

**XIII Congreso Español
de Metaheurísticas,
Algoritmos Evolutivos y
Bioinspirados
(XIII MAEB)**

MAEB 4: SESIÓN ESPECIAL:
SMART CITIES





Monitorización y clasificación de estancias en edificios empleando captación de comunicaciones inalámbricas de dispositivos inteligentes.

Antonio Fernández Ares
Universidad de Granada
Granada, Spain
antares@ugr.es

María Isabel García Arenas
Universidad de Granada
Granada, Spain
mgarenas@ugr.es

Pedro A. Castillo Valdivieso
Universidad de Granada
Granada, Spain
pacv@ugr.es

Resumen—La monitorización de personas y sus movimientos es un área de interés en las *smart cities* ya que estas enfrentan el problema de la superpoblación. Los actuales sistemas de monitorización de personas, basados en cámaras de vigilancia, son muy intrusivos, atentan a la privacidad y son costosos. En este artículo se propone el empleo de la captación de las comunicaciones inalámbricas emitidas por los dispositivos inteligentes portados por los ciudadanos como fuente de datos para los estudios de la monitorización de personas empleando inteligencia computacional para la extracción de patrones de comportamiento habituales. Para ello, se presenta un caso de uso aplicado a la monitorización del acceso a un edificio público y la obtención de los momentos que garantizan la mayor disponibilidad para emplazar eventos no regulares. Los resultados muestran que la captación de comunicaciones inalámbricas y la aplicación de técnicas de aprendizaje máquina permiten estudiar y extraer conocimiento sobre el los movimientos de las personas.

Index Terms—Ciudades Inteligentes, Monitorización de personas, Comunicaciones inalámbricas, WiFi, Smartphones

I. INTRODUCCIÓN

La monitorización de personas ha cobrado una gran importancia en el entorno de las *smart cities* debido a que muchos de los problemas a los que se enfrentan las ciudades están originados por el aumento significativo de la población, debido a la superpoblación de las áreas metropolitanas. Proveer de fuentes de datos a las ciudades del futuro sobre los movimientos de las personas que viven en ellas permite una mejor gestión de los recursos, lo que repercute positivamente en la calidad de los servicios ofertados. Este mismo principio puede aplicarse a los edificios de las *smart cities*, obteniendo información real sobre como interaccionan las personas con dicho edificio y emplear la información recogida para la extracción de información y conocimiento por medio de la inteligencia computacional.

Sin embargo, las tecnologías para el conteo y monitorización de personas actuales se encuentran bastante limitadas. Los sistemas que son capaces de reconocer individualmente a las personas son intrusivos y muy costosos tanto monetaria como computacionalmente, pues se basan principalmente en sistemas basados en reconocimiento de imágenes capturadas por cámaras. Los sistemas más simples y baratos, como placas de presión, tornos o haces láser, no son capaces de reconocer visitas sucesivas de un mismo individuo.

En este artículo se propone el empleo de la captación de las comunicaciones inalámbricas WiFi emitidas por los dispositivos inteligentes, como los smartphones, portados habitualmente por las personas. La información capturada permite identificar al dispositivo de forma única y unívoca, lo que permite reconocer al mismo dispositivo en futuras visitas. Las capturas sucesivas de un dispositivo en un lugar acotado, permite delimitar temporalmente la estancia de dicho dispositivo, y por tanto a su portador.

Empleando esta fuente de información se puede obtener la cuantía de visitantes así como la duración de cada visita. Mediante el empleo de técnicas de inteligencia computacional, se puede extraer conocimiento sobre el comportamiento y naturaleza de las distintas visitas, permitiendo obtener los hábitos o patrones distintivos de cada tipo de visitante.

Para estudiar el potencial de la fuente de datos se emplea el sistema *Mobywit* para monitorizar una de las entradas de la ETSIT¹ de la Universidad de Granada. Con los datos obtenidos mediante la detección de los dispositivos inteligentes, se realizan estudios que permitan cuantificar el número de visitantes recibidos en la escuela y mediante el empleo de algoritmos de agrupación ser capaces de extraer patrones que definan los comportamientos de los distintos visitantes.

Los resultados muestran que las estancias reincidentes a lo largo de varios días se pueden agrupar empleando algoritmos de *machine learning* o aprendizaje máquina para extraer patrones habituales de comportamiento, patrones que sirven para comprender mejor la ocupación y saturación del edificio en los distintos momentos del día.

El resto del artículo se estructura como sigue. En la Sección II se presenta un breve estado del arte sobre la monitorización de personas. En la Sección III se presenta la metodología del sistema así como el escenario empleado. En la Sección IV se recogen los experimentos realizados, incluyendo análisis de las estancias, así como los resultados del agrupamiento de las estancias. La Sección V debate sobre los resultados obtenidos en los experimentos. Finalmente la Sección VI alega

¹Escuela Técnica Superior de Ingenierías en Informática y Telecomunicaciones

la viabilidad de la captación de comunicaciones inalámbricas WiFi para el estudio de la naturaleza de las estancias de las personas en edificios y su importancia en el escenario de un edificio inteligente dentro de una *smart city*.

II. ESTADO DEL ARTE

Habitualmente la monitorización de personas ha estado asociada a los entornos de marketing y estudios de mercado [1] sin embargo en los últimos años ha cobrado vital importancia en los entornos de las *smart city* [2].

Las tecnologías para el conteo y monitorización de personas han evolucionado desde los sistemas de detección de vehículos [3], resultando en un problema más complejo. Esto se debe principalmente a que las personas no tienen elementos distintivos únicos fácilmente reconocibles por procesamiento OCR² de imágenes, al contrario que los coches con sus matrículas [4] y porque sus desplazamientos no se encuentran acotados por carriles y direcciones prefijados [5] como en los vehículos.

Distintas tecnologías han sido empleadas para la contabilización de personas, como los rayos infrarrojos [6] o las cámaras térmicas [7], [8]. Las imágenes captadas por cámaras de vídeo han sido igualmente empleadas [9], [10], siendo algunos sistemas incluso capaces de reconocer a la persona con tasas de acierto aceptables [11]. Sin embargo, estos sistemas tienen numerosos impedimentos como su alto coste [12] y la vulneración de libertades individuales y privacidad [13].

El empleo de los dispositivos inteligentes para la detección de personas es un área de investigación en auge, debido a que su proliferación en los ciudadanos es significativa. Se estima que en España el 83% de la población posee un smartphone³.

La detección de estos dispositivos por medio de sus comunicaciones Bluetooth ha sido empleada con éxito para la monitorización de personas tanto en pequeños edificios [14], grandes [15] o ante eventos culturales multitudinarios [16]. Esto es posible debido a que el protocolo Bluetooth dispone de mecanismos nativos para la detección de dispositivos cercanos, al estar diseñado para las comunicaciones *device-to-device*⁴.

Las comunicaciones WiFi, en cambio, carecen de mecanismos nativos para la detección de dispositivos, al estar basadas en comunicaciones por infraestructura. Sin embargo, es posible detectar los dispositivos que emplean WiFi capturando los paquetes que emiten [17]. Este tipo de estudios está menos extendido, pero ha sido empleado como pruebas de concepto en escenarios como el Sonar 2015⁵ o la Sagrada Familia⁶.

Desde el ámbito académico, la captación de comunicaciones WiFi para la monitorización de personas ha sido empleada con éxito para medir masas de personas [18], tanto en exteriores [19] como interiores [20].

²Optical Character Recognition o Reconocimiento óptico de caracteres

³Según el último informe anual sobre la sociedad de la información en España elaborado por telefónica <http://www.fundaciontelefonica.com/artecultura/publicaciones-listado/pagina-item-publicaciones/itempubli/483/>

⁴Comunicaciones entre dispositivos sin necesidad de estaciones base.

⁵Por el BSC en el proyecto "We know what you did last Sonar" <http://www.bsc.es/viz/whatyoudid/now.html>

⁶D-lab <https://d-lab.tech/project-1/>

III. METODOLOGÍA

Para la captación de las comunicaciones inalámbricas WiFi se emplea el sistema *Mobywit* [19], que permite identificar cada dispositivo de forma unívoca mediante un hash de la dirección MAC de la interfaz de red WiFi. La detección de un dispositivo en las inmediaciones del nodo es denominada *paso*, al cual se le asigna una ventana temporal con las marcas de tiempo de la primera y última detección consecutiva. En la Figura 1 se puede observar un diagrama del nodo *Mobywit* emplazado. Es esperable que un mismo dispositivo pueda pasar a lo largo del día varias veces por las inmediaciones del nodo. Para cada dispositivo y cada día, se calcula en base a los sucesivos pasos, el instante de tiempo de entrada y salida al centro, siendo el primer y último paso detectado por el nodo.

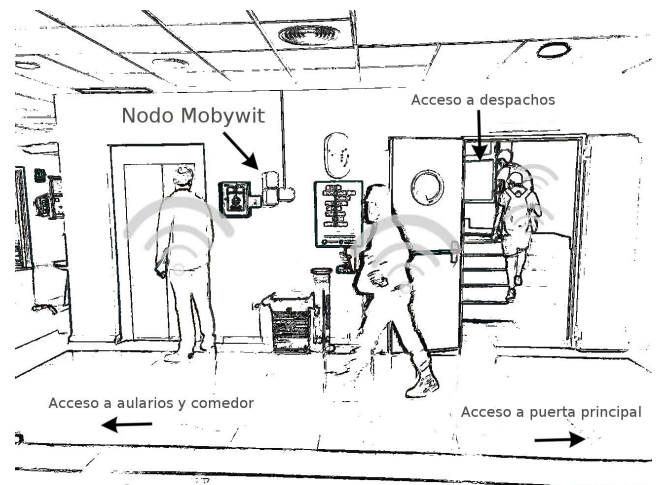


Figura 1. Emplazamiento del Nodo Mobywit en la Escuela. En la parte derecha de la fotografía, fuera de plano, se encuentra la puerta principal. La puerta en plano de la derecha

El escenario de estudio es la Escuela Superior de Ingenierías en Informática y Telecomunicaciones de la Universidad de Granada. Un nodo *Mobywit* fue emplazado en la entrada principal del edificio⁷. Dicha escuela tiene matriculados 2119 alumnos y en ella trabajan unas 250 personas⁸. Las clases tienen una hora de duración, y comienzan a las 8:30.

Durante un periodo de un mes comprendido entre el 3 de Marzo y el 3 de Abril de 2018, incluyendo tanto periodos lectivos como no lectivos, el nodo *Mobywit* capturó 1884219,13MB de tramas WiFi. A partir del procesamiento en tiempo real, el nodo estableció los *pasos* de cada dispositivo por las inmediaciones del nodo.

En base a la información en *pasos* obtenida por el procesamiento de dichas tramas, se propone realizar estudios y aplicar técnicas de aprendizaje máquina para obtener información sobre los hábitos y patrones de comportamiento de los visitantes.

⁷Aunque una puerta secundaria, al ser totalmente exterior, no pudo ser monitorizada al no poder emplazarse el nodo.

⁸Según la memoria académica del curso 2017/2018 https://secretariageneral.ugr.es/pages/memorias/academica/20162017/docencia/centros/_doc/24/!/download



IV. EXPERIMENTOS

Los pasos obtenidos se resumen y contabilizan en base a un intervalo de muestreo. La Figura 2 muestra el resultado de este resumen para tamaños de muestreo de un día y una hora. De esta forma, se puede observar como el periodo vacacional de Semana Santa la afluencia de gente se reduce drásticamente, de igual manera que durante los fines de semana.

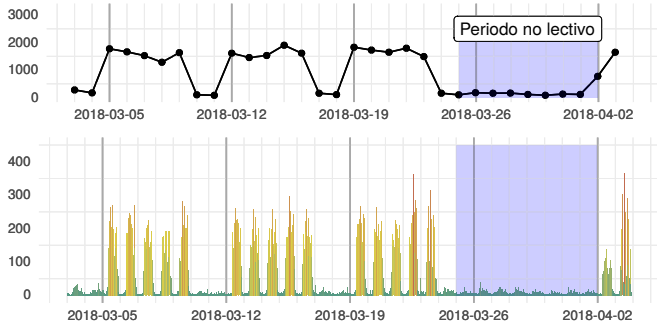


Figura 2. Conjunto de datos obtenidos durante un mes de monitorización. La gráfica superior presenta el número de personas detectadas cada día. La inferior presenta la misma información pero por cada hora.

Esta misma información puede ser mostrada para cada día. La Figura 3 el número de personas detectadas para distintos intervalos de muestreo de un día. Se observa como los cambios de clase afectan al número de personas detectadas.

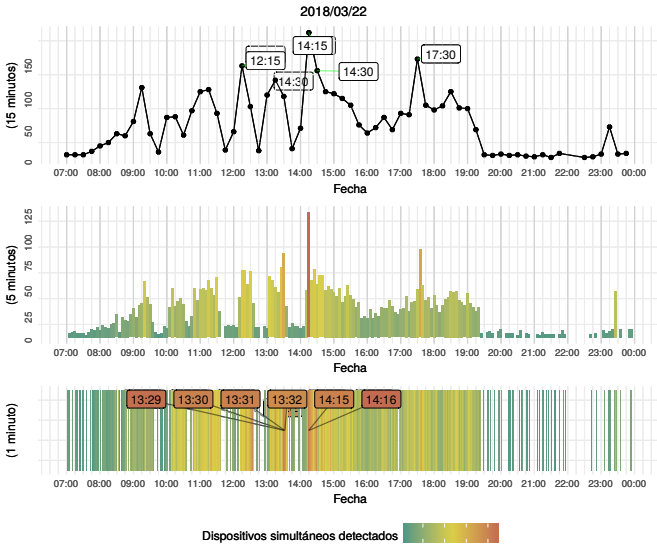


Figura 3. Vista detallada de un día monitorizado, empleando tres tamaños de muestreo: 15 minutos, 5 minutos y 1 minuto. Se han remarcado los valores que han sido considerados anómalos o extremos en las series. Para la ventana de 1 minuto, se emplea únicamente el grado de color como magnitud.

Se espera que cada visitante al edificio registre al menos una entrada y una salida del mismo, siendo registradas ambas en el caso de realizarlas por la puerta monitorizada. Esto permite calcular para cada día el tiempo medio de visitas. La Figura

4 registra la duración de todas las visitas en las que se ha registrado tanto la entrada como la salida.

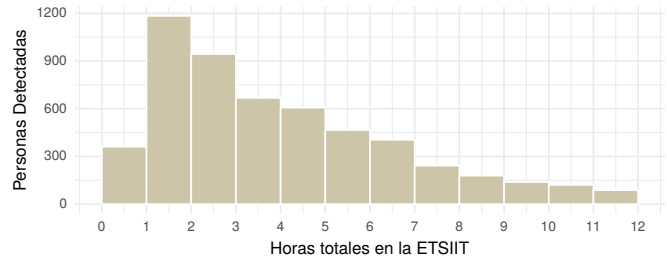


Figura 4. Duración de las visitas registradas. La mayoría de los visitantes están entre una y dos horas en la escuela.

Conociendo tanto la hora de entrada como la hora de salida de cada visitante, se ejecuta el algoritmos de agrupamiento *FarthestFirst* [21]. Para facilitar la interpretación de los resultados del algoritmos de agrupamiento lidiando con fechas y horas, las horas de entrada y salida se resumen en 5 categorías divisorias obtenidas como *boundaires* o fronteras del agrupador. Estas divisiones se recogen en la Tabla I.

Tabla I
INTERPRETACIÓN DE LOS DIVISORES DE HORAS DEL AGRUPADOR.

Madrugada	$\in (0, 1, 2, 3, 4, 5, 6]$
Mañana	$\in (7, 8, 9, 10, 11]$
Mediodía	$\in (12, 13, 14, 15]$
Tarde	$\in (16, 17, 18, 19, 20]$
Noche	$\in (21, 22, 23]$

El algoritmo de agrupamiento establece cinco clases de visitantes (notadas mediante \square , \circ , \triangle , \diamond y \boxtimes), más una clase auxiliar para las visitas no clasificables o sin un patrón común extraído notado mediante $+$. Adicionalmente, en este artículo se notará con el símbolo \times aquellas visitas registradas mediante un único paso cercano al sensor. En la Tabla II se recogen la asignación de cada divisor de hora a cada clase calculada mediante del algoritmo de agrupamiento.

Tabla II
CLASES RESULTANTES DEL AGRUPAMIENTO DE CATEGORÍAS

Clase	Entrada	Salida
\square g1	Mañana	Mediodía
\circ g2	Mediodía	Mediodía
\triangle g3	Tarde	Tarde
\diamond g4	Mediodía	Tarde
\boxtimes g5	Mañana	Tarde
$+$	-	*

Para la visita de cada día de cada persona distinta, se dispone de la hora de entrada, la hora de salida y la categoría que le ha asignada mediante el algoritmo de agrupamiento. En la Figura 8 se puede ver como se relacionan las tres variables entre ellas. Las regiones con patrones comunes de entrada y salida han sido asignadas a la misma clase por el algoritmo de agrupamiento. Debido al libre albedrío de las personas, hay visitas que no corresponden a un patrón compartido.

Dichas visitas son notadas mediante **+** y descartadas por el agrupador, en lugar de intentar a ajustarlas de forma artificial a algún patrón o clase ya existente.

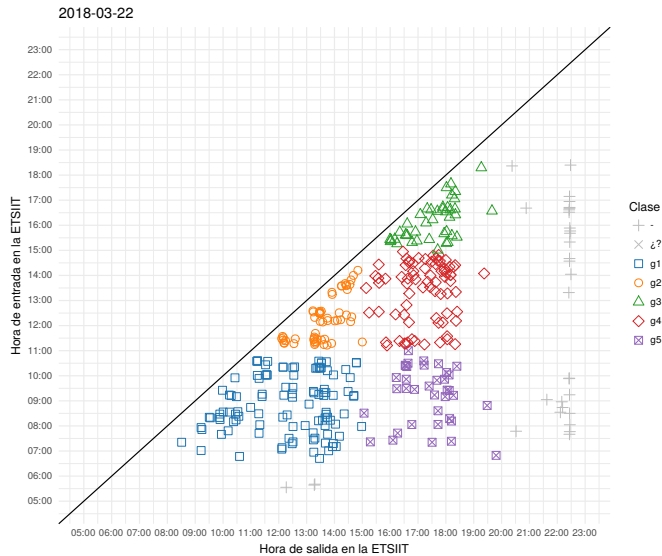


Figura 5. Relación entre la hora de entrada, la hora de salida y la clase asignada por el agrupador para todas las visitas de un día determinado.

El algoritmo de agrupamiento se aplica a todas las visitas de todos los dispositivos detectados a lo largo de todo el periodo. La figura 6 recoge la relación entre la hora de entrada, la hora de salida y la categoría asignada a la visita para todos los días que han sido monitorizados por el nodo.

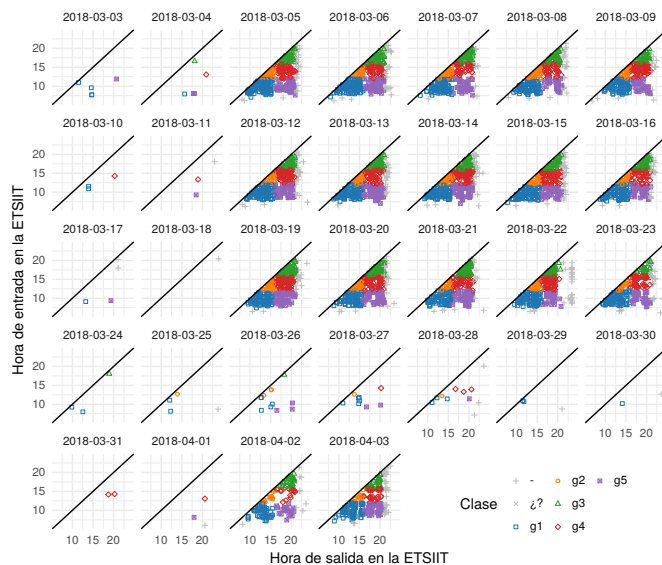


Figura 6. Relación entre la hora de entrada, la hora de salida y la clase asignada por el agrupador para todos las visitas del conjunto de datos.

Es esperable que los visitantes acudan de forma regular a la escuela, ya sea porque estudien o trabajen en ella. En la Figura 7 se puede ver el número de visitantes que han sido

detectados varios días distintos o de forma reincidente. Esta regularidad individual, supone a su vez un patrón en si mismo.

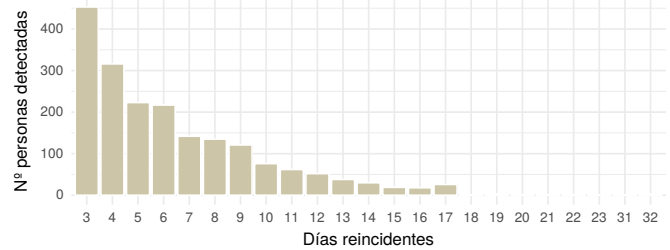


Figura 7. Representación del numero de visitantes reincidentes.

Dada la capacidad del sistema *Mobywit* para ser capaz de reconocer de forma reincidente al mismo dispositivo y por tanto a la misma persona, se pueden extraer patrones en los hábitos de una misma persona a lo largo de varios días. La Figura 8 muestra para un subconjunto de personas reincidentes más de 10 veces, la clase asignada a cada una de sus diferentes visitas en diferentes días. En dicha gráfica se puede como las personas se ajustan en su mayoría a un patrón común de comportamiento a lo largo del tiempo.

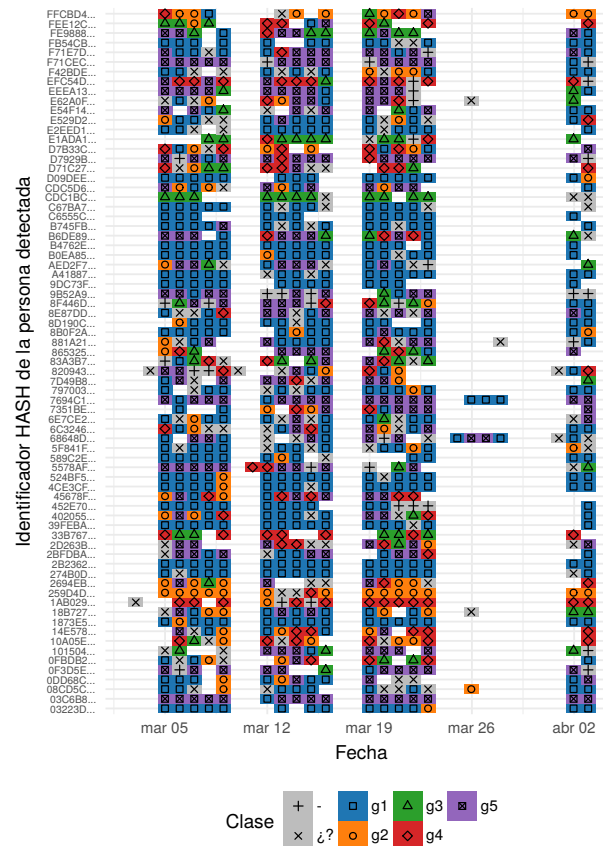


Figura 8. Representación de la clase asignada en diferentes visitas para un subconjunto de visitantes los reincidentes más de diez veces. Se observa como las visitas sucesivas de la mayoría de los visitantes son de la misma clase.



No todas las visitas de una misma persona tiene porque ser clasificadas con la misma categoría, sin embargo si puede ser definida una categoría predominante para cada persona de la cual se extraiga un patrón de comportamiento habitual, aunque debido a circunstancias externas ese patrón no sea perfecto. En la Tabla III se presenta la clase predominante de todos los visitantes ($n \geq 1$) como aquellos reincidentes al menos 10 veces ($n \geq 10$).

Tabla III
PORCENTAJE DE PERSONAS PARA CADA CLASE PREDOMINANTE.

Clase	Predominancia		
	Nombre	$n \geq 1$	$n \geq 10$
□	g1	28,5 %	43,8 %
○	g2	17,9 %	4,1 %
△	g3	18,1 %	4,1 %
◇	g4	11,3 %	10,9 %
⊠	g5	8,1 %	31,5 %
+	-	15,8 %	5,4

Estos patrones de comportamiento pueden ser estudiados a lo largo del tiempo para ver la influencia de factores periódicos, como los días de la semana, los periodos de exámenes o los cuatrimestres en los que se divide el curso. Por ejemplo la figura 9 presenta los horarios habituales para cada persona detectada a lo largo de las distintas semana.

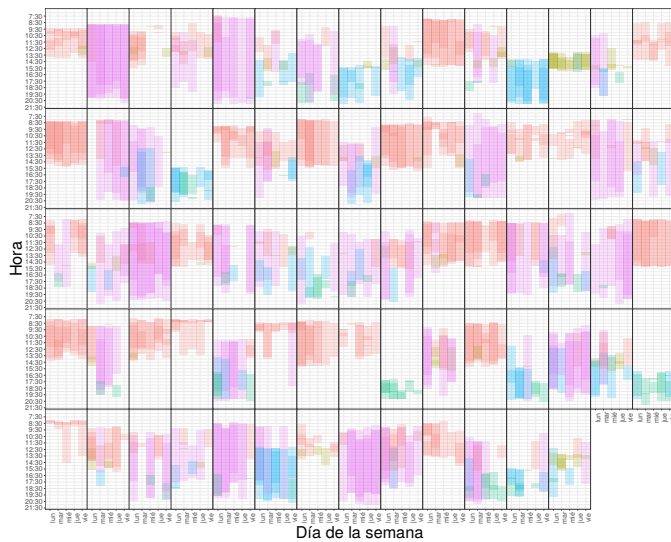


Figura 9. Horarios semanales de las 73 personas detectadas reincidentes más representativas. El color indica la clase en la que la que cada estancia ha sido clasificada. Cada visita se superpone con una capa de transparencia del 25%. La codificación de colores es la empleada en la Tabla III.

Esta información puede ser útil para decidir a que hora emplazar eventos maximizando la disponibilidad de los alumnos, escogiendo aquellas horas en las que habitualmente estén habitualmente a punto de irse de la facultad. Empleando la información del agrupador, se puede elegir intervalos de tiempos libres que beneficien a un tipo de visitante concreto, como se muestra en la Figura 10.

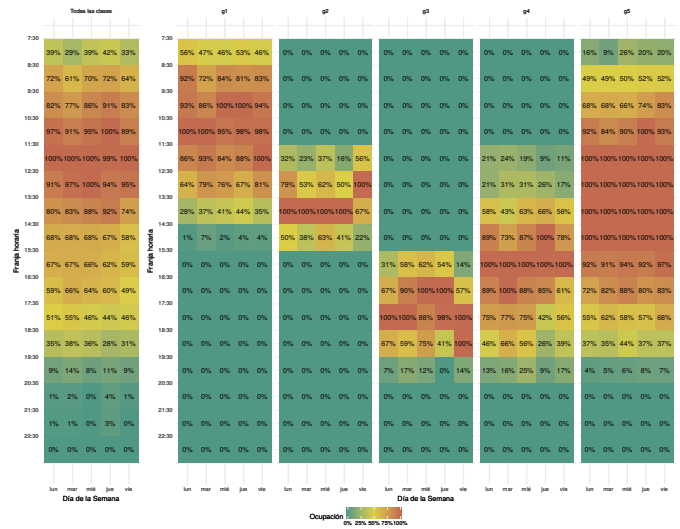


Figura 10. Disponibilidad de horarios para la realización de eventos en base a la asistencia a la escuela, tanto genérica como en base los distintos grupos de usuarios detectados. La ocupación se ha normalizado para cada día a una escala entre el 0 y el 100% del máximo diario.

V. RESULTADOS

Se observa que la clase □ asignada a las personas que entran por la mañana y salen al medio día es la que ha sido asignada a un mayor porcentaje de gente, indicando que la mayoría de las personas que visitan la escuela lo hacen solo en horario de mañana. Además este patrón se muestra tanto en aquellas visitas puntuales, y se acentúa en las visitas reincidentes. Para los reincidentes, la segunda clase con mayor porcentaje (⊠) es aquella asignada a los visitantes que llegan por la mañana y se marchan por la tarde, constituyendo aproximadamente un tercio de los visitantes reincidentes. Un 5% de las visitas, se produce únicamente al mediodía (○) posiblemente por gente que únicamente emplea los servicios de comedores del edificio. Los visitantes que llegan al mediodía y se van por la tarde (◇) suponen un 10% de las visitas. Por último, sólo un 5% de los visitantes realizan su visita únicamente durante horario de tarde (△).

En cuanto a la disponibilidad de horario, en términos generales la mayor disponibilidad se presenta los jueves a partir de las 18:30. Para los distintos turnos, la asistencia por la mañana presentará mayor disponibilidad los martes de hasta las 11:30, en los turnos de tarde los jueves a partir 18:30. Estos rangos de horas y días son por tanto el momento idóneo para emplazar cualquier evento que no interrumpa la docencia habitual, obtenida en base a los patrones de comportamiento habituales de los alumnos reincidentes. O en caso de querer atraer a toda la audiencia posible, elegir dos periodos de tiempo que permitan la afluencia de ambos tipos de visitantes disjuntos, repitiendo el evento en los dos intervalos de tiempo obtenidos.

VI. CONCLUSIONES

Si bien la información obtenida no es más que una muestra de la viabilidad de la captación de comunicaciones inalámbricas WiFi para la monitorización y extracción de patrones de comportamientos de las estancias de las personas en edificios, puede resultar una herramienta muy útil para la gestión eficiente de los recursos de dicho edificio. Así como para estudiar el impacto de nuevas medidas en los patrones de comportamiento de los visitantes.

A modo de ejemplo, en este artículo se han determinado los periodos de mayor disponibilidad de los alumnos para la posible asistencia a actividades no regladas, como charlas o conferencias, en base a los patrones de comportamiento extraídos de las visitas reincidentes.

La potencialidad de la aplicación de técnicas de inteligencia computacional y análisis de patrones a los datos obtenidos mediante la captación de comunicaciones inalámbricas permite dotar de fuentes de información baratas, fiables y anónimas a las ciudades inteligentes. Fuentes de datos que permitirán optimizar los recursos ajustándolos a los patrones de comportamiento aprendidos de las personas que los emplean.

La aplicabilidad del sistema *Mobywit* en multitud de escenarios de escenarios, supone también una futura línea de trabajo. El sistema puede ser empleado en cualquier punto geográfico donde se concentren masas de personas, y extraer información sobre el comportamiento de las mismas. Por ejemplo, ofreciendo tiempos de espera hasta ser atendidos en tareas administrativas, monitorizando los recintos habilitados para fumadores o el impacto de la calidad del menú del día en la afluencia de gente en los servicios de comedores.

AGRADECIMIENTOS

Este trabajo ha sido financiado parcialmente por los proyectos MOMOFES SPIP2017-02116 (Dirección General de Tráfico de España), EphemeCH TIN2014-56494-C4-3-P, DeepBio TIN2017-85727-C4-2-P y TEC2015-68752. Especial agradecimiento al cuerpo directivo de la ETSIT por la concesión del permiso para la implantación del nodo *Mobywit*.

REFERENCIAS

- [1] W. J. Stanton, "Fundamentals of marketing," 1967.
- [2] G. Pan, G. Qi, W. Zhang, S. Li, Z. Wu, and L. T. Yang, "Trace analysis and mining for smart cities: issues, methods, and applications," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 120–126, 2013.
- [3] R. Schneider, R. Patten, and J. Toole, "Case study analysis of pedestrian and bicycle data collection in us communities," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1939, pp. 77–90, 2005.
- [4] R. Parisi, E. Di Claudio, G. Lucarelli, and G. Orlandi, "Car plate recognition by neural networks and image processing," in *Circuits and Systems, 1998. ISCAS'98. Proceedings of the 1998 IEEE International Symposium on*, vol. 3. IEEE, 1998, pp. 195–198.
- [5] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.
- [6] K. Hashimoto, C. Kawaguchi, S. Matsueda, K. Morinaka, and N. Yoshiike, "People-counting system using multisensing application," *Sensors and Actuators A: Physical*, vol. 66, no. 1-3, pp. 50–55, 1998. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0924424797017159>

- [7] J. W. Davis and V. Sharma, "Robust background-subtraction for person detection in thermal imagery," in *CVPR Workshops*, 2004, p. 128.
- [8] K. Hashimoto, K. Morinaka, N. Yoshiike, C. Kawaguchi, and S. Matsueda, "People count system using multi-sensing application," in *Solid State Sensors and Actuators, 1997. TRANSDUCERS '97 Chicago., 1997 International Conference on*, vol. 2, Jun 1997, pp. 1291–1294 vol.2.
- [9] H. Celik, A. Hanjalic, and E. Hendriks, "Towards a robust solution to people counting," in *Image Processing, 2006 IEEE International Conference on*, Oct 2006, pp. 2401–2404.
- [10] A. Schofield, P. Mehta, and T. Stonham, "A system for counting people in video images using neural networks to identify the background scene," *Pattern Recognition*, vol. 29, no. 8, pp. 1421 – 1428, 1996. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0031320395001638>
- [11] C.-H. Kuo and R. Nevatia, "How does person identity recognition help multi-person tracking?" in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 1217–1224.
- [12] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM computing surveys (CSUR)*, vol. 35, no. 4, pp. 399–458, 2003.
- [13] M. L. Gras, "The legal regulation of cctv in europe," *Surveillance & Society*, vol. 2, no. 2/3, 2002.
- [14] T. Nicolai and H. Kenn, "About the relationship between people and discoverable bluetooth devices in urban environments," in *Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology*. ACM, 2007, pp. 72–78.
- [15] J. Weppner and P. Lukowicz, "Bluetooth based collaborative crowd density estimation with mobile phones," in *Pervasive Computing and Communications (PerCom), 2013 IEEE International Conference on*, March 2013, pp. 193–200.
- [16] A. Morrison, M. Bell, and M. Chalmers, "Visualisation of spectator activity at stadium events," in *Information Visualisation, 2009 13th International Conference*. IEEE, 2009, pp. 219–226.
- [17] Y. Chon, S. Kim, S. Lee, D. Kim, Y. Kim, and H. Cha, "Sensing wifi packets in the air: practicality and implications in urban mobility monitoring," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2014, pp. 189–200.
- [18] W. Xi, J. Zhao, X.-Y. Li, K. Zhao, S. Tang, X. Liu, and Z. Jiang, "Electronic frog eye: Counting crowd using wifi," in *INFOCOM, 2014 Proceedings IEEE*, April 2014, pp. 361–369.
- [19] A. Fernández-Ares, M. García-Arenas, P. A. Castillo, and J. J. Merelo, "Impact of protests in the number of smart devices in streets: A new approach to analyze protesters behavior," in *International Conference on Smart Cities*. Springer, 2017, pp. 75–85.
- [20] A. Fernández-Ares, A. Mora, M. Arenas, P. García-Sánchez, G. Romero, V. Rivas, P. Castillo, and J. Merelo, "Studying real traffic and mobility scenarios for a smart city using a new monitoring and tracking system," *Future Generation Computer Systems*, vol. 76, pp. 163–179, 2017.
- [21] Hochbaum and Shmoys, "A best possible heuristic for the k-center problem," *Mathematics of Operations Research*, vol. 10, no. 2, pp. 180–184, 1985.



Reduction of the Size of Datasets by using Evolutionary Feature Selection: the Case of Noise in a Modern City*

***Note:** The full contents of this paper have been published in the volume *Lecture Notes in Artificial Intelligence 11160* (LNAI 11160)

Javier Luque, Jamal Toutouh, Enrique Alba
Departamento de Lenguajes y Ciencias de la Computación
Universidad de Málaga
Málaga, Spain
javierluque@uma.es, {jamal,eat}@lcc.uma.es

Abstract—Smart city initiatives have emerged to mitigate the negative effects of a very fast growth of urban areas. Most of the population in our cities are exposed to high levels of noise that generate discomfort and different health problems. These issues may be mitigated by applying different smart cities solutions, some of them require high accurate noise information to provide the best quality of serve possible. In this study, we have designed a machine learning approach based on genetic algorithms to analyze noise data captured in the university campus. This method reduces the amount of data required to classify the noise by addressing a feature selection optimization problem. The experimental results have shown that our approach improved the accuracy in 20% (achieving an accuracy of 87% with a reduction of up to 85% on the original dataset).

Index Terms—Smart city, Genetic algorithm, Feature selection, Noise



Studying Solutions of the p -Median Problem for the Location of Public Bike Stations*

***Note:** The full contents of this paper have been published in the volume *Lecture Notes in Artificial Intelligence 11160* (LNAI 11160)

Christian Cintrano <i>E.T.S. Ingeniería Informática</i> University of Málaga Málaga, Spain cintrano@lcc.uma.es	Francisco Chicano <i>E.T.S. Ingeniería Informática</i> University of Málaga Málaga, Spain chicano@lcc.uma.es	Thomas Stütze <i>Université Libre de Bruxelles</i> Brussels, Belgium stuetzle@ulb.ac.be	Enrique Alba <i>E.T.S. Ingeniería Informática</i> University of Málaga Málaga, Spain eat@lcc.uma.es
--	--	--	---

Abstract—The use of bicycles as a means of transport is becoming more and more popular today, especially in urban areas, to avoid the disadvantages of individual car traffic. In fact, city managers react to this trend and actively promote the use of bicycles by providing a network of bicycles for public use and stations where they can be stored. Establishing such a network involves the task of finding best locations for stations, which is, however, not a trivial task. In this work, we examine models to determine the best location of bike stations so that citizens will travel the shortest distance possible to one of them. Based on real data from the city of Malaga, we formulate our problem as a p -median problem and solve it with a variable neighborhood search algorithm that was automatically configured with irace. We compare the locations proposed by the algorithm with the real ones used currently by the city council. We also study where new locations should be placed if the network grows.

Index Terms—Bike station location, p -Median problem, Variable neighborhood search