Mobile Crowd-Sensing Wireless Activity with Measured Interference Power

Weisi Guo and Siyi Wang

***Abstract*—In this paper, we present a novel method for sensing the volume of wireless throughput of other users by measuring the passive interference power received on mobile devices. The benefit of this mobile crowd-sensing approach is that it offers a non-intrusive way of inferring the level of wireless traffic, without requiring to extract data from commercial or private devices. The availability of cross-network wireless traffic data can allow individual operators to deploy femto-cells effectively. The proposed technique requires only approximate location data and is independent of the traffic pattern. The results show that the sensing accuracy for cell-level activity is high ($< 5\%$ error). The paper further demonstrates how this novel method can allow researchers to rapidly build up a map of the wireless activity level in a large region using crowd-sensing.**

***Index Terms*—Capacity planning, sensor network, stochastic geometry, urban sensing, wireless traffic.**

I. Introduction

OVER the past decade, there has been an exponential growth in the volume of wireless data transfer, but the spatial- and temporal-variation of wireless activity volume is not well understood for any given region [1]. This understanding has profound implications to both telecommunication operators in femto-cell planning [2], and social-economic policy makers. Whilst mobile voice call patterns are well modelled, less is understood regarding data demand across operators and networks (Wi-Fi, Cellular, and Wi-Max).

Existing work has shown that the data traffic volume in 3G cellular networks follow complex distributions [1], [3], [4], and there is a close association between the traffic volume and the capacity of the serving cell. However, such studies are usually performed over a relatively short period, for a specific network and location area. For a given location, it is difficult to maintain an updated model of the wireless traffic across multiple networks. Currently, the only approach is to extract usage data from a large sample of basestations, access-points, and mobile phones. This is clearly a difficult process that carries privacy and complexity issues.

More recently, inferring mobile data traffic using publicly available application data such as Twitter has been employed. However, there is a sampling bias related to application specific data mining. This paper proposes an inference method that uses the received interference power on a single mobile device to infer the activity intensity of wireless data transfer in an area (as illustrated in Fig. 1). A number of mobile sensors that can measure the aggregate interference power on individual sub-bands can map out the wireless activity (traffic) pattern of a large area. The method is neutral to the traffic pattern and does not incur privacy risks.

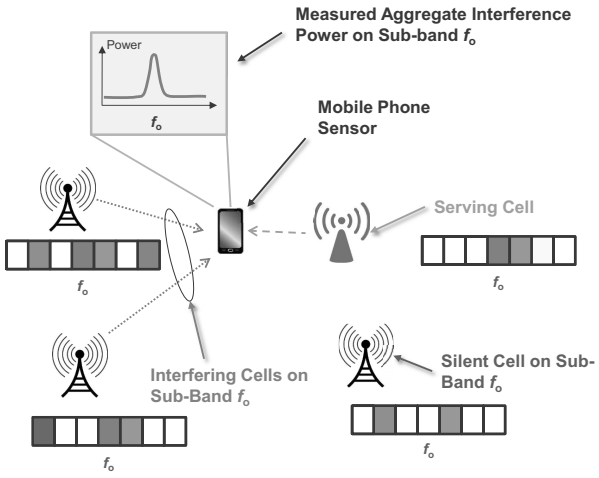

Fig. 1. Illustration of a mobile device that receives aggregate interference power on sub-band f_0 from multiple interfering cells. Some potentially interfering cells are silent on this particular sub-band. The aggregate traffic level in this area can be inferred from the interference power measured.

*The **ambiguity** this paper resolves is that several transmitters far away can cause the same aggregate interference power as a single transmitter nearby. The challenge is how to discern how many transmitters are transmitting accurately?* The analysis uses statistical methods to discern the number of transmitters that are active on each sub-band.

II. Statistical Framework

A. Aggregate Interference Power

The paper considers a mobile device at an arbitrary location using an OFDMA based communication system, but this can also be applied to CDMA systems. The device receives a signal power from the nearest serving cell on a frequency sub-band and aggregate interference power from all other cells transmitting on the same sub-band (Fig. 1). The analysis now derives an analytical expression for the probability distribution of the aggregate interference power of a sub-band using spatial poisson point process (SPPP) denoted as Φ with active transmitting cell intensity λ [5]. It is worth noting that the interference power derived is not for a single user location, but rather the average across all users distributed uniformly and randomly.

Consider the particular value I_r of random variable (R.V.) I_R denoted as the aggregate interference power received at the

Manuscript received June 28, 2013. The associate editor coordinating the review of this letter and approving it for publication was I. Guvenc.

W. Guo is with the School of Engineering, University of Warwick, United Kingdom (e-mail: weisi.guo@warwick.ac.uk).

S. Wang is with the Institute for Telecommunications Research, University of South Australia, Australia (e-mail: siyi.wang@mymail.unisa.edu.au).

Digital Object Identifier 10.1109/WCL.2013.070913.130335

mobile device:

$$I_r = \sum_{\substack{i \in \Phi \\ i \neq l}} h_i P \Lambda r_i^{-\alpha}, \tag{1}$$

where h_i is the multi-path fading gain, l is the serving cell index in SPPP Φ, P is the transmit power and is assumed to be same for all cells, Λ is the pathloss constant, r_i is the distance from interfering transmitter to receiver, and α is the pathloss exponent [6]. The letters will denote R.V. as capital letters and particular values as lower-case letters. The paper considers Rayleigh fading with mean 1 and defines R.V. $G = HP\Lambda$, which therefore follows the exponential distribution with mean β denoted by $G \sim \exp(\beta)$, where $\beta = 1/P\Lambda$.

The **moment generating function** of I_R is defined as follows:

$$M_{I_R}(s) \triangleq \mathbb{E}(e^{sI_r}) \quad (I_r > 0). \tag{2}$$

Since the Laplace transform of a function $f(t)$, defined for all real numbers $t \geqslant 0$, is the function $F(s)$, given by:

$$F(s) \triangleq \mathscr{L}[f(t)](s) = \int_0^{+\infty} e^{-st} f(t)\ \mathrm{d}t, \tag{3}$$

Equation (2) evaluated at $-s$ can be re-written as the Laplace transform of the PDF of I_R:

$$\begin{aligned} M_{I_R}(-s;\alpha) &= \mathscr{L}[f_{I_R}(I_r)](s) = \mathbb{E}(e^{-sI_r}), \\ &= \exp\left[-\frac{2\pi^2\lambda \csc\left(\frac{2\pi}{\alpha}\right)}{\alpha}\left(\frac{s}{\beta}\right)^{\frac{2}{\alpha}}\right], \end{aligned} \tag{4}$$

where the full proof is given in the Appendix. Therefore, the probability density function (PDF) of I_R is obtained on taking the inverse Laplace transform as:

$$\begin{aligned} f_{I_R}(I_r;\alpha) &= \mathscr{L}^{-1}[M_{I_R}(-s)](I_r), \\ &= \mathscr{L}^{-1}\left\{\exp\left[-\frac{2\pi^2\lambda \csc\left(\frac{2\pi}{\alpha}\right)}{\alpha}\left(\frac{s}{\beta}\right)^{\frac{2}{\alpha}}\right]\right\}(I_r), \\ &= \frac{\pi^{3/2}\lambda \csc\left(\frac{2\pi}{\alpha}\right)\exp\left[-\frac{\pi^4\lambda^2\csc^2\left(\frac{2\pi}{\alpha}\right)}{\alpha^2\beta I_r}\right]}{\alpha\sqrt{\beta}I_r^{3/2}}. \end{aligned} \tag{5}$$

The cumulative density function (CDF) of I_R is thus given by:

$$\begin{aligned} F_{I_R}(\zeta;\alpha) &= \int_0^{\zeta} f_{I_R}(I_r;\alpha)\ \mathrm{d}I_r, \\ &= \int_0^{\zeta} \frac{\pi^{3/2}\lambda \csc\left(\frac{2\pi}{\alpha}\right)\exp\left[-\frac{\pi^4\lambda^2\csc^2\left(\frac{2\pi}{\alpha}\right)}{\alpha^2\beta I_r}\right]}{\alpha\sqrt{\beta}I_r^{3/2}}\ \mathrm{d}I_r, \\ &= \mathrm{erfc}\left[\frac{\pi^2\lambda \csc\left(\frac{2\pi}{\alpha}\right)}{\alpha\sqrt{\beta\zeta}}\right]. \end{aligned} \tag{6}$$

B. Non-Line-of-Sight (NLOS) Environment

In NLOS environments, the pathloss distance exponent is typically of a value $\alpha = 4$ [6]. In this special case, the CDF of the aggregate interference power can be simplified to:

$$F_{I_R}(\zeta;4) = \mathrm{erfc}\left(\frac{\pi^2\lambda}{4\sqrt{\beta\zeta}}\right), \tag{7}$$

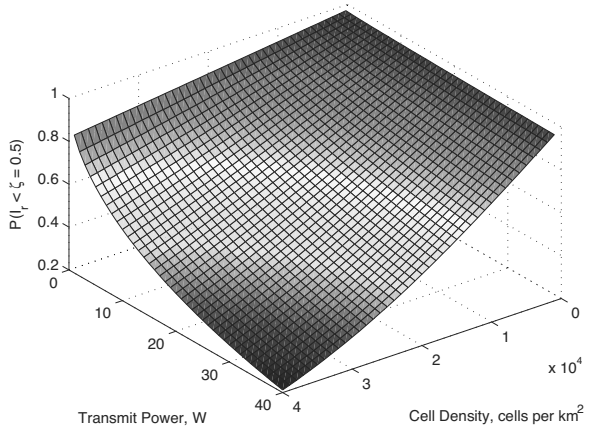

Fig. 2. Probability of aggregate interference smaller than a certain value ($I_R < \zeta$), for varying values of cell density λ and cell transmit power P, from Eq.(7) with $\zeta = 0.5$.

and the PDF is simplified to:

$$f_{I_R}(I_r;4) = \frac{\lambda}{4\sqrt{\beta}}\left(\frac{\pi}{I_r}\right)^{\frac{3}{2}}\exp\left(-\frac{\pi^4\lambda^2}{16\beta I_r}\right). \tag{8}$$

It can be observed that the probability distributions have a tight relationship with the following parameters:

- Active Cell Density λ: as the density increases, the expected interference power increases;
- Transmit Power $\beta \propto P$ (Λ is a constant): as the transmit power increases, the expected interference power increases;

which are both intuitive results and shown in Fig. 2.

The expectation of the interference power can be found using the complementary CDF:

$$\mathbb{E}(I_R(\zeta;4)) = \int_0^{+\infty} \mathrm{erf}\left(\frac{\pi^2\lambda}{4\sqrt{\beta\zeta}}\right)\mathrm{d}\zeta, \tag{9}$$

but it does not converge absolutely. An easier average to find is the median interference power $\widetilde{\zeta}$, which can be found by setting $F_{I_R}(\widetilde{\zeta};4) = 0.5$:

$$\widetilde{\zeta} = \left[\frac{\pi^2\lambda}{4\sqrt{\beta}\mathrm{erfc}^{-1}(0.5)}\right]^2. \tag{10}$$

In the next section, we estimate the wireless data transfer volume using the estimated interference power.

III. Wireless Activity Estimation

A. Theory

The paper considers an OFDMA downlink network with a fixed deployment of co-frequency cells with density χ. Assuming a random uniform user distribution, then the maximum achievable throughput of a cell C can be found using Monte-Carlo simulation or stochastic geometry methods [5], [6]. This is the *fully loaded* throughput where each sub-band radio resource is utilized. The load of a cell is defined as the ratio between the traffic demanded in a cell T and the maximum

TABLE I
SIMULATION PARAMETERS FOR A 4G NETWORK IN A TYPICAL METROPOLITAN AREA.

Parameter	Value
LTE Cellular System	OFDMA Downlink
Deployed Cell Density, χ	1–10 per km^2
Coverage Area	9 km^2
Transmission Frequency	2.6 GHz
Transmission Bandwidth	20 MHz
Number of Sub-bands	100
Pathloss Model	Urban Micro [6]
Frequency Dependent Pathloss, Λ	40 dB
Pathloss Distance Exponent, α	4
User Arrival Process	Poisson
Data Demand per User Session	[1]
Fading Gain, h	Rayleigh
Shadow Fading	Log-normal
Transmit Power, P	40 W
Sub-band Power Samples	50 per sub-band
Location Samples	400 per km^2

achievable throughput of a cell: $L = \frac{T}{C}$. The average load of multiple cells is the activity level.

Statistically, the load ratio L also corresponds to the proportion of sub-band radio resources utilized. Consider a particular transmission sub-band f_0, such that at any particular time, there are λ_{f_0} of the χ cells transmitting on the particular sub-band over a particular area (as shown in Fig. 1). Therefore, the paper defines the real network activity level on a particular sub-band as the ratio between actively transmitting cells and total deployed cells:

$$A_{\text{real}} = \frac{\lambda_{f_0}}{\chi}. \tag{11}$$

Let us assume a mobile device measures a certain average aggregate interference power of $\overline{\zeta}$. Then the sensed activity level at that frequency band is given by combining Eq.(10) and Eq.(11):

$$\boxed{A_{\text{sensed}} = \frac{4\sqrt{\overline{\zeta}\beta}\text{erfc}^{-1}(0.5)}{\pi^2\chi}}. \tag{12}$$

It is worth noting that the analysis has not resolved the ambiguity of which specific set of interference cells are actively transmitting. Instead, the stochastic analysis reveals the percentage of cells that are actively transmitting. The paper now considers a proof of concept demonstrator to compare theory with practice.

B. Validation and Accuracy

In the investigation, the simulation considers a realistic 4G network with deployment parameters given in Table I. The paper first compares real simulated network activity level (A) with the sensed activity level (A') for different number of interference power samples per sub-band and deployed cell densities. The results in Fig. 3 show that in order to achieve below 5% error with 2 cells per km^2, at least 40–50 samples per sub-band are needed. When shadow fading (variance 9 dB) is considered, the sensing errors compared to that of no fading is approximately 20% larger.

Figure 4 shows simulated real network activity and sensed network activity for 2 different times during the day with

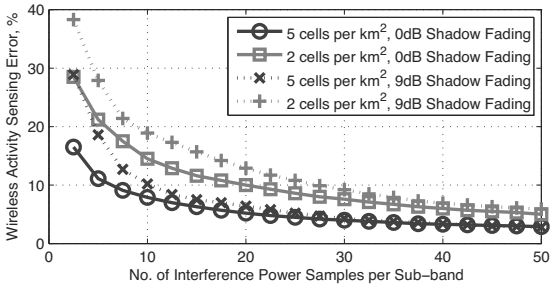

Fig. 3. Sensing error as a function of number of interference power samples per sub-band and the deployed cell density χ.

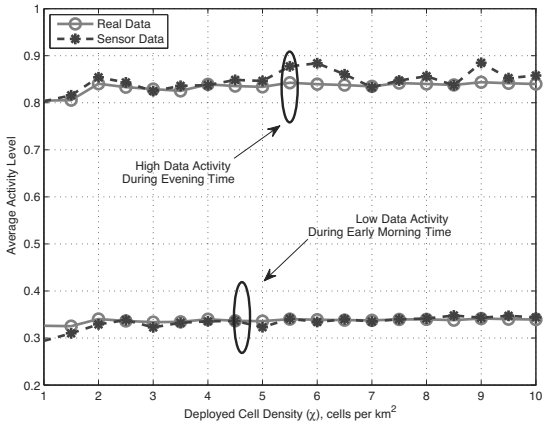

Fig. 4. Simulation of real network activity level (A) and sensed network activity level (A') for different times during the day with varying network deployment density.

50 samples per sub-band. Given that the average data-packet session of a user is approximately 6 seconds [1], the sensing time required is in the order of 5 minutes to obtain an accurate portrayal of the traffic model used. This sensing time can be divided amongst multiple sensors in a single cell coverage area. Provided that the sensor knows which cell it is served by, the mobility of the sensor is not an issue. Therefore, this crowd-sourcing of information can be performed by a number of smartphones with low resolution location information.

As shown by (12), the estimation requires knowledge of certain parameters. The number of cells deployed χ can be treated as a relatively static parameter, which can be found using existing cell-site databases. The pathloss distance exponent α and pathloss constant β are statistical parameters derived from empirical measurements of certain pathloss scenarios [6]. Whilst this is susceptible to non-uniform effects in propagation (shadow fading and antenna patterns), the results in the next section will show that statistical averaging of multiple readings can reduce the sensing errors.

C. Crowd Sensing

The paper now considers a 25 km^2 area, of which 18 cells are deployed in accordance to a Vodafone London network configuration. Over 10,000 uniform spatial samples are taken, where each spatial sample consists of 50 samples per sub-band over 100 sub-bands to yield a sensing error of less than 5%.

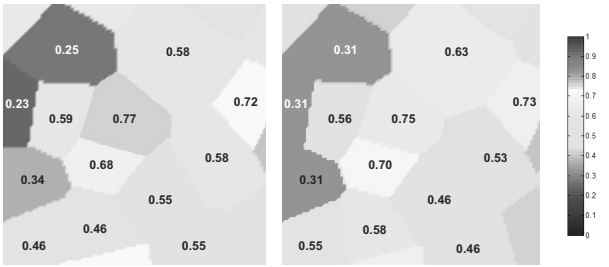

Fig. 5. Wireless data traffic activity rate for a particular period in time: a) Real simulated data based on a user arrival process in cells [1]; b) Sensed data using Eq. (12) and taking averages.

Figure 5a shows the real simulated data based on a user arrival and data demand process [1]. It is worth stressing that the sensing technique is independent of the traffic model and the model in this paper is for demonstrative purposes. Fig. 5b shows the sensed activity data using Eq. (12). The errors due to non-linear effects are reduced by taking averages of the interference power. Fig. 5 shows that the sensing data can offer a sufficiently accurate understanding of how the traffic level varies both in space, and also in time (when this is re-sampled at another time interval). The resolution is poorer on the edge of Fig. 5 due to the lack of sensing data. Over the coarse of a longer period of time, one is able to obtain interesting statistics about the spatial- and temporal-distribution of wireless traffic. This study is beyond the scope of this paper, which only serves to introduce the possibility of this concept. The wider applications of understanding wide band wireless traffic patterns include sustainable small-cell deployment [7].

IV. Conclusions

In this paper, we presented a novel method for sensing the volume of wireless activity at any frequency using the passive interference power. The benefit of this approach is that it offers a non-intrusive way of accurately inferring the level of wireless activity, without requiring to extract data from commercial operators or tapping each wireless device. The results show that given sufficiently accurate interference power measurements, the sensing accuracy for cell-level activity is high ($< 5\%$ error) across a wide range of cell density and activity levels. The paper further demonstrates how this novel method can allow researchers to rapidly build up a map of the wireless activity level in a large region through crowd-sensing. The availability of cross-network wireless traffic data can allow operators to deploy femto-cells effectively and allow better understanding of the digital economy.

Appendix

The moment generating function of I_R is:

$$\begin{aligned} M_{I_R}(-s;\alpha) &= \mathscr{L}[f_{I_R}(I_r)](s) = \mathbb{E}(e^{-sI_r}), \\ &= \mathbb{E}_\Phi\left\{\mathbb{E}_G\left[\exp\left(-s\sum_{\substack{i\in\Phi\\ i\neq l}} g_i r_i^{-\alpha}\right)\right]\right\}, \\ &= \mathbb{E}_\Phi\left\{\prod_{\substack{i\in\Phi\\ i\neq l}} \mathbb{E}_G\left[\exp\left(-sg_i r_i^{-\alpha}\right)\right]\right\}, \\ &= \mathbb{E}_\Phi\left(\prod_{\substack{i\in\Phi\\ i\neq l}} \frac{\beta}{\beta + sr_i^{-\alpha}}\right). \end{aligned} \tag{13}$$

The next step follows from the i.i.d. distribution of G and its further independence from the SPPP Φ:

$$\begin{aligned} M_{I_R}(-s;\alpha) &= e^{-2\pi\lambda\int_0^{+\infty}\left(1-\frac{\beta}{\beta+sv^{-\alpha}}\right)v\ \mathrm{d}v}, \\ &= \exp\left[-\pi\lambda\left(\frac{s}{\beta}\right)^{\frac{2}{\alpha}}\int_0^{+\infty}\frac{1}{1+u^{\frac{\alpha}{2}}}\ \mathrm{d}u\right], \\ &= \exp\left[-\frac{2\pi^2\lambda\csc\left(\frac{2\pi}{\alpha}\right)}{\alpha}\left(\frac{s}{\beta}\right)^{\frac{2}{\alpha}}\right], \end{aligned} \tag{14}$$

where the last step follows from the probability generating functional (PGFL) [8] of the SPPP and the manipulation results from converting from Cartesian to polar coordinates and the variable u is substituted for $v^2\left(\beta/s\right)^{\frac{2}{\alpha}}$ at the penultimate step above.

References

[1] M. Laner, P. Svoboda, S. Schwarz, and M. Rupp, "Users in cells: a data traffic analysis," in *Proc. 2012 IEEE Wireless Communications and Networking Conference*, pp. 3063–3068.
[2] W. Guo and S. Wang, "Interference-aware self-deploying femto-cell," *IEEE Wireless Commun. Lett.*, vol. 1, pp. 609–612, Nov. 2012.
[3] M. Shafiq, L. Ji, A. Liu, J. Pang, and J. Wang, "Characterizing geospatial dynamics of application usage in a 3G cellular data network," in *Proc. 2012 IEEE INFOCOM*, pp. 1341–1349.
[4] Y. Zhang and A. Arvidsson, "Understanding the characteristics of cellular data traffic," in *Proc. 2012 ACM SIGCOM*, pp. 13–18.
[5] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," in *IEEE J. Sel. Areas Commun.*, vol. 28, pp. 1029–1046, Sep. 2009.
[6] 3GPP, "TR36.814 V9.0.0: Further Advancements for E-UTRA Physical Layer Aspects (Release 9)," 3GPP, Technical Report, Mar. 2010.
[7] W. Guo and T. O'Farrell, "Capacity-energy-cost tradeoff for small-cell networks," in *2012 IEEE Vehicular Technology Conference – Spring*.
[8] D. Stoyan, W. S. Kendall, J. Mecke, and L. Ruschendorf, *Stochastic Geometry and its Applications, Vol. 2*. Wiley, 1987.