Misuse Detection in Dynamic Spectrum Access Networks

Abhay Rao Bhadriraju

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Engineering

Yaling Yang, Chair
Y. Thomas Hou
Jeffrey H. Reed

May 06, 2014
Blacksburg, Virginia

Keywords: Wireless Networks, Dynamic Spectrum Access, Linear Optimization, Game Theory, Security, Privacy

Misuse Detection in Dynamic Spectrum Access Networks

Abhay Rao Bhadriraju

(ABSTRACT)

With dynamic spectrum access emerging as an important paradigm for efficient spectrum use, mechanisms are required to ensure disciplined spectrum access by secondary users. This must be done without requiring secondary users to disclose private data, such as their exact usage pattern or identities of parties involved. We formulate, design and evaluate a mechanism to collect spectrum activity information using a set of CPEs. A system design is presented which uses a number of techniques to address mobility and security issues involved in relying on CPEs to collect spectrum activity information. The system imposes an observation probability such that a rational cheater is dissuaded from spectrum misuse. The minimum number of CPEs required to impose this observation probability is determined by formulating it as an integer linear program. The security and privacy of this system is analyzed, along with simulation results to evaluate the quality of the solution. Based on the current design, directions for future work are identified and preliminary approaches are presented.

Acknowledgments

As I prepare this document, I would like to thank Dr. Yaling Yang for her invaluable insights and guidance during the development of this work. I would like to thank Dr. Thomas Hou and Dr. Jeff Reed for being such open guides for this work. I would like to thank Dr. Yang for her time and support in my research efforts, and Dr. Hou for his interesting and eye-opening classwork. I would also like to thank Dr. Jung-Min Park for his guidance during my teaching efforts, and for his classes on security.

These acknowledgements would not be complete without thanking my fellow labmates, classmates and undergraduate students at Virginia Tech. Discussions with them have helped me expand my horizons.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Telecommunications and related areas have become the foundation for social stability and growth. The path has been fraught with challenges, and many have been overcome - yet new ones are brought on by the very nature of the technology infrastructure built by humankind.

One such challenge is spectrum scarcity. Increasing the spectrum available to a device, coupled with the ability to choose between bands, would always bring significant throughput improvements and would complement any other techniques developed to increase throughput. Yet currently, a device is constrained to operate within specified bands in order to prevent harmful interference to critical or primary wireless systems. This constraint is applied irrespective of whether a band is currently in use, thus leading to low spectrum efficiency in bands which are used only intermittently. Cognitive Radio (CR) and Dynamic Spectrum Access (DSA) technologies have been developed solve these very issues - to allow frequent users to utilize fallow bands as and when they require. They aim to improve the utilization of the scarce resource that is electromagnetic spectrum.

Having said this, DSA systems must implement techniques which prevent harmful interference to primary users, and fair use amongst secondary users - especially to prevent harmful interference to primary users when they are using the bands licensed to them. In-spite of measures to prevent such interference, attacks which cause malicious behavior in Cognitive Radio devices are still a threat. New business models, such as usage based charging, cannot be realised if no system for detecting spectrum misuse is put in place. A system to detect spectrum misuse would prove useful in enforcing spectrum usage policies by monitoring secondary users for adherence.

In this work, a system called SysMet for detection of spectrum misuse and verification of usage reports is presented. In this work, we target the issues related to misuse detection beyond the PHY and MAC layers, and develop SysMet as a solution for the same. This system functions both as an ex ante (preventive) technique as well as an ex post (punitive) technique [33] for ensuring disciplined secondary device behavior. The game theoretic formulation dis-

suades spectrum misuse, while the network of monitors not only provides a preventive misuse reporting system, but also aides in localizing the offending device. Through an empirical study of the system, we identify avenues using which some of the assumptions regarding the architecture model may be relaxed.

1.1 Security in DSA Networks

As devices become more powerful and demand more bandwidth, spectrum is becoming a precious resource. Cognitive Radio (CR) technologies and Dynamic Spectrum Access(DSA) will be an efficient way to manage spectrum access in licensed bands. As an example, the IEEE 802.22 standard on Wireless Regional Area Networks[47][1] discusses an architecture for DSA in the Television bands using 2 methods - spectrum sensing techniques and using geo-location/databases. The reference architecture describes Customer Premises Equipment (CPE) such as cellular handsets, tablets or laptops, along with Base-Station (BS) devices which serve CPEs, aiding or directing spectrum access at each CPE. DSANs are assuming a common architecture, similar in many ways to existing telecommunications networks

Apart from designing for failures, as presented in [19], DSANs can achieve widespread adoption only once a number of security aspects are considered. These are mainly related to enforcing spectrum access rules among secondaries, and have been categorized as follows [21] [33] :

- Confidentiality : Network communications can only be comprehended by those authorized to do so.
- Availability : The network can be accessed when needed by anyone authorized to do so.
- Integrity : Changes occurring to data (intentional or unintentional) are detectable.
- Access Control : Certain resources on the network are only available to privileged users.
- Non-Repudiation : Techniques to allow users to be held accountable for actions on the network.
- Compliance : Users must comply to rules of use set up for a resource.
- Privacy : Sensitive information about primary or secondary users must be protected.

These aspects include verifying the integrity of Primary User identity and transmissions, GPS location integrity, integrity of spectrum sensing, etc. In [33], a survey of current security

issues is presented. It includes Primary User Emulation (PUE) Attacks, Spectrum Sensing Data Falsification (SSDF) Attacks, along with questions on primary and secondary user privacy. Approaches to address these issues have also been presented.

1.2 Disciplined SU Behavior in DSANs

Dynamic Spectrum Access allows spectrum alloted to a primary user to be used opportunistically by secondary users. This is under the guarantee that the primary user maintains control of the spectrum - the primary user may choose her use of the spectrum freely, and can enable new business models where secondary users are charged based on their usage.

For the successful deployment of DSANs, the primary user needs a way to ensure disciplined spectrum use by secondaries. By "disciplined", we mean that secondary users must access spectrum while adhering to spectrum access policies mandated by the primary user, or by other empowered authorities. This is both to ensure prevent harmful interference to primary or secondary users, and to enable applications such as usage-based charging. This must be done with low overhead, and while maintaining the expected level of privacy for the BSs and the CPEs.

In the state-of-the-art, this has been enabled through hardware and software techniques native to a CR device, such as Policy Conformance Components (PCCs) [41] and policy reasoning mechanisms [6], where PCCs are used to prevent violations of spectrum access policies, as explained further in Chapter 2.

With a reference architecture similar to the one used for IEEE 802.22, we propose SysMet - a system which utilizes a subset of CPEs, called the Monitor set, to verify spectrum access conducted by BSs. We exploit the mobility of CPEs in order to efficiently monitor BS spectrum activity in a large area, and impose a minimum probability of observing a BS. We find the minimum quantity of monitors required to impose a given minimum probability of observation, and the density of monitors required at different points in a given area.

The mobility of CPEs is modeled using the Truncated Levy Walk(TLW) model, but may be substituted with empirical or scenario-specific mobility models. Assuming a square tessellation on the area of operation, Monte-Carlo simulations are used to obtain the Occupation Probability for each cell. An Integer Linear Program is formulated to obtain the monitor densities required to impose a given probability of observation in each grid unit.

This thesis is organized as follows - Chapter 2 briefly reviews related work in the state-of-the-art, as well as work that presents techniques used in SysMet . Chapter 3 gives an overview of Dynamic Spectrum Access Network Architectures, and presents the architecture model used in this work. Chapter 4 presents preliminaries and primary formulations for SysMet , along with the details of the system design and design analysis. Chapter 5 presents formulations and results for the specific issue of monitor allocation, along with an empirical study and

evaluation of the solution obtained. Chapter 7 presents the current limitations of SysMet , directions that may be taken for building further upon this work, as well as general directions for misuse detection systems in the future.

The main contributions of this work are as follows :

- We present a system design which allows the monitoring of spectrum activity by BSs
- We present a game theoretic formulation for the problem of monitoring BS spectrum activity
- We present a polynomial time algorithm to compute the monitor densities required to impose a given probability of observation
- We present the main architectural, security and privacy issues for such a system, and propose solutions for the same

Chapter 2

Related Work

This chapter briefly reviews the works and techniques relevant to this work, presented in the state of the art in different areas. We review works on game theoretic formulations for security, followed by works dealing with the main security issues in cognitive radio networks. These topics are the basis for our approach to misuse detection, and include works on Agent Based systems for security. We then move to review work in the area of Privacy for Location Based Services, which give us insights into the main issues and techniques for addressing privacy issues for scenarios where location data must be protected to prevent unauthorized use. We also review random geometric graph models, since wireless network models are in many ways similar in nature - results regarding mobility and guided link formation are relevant to this work as well.

2.1 Game Theoretic Approaches to Security

Problems in the Security domain are amenable to solutions through game theoretic formulations, since entities in such problems can be modelled as rational selfish players who optimize a given objective quantity such as cost. Depending upon the properties of a specific scenario, a problem may be formulated into one of many types of games.

In [36], the authors formulate an intrusion detection problem as a 2-player non-cooperative game, with multiple player types. Here, each player picks strategies based upon her belief about the type of the other player. This is an example of a *multi-stage game with incomplete information*, where players play according to their beliefs, and revise their beliefs after each interaction.

In contrast, many situations may be modelled as games where one player has full or close to full knowledge of the strategy used by the other. For instance, in [34], tasks such as scheduling police activities have been formulated as 2-player Bayesian Stackleberg games,

where one player commits to a randomized strategy and the other player responds, given this knowledge.

Games have also been used to evaluate the privacy risk faced by users. For instance, in [45] and [46], the privacy risk to users of a location based service is formulated as a Stackelberg game, where the users commit to given strategy of privacy protection, and the knowledge of their movements that an attacker can gain is evaluated.

Game theory has been applied to design systems for SUs to equitably decide the usage of fallow bands. It has been used to design secondary auction markets, where SUs may participate in auctions where they trade the right to available bands based on their own valuations. For instance, in [12] the authors present an auction framework where SUs exchange "credit tokens" - those SUs which are offering bands optimize a social welfare objective, such that band rights are alloted to non-interfering SUs. The system also allows each requesting SU to optimize a personal objective function when valuating and bidding for a band.

For a extensive survey on the topic, refer [32].

2.2 Security in Cognitive Radio networks

Security in CR networks and DSANs is being studied with respect to every part of the cognitive cycle [21]. Threats have been identified in both spectrum sensing based CRs, as well as database driven CRs [33].

In spectrum sensing based CRs, threats to PHY and MAC layer mechanisms enable attacks which exploit naive spectrum sensing to gain benefit. For instance, in Primary User Emulation Attacks [18], malicious SUs emulate PUs during transmission, thus forcing other SUs to vacate the band. Attacks on control traffic, such as Beacon Falsification [11], enable brute Denial of Service attacks [] or sophisticated attacks to gain benefit. For Database Driven CRs, inference attacks on the data present in the database, as well as on the queries received from SUs, are possible. Other threats, exploiting weaknesses in the Database Access Protocol, may lead to unauthorized access to data.

One of the main techniques explored to prevent policy violations is the use of Policy Conformance Components (PCCs), along with a policy language which can express the full breadth of policies which may be needed at a device. In this model, different subsystems in a device submit transmission requests to the radio subsystem which then performs a policy match to check if the transmission is allowed under currently known policies. The Radio subsystem would be equipped with a Reasoning Engine or a Policy Reasoner, which must be capable of interpreting complex policies and producing decisions quickly and correctly. For instance, a policy reasoner should be capable of providing decisions for partial or incorrect transmission requests. Ontology based Policy Reasoners are particularly useful, since they allow a wider range of policies to be used.

The idea of "tattling" on neighboring nodes has been mentioned in [33], and is similar to the premise of SysMet . This work explores this idea more deeply. Such agent-based security has been explored in the context of Anomaly and Intrusion Detection for wireless ad-hoc networks. Works in these areas present metrics, learning and classification algorithms to classify scenario-specific objects such as packets, events, etc. Some surveys on these areas are [14] [37].

2.3 Privacy in Location Based Services

Location Based services available to users are vulnerable to privacy violations, more so than normal. This is because most location based services today operate without a profit or contractual obligation to protect user privacy [44]. A rigorous framework for designing and evaluating privacy mechanisms is presented in [44].

The question of PU and SU privacy has been raised in the context of database driven CRs. Many techniques such as k-anonymity, t-closeness [31], etc. have been proposed to anonymize a requesting user. k-anonymity address *identity disclosure* attacks by ensuring that any record can associated with any user in a k user class. t-closeness addresses *attribute disclosure* attacks, by ensuring that the distribution of an attribute within a class is not far from the distribution over the entire dataset. This prevents attacks which discover a most likely estimate of an attribute for users in a class, where the estimate is improved by the class structure.

As opposed to intuitive approaches to privacy, a systematic approach to analyzing privacy has been presented in [44], and can readily be applied to the problem of PU and SU privacy, to obtain optimal strategies for privacy preservation. The framework developed in [44] allows a used to find the optimal privacy preserving mechanism when attacked by a rational attacker, under model assumptions such as mobility behavior obeying the Markov property. The author presents a definition for several inference attacks on partial, anonymized or obfuscated traces from users of a location-based service. A Markov chain where each state is a location is used, and different metrics for privacy are discussed. Given the traces of a user - the attacker estimates the most likely mobility profile of the user, in the form of an estimate of transition probabilities or steady-state probabilities. With this basis, a number of attacks are defined and optimal protection mechanisms against an optimal attack are found.

2.4 Random Geometric Graph Models

The idea of coverage has been explored in the body of work on random geometric graphs [9]. For a substantial introduction, see [38]. Intuitively, it can be said that geometric random graph models are relevant since they include a sense of proximity between nodes, while pre-

serving the analytical tractability of random graph models. For instance, geometric random graph models have been used to derive theoretical limits on connectivity and throughput in a wireless ad-hoc network [2] [51]. Further, they have been used to analyze the spread of contagions on geometric networks [48], and to analyze novel models which solve known security issues in ad-hoc networks. For wireless networks with well-known link properties, it is also possible to make assumptions of a clique over small areas, and analyze these networks using well-known models such as Erdos-Renyi model, or the Watts-Strogatz model.

Mobility, too, has been considered in this context, and results for interesting metrics such as detection time of a new node have been obtained for certain graph models [40]. One must keep in mind that these results make certain assumptions which may not always be realistic. For example, [40] assumes that nodes undergo Brownian motion. This may make analysis tractable, but such results cannot be used in real-world scenarios.

Chapter 3

Dynamic Spectrum Access Networks : An Overview

CRN (Cognitive Radio Network) and DSAN (Dynamic Spectrum Access Network) technologies aspire to enable a device to choose the best possible band to operate in[3]. In this Chapter, an introduction to network architectures, mechanisms and issues in CRNs and DSANs is presented. We review popular reference architectures for CRNs, and then present the model and the relevant assumptions used in this work. We also review work on mobility in this context, identifying relevant models as we proceed. We then do an in-depth review of security in CRNs, looking at different approaches and an attack model. We then motivate misuse detection for CRNs and DSANs, with this model as basis.

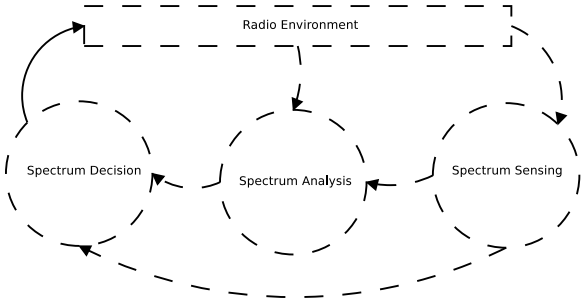


Figure 3.1: The Cognitive Cycle: Information Flow

Dynamic Spectrum Access is enabled by CR devices which can conduct spectrum sensing and manipulate transmission parameters in software. Each device runs through the "cognitive cycle" during operation, as shown in Figure 3.1 - obtaining spectrum information through spectrum sensing, analyzing spectrum information to decide upon the "best" band to operate in, and deciding upon the corresponding operating parameters[21]. To carry out the cognitive cycle, each device is equipped with reasoning and learning mechanisms[20] which allow it to decide upon a band on the basis of well-known policies, as well as policies learnt during

operation.

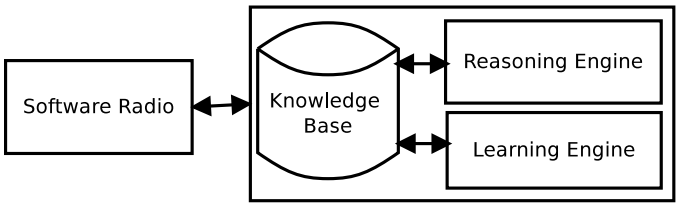


Figure 3.2: Cognitive Radio: Component View

The spectrum decision can be based upon a number of factors - for instance, the current state of the band or required signal power. Since the operating parameters are unrestricted in hardware, allowing CR devices to make decisions purely selfishly could result in harmful interference to incumbent primary transmissions, among other issues. Thus, mechanisms to issue operating policies and parameters have been developed. The CR device must consult these Spectrum access policies before transmitting on a given band. The reasoning engine in Figure 3.2 is responsible for interpreting spectrum access policies required by primary users and other empowered regulatory entities. Once operating parameters are decided, matching policies are consult to ensure that the transmission will not violate any policy.

3.1 The IEEE 802.22 Reference Architecture

The IEEE 802.22 standard on Wireless Regional Area Networks[47][1] is a standard for the use of TV broadcast bands in rural areas, without interference to incumbent primary transmissions. The reference architecture put forth in this standard provides a basis for envisioning the architecture of Dynamic Spectrum Access based systems and networks in the future. Thus, it gives an idea about the architecture of DSANs of the future, and is a starting point for the DSA network architecture model that is assumed in this work. The 802.22 standard is aimed at providing a WRAN in rural areas using TV bands. Each cell is typically 15-30 KM in radius, and is managed by a single BS. Each cell typically contains 255 CPEs, whose use of spectrum resources in dictated or guided by the BS. These CPEs may be terminating user equipment, or may be relay CPEs. It also stipulates that the location of each BS must be known within 10 meters of the true location, while the location of a CPE must be known within 100 meters of the true location.

Works such as [15] present architectures similar to the IEEE 802.22 reference architecture, and present additional paradigms such as infrastructure based, ad hoc and mesh network architectures for CR networks.

Dealing with the full breadth of the model set out by such works, is currently not in the scope of our work. A subset of network architectures are considered, by imposing additional restrictions on the nature of BSs and CPEs.

3.2 A Model of CR Network Architectures

With these works as inspiration, we assume the following model for DSAN and CR network architectures. We use the following model for the architecture of a DSA network in our analysis :

- A Base-Station (BS) manages a cell of arbitrary size, and guides the spectrum access of several Customer Premises Equipment (CPEs)
- A CPE may act as a relay, but its spectrum access is still guided by the BS
- The BS does not change location frequently, and its location is known to within 10 meters of its true location
- The CPE is mobile - that is, its location changes frequently, but its location is known to itself, and may be conveyed whenever required
- Both the BSs and CPEs have access to a Common Control Channel (CCC), which is used as required.
- In order to enable correction or compensation for improper BS behavior, a CPE managed by a BS may collect information which would allow the BS to be identified, in a form that is verifiable by a third party. This assumption is made purely to enable correction or compensation for improper BS behavior.

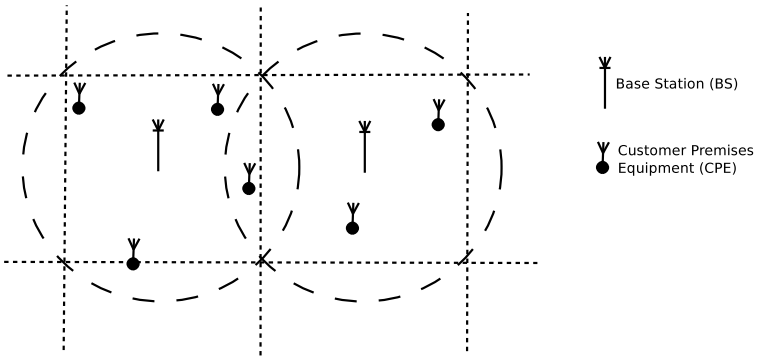


Figure 3.3: DSA Network Architecture Model

We state model assumptions, along with justifications, below :

- The ability of a CPE to report its location in a form verifiable by a third party, is a bold assumption, but reasonable in light of works such as [53] and [27]. These works allow location verification by judging the proximity to well-known devices, and may be adapted to allow location verification for this model, as explained in Chapter 7.

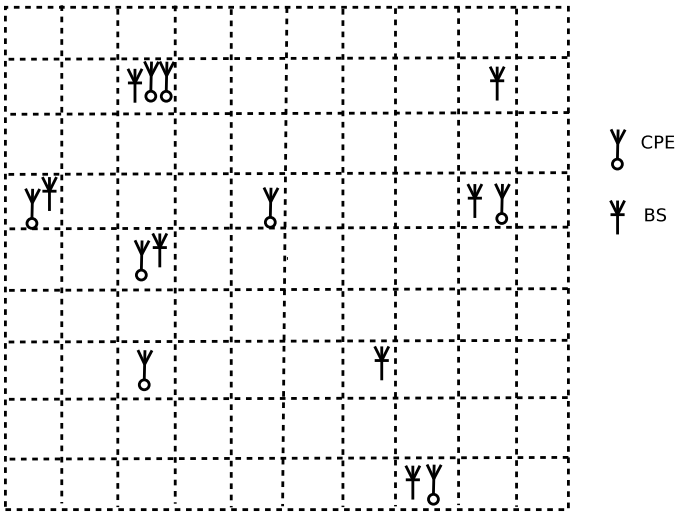


Figure 3.4: DSA Network Architecture Model

- A CPE has the ability to detect "misuse" of spectrum at least within its cell. This assumption is bold, since metrics such as power estimates made from large distances (from a transmitter) must compensate for the attenuation experienced by the transmission. This assumption can be considered reasonable, since we assume a small cell size for now. Ongoing research in the area could possibly identify metrics which can be measured at reasonable distances, and such work would increase cell size, and would only simplify the problem formulation.

- Each CPE has the required hardware and software to collect spectrum activity information while going about its regular activities. This could include additional radio components, but is not limited to this. The CPE is at-least capable of multiplexing requests to radio subsystem such that required probability of collection $p_{collect}$ is met. This assumption is reasonable since at-least some resources must be devoted to the collecting spectrum activity information. For instance, if the required $p_{collect}$ is 0.3, then the CPE can carry out other activities using 0.7 of its overall radio capabilities.

- A CPE can deduce the identity of a transmitter from its transmissions. Knowledge of a BS's identity can also be considered available, since CPEs report information such as their current location, and meta-data about their interaction with the BS. Works such as [28], [43] and [52] advocate PHY-layer measures to allow identification of transmitters, and thus this assumption is reasonable.

- We assume BSs which move infrequently. This can be considered reasonable, since mobile BSs are unlikely to be harmful to primary transmissions over significant areas. This assumption is made purely to aid problem formulation, and can be removed with minor modifications to the formulation.

- Beyond the above model, it is also assumed that the CPE can access a third-party

system server intermittently. The messaging between the system server and a CPE is delay tolerant to a large extent, and the CPE need not be continuously in contact with the system server.

To aid our problem formulation, we assume a square tessellation over which BSs and CPEs are distributed. A BS or CPE can be present any cell, though the BSs are confined to a single cell, while CPEs may move from one cell to another. It is assumed that the number of CPEs which can be considered be "within" the tessellation, does not vary on the timescales considered.

This model may be extended to include scenarios which involve mobile BSs, and ad hoc network architectures. This is not in the scope of this work, but directions for the same are presented in Chapter 7. Another direction that is not explored in this work is the information (such as location) used to actuate a mechanism to prevent or correct BS behavior. We simply assume that the CPE is privy to the identity of the BS, in a form verifiable by a third party.

3.3 Modelling CPE Mobility

Human mobility has been studied in-depth, in a variety of fields like urban planning and wireless networks. Human mobility may be studied at various granularities - at the scale of cities, as done often in urban and environmental planning, and at the scale of meters and seconds, as done in wireless networking literature. Human mobility exhibits distinct power laws in many aspects, modeling which has been difficult. It was Paul Krugman remarked in his book that though power laws give rise to a seemingly simple reality, they lead to complex and unwieldy models. Many models which have expounded simple rules remain analytically intractable. Historically, human mobility has been analyzed with a view that collecting real mobility traces is difficult. But recently, statistically accurate and practical models for human mobility have been discovered by analyzing real GPS based human mobility traces. [42]

3.3.1 Mobility at the Scale of Individuals

Wireless networking literature has expounded a number of models to model different aspects of human mobility. These models have been used primarily in the context of ad-hoc protocols for MANETs. These models have been used for 2 main purposes - as means of evaluation through simulation, and as theoretical basis for results. A number of models have been proposed and used [13] [7], but many are limited by the fact that full trace collection and trace based models were seen as difficult to accomplish.

Some well known models for modeling human mobility for simulation include :

- Random Walk : A speed and a direction are chosen uniformly randomly. The node then "moves" for a predefined time t or for a predefined distance d.
- Random Way-point : A speed is picked uniformly randomly, and destination is picked uniformly randomly from a set of predefined destinations. Once a node reaches a destination, it chooses a pause time uniformly randomly
- Gauss-Markov Mobility : A speed and direction are computed using a weighted sum of the previous, the average and a random variable with a Gaussian distribution

Additionally, assumptions to prevent edge effects, such as a toroidal simulation area or rebounding are used.

Brownian motion [] has been used in many works [39] due to its analytical tractability, but it fails to explain the heavy-tailed nature of human mobility.

A recent model is the Truncated Levy Walk model, a fine-grained model based on real GPS traces of humans[42]. This model has many of the characteristics of real human mobility, such as heavy tailed flight times and pause times. This model is used as the primary mobility model is this work, and is explained further in Section 4.1. Supplementary models like [29] have also been proposed, which introduce way-points and path optimization carried out by humans.

In the absence of works which perform a rigorous analysis of human mobility at the scale of a city, we are forced to assume that models such as the TLW will not remain valid when considering scales beyond those found in [42]. As a workaround, we design an additional refresh mechanism for SysMet , which triggers several activities which correct for this issue.

3.3.2 City-Wide Mobility

Anas et al [4] provides a dated, yet extensive survey of human mobility analysis existing in urban planning literature. They also provide a good survey on early models, such as Getis's analysis of population clusters in the Chicago Area. Some basic results from Getis's work mentioned in [4] show that up to scales of the order of 0.7 miles, population densities remained uniform. But at scales of 8 miles, densities can be modeled by a Poisson process - concentrated at a center and decreasing outwards. Other studies have included those that analyze the population distribution across cities, an example being [26]. In [10], the authors attempt to characterize the nature of human mobility using mobile phone traces, at the scale of economic centers within a city or town.

Such studies provide a general notion of human mobility at coarse granularities, but have not been rigorous enough to point towards a concrete model at these scales. Intuitively, we can say that movement at such scales is governed strongly by city structure, time of day,

etc., and thus a model that successfully accounts for such a variety of aspects would be hard to formalize.

3.4 Roles of Secondary Users

SUs are varied in nature - from individual or terminal SUs, to relay SUs, to SUs which provide enhanced quality of service to other SUs in subtler ways. With this in mind, SUs must utilize fallow spectrum in an efficient manner - with their access coordinated to provide a reasonable and consistent quality of service for all. Beyond MAC or PHY layer measures such as contention protocols, additional coexistence mechanisms such as secondary markets [50] and token based spectrum etiquette frameworks [12] have been proposed. Another paradigm for the use of fallow bands is through contracts where a SU pays the PU on the basis of the SU's usage of the band.

Such paradigms provide a framework for SU cooperation, and can also provide an economic incentive to PUs to support access to a band licensed to them. Such paradigms also bring SU has an inherent motivation to detect and prevent spectrum misuse - primarily because they want, or have paid a cost for good quality of service. This is especially true for small, mobile CPEs which may not have the resources to compensate for the loss in quality of service they experience.

The state-of-the-art uses PCCs which enforce policies with multiple levels of priority would be a good measure to control spectrum access. These paradigms would require several additional mechanisms to be viable - the most important ones being the ability to observe SU spectrum access without foreknowledge of SU activities and providing SU privacy. SysMet can provide such a system, by enabling a mechanism to report information about spectrum activity, such as identity information of devices which are transmitting on a given band.

3.5 Security in Dynamic Spectrum Access/Cognitive Radio Networks

Security for CR networks may be analyzed using 2 approaches - by presenting solutions to known attacks or threats on current designs, or by considering preventive measures such as Policy Conformance Components (PCCs).

3.5.1 Attack-based Security Measures

As presented in [33], security threats in CR networks in the state-of-the-art may be categorized into those related to a spectrum sensing based approach for CR systems, and those

related to a database-driven approach.

For spectrum sensing based approaches, the identity presented by an incumbent primary must be robust. If it is not so, then the systems becomes vulnerable to Primary User Emulation(PUE) attacks, where a malicious SU may force other SUs to vacate a band, thus making spectrum contention between SUs unfair. In [17], the authors present a localization based approach as a countermeasure to the PUE attack.

Spectrum Sensing Data Falsification attacks exploit the cooperative nature of spectrum sensing to inject false spectrum sensing information into the network and gain benefit. In [16], the authors present a robust framework which, given attacker and normal user reporting strategies, weights spectrum sensing results from a node's neighbors. A Weighted Sequential Probability Ratio Test is used update the weights assigned to neighbors. [21] [35]

3.5.2 An Attack Model for DSANs

We consider the following attack model to design and evaluate approaches for misuse detection. Before using a particular band, an honest BS would obtain availability or policy information for the band from the primary user, using either a database query, or local spectrum sensing. It may also enter a usage-tied contract with the primary user, if required. A dishonest BS could either ignore the primary user's requirements altogether, or more subtly misreport usage information for benefits on a usage-tied contract.

Thus, the attack model consists mainly of 3 kinds of attacks :

- Violations of spectrum access policies
- Attacks which are not direct violations of spectrum access policies, yet cause anomalous behavior
- Mis-reporting usage information to gain benefits on usage-tied contracts

Spectrum Access policies contain a description of the parameters to be used while transmitting on a given band. These could include "prohibitive" policies, which mandate zero transmit power for certain bands under some conditions. A BS may violate spectrum policies intentionally, or may do so unintentionally if it is under attack.

If a BS enters a usage-tied contract with the the primary user, it must pay the primary user according to its use of the band. The BS may not be honest about its use of the band, and may under-report usage. The primary user would have no way to verify usage reports made by the BS.

Additionally, some attacks may cause detectable anomalous behavior, but do not cause violations of any particular access policy. For instance, a Primary User Emulation attack would

not violate any spectrum access policy, yet would cause anomalous behavior. Any attack which causes anomalous behavior detectable by the SUs in the vicinity may be included in the attack model.

3.5.3 Preventive Security Measures

Under the second approach, Policy Conformance Components on-board each CR device are used to enforce spectrum access policies by checking every potential transmission against policies transmitted to the device. To this end, Policy Reasoning systems [] are part of the state-of-the-art.

A system to detect and identify policy violators or "misuse", would also be required to detect any attacks which cause policy violations, or exploit vulnerabilities in policy conformance components. Such a system was mentioned cursorily in [33], but we present a complete system for "misuse" detection.

3.6 Misuse Detection - Approaches

For misuse detection in general, a number of approaches may be taken. Trusted components within CR devices is one, but is limited by the scope and practicality of foolproof trusted components. Also, they can a good measure to counter en mass violations of policy, but measures to detect localized and byzantine violations are also required. An observation based system for misuse detection is thus important to enforce disciplined spectrum access.

Within the category of observation based systems, we consider the following approaches. A first approach would be allow CPEs to perform misuse and anomaly detection independently. This approach would be limited in its scope as detecting misuse may require correlating multiple measurements or observations. Trusting CPEs with certain anomaly detection capabilities is also possible, but would be resource intensive. CPEs would be unable to perform analysis which correlate inputs from multiple CPEs, without a resource intensive cooperation protocol.

A second approach would be to enlist every CPE as a monitor, and mandate that all CPEs must report all transmissions encountered during spectrum sensing, together with all CCC messaging. This information may be fed to an anomaly detection system which flags any indications of spectrum misuse. This approach would be highly resource intensive too, and would require real-time transmission and processing of CPE data. It could also potentially violate the privacy of both the BS and CPE as well, even if some anonymity mechanisms are placed. Another issue would be the need to compensate a large number of CPEs for resources expended.

An improved approach would be to place some trust in a subset of CPEs, and employ them

to observe spectrum activity under some sampling regime. This would allow a desirable degree of control over the data collected, as well as provide avenues for maintaining BS and CPE privacy. This third approach is the inspiration for SysMet . It allows a central system to impose a required observation probability in each cell, while requiring CPEs to expend only a given fraction of its spectrum sensing or transmission opportunities.

Under the assumptions of our model, mechanisms for misuse detection can be characterized as mechanisms which ensure disciplined BS spectrum access behavior. A system which implements the 3rd approach discussed above, must be designed with the following aspects in mind :

- Utility : A Misuse Detection System should provide the ability to detect and locate a policy violator
- Privacy : It should provide mechanisms to reduce the privacy risk posed by an observation based system. It should not expect information which may be considered sensitive or critical to a PU or an SU.
- Trust : It should have a mechanism, a reward mechanism or otherwise, to support or affect the prescribed behavior from selfish CPEs and SUs.
- Robustness : A Misuse Detection System should be robust to byzantine failures or malicious attacks. For instance, heavy reliance on a small number of fixed infrastructure devices or CPEs may make the system vulnerable to malicious devices which masquerade as trusted devices.
- Efficiency : The system must be reasonable in its resource consumption. Measures such as shifting resource intensive tasks away from resource constrained devices should be considered.

3.7 Conclusion

To conclude, we consolidate and re-state the most important model assumptions developed in this section :

- CPEs have the ability to detect misuse and record relevant transmissions, at-least within 1 cell of the tessellation it is present in. The recorded information allows the identification of the violating transmitter.
- CPEs are willing to devote a given fraction of their resources to spectrum misuse detection activities. This would include the use of a given fraction of spectrum sensing or transmission opportunities.

- CPEs can report their current location in a form verifiable by a third party.

It can be concluded that SUs, particularly mobile CPEs, have an inherent motivation to report or prevent spectrum misuse. Additional rewards could be used in case the CPE's privacy risk is not compensated by the requirement of good QoS. Penalties, in terms of a trust metric or otherwise, could be used as countermeasures against malicious CPEs.

SysMet employs an approach for misuse detection which gives the system designer control over the information collected, as well as provides a theoretical basis for preserving CPE and BS privacy. The design for SysMet must consider the privacy risk of PUs and SUs, while being robust to attackers itself. CPE trust must also be considered. A CPE's reports may not always be trustworthy, and thus some mechanism for accounting for this must be used. This may be done in many ways - providing redundancy, or by using a trust metric, which is updated periodically. In its current form, SysMet does not provide for this, as explained in Chapter 7.

Chapter 4

Misuse Detection in Dynamic Spectrum Access Networks

As discussed in Chapter 3, misuse detection and spectrum access verification are important, in order to ensure that spectrum access policies mandated for a band are enforced, and also to locate and identify policy violators. SysMet carries out verification and misuse detection activities by employing a network of "Monitors" - a set of CPEs which expend resources to carry out monitoring, in return for rewards of some kind.

The monitors are selected such that a given probability of observation P_{caught} is enforced in the given area. The probability of observation is obtained from a game formulation involving the system and a malicious BS. A formulation of a similar flavor was presented in [49] and [5]. The monitor requirements for a given area are computed using a Integer Linear Program, which obtains the minimum monitors required to enforce P_{caught}.

CPEs interested in becoming monitors must request the same from an allocation server. Not every CPE has the ability to become a monitor in a given epoch - It must possess a valid "seed", which is provided by the allocation server during registration. An overview of SysMet is presented is Section 4.3.

This Chapter is organized as follows - We first briefly overview the preliminaries required for the problem formulation, followed by the game theoretic formulation and a detailed system design. We then perform an analysis of the security and privacy aspects of this design.

4.1 Preliminaries

In this section, we discuss mobility model assumed for the CPEs, as well as the basics of additional tools for our analysis. As explained in this section, the Truncated Levy Walk (TLW) Model [42] is used to model the mobility of CPEs.

4.1.1 Stackelberg Games for Security applications

As discussed in Chapter 2, Game theory has been used in the context of Security since it can be used to model the behavior of a rational attacker. Specifically, a Stackelberg game models scenarios where the Security mechanism must commit to strategy which is well-known.

In a Stackelberg game, player one - called the "leader" - commits to a strategy and makes it well-known. Player two - called the "follower" - must now play a strategy which maximizes his expected reward. Though the leader revealed its strategy and forfeited his right to change its strategy can force the follower into a strategy which maximizes expected reward for this particular strategy.

A Stackelberg game would be a *2-player game with complete information*. The ability of player one to move first and become a leader is critical, and is made possible by complete knowledge of the nature of the game - such as the both the strategies as well as the rewards for the follower. For instance, in [34] the authors formulate a 2 player Stackelberg Game with multiple player types to model the scenario of optimizing spot checks at airports. Note that there are no requirements of *perfect information*, in that full knowledge of each leader move is not required.

4.1.2 Integer Linear Programming

Integer Linear Programs (ILPs) maximize the an objective function for scenarios each of the variables in the takes integer values. ILPs and Mixed Integer Linear Programs (MILPs) are a large class of problems that map well to real-world situations, such as resource management and scheduling. In the context of wireless networks, they have been used in problems of flow routing and relay placement in sensor networks [24] and efficient spectrum sharing[23].

The body of work in ILPs,MILPs and MINLPs provides a robust framework to develop techniques for other fields, such as Wireless Networks and Machine Learning. ILPs, are NP-hard if an exact solution is required, approximate solutions which are provably close to the optimal can be found in polynomial time. The main technique to accomplish this is to relax the ILP to a Linear Program (LP), and then iteratively constrain the solution space. These constraints can be in the form of additional inequalities, or equalities for certain variables.

To elaborate, there are primarily 3 kinds of methods used to solve ILPs :

- Cutting Plane methods : This class of methods divide the given solution space iteratively, and attempt to find integer solutions which are optimal in these smaller spaces. The inequalities used to divide the solution space could be created in a number of ways, usually using the optimal solution of the relaxed LP from the current iteration.
- Branch & Bound methods : These methods explore systematically explore the solution space, keeping track of the branches already explored and using conditions to judge

whether a candidate solution may be found in a branch. These methods, particularly branch and cut methods, divide the solution space by choosing a variable, and re-solving the LP with added constraints such that the current solution becomes infeasible. If the value of the objective function along any branch is worse than the best known integer solution, that branch is not pursued further. Versions of this method can find solutions which are provably close to the optimal.

- Sequential Fixing methods : These methods constrain the solution space by adding equalities which fix the values of certain variables, and then iteratively solve the resulting LPs. These methods give good solutions in practice, but no bounds on the quality of the solution can be obtained.

As explained in Chapter 5, experiments found that branch and bound methods were inefficient for problems which have a large number of variables and where the importance of each variable to the objective function is similar. It was noticed that adding bounds for one variable would result in a mirroring of the previous optimal value into one of the other variables.

4.1.3 The TLW Mobility Model

Human mobility has been found to exhibit Heavy-Tailed distributions of flight length and pause time[22][42]. This makes the nature of human mobility different from Brownian Motion, in terms of diffusivity. Mobility models such as Random Way-point (RWP) [25] exhibit diffusivity and heavy-tailed distributions similar to human mobility, but typically make analysis difficult. The Truncated Levy Walk (TLW) model [42] tries to achieve an accurate characterization of the statistical properties of human mobility, while being analytically tractable.

The TLW model models human mobility as a 2 dimensional random walk, where each step is characterized by the tuple S = $(l, \theta, \Delta t_f, \Delta t_p)$ - the flight length (l), direction (θ), flight time (t_f) and pause time (t_p). The direction is chosen uniformly randomly, while t_f and t_p are obtained from the distribution (expressed below as its Fourier transform) :

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx - |ct|^\alpha}$$

where c, α and β are simulation parameters which can be considered analogous to diffusivity. Note that setting $\alpha = 2$ reduces the distribution to that of Brownian motion (diffusivity is 1). Δt_f and Δt_p generated from this distribution are truncated to values τ_f and τ_p - that is, for samples greater than τ_f and τ_p or less than 0, the value is set to τ_f and τ_p.

A speed model of the form $\Delta t_f = k * l^{(1-\rho)}; 0 \leq \rho \leq 1$, is used to obtain l.

The TLW model was obtained by analyzing real user GPS traces for multiple locations. Using the TLW model, an analysis of occupation probability and upper bounds on critical delays for mobile networks can be found in [30].

4.2 Problem Formulation

4.2.1 Game Theoretic Formulation

If a BS intends to misuse spectrum, it would have to take into account both the pay-off from an instance of misuse, as well as the penalty it faces if its misuse were discovered. Intuitively, given a large enough penalty for misuse, a dishonest BS will refrain from it because the risk outweighs the reward to be gained from misusing spectrum. We formalize this notion by formulating a 2-player zero-sum Stackelberg game between the BS and the verification system. Such as formulation can be used to determine what the probability of observation (P_{caught}) should be, given the pay-offs that a BS receives from misusing spectrum.

Suppose a verification system was able to monitor BS usage, such that any particular instance of spectrum activity would be observed with a probability P_{caught} . In this scenario, we have the following zero-sum Stackelberg game, where the verification system is the leader :

1. The verification system imposes the probability P_{caught} , and declares the same.
2. When a BS cheats when not observed, the verification system must pay the BS a reward C_{misuse}. This cost represents the loss incurred by the primary user, which we model by a cost to the verification system.
3. If a BS is observed when cheating, then the verification system earns a reward R_{caught}. This represents the penalty the BS must pay for misuse - it may or may not be monetary in nature.

Let the BS strategies be "Cheating"(Ch) and "Not Cheating"(NCh). Let the Verification System Strategies be "Check" (Ck) and "No Check" (NCk). The Verification System is known to play Ck with probability P_{caught} .

Thus, the pay-off matrix is as shown in 4.1.

For now, we assume that there is not additional cost for observing key facts when no misuse is in progress. This cost will be modelled later, using an LP formulation for the cost of imposing P_{caught} .

The expected pay-off for the BS for each of it's strategies (Ch and NCh) would be :

$$Payoff_{Ch} = P_{caught} * (-R_{caught}) + (1 - P_{caught}) * (C_{misuse})$$

SysMet		BS: Ch	BS: NCh
	Ck	$-R_{caught}$ R_{caught}	0 0
	NCk	C_{misuse} $-C_{misuse}$	0 0

Table 4.1: Cheater's Choice: Payoff Matrix

$$Payoff_{NCh} = 0$$

For a given value of P_{caught} , a rational BS would play Ch only if the expected pay-off is positive. If the expected pay-off is negative, it will never play Ch.

Thus, the minimum required P_{caught} to prevent misuse would be such that :

$$P_{caught,reqd} * (-R_{caught}) + (1 - P_{caught,reqd}) * (C_{misuse}) \leq 0$$

$$P_{caught,reqd} \geq \frac{C_{misuse}}{C_{misuse} + R_{caught}}$$

Depending upon the specific values of the pay-offs, a $P_{caught,reqd}$ can be determined.

4.2.2 Imposing P_{caught}

In order to impose a given P_{caught}, a subset of CPEs - called the Monitor set - is employed to verify spectrum access conducted by BSs that they encounter as they move through a given area.

As monitors connect to BSs and obtain services, they record information about the channel state and the interaction, etc. This information could include facts such as the bands being used, the start and end time of the interaction, etc. At the end of the day, they report these facts to the primary user. The mobility of CPEs is exploited to efficiently monitor BSs in a large area, and impose a minimum probability of observation (P_{caught}).

The area under consideration is divided into a square tessellation, with each BS's range covering an area at least equal to the area of 1 cell . To obtain the minimum number of monitors required to impose a given P_{caught} , we solve an ILP in the densities of monitors at each cell from which CPEs can send requests, and the visit probabilities at each cell . We call this ILP ILP_1 . This ILP gives us both the optimal value of the total number of monitors required (N_m), as well as the corresponding set of Monitor densities.

In conclusion, our approach can be divided as follows:

- Obtain Visit Probability for each cell
- Solve ILP_1 to obtain $N_{m,optimal}$ and monitor densities for each cell from which monitor requests were recieved

4.3 System Design

SysMet is a mechanism which allows CPEs to request monitor status from the SysMet Allocation Server. A CPE which is interested in becoming a monitor, must first register with the Allocation Server , as shown in Figure 4.3. During registration, the Allocation Server associates unique session parameters with this CPE, such as a monitor ID, cryptographic keys, etc. to ensure that any further communication carrying this monitor ID is with this CPE. The CPE too has incentive to prevent sharing of these keys, as its reward is at stake. Even if the CPE shares these keys, every report made by the CPE must be accompanied by verification of the locations reported. The session parameter allocation includes a seeding process, where the newly registered CPE is allocated seeds , which it must present when requesting monitor status in an epoch.

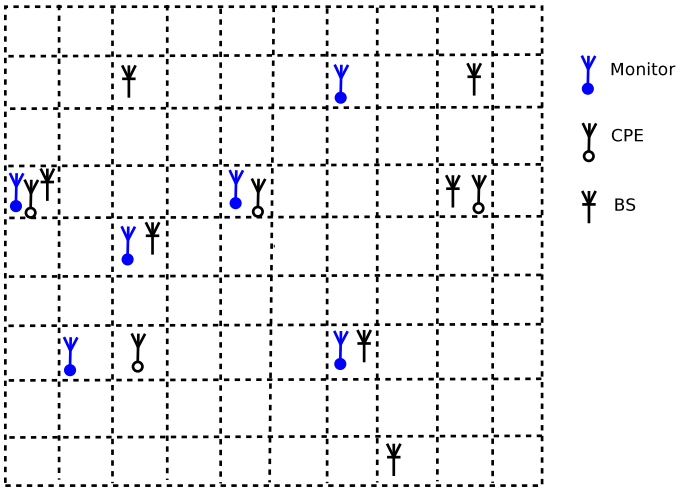


Figure 4.1: SysMet : Monitors

We consider time divided into *epochs*. Within each epoch, we assume that the assumptions of the mobility model are satisfied - for example, we assume that during an epoch, no monitor will leave the tessellated area. Each CPE is informed about the start and end time of each epoch during registration.

The Allocation Server solicits applications for monitors for each cell in each Tessellation , for the given area, during a predefined time towards the end of the current epoch. We call this

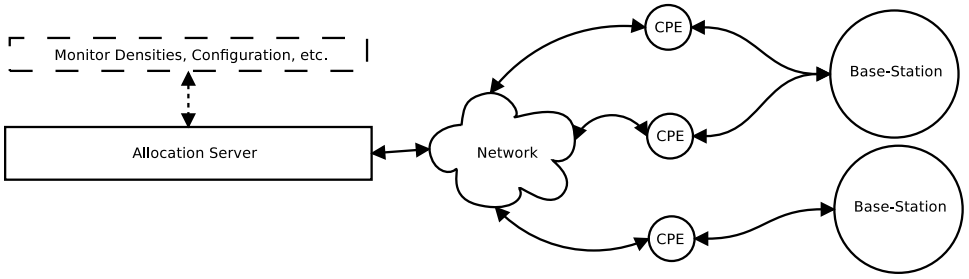


Figure 4.2: CPE-Allocation Server Network
In the model assumed here, The CPE must be able to access the Allocation Server intermittently

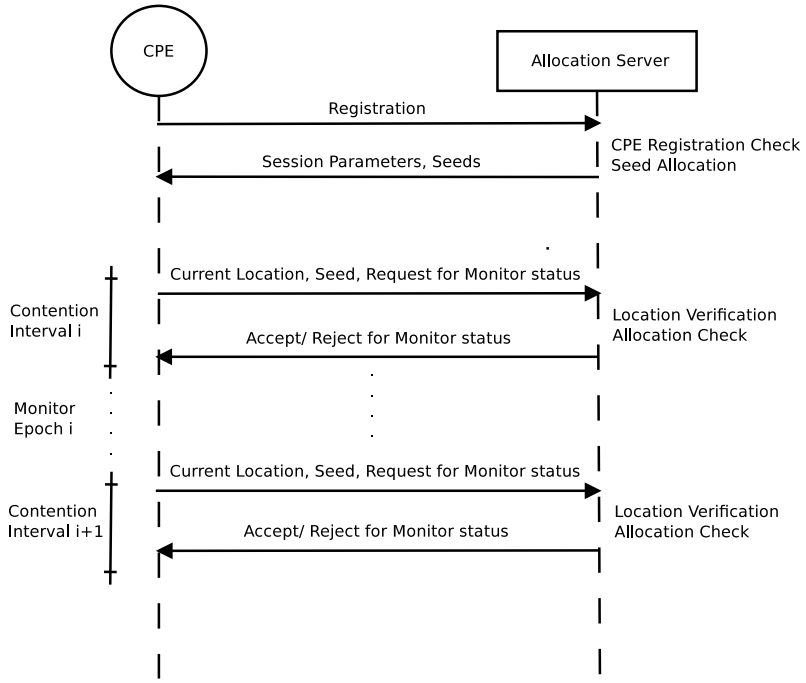


Figure 4.3: CPE-Allocation Server Protocol
The CPE and server communicate during registration, and during contention intervals if the CPE wishes to become a monitor

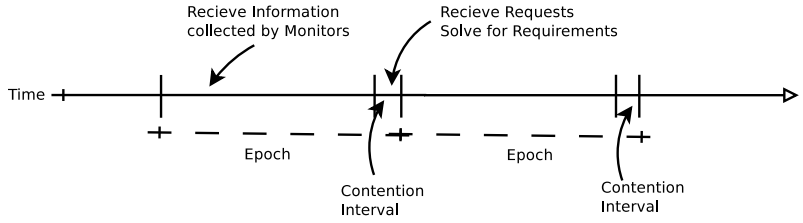


Figure 4.4: Server Activities During System Operation
During each cycle, the Allocation Server receives traces from monitors during the monitor epoch, and receives requests for awarding monitor status during the contention interval. It must use these requests to compute monitor requirements, and then send out accept/reject messages for each request

the contention interval. Once the contention interval has ended, the Allocation Server uses the requests received to compute the number of monitors required on the basis of the solution

of the Monitor allocation ILP, presented in chapter 4.

If a CPE would like to become a monitor, it must request monitor status, along with its current location and additional information (such as a valid seed), in a form verifiable by the Allocation Server . The ability of a CPE to do so is assumed in our model, and justified previously in Chapter 3, Section 3.2. When a CPE requests monitor status, the Allocation Server consults precomputed CPE density requirements to check if a monitor position is available for the requesting CPE's current location.

4.3.1 Registration & Seeding

A CPE must register itself with the Allocation Server in order to be considered for a monitor position. This registration can be optional, or obligated by a service contract. The registration provides the CPE with a monitor ID, which uniquely identifies in the system. It is also alloted "seeds", which are chosen uniformly randomly from the seed space. During its monitor status request, the CPE must transmit this seed to the Allocation Server . The Allocation Server will first check if this seed is valid for this epoch, and whether it was alloted to this CPE. It proceeds only if both these conditions are met.

This seeding mechanism is used in order to ensure that monitor status is not given to CPEs who request monitor status aggressively.

The registration also includes a mechanism to determine if the CPE has registered before in the recent past, and continue with the registration only if it has not. Uniqueness is necessary in order to prevent malicious CPEs from registering multiple times, for increased rewards. In our model, we assume that each CPE possesses a unique private-public key pair, similar to a key distributed on SIM modules today. This key pair can be used to check if the CPE has registered before, by asking the CPE to perform a keyed hash on a nonced string and then comparing the digest to those received from registered CPEs.

Another mechanism would be a zero-knowledge proof protocol towards the secret key on an SIM module given to all CPEs. This zero knowledge proof procedure would also have to distinguish between secret keys.

4.3.2 Requests for Awarding Monitor status

If a registered CPE wishes to be a monitor, it must send a request for awarding monitor status to the Allocation Server . This request would include their current location, location verification information, along with the seed which the was alloted to the CPE for the epoch being contented. The Allocation Server may also initiate a verification of the identity of the CPE. The availability of location verification information and location verifiability is an assumption in our model, but this assumption can be considered

reasonable, as explained in Chapter 3, Section 3.2. Upon receiving this request during a contention interval, the Allocation Server carries out a location verification, a seed check, and possibly an identity verification. It then adds this request to the list of pending requests for that cell, till the contention interval ends. Once the contention interval has ended, the Allocation Server computes the number of monitors required, and uniformly randomly chooses requests to accept from the pending list for each cell. The Allocation Server maintains a list of active monitors for a given Tessellation . At the end of every epoch, the server purges this list. The end of every epoch is predefined for a given area, and is well-known to all registered CPEs.

4.3.3 Awarding Monitor status

Under the regime of a contention interval and a monitor epoch, the Allocation Server has the freedom to obtain all monitor requests for the upcoming epoch. The Allocation Server thus waits for the contention interval to complete, and then uses the number of requests received as input to the ILP formulation described in Chapter 5.

From a system design perspective, monitor status can be awarded by the Allocation Server in a number of ways. Apart from soliciting monitor requests at the beginning of an epoch, Monitor status could also be awarded as the requests come in. Such a mechanism would be useful in situations where monitors leave the given area during an epoch. Though such a mechanism is possible, it is outside the scope of this work. A brief description of a direction for the same is provided in Chapter 7.

4.3.4 Observing Spectrum Activity

Once a CPE has received monitor status, it will be rewarded for spectrum activity information it collects and relays to the Allocation Server . This information, also called a “trace”, would be verified and then analyzed for spectrum misuse. Each trace can contain multiple data-points, or “reports, pertaining to spectrum activity regarding different bands, at different times or for different misuse detection techniques. This is under the assumption that it is possible for such information to be collected. For instance, one kind of misuse would be in the form of transmission using unacceptable signal strength - since signal strength decays quickly, only monitors close by will be able to detect this kind of violation.

Traces submitted by monitors includes the location of the monitor when was collected, as shown in Figure 4.5. This location information must be verifiable, so that a monitor cannot falsify his presence at a location. As we see later, location verification is an open question with several possible approaches. The specific method of trace submission is tied to the method of granting rewards, along with the privacy requirements of CPEs.

Though a monitor may have the opportunity to observe all instances of spectrum activity in

a cell, it may be utilizing transmission or spectrum sensing opportunities which are required for regular operation. To enable the monitor to multiplex between regular operation and spectrum activity observation, the monitor decides to collect a utilize an opportunity to observe spectrum activity according to a collection probability $p_{collect}$. Tha is, if an instance of spectrum activity is observable by the monitor, the instance is collected by the monitor with a probability $p_{collect}$.

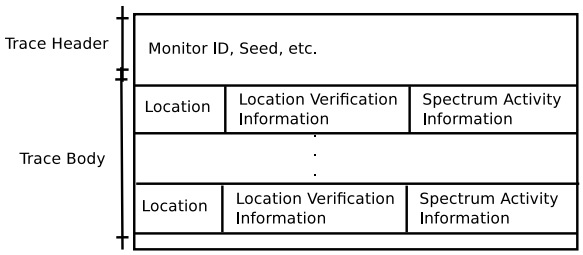


Figure 4.5: SysMet : Trace Format

In the case of usage-tied contracts, the nature of the BS-CPE relationship is that the BS provides services to passing CPEs. This allows a monitor to collect additional information, and it be able to collect specific spectrum usage information which the BS is aware of as well. In this scenario, the information collected by the monitor has the property of non-repudiation, and can be used in a challenge system to ensure that the BS is reporting all usage information to the band license holder or spectrum broker.

4.3.5 Incentives for CPEs

Since CPEs are mobile, they do not have any specific commitment of trust to a BS they encounter. Thus, it is reasonable to assume that CPE would be inclined to participate in a system like SysMet , or that they can be given incentives to do so. As noted in Chapter 3, a misuse detection mechanism is in the interest of CPEs, as it enables them to report and correct issues which effect the quality of service they receive.

Incentives offered to CPEs to become monitors could be a simple absolute reward, or tied to their performance as monitors. As we see later, providing mechanisms of trust or performance for a given CPE is difficult, given the privacy requirements for a system like SysMet . If reasonable privacy is required, then an absolute reward combined with methods to eliminate false trace information would be a better approach. There is also the option of allowing CPEs to accept some well-defined terms which entail which aspects of their privacy may at risk when using this system.

The incentive mechanism for SysMet is one where traces can be associated with a given monitor, but the traces are protected such that the Allocation Server must query the monitor in order to unlock a trace. This is similar to the privacy preserving toll pricing system proposed in [8]. This method is shown in Figure 4.6. The monitor is given additional rewards

according to the number of traces unlocked, and the monitor is free to limit the number of traces unlocked, thus enabling it to decrease its privacy risk as it sees fit. This also provides incentives to the Allocation Server or anomaly detection mechanisms to use as few traces as possible. This method has an additional benefit - since traces can be associated CPEs, a reputation or trust metric may be computed for the CPE. Since the Allocation Server may ask the monitor to unlock any of its traces, a rational monitor would try to keep all traces authentic.

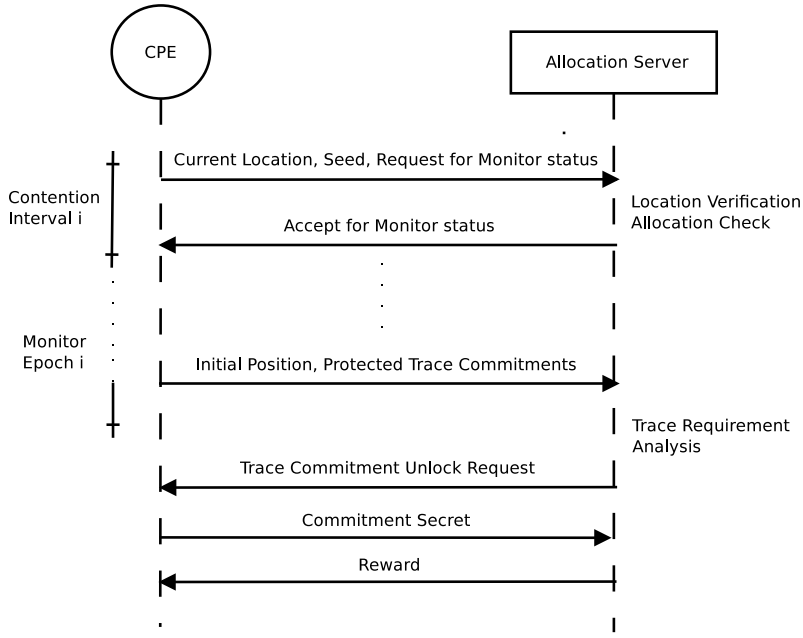


Figure 4.6: Trace Submission & Incentive Scheme

The loss of incentive can be a sufficient motivator for honest monitor behavior, but other schemes can be used as supplements. For example, the introduction of trust metric would allow dishonest actions to effect the monitor selection process, and bring a notion of punishment for dishonesty.

4.4 Security and Privacy Aspects

4.4.1 Security Analysis

We present a security analysis of SysMet described previously, and consider the security consequences and rationale of SysMet .

Monitor Detectability

The ability of a monitor to be indistinguishable from other CPEs is essential. If the monitor reveals this fact in any way, then the game formulation would no longer hold - a rational cheater would try to detect the presence of a monitor before attempting to cheat. For example, a location verification mechanism which uses monitors to generate proof of location would reveal the presence of monitors in a given area.

The Allocation Server must also reveal only limited information about other monitors present in the area. For example - a high granularity system, where CPEs request monitor status for each cell, would reveal the presence of other monitors in the cell by rejecting requests frequently.

Requests for monitor status

A system such as SysMet must consider the security of allowing CPEs to request monitor status at will. Choosing monitors non-deterministically is important, since the system should not favor malicious CPEs which apply aggressively for monitor positions, nor should the system choose the same CPEs as monitors.

With this issue in mind, SysMet requires all candidate CPEs to register at the beginning of the day, and includes a refresh mechanism which allows randomly seeding all registered CPEs with the ability to be a monitor. The selected monitor set and random seeding is carried out periodically, primarily to prevent a breakdown of the model assumptions - such as deviations from the mobility model, variations in CPE density etc. Random seeding allows the system to limit the number of times a CPE becomes a monitor, and also allows the system to control the epochs in which a CPE can become a monitor.

The seeding is possible in a number of ways - a distributed approach, or a centralized approach. The centralized approach would be one where the server distributes seeds to all registered CPEs. A distributed methodology - where monitors distribute seeds - would put an excessive amount of trust in the monitor's behavior. Anomaly detection methods using the CPE graph may be used to check whether a monitor is distributing seeds properly - but in the opinion of the author, the overhead of such a method would probably exceed the overhead of a centralized approach.

With a centralized approach, seeds are simply associated with a CPE at the Allocation Server , and seeding is local to a administrative unit which can be considered self-contained - a city, for example. Seeds are distributed uniformly randomly over the set of registered CPEs, and thus areas with larger CPE densities will contain more seeded CPEs. We make an initial estimate of the fraction of seeded CPEs which will apply to become monitors (f_s), and update our estimate of f_s using real-time monitor information.

Registration

The security of the registration process is also important, since multiple registrations for the same CPE must be prevented. Simultaneously, the registration process must also be anonymous in nature, such that previous traces submitted by this CPE cannot be associated with it.

Trace submission

The system must be able to check the authenticity of traces submitted by a monitor. For this to be possible, every trace must contain proof that the monitor was actually present in the locations mentioned in the trace. This proof can be in the form of tokens from trusted CPEs or BSs, etc. Any such method must not reveal the presence of the monitor to other CPEs or BSs. The proof mechanism is an assumption in our network model.

Location verification may not be necessary if sufficient redundancy of monitors is provided.

Location Verification

Location Verification is currently an open problem, with one approach being the use of fixed or trusted infrastructure as an anchoring mechanism[27]. In such mechanisms, fixed or trusted infrastructure initiates a verification cycle by transmitting beacons which contain information or tokens verifiable by a third party. The main premise is that the only nodes present in the proximity of a fixed node will receive the information/token. If implemented in the context of mobile nodes, such a system would require the trust associated with particular nodes to be high, so that no false beacons are transmitted in the network. Additional complexity is added when nodes are mobile, since their positions would be ambiguous and tokens from one part of the network may be introduced into another.

The nature of such mechanisms alludes to a random geometric graph model to study and engineer the transmission of beacons. One way to control the introduction of beacons from one part of the network into another could be forcing the creation of a graph which has *disconnected* components. If this is the case, then only nodes located in one part of the network can be associated with each other.

In the state-of-the-art, the notion of a verifiable location has been implemented only relative to trusted reports of location. Any location verification mechanism is thus vulnerable to an attack where all relevant values are transferred between colluding users. In general, real-time schemes would be vulnerable to variants of such a “cloning” attack. Thus, a better scheme would be to use a trust metric for each CPE, computed on the basis of the quality of traces submitted by the CPE in the past.

Trace Quality

Redundancy across traces can be used to quantify the quality of traces submitted by a CPE. For instance, if a majority of datapoints in a trace agree with datapoints for those location from other traces, then the quality of the trace is high.

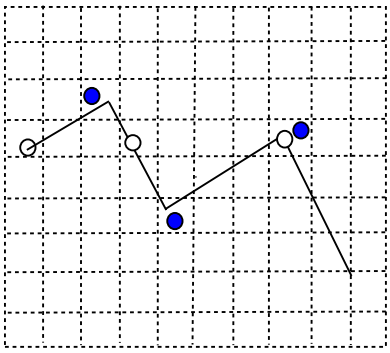


Figure 4.7: Quantifying Trace Quality
As a SU (white node) move through the Tessellation , monitors can record its identity. They submit all recorded identifiers when submitting traces to the Allocation Server

Another method for quantifying the "quality" of a trace would compute the intersection of the device identifier sets observed by a monitor when the trace was being collected. A majority mechanism where each device identifier reported for a Tessellation is considered "authentic" only if a majority of monitors in the area report it. Tests like the WSPRT can be used to wieght each trace, and the wieght computed would be a measure of the quality of the trace. More advanced learning algorithms which treat device identifiers as features, while compensating for incorrect data could also be used.

If the space of device identifiers is large enough, the likelihood of naturally occuring fake identifier reports can be made small. Additionally, monitors which send false reports run the risk of being proved dishonest, since they cannot estimate which other monitors are dishonest. In such situations, a monitor may attempt to learn a belief on the number of dishonest monitors, but with a large device identifier space, such scenarios can be dealt with. The algorithm used by honest monitors to collect device identifiers can be created such that geometric random graph formed favors device identifiers which have already been reported can be done.

Uniqueness of Datapoints and Traces

It is important that monitors be prevented from submitting the same datapoints or traces repeatedly, since this would effect the anomaly detection process which must be performed once SysMet collects traces. This would require that traces be associated with the monitor which submits it, so that it does not submit multiple traces for the same epoch. Also, the registration process must prevent multiple registrations by the same CPE.

4.4.2 Privacy Aspects

The privacy the primary user and secondary users has been noted as an important security issue for CR devices in the state-of-the-art [33]. Inference attacks on database which drive spectrum access can reveal information about the activities of PUs, while inference attacks on the querying patterns can reveal information about the activities of SUs.

For SUs and PUs, BSs or CPEs, given some background information about CPEs - such as visit probability matrix for each user, or the transition probability matrix for each user - Inference attacks can be categorized as follows [44] :

- De-anonymization - Given anonymized traces for users, finding the most probable mapping between mobility traces and pseudonyms associated with users
- Localization - Given anonymized, obfuscated or incomplete traces from users, finding the most probable location for a user at a given time instant
- Tracking - Given anonymized, obfuscated or incomplete traces for users, find the most probable trace for a user

These inference attacks may be carried out by Bayesian inference methods, as presented in [44].

Another issue would be the privacy of location of CR devices in general. If techniques such as PHY layer identifiers are implemented, they must be implemented in a manner which does not allow unauthorized tracking of CR devices. In much the same way as IPv6, CR device identifiers would uniquely identify a device, but pose a bigger threat to privacy since both the high granularity location and the identity of the device are available. Thus a first measure would be use a protected, dynamic identifier which is only decipherable by those with the required key.

BS Privacy

The privacy of BS's activities must not be violated by the monitoring system. One of the ways of achieving this would be an adaptive monitoring scheme for areas where misuse is detected. The system would also be able to maintain a low probability of observation in neutral areas, in order to preserve the privacy of honest BSs.

Another way of preserving BS privacy would be use an architecture where the responsibility of correcting misuse lies with a organized SU, rather than an individual. In such a situation, it would not be necessary to reveal specific information about an individual BS's identity - only meta-information, along with a generic pseudonym for the organized SU needs to be revealed.

CPE Privacy

When a CPE submits spectrum activity information collected during its time as a monitor, the location information obtained from the trace may reveal information about the CPE's whereabouts. Currently, SysMet uses a collection probability $p_{collect}$, which can be kept low such that only coarse location information about a monitor is revealed. Under location inference attacks described previously, the number of data-points within a trace, as well as the number of traces submitted by a monitor directly effect the quality of the inference made. If a trace contains few and well-separated data-points, then a tracking attack would be made difficult. A low collection probability $p_{collect}$, would cause a monitor to record a low number of datapoints as part of each trace. A completely randomized method of trace collection would report more data-points for locations where the CPE spends more time - and may be vulnerable to the a localization attack. Thus, a trace collection method with a random initial offset, followed by periodic sampling would be best to maintain CPE privacy within a Tessellation .

Another set of countermeasures would be the use of anonymization methods, such as the use of pseudonyms. This would make inference attacks against a particular CPE difficult. Location obfuscation could also be used, provided that it does not degrade the usefulness of the traces collected by the CPE significantly. For a CPE, Anonymity during Trace Submission can be provided in the following forms:

- Partial Anonymity - Traces submitted by a monitor cannot be associated with it
- Complete Anonymity - A new registration during a new cycle cannot be associated with a registration in a previous cycle

With Partial anonymity, traces submitted by a monitor cannot be associated with it. A trust metric for each trace would be computed on the basis of location verification information and cross-verification using redundancy across traces. Other mechanisms for this would involve the use of pseudonyms for k-anonymity - pseudonyms for which a CPE can verify the anonymity that is offered. Computing a trust metric on a class of pseudonyms could allow a trust based mechanism for awarding monitor status. Complete anonymity would make any attribute based attacks which attempt to associate traces with monitors impossible.

Using anonymity would have disadvantages which would be difficult to mitigate. For instance, anonymity would make the system vulnerable to repeated submissions of traces. Also, anonymity comes at the cost of the inability to hold a monitor accountable for traces submitted by it, or reward it on the basis of the quality of a trace. No notion of trust can be associated with a particular CPE because any evaluation of the quality of a trace or location information cannot be persistently associated with a given monitor. In general, redundancy in the monitor network could reduce the need for a trust system, if we can assume that a certain number of monitors will be truthful.

For CPE privacy as a CPE moves from monitor status in one Tessellation into monitor status in another Tessellation - if anonymity of the CPE is maintained, then trace reporting poses minimal risk to CPE privacy. In general, a first measure of privacy would be to ensure that policies themselves do not require sensitive information to operate. For example, policies could require certain power level for all transmitters, but mandating a certain power level depending upon the data being transmitted may violate the transmitter's privacy by revealing the contents of the transmission.

The privacy concerns for non-monitor CPEs are similar to those of a BS, and would be dealt with using similar mechanisms. An additional concern would be tracking attacks on CPE device identifiers recorded during a trace. This could be checked by using slow-changing protected device identifiers, which can only be decrypted by authorized persons.

Conclusion

In a system such as SysMet - where CPEs report traces which contain location information - CPE and BS Privacy become important concerns. Traces submitted by monitors must be made robust to inference attacks, and a low $p_{collect}$ can be used so that each trace leaks only coarse location information. If CPEs agree to a certain extent of risk to their privacy in exchange for rewards, then the risk can be minimized by the incentive scheme, as discussed in Chapter 3.

The privacy risk of a BS and non-monitor CPEs can be reduced using a low $p_{collect}$. An adaptive monitoring scheme can also be used, which would increase the data-points reported for an area where misuse was detected, while maintaining a low reporting probability in neutral areas.

We note that many of the security and privacy issues represent a trade-off between cost and redundancy across traces, and this trade-off can be explored to alleviate some of these issues. Redundancy can be used to compute the trust which can be associated with a given trace. It increases system cost, but provides advantages in terms of trace cross-verification and a reduction in the number of data-points and traces that monitors must report. The idea of redundancy can be incorporated easily into SysMet , though the specifics of implementation are left as future work.

Anonymity within the registration or the trace submission process could also be used to preserve CPE Privacy - but preventing repeated submissions would be difficult, and no trust metric for CPEs would be possible. Redundancy becomes important when anonymity is brought into the picture, since a strong notion of trust becomes difficult to provide.

4.5 Conclusion

In conclusion, we re-state some of the important assumptions made during this formulation, followed by a brief summary of the formulation and system design.

The important assumptions made during the problem formulation and system design were :

- It is assumed that the system has complete knowledge of the strategies, payoffs, etc. of a BS that intends to misuse spectrum.
- The system has an imperative to prevent misuse - it has the resources to meet the cost of monitors

A Game Theoretic formulation models the decision that the BS must make - to misuse spectrum or not. An observation probability P_{caught} obtained from this formulation is required to dissuade spectrum misuse. The system design for SysMet imposes this observation probability by soliciting for the role of a "monitor". CPEs who wish to be monitors must register with the Allocation Server , and subsequently request monitor status. SysMet uses a "seeding" procedure to ensure that the monitor selection mechanism is fair. A given collection probability $p_{collect}$ requires a monitor to expend only a $p_{collect}$ fraction of its resources, while reducing its privacy risk.

A number of security and privacy aspects are analyzed for this system design. Apart from k-anonymity, other methods such as a trust metric based on redundancy, adaptive observation probability or adaptive traces, etc. should be used. Redundancy can be used in a number of ways - for location verification and trace cross-verification, among others.

The system has an imperative to catch misuse, and pays no penalty if it observes an instance when misuse is not occurring. Also, there is no payoff to the BS in that case, and the game remains a zero-sum game. The minimization of cost occurs through the optimal monitor allocation procedure. This shifts the cost of observation to the scale of a whole trace, rather than every data-point in the trace. It also removes the requirement that a belief on the strategy of the BS be maintained. Intuitively, a privacy cost $C_{privacy}$ may be considered for each observed instance where no misuse is occurring. This will change the required P_{caught}, and can be compensated for by a change in the reward if misuse is detected.

Chapter 5

Optimal Monitor Allocation

Given a required Observation probability P_{caught} for each cell in a Tessellation and a collection probability $p_{collect}$, and given which cells monitors can be placed in, we would like to find the optimal allocation of monitors, such that the given Observation probabilities are imposed. We would like to use the smallest number of monitors possible, in order to minimize the cost incurred due to monitors, as well as minimize the number of CPEs whose privacy is risked.

In order to do this, we formulate an Integer Linear Program which minimizes the number of monitors required in the Tessellation . The formulation contains one constraint for each cell in the Tessellation , constraining the probability of observation for each cell to be above the required P_{caught}. In order to realize the probability of observation imposed by a monitor on any cell in the Tessellation , we use an estimate of the Visit Probability Matrix P, which represents the probability of visiting a particular cell in the Tessellation . We assume the existence of such probabilities, but the procedure for obtaining this matrix would be specific to the model of human mobility used - the decision of the model of mobility would rely on what properties we require from the model.

Upon solving this formulation, we obtain a monitor density for each cell in the Tessellation . If the visit probability matrix P truly exists and does not change with time, then over sufficient number of monitor epochs, the given P_{caught}s will be imposed in each cell. The subtler issues of the visit probability matrix, and how to obtain it, are discussed in the sections that follow. One must note that, if such an optimization procedure is not used, then there is no guarantee as to how what Observation probabilities will be imposed over time.

This chapter is organized as follows - Section 5.1 presents the specifics of the problem formulation, while Section 5.2 presents the solution technique. Section 5.3 presents the results of a simulated solution of the monitor allocation problem, along with other experiments to evaluate the quality of the solution obtained.

5.1 Problem Formulation

Visit Problem Formulation

Let us consider a cell (i,j) in a given tessellated area. For a monitor which begins in (i,j), let $X_{k,(i',j')}$ be a vector of length equal to T. Let $X_{k,(i',j')}(t) = 1$ if the monitor k is present in cell (i,j) at time t. For example, if $X_{k,(i',j')}(1) = 1$, then monitor k was in cell (i,j) at that time t = 1. The expectation $E(X_{k,(i,j)}(1))$ can be interpreted as the probability that the monitor k is at cell (i,j) at t = 1.

$X_{k,(i',j')}$ is a random vector drawn from the distribution H_k over all possible vectors with binary elements of length T.

Thus, if cell (i,j) needs to be observed P_{caught} of the time, then

$$\sum_{t=0}^{T} \vee_{k=1}^{N_m} X_{k,(i',j')}(t) \geq P_{caught} * T$$

We assume that the distribution $P(X_{k,(i',j')} = x)$ makes it unlikely that $X_{k_1,(i,j)}(t) = X_{k_2,(i,j)}(t) = ..X_{k_m,(i',j')}(t) = 1$. That is, it is unlikely that m monitors are present in the same cell at the same time, where m is a small integer. Due to TLW mobility and the large number of cells in the tessellation, we can also assume that only a small number of elements in any $X_{k,(i',j')}$ are 1. We also assume that T and N_m are large.

We can impose 2 kinds of conditions on this sum - either that it achieves the required P_{caught} in expectation, or that it achieves P_{caught} with some confidence interval. Intuitively, it can be said that the latter condition is stronger, and would cause the expectation to be much larger than $P_{caught} * T$. But since the distribution of this sum is difficult to write in closed form, we take the former approach. If a monitor collects observations with a probability $p_{collect}$, then the expectation on the sum for each $X_{k,(i',j')}$ becomes a p_{report} fraction of its current value.

Thus, in expectation,

$$\sum_{k=1}^{N_m} E(p_{collect} * \frac{\sum_{t=0}^{T} X_{k,(i',j')}(t)}{T}) \geq P_{caught}$$

Let Y be a random variable, where

$$Y_{k,(i,j)} = \frac{p_{collect} * \sum_{t=0}^{T} X_{k,(i',j')}(t)}{T}$$

By the law of large numbers, the mean of samples from the distribution of Y will converge to a value greater than P_{caught} if the condition on the expectation is met. Thus, we can say

Variable	Description
D_i	ith dimension of area under consideration
T	Length of a monitor epoch
N_m	Number of monitors required
d_i	ith dimension of cell
a	area of cell
$r_m^{(i,j)}$	The number of requests to award monitor status received for cell (i,j)
$\lambda_m^{(i,j)}$	density per unit area of monitors in a cell (i,j)
$P_{k,t}$	Occupation probability matrix for kth CPE at time t
P_k	Visit probability matrix for kth CPE
$p_{(i',j')}^{(i,j),k}$	Probability of kth CPE, which started in cell (i,j), to occupy cell (i',j')
$p_{collect}$	Probability that a given CPE collects an observable instance of spectrum activity

Table 5.1: Problem Formulation: Table of Symbols

that over time, the average number of observed time instances per epoch for each cell will go to a value greater than P_{caught}. It should be noted that this does not prevent the observed instance from a particular epoch to be less than required.

Thus, for each cell, we would like to impose the following condition in each epoch :

$$\sum_{k=1}^{N_m} E(Y_{k,(i',j')}) \geq P_{caught}$$

We call $E(Y_{k,(i',j')})$ the probability of visiting a cell (i',j') for a monitor k starting in a cell (i,j). This condition states that the cumulative visit probability imposed by each monitor in the area should meet the required observation probability.

ILP Formulation for densities

We consider an area $D_1 * D_2$, divided by a Tessellation , with each cell of dimensions $d_1 * d_2$, and area a. In each cell (i,j) - let monitors be present with a density $\lambda_m^{(i,j)}$ and CPEs be present with a density $\lambda_{CPE}^{(i,j)}$. Let the mobility of a CPE u_i be defined in terms of its visit probability matrix P_k, where each element represents the visit probability of cell (i',j'), for an CPE k from cell (i,j).

The visit probability for the cell (i',j') is defined as

$$p_{C,(i',j')}^{(i,j),k} = \sum_{t=0}^{T} p_{(i',j'),t}^{(i,j),k}$$

$$P_k = \begin{pmatrix} p_{(1,1)}^{(i,j),k} & p_{(1,2)}^{(i,j),k} & \cdots & p_{(1,D_2/d_2)}^{(i,j),k} \\ p_{(2,1)}^{(i,j),k} & p_{(2,2)}^{(i,j),k} & \cdots & p_{(2,D_2/d_2)}^{(i,j),k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{(D_1/d_1,1)}^{(i,j),k} & p_{(D_1/d_1,2)}^{(i,j),k} & \cdots & p_{(D_1/d_1,D_2/d_2)}^{(i,j),k} \end{pmatrix}$$

Now, $\lambda_m^{(i,j)}$ must be such that the system can guarantee a given $P_{caught,(i',j')}$, at each cell (i',j'), while minimizing the total number of monitors used.

We can thus define the cumulative cell visit probability as,

$$p_{C,(i',j')}^{(i,j),k} = \sum_{(i,j) \in grid} a * \lambda_m^{(i,j)} * p_{(i',j')}^{(i,j),k}$$

To simplify out LP formulation, we define :

$$P = \begin{pmatrix} p_{(1,1)}^{(1,1)} & p_{(1,2)}^{(1,1),k} & \cdots & p_{(2,1)}^{(1,1),k} & \cdots & p_{(1,n_2)}^{(1,1)} \\ p_{(1,1)}^{(1,2),k} & p_{(1,2)}^{(1,2),k} & \cdots & p_{(2,1)}^{(1,2),k} & \cdots & p_{(2,n_2)}^{(i,j),k} \\ \vdots & \vdots & \ddots & \vdots & & \\ p_{(1,1)}^{(n_1,n_2)} & p_{(1,2)}^{(n_1,n_2)} & \cdots & p_{(2,1)}^{(n_1,n_2)} & \cdots & p_{(n_1,n_2)}^{(n_1,n_2)} \end{pmatrix}$$

$$\lambda = \begin{pmatrix} \lambda^{(1,1)} & \lambda^{(1,2)} & \cdots & \lambda^{(2,1)} & \cdots & \lambda^{(n_1,n_2)} \end{pmatrix}$$

$$n_1 = \lceil D_1/d_1 \rceil; n_2 = \lceil D_2/d_2 \rceil$$

Thus, the ILP 1, as shown below, must be solved.

$$Minimize\ N_m$$

subject to

$$\sum_{(i,j) \in grid} a * \lambda_m^{(i,j)} = N_m$$

$$p_{collect} * p_{C,(i',j')}^{(i,j),k} \geq P_{caught,(i',j')} \forall (i',j') \in grid$$

$$a * \lambda_m^{(i,j)} \in \{0, 1, 2, ..., r_m^{(i,j)}\} \forall (i,j) \in grid$$

In matrix notation,

$$Minimize\ N_m$$

subject to

$$a * \lambda * \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = N_m$$

$$p_{collect} * \lambda * P \geq \begin{pmatrix} P_{caught,(1,1)} \\ \vdots \\ P_{caught,(n_1,n_2)} \end{pmatrix} \forall (i', j') \in grid$$

$$\{0\}^{n1*n2} \leq a * \lambda \leq \{r_m^{(1,1)}, r_m^{(1,2)} ..., r_m^{(2,1)} .. r_m^{(n1,n2)}\}$$

We relax the ILP above to an LP by allowing all variables be a real numbers, rather than integers. We elaborate on the solution approach in Section 5.2.

5.1.1 Variations on ILP 1

Over and above the ILP formulation stated previously, certain additional constraints may be added to the formulation to include specifics of the given scenario. For example, the P_{caught} may be equal for each cell . Another example would be one where some λ_m or λ_{CPE} are constrained to a zero value, so that no monitors are selected from that particular cell .

5.1.2 Estimating the Visit Probability Matrix

The nature of the visit probability matrix P is specific to the model of mobility used. The model may use no information about the destinations within the Tessellation , or may use some partial information about the movements of CPEs or destinations inside the Tessellation . Depending upon the model used, P may be obtained in a number of ways - analytically, experimentally or using learning methods. Rhee et al [30] develop an analysis of the TLW mobility model to obtain the occupation probability, and hence the critical delay for a network of nodes moving under the TLW model. This analysis is asymptotic, but a similar analysis may be conducted to obtain the occupation Probability in a form suitable for our analysis.

Another method for estimating P would be using Monte-Carlo simulations. Monte-Carlo simulations can be used to sample from probability distributions through repeated simulation, in situations where a closed-form solution is difficult to obtain. These samples can then be used to obtain an estimate of the distribution.

P may also be obtained by an inference method, which takes traces collected by Monitors to improve the estimate of P. Let the motion of each CPE be described by a Markov chain, where each state represents a cell in the Tessellation . For a SU u_i, some background information may be known in the form of an estimate of the steady state probability matrix $\hat{\pi}$ for its Markov chain, or an estimate of the transition probability matrix $\hat{P}$. A Markov chain model for mobility may not be strictly correct, since human mobility is not memoryless. Also, using P obtained from an inference method could lead a monitor allocation which re-enforces the current beliefs about transition probabilities, leading to unexpected behavior.

5.2 Integer Linear Programs and Sequential Fixing (SF)

Having formulated ILP 1 as given in the previous section, a solution must now be obtained for the ILP. An ILP can be solved using many techniques as described in brief previously. The basic approach is to utilize the a relaxed version of the ILP to obtain an optimal, possibly infeasible solution. One may then obtain a feasible solution by iteratively adding constraints which would drive the solution towards an integer solution.

5.2.1 Sequential Fixing Methods

For the particular experiment we conducted, it was found that the large number of monitor density variables caused progress to be slow. This was due to the fact that each density variable contributed equally to the objective function. When a new constraint is added, it simply causes some unconstrained variable to assume the value which was eliminated by the new constraint. Thus a large number of constraints were required to make progress towards an integer solution.

Due to these facts, provably competitive methods became impractical for this ILP. Thus a variant of the Sequential Fixing (SF) method for binary integer variables was used to solve the ILP. Though this is a heuristic method, it found a solution which was close to the LP relaxation solution. In fact, the solution found by the SF method was 3 times as close as the baseline (a simple rounded up solution). We describe the Sequential Fixing Algorithm used in Algorithm 1, and the symbols in the LP description shown below. In the Algorithm 1, the largest variable in LP relaxation solution is fixed. This variable's value is first rounded down, and a new LP with this fixed value is solved. If this LP is not feasible, then the same value is rounded up, and the LP resolved. It should be noted that since there at-least one of these LPs will be feasible, since the rounded up value can never exceed the upper bound, and thus will always give a feasible solution.

$$Minimize\ f * x$$

subject to

$$A * x \leq b$$

$$Aeq * x = beq$$

$$lb \leq x \leq ub$$

Where x is a n * 1 vector

Algorithm 1 Sequential Fixing

```
1: procedure SF(f, A, b, Aeq, beq, n)                                   ▷
2:     nd ← zeros(n)
3:     x ← Solve_LP(f, A, b, Aeq, beq)
4:     objc ← f * rounded_up(x)
5:     xc ← x
6:     while Zeros remaining in nd do                                   ▷
7:         i ← Find_Largest_in(xc, nd)
8:         nd[i] ← 1
9:         Aeq, beq ← Add constraint x_i = ⌊xc[i]⌋
10:        x ← Solve_LP(f, A, b, Aeq, beq)
11:        if then LPfeasible
12:            xc ← x
13:            objc ← f * x
14:        else
15:            Aeq, beq ← Remove constraint x_i = ⌊xc[i]⌋
16:            Aeq, beq ← Add constraint x_i = ⌈xc[i]⌉
17:            x ← Solve_LP(f, A, b, Aeq, beq)
18:            xc ← x
19:            objc ← f * x
20:        end if
21:    end while
22:    return xc
23: end procedure
```

5.2.2 Solution Fairness

Even though the solution obtained from ILP 1 using Sequential Fixing may give good results in terms of the number of monitors, the SF method does not eliminate any inherent unfairness that may have been introduced by LP relaxation. The solution may be unfair in many ways - unfairness in monitor density, unfairness in the distribution of observed instances, among others. It should be noted that due to the properties of TLW mobility, fairness in monitor density would imply a certain degree of fairness in distribution of observed instances. Low initial monitor density, combined with uniformly randomly chosen direction would bring fairness to the distribution of observed instances. Additionally, a high P_{caught} would make unfairness in observed instances irrelevant, since a large fraction of instances will be observed.

In this section, we consider fairness in monitor densities. The solution obtained may add a high density of monitors to one cell, while keeping monitor density in other cells low. Though this unfairness was not found in the experiments currently performed, it is possible that it becomes prominent if requests for awarding monitor status are only received from a small number of cells.

This problem may be solved by using a Lexicographic Max Min method to obtain monitor densities, once an optimal number of monitors is obtained from the solution of ILP 1. In this procedure, the smallest Observation probability is maximized as shown below. This is done repeatedly until the solution no longer changes.

$$Maximize\ p_{min}$$

under

$$\sum_{(i,j)\in grid} a * \lambda_m^{(i,j)} = N_{m,optimal}$$

$$p_{collect} * p_{C,(i',j'),t}^{(i,j),k} \geq p_{min} \forall t \in (0,T], and (i',j') \in grid$$

$$p_{min} \geq P_{caught}$$

The solution and experimental evaluation of this formulation is beyond the scope of this work, though a iterative sequential fixing method would suffice.

5.3 Empirical Study & Evaluation

The empirical study of SysMet was conducted with the aim of evaluating the quality of the solution, the practicality of the method, and to identify avenues which would allow the relaxation of some the assumptions made during system design. The empirical study of SysMet consisted of many parts :

- *Obtaining P empirically* - There are a number of ways to estimate the visit probability matrix P. To estimate the matrix P experimentally, Monte Carlo simulations using the TLW mobility model were used to obtain each element of the matrix.
- *Evaluation of Solution* - The $N_{m,optimal}$ obtained was compared to a baseline found by rounding up each value in the LP relaxation solution. The monitor densities obtained and the resulting Observation probabilities were also compared to a uniform random distribution of $N_{m,optimal}$ monitors.
- *Evaluation through Simulation* - The motion of monitors and Observation probabilities were studied through simulation, with monitors distributed according to the monitor densities obtained using ILP 1 with sequential fixing. The redundancy generated by the motion of monitors was also studied.

In order to obtain an estimate of the visit probability matrix P, 1000 simulated traces using the Truncated Levy Walk model were generated, each trace lasting a length of 3 hours, at the granularity of one report every minute. The matrix P was obtained for an square area

Table 5.2: Empirical Study: Values used

Variable	Value	Description
D_1	800m	1st dimension of area under consideration
D_2	800m	2nd dimension of area under consideration
T	180	Length of a monitor epoch
d_1	40m	1st dimension of cell
d_2	40m	2nd dimension of cell
P_{caught}	0.7	Observation Probability for all cells
$p_{collect}$	1	Probability that a given CPE collects an observable instance of spectrum activity
α	0.6	Levy Distribution Parameter for flight length
β	0.6	Levy Distribution Parameter for pause time

800*800 meters in dimensions, tessellated with cell sizes of 40* 40 meters. An assumption of a toroidal simulation area was made, which removed edge effects by allowing nodes to drift into one edge, and drift out the opposite edge. Both the α and the β parameters were set to 0.6, which are the lowest values found for human mobility [42]. This makes the diffusivity for these traces worse than the diffusivity of Brownian motion. In-spite of this, the ILP 1 was able to find a good set of monitor densities. The Sequential Fixing method for solving the ILP was applied to obtain $\lambda^{(i,j)}_{m,optimal}$s and $N_{m,optimal}$.

The ILP 1 was then solved, assuming that requests for awarding monitor status have been received from every cell in the Tessellation . The solution obtained from ILP 1 was evaluated by comparing it to the LP relaxation solution, and several scenario-specific baseline solutions.

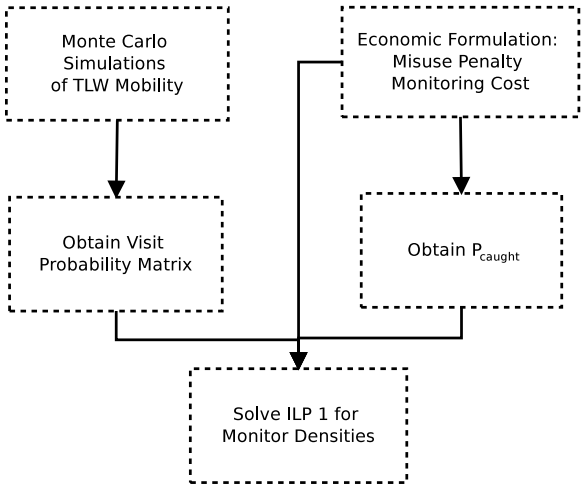


Figure 5.1: Empirical Study : Work-flow

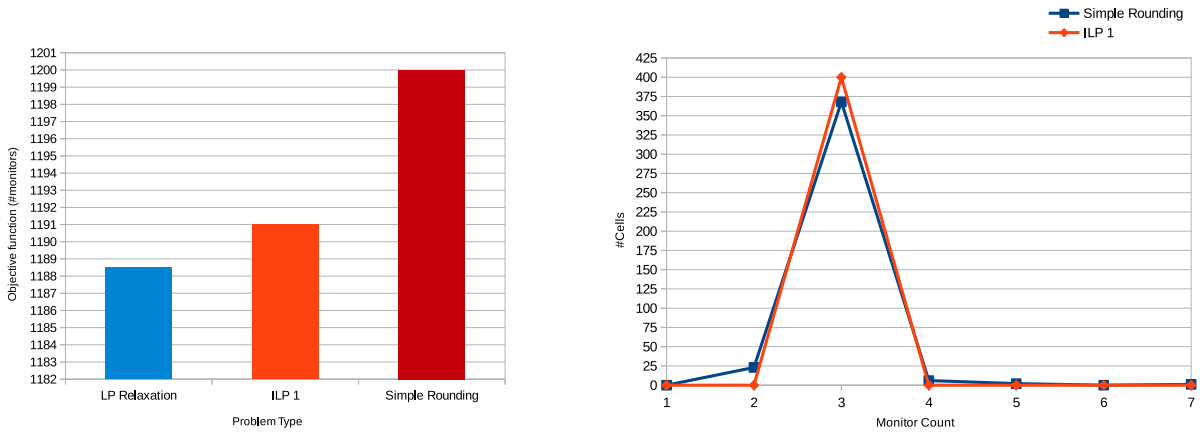
There are several points which must be kept in mind while attempting to generate the visit probability matrix using Monte Carlo simulations. For the purposes of the Monty Carlo

simulations, it was noted that Tessellation dimensions and trace lengths should be set such that a true picture of the visit probability matrix is obtained. If the trace lengths are too large or if each cell is large, then due to the boundary conditions each simulated node would visit a majority of the Tessellation . This causes the visit probability matrix to contain values which are more or less equal. This may also occur of the number of simulated nodes is not large enough.

5.3.1 Experiment 1: Monitor Counts

Experiment 1 was conducted to evaluate the quality of the solution obtained for ILP 1 using Sequential Fixing. For Experiment 1, the LP relaxation and ILP 1 were solved independently. The optimal number of monitors obtained from ILP 1 was compared to the number of monitors obtained from the LP Relaxation solution, and from a simple rounding up of the LP relaxation solution. The solution from ILP 1 was found to be much closer to the LP relaxation solution, as shown in Figure 5.2a - while the LP relaxation required 1188.5 monitors, ILP 1 required 1191 monitors, and the simple rounding solution required 1200.

Figure 5.2b shows a plot of the number of cells versus the number of monitors in the cell. This figure shows that the solution obtained from ILP 1 gives a better solution by rounding down values in the LP relaxation solution, while increasing the number of monitors in a small number of cells to compensate. This can be concluded from the count of cells which required 3 monitors - the simple rounding solution has 400 such cells, while the solution from ILP 1 has 375.



(a) (b)

Figure 5.2: Experiment 1: Monitor Counts and Cell Counts
(a)Monitors Required:SysMet vs Baselines - The number of monitors required by the solution of ILP 1, as obtained using Sequential Fixing is shown. It is compared to the monitor requirements shown for the LP relaxation Solution, and the Simple Rounding Solution (b)Cell Counts for each Monitor Count - The number of cells which required a given monitor count are shown of the Sequential Fixing solution, the LP relaxation solution and the Simple Rounding Solution

5.3.2 Experiment 2: Monitor Densities & Fairness

Experiment 2 was carried out to evaluate the value of the monitor configuration obtained from solving ILP 1 using the Sequential Fixing technique. We compared the monitor densities obtained from ILP 1 to a baseline wherein $N_{m,optimal}$ monitors are distributed uniformly randomly.

To evaluate the Observation probabilities imposed by different configurations, the Observation probabilities imposed by optimal number of monitors were compared, when densities were set by either the LP relaxation solution, the solution of ILP 1 or a uniform random distribution. It was found that the majority of densities obtained using uniform random distribution matched those obtained by the ILP 1 with Sequential Fixing.

It can be concluded that the SF solution is fair in its distribution of monitors, closely matching a uniform random distribution of the optimal number of monitors. We can conclude that the natural uniform structure of CPE populations in a small area is well suited to random seeding and a request-based monitor allocation mechanism, but a few cells require a slightly larger number of monitors. This implies that the task of the Allocation Server would be simple - it would usually have to grant a uniform fraction of requests, and grant more requests for CPEs from a small number of cells.

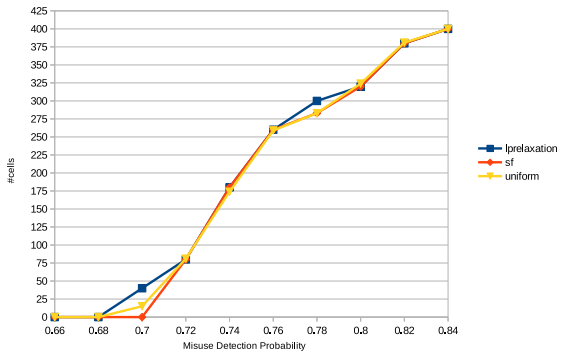
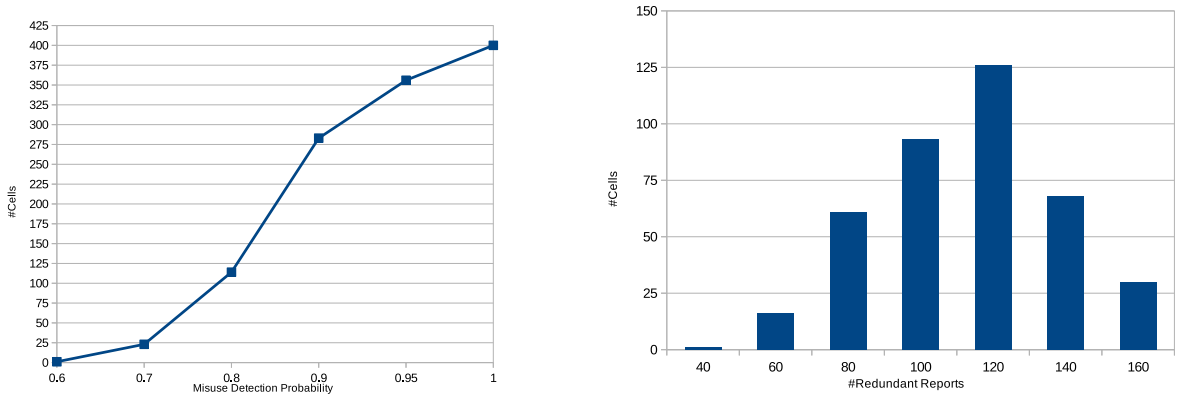


Figure 5.3: Experiment 2: Variations in Observation Probabilities
Variations in Observation probabilities between a uniform random distribution, distribution as given by the LP relaxation solution, and distribution as given by the Sequential Fixing (SF) solution

5.3.3 Experiment 3: Observation Probability & Redundancy

Experiment 3 is targeted at evaluating the solution presented using simulations of CPE mobility. The monitor densities obtained from the SF solution of ILP 1 were used to place monitors in a Tessellation of 800 * 800 meters, with each cell of dimensions 40 * 40 meters. The Observation probabilities imposed were computed by counting the number of time-slots in which a monitor was present in a cell . The simulation showed that 24 cells had



(a) (b)

Figure 5.4: Experiment 3: Observation Probability & Redundant Reports
(a)Observation Probability - The number of cells which have a Observation probability of less than the x-axis label (b) Redundant Datapoints in a trace - The number of cells which had redundant datapoints

Observation probability between 0.6 and 0.7, and were less than the required probability of 0.7.

To evaluate the extent of redundancy - that is, the instances where more than 1 monitor was present in a cell - the number of time-slots in which multiple monitors were present were counted for each cell. With 180 time-slots in the simulation (equivalent to 180 minutes), It was found that over 126 cells of 400 had between 100-120 time-slots in which multiple monitors were present in the cell. This result shows that redundancy can be actively utilized to enable the detection of dishonest CPEs.

5.4 Conclusion

In conclusion, we re-state the main assumptions made under the optimal monitor problem formulation, followed by a brief summary of the solution and evaluation :

- A monitor remains within the given tessellated area for the length of the monitor epoch T.
- Each monitor operates under the TLW mobility model. Due to heavy tailed nature of the pause time distribution, only a small fraction of monitors could wait for a significant amount of time at a given cell.
- Due to the uniform random selection of direction, only a small fraction of monitors can be present in a cell at the same time.

The mobility of CPEs/monitors was first analyzed to obtain a condition on the cumulative visit probability which would impose a given observation probability. These conditions were then used as part of an Integer Linear Program, in order to obtain the monitor densities required. The ILP was solved using a Sequential Fixing method, and the solution was found to be impose the required observation probability for more than 95% of cells in the tessellation.

It was also found that the densities obtained were fair, but an unfair allocation is not precluded. A supplementary ILP may be formulated which uses a Max-Min fairness criteria to obtain monitor densities, but the solution and evaluation of this problem is beyond the scope of this work. Unfairness in the distribution of observed instances may also occur, but the mobility model implies that a solution which is fair in terms of monitor densities would also be fair in the distribution of observed instances over time. Additionally, a high P_{caught} makes this kind of unfairness irrelevant.

Chapter 6

Conclusions

Misuse detection and monitoring spectrum activities is an important tool to ensure disciplined CR device behavior. We consider the specific case of enforcing mechanisms for disciplined BS behavior. SysMet uses a network of CPEs to monitor spectrum activities by reporting traces of activity. These CPEs, called monitors are chosen such that a minimum number of monitors are required to enforce a required probability of catching BSs which may be cheating.

Given the economic formulation, SysMet computes a probability of Observation P_{caught} which must be imposed, in order to dissuade a rational cheater. It allows CPEs to become monitors, in return for rewards. During a contention interval, CPEs can request to be awarded monitors status, by contacting the Allocation Server and providing their current location. The Allocation Server uses the requests received during the contention interval to compute the number of monitors required to impose P_{caught} for each cell, by solving an ILP to minimize the number of monitors required.

The network model used as the basis for SysMet makes certain assumptions which are still open questions in the research community. The ability to detect spectrum policy violations or misuse, and the ability of CPEs to provide location proofs are the two bold assumptions. These are justified by the fact many works in the state-of-the-art propose mechanisms to identify and conclusively detect policy violations.

A analysis of the security of SysMet revealed that CPE trust, location verification and privacy are important issues which must be dealt with. The issues of privacy and trust must be traded off appropriately, in order to maintain CPE privacy while allowing a metric of the veracity of traces submitted by monitors. Redundancy in the Monitor network can be used to cross-verify traces reported by monitors, as well as verify location information submitted with a trace.

Chapter 7

Directions for Future Work

Having explored the basic premise of SysMet , many avenues for future work can be identified. These include methods to implement assumed functionality, as well as design additional extensions to this system.

7.1 Limitations

The limitations of the SysMet can be considered for each of the following aspects of the current design :

- Scope of Observation
- Game Formulation
- ILP formulation
- Trace collection
- CPE trust and incentives
- BS and CPE privacy

We have currently considered monitoring on a single band, and the analysis must be extended to include monitoring on multiple bands. If monitors are dedicated to certain bands or if a randomized algorithm are used to monitor multiple bands, then simply adding additional constraints to impose a P_{caught} for each band at each location would suffice. The complexity of the problem would increase if cooperative or coordinated monitoring is required.

The game formulated in Chapter 4 assumes that the system is aware of the rewards of misuse. This also implies that the system is aware of every possible misuse strategy - that is, it is

aware of every means of misuse a cheater may use. These assumptions are reasonable when checking for violations of a known policy set, but are limited when it comes to accounting for new or unknown methods of misuse. In the scenario considered, the rewards of rewards can be quantified in terms of the cost of use or access.

The ILP formulation minimizes the number of monitors required, since the cost of each additional monitor is assumed fixed. Given a distribution on cheaters, the ILP formulation can be extended to include a dynamic cost of each additional monitor depending upon the number of traces unlocked.

For each trace collected, the observed activity information is assumed to be enough to identify the violator. But in the state-of-the-art, identification and location verification remain open questions, since PHY-layer and other identification techniques are still nascent, and are known to cause SNR degradation [33] . Non-repudiation of violations must also be considered when designing systems for identification. For instance, if the method of identification is an signed string, then any violating transmissions carrying this identifications are non-repudiable.

Currently, CPE trust is not considered during the monitor selection process. A trust metric maintained for each CPE could allow SysMet to account for Monitors which may not perform all required activities. We may not require a strict measure of trustworthiness, but a measure which can at least convey the likelihood of selecting a dishonest CPE as a monitor.

The incentive schemes proposed for CPEs allow selective unlocking of traces submitted by a monitor. Such a system has many advantages in terms of a middle ground between privacy and trust, but calls for mechanisms which can efficiently determine which traces to unlock, while minimizing the rewards given to CPEs for the trace information.

The primary method for CPE and BS privacy in the current design is the use of a collection probability $p_{collect}$, which causes a monitor to collect an observable instance of spectrum activity with probability $p_{collect}$. Further measures, such as adaptive increase of observation probability have not been address.

7.2 Future Work

In general, a system which provides a given probability of node presence can have many applications. Location verification for mobile nodes is one such open question - nodes can use the monitor network to obtain or convey a proof of location for use with various services.

The current game formulation assumes a limited number of cheater strategies, which helps in simplifying the formulation. The formulation can easily be extended for multiple cheater strategies, and could include different P_{caught} values for different attacks. Currently, requests for awarding monitor status are entertained only during the contention interval, but not during the monitor epoch. An optimal stopping game formulation for choosing monitors as

the requests are received can be explored.

With what was learnt from SysMet , a new system may be proposed. This system uses the same architecture as SysMet , but would not make the assumptions of location verifiability and CPE trust made before. Locations would be verified using common knowledge acquired from the device identities on the SUs in the area as location proof. A majority mechanism, such as WSPRT, would determine which location proofs are authentic and which are not. If a trace passes the location verification, the spectrum activity information gathered from the trace is analyzed for misuse. If misuse is found in a given cell or by a given device, then other redundant readings for the cell or device are gathered. Again, a majority mechanism would determine if misuse is truly occurring. This system would also have to take into account the strategies of a cheater which knows the majority mechanism for misuse detection. Monitor allocation would be done on an opportunistic basis, with requests being accepted or rejected as they come in.

7.2.1 Redundancy & Trust

Systems of trust can be used a metric to quantify the integrity of traces provided by a CPE. This would allow SysMet to account for Monitors which may not perform all required activities. We may not require a strict measure of trustworthiness, but a measure which can at least convey the likelihood of selecting a dishonest CPE as a monitor. Redundancy in the monitor network, along with methods to combine traces or reports from different monitors could also be explored. Methods such as WSPRT [] have been used to combine spectrum sensing information, while providing a weight-based mechanism to reduce the value of decisions provided by malicious users. Such a mechanism would require multiple reports for the same cell, and thus immediate misuse detection with such a system would not be possible.

A deeper analysis of the privacy afforded to monitors, along with the implications of reporting traces of spectrum activity for BS privacy must be studied. The stationary probability matrix P is estimated using the visit probability matrix P, but this estimate can be refined by learning stationary probabilities from traces provided by monitors for different areas.

7.2.2 Optimal Stopping Games

With a visit probability based approach, monitor status may also be awarded on an opportunistically. Given that seeds are distributed uniformly randomly, it is reasonable to assume that the same fraction is more or less maintained in each cluster of CPEs. Thus, given the density of CPEs present in a given area, the number of candidates for monitor status can be estimated. With this assumption, a simple *optimal stopping game with known number of candidates* may be formulated and solved to obtain a request acceptance rule for a single

monitor position. Common examples of optimal stopping games are the *Secretary problem* and the *Dice problem*. A *optimal multiple stopping problem* may also be formulated, under certain assumed conditions for the distribution of requests.

In an optimal stopping game, an optimal stopping rule is formed by checking which strategy in the decision has a higher expected payoff. The expected payoff is the most likely estimate from either decision, and thus becomes the best criterion for the decision. Very often, such games involve decisions where strategies involve instances of the same game. This makes such optimal stopping games amenable to dynamic programming solutions for small instances of the game.

Chapter 8

Bibliography

[1] IEEE Draft Standard for Information Technology -Telecommunications and information exchange between systems - Wireless Regional Area Networks (WRAN) - Specific requirements - Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Policies and procedures for operation in the TV Bands. *IEEE P802.22/D1.0, December 2010*, pages 1–598, Dec 2010.

[2] A. Agarwal and P. R. Kumar. Capacity bounds for ad hoc and hybrid wireless networks. *Computer Communication Review*, 34(3):71–81, 2004.

[3] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty. NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey. *Comput. Netw.*, 50(13):2127–2159, Sept. 2006.

[4] A. Anas, R. Arnott, and K. A. Small. Urban Spatial Structure. *Journal of Economic Literature*, 36(3):1426–1464, September 1998.

[5] G. Atia, A. Sahai, and V. Saligrama. Spectrum enforcement and liability assignment in cognitive radio systems. In *New Frontiers in Dynamic Spectrum Access Networks, 2008. DySPAN 2008. 3rd IEEE Symposium on*, pages 1–12, Oct 2008.

[6] B. Bahrak and J.-M. Park. Security of Spectrum Learning in Cognitive Radios. *CoRR*, abs/1304.0606, 2013.

[7] F. Bai and A. Helmy. Chapter 1 A SURVEY OF MOBILITY MODELS in Wireless Adhoc Networks.

[8] J. Balasch, A. Rial, C. Troncoso, C. Geuens, B. Preneel, and I. Verbauwhede. Pretp: Privacy-preserving electronic toll pricing. In *19TH USENIX SECURITY SYMPOSIUM*, pages 63–78. USENIX Association, 2010.

[9] P. Balister, A. Sarkar, and B. Bollobás. Percolation, Connectivity, Coverage and Colouring of Random Geometric Graphs. In B. Bollobás, R. Kozma, and D. Miklós, editors, *Handbook of Large-Scale Random Networks*, volume 18 of *Bolyai Society Mathematical Studies*, pages 117–142. Springer Berlin Heidelberg, 2008.

[10] R. Becker, R. Cáceres, K. Hanson, S. Isaacman, J. M. Loh, M. Martonosi, J. Rowland, S. Urbanek, A. Varshavsky, and C. Volinsky. Human mobility characterization from cellular network data. *Commun. ACM*, 56(1):74–82, Jan. 2013.

[11] K. Bian and J.-M. J. Park. Security vulnerabilities in ieee 802.22. In *Proceedings of the 4th Annual International Conference on Wireless Internet*, WICON '08, pages 9:1–9:9, ICST, Brussels, Belgium, Belgium, 2008. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

[12] Y. Y. Bo Gao and J.-M. Park. A credit-token-based spectrum etiquette framework for coexistence of heterogeneous cognitive radio networks. In *(to appear) INFOCOM 2014. The 33rd Conference on Computer Communications. IEEE*, pages –, April 2014.

[13] T. Camp, J. Boleng, and V. Davies. A survey of mobility models for ad hoc network research. *Wireless Communications and Mobile Computing*, 2(5):483–502, 2002.

[14] V. Chandola, A. Banerjee, and V. Kumar. Anomaly detection: A survey. *ACM Comput. Surv.*, 41(3):15:1–15:58, July 2009.

[15] K.-C. Chen, Y.-J. Peng, N. Prasad, Y.-C. Liang, and S. Sun. Cognitive Radio Network Architecture: Part I – General Structure. In *Proceedings of the 2Nd International Conference on Ubiquitous Information Management and Communication*, ICUIMC '08, pages 114–119, New York, NY, USA, 2008. ACM.

[16] R. Chen, J.-M. Park, and K. Bian. Robust distributed spectrum sensing in cognitive radio networks. In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pages –, April 2008.

[17] R. Chen, J.-M. Park, and J. Reed. Defense against primary user emulation attacks in cognitive radio networks. *Selected Areas in Communications, IEEE Journal on*, 26(1):25–37, Jan 2008.

[18] R. Chen, J.-M. Park, and J. H. Reed. Defense against primary user emulation attacks in cognitive radio networks. *IEEE J.Sel. A. Commun.*, 26(1):25–37, Jan. 2008.

[19] R. Chen, J.-M. J. Park, and K. Bian. Robustness Against Byzantine Failures in Distributed Spectrum Sensing. *Comput. Commun.*, 35(17):2115–2124, Oct. 2012.

[20] T. Clancy and N. Goergen. Security in Cognitive Radio Networks: Threats and Mitigation. In *Cognitive Radio Oriented Wireless Networks and Communications, 2008. CrownCom 2008. 3rd International Conference on*, pages 1–8, May 2008.

[21] A. G. Fragkiadakis, E. Z. Tragos, and I. G. Askoxylakis. A Survey on Security Threats and Detection Techniques in Cognitive Radio Networks. *IEEE Communications Surveys and Tutorials*, 15(1):428–445, 2013.

[22] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, June 2008.

[23] Y. Hou, Y. Shi, and H. Sherali. Spectrum sharing for multi-hop networking with cognitive radios. *Selected Areas in Communications, IEEE Journal on*, 26(1):146–155, Jan 2008.

[24] Y. Hou, Y. Shi, H. Sherali, and S. Midkiff. On energy provisioning and relay node placement for wireless sensor networks. *Wireless Communications, IEEE Transactions on*, 4(5):2579–2590, Sept 2005.

[25] E. Hyytiä, P. E. Lassila, and J. T. Virtamo. Spatial Node Distribution of the Random Waypoint Mobility Model with Applications. *IEEE Trans. Mob. Comput.*, 5(6):680–694, 2006.

[26] B. Jiang and T. Jia. Zipf's law for all the natural cities in the united states: A geospatial perspective. *Int. J. Geogr. Inf. Sci.*, 25(8):1269–1281, Aug. 2011.

[27] S. K. R. Kexiong Zeng and Y. Yang. Location robustness in database-driven white spaces network. In *Dynamic Spectrum Access Networks (DYSPAN), 2014 IEEE International Symposium on*, April 2014.

[28] V. Kumar, J.-M. Park, T. Clancy, and K. Bian. PHY-layer authentication by introducing controlled inter symbol interference. In *Communications and Network Security (CNS), 2013 IEEE Conference on*, pages 10–18, Oct 2013.

[29] K. Lee, S. Hong, S. J. Kim, I. Rhee, and S. Chong. Slaw: A new mobility model for human walks. In *INFOCOM 2009, IEEE*, pages 855–863, April 2009.

[30] K. Lee, Y. Kim, S. Chong, I. Rhee, Y. Yi, and N. B. Shroff. On the Critical Delays of Mobile Networks Under Levy Walks and Levy Flights. *IEEE/ACM Trans. Netw.*, 21(5):1621–1635, Oct. 2013.

[31] N. Li, T. Li, and S. Venkatasubramanian. t-closeness: Privacy beyond k-anonymity and l-diversity. In *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on*, pages 106–115, April 2007.

[32] M. H. Manshaei, Q. Zhu, T. Alpcan, T. Bacşar, and J.-P. Hubaux. Game Theory Meets Network Security and Privacy. *ACM Comput. Surv.*, 45(3):25:1–25:39, July 2013.

[33] J.-M. Park, J. Reed, A. Beex, T. Clancy, V. Kumar, and B. Bahrak. Security and enforcement in spectrum sharing. *Proceedings of the IEEE*, 102(3):270–281, March 2014.

[34] P. Paruchuri, J. P. Pearce, J. Marecki, M. Tambe, F. Ordóñez, and S. Kraus. Efficient Algorithms to Solve Bayesian Stackelberg Games for Security Applications. In D. Fox and C. P. Gomes, editors, *AAAI*, pages 1559–1562. AAAI Press, 2008.

[35] S. Parvin, F. K. Hussain, O. K. Hussain, S. Han, B. Tian, and E. Chang. Cognitive radio network security: A survey. *J. Network and Computer Applications*, 35(6):1691–1708, 2012.

[36] A. Patcha and J.-M. Park. A game theoretic approach to modeling intrusion detection in mobile ad hoc networks. In *Information Assurance Workshop, 2004. Proceedings from the Fifth Annual IEEE SMC*, pages 280–284, June 2004.

[37] A. Patcha and J.-M. Park. An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer Networks*, 51(12):3448 – 3470, 2007.

[38] M. Penrose. *Random geometric graphs*, volume 5. Oxford University Press Oxford, 2003.

[39] Y. Peres, A. Sinclair, P. Sousi, and A. Stauffer. Mobile Geometric Graphs: Detection, Coverage and Percolation. In D. Randall, editor, *SODA*, pages 412–428. SIAM, 2011.

[40] Y. Peres, A. Sinclair, P. Sousi, and A. Stauffer. Mobile geometric graphs: detection, coverage and percolation. *Probability Theory and Related Fields*, 156(1-2):273–305, 2013.

[41] F. Perich and M. McHenry. Policy-based Spectrum Access Control for Dynamic Spectrum Access Network Radios. *Web Semant.*, 7(1):21–27, Jan. 2009.

[42] I. Rhee, M. Shin, S. Hong, K. Lee, S. J. Kim, and S. Chong. On the Levy-walk Nature of Human Mobility. *IEEE/ACM Trans. Netw.*, 19(3):630–643, June 2011.

[43] A. Sahai, K. A. Woyach, G. Atia, and V. Saligrama. A Technical Framework for Light-handed Regulation of Cognitive Radios. *Comm. Mag.*, 47(3):96–102, Mar. 2009.

[44] R. Shokri. Quantifying and protecting location privacy, 2013.

[45] R. Shokri, G. Theodorakopoulos, G. Danezis, J.-P. Hubaux, and J.-Y. Le Boudec. Quantifying Location Privacy: The Case of Sporadic Location Exposure. In *Proceedings of the 11th International Conference on Privacy Enhancing Technologies*, PETS'11, pages 57–76, Berlin, Heidelberg, 2011. Springer-Verlag.

[46] R. Shokri, G. Theodorakopoulos, C. Troncoso, J.-P. Hubaux, and J.-Y. Le Boudec. Protecting Location Privacy: Optimal Strategy Against Localization Attacks. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, CCS '12, pages 617–627, New York, NY, USA, 2012. ACM.

[47] C. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. Shellhammer, and W. Caldwell. IEEE 802.22: The first cognitive radio wireless regional area network standard. *Communications Magazine, IEEE*, 47(1):130–138, January 2009.

[48] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos. Epidemic spreading in real networks: an eigenvalue viewpoint. In *Reliable Distributed Systems, 2003. Proceedings. 22nd International Symposium on*, pages 25–34, Oct 2003.

[49] K. Woyach, A. Sahai, G. Atia, and V. Saligrama. Crime and punishment for cognitive radios. In *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, pages 236–243, Sept 2008.

[50] H. Xu, J. Jin, and B. Li. A secondary market for spectrum. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–5, March 2010.

[51] F. Xue and P. R. Kumar. The Number of Neighbors Needed for Connectivity of Wireless Networks. *Wireless Networks*, 10(2):169–181, 2004.

[52] Z. Zhang, L. Yang, Y. Zhu, B. Zhao, and H. Zheng. Enforcing dynamic spectrum access with spectrum permits. In *Dynamic Spectrum Access Networks (DYSPAN), 2012 IEEE International Symposium on*, pages 278–279, Oct 2012.

[53] Y. Zheng, M. Li, W. Lou, and Y. Hou. Sharp: Private proximity test and secure handshake with cheat-proof location tags. In S. Foresti, M. Yung, and F. Martinelli, editors, *Computer Security ESORICS 2012*, volume 7459 of *Lecture Notes in Computer Science*, pages 361–378. Springer Berlin Heidelberg, 2012.