, donde normalmente se utilizan mapas de cobertura de parámetros radioeléctricos [7], pueden ser sustituidos por mapas de calidad de usuario, haciendo más eficientes las futuras estrategias de colocación de antenas. Dado que los sistemas de recogida de muestras no proporcionan un mapa de calidad continuo, se pueden utilizar técnicas de interpolación espacial como el Kriging bayesiano para obtener mapas completos [7]. Esto reduce considerablemente el uso de pruebas de conducción.

Además, el uso de sistemas más avanzados basados en Inteligencia Artificial puede sustituir a la tradicional, ineficiente e inexacta detección de agujeros basada en umbrales [8].

La sección II introduce el sistema propuesto, explicando en subapartados cada una de las secciones de dicho sistema, específicamente el tipo de trazas obtenidas, preprocesado y selección de variables, imputado de datos, normalización y mecanismos de regresión. Posteriormente en la sección III se discute sobre la evaluación del sistema. Finalizando con el apartado IV, se realiza una conclusión y se indican las líneas

futuras.

II. SISTEMA PROPUESTO

El sistema propuesto, representando en la Fig. 1 comienza por la obtención de las trazas móviles. Estas son tomadas de diferentes dispositivos en los que se ha añadido una aplicación para este propósito. La información, formada por una combinación de mediciones de capa baja, sistema operativo y experimentos, se recoge en una base de datos.

A. Trazas móviles de la capa de aplicación

Estos datos se basan en la minimización de los Drive Test (Minimization of Drive Test, MDT), fue introducida por la versión 10 del 3GPP para proporcionar a los operadores herramientas de optimización del rendimiento de la red de forma eficiente en cuanto a costes.

En comparación con los Drive Test, el uso de indicadores tiene algunas ventajas relacionadas con la precisión, el coste y el tiempo, entre otras. Los Drive Test normalmente no son representativos de la calidad del usuario, el 85 por ciento de las llamadas se realizan desde el interior de los edificios, incluso si la carretera donde se toman las medidas está muy cerca de un edificio, las medidas no serán tan precisas como las que se tomarían desde un móvil dentro del edificio. La necesidad de conducir el coche en el que se toman las medidas hace que todo el proceso sea lento y con altos costes.

Las trazas móviles de la capa de aplicación se basan en el sistema operativo (S.O.), los datos dinámicos se combinan con los datos de configuración de la red. El hecho de estar basado en el sistema operativo hace posible añadir una rica información contextual que carecen los Drive Test, estimación de interior/exterior, estimación de altitud, actividad del usuario, localización GPS...

B. Preprocesado de datos y selección de variables

El marco propuesto se encarga de leer esta información para separarla en KPI (entradas) y KQI (salidas), se realizan estudios estadísticos sobre los datos para eliminar posibles fallos que puedan ser causados por la aplicación o el propio sistema operativo. El sistema toma en consideración otros factores como los valores atípicos, que podrían reducir la precisión de los resultados finales, por lo que las muestras que no son estadísticamente relevantes se eliminan del KQI. Se tienen en cuenta otros procesos, como el uso de mecanismos de selección de variables (Feature Selection) que permiten conocer el efecto de las variables de entrada con respecto a las variables de salida. En un principio se utilizan métricas como el coeficiente de correlación de Pearson para establecer el nivel de relación entre las entradas y la salida que se pretenden predecir. Posteriormente se hace uso de otras técnicas más elaboradas convirtiendo el coeficiente de correlación de Pearson en una "F-score" y luego en "P-value", o utilizando la información mutua (Mutual Information, MI) para medir la dependencia entre las variables.

C. Imputado de datos

Al igual que hay valores que pueden ser eliminados por ser estadísticamente insignificantes o ilógicos por su naturaleza ("Reference Signal Received Power", RSRP inferior a -100 dBm o "Reference Signal Received Quality", RSRQ inferior

a -20 dB), hay otro tipo de valores que indican que ha habido un fallo con la muestra. Estos valores, generalmente llamados valores NaN (Not a Number) o Null, pueden causar verdaderos problemas ya que muchos mecanismos de aprendizaje automático no permiten la entrada de este tipo de datos, por lo que deben ser eliminados antes de entrenar los modelos.

Se denomina "muestra" a la composición de unos inputs (RSRQ, RSRP, timestamp, latitud, longitud, os_version...) y el output (download_time_2_mb, ping_latency_time, ping_jitter_time u otros). El problema es que estos valores NaN son aleatorios, y si hay una sola entrada no válida, hay que eliminar la muestra completa. Si se utilizan los datos en su totalidad no es un problema, pero reduciendo el rango por latitud y longitud con motivo de buscar correlaciones más altas puede que no se obtengan suficientes muestras para entrenar un modelo. Utilizando las técnicas de selección de variables mencionadas anteriormente, se pudo comprobar que si se reducía la región de los datos en uso, la relación entre las variables tendía a mejorar. Esta componente espacial se redujo creando una rejilla hexagonal con la que fue posible encontrar zonas de alta densidad de muestras de diferentes resoluciones, la resolución indica el tamaño del hexágono, a mayor resolución, menor tamaño del hexágono. Después, se aplica una técnica de clustering al conjunto de datos restante para separar por latitud y longitud grupos de datos que puedan comportarse de forma similar. En este caso se utiliza DBSCAN [9] (Density-Based Spatial Clustering of Applications with Noise) que es un algoritmo no paramétrico de clustering basado en la densidad, en este caso espacial, que agrupa las muestras que están próximas entre sí con la posibilidad de marcar valores atípicos (puntos negros en la Fig. 2.). Por último y sólo para la representación se realiza un "Convex-Hull" para representar la región de cada cluster.

Tras realizar esta aproximación espacial, se debe hacer frente al problema de no obtener suficientes datos mencionado anteriormente. Se buscan alternativas, como reducir la resolución de la rejilla para obtener un mayor número de muestras o tomar datos que pertenezcan sólo al mismo CGI (Cell Global Identity, concatenación de MCC-MNC-LAC-CID) o al mismo TAC (Tracking Area Code). Dado que no se obtienen suficientes datos, se utilizan técnicas de imputado de datos para rellenar las entradas vacías y no descartar las muestras completas. Las técnicas simples utilizan la imputación de constantes normalmente estadísticas como la media o la mediana, la ventaja de estos métodos sobre otros es que sólo necesitan los datos de una entrada y no el resto de las entradas para realizar el imputado. Sin embargo, técnicas más avanzadas, como los mecanismos de aprendizaje automático no supervisado, pueden proporcionar un mejor resultado utilizando el resto de las entradas. Son especialmente interesantes las técnicas como "K-Nearest Neighbors" (KNN), en las que los valores se imputan utilizando el valor medio de los k vecinos más cercanos encontrados en el conjunto de entrenamiento, y "Multivariate Imputation", que es una estrategia para imputar los valores perdidos utilizando cada variable con valores perdidos en función de las otras variables de forma iterada con una planificación "Round-Robin".

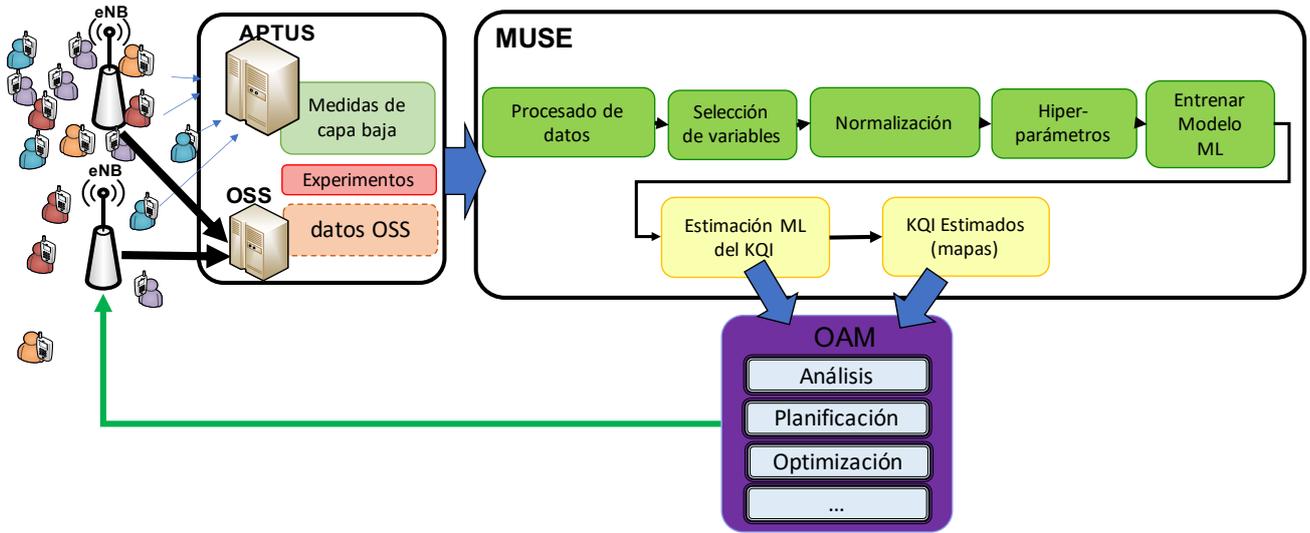


Fig. 1. Esquema propuesto para el sistema de predicción de KQI.

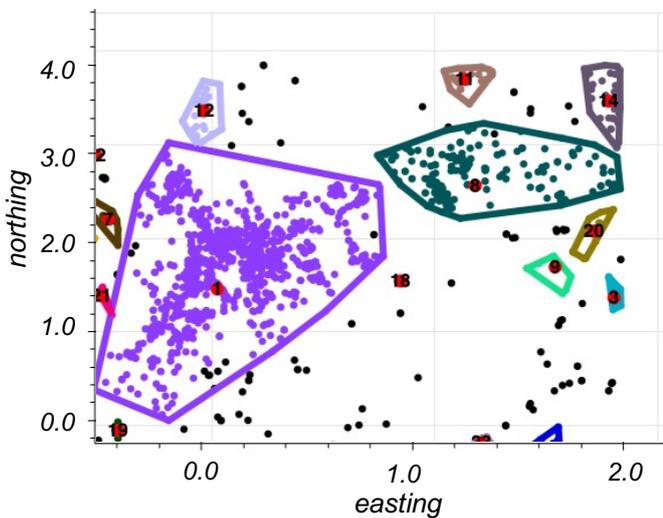


Fig. 2. Separación de los datos por latitud y longitud usando la técnica de clusterización DBSCAN.

D. Normalización

Cuando se comparan variables con unidades diferentes, puede ocurrir que estas variables no contribuyan de forma equilibrada al análisis, lo que produce cierto sesgo para las técnicas de aprendizaje automático en las que se tiene en cuenta la distancia entre muestras, como KNN, K-means y SVM, siendo la distancia euclidiana la más utilizada en estos casos. Aunque para otros algoritmos, en particular los basados en árbol de decisiones, implementados de forma que se trata de buscar la mejor partición en función de cada variable, la normalización no tiene ningún efecto, pero hay ciertos tipos que pueden beneficiar el aumento de la precisión en términos de error al utilizarlos. La normalización de las variables también puede provocar un aumento de la velocidad de entrenamiento de los modelos.

Entre algunas de las técnicas de normalización existentes, se opta por utilizar la transformación por cuantiles (Quantile Transformation), ya que es apreciable una mejora en la preci-

sión de los modelos, independientemente de que estén basados en árbol de decisiones, esto no ocurre con otras técnicas de escalado y normalización, como el escalado entre un mínimo y máximo (típicamente entre 0 y 1), la normalización estándar (Z-score) u otros métodos de normalización robustos.

La transformación por cuantiles es un tipo de esquema de preprocesamiento robusto en el que se reduce el impacto de los valores atípicos, los valores más frecuentes se distribuyen de forma que la variable sigue una distribución uniforme o normal en este caso. Esta transformación puede distorsionar las correlaciones lineales entre variables con la misma escala, pero hace que las variables con escalas diferentes sean más comparables.

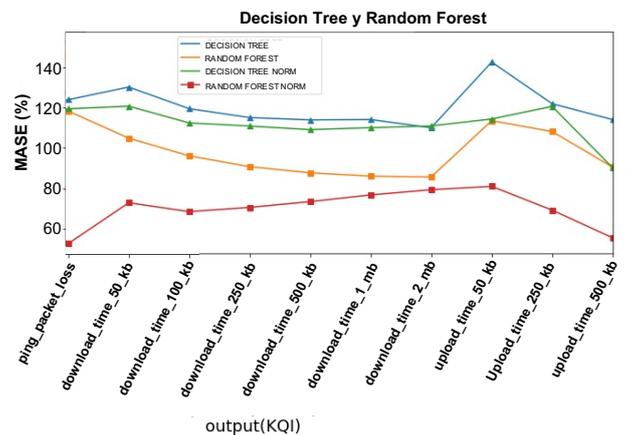


Fig. 3. Efectos de la normalización "Transformación por cuantiles" en KQIs cuando usamos Random Forest y Decision Tree Regressor. MASE mayor a 100 % indica peor comportamiento que imputando la media.

Utilizando el conjunto de datos completo, se observa en la Fig. 3 una mejora en términos de MASE (Mean Absolute Scaled Error) en todos los KQIs tanto para Decision Tree como para Random Forest. También se observa que se obtienen resultados considerablemente mejores con este último mecanismo de regresión.

E. Regresión

Como mecanismos de regresión se utilizan algoritmos comunes utilizados en otros problemas similares [5], añadiendo mecanismos más complejos como Random Forest o el Perceptrón Multicapa, un modelo de red neuronal supervisada.

Se puede comprobar que no existe una relación lineal entre las variables de entrada y salida mediante análisis visuales como los gráficos de dispersión y valores estadísticos como el coeficiente de correlación de Pearson. Esto indica que ciertos métodos como la regresión lineal o la regresión lineal por pasos pueden no ser muy eficaces en las predicciones y que, sin embargo, otros sistemas capaces de detectar relaciones no lineales como Decision Tree Regressor (DTR), Random Forest (RF) o el Perceptrón Multicapa (Multi-Layer Perceptron, MLP) probablemente sean más adecuados. Un primer análisis refleja una clara confirmación de la hipótesis anterior, los mecanismos basados en árboles de decisión se comportan mejor en términos de precisión en las métricas de error que los lineales, con lo que se centrará el estudio en los primeros a la hora de evaluar el sistema propuesto.

III. EVALUACIÓN

Las muestras obtenidas son de un amplio rango temporal y espacial, además, al estar recogidas de diferentes dispositivos la correlación entre las variables obtenidas tiende a ser pequeña. La cantidad de muestras utilizada, 400 millones, se reduce si solo se consideran las muestras que tienen experimento (250000 muestras), aunque pueda parecer un número elevado, cuando se tiene en consideración filtros espaciales o temporales con el objetivo de entrenar modelos que ofrezcan mejores resultados, el número de muestras es insuficiente. A pesar de esto, con métodos como Decision Tree o Random Forest se obtienen unos resultados considerablemente mejores de los que se conseguirían si se realizara imputación con la media de los datos, ya que el error, como se aprecia en la Fig. 3, es menor que 100%. Esto es lo que representa la métrica MASE utilizada, que ha sido adaptada para ser usada con datos que no son series temporales.

$$q_j = \frac{e_j}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|} \rightarrow q_j = \frac{e_j}{\frac{1}{N} \sum_{i=1}^N |y_i - \bar{y}|} \quad (1)$$

Estas medidas de error normalizadas y relativas se usan para calcular el MAE (Mean Absolute Error), obteniendo así una medida que puede ser comparada entre diferentes mecanismos de regresión y KQIs de una manera sencilla e intuitiva.

En entornos donde la correlación entre variables es reducida, lo que implica casi aleatoriedad, conseguir resultados mejores al imputado de la media es complejo, sin embargo, podemos ver que mediante el uso de las técnicas de regresión mencionadas y los mecanismos de preprocesamiento de datos, entre ellos, recalcar la mejora ofrecida por la normalización, se obtienen muy buenos resultados en torno al 70 y 80 por ciento, Fig. 3. El comportamiento descrito ocurre con todos los KQIs, aunque son los de subida de ficheros de 250kb y 500kb con los que se aprecia una mejor precisión.

IV. CONCLUSIÓN Y LÍNEAS FUTURAS

El gran efecto de las técnicas de preprocesamiento (imputación de datos, normalización y análisis espacial) provoca

una mejora considerable en la correlación entre las variables de entrada y salida, así como en los resultados de predicción obtenidos. Aunque los resultados podrían considerarse imprecisos debido al bajo R-cuadrado obtenido, el error alcanzado es relativamente bajo. Además, se ha obtenido un rico análisis estadístico que permite observar el efecto de estas técnicas de preprocesamiento de datos junto con el análisis espacial. Cabe destacar que, hasta donde los autores conocen, no se han realizado antes predicciones de este tipo, ya que debido a la intrincada cantidad de procesos que existe en el trasfondo de una muestra, la complejidad de la red es notablemente alta, esto implica que normalmente la correlación entre las variables sea demasiado baja para considerar el uso de mecanismos de Machine Learning. En futuros estudios donde la densidad de muestras no sea un inconveniente, se pretende profundizar en el análisis de los efectos en estas regresiones, siguiendo las técnicas de separación espacial mencionadas anteriormente, así como el posible uso de técnicas de Machine Learning más avanzadas como de redes neuronales más complejas.

AGRADECIMIENTOS

Este trabajo ha sido financiado por la Junta de Andalucía y FEDER en el marco del proyecto MUSE - “Massive User Experience Assessment and Prediction for Mobile Networks” (Ref. UMA-CEIATECH-13, “Proyecto singular de actuaciones de transferencia del conocimiento Campus Excelencia Internacional Andalucía TECH. Ecosistema innovador con inteligencia artificial para Andalucía 2025”) y beca postdoctoral (Ref., DOC_01154, “selección de personal investigador doctor convocado mediante Resolución de 21 de mayo de 2020”, PAIDI 2020). También ha sido parcialmente financiado por la Universidad a través del I Plan Propio de Investigación y Transferencia de la Universidad de Málaga.

REFERENCIAS

- [1] J. Bejarano-Luque, M. Toril, M. Fernández-Navarro, R. Acedo-Hernández, and S. Luna-Ramírez, “A Data-Driven Algorithm for Indoor/Outdoor Detection Based on Connection Traces in a LTE Network,” *IEEE Access*, May 2019.
- [2] I. Saffar, M.-L. Alberi-Morel, K. Deep Singh, and C. Viho, “Machine Learning with partially labeled Data for Indoor Outdoor Detection,” *16th IEEE Consumer Communications Networking Conference*, no. pp.1-7, 2019. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-02011454>
- [3] N. Kanazawa, A. Nagate, and A. Yamamoto, “Field experiment of localization based on machine learning in lte network,” in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, 2018, pp. 1–6.
- [4] D. Mulvey, C. H. Foh, M. A. Imran, and R. Tafazolli, “Cell Fault Management Using Machine Learning Techniques,” *IEEE Access*, vol. 7, pp. 124 514–124 539, 2019.
- [5] C. Baena, S. Fortes, E. Baena, and R. Barco, “Estimation of Video Streaming KQIs for Radio Access Negotiation in Network Slicing Scenarios,” *IEEE COMMUNICATION LETTERS*, vol. 24, no. 6, March 2020. [Online]. Available: <https://doi.org/10.1109/LCOMM.2020.2979713>
- [6] S. Xie, K. Li, M. Xiao, L. Zhang, and W. Li, “Key Quality Indicators Prediction for Web Browsing with Embedded Filter Feature Selection,” *Applied Sciences*, vol. 10, no. 6, 2020. [Online]. Available: <https://www.mdpi.com/2076-3417/10/6/2141>
- [7] A. Galindo-Serrano, B. Sayrac, and S. Ben Jemaa, “Automated Coverage Hole Detection for Cellular Networks Using Radio Environment Maps,” *9th International Workshop on Wireless Network Measurements 2013*, May 2013.
- [8] A. Gómez-Andrades, R. Barco, and I. Serrano, “A method of assessment of LTE coverage holes,” *EURASIP Journal on Wireless Communications and Networking*, vol. 236, 2016.
- [9] G. Xingyu, Y. Peng, L. Wenjing, and Q. Xuesong, “Clustering-based KPI Data Association Analysis Method in Cellular Networks,” *IEEE/IFIP NOMS 2016 Workshop*, april 2016. [Online]. Available: <https://www.researchgate.net/publication/304796718>