Abstract—The defining characteristic of wireless and mobile networking is user mobility, and related to it is the ability for the network to capture (at least partial) information on where users are located and how users change location over time. Information about location is becoming critical, and therefore valuable, for an increasingly larger number of location-based or location-aware services. One key open question, however, is how valuable exactly this information is.

Our goal in this paper is to develop an analytic framework, namely models and the techniques to solve them, to help quantify the economics of location information. Our aim is to derive models which can be used as decision making tools for entities interested in or involved in the location data economics chain, such as mobile operators or providers of location aware services (mobile advertising, etc). We consider in particular the fundamental problem of quantifying the value of different granularities of location information, for example how much more valuable it is to know the GPS location of a mobile user compared to only knowing the access point, or the cell tower, that the user is associated with. We illustrate our approach by considering what is arguably the quintessential location-based service, namely proximity-based advertising.

We make three main contributions. First, we develop several novel models, based on stochastic geometry, which capture the location-based economic activity of mobile users with diverse sets of preferences or interests. Second, we derive closed-form analytic solutions for the economic value generated by those users. Third, we augment the models to consider uncertainty about the users’ location, and derive expressions for the economic value generated with different granularities of location information.

To our knowledge, this paper is the first one to present and analyze economic models which can help understand the economic value generated by mobile users with location based services, for different granularities of location information in wireless networks.

I. INTRODUCTION

The Internet has become a fundamental component of modern economies, and it provide services, starting with connectivity, that are strategic to companies, governments, families and individual users, and in general to the well functioning of modern life. A growing fraction of those services are accessed by mobile users. Indeed, the size and strategic importance of the mobile Internet, i.e. the Internet as accessed via mobile devices such as laptops or cell phones, is rapidly increasing. Recent reports indicate that the mobile Internet is ramping up in size faster than the “desktop Internet” did in the 80’s and 90’s; in fact, the estimated total value of the mobile data industry grew by 20% in 2009 - a year of major economic crisis when the global economy decreased by 5% - and mobile data revenues reached $284B [1]. This is now larger than the total PC Internet economy, including Internet content and advertising revenues plus all subscription fees such as monthly dial up and broadband access fees. Furthermore, the number of users of the mobile Internet (measured by the number of users accessing browser-based services on cell phones only) is estimated at between 500 million and 1 billion, almost on par with the total number of PCs connected to the Internet [1], [2].

A key characteristic of mobile networks and devices is their ability to capture and analyze (at least partial) information on the location of mobile users. For example, cellular operators have routinely captured large scale location data for billing purposes, but also to improve location management or satisfy legal requirements such as E911. More recently, they have started exposing (often for a fee) large scale location information to application developers. In parallel, a significant fraction of mobile devices is now GPS-enabled and captures and (sometimes for a fee) provides access to real-time GPS data. Note that the capture and availability of location data is part of a larger trend, where data of various kinds such as location data but also social network data or spectrum usage data is seen as an extremely valuable, even strategically important, asset by the carriers in particular, but also by the general mobile industry.

In any case, the capture and availability of location data enables the development of a wide range of location-based or location-aware services, and indeed an rapidly increasing number of such services is now available, ranging from navigation to location-aware advertising, friend finder, etc, and many more are announced or launched on a daily basis. As noted above, this location data, since it enables new services and new economic activities, is also seen as economically valuable. This raises the question then of how valuable it is, and how to quantify that value. This question is precisely at the core of the paper.

We start with two observations. First, we expect that developing models to study and quantify the economic value of location data to be a complex task, because the models need to capture a number of different factors such as i) spatially distributed users, ii) spatially distributed businesses and/or entities that trigger economic transactions, iii) transactions dependent on user location and, likely iv) transactions dependent on specific user preferences and interests. Furthermore, the models need to capture the granularity at which location information is available. We do not present in this paper models that capture all the properties of mobile users involved in and responding to proximity-based advertising (for example we do not consider non-Poisson or non-homogeneous user distributions). Deriving and especially solving such a general
model is still an extremely challenging problem. This paper is a first step, which presents what we believe is a useful and very promising approach to tackle and formalize the general problem. Therefore, our focus here will be on describing our approach and illustrating its power and potential. We discuss the limitations of the model in more detail in Section 3, and we leave the validation of the model (itself a challenging problem) to future research.

Second, we do not expect models to provide specific dollar-and-cents answers to questions such as “What is the value of one user’s daily GPS data worth?” Instead, we aim to develop models which can be used as decision making tools for entities interested in or involved in the location data economics chain, such as mobile operators or providers of location aware services (e.g. mobile advertising). Note that the availability of such models provides quantitative answers to “real-life” questions with enormous economic impact. For example, we consider in this paper what is arguably the quintessential location-based service, namely proximity-based advertising: a mobile user, with a given set of preferences or interests, is offered an advertisement (in the form of an SMS coupon, or a discount of some kind) when getting close to a business or a store which offers products and services that match (at least partially) with the interests of the user - in practice, a hungry user who likes Italian food would get a coupon from a pizzeria for discounted pizza when getting within some range of that pizzeria. Mobile coupon usage is expected to triple by 2014, exceeding 300 million people; almost 3 billion coupons are expected to be issued by 2011 with coupon redemption value (the amount of discounts redeemed) expected to approach $7 billions globally [3]. An economic model of this situation helps answer the following questions

- What is the overall economic value generated by such transactions?
- What is the additional value created by fine-granularity location data (such as GPS data) compared to coarse-granularity data (such as cell tower or access point-level data)?
- The location-aware advertisement situation described above typically involves three parties: 1) the mobile operator which provides or exposes the location data, 2) an entity which provides or exposes the interests or preferences of the user (for example Google, based on past search queries of the user), and 3) the source of the ad (the restaurant). Given that the user responds to the ad and walks into the restaurant to buy a pizza, how should the mobile operator, Google, and the restaurant (all of which provided component of the transaction) share the revenues created by the transaction?

We make three main contributions in the paper. First, we introduce and develop a family of novel models which capture the proximity-based advertising situation described above, and more generally the location-based economic activity of mobile users. Our second contribution is to derive closed-form analytic solutions for the models by using stochastic geometry tools, in particular for the economic value generated by the mobile users. Third, we augment the models to consider uncertainty about the users’ location, and derive expressions for the economic value generated with different granularities of location information.

This paper ties in well and complements other work on location data generation and analysis. We believe that the paper provides an important first step, especially given the increasing economic importance of location data for both cellular/wireless operators and the providers of location-based services, to explore the economics of the mobile Internet.

The rest of the paper is organized as follows. In Section 2, we describe related work. In Section 3, we develop the first model and analyze it in the case of perfect (i.e. GPS quality) location information. In Section 4, we consider specific cases of the model when we can solve it analytically, and we examine the question of quantitatively comparing the value of fine-grained (GPS) versus coarse-grained (cell-level) location data. In Section 5, we discuss the relative gain brought by the rating of services by other users. Finally, Section 6 describes an alternative model which uses a general attenuation of interest with distance rather than a vicinity radius as in the first models. We conclude and comment on future work in Section 7.

II. RELATED WORK

A vast amount of research has been published in the area of economic modeling of course (see e.g. [4] and many other references) on one hand, and on the capture and analysis of location and mobility data on the other hand (refer to the large body of research published in earlier proceedings of Mobicom and related workshops) and elsewhere e.g. [5], [6], [7]).

Comparatively less has been carried out or published at the intersection of economics and mobility in the Internet. We can divide relevant contributions in two areas, namely Internet economics (without mobility) and the economics of location, or spatial economics (not focusing on the location or mobility of Internet users).

Research on Internet economics aims at increasing our understanding of the Internet as an economic system and at developing policies and mechanisms to achieve desirable economic goals (much the same way early research on the Internet aimed at developing policies and mechanisms - such as the IP protocol - to achieve desirable design goals such as those described in [8], or more recent research aims at developing clean-slate policies and mechanisms to achieve the desired goals of the future Internet). The importance of the economic aspects of the Internet was recognized very early on. Kleinrock in 1974 mentioned that “[H]ow does one introduce an equitable charging and accounting scheme in such a mixed network system. In fact, the general questions of accounting, privacy, security and resource control and allocation are really unsolved questions which require a sophisticated set of tools” [9]. More recently, Clark et al [10] mention economic drivers as key drivers to revisit old design principles and suggest new ones. Research in Internet economics has examined several
issues, such as the economics of digital networks (refer to [11] for pointers to recent work in the area, and e.g. [12] for the analysis of a point problem, specifically the impact of layering), pricing models and incentive mechanisms for resource allocation that align the interests of possibly selfish users with the interests of the network architect [13], and the economics of security (refer to [14] for a recent survey and references, also the proceedings of the Workshop on Economics of Information Security).

A different body of research has considered spatial economics and the economics of location. Spatial economics is generally defined as being concerned with the allocation of resources over space and with the location of economic activity. That definition can be thought of as either covering much of economics (since virtually all economic activities take place in some point in space), or instead as narrowing down the issue to that of the optimal choice of location for economic activity (e.g. optimal placement of store or city) or the impact of location on economics. Overall, though, much of the theoretical work focuses on reasons for the growth or decline of cities or regions or the clustering and spatial dynamic of industries (see e.g. [15]), with work in parallel focusing on data collection, statistical analysis and validation (e.g. [16]). The related work of economic geography focuses on the processes generating cities and other economic landscapes [17]. In general, then, spatial economic theory as it is considered in economic research is not particularly tied to our problem of interest here.

III. Model with Perfect Knowledge on User Localization and Preferences

Recall that our goal is to develop models which capture the location-based economic activity of mobile users. In order to make sure that the model can help address realistic questions, we want in particular to be able to analyze situations such as the proximity-based advertising situation described earlier. Thus, the question we consider here is that of the evaluation of the economic value of proactive commercial services that use the following ingredients:

- The knowledge that a wireless operator has to locate mobile users;
- The knowledge that a search engine company may have of (i) the list of user preferences as obtained from e.g. the history of their searches; (ii) the rating of each individual service as obtained by e.g. the ranking made by other users;
- Information proactively sent through the wireless operator to users on the set of services matching their preference list and located in the vicinity of their current geographic position.

A. Assumptions

In order to develop our models, we make the following assumptions:

- The economic value of such a service is proportional to the revenue of the commercial transactions it generates.
- The likelihood for a user to respond to the incentive to stop and purchase is an increasing function of the number of commercial services matching his/her list of preferences.

Going back to the pizza ad mentioned earlier, assumption # 2 above states that the likelihood for a mobile user to step into a pizzeria increases with the number of pizzerias in the vicinity.

We next model the spatial distribution of businesses (such as the pizzeria above) which offer services that might be of interest to the mobile users (and which therefore might lead them to send mobile ads or coupons). We assume that there is a denumerable number of businesses or services which are randomly deployed in an infinite plane. We assume that services of type n (e.g. pizzeria versus movie theater vs coffee shop) are deployed according to a homogeneous Poisson point process \( \Phi_n \) of intensity \( \lambda_n \) and that all these Poisson processes are independent (refer here and throughout the rest of the paper to [18] for background and details on stochastic geometry). Let us stress that the Poisson assumption is adopted for tractability of the analysis and that it is beyond the scope of the present paper to validate it in statistical terms. The homogeneity assumption should also be challenged as it is clear that urban densities vary from city centers to suburban environments. There are two ways of addressing this last question. The simplest way consists in considering the model to represent a large but homogeneous area. The second way consists in extending the analysis to non-homogeneous Poisson point processes, which seems feasible. This last line of thought is left for future work.

Users are characterized by a random preference list which is a list of services, namely a subset of the integers. Users of class \((k, i_1, \ldots, i_k)\) have \(k\) elements in their list and these elements are the services \((i_1, \ldots, i_k) = i\) they are interested in. The following notation is used:

- The probability to have a user of this class is denoted by \(\pi(k, i)\).
- The radius of the ball defining the vicinity of a user is denoted by \(R\); this vicinity range will be used in Model 1 and Model 2 below; this will be replaced by a function representing the attenuation of interest with respect to distance in Model 3.
- The spatial density of users is denoted by \(\nu\).

In the most basic model (referred to as Model 1), we adopt the following assumptions:

- The propensity for users to stop and check out the available services depends on the total number of services \(m\) matching their list and located in their vicinity; this is quantified by a function \(f(m)\); it makes sens to assume that \(f\) is non-decreasing.
- Given that a user stops, the revenue generated is proportional to the number of different services located in his/her vicinity.

The \(f\) function is central in the model in that it captures the essence of the physical/psychological process in action: because mobile users are localized (by the wireless network
operator), and because their interests are known (thanks to Google), they can be informed on how well their current location matches their interest (through the variable \( m \) in Model 1; other and more elaborate scenarios based on rating will be considered in Model 2 and Model 3), and this triggers (through some psychological process) a decision of the user to check the services with the probability \( f(m) \) and eventually to purchase.

Now that the bases for Model 1 are in place, we can examine how to quantify the economic value generated by the users responding to the location-aware advertisements in this case.

**B. Analysis**

Pick a typical user. Given this user is of the \((k, i)\) type, what matters for his/her is the Poisson point process of intensity \( \lambda(k, i) = \sum_{j=1}^{\infty} \lambda_{ij} \), which is that of services present in his/her preference list.

It is easy to see that in case of a perfect localization, a user of type \((k, i)\) could be sent a proactive message informing him/her of the presence of \( m \) services matching his/her preferences iff his/her location belongs to the \( m \)-coverage region of the Boolean, or germ-grain model with germs (the points of the Poisson point process) of intensity \( \lambda(k, i) \) and with grains equal to balls of radius \( R \), centered on these points. By definition, a location belongs to this \( m \)-coverage region if exactly \( m \) balls cover it.

Because the number of grains that cover a given point follows a Poisson law on the integers of parameter \( \lambda(k, i) \pi R^2 \), the probability that a typical location is \( m \)-covered is

\[
p(m, k, i) = e^{-\lambda(k, i) \pi R^2} \left( \frac{\lambda(k, i) \pi R^2}{m!} \right)^m.
\]

Hence the mean revenue generated per unit of space under Model 1 is

\[
\rho = \nu \sum_k \sum_i \pi(k, i) \sum_m f(m) g(m, k, i) e^{-\lambda(k, i) \pi R^2} \left( \frac{\lambda(k, i) \pi R^2}{m!} \right)^m,
\]

where \( g(m, k, i) \) denotes the mean number of different services among the \( m \) for a user of type \((k, i)\). Using the fact that the probability that there is no service of type \( p \) among the \( m \) is \( (1 - \lambda_p / \lambda(k, i))^m \), with \( \lambda(k, i) = \sum_{q=1}^{k} \lambda_{iq} \), we get that

\[
g(m, k, i) = \sum_{p=1}^{k} \left( 1 - \left( 1 - \frac{\lambda_p}{\lambda(k, i)} \right)^m \right) = \frac{k - \sum_{p=1}^{k} \left( 1 - \frac{\lambda_p}{\lambda(k, i)} \right)^m}{m!}.
\]

To get (1), we used: (i) the assumption that the probability of stopping given an \( m \)-match depends only on \( m \) through a function that we denote by \( f(m) \); (ii) the assumption that the mean revenue given an \( m \)-match is proportional to the number of different services among the \( m \) (we take the mean revenue per service given that this service is visited by the user equal to 1). Note that in spite of the fact that the tagged user has a list of cardinality \( k \), the sum over \( m \) ranges over all integers. The reason is that there may be several services of the same type in his/her vicinity.

We describe now a special case which will be of special interest throughout the paper (it will be adapted to Model 2 and Model 3 below) and where closed forms can be obtained.

**Computational Example.** We assume that:

- There are \( N \) types of services.
- \( f(m) = 1 - \alpha^m \) with \( 0 < \alpha < 1 \), where the constant \( \alpha^{-1} \) can be seen as the propensity to react to proactive information.
- \( \pi(k, i) = \beta^k (1 - \beta) \left( \frac{N}{k} \right)^{-1} \), i.e. we have a geometric size for the list of preferences and a uniform law on the services; the parameter \( \beta \) determines the mean size of the list of preferences as provided by e.g. Google.
- \( \lambda_n = \lambda \) for all \( n \); \( \lambda \) is the spatial density of services.

These assumptions are not meant to be realistic and should be refined. They are again chosen for tractability reasons. In this case, \( g(m, k, i) = k \left( 1 - \left( 1 - \frac{k}{m} \right)^m \right) \), so that

\[
\rho = \nu \sum_k \sum_{i} \beta^k (1 - \beta) \sum_m (1 - \alpha^m) (k - k(1 - 1/k)^m) \frac{e^{-\lambda k \pi R^2} \left( \lambda k \pi R^2 \right)^m}{m!}.
\]

The sum w.r.t. \( k \) is from 0 to \( N \). Below, we replace it by a sum from 0 to \( \infty \) in order to simplify the expressions (exact formulas can be derived along the same lines but they are close to the simplified expressions provided \( \beta < 1 \) and \( N \) is large).

In this case, we can derive using straightforward calculations the mean revenue per unit of space, which is given by the following closed form:

\[
\rho = \nu \beta \left( \frac{1 - e^{-\lambda \pi R^2}}{1 - \beta} - \frac{(1-\beta) \left( e^{-\lambda (1-\alpha) \pi R^2} - e^{-\lambda \pi R^2} \right)}{(1-\beta e^{-\lambda (1-\alpha) \pi R^2})^2} \right).
\]

We plot this function in Figure 1 and discuss its properties below.

We now compute two quantities associated with this model and to be used later:

- The probability that the user stops is

\[
p_s = \sum_k \sum_i \pi(k, i) \sum_m f(m) e^{-\lambda(k,i) \pi R^2} \left( \frac{\lambda(k,i) \pi R^2}{m!} \right)^m
\]

and in the above example this can be evaluated as

\[
p_s = \sum_k \beta^k (1 - \beta) \sum_m (1 - \alpha^m) e^{-\lambda k \pi R^2} \left( \frac{\lambda k \pi R^2}{m!} \right)^m
\]

\[
= \frac{\beta (1 - e^{-\lambda \pi R^2 (1-\alpha)})}{1 - \beta e^{-\lambda \pi R^2 (1-\alpha)}}.
\]
of the list is \( k \), is equal to
\[
p_s(k) = \sum_{m} (1 - \alpha^m) e^{-\lambda k \pi R^2} \frac{(\lambda k \pi R^2)^m}{m!} \]
\[
= 1 - e^{-\lambda k \pi R^2 (1 - \alpha)}. \tag{6}
\]

- The mean potential revenue per unit of space is
\[
\phi = \nu \sum_{k} \sum_{i} \pi(k, i) \sum_{m} g(m, k, i) e^{-\lambda (k,i) \pi R^2} \frac{(\lambda (k,i) \pi R^2)^m}{m!} \tag{7}
\]
The last quantity should not be confused with the mean revenue per unit of space defined in (1); here, \( g(m, k, i) \) is not multiplied by \( f(m) \), so that the last expression is the mean revenue per unit of space when all users stop with probability 1. In the example, this evaluates to
\[
\phi = \sum_{k} \beta^k (1 - \beta) \sum_{m} (k - k(1 - 1/k)^m) \frac{e^{-\lambda k \pi R^2} (\lambda k \pi R^2)^m}{m!} \frac{1}{1 - \beta}.
\]
The inner sum, which is the mean potential revenue per unit of space conditional on \( k \), is
\[
\phi(k) = \sum_{m} (k - k(1 - 1/k)^m) e^{-\lambda k \pi R^2} \frac{(\lambda k \pi R^2)^m}{m!} \frac{1}{1 - \beta}. \tag{9}
\]

IV. THE VALUE OF PRECISE KNOWLEDGE ON USER LOCALIZATION/PREFERENCES
The aim of this section is to compare the revenue generated by the exact localization as evaluated in the last section and that of an error-prone localization.

A. ERROR-PRONE LOCALIZATION MODELS

We consider 2 basic models.

1) The first is the case where the user is mistakenly localized at a distance \( r \) from its true location. Then he/she stops with a propensity which is still that of his/her estimated location (being mistakenly advertised), but the commercial value of his/her transactions is proportional to the number of services matching his/her interests at the location where he/she stops, \( r \) miles away from the location that triggered his/her interest. There are two subcases:

1.a If \( r > 2R \), the services at the location where the user stops are independent of the services at his/her estimated position. If view of this conditional independence, we can then evaluate the mean revenue generated in case of an error prone localization as
\[
\tilde{\rho}_1 = \nu \sum_{k} \beta^k (1 - \beta) p_s(k) \phi(k), \tag{10}
\]

where \( p_s(k) \) and \( \phi(k) \) are defined in the last section.

In the example,
\[
\tilde{\rho}_1 = \nu (1 - \beta) \left( 1 - e^{-\lambda \pi R^2} \right) \sum_{k} \beta^k k \frac{1}{2} \frac{1}{1 - \beta} \left( \beta - (\beta - e^{-\lambda \pi R^2 (1 - \alpha)})^2 \right), \tag{11}
\]

1.b If the localization error is \( r < 2R \), a general expression for the mean revenue \( \tilde{\rho}(r) \) per unit of space is given in Appendix VIII, where we also show that in the example,
\[
\tilde{\rho}(r) = \nu \sum_{k} \beta^k (1 - \beta) \sum_{m,n} (1 - \alpha^m) P_s(m, n) (k - k(1 - 1/k)^n), \tag{12}
\]

with
\[
P_s(m, n) = \frac{1}{m!} \sum_{l \leq \min(m, n)} \frac{m!}{l!} \left(1 - \alpha^m\right) \frac{(\lambda k \pi R^2)^m}{m!} \times q^l (1 - q \beta)^{m-l} e^{-\lambda \pi R^2 (1 - q \beta)} \left(\frac{\lambda \pi k R^2 (1 - q \beta)}{n - l} \right)^{n-l}.
\]

2) The second case is that where stopping is at the same rate as above, but is completely independent of value and preference list, namely \( \tilde{\rho}_2 = p_s \phi \), where \( p_s \) and \( \phi \) are the probability to stop and the mean potential revenue per unit of space, respectively (these quantities are evaluated in the last section). This scenario represents in a sense the situation without localization nor preference list. In the example, this ‘independent case’ leads to the following expression:
\[
\tilde{\rho}_2 = \frac{\nu \beta^2 (1 - e^{-\lambda \pi R^2}) \left( 1 - e^{-\lambda \pi R^2 (1 - \alpha)} \right)}{(1 - \beta) (1 - \beta e^{-\lambda \pi R^2 (1 - \alpha)})}, \tag{13}
\]

B. NUMERICAL RESULTS

We now use these expressions to quantify how the exact localization process generates revenue.

In Figure 1, we plot the mean revenue in the precise localization case (Equation (3)), the erroneous localization case (Equation (11)) and in the independent case (Equation (13)) for comparison. Three different values of \( \lambda \) are considered. Not surprisingly, the combination of the list and of the location in the independent localization case creates the largest revenue per unit space, primarily by providing the best positive correlation between the users’ decision of stopping and their interests. More information creates more value: the list and exact information case (e) creates more value than case (1.a), which in turn creates more value than the no knowledge (independent) case (2).

In Figure 1 bottom-right and in Figure 2, we plot various types of relative gains. For instance, the relative gain of (e)
Note that the impact of the parameter $\alpha$ on the model parameters is identified by appropriate statistical tests. This of course requires that the basic relative gains brought by exact localization and by the list information respectively. This spatial shot-noise field will be referred to as the Interest Field for this user at location $x$. Due to stationarity, the law of this field does not depend on $x$ and this argument will hold for the chosen set of parameters. Even for a moderate mean list of the order of 3 ($\beta$ of the order of .75):

- The relative gain of precise localization alone (i.e. gain of (e) w.r.t. (1.a)) is substantial (of the order of 40 % as shown by Figure 2, left);
- The relative gain of the list information alone (i.e. gain of (1.a) w.r.t. (2)) is bigger (of the order of 100 %, see Figure 2, center);
- The gain of the combination of the list information and the exact location information (i.e. gain of (e) w.r.t. (2)) is huge (of the order of 200 % - Figure 2, Right).

Note that the impact of the parameter $\alpha$, the propensities to stop, is moderate. Of course, the last figures have no universal validity and only hold for the chosen set of parameters. However, they suggest a methodology to answer the question of the sharing of the revenues created by the transactions between the mobile operator and e.g. Google: it makes sense to propose that this revenue be shared proportionally to the relative gains brought by exact localization and by the list information respectively. This of course requires that the basic model parameters are identified by appropriate statistical tests.

To conclude this section, let us look at case 1.b. Let $P(r) := \frac{\rho(r) - \rho(\infty)}{\rho(\infty)}$ be the relative gain of a $r$-miles away localization compared to very poor localization. Here $\lambda = .1, \nu = 1, \beta = 0.7, x$ and $\alpha$ vary.

Fig. 1. Top figures and bottom left figure: the mean revenues per unit of space in the precise localization case (e) (always the top curve), in the erroneous localization case (1.a) (always the intermediate curve) and in the independent case (2) (bottom curve) for three different values of $\lambda = .01, .1$ and 1. Right bottom figure: the relative gain of (e) w.r.t. (1.a) when $\lambda = .1$ and $r > 2R$. Here $R = 1, \nu = 1, \beta = .7$ and $\alpha$ varies.

Fig. 3. The relative gain of revenue due to a $r$-miles away localization compared to very poor localization. Here $\lambda = .1, \nu = 1, \beta = 0.7, x$ and $\alpha$ vary.

V. ADDED VALUE CREATED BY SERVICE RATING

A. Model 2 and its Analysis

In Model 2, introduced below, each service of type $n$ has a random weight, which is assumed to be an independent random variables with a positive value and a distribution function $F_{n}$. This random variable is meant to represent the rating of this specific instance of service $n$. The main motivation for this section is a quantitative evaluation of the added value created by the knowledge of such a rating compared to the situation without rating.

The general assumptions in Model 2 are that: (i) the propensity for a user to stop is a function $f$ of the sum of the weights of the services in his/her vicinity; (ii) The revenue is proportional to the sum of the weights. Under these assumptions, the mean revenue per unit of space is

$$\rho = \nu \sum_{k} \sum_{i} \pi(k,i) E[f(I(k,i))]I(k,i)],$$

where $I(k,i)[x]$ is the following sum:

$$I(k,i)[x] = \sum_{i_j \in \{i_1, \ldots, i_k\}} \sum_{X_k \in \Phi_{i_j}} W_{ij} I_{X_k-x} \leq R,$$

where $W_{ij}(k)$ is the weight (ranking) of service $k$ of type $i_j$. This spatial shot-noise field will be referred to as the Interest Field for this user at location $x$. Due to stationarity, the law of this field does not depend on $x$ and this argument will...
be dropped when dealing with the law of this field in what follows.

It is well known that the Laplace transform of $I(k, i)$ is

$$E\left(e^{-sI(k, i)}\right) = \Psi_{I(k, i)}(s) = \prod_{i_j \in \{i_1, \ldots, i_k\}} \exp\left(-\pi \lambda_{i_j} R^2 (1 - \Psi_{F_{i_j}}(s))\right)$$

where $\Psi_{W_n}(s)$ denotes the Laplace transform of $W_n$ (with law $F_n$). We see that the law of the interest field is fully determined by that of the ranking variables and the other model parameters.

B. Example

The assumptions are the same as in the example for Model 1. In addition, we assume that $F_n = F$ for all $n$ (in what follows, $W$ denotes a random variable with law $F$). Then $I(k, i)$ has the same law for all $i$ and $I(k)$ denotes a random variable with this law. We have

$$\rho = \nu \sum_k \beta^k (1 - \beta) E\left[(1 - \alpha I(k)) I(k)\right]$$

$$= \nu \sum_k \beta^k (1 - \beta) \left(-\Psi_{I(k)}(0) + \Psi_{I(k)}(-\log(\alpha))\right),$$

Since in this case $\Psi_{I(k)}(s) = e^{-\pi \lambda k R^2 (1 - \Psi_{W}(s))}$, this gives

$$\rho = \nu \lambda \pi R^2 \sum_k \beta^k (1 - \beta) k \left(E(W) + \Psi_{W}(-\log(\alpha)) e^{-\pi \lambda k R^2 (1 - \Psi_{W}(-\log(\alpha)))}\right)$$

that is

$$\rho = \nu \lambda \pi R^2 \left(E(W) \frac{\beta}{1 - \beta} + \frac{\Psi_{W}(-\log(\alpha)) \beta(1 - \beta) e^{-\pi \lambda R^2 (1 - \Psi_{W}(-\log(\alpha)))}}{1 - \beta e^{-\pi \lambda R^2 (1 - \Psi_{W}(-\log(\alpha)))}}\right).$$

In order to evaluate the value of rating, we compare below the case with $F$ exponential of mean 1, i.e. $\Psi_{W}(s) = 1/(1 + s)$, and the case with $F$ deterministic of mean 1, i.e. $\Psi_{W}(s) = e^{-s}$, which represents the situation without rating. As we see on the rightmost plot of Figure 2, this gain is now very sensitive to the $\alpha$ parameter; the larger the list, the larger the relative gain. For moderate list sizes and optimal values for $\alpha$ this gain ranges from 5 to 20%.

C. The Value of Precise Localization

We follow the same path as in the previous section. If there is an error in the localization of more than $2R$, then, conditionally on the type of the customer, his/her chance to stop and the economic value generated by his/her stop are independent. So the mean revenue per unit of space is

$$\tilde{\rho} = \nu \sum_k \sum_i \pi(k, i) E[f(I(k, i)) I(k, i)].$$

Hence in the example,

$$\tilde{\rho} = \nu \sum_k \beta^k (1 - \beta) E\left[(1 - \alpha I(k)) I(k)\right]$$

$$= \nu \sum_k \beta^k (1 - \beta) \left(-\Psi_{I(k)}(0) + (1 - \Psi_{I(k)}(-\log(\alpha)))\right),$$

that is

$$\tilde{\rho} = \nu \lambda \pi R^2 E(W) \left(\frac{\beta}{1 - \beta} - \frac{\beta(1 - \beta) e^{-\pi \lambda R^2 (1 - \Psi_{W}(-\log(\alpha)))}}{1 - \beta e^{-\pi \lambda R^2 (1 - \Psi_{W}(-\log(\alpha)))}}\right).$$

Consider now the case where the true location $y_t$ is $r$-miles away from the estimated location $y_e$, i.e. $|y_t - y_e| = r$, with $r < 2R$. Due to stationarity, the mean revenue generated depends only on $|y_t - y_e|$ and not on $(y_e, y_t)$, so that

$$\tilde{\rho}(|y_t - y_e|) = \nu \sum_k \sum_i \pi(k, i) E\left[f(I(k, i)|y_e) I(k, i)|y_t]\right]$$

is

$$= \nu \sum_k \beta^k (1 - \beta) E\left[(1 - \alpha I(k)|y_e) I(k)|y_t]\right].$$

So we need to evaluate the joint law of the interest field at two locations of the space. We obtain (we do not show the
where \( I \) and \( \beta \) are the Laplace transform of the function for some \( A > 0 \).

The mean revenue per unit of space is

\[
\rho = \nu \sum_k \beta^k (1 - \beta) \int \frac{P_i}{l(i|x)} dx.
\]

In this case, the Laplace transform of \( I \) is

\[
E(e^{-\lambda |I|}) = \Psi_I(s) = \prod_{i,j} \exp\left(-2\pi \lambda_i \int_0^\infty r \left(1 - \Psi_{P_j} \left(\frac{s}{l(r)}\right)\right) dr\right).
\]

**B. Example**

We take the same assumptions as in Model 2. Then the mean revenue generated is

\[
\rho = \nu \sum_k \beta^k (1 - \beta) \left(-\Psi_I(0) + \Psi_I(-\log(\alpha))\right).
\]

The following examples of proximity functions will be considered: (PF1): \( l(r) = (Ar)^\beta \); (PF2): \( l(r) = (A \max(r_0, r))^\beta \), for some \( A > 0 \), \( r_0 > 0 \) and \( \beta > 2 \). Note that both cases give similar values for \( r > r_0 \). We assume an exponential distribution for \( F_n = F \); i.e. \( \Psi_W(s) = \frac{\mu}{\mu+s} \), with \( \mu \geq 0 \).

Assuming (PF1) for \( l \), one obtains

\[
\Psi_I(s) = \exp\left(-\lambda k \frac{s}{\mu}\right) \frac{2\beta}{A^2\beta \sin(2\pi/\beta)}
\]

Note that \( E[I_l] \) is infinite in this case for \( \beta > 2 \) due to the pole at the origin. In consequence, this function cannot be used to model the interest field. We now assume (PF2) for \( l \) with \( \beta = 4 \). We have

\[
\Psi_I(s) = \exp\left(\frac{\lambda k \pi}{A^2} \sqrt{\frac{r}{\mu}} \arctan\left(\frac{(Ar_0)^2}{\mu}\right) - \frac{\lambda k \pi^2}{2A^2} \sqrt{\frac{s}{\mu}} + \frac{\lambda k \pi r_0^2 (1 - \mu)}{s + (Ar_0)^4 / \mu}\right).
\]

In this case we have \( E[I_l] = -\Psi_I(0) = \frac{2\pi \lambda}{\mu A^2 r_0^2} \).

**C. The Value of Precise Localization**

Suppose that the user is mistakenly localized at the point \( y_e \) but that his/her true location is the point \( y_t \). We can then evaluate the mean revenue generated in case of such an error prone localization as

\[
\rho(y_t, y_e) = \nu \sum_k \beta^k (1 - \beta) E[f(I_l(y_e)) I_l(y_t)].
\]

By stationarity, this value depends on \( |y_t - y_e| = r \) and not on \( (y_e, y_t) \). In the case of \( f(x) = 1 - x^2 \), we find (we do not show the details of the derivation for lack of space) that the mean revenue per unit of space in Model 3 is

\[
\rho(r) = \nu \sum_k \beta^k (1 - \beta) \left(-\Phi_{I_l}(0) + \Phi_{I_l}(-\log(\alpha))\right).
\]

**VII. C. The Value of Precise Localization**

We have presented and analyzed three simple parametric models allowing one to capture how to jointly leverage the combination of three types of basic information: 1) that of the geographic location and mobility of users of a cellular phone network 2) that of their needs and interests as obtained from their preference lists and 3) the services available at all locations of space and their rating by the users. We have shown how these models could be used to give estimates of the potential revenue of this combination of informations and a potential ground for sharing this revenue.

As mentioned earlier, we did not present in this paper models that capture all the properties of "real-life" mobile users involved in and responding to proximity-based advertising. Deriving and especially solving such a general model is clearly a challenging area for future research (we indicate a few next steps below). This paper is a first step, which presents what we believe is a useful and very promising approach and our focus in the paper has been on on describing our approach and illustrating its power and potential.

The models we presented in the paper can be enhanced in several ways. We foresee three main types of extensions:

- Those associated with the assumptions which were made to get a computational framework: (i) the locations of services should allow one to represent more than homogeneous Poisson point processes. Clustering or inhomogeneity can fortunately be considered within the
present setting (e.g. with non-homogeneous Boolean or Shot Noise processes). (ii) In either Model 1 or Model 2, it would be interesting to consider more general propensities to stop; for instance \( f(m) = 1 - \frac{1}{m} \) with \( \alpha > 0 \) would describe a weaker propensity to stop compared to the geometric model described above.

- Those describing customer classes and states: some state should be defined for each user: roaming, commuting, working, ..., and some parameters like e.g. the radius of the vicinity ball or the propensity parameter should then dependent on the user class/state.
- Those associated with the weaknesses of these first models, which have many subjective assumptions. This mainly concerns the mechanisms through which the user stops and purchases. Other schemes should be considered. For instance, in Model 1, a refined model could address the case where the propensity to stop is a function of the number of different services in the user’s vicinity. In Model 2, the revenue could be proportional to the number of different services, or to the sum over the service types \( n \) of the max of the weights of the services of class \( n \).

There is also a rich set of open questions on the simple models introduced in the present paper. Here are a few examples:

- For a given user class, can one evaluate the stationary distribution function, or the variance of the generated revenue?
- In either Model 1 or Model 2, users could also rank each service type, and the propensity to stop as well as the generated revenue could depend on this ranking. In Model 1, this could mean that given the ranking of a user, one associates deterministic weights to each service type. The main question is of course that of the the added value of the knowledge of service ranking compared to the situation without ranking.
- It would be desirable to incorporate user motion in the model (e.g. Brownian motion, random walk or random waypoint) so as to be in a position to analyze the time/space structure of the commercial transactions of a given user class.

Finally, there is a clear need to calibrate the models by identifying via appropriate measurements, or solutions of inverse problems, what are the typical values for the key model parameter (propensity to stop, structure of the preference list, density of services, etc.); this challenging task will require a multidisciplinary work involving operators, search engine specialists and economists.

ACKNOWLEDGMENTS

The authors thank H. Amini for his contributions to the results of the case 1.b of \( \S \) IV-A.

REFERENCES


VIII. APPENDIX

Proof of Formula (12)

Let \( \hat{P}(r) \) be the mean revenue generated in case of a localization \( r \) miles away from the true one. For \( m, n \in N \), we denote by \( P_r(m, n) \) the probability that the estimated location is \( m \)-covered and the true location is \( n \)-covered. We have

\[
\hat{P}(r) = \nu \sum_{k=1}^{\infty} \pi(k, i) \sum_{m, n} f(m) P_r(m, n) g(n, k, i)
= \nu \sum_{k} \beta^k (1 - \beta) \sum_{m, n} (1 - \alpha^m) P_r(m, n) (k - (k - 1 - 1/k)^n).
\]

Let \( A_r \) denote the overlapping area of two discs with radius \( R \) and distance \( r \) between their respective centers and by \( q_r \) the proportion of this area to the area of the disc, namely we have \( q_r = \frac{A_r}{\pi R^2} \). We have

\[
q_r = \frac{1}{\pi} \left( 2 \arccos \left( \frac{r}{2R} \right) - \frac{r}{2R} \sqrt{1 - \left( \frac{r}{2R} \right)^2} \right),
\]

so that

\[
P_r(m, n) = e^{-\lambda \pi \rho R^2} \left( \frac{\lambda \pi \rho R^2}{m!} \right)^{m} \sum_{l \leq m, n} \left( \frac{m}{l} \right)^{m} q_r^{l} (1 - q_r)^{m-l} \frac{\exp \left( -\lambda \pi \rho R^2 (1 - q_r) \right)}{(n-l)!}.
\]