

O impacto dos *smartphones* nos modelos de mobilidade tradicionais

Nuno Cruz^{1,2}, Hugo Miranda², and Pedro Ribeiro¹

¹ Instituto Politécnico de Lisboa
Instituto Superior de Engenharia de Lisboa
ADEETC

² Universidade de Lisboa - Faculdade de Ciências
LaSIGE - Large-Scale Informatics Systems Laboratory

Resumo O desenho e avaliação de protocolos, aplicações e *frameworks* para redes sem fios recorre frequentemente às premissas ditadas por modelos de mobilidade sintéticos. Os modelos de mobilidade permitem ultrapassar as dificuldades em colecionar o grande número de registos de mobilidade dos utilizadores que seriam necessários para dar relevância estatística às avaliações, mas falham por não representar movimentos e padrões de conectividade observados.

Este artigo apresenta uma análise sobre os registos de mobilidade dos cerca de 36000 dispositivos que estiveram ligados aos quase 240 APs que suportam a rede WiFi eduroam do Instituto Politécnico de Lisboa no ano 2013. Os registos são comparados com modelos sintéticos utilizando um conjunto de métricas. A comparação levanta algumas questões relativamente às suposições feitas nos cenários típicos de avaliação com modelos sintéticos.

1 Introdução

Os simuladores têm um papel importante na validação e avaliação de aplicações e protocolos para redes sem fios, uma vez que permitem que se ultrapassem as dificuldades em criar experiências reais de larga escala com uma grande duração. Os simuladores testam uma implementação da aplicação utilizando uma série de abstrações do ambiente, tráfego e movimento dos nós. Assim, a fiabilidade das experiências é fortemente influenciada pela qualidade com que cada uma destas abstrações reproduz a realidade.

A simulação do movimento dos nós tem seguido duas aproximações: modelos de mobilidade sintéticos puros e modelos de mobilidade sintéticos baseados em registos dos dispositivos. Nos modelos de mobilidade sintéticos puros, os nós movimentam-se de acordo com uma função estatística pré-definida numa tentativa de representar de forma realista o movimento humano ou do dispositivo. A qualidade com que estes modelos replicam corretamente os padrões de movimento dos utilizadores tem sido criticada pela sua distância à realidade quando avaliada por métricas como a não homogeneidade [4].

Os modelos de mobilidade sintéticos baseados em registos utilizam distribuições estatísticas que tentam retratar as propriedades de mobilidade observadas previamente (e.g. [5, 8, 13]). Os registos são fornecidos por voluntários, que disponibilizam a sua localização ou por terceiros que a observam passivamente. Infelizmente, os registos disponíveis são escassos, apresentam durações temporais curtas e/ou poucos utilizadores, afetando a qualidade da modelação.

Para ultrapassar as limitações dos modelos de mobilidade baseados em registos, este artigo propõe um novo modelo que consiste na reprodução efetiva dos registos dos utilizadores. O estudo incide na observação dos 36000 dispositivos que se ligaram à rede Eduroam do Instituto Politécnico de Lisboa durante o ano de 2013. O estudo permitiu identificar 17714 dispositivos a correrem especificamente um sistema operativo para dispositivos de pequenas dimensões (iOS, Android e Windows Mobile/Phone) e 15532 dispositivos a correr sistemas operativos normalmente encontrados em portáteis. Os dados incluem momentos de associação e dissociação dos dispositivos móveis a pontos de acesso e as coordenadas geográficas dos pontos de acesso.

As expectativas são de que a escala do conjunto de dados contribua para a definição de modelos de mobilidade de conjuntos de registos, uma nova classe que, em contraste com aproximações anteriores, crie cenários de mobilidade que utilize exclusivamente dados reais sem utilização de modelos estatísticos.

Este artigo apresenta duas contribuições. Primeiro, apresenta o MobIPLity, um modelo de mobilidade que utiliza conjuntos de registos de mobilidade reais. Em segundo, caracteriza e compara os registos do MobIPLity com modelos sintéticos baseados em registos encontrados na literatura.

O artigo está organizado da seguinte forma. A Sec. 2 apresenta brevemente os diferentes modelos de mobilidade e os esforços para colecionar dados de mobilidade existentes na literatura. Adicionalmente, introduz as métricas utilizadas para caracterizar a mobilidade humana. A Sec. 3 e 4 apresentam respetivamente a metodologia utilizada para converter o conjunto de dados em bruto em cenários de mobilidade baseados em registos e comparados com outras aproximações que podem ser encontradas na literatura. As conclusões surgem na Sec. 5.

2 Trabalho relacionado

Os modelos de mobilidade sintéticos puros utilizam distribuições aleatórias para modelar o movimento dos nós. Um exemplo disso é o Random Way Point Mobility Model (RWP) cuja simplicidade na geração de cenários de mobilidade facilita a avaliação de aplicações e protocolos. Contudo, tem sido demonstrado que, adicionalmente à sua modelação simples do comportamento humano, os modelos de mobilidade sintéticos distribuem os nós numa forma não natural [4]. Para colmatar estes problemas, em [8] é proposta uma variação do RWP utilizando padrões que apresentam a não homogeneidade encontrada na mobilidade humana.

Os modelos de mobilidade sintéticos baseados em registos, por outro lado, tentam reproduzir os padrões observados no movimento humano modelando o comportamento dos nós de acordo com uma função de distribuição probabilís-

tica. Os mecanismos utilizados para colecionar dados que inspiraram os modelos de mobilidade baseados em registos podem ser divididos em duas categorias. Aproximações intrusivas (por exemplo [15, 18]) são aquelas que obtêm os seus dados diretamente do dispositivo transportado pelo utilizador. Estas aproximações beneficiam da precisão dos dados, capturados por software ou hardware dedicado. Infelizmente, estes estudos estão limitados pela quantidade considerável de recursos envolvidos, que limitam a duração e o número de participantes e que podem influenciar as conclusões relacionadas com a identificação de padrões.

Aproximações não intrusivas utilizam registos obtidos por dispositivos externos (como pontos de acesso ou dispositivos de localização interior), para produzir registos com a localização dos utilizadores a cada instante. Devido aos problemas de privacidade criados pela extração dos dados, as aproximações não intrusivas são as mais populares e as que apresentam uma maior escalabilidade no número de utilizadores e na duração temporal. Infelizmente, a pesquisa por modelos de mobilidade [3, 11] sugere uma escassez de registos a partir de meados de 2008, excluindo assim a época de massificação dos *smartphones* e *tablets* e que poderiam apresentar novos padrões de mobilidade e contactos entre utilizadores.

A rede WiFi do *Dartmouth College* tem sido utilizada para extrair um número considerável de registos, por exemplo durante as 17 semanas do semestre de inverno do ano lectivo de 2003/2004 [9]. O método para extrair os registos é muito semelhante ao utilizado neste artigo, e que é descrito na Sec. 3. Neste trabalho, os autores utilizaram os registos para modelar movimentos reais dos utilizadores e definiram um limiar de velocidade, abaixo do qual se assume que os utilizadores estão parados antes de se dirigirem até ao destino. Os registos de mobilidade de 198 telefones VoIP sem fios foram utilizados para definir um modelo de mobilidade sintético baseado em registos [12]. A modelação aborda características sociais, espaciais e temporais, considerando ainda *hotspots* (grupos densos de utilizadores), distinção entre dias úteis e fins de semana e conjuntos de dispositivos móveis e estacionários. Em comparação com o *MobIP*Lity, o estudo de 2003/2004 analisa um conjunto mais elevado de pontos de acesso, mas um número mais baixo de utilizadores e durante um período mais curto. Infelizmente, a sua idade impede-o de considerar um número de avanços recentes, como o crescimento de dispositivos móveis de pequenas dimensões, tais como o iPhone e os *smartphones* baseados em Android (disponibilizados ao público em 2007³ e 2008⁴, respetivamente).

Os resultados de um estudo de 2 meses da rede sem fios *eduroam* das universidades de Minho e Vigo pode ser encontrado em [14]. A metodologia seguida é bastante semelhante à utilizada no nosso trabalho. A associação dos pontos de acesso ao espaço físico permitiu separar o tráfego originado em áreas residenciais das áreas académicas. Durante este trabalho os autores descobriram que os APs com mais utilizadores não são necessariamente aqueles com mais tráfego. Tal como esperado, o artigo evidencia um padrão semanal de utilização da rede, que indica que a vasta maioria dos utilizadores apenas se liga durante dias úteis. Em

³ <http://en.wikipedia.org/wiki/IPhone>

⁴ [http://en.wikipedia.org/wiki/Android_\(operating_system\)](http://en.wikipedia.org/wiki/Android_(operating_system))

termos de mobilidade, os autores concluem que 90% dos utilizadores ligam-se a mais que um AP mensalmente, com cerca de 35% dos utilizadores a visitar pelo menos 5 APs. Infelizmente, a duração curta do estudo cria a noção de uma mobilidade dispersa no tempo e de pouca relevância na caracterização da mobilidade real.

A mobilidade humana tem sido caracterizada nos eixos espaciais, temporais e sociais [11] por um conjunto de métricas. A densidade da distribuição é caracterizada no eixo espacial por métricas como *comprimento do salto* (também referida como *voe*) que caracteriza a distancia média de cada movimento. Os modelos de mobilidade baseados em registos modelam o *comprimento do salto* utilizando distribuições log-normal [12] ou lei de potência truncada [5, 13]. O comprimento do salto depende da área física onde a simulação ocorre, sendo influenciado, por exemplo, pela distância entre edifícios. A métrica de *não homogeneidade* [17] avalia as distribuições dos nós no espaço, que se reflete na existência de pontos mais populares. A variação do Random Way Point apresentada em [8] e o Disaster Area [2] são bons exemplos de modelos de mobilidade que forcem uma distribuição heterogénea dos nós. Um valor baixo de não homogeneidade é esperado em distribuições aleatórias, onde os utilizadores não formam grupos.

As propriedades espaciais estão normalmente ligadas às temporais, associando o tempo e a distância percorrida entre dois pontos. As propriedades da mobilidade humana que variam no tempo caracterizam padrões como as variações entre dias úteis e fins-de-semana e o tempo despendido num local específico (conhecido como tempo de pausa), utilizados para modelar padrões de pausa encontrados, por exemplo, na rede do campus do Dartmouth college [12].

O eixo social caracteriza os encontros entre pares. Na combinação com o eixo temporal, contribuem para determinar durante quanto tempo e com que frequência dois ou mais nós se encontram. Múltiplos modelos centrados fortemente nas relações sociais estabelecidas entre indivíduos, foram propostos considerando métricas de atração (encontrada por exemplo em [5]) mas também de repulsão. Ambas as propriedades são exploradas em [7] através da combinação dos modelos de relações entre deslocações individuais e viagens em grupo. O tempo entre contactos (do inglês *inter-contact time* (ICT)) é uma métrica frequentemente utilizada para avaliar o eixo social de um modelo de mobilidade. É definido pelo intervalo de tempo entre contactos consecutivos entre dois nós. Os modelos sintéticos utilizam normalmente uma lei de potência truncada para modelar o ICT [6, 10, 13].

3 MobIPLity

3.1 Caracterização do conjunto de dados

O conjunto de dados utilizado neste estudo contém os registos de acesso produzidos por todos os Pontos de Acesso (APs) da rede sem fios eduroam do Instituto Politécnico de Lisboa (IPL), gerados entre 1 de Janeiro e 31 de Dezembro de 2013. Um total de 35930 dispositivos e 21162 utilizadores acederam à rede durante este período, gerando cerca de 15 milhões de registos.

O IPL é a 7^a maior instituição de ensino superior em Portugal com aproximadamente 1300 docentes e 15000 alunos, distribuídos por 10 localizações distintas na área metropolitana de Lisboa (ver Fig. 1). A rede eduroam do IPL é suportada por 236 APs Cisco Systems, cobrindo um total de 26 edifícios e o campus circundante a estes. Os registos têm como origem todos os utilizadores a aceder à rede, incluindo os visitantes de outras instituições nacionais ou internacionais.

O nosso estudo distribui os dispositivos em duas classes: *Dispositivos de Pequenas Dimensões* (DPD) pequenos e utilizados em andamento. Exemplos destes dispositivos são os *smart-phones*, PDAs e *tablets*. A segunda classe, *Portáteis*, agrupa os dispositivos maiores, executando um sistema operativo clássico (Linux, Windows ou MacOS). Os registos de DHCP permitiram identificar 17714 *DPDs* e 15532 *Portáteis*. A análise dos resultados considera uma terceira classe, *Todos*, que agrega *DPDs*, *Portáteis* e os 2684 dispositivos para os quais não foi possível uma classificação.



Figura 1: Localizações do IPL

Os dados de mobilidade são obtidos a partir dos registos produzidos pelo protocolo RADIUS [16]. Estes registos reproduzem o conceito de sessão RADIUS considerando assim a associação do utilizador a um AP distinto. Os registos contêm o endereço MAC do dispositivo, o identificador do AP, nome de utilizador, duração, início e fim da sessão. Os registos foram editados para: *i*) remoção de todos os registos com hora de entrada igual à de saída. As sessões com estas características são criadas quando o utilizador tem algum problema ao ligar-se à rede, apesar da rede considerar o utilizador autenticado (criando efetivamente o registo RADIUS). Tipicamente uma nova sessão é criada logo de seguida; *ii*) correção de sessões concorrentes do mesmo dispositivo a APs distintos. Esta impossibilidade só pode ser explicada se o dispositivo não se dissociou corretamente de um AP antes de se associar ao próximo, tendo a hora de saída sido estabelecida artificialmente pelo serviço RADIUS por *timeout*. Neste caso, a saída da primeira sessão é corrigida para ocorrer imediatamente antes do início de sessão do próximo registo; *iii*) ajuste da hora de saída sempre que a duração determinada pela diferença entre a hora de início e a hora de saída tenha uma diferença superior a 30 segundos ao registo da duração de sessão. A diferença dos campos ocorre devido ao utilizador poder sair por *timeout* e por a atualização da duração de sessão ser responsabilidade do AP, que a envia periodicamente, podendo levar a que os valores tenham diferenças elevadas por exemplo em situações que o AP reinicie. Uma vez que a duração de sessão é o valor mais fidedigno, optou-se por ajustar a hora de saída.

A Fig. 2 mostra o número de dispositivos observado ao longo do ano, onde são contabilizados os dispositivos que se ligaram pelo menos uma vez em cada dia. Nesta é possível observar a variância entre dias úteis, fins de semana e a variação entre feriados/férias e períodos letivos. De notar também que entre 27

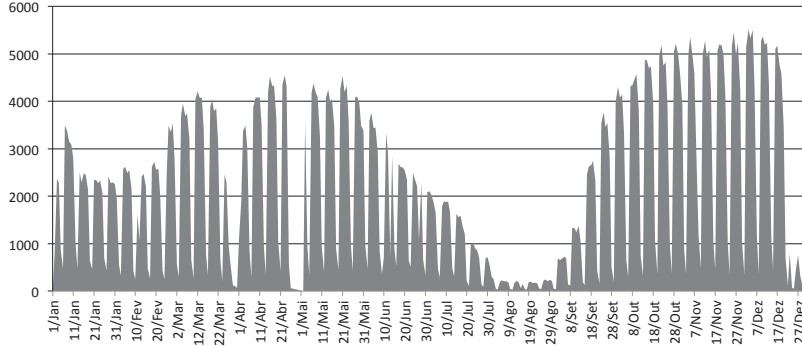


Figura 2: Dispositivos ligados por dia

de Abril e 1 de Maio não ocorreram registos de acessos de RADIUS devido a uma indisponibilidade do sistema.

3.2 Algoritmo para geração dos registos de mobilidade

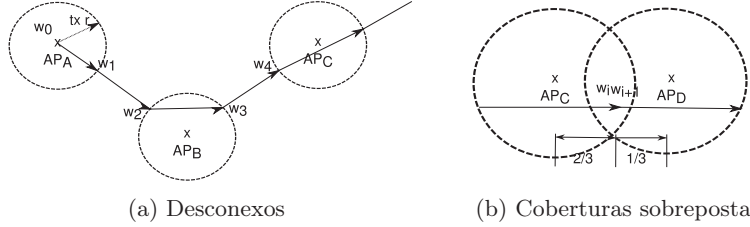
O conjunto de registos de mobilidade produzidos pelos dados é criado a partir de um conjunto $E \subseteq D \times A \times \{in, out\} \times T$ onde D é o conjunto de dispositivos móveis, A é o conjunto de pontos de acesso da rede anotado com a sua coordenada geográfica e T são estampilhas temporais. O conjunto é preenchido por um par de eventos da forma $(d, a, in, t_1), (d, a, out, t_2)$, para cada registo de acesso RADIUS, sendo t_1 e t_2 respetivamente o momento de associação/dissociação de d ao AP a . Finalmente, $E_d \subseteq E$ é o subconjunto de E que contém todos os eventos produzidos pelo dispositivo d .

É esperado que E_d respeite duas invariantes: *i*) os dispositivos estão sempre associados com um ponto de acesso antes de se dissociarem deste; e *ii*) em qualquer momento, um dispositivo está associado no máximo a um ponto de acesso. De notar que a invariante *i*) é trivialmente assegurada pelo software dos pontos de acesso e a invariante *ii*) pelas correções aplicadas aos registos RADIUS descritas na Sec. 3.1.

Definimos $E'_d = e_{d,0}, e_{d,1}, \dots, e_{d,n}, d \in D \wedge e_{d,i} \in E_d \wedge i > 0$ como o conjunto de eventos ordenados no tempo para o dispositivo d . De notar que a invariante *i*) assegura que $e_{d,2j}, j \geq 0$ são eventos do tipo *in* e por consequência, $e_{d,2j+1}, j \geq 0$ são eventos do tipo *out*.

Uma rota $W_d = w_0, w_1, \dots, w_{2n-1}, w_i = F(e_{d,2j+i}), 0 \leq i \leq 2n-1 \wedge j \geq 0 \wedge n \geq 1$ para um dispositivo d , é definida por uma sequência de *way-points* w_i . Os *way-points* são definidos por uma coordenada geográfica e uma estampilha temporal, obtida por uma função F aplicada a eventos consecutivos (não necessariamente iniciada em $e_{d,0}$) em E'_d .

A função F depende da posição do *way-point* na sequência e no tipo (*in, out*) do evento. O caso geral é retratado na Fig. 3a. w_0 é um *way-point* com a estampilha temporal e as coordenadas do ponto de acesso em $e_{d,2j}$, o primeiro



(a) Desconexos (b) Coberturas sobrepostas

Figura 3: Exemplos de extração de registos de mobilidade

da sequência de eventos escolhidos. As transformações subsequentes de pares de eventos em pares de *way-points* $w_{2i+1}, w_{2i+2}, i \geq 0$ irão devolver as coordenadas numa reta dada por um vetor $\vec{AP_A AP_B}$, onde AP_A, AP_B são as coordenadas dos pontos de acesso dos eventos correspondentes $e_{d,2j+2i+1}, e_{d,2j+2i+2}$. A localização precisa é determinada pelo alcance de transmissão dos pontos de acesso, de tal forma que w_{2i+1} (resp. w_{2i+2}) será colocado na intersecção do vetor com o raio de transmissão do AP_A (resp. AP_B). As estampilhas temporais de w_1 e w_2 são copiadas dos eventos correspondentes. De notar que, de acordo com a definição acima de E'_d , os eventos $e_{d,2j+2i+1}, e_{d,2j+2i+2}$ são respetivamente um registo de *out* e *in*, sinalizando o momento em que d abandonou a área coberta por AP_A e o momento em que d se associou a AP_B . O algoritmo é sucessivamente repetido para cada par de eventos e *way-points*, resultando numa rota como é apresentada na Fig. 3a.

No caso particular em que a área coberta por dois pontos de acesso consecutivamente acedidos por um dispositivo se sobrepõe (Fig. 3b), o algoritmo reflete a aproximação conservadora dos drivers dos dispositivos. Os dois *way-points* recebem a mesma estampilha temporal do registo *in* e são colocados a $2/3$ da distância entre pontos de acesso.

Os registos de mobilidade são terminados por um evento *out* com a criação de um *way-point* nas coordenadas do ponto de acesso. Assume-se que o dispositivo abandona a rede eduroam e portanto que o registo deve ser terminado quando a velocidade para percorrer a distância entre dois-pontos de acesso consecutivos desce abaixo de um limiar. Em alternativa, os registos são terminados quando duas ligações consecutivas para o mesmo AP excedem um limiar temporal, sugerindo que os utilizadores abandonaram o local e voltaram posteriormente.

3.3 Extração dos registos de mobilidade

Os registos de mobilidade são extraídos a partir de um conjunto E armazenado numa base dados local usando o algoritmo descrito na Sec. 3.2. A Tabela 1 lista os diferentes parâmetros de configuração do algoritmo.

Um período de *warm up* e *cool off* são definidos imediatamente antes e depois do período selecionado. Estes períodos asseguram que apenas registos de mobilidade que permanecem ativos durante a totalidade da duração da simulação são selecionados, uma vez que têm de incluir pelo menos um *way-point* em cada

Tabela 1: Opções para extração do registo de mobilidade

Data de início/fim	Intervalo de registos
# Dispositivos	Número de dispositivos
Tipo de dispositivo	Portátil/DPD/Todos
Pontos	Número mínimo de APs num registo
Duração	Duração do registo de mobilidade
Eixos	Representação 2D/3D
Localização	Escola para extração
Warm Up	Duração do período de <i>warm up</i>
Cool Off	Duração do período de <i>cool off</i>
Interrupções	Incluir os registos interrompidos
Alcance do AP	Raio de alcance do AP
Velocidade	Velocidade mínima para APs consecutivos
Tempo	Tempo máximo entre restabelecimento de ligações

um dos períodos. Alternativamente, a opção *Interrupções* permite considerar registos que começam e/ou terminam durante o cenário.

O algoritmo produz registos de mobilidade no formato do gerador e analisador de cenários Bonnmotion [1], ajustando as estampilhas temporais e criando um *way-point* para cada registo com a localização prevista para cada dispositivo do cenário no instante 0.

3.4 Privacidade dos utilizadores

Apesar de estar planeada a disponibilização à comunidade das rotas geradas pelo algoritmo, para proteger a confidencialidade dos dados de movimento dos utilizadores, os registos originais são guardados numa localização segura e não podem ser disponibilizados. Este modelo protege a privacidade dos utilizadores uma vez que o algoritmo produz exclusivamente pares (data, coordenada) e portanto, nenhuma identificação que possa ser associada diretamente com um utilizador ou dispositivo. Adicionalmente, os dados são ofuscados por: *i*) reposicionamento das coordenadas dos *way-points* em cada nova geração de cenários; e *ii*) início de todos os cenários no instante temporal 0, sem disponibilizar a diferença entre a data de início pedida e o início efetivo do cenário, a qual está sujeita às condicionantes impostas pelas opções e à disponibilidade de dados.

Para limitar potenciais análises dos dados que pudessem ser cruzadas com informação tornada pública a partir de outras fontes, todos os pedidos serão moderados. Limites à duração dos cenários poderão também ser aplicados.

4 Análise da mobilidade

Esta secção apresenta e discute as características dos padrões de mobilidade encontrados no conjunto de registos de mobilidade MobIPLity utilizando como

Tabela 2: Sumário do conjunto de registos de mobilidade de 2013

	Todos		Portáteis		DPDs	
	IPL	ISEL	IPL	ISEL	IPL	ISEL
Dispositivos	35930	12559	15532	6507	17714	5544
Registos de Mobilidade	3035396	1236658	990048	525289	1997136	697065

métricas o tempo entre contactos, comprimento do salto e tempo de pausa respectivamente (do inglês *inter-contact time* (ICT), *jump size* e *pause time*, descritos na Sec. 2). Apesar dos nossos registos conterem os dados desde 2005, este artigo aborda exclusivamente o conjunto de registos de mobilidade de 2013, escolhido por ser o mais recente, por refletir a evolução da tecnologia e por ser o mais rico em termos de número e variedade de tipos de dispositivos. A secção estuda 3 diferentes opções de configuração para os tipos de dispositivos, permitindo confirmar a existência de padrões distintos de mobilidade para utilizadores que transporta dispositivos móveis grandes (“portáteis”) e pequenos (“DPD”).

Para aumentar a diversidade, a análise considera duas localizações: o ISEL e o IPL. A escola da engenharia do IPL (apresentada como ISEL) está localizada num único campus e fornece o maior número de dispositivos disponíveis num único campus. A localização IPL considera os registos obtidos da totalidade dos pontos de acesso de todas as escolas (incluindo o ISEL), contrastando em densidade com o ISEL uma vez que os campus estão distribuídos pela totalidade da área metropolitana (Cf. Fig. 1).

4.1 Métricas genéricas do conjunto de dados

A tabela 2 mostra a dimensão do conjunto de registos estudados. Como pode ser observado, o ISEL totaliza aproximadamente 35% dos dispositivos e 40% do número de registos de mobilidade. Na avaliação são consideradas duas abordagens ao conjunto de dados. Uma global “Por registo de mobilidade” que considera cada registo individual, e outra “Por dispositivo” que agrega os registos de mobilidade de cada dispositivo, avaliando as médias das suas métricas.

Duração dos registos de mobilidade A Fig. 4 apresenta a função de distribuição cumulativa complementar (CCDF) da duração de cada registo de mobilidade. As durações elevadas, especialmente em Portáteis no IPL, são atribuídas à cobertura da rede eduroam às residências de alunos. A figura também mostra que menos de 17% dos registos de Portáteis excedem 2 horas e que para DPDs este valor diminui para cerca de 10% dos registos no IPL e 8% no ISEL.

Esta pequena quantidade de “registos de elevada duração” é surpreendente. De notar por exemplo que, em oposição às nossas expectativas, não existe um conjunto de 6 horas para os quais 100 DPDs estejam permanentemente ligados à rede na totalidade do conjunto de registos de mobilidade de 2013. Era de esperar que o padrão de utilização refletisse a crescente utilização de dispositivos móveis no campus e, portanto, que as durações dos registos de mobilidade fossem

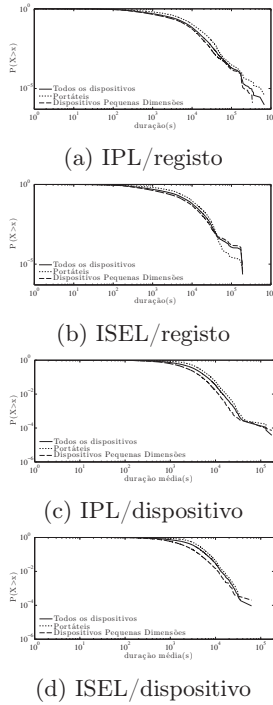


Figura 4: Duração

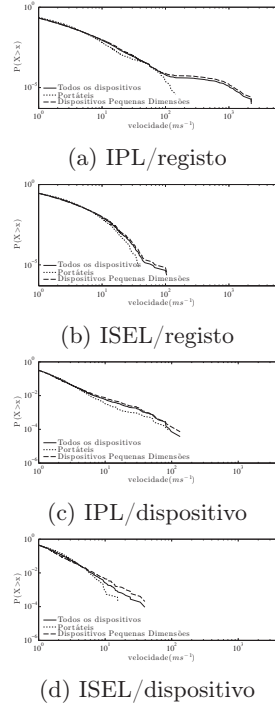


Figura 5: Velocidade

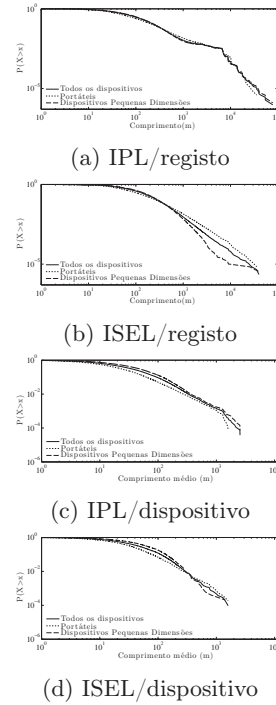


Figura 6: Comprimento

também superiores. De salientar que estamos perante um cenário em que existem mais DPDs que portáteis no caso da totalidade do IPL, e uma diferença pequena no número dos mesmos dispositivos no caso do ISEL.

A pequena duração dos registos é atribuída aos mecanismos de poupança de energia que podem ser encontrados em dispositivos móveis e que desabilitam a interface de rede sem fios quando o dispositivo não está em uso ou o ecrã está desligado. É interessante observar que este aspeto tem sido constantemente ignorado pelos modelos de mobilidade baseados em registos e pelos simuladores de rede. Contudo, esta funcionalidade tem um impacto não negligenciável no desenho e avaliação de muitos protocolos para redes ad hoc e tolerantes à latência (do inglês Delay Tolerant Networks) que assumem que as interfaces de redes sem fios se mantêm ativas para estabelecer ligações entre dispositivos. Investigar como é que a assunção de conectividade permanente tem influenciado os resultados das simulações sai fora do alcance deste artigo e é relegado para trabalho futuro.

Velocidade Os gráficos da velocidade (Fig. 5) exibem alguns resultados marginais atribuídos ao efeito *ping-pong* que resulta de uma combinação de mudanças rápidas de pontos de acesso sobrepostos com o algoritmo de geração de registos (Cf. Fig. 3b). Este é um problema que tem sido encontrado noutros modelos

(e.g. [12]) e tem um impacto negligenciável uma vez que estas velocidades ocorrem durante pequenos períodos de tempo e distâncias muito curtas. É de relevar que uma porção de 15% dos registos de mobilidade têm uma velocidade média acima da velocidade típica de um ser humano no conjunto de dados do IPL, que é atribuída à circulação dos utilizadores entre localizações do IPL. Ao mesmo tempo, cerca de 70% dos dispositivos têm velocidades médias inferiores a 1m/s no IPL. Para o ISEL o mesmo conjunto de dispositivos desce para 55%, o que sugere uma mobilidade superior no ISEL. As figuras mostram também uma maior velocidade média dos DPDs, o que não surpreende e confirma a possibilidade de utilização destes dispositivos com os utilizadores em movimento.

Distâncias percorridas A Fig. 6 apresenta a distancia percorrida em cada registo. Um número muito limitado dos registos (0,001%) atingem valores surpreendentes de quase 80 Km e que são atribuídos a viagens entre os campus do IPL. As médias por dispositivo são mais previsíveis e apenas atingem 2500m para o IPL e 1500m para o ISEL. A já esperada maior mobilidade dos DPDs é confirmada por registos de mobilidade com dimensões médias maiores em conjunção com durações mais curtas.

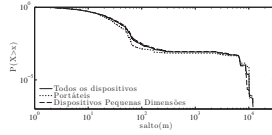
Os resultados mostram que cerca de 27% de todos os dispositivos do IPL e 16% do ISEL são estáticos, contribuindo para cerca de 65% de registos sem movimento. Contudo, o tamanho do conjunto de dados é suficiente para atenuar o seu impacto uma vez que, em absoluto, mais de 1 milhão registos de mobilidade para o IPL apresentam movimento entre pelo menos 2 pontos de acesso. Os autores acreditam que este valor elevado deve ser considerado suficiente para suportar um modelo de mobilidade utilizando exclusivamente registos de acesso.

A distribuição dos saltos (i.e. a distância viajada entre *way-points*) é visível na Fig. 7. A CCDF em “escada” é atribuída à distribuição geográfica dos campus do IPL, com cada degrau a representar a distância entre um par de campus.

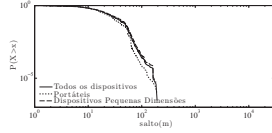
Tempos de pausa A Fig. 8 mostra que os DPDs exibem generalizadamente tempos de pausa mais curtos. Isto suporta o senso comum de que os DPDs terão uma maior mobilidade, que contrasta com os esperados tempos de pausa elevados dos portáteis, tipicamente utilizados de forma estática. Os tempos de pausa no IPL são ainda mais amplificados pela conectividade dos dispositivos nos quartos das residências de alunos.

4.2 Comparação com modelos sintéticos

Para facilitar a comparação com modelos de mobilidade sintéticos baseados em registos de acessos, foi utilizado o teste de Akaike para comparar a adequação dos registos de mobilidade do ISEL a distribuições bem conhecidas. Os registos de mobilidade do ISEL foram escolhidos devido ao facto de estarem limitados a um único campus, refletindo assim um ambiente mais próximo do utilizado noutros estudos. Os resultados apresentados na Tabela 3 foram obtidos utilizando o Matlab para efetuar a adequação de entre todas as distribuições suportadas,

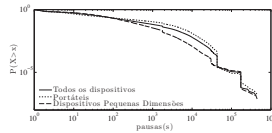


(a) IPL/registro

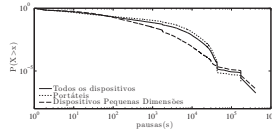


(b) ISEL/registro

Figura 7: Saltos

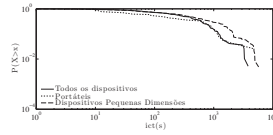


(a) IPL/registro

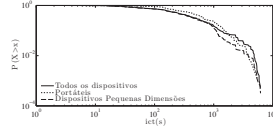


(b) ISEL/registro

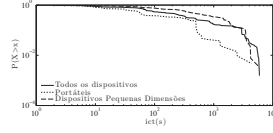
Figura 8: Pausas



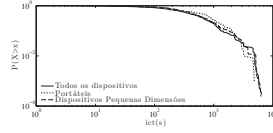
(a) IPL (3 dias)



(b) ISEL (desconexo)

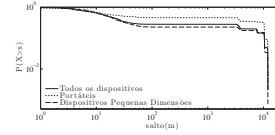


(c) ISEL (3 dias)

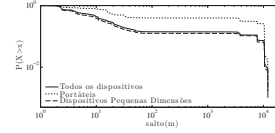


(d) ISEL (desconexo)

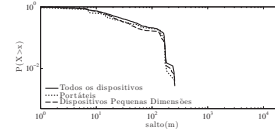
Figura 9: ICT



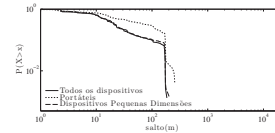
(a) IPL (3 dias)



(b) ISEL (desconexo)



(c) ISEL (3 dias)



(d) ISEL (desconexo)

Figura 10: Saltos

ordenadas pelo critério Akaike. Salienta-se a diferença da adequação das diferentes distribuições estatísticas aos diferentes tipos de dispositivos onde na maioria dos casos para os DPDs a Pareto é a distribuição mais adequada, mas no caso dos portáteis esta já varia consoante a métrica, isto confirma as nossas suspeitas da existência de modelos de mobilidade distintos para estes equipamentos. A tabela apresenta ainda a distribuição modelada pelos modelos SLAW [13] e Dartmouth [12] para as métricas onde foi possível obter essa informação. São visíveis as diferenças de modelação por exemplo nos tempos de pausa e distância de saltos. A modelação do ICT não foi considerada uma vez que a extração desta métrica para o conjunto de dados totais para 2013 não é possível em tempo útil.

4.3 Métricas do cenário

Esta secção avalia as métricas de 2 cenários distintos, com características comparáveis às encontradas no trabalho relacionado. O cenário “3 dias”, agrega os registos de mobilidade de 3 dias onde foram observados mais DPDs no ISEL, nomeadamente 9 de Outubro, 20 de Junho e 19 de Dezembro. Este critério foi

Tabela 3: Distribuições observadas

	Todos os dispositivos (SLAW/Dartmouth)	Portáteis	DPDs
Duração	Log-normal	Weibull	Log-normal
Duração/disp	Weibull	Weibull	Pareto
Comprimento	Pareto	Pareto	Pareto
Comprimento/disp	Pareto	Pareto	Pareto
Velocidade	Pareto	Weibull	Pareto
Velocidade/disp	Pareto (-/Exponential)	Pareto	Pareto
Pausa	Log-normal (Pareto/Log-normal)	Log-normal	Pareto
Salto	Pareto (Pareto/-)	Weibull	Pareto

utilizado num esforço para ultrapassar as dificuldades em encontrar registos de pelo menos 2 horas, discutido na Sec. 4.1. O MobIPLity foi configurado para selecionar registos de 2 horas com os períodos de *warm up* e de *cool off* configurados com igual duração. A velocidade mínima para saltos foi estabelecida em 0.5ms^{-1} e o tempo máximo de restabelecimento de ligações em 120s. É assumido igualmente que tanto os pontos de acesso como os dispositivos têm um raio de alcance de 50m. Com estes parâmetros, um segundo cenário, chamado “desconexo”, foi extraído para apenas o dia 9 de Outubro, e inclui os registos de mobilidade que terminam ou têm início durante o período de 2h.

As métricas apresentadas consideram todos os registos de mobilidade dos DPDs extraídos para os dias em análise (respetivamente, 46, 43 e 42 para o ISEL e 90, 76 e 100 para o IPL). Os cenários “desconexo” e “3 dias” partilham todos os restantes parâmetros de configuração com a exceção da opção para incluir registos interrompidos durante o período. A expectativa é que as diferenças observadas nos forneçam algumas pistas sobre o impacto dos mecanismos de poupança de energia dos dispositivos nos modelos utilizados atualmente.

ICT A Fig. 9 apresenta a CCDF para os tempos entre contactos dos 6 cenários em análise. O limiar do alcance de rádio para determinação do contacto é de 50 metros. A irregularidade dos gráficos para os cenários “3 dias” é atribuída ao número pequeno de dispositivos que permanecem ligados durante o período de 2h. Os resultados mostram tempos entre contactos semelhantes entre o ISEL e o IPL, algo já defendido pelos autores do SLAW [13]. Isto sugere que apesar das diferenças geográficas entre ambos os cenários, a distribuição de ICTs mantém-se. É interessante notar que estes resultados estão na linha dos ICTs para cenários criados utilizando o SLAW e configurados para emular o ISEL e o IPL e que são apresentados na Fig. 12. Infelizmente, o SLAW não distingue entre tipos de dispositivos apesar dos nossos resultados mostrarem que o tipo de dispositivo tem um papel importante no ICT, com os DPDs tipicamente a apresentarem ICTs mais curtos, o que é atribuído à sua maior mobilidade.

Comprimento do salto A CCDF para os comprimento dos saltos para o IPL (Fig. 10) mostra um padrão incomum em degraus atribuído às distâncias consideráveis entre as diferentes escolas do instituto e à necessidade de alguns utiliza-

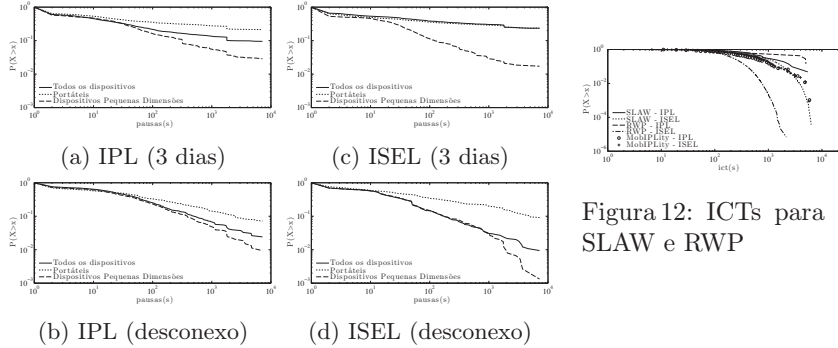


Figura 11: Pausas

dores circularmente entre algumas delas. Como esperado, a área consideravelmente mais pequena do ISEL restringe a distância máxima viajada em comparação com o IPL. É interessante reparar nas diferenças entre DPDs e portáteis, onde os saltos dos primeiros têm uma probabilidade maior de serem mais curtos que os dos últimos, algo que pode ser atribuído à mobilidade dos dispositivos, que permite ligações mais frequentes a APs distintos.

A Fig. 10c) apresenta uma exceção interessante à relação das curvas apresentadas por DPDs e portáteis dado que estes têm probabilidades semelhantes de se moverem ao longo do ISEL. Contudo, a Fig. 10d) contradiz os resultados. Uma vez que as diferenças entre ambos os registos de mobilidade é restrito à velocidade mínima do percurso (que na Fig. 10c) tem de ser superior a 0.5m/s), é razoável afirmar que o comportamento anormal é devido à velocidade a que os dispositivos viajam. Em geral, o resultados no comprimento dos saltos tende a justificar a afirmação de que os DPDs têm uma mobilidade superior, que produz registos de mobilidade maiores, atravessando múltiplos APs, enquanto os portáteis são desligados e ligados de novo numa outra localização.

O comprimento dos saltos tem uma relação direta com a dispersão física do IPL, que cria grupos de utilizadores, os quais permanecem próximos numa única escola ou viajam ao longo de várias. Isto já foi anteriormente identificado por exemplo, no HCMM [5] onde a geração de cenários considera a possibilidade de definição de um número de grupos e a criação de distribuições normais para modelar o comprimento do salto para cada grupo.

Tempos de pausa Os tempos de pausa são particularmente pequenos uma vez que a metodologia seguida para definição de registos de mobilidade tende a manter um nó em movimento mesmo que a uma velocidade muito baixa. Nos resultados apresentados na Fig. 11, os dispositivos são considerados parados se a distância entre *way-points* for inferior a 1 metro. Mesmo assim, é interessante observar que os DPDs têm distribuições diferentes de tempos de pausa, com probabilidades mais baixas de terem tempos de pausa superiores, algo que suporta as características de mobilidade esperadas dos DPDs.

Tabela 4: Média e desvio padrão dos valores de não homogeneidade

	Todos (σ)	Portáteis (σ)	DPDs (σ)		Todos (σ)
IPL 1-dia	0,84 (0,01)	0,96 (0,006)	0,89 (0,006)	IPL RWP	0,36 (0,03)
ISEL 1-dia	0,64 (0,02)	0,69 (0)	0,58 (0,07)	ISEL RWP	0,35 (0,03)
IPL (desc.)	0,71 (0,02)	0,71 (0,01)	0,75 (0,02)	IPL SLAW	0,65 (0,02)
ISEL (desc.)	0,62 (0,005)	0,64 (0,02)	0,63 (0,03)	ISEL SLAW	0,5 (0,05)

Os nossos resultados contrastam com os apresentados em [12] (Cf. 4.2) onde os tempos de pausa eram definidos pela deteção de utilizadores que caminhavam a uma velocidade muito baixa.

Não homogeneidade A Tabela 4 apresenta a métrica de não homogeneidade, apresentada em [17]. A amostra para esta métrica foi selecionada em 3 instantes (início, 1/3 e 2/3 do tempo decorrido), para o dia com mais DPDs na rede. Os portáteis apresentam uma maior não homogeneidade, indicando uma maior concentração e distribuição irregular para estes dispositivos, sugerindo que estão agrupados, por exemplo em salas de aula ou bibliotecas. O valor mais baixo da não homogeneidade para os DPDs confirma a disposição pseudo-aleatória dos dispositivos móveis. De novo, a metodologia para extração dos registos de mobilidade que excluí as trajetórias que foram interrompidas durante o período, produz algumas incongruências com o esperado, como por exemplo para o cenário IPL. A Tab. 4 apresenta igualmente os resultados da não homogeneidade para cenários gerados pelo SLAW que tentam replicar as condições do IPL e do ISEL (área, duração, dispositivos e número de *hotspots*/pontos de acesso). Os resultados dos nossos registos de mobilidade são semelhantes aos obtidos pelo SLAW, e, como esperado divergem da aleatoriedade encontrada no RWP onde o valor da métrica de não homogeneidade é baixo [8].

5 Conclusões e Trabalho Futuro

A utilização registos de mobilidade reais contribui para a qualidade da avaliação e para a modelação de novos protocolos, aplicações e algoritmos. Este artigo apresenta um conjunto de dados composto pelos registos de acesso à rede eduroam do Instituto Politécnico de Lisboa durante o ano de 2013 e uma metodologia para extrair cenários de mobilidade a partir deste conjunto.

A recência do conjunto de dados é uma mais valia por refletir os padrões de mobilidade atuais e a popularidade dos dispositivos de pequenas dimensões. Em comparação com modelos de geração sintéticos, verifica-se que os mais populares tendem a reproduzir com precisão razoável os padrões efetivamente observados. Contudo, a avaliação mostra que os mecanismos de poupança de energia que são comuns nestes dispositivos reduzem as possibilidades de comunicações espontâneas entre dispositivos e com o ambiente. Este é um aspeto que tem sido ignorado por modelos de mobilidade, simuladores de rede e designers de aplicações e que pode ter um impacto não negligenciável em ambientes de produção.

Como trabalho futuro os autores irão investigar os padrões de contactos entre utilizadores e as mudanças na utilização da rede e dispositivos ao longo do tempo.

Referências

1. Aschenbruck, N., Ernst, R., Gerhards-Padilla, E., Schwamborn, M.: Bonnmotion: A mobility scenario generation and analysis tool. In: *Procs. of the 3rd Int'l ICST Conf. on Simulation Tools and Techniques* (2010)
2. Aschenbruck, N., Gerhards-Padilla, E., Gerharz, M., Frank, M., Martini, P.: Modelling mobility in disaster area scenarios. In: *Procs. of the 10th ACM Symp. on Modeling, Analysis, and Simulation of Wireless and Mobile Systems* (2007)
3. Aschenbruck, N., Munjal, A., Camp, T.: Trace-based mobility modeling for multi-hop wireless networks. *Computer Communications* (2011)
4. Bettstetter, C., Resta, G., Santi, P.: The node distribution of the random waypoint mobility model for wireless ad hoc networks. *Mobile Computing, IEEE Trans. on* (2003)
5. Boldrini, C., Passarella, A.: Hcmm: Modelling spatial and temporal properties of human mobility driven by users' social relationships. *Computer Comms.* (2010)
6. Chaintreau, A., Hui, P., Crowcroft, J., Diot, C., Gass, R., Scott, J.: Impact of human mobility on opportunistic forwarding algorithms. *Mobile Computing, IEEE Trans. on* (2007)
7. Fischer, D., Herrmann, K., Rothermel, K.: Gesomo: A general social mobility model for delay tolerant networks. In: *Mobile Adhoc and Sensor Systems (MASS), 2010 IEEE 7th Int'l Conf. on* (2010)
8. Gyarmati, M., Schilcher, U., Brandner, G., Bettstetter, C., Chung, Y.W., Kim, Y.H.: Impact of random mobility on the inhomogeneity of spatial distributions. In: *IEEE Global Telecommunications Conf. (GLOBECOM)* (2008)
9. Henderson, T., Kotz, D., Abyzov, I.: The changing usage of a mature campus-wide wireless network. *Computer Networks* (2008)
10. Karagiannis, T., Le Boudec, J.Y., Vojnovic, M.: Power law and exponential decay of intercontact times between mobile devices. *Mobile Computing, IEEE Trans. on* (2010)
11. Karamshuk, D., Boldrini, C., Conti, M., Passarella, A.: Human mobility models for opportunistic networks. *Communications Magazine, IEEE* (2011)
12. Kim, M., Kotz, D., Kim, S.: Extracting a mobility model from real user traces. In: *INFOCOM 2006. 25th IEEE Int'l Conf. on Computer Comms. Procs.* (2006)
13. Lee, K., Hong, S., Kim, S.J., Rhee, I., Chong, S.: Slaw: A new mobility model for human walks. In: *INFOCOM 2009, IEEE* (2009)
14. Mulhanga, M., Lima, S., Carvalho, P.: Characterising university wlans within eduroam context. In: *Smart Spaces and Next Generation Wired/Wireless Networking. Springer* (2011)
15. Piorowski, M., Sarafijanovic-Djukic, N., Grossglauser, M.: CRAWDAD data set epfl/mobility (v. 2009-02-24). <http://crawdad.cs.dartmouth.edu> (Feb 2009)
16. Rigney, C.: RADIUS Accounting. RFC 2866 (Jun 2000)
17. Schilcher, U., Gyarmati, M., Bettstetter, C., Chung, Y.W., Kim, Y.H.: Measuring inhomogeneity in spatial distributions. In: *Vehicular Technology Conf., 2008. VTC 2008. IEEE* (2008)
18. Thiagarajan, A., Ravindranath, L., Balakrishnan, H., Madden, S., Girod, L.: Accurate, low-energy trajectory mapping for mobile devices. In: *Procs. of the 8th USENIX Conf. on Networked Systems Design and Implementation* (2011)