Chapter 1

Introduction

In the last hundred years, wireless transmission has revolutionized the way we communicate. In the first half of the last century, broadcast radio transmission completely changed the way information was distributed on a global scale. In the second half of the century, wireless networks greatly lowered the barrier to information sharing between computer systems. Perhaps because of the level of integration they have obtained in our day-to-day lives, today's wireless networks are simultaneously amazing and disappointing. They allow us to do something that seems almost unbelievable: effortlessly moving bits from one computer to another through the air. Yet, they never seem to work quite as well as we would like in the places we would most like them to.

This thesis focuses on a specific problem that is at the center of many other problems with wireless networks: *there are not good methods for determining how well a given network works over a given area and presenting this information in a meaningful way*. The seemingly fundamental task of drawing a meaningful and accurate picture of the "usable" coverage of an existing wireless network is an open question that this thesis will address.

Solving the general coverage problem involves advancing the state-of-the-art in four integral sub-problems:

(1) **Prediction**: How can the signal quality at a given point (or many points) be predicted if only information about the environment and the transmitter is known?

(2) **Measurement**: Assuming measurements are to be made to correct or evaluate a model, how and

where should they be made? How many are necessary? What are the tradeoffs in terms of cost and accuracy for these measurements?

(3) **Interpolation**: When measurements are in hand, how can they be used to make inferences about the coverage at locations where measurements have not been made? Which of the numerous interpolation strategies is most appropriate for mapping wireless network coverage?

(4) **Presentation**: How can the resulting maps and spatial processes be presented in a clear and meaningful way, while not over- or under-estimating the network's abilities? How does the ideal visualization differ for end-users, designers, and operators of these networks?

To provide answers to all of these questions is a substantial undertaking. However, this thesis makes strides in each subproblem. First, to understand the practical accuracy of existing methods, a great number of prior models have been carefully analyzed, implemented, and performance-tested. Measurement methodologies are advanced by applying statistically robust, domain-appropriate spatial sampling methodologies. Methods for optimizing multi-phase sampling schemes are developed so that those measurements that are made have been placed to enable the largest gains with the least amount of work. This thesis proposes the use of geostatistical interpolation and modeling for wireless coverage mapping, in large part because this method embraces the intrinsic variability of the radio environment and allows for residual error and variance to be modeled explicitly. These techniques are adapted as necessary and applied to the problem through two novel case studies, and a set of best practices are extrapolated from the lessons learned. Finally, a simple visualization scheme is developed that presents wireless coverage using a color-mapping scheme adapted from medical imaging, and can be interactively used with popular mapping software. By providing a complete method for measurement-based coverage mapping, this thesis seeks to provide a novel, systematic, well-defined, and thoroughly evaluated approach to the very important problem of measurement-based coverage mapping.

1.1 Motivation and Applications

A general solution to the coverage mapping problem could have substantial impact on both current and next generation wireless technologies. This section briefly discusses five important applications of measurement-based coverage mapping.

1.1.1 Cognitive Networks

Cognitive networks are considered by many to be the next step in intelligent spectrum use [145]. Although much of the available (useful) wireless spectrum has been auctioned to particular primary users, there has been a great deal of work showing that not all of this spectrum is being fully utilized. Indeed, there are large areas distributed in space and time that can be gleaned by a secondary user for communication (e.g., [229]). However, for the cognitive radio model to work, accurate coverage maps, or "radio environment maps", are necessary to provide insight into locating and avoiding primary users. Some proposals suggest that predictive models be used to estimate transmission boundaries (e.g., [89]). However, as will be shown in chapter 3, the error associated with these models can be substantial and unpredictable. More recently, some researchers are investigating the possibility of measurement-based mapping in this domain using fixed or mobile sensors (e.g., [195, 56, 81]), but there are still many open questions. This thesis makes some first steps to answer some of them. Bounding the error of predictive models will help motivate a measurement-based solution. Then, when studying the effect of spatial sampling strategies, optimized sampling provides insight into where sensors must be deployed to create accurate maps of existing usage, to determine how many sensors are necessary, and to establish whether user-collected (i.e., "crowd-sourced") data is useful for generating maps.

1.1.2 Self-optimizing Networks

Networks that make decisions about their channel usage and other configurations benefit from accurate information about the channel in terms of both current coverage and interference. For instance, in [116], Kanade *et al.* propose a network optimization strategy where routing decisions are made based upon inferences of link quality taken from a measured signal map, which they call the "wireless manifold". Understanding how and where (and even when) measurements must be made to model the Radio Frequency (RF) environment with sufficient precision is an an important open question that underlies these proposals. The robust geostatistical approaches developed in this thesis begin to answer these questions. In an environment with highly mobile transmitters, where the surrounding RF environment cannot be assumed to be optimal, optimized multi-phase sampling of the sort desribed in chapter 8 might also be used to iteratively learn and correct a coverage map over time.

1.1.3 General Network Evaluation and Planning

Network operators require an understanding of the extent of coverage of wireless networks in such a way that it can be used to repair problems (holes), expand the network, and communicate coverage to users and marketers. By identifying areas of potential inter-node interference or coverage gaps, a network operator can choose to tune the antenna orientation and tilt of a given Base Station (BS) antenna, or add transmitters where they are needed. For instance, by identifying the coverage holes in microcell outdoor networks, a cell network operator might choose where nanocells could be installed to address local regions of poor connectivity.

Coverage maps can also be used in the planning and build-out phases of wireless network deployment. For instance, in [95], Hills discusses the wireless network at Carnegie Mellon University (CMU) and argues for an iterative deployment process where coverage testing feeds back into deployment decisions. In [115], the authors propose an Access Point (AP)-placement algorithm which uses ray-optical measurements as input. In practice, network operators often obtain information about coverage of their network by collecting data with mobile vehicles, colloquially called "drive-testing" or "war-driving". For instance, in [43], Byers and Kormann provide a good overview of AP mapping and in [97], geography researchers provide their mapping technique for the unplanned networks of Salt Lake City, Utah. Despite the prevalence of this technique, seldom are important concerns such as sampling bias, completeness, choice of performance metric, or statistical significance considered. This thesis investigates methods for principled spatial sampling in wireless coverage mapping, and appropriate interpolation techniques, which can provide insights into how “drive-testing” and investigative coverage mapping (for planning or diagnosing problems) might be made more robust with the use of appropriate statistical techniques.

1.1.4 Contractual Verification

Being able to make a strong statement about the extent of coverage for a wireless network is crucial for verifying contractual requirements for network deployments. In a typical scenario a contract will be drafted between the client wishing to build a wireless network, and the company they have hired to build it. In this contract, the goal coverage criterion can be specified along with an appropriate testing methodology to determine when the contractual obligations have been fulfilled. Chapter 4 will look at a municipal wireless network in Portland, Oregon that was substantially harmed by the operators’ inability to identify coverage gaps [121, 164, 232]. In the business of wireless telephones, communicating wireless coverage to end users has become an important business practice, which educated consumers use to choose their provider [227, 140, 31].

There are no shortage of companies, both big and small, that offer contractual coverage testing as a service (e.g., [239, 4]). However, there are no universal standards or best practices for how coverage testing should be performed, and in some cases the techniques used by consulting firms may lack statistical or procedural rigor. Due to variation in the methods, results from tests might not be comparable or reproducible. Section 4 will look at the coverage testing problem in the context of municipal wireless contractual verification. By proposing a straightforward and robust method for coverage testing, hopefully the uncertainty in this domain can be mitigated. In some networks, contractual verification may involve more rigid specifications including varying coverage requirements in different regions. The coverage mapping methods described in chapter 5 can help to address these issues.

1.1.5 Detecting Spurious Emissions

A related problem involves detecting spurious emissions. In some scenarios, regulatory enforcement agencies may be required to determine whether rogue transmitters are creating harmful interference or operating outside of their band. For instance, in a recent decision, the Federal Communications Commission

(FCC) determined that emissions from a neighboring commercial network might impact Global Positioning System (GPS) devices [216]. In this scenario, the same coverage mapping problem presents itself. The *de facto* approach used today in this scenario involves point-testing and *predicting* out-of-band emissions using a model. A generalized coverage mapping method, adapted to making inferences about the RF environment could be useful in this scenario to determine not only whether spurious emssions are present, but also the source and exact extent of interference.

1.2 The Case for Hybridization

There are two approaches to coverage mapping that dominate the state of the art today: direct measurement and *a priori* modeling. It stands to reason that these themes would prevail, as they comport to Occam's razor, but *how well do they work*?

Direct measurement is straightforward: visit a large number of points in the area of interest, measure the signal strength at those points (and ideally perform higher-layer tests as well), and then use these measurements to draw a coverage map. The problem with this approach is that it scales poorly and becomes stale quickly. Exhaustive measurement is very laborious for small networks, and for networks that are city-sized, it is likely cost prohibitive. When measuring a large area, one must choose a subset of points to measure, and appropriate sampling strategies are not well known, nor is it well understood what sources of bias can stem from inappropriate sampling. For instance, many cell carriers typically rely on "drive-test" measurements, where measurements are made exclusively with a mobile (vehicular) tester along streets, without much concern for how the atypical propagation environment created by streets, as well as sampling bias, may effect the validity of measurements. Once the data is collected, it is not clear how well this data will age. There can be substantial small-scale and large-scale temporal variation in the radio channel. At what point does it become too stale to use? How often should it be updated?

Another problem with this approach is that the best way to interpolate between measurement points is not well known, nor are there standards for how to present this data in a useful way. Linearly (or exponentially) interpolating between neighboring measurements causes an uncertain amount of error as a result of this smoothing. How does one make a map that is actually meaningful for the person using it? Mapping schemes that only plot the expected received signal strength (or signal to noise ratio) neglect a number of channel attributes that may significantly contribute to the usability of a network at that point, such as neighboring interference or link asymmetry. In short, despite its attractive simplicity, direct measurement alone is not enough—it does not properly treat the problems of sampling, interpolation, and presentation.

The other popular way to make a map is using a predictive model. This is especially useful for planning networks that are not yet built, and hence, cannot be directly measured. Maxwell's equations describe the propagation of plane waves. It is well known that signal degrades approximately proportionaly to the square of the distance. Is it possible, then, to use an analytical model to predict the coverage in a given environment *a priori*? There is no shortage of existing models that try to predict signal attenuation as a function of distance between points, using any number of other variables and parameters. These models come in every shape and size, but it is not known how well they work in general. Certainly, picking a propagation model from the literature and using it to create a coverage map is one approach. However, saying how accurate such a map is without making direct measurements is impossible. And, choosing incorrectly can have penalties whose severity cannot be determined *a priori*. For instance, in [44], Camp *et al.* show that small changes in model parameters used for planning a wireless mesh network can result in massively under- or over-provisioned networks. Chapter 3 will attempt to define bounds for the error associated with these *a priori* models in practical applications, and show that on their own, they cannot sufficiently model the channel. There is no silver bullet here, either.

Because neither approach works well enough on its own, the approach investigated here is a hybridization of these two: measurement-corrected modeling. The starting assumption is that creating an accurate coverage map for a given network requires some direct measurement. However, because measurement is costly, it is desireable to make *as few measurements as possible to generate a map within application-appropriate accuracy bounds*. To address this problem, this thesis will look to the geostatistics literature, which has much to say about rigorously sampling random fields. Chapters 5, 6, and 7 provide an adapatation of these methods to the coverage mapping problem. Then, chapter 8 will suggest optimized sampling strategies to generate a single system that is able to create a more accurate coverage map than is possible with either approach alone, while requiring less work than an exhaustive measurement campaign.

1.3 The Case for Geostatistics

In 1951, Daine Krige revolutionized the field of mine valuation by developing new statistical methods for spatial processes. Mine valuation is the task of determining the grade of ore or the amount of precious metal over some region. Predicting the shape and distribution of this field is essential to planning mines, and hence there are substantial penalties for inaccuracy. Prior to his proposals, the task of mapping the grade of ore in mines was in a similar state to how things are today with mapping the coverage of wireless networks. It is well known that creating maps is necessary, but there is little consensus about the best way to do it, and sources of measurement bias and modeling error are not well understood. In [123], Krige made the case for statistics:

> The need for greater uniformity in valuation procedures, and for the limitation as far as possible of the personal element cannot be disputed. The solution to this problem lies, in my opinion, in the extensive application of statistics. I do not wish to imply, however, that statistics is a miracle tool with rigid procedures that can be applied indiscriminately on any mine without a proper appreciation of local conditions. On the contrary, a clear concept of the problems involved is essential, and this can emanate only from practical experience. Once the necessary spade work has been done, however, the routine application of statistics on any mine will involve only simply arithmetical calculations well within the scope of the average surveyor and sampler.

This thesis will take a “Krigian” approach. A fundamental assumption here is that the solution to the coverage mapping problem is, like mine valuation, a task involving costly sampling of points in a random field and using these samples to infer the shape of the field overall. However, as Krige eloquently points out, reckless application of statistical methods here leaves us no better off than we started—it is necessary to marry practical knowledge about how wireless networks work with proper statistical methods. In the end, as Krige did with mine valuation, this thesis will provide a complete method that brings the currently costly and complex task of coverage mapping into the domain of the “average surveyor”.

1.4 Thesis Statement

The following thesis is asserted:

> Domain-appropriate geostatistical methods can provide a solution to wireless coverage mapping that (a) is more accurate than is possible with *a priori* modeling approaches and (b) requires fewer measurements than explicit, undirected measurement-based approaches.

Chapter 3 will show that the minimum practical accuracy of existing approaches to *a priori* modeling is on the order of 9 to 12 dB. Chapters 5 and 6 present a geostatistical method for wireless coverage mapping which more than halves this error in practice. Chapter 7 shows that this method can produce results that compete with the state-of-the-art methods for finding coverage holes using drive-test measurements, both in terms of minimizing the number of measurements and predictive accuracy. Finally, chapter, 8 will extend the methodology with a iterative optimized sampling system that can further tune coverage maps with intelligently placed second-phase samples.

Because the landscape of all wireless networks is large and varied, this evaluation focuses on a type of network of particular interest: large-scale, microcell, outdoor, urban wireless networks operating in the UHF band (300 MHz to 3 GHz). This sort of network is widely used for providing wireless network access to cities and towns (e.g., [146]), university campuses, mobile cell networks, commercial and industrial campuses, and military sites (e.g., [95]). Because of their relatively low cost and accessibility, these networks have also gained some traction in sparse rural applications (e.g., [9, 180]). Besides existing networks, the results here are applicable to future networks that operate using a similar microcell design in urban areas and at similar frequencies. And, as metrics and models are improved with time, they can be "plugged in" to the methods proposed here. Although these sorts of networks are the focus of this thesis, the methods developed are likely applicable to a large class of similar networks and frequencies.

1.5 Dissertation Outline

The next chapter provides background on the problem, including a discussion of practical wireless path loss prediction, and describes the state of the art in terms of coverage mapping. Chapter 3 puts the state of the art to the test by performing a rigorous evaluation of 30 analytical and empirical propagation models from the literature and puts practical bounds on the prediction error of these models with respect to ground-truth data. These results show that "simple" approaches to path loss prediction are troublesome and choosing amongst them is precarious. This observation motivates the principled measurement-based coverage mapping method developed in the subsequent chapters. Chapter 4 begins with the smaller problem of coverage *testing*, where the goal is to determine the *percentage* of area covered at some level of performance. These methods are applied to the problem of testing the coverage of a large municipal wireless network in Portland, Oregon. Chapter 5 introduces the core proposal of the thesis, which provides a geostatistical method for sampling, measurement, and coverage *mapping*. The efficacy of this method is evaluated through two case studies provided in chapter 6, which seek to map the coverage of WiMax and Long Term Evolution (LTE) networks in Boulder, Colorado. As a practical extension to this proposal, chapter 7 analyzes re-sampling approaches which allow for commonly available drive-test data to be used with the geostatistical methods proposed here. Chapter 7 also introduces the topic of crowd-sourced coverage mapping where end-users volunteer to collect measurements of network coverage. A feasibility study is performed using a mobility model, as well as a case study using data collected with smart phones in greater Colorado by the company Open Signal Maps. Chapter 8 presents a method to fine-tune maps generated with this system using iterative refinement with optimized multi-phase sampling. Finally, chapter 9 will conclude the thesis with a recap of results and a discussion of possible future directions.

The main text is accompanied by several appendices which provide results from related experiments. Appendix A describes the Effective Directivity Antenna Model (EDAM), a path loss model for directional antenna systems. Appendix B describes an experiment seeking to understand the numerical stability of path loss models that make a large number of numerical approximations using the Uniform Theory of Diffraction (UTD). And, appendix C provides source-code listings for the most important algorithms produced by this thesis.