2020-05-22

On the Distribution of Inter-Arrival Times of 911 Emergency ResponseProcess Events

Blake Cameron Moss
Brigham Young University

On the Distribution of Inter-Arrival Times of 911 Emergency Response Process Events

Blake Cameron Moss

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Sean Warnick, Chair
Daniel Zappala
Casey Deccio

Department of Computer Science

Brigham Young University

ABSTRACT

On the Distribution of Inter-Arrival Times of 911 Emergency Response
Process Events

Blake Cameron Moss
Department of Computer Science, BYU
Master of Science

The 911 emergency response process is a core component of the emergency services critical infrastructure sector in the United States. Modeling and simulation of a complex stochastic system like the 911 response process enables policy makers and stakeholders to better understand, identify, and mitigate the impact of attacks/disasters affecting the 911 system. Modeling the 911 response process as a series of queue sub-systems will enable analysis into how CI failures impact the different phases of the 911 response process. Before such a model can be constructed, the probability distributions of the inter-arrivals of events into these various sub-systems needs to be identified. This research is a first effort into investigating the stochastic behavior of inter-arrival times of different events throughout the 911 response process. I use the methodology of input modeling, a statistical modeling approach, to determine whether the exponential distribution is an appropriate model for these inter-arrival times across a large dataset of historical 911 dispatch records.

Keywords: 911, emergency response, critical infrastructure protection, statistical modeling, queuing theory, input modeling

ACKNOWLEDGMENTS

My sincere appreciation to my advisor Dr. Sean Warnick, for helping me develop in technical areas that have not always come easy for me. My deepest gratitude to my parents, friends and siblings who have supported me in all my educational endeavors. And of course, to my wife Madi and daughter Naomi, who are my everything. For them, the words of Twain are true in my life: “wheresoever [they were], there was Eden.”

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction and Motivation

Research into critical infrastructure (CI) protection in the United States and across the world has received an increased focus from academic, industrial, and government professionals. Much of this attention can be traced to recent and notable disasters and attacks that have affected CI systems. Modeling and simulation techniques have shown promise in aiding in evaluating the impact of future attacks or disasters on CI systems. The 911 emergency response process (911) in the United States represents a vital component of the emergency services critical infrastructure sector as defined by the Department of Homeland Security [25]. 911 is the primary method citizens request and receive emergency services in times of need. Recent cyberattacks and natural disasters throughout the country have demonstrated the fragility of the system and brought into light questions regarding 911's resiliency. Also of grave concern is the potential for cascading failures among inter-related components of 911 in the face of targeted attacks or large-scale disasters. Modeling and simulation efforts are necessary to aid in quantifying the impact that future attacks or disasters could have on the 911 system, its participants, and stakeholders (i.e. the citizens the system serves). Since 911 is primarily focused on responding to emergencies throughout a community, some of the most important performance metrics are queue-related measurements. These measurements include estimates of customer queuing time (e.g. how long it takes from the time of call-placement to first responders arriving at an incident scene), system utilization, and expected service time for incidents. With regards to attacks or disasters affecting key 911 participants and/or infrastructure, assessing potential impact to these performance measures is vital to building appropriate mitigation policies.

Queuing models and the corresponding field of queuing theory provides the mathematical foundations to begin developing an appropriate model that will allow for the impact analysis as described above. Developing a valid and useful model of this nature is not a trivial task. This is especially true for complex, non-homogeneous, non-linear, stochastic systems like 911. There are several necessary steps that are required in order to compose such a model. When modeling systems with stochastic properties like 911, one of the critical steps is identifying and selecting appropriate probability distributions that best describe the stochastic properties of the system. This effort is known as *input modeling*.

For queuing models, one of the key stochastic properties is the *arrival process* and subsequent distribution of *inter-arrival times* of various events in the system. This research uses the methodology of input modeling to identify probability distributions that most appropriately describe the inter-arrival processes of events that occur during the 911 response process. This analysis is being performed primarily to determine whether the theoretical exponential distribution accurately reflects historical inter-arrival observations from multiple agencies' 911 environments throughout the country. This work is an important first step towards modeling the impact of disasters or attacks on the end-to-end 911 process in terms of vital queue-related performance metrics.

The contributions of this thesis are as follows:

1. Proposing a queuing model for impact analysis of CI failures on the behavior of 911 system processes.
2. Collection and processing of a large *publicly available* dataset of over 3.3 million computer-aided 911 dispatch events from nine agencies across eight states.
3. Synthesis of recorded events in the agencies' datasets into a coherent sequential end-to-end 911 event model.

4. Demonstrating that the inter-arrival times of all 911 response process events can be appropriately modeled using an exponential distribution.

5. Demonstrating that a Weibull distribution performs more accurate fitting for inter-arrival times of 911 response process events than the exponential distribution.

The organization of this research is as follows. Chapter two will introduce and motivate this research work. I first introduce the system of study: the 911 response process. A historical overview of the development of 911 in the United States will be detailed. 911's vital role as part of the Emergency Services Critical Infrastructure sector will also be treated. Notable attacks or disasters on the 911 system will be discussed. Finally, a reasoning as to the usefulness of a queuing model to perform impact analysis of this system will be presented.

Chapter three will provide preliminaries in the form of an introduction to queue models and the process of input modeling. The importance of the distribution of inter-arrival times in a queue model will be detailed.

Since this work falls under the research domain of modeling and simulation for critical infrastructure protection, chapter four will provide a literature review of modeling approaches for critical infrastructure protection. In addition, specific research into modeling efforts for emergency response systems will be detailed, as well as past research into the 911 system in the United States.

Chapters five through eight discuss the methodology, results and discussion of each step of the input modeling effort with regards to the 911 response process. The final chapter concludes with a discussion on the implications of the findings and path to further model development and future research.

Chapter 2

The 911 Response Process

2.1 Overview and History of 911

One of the most important public services provided in the United States is what is colloquially known as 911[1]. This system is best described as the process by which members of the public request and receive emergency services in the form of emergency medical, fire, or law enforcement resources. Since its inception over 50 years ago, 911 has proven to be incredibly valuable in assisting emergency response personnel, law enforcement, safety officials, and most importantly the public these entities serve. Its widespread use today has made the underlying processes that power 911 indispensable in emergency response. It is no secret that both the government and members of the public expect a level of reliability and resiliency from 911 services that far exceeds reliability requirements of most other complex socio-technological systems. It is estimated about 240 million calls to 911 are placed per year, with incidents ranging from violent crimes to natural disasters [62]. Those placing such calls unequivocally expect that their requests for service will be answered, processed, and appropriate resources dispatched, if needed, in a timely manner. The consequences of failing to provide these services is not merely dissatisfied customers, as is the case with many technological platforms, but rather lost lives and vast economic consequences in terms of increased healthcare/emergency response costs. It is of little surprise that the Federal Communications Committee (FCC) stated the 911 system is a "vital part of our nation's emergency response and disaster preparedness system"[30].

[1]Although countries throughout the world have similar emergency response systems, our work is limited in scope to the United States.

Although the emergency number "911" is a relatively new component of emergency response in the United States, systems and procedures designed to process emergencies and dispatch appropriate resources has existed for some time. In the days of colonization, policing was mainly a community task or provided by for-profit entities [70]. Fire-fighting organizations at the time were also primarily made up of citizen volunteers [3, 69]. However towards the mid-late 19^{th} century, jurisdictions began creating official, salaried police and fire departments [3, 70].

The development of communication technology, in particular the telephone, dramatically altered emergency response in the United States. The near-instant vocal communication allowed citizens to call their local fire/police departments to report emergencies and request services. Despite this convenience, many problems could abound. For instance, since local precincts/departments had to be called directly, visitors unfamiliar with the numbers were unable to report emergencies or crime [61]. Similarly, the lack of centralization in ever-increasing populations meant that prioritizing and coordinating resources was difficult [40].

Likely prompted by the above-mentioned concerns, the President's Commission on Law Enforcement and Administration of Justice in 1967 issued a report entitled "The Challenge of Crime in a Free Society." The report recommended that "wherever practical, a single police telephone number should be established, at least within a metropolitan area and eventually over the entire United States" [71]. In response to the recommendation, the FCC in conjunction with the American Telephone and Telegraph Company (AT&T) established "911" as the primary emergency number throughout the US. The initial test system was developed at Haleyville, Alabama and in 1968, the first 911 call was placed by Alabama Senator Rankin Fite [40].

Since then, the communications and technological infrastructure supporting various 911 systems throughout the country have improved tremendously. Several notable laws and regulations have been developed on both a federal and state-wide level to organize funding, establish quality standards, and otherwise assist in the further development of 911 [31]. All the while the use of the 911 system as the primary method of emergency response has grown significantly. According to the National Emergency Number Association (NENA), around 240 million calls are made to 911 every year [62].

2.2 Overview of Current 911 Process

Since there is no coordinating body for specific 911 implementation across the country, each 911 governing entity (i.e. state, city, local) decide to a large degree the specific policies, procedures and design of 911 systems. Even within these governing entities, different emergency service agencies (law, fire, EMS, etc.) can have their own policies and procedures for receiving, processing, and responding to 911 calls. Despite the heterogeneous nature of 911 implementations throughout the country, the high level behavior and end-result (i.e. providing emergency service to the public) is fairly consistent. Using the structure developed in [64], I present an overview of the notable events that occur during the 911 process.

2.2.1 The Call for Service

The call-for-service (CFS) is the initial event in the 911 response process. The term "call" in most cases is a voice call delivered by a variety of different mediums such as cellular networks, the public-switched telephone network (PTSN) or Internet infrastructure to a 911 call-taker stationed at a public-service answering point (PSAP). PSAPs themselves are heterogeneous and can be governed by state, city, or local governments [64]. In many cases, employees of specific agencies within these governments (i.e. law enforcement, fire, etc.) operate the PSAP. The calls are placed to 911 when emergency service is requested[2] Although voice calls are the most prevalent and preferred medium for emergency requests to 911, today over 2000 PSAPs support text-to-911 capabilities (e.g. sending SMS messages to request emergency service) [32].

[2]In reality, many calls placed to 911 are not for emergency service. For an illustration of this problem, see [74].

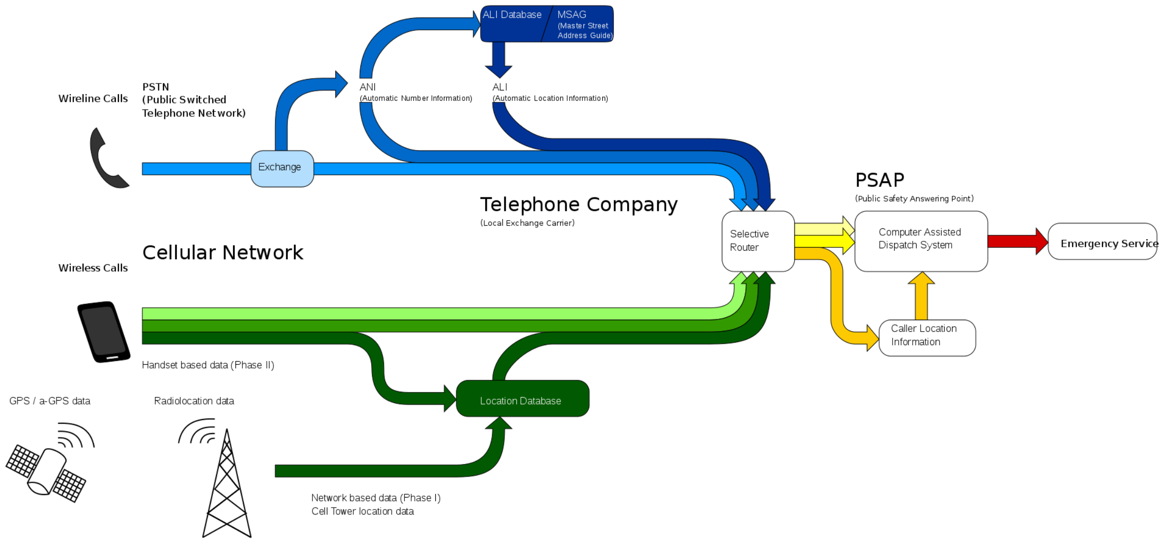

Figure 2.1: 911 Call Delivery Process, taken from Evan Mason, "9-1-1 System," via Wikimedia Commons. Licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license [57].

2.2.2 Call Delivery and Routing

As mentioned, once the call is placed, it is routed through a series of communication infrastructures to the appropriate PSAP. Depending on the source of the call (i.e. landline, cell-phone, Voice-over-IP, etc.), this routing step takes different forms. In contrast to specific point-to-point communication processes, 911 is the emergency number throughout the country. That means that instead of the caller needing to know the specific telephone number of the appropriate PSAP, it is incumbent on the 911 communications infrastructure to route the call based on location of the caller to the correct PSAP. Additionally, for the most effective emergency response, the call-taker at the PSAP ideally automatically receives the caller's current location. Accurate and automatic location information (ALI) delivery is obviously vital to the 911 response process, as many callers to 911 are unsure of their exact location. Figure 2.1 details some of the components of the call routing including automatic location/number identification (ALI/ANI).

This location-based routing and delivery has been a source of consternation for 911 stakeholders [48]. For landline based calls, the routing process is relatively simple. Phone companies are required to associate a landline number with a physical address. This associated physical address is used to map the landline caller to the geographically appropriate PSAP. Wireless or cellular calls are obviously more challenging to route, as the caller's physical location is not fixed to any particular registered address. In this case, the call location is usually determined via GPS on the wireless device or cellular towers via trilateration. This is not a flawless process, as calls can often be routed to the wrong PSAPs. One reason for this could be the non-uniform geographical boundaries for PSAP jurisdictions as shown in figure 2.2. For internet-based calls (i.e. VoIP), location of a caller is even more challenging. Delivery of accurate location information has been the subject of many of the laws and regulations around 911 [50].

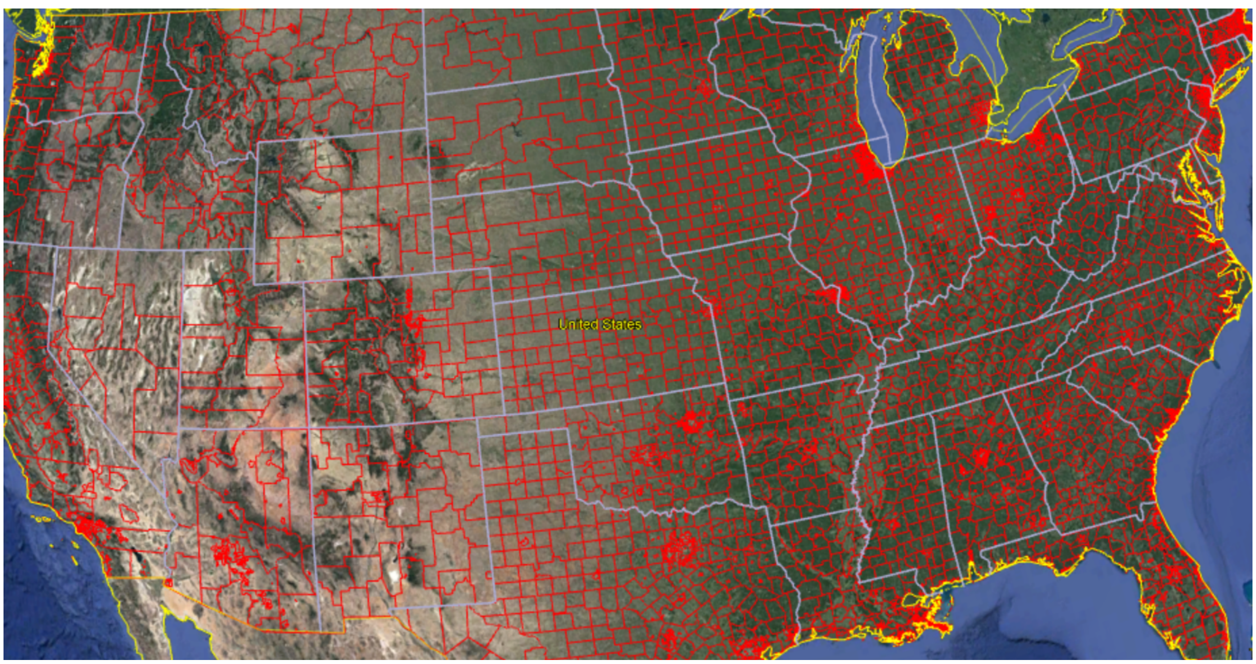

Figure 2.2: Primary PSAP Geographical Boundaries (outlined in red) in the Continental United States [27].

2.2.3 Call Processing

Although the routing and delivery of the call and location information differ depending on the communication medium and local 911 implementation, the next event in the response process is call-processing performed by a 911 call-taker housed at a PSAP. The responsibilities and procedures for 911 call-takers differ throughout PSAPs. At a high level, however all call-takers communicate with the caller to ascertain the nature of the emergency. Even if location information is automatically delivered, most 911 call-takers will confirm or ask for the location of the caller. The details of the incident are entered by the call-taker into the *computer-aided dispatch* (CAD) system [50, 64]. Modern CAD systems are a combination of various software and hardware components that enable several important incident-management capabilities. CAD implementations also differ widely as there are several prominent CAD vendors throughout the world [56]. CAD implementations play a major role in the measurement capabilities of PSAPs and coordinating first responder agencies.

2.2.4 Incident Dispatch

If it is determined by the call-taker that emergency services are required, the appropriate emergency response units are dispatched. This actual dispatching is carried out by another PSAP or agency-specific employee called the *call-dispatcher*. In some jurisdictions however, the call-taker and dispatcher may be the same person [64]. In either case, the dispatcher coordinates and facilitates the dispatch of emergency units to the geographical scene of the incident. As part of the CAD system, automatic vehicle location (AVL) equipment in conjunction with mobile terminal units (MTU) in response vehicles can record measurements such as the timestamps of when the unit is enroute and arrives at the scene [58].

2.2.5 Performance of Emergency Service and Closing of Incident

Without detailing the specifics of one particular agency or jurisdiction, the next event that occurs in the process is the performance of whatever emergency service is required at the incident scene. This service can take many different forms, from apprehending criminals to providing medical services. Theoretically though, the response units spend a duration of time performing the required service and then indicate to the PSAP (most likely through CAD components) that the incident can be considered closed or that the communicating response unit is ready to be dispatched to another incident. This incident closure represents the final event in the 911 process.

2.3 911 as a Critical Infrastructure Component

The 911 response process makes up a core component/capability of the Emergency Services Critical Infrastructure Sector (ESS) as defined by the Department of Homeland Security (DHS) and the Cybersecurity and Infrastructure Security Agency (CISA). According to a CISA sector overview: "ESS is a community of millions of trained personnel along with the physical and cyber resources that enable them to provide a wide range of prevention, preparedness, response, and recovery services during both steady-state and incident management operations"[25].

2.3.1 Dependencies of the 911 Process

Although logically placed within the emergency services CI sector, it is obvious that the efficacy of the 911 response process hinges on the intra-dependencies within the ESS as well as inter-dependencies between other critical infrastructure sectors. As detailed in [25], several notable intra/inter-dependencies include:

1. Reliance on performant communication channels that deliver CFS to appropriate PSAPs and allow for inter-agency incident communication.

2. Electric and other energy production/distribution systems to power communication infrastructure, buildings, transportation units, and other vital 911 components.
3. Adequate transportation networks and resources to efficiently facilitate emergency resources to varying geographical locations.
4. Efficient internal information technology components like CAD, that aid in data input, storage, manipulation and distribution.
5. Water supply and distribution network for fire-fighting response capabilities.

As 911 systems have grown in complexity, dependencies such as those listed above have become even more entrenched. An apt example is the proposed Next-Generation 911 (NG-911) design that is being actively implemented throughout the country. NG-911 relies primarily on Internet-based (IP) technologies and protocols to perform enhanced call-processing and delivery capabilities. The use of IP technologies allows for more robust and efficient functionality but adds an additional dependency for the system (i.e. the public Internet infrastructure) [73].

These increasingly reliant dependencies help to illustrate devastating potential impact on the 911 process should these dependencies be compromised (even partially) by attacks or disasters. While assessing the ESS sector, CISA listed the four most significant risks facing ESS and consequently, the 911 response process. These risks included:

- cyber-infrastructure attacks or disruptions
- natural disasters and extreme weather
- violent extremist and terrorist attacks
- chemical, biological, radio-logical and nuclear incidents

In addition, DHS/CISA identified priority research and development (R&D) topics to address these risks. Among these R&D topics, situational awareness and training/simulation capabilities were noted. Despite the ubiquity of the 911 system and its core importance to the ESS sector, specific progress with respect to these R&D priorities for 911 has been lacking.

2.4 Impact of CI Attacks/Disasters on the 911 Response Process

The impact of CI dependency failures specifically on the 911 response process ultimately reduce down to three primary effects:

- Reduced capacity for callers to request and receive emergency service.
- Increased delays between intervals of 911 response process events (i.e. longer periods necessary to take down caller information, relay call information to first responders, travel to incident location, etc.)
- Increased frequency of requests for emergency service resources.

Since the 911 response process is relied on by participants within ESS as well as the general public, the results of these effects can lead to distrust of 911, the waste of precious emergency service resources and in extreme cases, increased injury or loss of life.

Although real-world examples abound, I present three historical events that highlight how the previously mentioned risks to the 911 system have the potential to cause the discussed effects.

Terrorist Attack: September 11 World Trade Center Attacks

The hijacking and targeted crashing of two commercial airliners into the World Trade Center Towers in New York City (NYC) on September 11, 2001 had severe and immediate impact on components of the city's 911 system including NYC's primary PSAP, New York Fire Department (NYFD), and New York Police Department (NYPD). Almost instantly after the first tower was struck, calls immediately flooded the 911 system and many callers received busy signals indicating that all telephone circuits were in use. At the time the NYPD was the primary agency managing the 911 call centers and had a policy in place to transfer fire-related calls to a NYFD dispatcher. In many instances, these transfers encountered delays and many calls were pre-maturely disconnected. In addition, "because the [911] operators were not informed of NYPD Aviation's determination of the impossibility of rooftop rescues from the

Twin Towers on that day, they could not knowledgeably answer when callers asked whether to go up or down"[43].

Natural Disaster: Hurricane Katrina

The 2005 Hurricane Katrina devastated many parts of the Gulf Coast, including the densely populated city of New Orleans, Louisiana. Due to floodwaters, the initial and backup PSAPs in New Orleans were abandoned. After "hours of confusion" and most likely unanswered emergency service requests, calls were then subsequently routed to the Louisiana State Police emergency operations center, more than 70 miles away from New Orleans. Operators at the center were unaware of the routing decision and were unprepared for the crush of calls in the following days, resorting in some cases to "scribbling [CFS] information on sheets of paper". Makeshift operations heroically answered more than 22,000 emergency calls but "only a fraction could be handled at a time" [19].

Cyber Attack: Telephony Denial of Service

In October 2016, 18-year-old Meetkumar Desai, a resident of Maricopa County, Arizona, was indicted on charges of carrying out a cyber-attack against 911 systems. A friend of Desai's had shared a bug in Apple's iOS software (the operating system that runs iPhone devices). This bug had the potential to maliciously activate the telephone-dialing functionality of the phone. Desai created an exploit of the bug in the form of a malicious javascript embedded link. When unsuspecting users clicked on the link, the exploit would continuously call 911. Desai then posted the link to both his Twitter and Youtube accounts.

Although only a limited amount of users clicked on the links, the impact of the attack was felt by 911 call centers throughout Arizona and possibly the southwest. One PSAP received over 100 calls within the space of a few minutes and was in "immediate danger of losing service" to vital communication infrastructure [2]. Other PSAPs were also affected and immediate action in the form of shutting down the hosted exploit links had to be taken in order to maintain PSAP capability to receive true emergency calls [52].

These historical incidents serve as proof that impacts to CI systems can have major cascading effects on the various components of the 911 response process.

2.4.1 Queue Models to Assess Impacts of CI failures on 911 Process

The purpose of modeling and simulating the impacts of attacks/disasters on critical infrastructure is in part to analyze how failures in dependency links can change the performance of the impacted systems. One of the most important benefits of serious modeling efforts is the potential for "what if" or predictive analysis. The reasoning behind this is quite simple: real-world disasters/attacks like those listed in the previous section are undesired but inevitable. In order to mitigate the consequences of these events, it is imperative to understand the scope of those consequences. Predictive analysis allows for insight into how a system might respond during an attack/disaster scenario. This type of analysis is obviously much less costly and more flexible than retrospective assessment that takes place *after* an actual disaster/attack has occurred.

For the 911 response process, the modeling effort to inform this type of impact analysis could take many forms and could result in several relatively dissimilar, though useful models. However, considering the main functions and performance metrics of 911 helps to provide insight into why modeling the 911 response process as a series of queue sub-systems would prove fruitful in quantifying the potential impact of CI failures.

The main function of the 911 response process is to provide emergency services to those who request it. Those requesting service can be defined as *customers* of the system. Customers demand service or *arrive* at the system randomly. *Servers* in the system are the abstract entities that provide service to customers. In the 911 process, servers could include communication infrastructure (e.g. telephone circuits), PSAP call-takers/dispatchers, and emergency response units. Providing service to customers of the 911 system takes time (i.e. processing the call, dispatching resources, traveling to the incident, performing service, etc.). Due to this and the potential that customer demand could presumably exceed server resources, a *queuing delay* could be incurred. In other words, customers are forced to wait for a period of time to receive emergency service.

Time is usually the measurement by which 911 systems are evaluated. For example, since emergencies are generally urgent in nature, the *response time* of a 911 system, or how long it takes a dispatched emergency response unit to arrive at the scene of an incident, is an important metric used to gauge a certain emergency service agencies' effectiveness. Beyond just response time, the total time from when a customer places a 911 call to when emergency service concludes is a gravely consequential interval, where in extreme situations even a few seconds may make the difference between life and death.

Queuing models and the mathematical field of *queuing theory* help to identify the behavior of stochastic systems like 911 as described above. Queuing models are primarily developed to describe the nature of waiting in a particular system. Given that the performance of 911 is commonly evaluated by measurements of time intervals, queue models represent an appropriate model class to represent the various sub-processes or sub-systems of 911. Furthermore, as detailed in section 2.4 CI failures tend to affect the 911 process in terms of increased *arrival rate* of customers, decreased *service time* of serving resources, and reduced *server capacity*. These properties are critical inputs into queuing models. Thus queuing models provide a promising avenue to explore the rippling effects of CI failures throughout the 911 system.

To further illustrate, consider figure 2.3. Each 911 sub-process/event described previously is modeled by a series of queue sub-systems. The reasoning here is intuitive. As a customer (CFS) progresses or *arrives* at different phases of 911, they must receive *service* at that phase before continuing on to the next sub-system. These sub-systems have their own inter-arrival and service rate parameters: λ and μ, respectively. I define inter-arrival of events to mean the time-interval between sequential customers requesting service from the

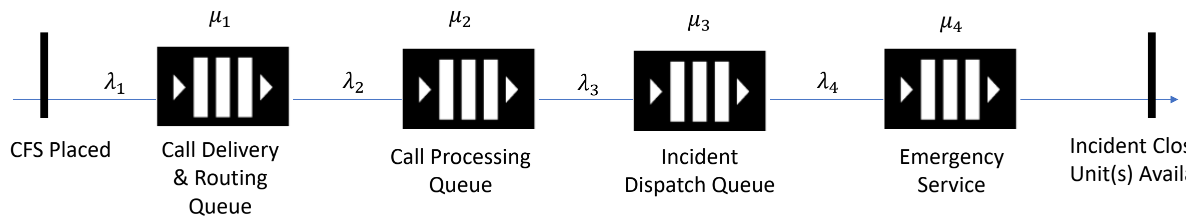

Figure 2.3: 911 Response Process Events as a Series of Queue Sub-Systems

sub-system. For example, after receiving service at the Call Delivery & Routing sub-system (i.e. the call is delivered through the communication infrastructure to a PSAP), the call enters the Call Processing subsystem. Depending on the availability of servers (i.e. call-takers), the customer may have to wait on the line (demonstrating queuing behavior) or if the situation is extreme enough (i.e. the queue is too large), that customer could be rejected from the sub-system altogether.

Now consider an attack or disaster resulting in a CI failure that impacts one of these sub-systems. Using Hurricane Katrina as an example scenario, let us conjecture as to the effects on the parameters of these 911 sub-systems. During Hurricane Katrina, call-takers were forced to abandon their primary and backup facilities. Subsequently, calls were routed to a different facility, where call-takers were unprepared. Further technological impacts reduced the call-takers service capability (i.e. decreased service rate parameter), as they resorted to manually writing down incident details. Incident dispatch and emergency service service rates were also obviously reduced, since it took longer to communicate with first-responders as well as to travel to those requesting emergency service (due to flooding in the streets). As shown in figure 2.4, this scenario can be modeled by a decreases in service rates at the respective sub-systems. These reduced service rate would also most likely have other cascading effects as well. For example, since incidents are not being processed as quickly, the inter-arrival times of incidents arriving at the Incident Dispatch sub-system decreases.

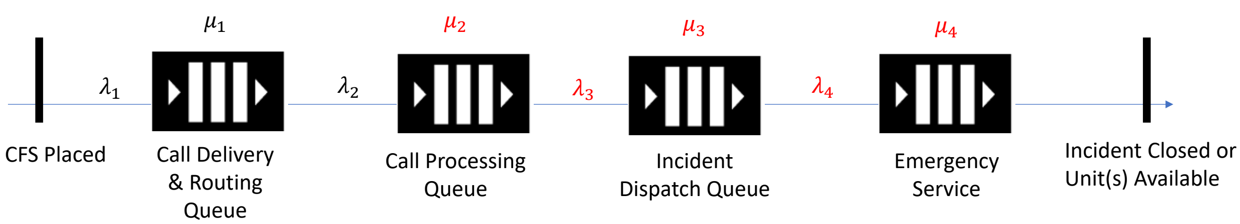

Figure 2.4: 911 Response Process Events as a Series of Queue Sub-Systems during CI Failure (Impacted Parameters Highlighted in Red)

As will be discussed, queuing theory provides the mathematical foundation for developing a model like the one proposed. However, before any such model can be constructed, a critical first step is to identify the stochastic properties of each 911 sub-system as it currently operates. Only with reasonable and validated assumptions regarding the behavior of these processes will the impact-analysis (i.e. changing λ, μ parameters) be useful. Input modeling is the methodology for identifying these stochastic properites and is introduced in the next chapter.

Chapter 3

Preliminaries: Introduction to Input Modeling for Queues

This chapter will introduce input modeling, its importance and methodology. First though, an overview of modeling and simulation, as well as an overview of queuing theory, will be given.

3.1 Modeling and Simulation

Modeling provides an avenue to scientifically study real-world processes, phenomena, and facilities. These processes, phenomena and facilities are usually termed *systems*. Even in the most basic cases, real-world systems are infinitely complex. In light of this complexity, simplifying *assumptions* in the form of mathematical or logical relationships are made to develop a limited understanding of the behavior of a system. These mathematical/logical relationships constitute a *model*. In some cases, the model may be simple enough to ascertain through the exact solutions for particular properties of interest for a system mathematical techniques (e.g. differentiation, integration, etc.). Solving for properties of a model in this way is known as an *analytical* solution. In many cases, the relationships that define the model can be too complex for analytic methods. In these cases, the model is often examined via *simulation*. One author has defined simulation as "numerically exercising the model for the inputs in question to see how they affect the output measures of performance"[49]. The advent of computing technology has made simulation an attractive alternative to analytic techniques.

In many cases it is desired that a system be analyzed *dynamically* or as it evolves over time. Models that account for this time-evolution are considered *dynamic models*. This in contrast to *static models* where a system is being analyzed at a fixed point in time or time does not play any role in the behavior of the system.

In the real-world, many systems also display properties of randomness or unpredictability. To account for this randomness, *probability* is employed in the modeling process to characterize the random components of the system. The inclusion of randomness transform a *deterministic* model into a *probabilistic model*. If a probabilistic model also seeks to describe the evolution of a system's random behavior over time, this constitutes what is known as a *stochastic model*. Since this research primarily focuses on modeling efforts relating to a stochastic system, a brief review on important probability terms will be given here.

3.1.1 Probability Spaces

Probabilistic models rely on the mathematical construct of *probability spaces* to characterize the behavior of random components. A formal description of a probability space can be given by $(\Omega, \mathbf{F}, \mathbf{P})$. Ω represents the sample space or all possible values of the variable in question. $\mathbf{F}$ describes the event space which is usually thought of as subsets of the sample space. $\mathbf{P}$ is the probability function where:

$$\mathbf{P} : \mathbf{F} \to [0, 1]$$

3.1.2 Random Variables

A real-valued *random variable* X is a function which maps a probability space into the real line such that:

$$X : (\Omega, \mathbf{F}, \mathbf{P}) \to \mathbb{R}$$

A random variable whose values can be finitely defined is considered a *discrete* while a random variable that can be mapped to any value on a continuous interval like $\mathbb{R}$ is considered *continuous*.

Probability Distributions

In the univariate case, a *probability distribution* is a mathematical function that describes the probabilities of the values a random variable X can take on. Probability distributions are also either discrete or continuous, since they correspond with the nature of the value of X. There are different ways to characterize the probability distributions of random variables but two common functions are the *probability density function (PDF)* and *cumulative density function (CDF)*. PDFs are used for continuous random variables while *probability mass functions (PMFs)* are used for discrete random variables. Formally these functions can be defined as follows, where X is a real-valued random variable and $\mathbf{P}$ is the probability measure:

- CDF: $f(x)$ such that $f(x) = \mathbf{P}(X \leq x)$
- PMF: $f(x)$ such that $f(x) = \mathbf{P}(X = x)$
- PDF: $f(x)$ such that $\mathbf{P}(X \in [a, b]) = \int_a^b f(x)dx$

There are numerous formally defined theoretical continuous and discrete probability distributions. These theoretical distributions can be used as approximations to model the randomness of real-world systems/processes.

3.2 Queuing Theory

The origin of queuing theory is believed to have its roots in A.K. Erlang's work on modeling caller traffic for the Denmark telephone exchange. Emanating from stochastic systems theory, queuing theory has been given attention in many research disciplines, most notably operations research. However, given the prevalence of queues in almost every scientific and social discipline, applications of queuing theory results span across a diverse set of fields. In its most basic form, the main participants in a mathematical queue model are customers and servers. Customers arrive at the system in some fashion and receive service before exiting the system. In general, queues are considered stochastic systems governed by random variables that are defined by a number of key terms:

- Arrival Discipline: Since customers arrive at random to the system throughout time, this process is stochastic. The distribution of elapsed time between sequential customers requesting service can characterize this process.
- Service Discipline: The distribution of service times for customers. The most tractable and probably most studied service distribution is exponential with parameter μ.
- Queue Discipline: This term refers to how to system receives customers into the system. The most common queue discipline is first-come-first-serve (FCFS).
- Number of Servers: A distinct queuing system can have C parallel servers. In general most parallel systems assume that servers are homogeneous, meaning they possess identical service rate distributions.
- Buffer Length: In many real-world systems, there is a maximum length queues can reach before arriving customers either balk (refuse to enter the queue) or are otherwise blocked from entering the queue.

It is important to note that there do exist deterministic queues, where the queuing properties mentioned above are fixed (i.e. an arrival rate of exactly 4 customers per minute without exception). However in the real world uncertainty almost always exists, which necessitates stochastic modeling approaches. Queues can be studied in discrete or in special cases, continuous time. Additionally, in more complex and real-world scenarios single queuing stations are often combined to form networks of queues. Much research has been devoted to studying these networks, although significant results are mostly dependent on arrival and service disciplines since those properties largely determine the tractability of the network model.

3.2.1 Steady-State Analysis

Queuing behavior is usually analyzed in two paradigms: transient and steady-state. Usually transient analysis of queues is difficult to obtain. Furthermore, the results may be intractable or else too complicated to be used realistically [59]. For that reason, a large amount of research has been devoted to steady-state analysis of queues. Using the notation in [59], to further clarify, let $N(t)$, be the number of entities in the queuing system (customers both in the queue plus those being served). At time t the probability distribution of $N(t)$ is given

$$p_n(t) = PrN(t) = n, n = 0, 1, 2, ...$$
$$p_i(0) = 1$$

Where, at $t = 0$ the system had an initial number of customers i. To fully characterize this stochastic system, time-dependent solutions would be needed for every $p_n(t), n \geq 0$. Due to the difficulty and practical challenges this entails, one often seeks the limiting behavior, $p_n(t)$ as $t \to \infty$. When the limit exists, the system has reached steady-state, or equilibrium. This means that the probabilities p_n are no longer time-dependant which greatly simplifies the analysis. For the purposes of this paper, we assume queues to be in steady-state.

3.2.2 Performance Measures

Given a queuing system in steady-state, performance measures are generally sought to express statistics relevant to the queuing process. Besides P_n, the probability that n customers are in the system, several other performance measures exist. These measures (and their associated notation) include:

- L = expected number of customers in the queuing system (those waiting in the queue plus those being served)
- L_q = Expected number of customers in the queue (excluding those being served)
- W = Expected sojourn time for each customer. Sojourn here means wait time in the queue plus service time
- W_q = Expected wait time (in queue only)

These performance measures are important to characterize relevant attributes of a queuing system.

3.2.3 Kendall Notation

The most popular way for succinctly summarizing a queuing system was proposed by David Kendall in 1953 [44]. Kendall's notation, as it is referred to, describes a queuing system as follows:

$$A/B/C/D$$

where A denotes arrival discipline, B service discipline, C number of servers, and D finite buffer length. If D is left out, it is generally assumed the system has an infinite buffer length.

A commonly studied queuing model using Kendall's notation is the $M/M/C$ model. Generally regarded as the most simple queuing model, this system's properties include a Poisson arrival rate λ, exponential service rate with mean $\frac{1}{\mu}$, C parallel, homogeneous servers, and an infinite buffer/queue length.

3.2.4 Theoretical Exponential Distribution

The exponential distribution is arguably the most important distribution in queuing theory. In many cases it represents the time until an event occurs (like the next customer arriving to a system). The importance of the exponential distribution for many queue models is due in large part to the memory-less or Markovian property of the distribution. A brief discussion of some of the basic properties of this distribution and its importance to queuing theory.

Formally defined, an exponential random variable T is a continuous random variable with a PDF of:

$$f(t) = \lambda e^{\lambda t} \ (t \geq 0)$$

where $\lambda > 0$ is constant. For this work, T represents some quantity of time, usually the time until an event occurs. The parameter λ describes the rate with a unit of event per time. For example, if T represents the inter-arrival time of calls to 911, then λ is the number of calls that happen over the unit of time.

The exponential distribution has been proven to have the memory-less property. This property can be defined as:

$$PrT > t + s | T > s = PrT > t$$

Which can be interpreted as the probability of the time until an event occurs (T) being greater than $t + s$ where both t and s are quantities of time, given that s amount of time has already passed, is equivalent to the probability of the time until an event occurs being greater than t. In other words, the time until the next event does not depend on how much time has previously passed. It is important to note that exponential distribution is the only continuous distribution that displays this property [80].

Exponential inter-arrival times for queues are convenient because they allow the queue to be modeled by the theory of continuous-time Markov chains (CTMC). CTMC's are a way to represent the stochastic evolution (in terms of state transitions) of a system through time. The Markov property (in continuous-time) is defined as:

$$PrX(t + s) = J | X(t) = i, X(u), 0 \leq u < t = PrX(t + s) = j | X(t) = i$$

where given a current state $X(t)$, the future state $X(t + s)$ is unrelated to the past. This property is only possible because CTMC's have both exponential inter-arrival service rates [80].

Choosing to model the inter-arrival times of a system as exponential is commonly done to ease analytical evaluation. Validating that assumption on real-world system data however is critical in developing a model that actually represents the system in real-life, not just a mathematically convenient representation. It is for this reason that input modeling is necessary.

3.3 Input Modeling for Queues

In both analytical and simulation modeling efforts for stochastic systems, a necessary process is to understand and hopefully identify the "appropriate [probability] distributions to represent the stochastic mechanisms of the system" [80]. For queues, the two most important mechanisms are the inter-arrival and service processes. Determining the appropriate distribution to represent these two phenomenon is a process known as *input modeling*. In situations where observations of the system in study are available, input modeling usually consists of five activities as partly outlined in [49]:

1. Gathering relevant observations from the real-world system that is being modeled
2. Assessing the independence (and stationary) of the observations
3. Hypothesizing families of probability distributions that represent the observed data
4. Estimating parameters of the theoretical distributions in order to "fit" the distribution to the observed data
5. Determining how adequately the fitted distribution represents the observed data

For the construction of analytically or numerically evaluated models, the identification of the input probability distributions can have a major implications regarding the overall system model [49]. This is especially true when the theoretical distribution possesses convenient mathematical properties that allow for significantly more tractable analysis of the model. For example, a large amount of analytically tractable queuing models depend on the assumptions that inter-arrival and service times of customers are exponentially distributed.

Simulation modeling efforts also necessitate identifying input distributions, albeit with more flexibility. It may be the case that no theoretical distribution adequately describes the observed data. In this case an *empirical distribution* can be constructed from the observed data. This empirical distribution can subsequently be used in the simulation model to sample random variates as needed. However as argued in [49], a theoretical distribution, if appropriate, is preferred for several reasons including concise representation and easily modified parameters.

Input modeling can be viewed as a statistical modeling approach in that it uses empirical data (historical observations of the system) in order select the model that best represents the stochastic or probabilistic system properties of interest.

Chapter 4

Related Literature

4.1 Critical Infrastructure Protection: Modeling & Simulation

Critical infrastructure in the United States is defined by the Cybersecurity and Infrastructure Security Agency as "infrastructure sectors whose assets, systems, and networks, whether physical or virtual, are considered so vital to the United States that their incapacitation or destruction would have a debilitating effect on security, national economic security, national public health or safety, or any combination thereof" [21]. The sectors include diverse areas ranging from as communications and financial services, to energy and food/agriculture. The stability and security of these critical infrastructure sectors is necessary for the primary functioning of society as we know it. Of key concern in critical infrastructure security are cascading failures, where disfunction in one CI system carries over to another. These interdependent failures have been observed in the real-world in the wake of both natural disasters and plotted attacks [10, 86]. Technological advancements have led to critical infrastructures being more inter-reliant than ever before. Along with this advancement, the digitization of many critical infrastructure processes has led to an increased wariness of the potential of crippling cyber-attacks on critical infrastructure throughout the country. A key component to protecting and securing critical infrastructure systems is understanding the system's internal dynamics as well as its effects on other interdependent systems. Modeling and simulation (M&S) are well-established scientific methods that allow for the abstraction of complex systems in order to better understand their functioning. These methods have been used to assess the impact of potential cyber-attacks and other adversities to critical infrastructure systems. Research in this area has led to improved design choices and mitigation strategies for critical infrastructure security. This paper will survey a selection of those research efforts. First, however an introduction to critical infrastructure in the United States and its importance to the country's functioning will be detailed.

4.1.1 Origins of CI protection

Critical infrastructure protection in the United States, as it is known today, has been a rapidly changing field for the past two decades. In fact, until 1996, the words "critical infrastructure" had not been formally defined by an official government body. Lewis argues in [51] that the origins of CIP had its roots in the aftermath of the Cuban Missile Crisis. During the crisis, problems with telecommunication between the Soviet Union and the US had threatened to derail negotiations. Due to the newfound importance of robust telecommunication technology for matters of national urgency, President John F. Kennedy created the National Communications System (NCS) in 1963 to provide robust and resilient communications for the Federal Government in all situations.

Soon after the NCS, another telecommunications-focused agency, the National Telecommunications and Information Administration (NTIA) was formed in 1978. Lewis believes that these agencies were the first to be created primarily for CIP. The increasing reliance of technologies such as television, radio, telephone networks, and later the Internet, by American society (both public and private sectors) throughout the 1980s made telecommunications widely recognized as critical infrastructure (even though the term was not defined yet).

Although telecommunications agencies had been created, during the 1970's and 1980's natural disaster agencies like FEMA were mainly charged with responding to disasters. At this time there was no consistent definition of critical infrastructure and definitely no formalized agreement as to what assets throughout the country were considered to be critical. In 1988, Executive Order 12656, "Assignment of Emergency Preparedness Responsibilities" attempted to assign responsibility for the "safekeeping of essential resources" to the Federal departments and agencies, but was full of problematic statements and did not define further what constituted essential resources [51].

The 1990's were a period of rapid technological development, increased interdependence in critical societal functions in the United States. Lewis argues that the uptick in terrorist attacks between 1993 and 1995 exposed the taken-for-granted infrastructure like telecommunications, energy-plants, and notably emergency response systems, as poorly prepared for not only terrorist attacks, but even accidental failures. In 1996, Executive Order 13010, "Critical Infrastructure Protection" defined critical infrastructure for the first time as "an infrastructure so vital that its incapacity or destruction would have a debilitating impact on our defense and national security". It listed infrastructures such as "telecommunications, electrical power systems, gas and oil storage and transportation, banking and finance, and emergency services (including medical, police, fire, and rescue)" as critical and defined two types of threats to CI: physical and cyber.

The main impetus for the intense focus on CIP was the September 11, 2001 terrorist attacks [5, 51]. Beyond the terrible loss of life that occurred, the attack also had impact on several CI sectors like telecommunications, transportation and as a consequence, emergency response [41, 55, 63]. In response to the attacks, the Department of Homeland Security (DHS) was created. Today, the DHS is the federal agency most directly responsible for CIP. In 2018, President Trump signed the Cybersecurity and Infrastructure Security Agency Act of 2018. This act created the Cybersecurity and Infrastructure Security Agency (CISA) that is tasked with coordinating "security and resilience efforts using trusted partnerships across the private and public sectors" [28].

Today, the CISA has categorized critical infrastructure into the following sectors [21]:

- Chemical: Responsible for the creation, manufacturing, and distribution of chemical resources.

- Communications: Includes multiple communication mediums like telephone, radio, satellite, Internet, etc. Given the technological developments of recent years, this sector has high inter-dependencies with the Information Technology sector.
- Dams: In charge of hydroelectric power generation, flood-control, and other related services.
- Emergency Services: Includes law enforcement, fire services, and emergency medical services (EMS). Also included are public safety answering points where 911 emergency calls are received and processed.
- Financial Services: This sector entails banking and other financial markets that are critical for the purchasing and selling of goods and services.
- Government Facilities: Responsible for buildings and associated property, equipment, or individuals that are used by federal, state, and local government for official purposes.
- Information Technology: This sector includes a wide array of various computer-based technology like hardware and software components. It also shares responsibility with the communications sector for the Internet and its associated entities.
- Transportation Systems: This includes the all the various mediums used to transport goods and people throughout the country and world. This sector includes aviation, motorways, water transportation, and postal/shipping services.
- Commercial Facilities: This sector entails many buildings or places of gatherings that are used by the public. This includes gathering places like casinos, offices, apartments, and theme parks.
- Critical Manufacturing: This sector focuses on the manufacturing of products that power crucial industry in the United States. This includes metal, machinery, electrical and transportation equipment.

- Defense Industrial Base: This sector relates to entities that enable the function and progress of the US military. This includes a plethora of government and non-government companies, agencies, and contractors that perform national-defense related work.
- Energy: One of the most visible critical infrastructures, this sector covers entities and processes relating to energy supply. This sector is broken down into electricity, oil, and natural gas sections.
- Food and Agriculture: This sector entails farms, food services (i.e. restaurants), and food manufacturing and processing plants.
- Healthcare and Public Health: These two components form the sector that is responsible for local healthcare as well as overall population health (public health). This sector most likely consists of the myraid of healthcare professionals, manufacturers, medical treatment facilities throughout the country.
- Nuclear Reactors, Materials, and Waste Sector: Besides the energy producing nuclear reactors, this section also includes all efforts relating to nuclear materials and waste management.
- Water and Wastewater Systems: This sector is responsible for providing safe drinking-water as well as properly treating waste-water (sewage systems).

As technology and society has evolved, these sectors have become increasingly more complex and interdependent. For example, the communications sector, which was once primarily voice and basic text transmission (telegraph and telephone), is now a myriad of various communication technologies that span the country and mediums that range from satellite to WiFi. Additionally, the communications sector is critical to the proper functioning of a number of other sectors like emergency response and financial services.

Developing the understanding of system dynamics necessary to create informed policies, designs, and implementations of secure CI is extremely difficult due to the complexity and highly interdependent nature of these infrastructures. As a result, CIP has been an active field of research among both private and public sector institutions. Of particular importance to the field has been efforts in modeling and simulation. The Department of Defense glossary defines a model as an "approximation, representation, or idealization of selected aspects of the structure, behavior, operation, or other characteristics of a real-world process, concept or system" [26]. Simulation can be thought of to be the actual execution of a model with set inputs or controls. Modeling and simulation (M&S) have a long history in scientific research. M&S provides focused analysis of CI systems, which can and have been used for assessing security properties of CI [4]. Modeling exercises in CIP usually have a pre-determined objective such as assessing the potential impact of an attack, or evaluating the susceptibility of a CI system to certain vulnerabilities. Although there are numerous modeling objectives that are pursued in the CIP field, two themes in particular have emerged: resiliency and inter-dependency. Resiliency approaches seek to assess the robustness of a system in response to attacks. Inter-dependency refers to modeling efforts that seek to understand the relationship between CI sectors/systems and how perturbations or failings in one system can have cascading effects on others. Within those two objective paradigms, there are various modeling techniques that are applied. In recent survey papers, these modeling approaches have been partitioned into six categories: agent-based, system dynamics, empirical, network-based, economic, and others (techniques that do not fall in the previous five categories) [4, 67, 77]. This survey paper will attempt to review and organize CIP resiliency and inter-dependency M&S research efforts using the above-mentioned taxonomic system.

4.1.2 Empirical

Empirical methods, in the context of this paper, are modeling efforts that primarily use empirical (usually historical) data to create models for inter-dependencies of CI systems, especially cascading failures.. For CIP, empirical data that demonstrates significant inter-dependency or resiliency issues usually appear after CI failures. Generally, empirical data collection is a difficult task and incorporates several different qualitative and quantitative collection approaches. An example of this type of research includes [10], where authors sought to understand the interconnected failings of the greater Orange County, Florida utilities and transportation systems during the 2004 Hurricane Season. In this study, detailed field interviews were used to model the cascading failures in multiple CI sectors like energy and communications. Similarly, [86] examines CI inter-dependencies and associated disruptions that occurred during and after the 9/11 attacks. Surveys like [60], which detail different historical critical infrastructure attacks also fall into this category due to their *a posteriori* perspective. Since historical evidence and data are applied to modeling efforts, a key result of most empirical research is a retrospective "lessons learned" in which a summary of historical trends, patterns, or behaviors is identified to inform future designs or implementations of CI systems such as in [10]. Empirical analysis can also discover important concepts the can aid other modeling efforts as well. In [54], empirical analysis is used to illustrate deficiencies in other modeling approaches and how incorporating empirical evidence of complex inter-dependencies among CI systems can have beneficial impact in the modeling effort.

Another important note regarding empirical modeling, as it has been defined here, is that it is usually used to inform another modeling process. This is the case in [86] where empirical modeling informed a network-based approach.

Additionally, proposed improvements to CIP can be evaluated retrospectively using historical data. This method is used in [81]. This research focused on analyzing a Supervisory Control and Data Acquisition (SCADA) breach that impacted an Australian wastewater system. The authors who previously had proposed a SCADA security architecture, assessed whether their proposed security model would have prevented the breach. Obviously, empirical approaches to CIP modeling have appeal in the fact that some notion of ground-truth can be established for real-world robustness of CI systems. Compiling, organizing and reflecting on historical data points and experiences are important building blocks for improving systems. A drawback of this approach however is the cost of empirical data gathering. For instance, in [53] a study was undertaken to examine critical infrastructure dependencies in Europe. The effort compiled and analyzed over 2375 CI failures. Of course, the cost for gathering and analyzing these historical data sources is substantial. Even more than that though is the cost of the actual CI failings themselves. Obviously, it is desirable to design, implement and analyze secure CI without having to experience a failure. Although this is clearly impossible, as systems will continue to fail, it is probably desirable to reduce this cost and thus necessitates additional modeling efforts that do not rely solely on CIP failings. Clearly though, this modeling effort does have significant importance to the CIP field as failures occur.

4.1.3 Agent-Based

In large part because of the complexity and inter-dependencies of CI systems, agent-based modeling has been an fast-growing area of research in CIP. The basic idea of agent-based modeling involves simulating the interaction between components or agents of the system in order to observe emergent behavior. Each agent type has differing rules and behaviors that govern its interaction with other agents. The intuition is that the complexity in CI systems, particularly in reference to inter-dependency, has more to do with interactions between various systems rather than the complexity of system components [7]. One research effort uses an agent-based modeling approach to identify inter-dependencies in IT infrastructure [13]. One of the main arguments from this work are that complex phenomenons that occur in CI systems, like cascading failures, are based on the interactions of the different components. Thus even a complete understanding of multiple system's dynamics will not provide a clear picture of how each system is interconnected. Agent-based modeling and associated simulation for CIP provides flexibility to assess how small changes in individual agent types can affect emergent trends of the whole complex system. Sandia National Laboratories developed an agent-based model to using time-dependent Monte Carlo method to predict how perturbations to various critical infrastructures and associated agents, could result in economic impact throughout the country [7].

Agent-based modeling can also be used to address the security implications various policies will have on a system. Building off a game-theoretic model, Hare and Goldstein use an agent-based approach to demonstrate how social networks influence policy cyber-security investment in the defense industry (a CI sector). Like the previous work shows, agent-based modeling can be used to extend existing modeling efforts. Olivia et. al adopt this approach by incorporating agent-based modeling design to an Input-Output Inoperability Model (to be discussed later) in order to characterize inter and intra dependencies in CI systems [66]. Validation in agent-based modeling is a challenge however

4.1.4 Network-Based Models

Network or graph models are another popular approach to modeling interdependent CI systems. Network and graph-theoretic models constitute an extremely large body of work in many scientific fields. In CIP, nodes and edges can be used to model separate infrastructures and their associated dependencies. Guidotti et. al. use this approach in [37]. This work proposes a framework to model dependencies between two infrastructure networks using a graph theoretic approach. A six-part methodology is proposed and applied to assess the impact of cascading failures from a electrical power network to a water distribution network. Network modeling seems to be a natural fit in this context since individual components of a CI system can be connected to components of interdependent systems.

Using a network modeling approach can also be applied to intra-dependencies within a CI system. This is done by Grimsman et. al. who establish an attack-design based on a network model to characterize and assess the potential impacts and corresponding cascading failures of cyber-physical attacks to a water system [36]. Similarly, research proposed in [46] uses a graph-theoretic approach to formulate an "impact analysis framework" for a smart-grid that combines both cyber and physical elements. The spread of a particular disease or virus (biological or technological) can also be modeled by network approaches as in [65]. This method can be particularly informative to quantify the impact of a particular sub-network of the graph failing. According to Ouyang, network modeling for interdependent CI systems falls under two broad categories: topology and flow-based [67]. It appears that topology-based approaches attempt to capture topographical information such as the actual architecture of the nodes and edges in the graph. Flow-based modeling focuses more on the transmission of goods, services, or information between different CI systems (represented by nodes on the graph). Wallace et. al. present this approach for studying cascading effects on CI systems during the 9/11 terrorist attacks [86]. One of the principal applications of this work is to use network modeling to assess, respond to, and restore CI failures that occur due to an attack or disaster. This is done by modeling the CI systems as a network with flows, with various supply and demand nodes giving and receiving service. This service could be provide anything from physical products to electromagnetic signals (i.e. Internet connectivity). Using network flow solutions, a normal operations model can be created. Disruptions and failures in one infrastructure can affect the network model in terms of unmet demand from different nodes. A restoration model can be created once inter-dependencies are identified to prioritize a restoration strategy depending on demand constraints in the network.

4.1.5 System Dynamics

System dynamics (SD) refers to the mathematical modeling of dynamic systems. One of the key principles of SD applications is that complex and dynamic systems rely on “circular, interlocking, and time-delayed relationships among its components” [14]. SD began in the 1950’s with the work of Forrester who used this approach to analyze management dynamics [87]. Feedback and causal loops represent important principles in SD that define how components interact with each other. This approach can be contrasted with bottom-up approach like agent-based modeling where individual agent behaviors are defined and then emergent behaviors are identified as a result. SD instead seeks to define the relationships between components using this top-bottom method. Causal loops and feedback are obviously features of interdependent CI systems, so it is no surprise that this modeling approach has been an area of focus in the CIP community. A SD framework is proposed in [14] to develop a tool to assess serious impacts on interdependent CI systems. This tool would connect different infrastructure models and their associated sub-models to related infrastructures using causal and feedback loops. The Critical Infrastructure Protection Decision Support System (CIP/DSS) project uses SD techniques to "couple separate infrastructures to each other according to their inter-dependencies" [12]. This tool leverages causal loops to develop "consequence models" that are able to characterize the impact one CI system has on another. The causal loops identify inter-dependencies by evaluating the influence certain CI components have on others. In a related research effort Min et. al. use SD techniques in conjunction with functional modeling and nonlinear optimization to propose a framework capable of characterizing inter-dependencies (cyber, physical, geographical or logical) between CI systems. This framework is used to determine physical and economic consequences of critical infrastructure failures. Relationships between the variables in the SD model are defined by differential equations.

A weakness of SD modeling approaches is the vast complexity that can ensue when attempting to describe dynamical behavior of complex CI systems. Modeling only has benefits when the models are simple and understandable while also accurately representing aspects (only aspects) of real-world systems. Careful consideration needs to be made to SD approaches to ensure that models are tractable enough to use.

A related field to system dynamics is control theory. Although both SD and control theory fall under the larger umbrella of dynamical systems, control theory uses similar concepts such as feedback in order to model mechanical and physical processes primarily. Due to rapid technological developments, networked control systems (NCS) make up a large part of the entities that presently control critical infrastructure systems like water, electricity, etc. [75]. These systems are cyber-physical in nature meaning that attacks on the communication layer can impact physical processes. This is another example of cascading failures. For instance, in [17] malware propagation in a corporate network interconnected with a SCADA network is modeled. This work demonstrates how malware could propagate throughout the network and ultimately result in degraded levels of electrical power to customers. In [68], cyber-physical systems under attack are modeled using a linear time-invariant system and attack strategies, including those that are undetectable according to the system model, are formally defined. Although mainly focusing on modeling attacks to a specific target (usually a controlled system), this modeling approach can be extended to better understand cascading failures that can occur among infrastructures connected in a control loop. Consider for example the 911 emergency response system and telephone/Internet communications (Communications/Information Technology Sectors). It could be argued that the state of the 911 system is in large part determined or controlled by the input coming from the telephone system and internet (voice calls being transmitted via voice-over-IP, public-switched telephone networks, or cellular data towers). Identifying the processes that govern this control loop could help inform stakeholders regarding the impacts of potential attacks to either infrastructure component.

4.1.6 Economic Models

Since critical infrastructure sectors make up much of the entities that enable the economy to properly function, modeling that reflects the inter-dependencies of CI systems in terms of economics has a large body of research. The first effort in this area is largely recognized as Wassily Leontief's input-output economic model [67]. The model, analyzed in Leontief's work "Input-Output Economics", seeks to characterize the relationships between different sectors of the economy. The basic equation of Leontief's model is:

$$x_i = \{\sum_j a_{ij}xj + c_i\}\forall i \tag{4.1}$$

where x_i is the total output of production for an industry i, or in the context of this paper a CI sector. a_{ij} represent the input ratio of industry i to industry j, and c_i is end-user demand for industry i. This model helps to relate the interdependent economic nature of various CI sectors. This type of model was used in [76] to provide a framework for relating cybersecurity metrics to economic results in critical infrastructure sectors. This research used that framework to illustrate the potential economic impact of a cyber-attack in the oil sector. One of the main objectives of this modeling approach was to demonstrate how cyber-security disruptions in CI systems (like a SCADA system in the oil sector) can impact other interdependent CI sectors. Economic modeling efforts are appealing because they demonstrate the potential impacts of disasters and attacks in terms of economic metrics. Besides the loss of, or injury to human life, economic incentive would seem to be one of the greatest motivators for stakeholders to invest in CI security measures.

4.2 Modeling for Emergency Dispatch

Starting in the late 1960's and continuing into the 1980's there were several research efforts into constructing useful models for police, fire, and ambulance services [35]. These efforts usually related to the efficient allocation and distracting of emergency resources. One of the most notable of these efforts was the work of Larson in the development of the hypercube queuing model for determining the proper location and distracting for emergency services in an urban setting [47]. The hypercube model has become a popular tool for redistricting and facility location problems and extensions and applications of the hypercube model can be found in [11], [39], and [24]. Additional work in police car allocation that incorporated an $M/M/C$ based queue with priority classes was developed in [15] and further developed for multi-car dispatch in [34]. Simpler models like the square-root model for fire-department response developed in [45] were also influential in the New York Fire Department for allocation design. Most of the work in Operations Research with regards to emergency response has been with the purpose of optimizing resource allocation, not necessarily having to due with impact-analysis.

4.3 911 Related Research

4.3.1 Broad 911 Metrics/Reports

The National 911 Program, a government organization in charge of facilitating federal efforts to support 911 systems, releases yearly reports on broad 911 metrics for the country. The latest report in 2017 published 911 call volume for US states and territories [72]. This is one of the only nationally aggregated reports for 911 in the country. Various news outlets have covered deficiencies with the 911 system in various locales however. In [23] evidence is shown that shortages in 911 call-takers is leading to increased response time. Articles detailing slow 911 response time can be found in various news agency reports such as [29, 82].

4.3.2 Predicting 911 Call Volume

Predicting the volume of 911 calls has been an area of active research in the past twenty years. In [20] dispatch data from Portland, Oregon is utilized for hot-spot analysis. After identifying geographical areas with high caller-volume, this study used a regression-based model to predict the volume of 911 calls at a particular time. Researchers used traffic behavior to build a detection tool capable of identifying abnormal trends in call data. This work was mainly targeting identifying large-scale emergencies through increased call volume [42]. Researchers in [16] use New York City call data to train a machine-learning-based model to predict call volume and identify anomalous events in 911 call data. In a security-focused study, researchers showed how a small botnet of infected cell-phones could potentially generate a strong enough distributed denial-of-service (DDOS) to block emergency services state-wide for days [38]. The cited study has particular relevance to this research since botnets are an obvious threat-vector to 911 systems and result in increased inter-arrival times between calls.

4.3.3 911 Traffic Analysis

There has been some research into 911 call traffic and other 911 process modeling. Modeling the distribution of the inter-arrival of 911 calls to PSAPs was the subject of work in [6]. They show the fits of several different distributions to the inter-arrival times of calls arriving at a PSAP. They also show how a Weibull distribution best fits the inter-arrival time samples they collected from one PSAP in Germany. This work differs from the previous citation since I am looking at inter-arrival times of multiple 911 events, not just calls to a PSAP. Also, I am using multiple datasets from several different agencies throughout the United States. In [33], 911 call inter-arrival and service times throughout the 911 process are analyzed. This publication states that inter-arrival 911 calls fit best with a Beta distribution, however their methodology is not fully explained so it is difficult to determine what other distributions were tested and how. Additionally, it appears that the work only focused on EMS data, although they may have included all 911 calls in their inter-arrival time analysis. In contrast with this work, inter-arrival times of different 911 events are not analyzed and 911 data only comes from one location. Spatial analysis of 911 calls in a region of Tennessee was analyzed in [9]. Police response times from 40 police departments throughout the United States was used to create models of optimal-coverage in [8]. A series of recent blog posts detailed research efforts for compiling a nation-wide 911 dispatch record database [84]. These researchers were particularly focused on police activities but some of the datasets overlap with this research. These researchers also mainly focused on understanding demographics of 911 calls and their association to police response. A full survey of 911-specific research, particularly for policing, can be found in [64].

To the best of my knowledge, this is the first work that has analyzed the inter-arrival times of *all* the notable events in the 911 process (not just call inter-arrival times). In addition, there does not appear to be any study that utilizes data from different agencies (Fire, Police, EMS) and locations to assess distribution fits across environments. It my hope that the results of this research will help forward further study into the rich publicly available

911 datasets to help produce more general models capable of adapting to different locations and agencies.

Chapter 5

Gathering 911 System Observations

The first step in the input modeling process is to gather historical data from the system in study. I wanted to focus on gathering a robust set of publicly available 911 data. The reasoning for this was to make further analysis and reproducibility of my experiments convenient for researchers. Furthermore, I wanted to ascertain the quality of research that could be performed on publicly available datasets without a primary agency sponsor. This chapter will describe those data-gathering efforts and detail the publicly available datasets studied in this research. The methodology for processing those datasets, including synthesizing different datset attributes into a comprehensive 911 event model, will also be set forth.

A challenge with research in emergency response is aptly expressed by Green and Kolesar: "it is very difficult, if not impossible, to do without the sponsorship of a client organization" [35]. I initially sought out an agency partner, the Utah Department of Public Safety, who communicated they would be willing to participate in a data exchange agreement. However, due to the COVID-19 pandemic and the timing of this research, this collaboration was delayed. Instead I sought out publicly available 911 datasets from across the United States that could aid in input modeling, but also enable comparisons across different agencies and emergency service types (i.e. police, fire, etc.). Although the sentiment of Green and Kolesar are still in part true today, the advent of technology, in particular big-data platforms, has provided an additional avenue for research. As part of what seems to be an initiative from several state and local governments, large public-safety datasets are being made available through partnerships with well-established data pipeline/hosting companies like ArcGIS and Socrata or open-source data-portals like CKAN. These data-portals not only include public-safety information (i.e. 911 dispatch records) but also economic, health, and social datasets as well.[1]

Still, manually searching for adequate 911 datasets was a tedious process as there are only a limited number of agencies openly publishing such data. In [84] the researchers compiled 911 data from across five different police departments. However, the main goal of this research was to profile the demographic distribution of different 911 call types. Thus, many of these datasets did not have timestamps needed to look at inter-arrival processes and event-intervals. Furthermore, this data-gathering effort did not extend to fire departments or emergency medical services, both critical components to the overall 911 system. I will now detail the methodology undertaken to obtain and process the datasets used in this research.

5.1 Collection Methodology

The data-gathering process occurred in three steps:

1. Internet searches for any publicly available datasets relating to the 911 process.
2. Assessing potential datasets for desired properties like recorded timestamps of events and filtering out non-911 events like officer-initiated events.
3. Writing data retrieval scripts for automated data polling

Internet searches across different agencies' data-portals constituted the initial collection effort. It quickly became clear that although many agencies published 911-related data, there were only a select few that had detailed timestamps for events throughout the process. Since the these timestamps were necessary for this research, those that did not have at least two timestamp measurements were omitted.

It also became apparent that updates to the datasets differed by agency. For example, Tempe's Advanced Life Support (ALS) incidents had dispatch/response datasets partitioned

[1] For an example of such a data-portal see [18].

Table 5.1: Jurisdiction and Emergency Service Details for Collected Datasets

Jurisdiction	Estimated Population	Service Type(s)	Dataset Time Span
New Orleans, LA	391,006	Police	Jan 2019-Feb 2020
Detroit, MI	672,662	Police	Jan 2019-Jun 2019
Richmond, CA	110,146	Police	Jan 2014-Nov 2018
Boulder, CO	107,353	Fire and EMS	Jan 2015-Apr 2020
New York, NY	8,399,000	Fire and EMS	Jan 2019-Jan 2020
Cincinnati, OH	302,605	Police	Jan 2019-Jan 2020
San Francisco, CA	883,305	Fire and EMS	Jan 2019-Jan 2020
Tempe, AZ	192,364	Fire (ALS)[3]	Jan 2019-Dec 2018

yearly, with data available for the years 2012-2018. In contrast, San Francisco's Fire Department provides a consistently updated, single dataset (new entries added every day) that spans from 2000 to the present day.

It became important to limit the time spans of the datasets we studied, in order to provide somewhat of a control. It was determined that roughly the latest full-year of data available for each chosen jurisdiction would be studied. Exceptions to this measure was in the case of smaller incident datasets (i.e. Boulder, CO Fire& Rescue and Richmond,CA Police) where the amount of incident samples was minute compared to the other datasets. These datasets were augmented with additional years of data in order to boost the number of samples and thus compare more evenly across agencies. Additionally, due to a processing error only ~ 6 months of data was analyzed from the Detroit Police Department.

Given that a number of these datasets were hosted using reputable data-analysis platforms like Socrata, a data-retrieval script was written to automate gathering the selected datasets within the chosen time-windows. The format for these datasets was chosen to be comma-separated value (CSV) files due to the widespread support in data-science software packages like Dask and Pandas [22, 83]. It should be noted however that additional geographic file formats like GeoJSON and shapefiles were also collected for future use in studying geographical properties of the data.[2].

[2]Dataset retrieval, processing and analysis code and dataset links used for this research can be accessed via `https://github.com/bmoss6/911_modeling`

Ultimately, nine datasets, across eight US cities, as shown in table 5.1 were collected. The estimated population of the serviced jurisdiction, the source agency of the 911 dispatch records, and time span of the datasets are also detailed. It should be noted that New York City divided fire and EMS incidents into separate datasets. Additionally, Tempe's dataset was limited to only incidents labeled with an advanced life-support code.

5.2 Processing Methodology

Although each raw dataset contained timestamps with the same underlying semantics, correlating those various timestamps and their corresponding events was a significant undertaking. In addition, three out of the nine datasets did not have any further documentation regarding definitions of their fields. For these datasets, assumptions on the attribute semantics were made using the field name, sample values, and similarity to other datasets. In order to holistically analyze the overall data, a comprehensive sequential event model of the 911 response process that sufficiently captured all the various events recorded by the agencies was developed.

5.2.1 911 Event Model

As mentioned previously, the 911 response process involves many coordinated components. Even with a landline phone call to 911 (the most basic and increasingly outdated method), multiple components like the master street address guide (MSAG) and local telephone exchange carrier, work together to ensure that the call is routed to the geographically appropriate PSAP. From there, a call-taker answers the call and records the incident details in the computer-aided dispatch (CAD) system. They then coordinate with an agency dispatcher who assigns the appropriate emergency response units (police, fire, EMS) to the scene of the incident. The response units must then acknowledge the assignment, travel to the scene of the incident and perform whatever service is necessary. In many cases, especially when there are urgent medical needs at the scene, hospital transport is required.

Clearly, there are a plethora of events that transpire within the time an individual dials 911 to when emergency services consider the incident closed. However, since many of these events cannot be feasibly recorded by agencies (i.e. the time a caller actually dialed 911), the proposed event model only includes events that were actually measured in one or more of the datasets, with the exception of the "Call Placed" event. The timestamps of these events is assumed to be recorded automatically by the CAD systems at the PSAPs and MTU's in the various emergency response vehicles. We now describe in detail the events of the model illustrated in figure 5.1.

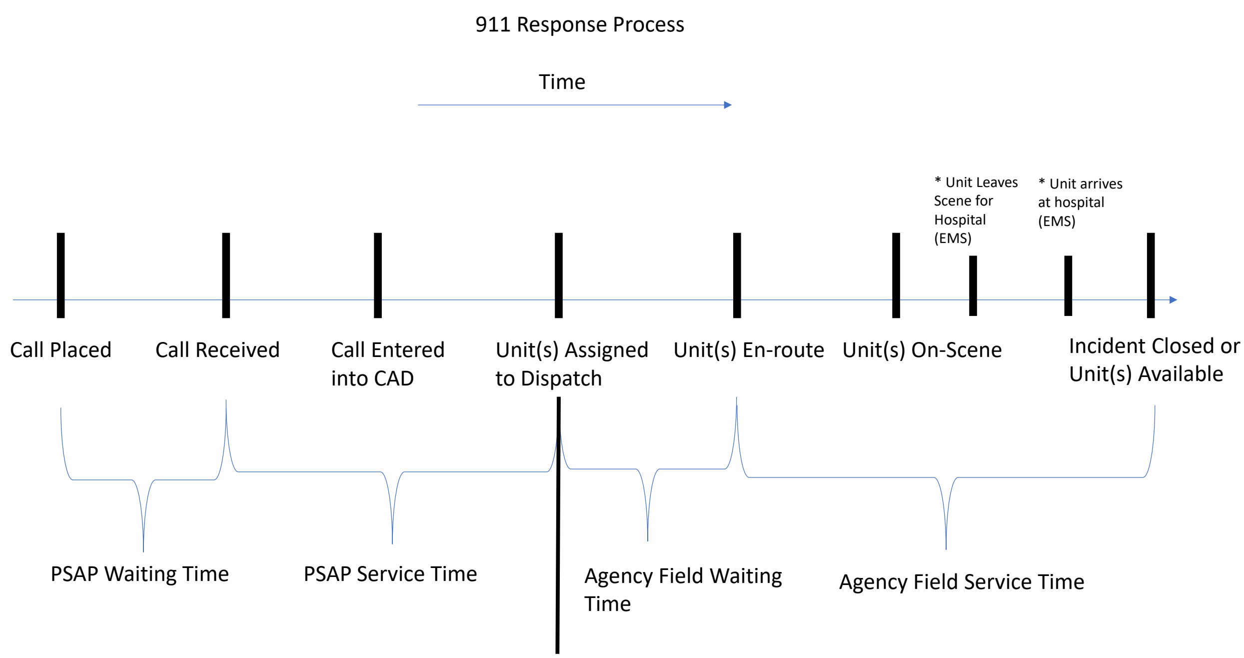

Figure 5.1: Event Model of End-to-End 911 Process

As stated previously, none of the datasets record the exact time of when a call is placed to 911. This is understandable, as it is outside the knowledge of the PSAP. However it is important to include in our model because this is where the 911 process starts (at least from the caller's perspective). The first event that is recorded in some of the datsets is when the call is received by the call-taker at the PSAP or dispatch center. Based on the descriptions of this field from several of the dataset's attribute dictionaries, this event can be interpreted as when either: 1) the call-taker answers the call and begins interaction with the caller or 2) the call is received into the automatic call distribution (ACD) system at the PSAP.

The next timestamped event occurs when the call-taker enters the incident information into the CAD system. This is the first event in the 911 process that is recorded in the datasets that do not record a "Call Received" event.

The next event recorded is when the call-taker or downstream dispatcher assigns one or multiple units to dispatch to the scene of the incident. Several datasets only record the timestamp of when the first unit was assigned to dispatch. This could be because jurisdictions are mainly interested in the elapsed time between when the first unit is assigned and when the first unit arrives to a scene, even if those units are not necessarily the same (i.e. unit #1 is assigned first but unit #2 arrives on-scene first). To standardize across the collected data, we modified datasets where multiple units' events were recorded to only assess the first sequential unit's timestamps.

At this point in the process, the measurements are most likely being recorded from MTUs in the assigned unit's vehicle. The participants in the process at this point also shift from the call-taker/dispatcher, to the first responders. The next event measured is the time the unit acknowledges they are en-route to the incident. Following this event, the next timestamp is when the unit arrives at the scene of the incident. In more detailed datasets like San Francisco Fire Department and New York City EMS, timestamps for when units (usually ambulances) depart and arrive at a hospital for incidents requiring further medical treatment are also recorded.

The final event in the model is when an incident is closed in the CAD system. This of course may not necessarily mean that the actual incident has been resolved, but most likely serves to indicate that the resources associated with the dispatch are available for a new CFS.

5.2.2 Mapping Dataset Attributes to Event Model

As mentioned previously, it became clear from the datasets that event recording protocols differed widely across agencies. Furthermore, each agency had its own unique naming conventions for different events. This illustrates the standardization problems raised in [84]. Since each CAD system recording this data was most likely tuned for local deployment in a specific jurisdiction and no centralized body exists to regulate naming protocols, correlating events across datasets was performed manually. In some instances, feature names that appeared semantically the same actually had different meanings and assumptions had to be made about how to map the specific dataset feature to the proposed comprehensive event model. A particularly confusing example of this is the assigned-to-dispatch (ATD) event and the en-route event. The New Orleans dataset includes a feature titled "TimeDispatch" which they define as the "entered time by Orleans Parish Communications Department or New Orleans Police Department when an officer was dispatched". It would seem that this field would most likely correspond to an ATD event. However further examination of the dataset values indicated otherwise. Table 5.2 shows a few incidents from this data set. In these events, the large time gap between the "TimeCreate" and "TimeDispatch" fields and the correspondingly short elapsed time between the "TimeDispatch" and "TimeArrival" fields indicate that the "TimeDispatch" field more appropriately corresponds to when the dispatched unit was considered en-route to the scene.

In constrast, San Francisco's Fire Department dataset included a field titled "Dispatch DtTm" which was defined as the "date and time the 911 operator dispatches this unit to the call." Also included in this dataset is a field titled "Response DtTm" defined as the "date and time [the] unit acknowledges the dispatch and records that the unit is en-route to the

Table 5.2: Sample Dispatch Data: New Orleans Police

NOPD_Item	TypeText	TimeCreate	TimeDispatch	TimeArrival	TimeClosed
E2961519	HIT & RUN	5/21/2019 9:42	5/21/2019 10:35	5/21/2019 11:01	5/21/2019 12:54
E3844619	HIT & RUN	5/27/2019 10:48	5/27/2019 12:03	5/27/2019 12:10	5/27/2019 13:27
E3879319	SIMPLE BURGLARY VEHICLE	5/27/2019 16:59	5/28/2019 0:04	5/28/2019 1:02	5/28/2019 2:48

Table 5.3: Sample Dispatch Data: Detroit Police

Call Description	Call Time	Intake Time (Minutes)	Dispatch Time (Minutes)	Travel Time (Minutes)	Time On Scene (Minutes)
PANIC / DURESS ALARM	2019-01-01 00:31:22	2.1	9	7.2	0
SUICIDE THREAT	2019-01-01 00:31:43	3.2	20.4	0	20.2

location of the incident." The presence of these fields indicate that the "Dispatch DtTm" field corresponds to the ATD event in the proposed model while the "Reponse DtTm" falls within the scope of the en-route event.

The actual format of timestamps was another issue in data processing. For example, the Detroit Police dataset only recorded the elapsed time of different intervals in the response process as seen in table 5.3. To handle this case, the actual timestamps of the events in the sequential model were computed by using each time interval as a time-delta added to the initial "Call Time" timestamp.

5.2.3 Limitations of the Datasets and Sequential Event Model

Data incongruities like those mentioned above meant that certain assumptions had to be made about how to map each dataset's timestamped features into the proposed sequential model. In some cases, this simplification may not truly reflect the dynamics of every particular incident. However, from thorough review of the literature and collected datasets, I believe that the event model is broad enough to reflect the essential sequential events that characterize the 911 response process. In addition, as seen in table 5.4, not every event was recorded by agencies. This also could impact the analysis by biasing input-modeling for certain events towards the agencies that recorded those events.

Table 5.4: Recorded Events by Agency Dataset (blank cell indicates event was not recorded)

Dataset	Call Received Event	Entered to CAD Event	First Unit Assigned Event	First Unit Enroute Event
Boulder Fire Rescue	X		X	X
Cincinnati Police		X		X
Detroit Police	X	X		X
New Orleans Police		X		X
New York EMS		X	X	X
New York Fire		X	X	X
Richmond Police		X		X
Tempe ALS	X		X	X
San Francisco Fire	X	X	X	X
Dataset	First Unit Onscene Event	First Unit Departs to Hospital	First Unit Arrives at Hospital	Incident Closed
Boulder Fire Rescue	X			X
Cincinnati Police	X			X
Detroit Police	X			X
New Orleans Police	X			X
New York EMS	X	X	X	X
New York Fire	X			X
Richmond Police				X
Tempe ALS	X			
San Francisco Fire	X	X	X	X

Chapter 6

Assessing Independence and Stationary of 911 Event Arrival Processes

This chapter will investigate the arrival process of 911 events outlined in the preceding section. I will detail how the arrival rates for each 911 event exhibit non-stationary behavior depending on the hour of the day. I argue that because of this non-stationary behavior, analyzing inter-arrival times across an entire day would be ineffective and that instead assessing the distributions of inter-arrival times during hourly segments is more appropriate.

6.1 Changes in 911 Event Occurrences by Hour

Understanding whether a system as a non-homogeneous arrival process is critical in input modeling. Fitting data to a theoretical distribution works best if that data forms independent and identically distributed observations. If an arrival process does change with time, then observations over that total time period cannot be independent. This implies that one probability distribution will not appropriately fit to all of the samples across time [49]. Although this chapter is analyzing arrival counts of different 911 events, non-homogeneous arrival processes have obvious implications for inter-arrival times. That is, if the arrival process changes throghout time, the samples of inter-arrival times will be non-stationary/non-independent as well. Since events occurring throughout the 911 process are dependent on customers calling for emergencies, I hypothesize that, depending on the hour of the day, there are varying levels of intensities of the number of 911 events that occur.

This hypothesis can be confirmed graphically by looking at average event counts per hour of the day for each of the events considered in the response process model. Consider for example figure 6.1, showing the average amounts of "Call Received" events for the San Francisco Fire Department. Clearly peaking at around noon on most days, there is obvious non-stationary behavior. This implies that there are more calls for fire and medical emergencies during the afternoon-evening period then during the morning hours. This would seem rather intuitive, since more people being awake most likely corresponds to increases for emergency services.

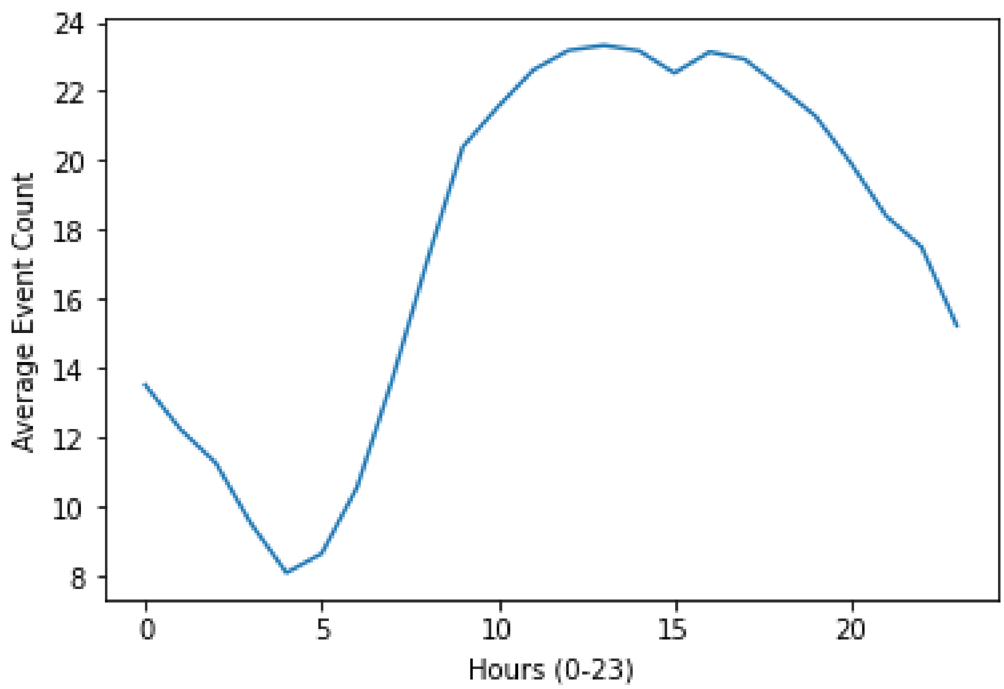

Figure 6.1: Number of Calls Received on Average, Per Hour of the Day (San Francisco Fire Department)

As seen in table 6.1, this global peak of events during the second half of the day (12PM-12AM) can be observed for nearly every agency's dataset, indicating that many agencies experience similar hourly trends. However, instead of a smooth cycle, many events have many local peaks and valleys. For example consider figure 6.2, showing the average number of "Incident Closed" events occurring per hour for the Richmond Police dataset. Although the global minimum occurs at 7AM, there is a notable drop at 3PM as well.

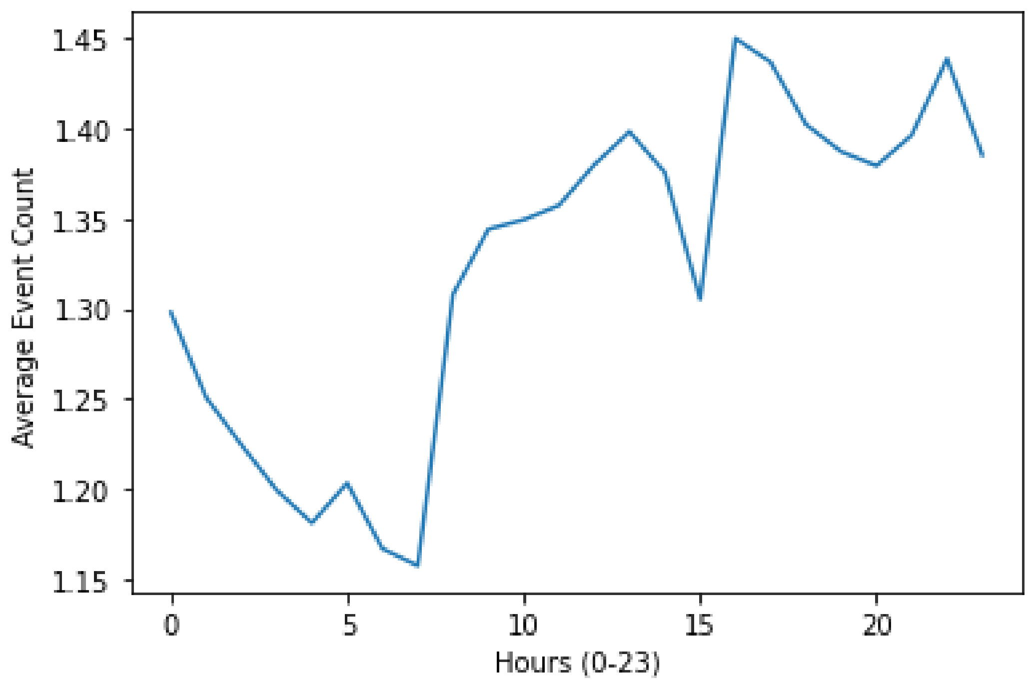

Figure 6.2: Average Number of Incident Closed Events, Per Hour of the Day (Richmond Police Department)

6.2 Auto-Correlation of 911 Events

Another clear indication of non-independence can be seen in the auto-correlation graphs for hourly event counts. Auto-correlation is a measure of the degree a time-series variable is correlated with a delayed copy of itself. Formally the auto-correlation function can be defined as:

$$r_k = \frac{\sum_{i=1}^{N-k}(Y_i - \bar{Y})(Y_{i+k} - \bar{Y})}{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}$$

Where Y_i is the time-series observations at time i,and $i+k$ is the lag value. There are different methods for calculating auto-correlation and I used the "statsmodels" auto-correlation function to plot these graphs [79]. The results indicate clear periodicity occurring throughout 24-hours in the datasets. For example, figure 6.4 shows the auto-correlation plot for hourly counts of the "Call Entered to CAD" event. The shaded blue region indicates the 95% confidence interval that the auto-correlation was caused by chance. This same pattern follows almost identically for all but two of the datasets, clearly indicating that in large part the

Table 6.1: Hours of Maximum and Minimum Average Event Counts Per Dataset

Dataset	Call Received (Max/Min)	Entered to CAD	First Unit Assigned	First Unit Enroute
Boulder Fire Rescue	12AM/11PM		12AM/11PM	12AM/11PM
Cincinnati Police		2PM/5AM		2PM/5AM
Detroit Police	9PM/10AM	9PM/10AM		9PM/10AM
New Orleans Police		5PM/5AM		3PM/6AM
New York EMS		3PM/5AM	4PM/5AM	4PM/5AM
New York Fire		5PM/4AM	5PM/4AM	5AM/4PM
Richmond Police		9PM/6AM		10PM/6AM
Tempe ALS	7PM/5AM		7PM/5AM	4PM/5AM
San Francisco Fire	1PM/4AM	1PM/4AM	1PM/4AM	1PM/4AM
Dataset	First Unit Onscene	First Unit Departs to Hospital	First Unit Arrives at Hospital	Incident Closed
Boulder Fire Rescue	12AM/11PM			1AM/1PM
Cincinnati Police	2PM/5AM			6PM/5AM
Detroit Police	9PM/10AM			10PM/11AM
New Orleans Police	4PM/6AM			5PM/7AM
New York EMS	4PM/5AM	1PM/5AM	1PM/6AM	4PM/6AM
New York Fire	6PM/5AM			6PM/5AM
Richmond Police				4PM/7AM
Tempe ALS	4PM/5AM			7PM/5AM
San Francisco Fire	1PM/4AM	2PM/5AM	1PM/4AM	3PM/5AM

number of 911 events that occur during an hour T are positively correlated with hours between T and $T- \sim 5$, negatively correlated with hours between $T- \sim 6$ and $T- \sim 17$, and finally positively correlated with hours between $T- \sim 19$ and $T- \sim 24$.

6.3 Independence/Stationarity of Hourly-Partitioned Data

Given these results, I assume that in general, 911 event rates occur differently depending on the time of the day. From this knowledge it is clear that fitting one probability distribution to the entire time-span of the data would not be appropriate. It is for this reason that I argue for partitioning the dataset by hour intervals and performing input-modeling on those hour intervals. I reason that although there is clear periodicity when looking at the data across a 24-hour interval, examining hourly partitions of the data samples would will provide for relatively stationary and independent samples. The intuition behind this reasoning is that although rates of events may differ depending on the hour of the day, within that hour (i.e. from minute to minute) they are unlikely to exhibit the same trend. For example, it is unlikely that events would have a significantly more or less intense rate of occurrence at minutes 30-35 during a particular hour compared to minutes 10-15 at that same hour. This hypothesis is supported by examining the average events occurring per minute as seen, for

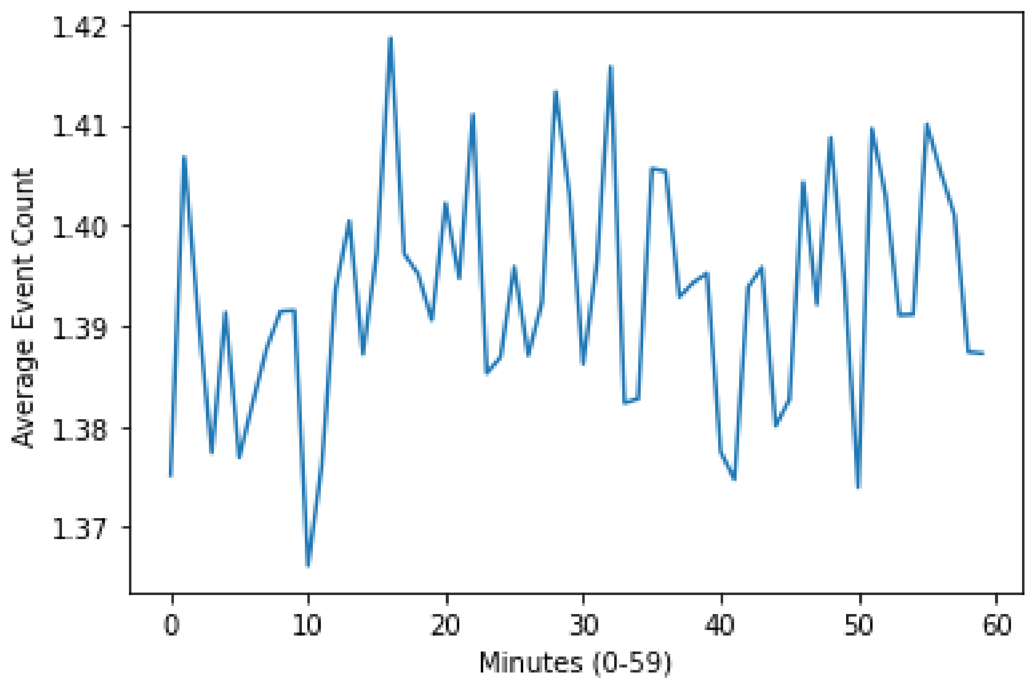

Figure 6.3: Average "Call Entered to CAD" events per minute (New Orleans Police Department)

example, in figure 6.3 for the "Entered to CAD" event for the New Orleans Police Department. In this figure although there are many spikes, the mean between intervals is relatively similar and there is no obvious correlation between one minute's event count and previous minute's. The vast majority of each dataset's minute event counts exhibit the same behavior. Choosing hourly partitions, is also reinforced by table 6.2 where all but one dataset exhibit higher on-average auto-correlation across all 911 events when examining event counts at the hour granularity compared to minute granularity.

Having established why performing input-modeling is more appropriate on hourly partitions of the 911 data, in the next chapter I proceed with step two of the process: identifying candidate distributions.

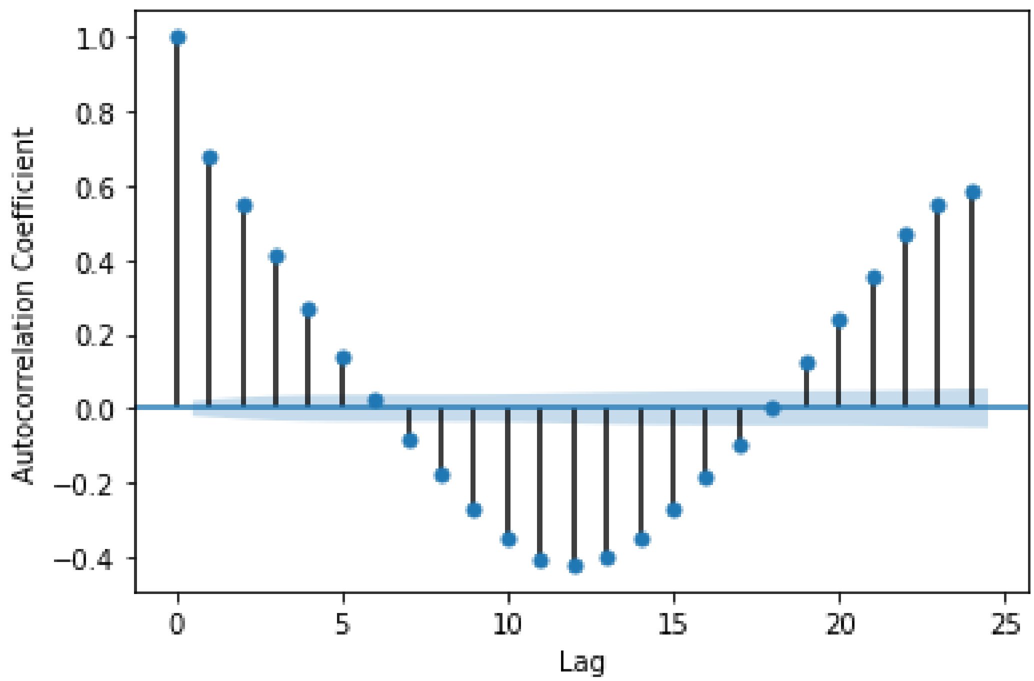

Figure 6.4: Auto-Correlation Plot for event (New Orleans Police Department)

Table 6.2: Average Autocorrelation Comparisons (Minutes vs Hours)

Dataset	Avg Minute Autocorrelation	Avg Hour Autocorrelation
Boulder Fire Rescue	**.017**	.060
Cincinnati Police	**.062**	.090
Detroit Police	**.039**	.161
New Orleans Police	**.038**	.084
New York EMS	.160	**.050**
New York Fire	**.100**	.128
Richmond Police	**.017**	.048
Tempe ALS	**.018**	.038
San Francisco Fire	**.024**	.052

Chapter 7

Identifying Candidate Distribution Families for Inter-Arrival Times

Now that it has been decided to analyze the 911 events according to hourly partitions and by so-doing provide more stationary/independent samples, I proceed with identifying candidate distributions for the inter-arrival times of the 911 events.

7.1 Summary Statistic Coefficient of Variation

As detailed in [49], "some distributions are characterized at least partially by functions of their true parameters." When samples are independent and identically distributed (iid), the sample estimates of these functions can be helpful in identifying the appropriate underlying distribution. The *coefficient of variation* (CV) is an important statistic when seeking to determine whether inter-arrival times can be modeled by an exponential distribution. This statistic describes the ratio of the standard deviation to the mean. The sample CV can be calculated by

$$\hat{cv}(n) = \frac{\sqrt{S^2(n)}}{\bar{X}(n)}$$

Where $\sqrt{(S^2(n)}$ is the sample standard deviation and $\bar{X}(n)$ is the sample mean.

The exponential distribution has a theoretical CV of 1. Thus if our events have sample CV's ~ 1 then could suggest the exponential distribution is a good candidate for the data. The sample CV's were computed for each hour partition for each 911 event across the datasets. The average CV values for 911 events and the variance of those average CV values are shown in table 7.1. Every 911 event appears to have CVs ~ 1 with a relatively

Table 7.1: Average CV Values and CV Variance for 911 events

911 Event	Average CV Values	Average Variance of CV Values
Call Received	1.03	.006
Entered to CAD	1.04	.004
First Unit Assigned	1.02	.006
First Unit Enroute	1.14	.372
First Unit On-Scene	1.13	.329
First Unit To Hospital	1.01	.001
First Unit At Hospital	1.00	.001
Incident Closed	1.20	.472

small variance. This suggests that the exponential distribution is indeed a good candidate for modeling the inter-arrival times.

7.2 Summary Statistic: Skewness

Another important summary statistic that is helpful for identifying candidate distributions is skewness. Skewness essentially helps to identify the symmetry of a distribution. For example, right-leaning distributions have a skewness of > 0. The exponential distribution has a theoretical skewness of 2. Sample skewness can be calculated using the Fisher-Pearson coefficient of skewness:

$$g_1 = \frac{m_3}{m_2^{\frac{3}{2}}}$$

where

$$m_i = \frac{1}{N}\sum_{n=1}^{N}(x[n] - \bar{x})^i$$

where N is the sample size and and $\bar{x}$ the sample mean. Average sample skewness values were computed for each 911 event using the same process as was done for CV. Table 7.2 show the results of these computations. These results indicate that the sample skewness for each

Table 7.2: Average Skewness across 911 Events

911 Event	Average Skewness
Call Received	2.39
Entered to CAD	2.82
First Unit Assigned	3.05
First Unit Enroute	5.13
First Unit On-Scene	5.00
First Unit To Hospital	2.71
First Unit At Hospital	2.72
Incident Closed	7.21

911 event is > 1 on average. Several events have an average skewness of ~ 2 as well. This again supports the hypothesis of exponentially distributed inter-arrival times.

7.3 Histograms

Histograms are another approach for graphically assessing candidate distributions. The main decision that must be made when creating a histogram is the number of bins (intervals) to use. Following the suggestion of [49] I compared different bin sizes to determine the best number that would create a smooth histogram. After experimentation, I concluded that using 100 bins created the best histograms across the different datasets. Using this fixed number of bins, the inter-arrival times according to bin intervals were computed for each 911 event across each dataset's hourly partitioned samples.

The resulting histograms were also promising in supporting the exponential distribution as the candidate parameter of choice. For example, see figure 7.1, a histogram of inter-arrival times for an hourly partition of "Call Recieved" events from the Detroit Police Department. All the histograms generated for each event/dataset/hour combination yielded a right-skewed histogram with very apparent exponential visuals.

From both the summary statistics and histograms, it can be argued that the exponential distribution appears to be an appropriate choice for the inter-arrival distributions. I now move to estimating distribution parameters and assessing the fitted theoretical distribution to the sample data.

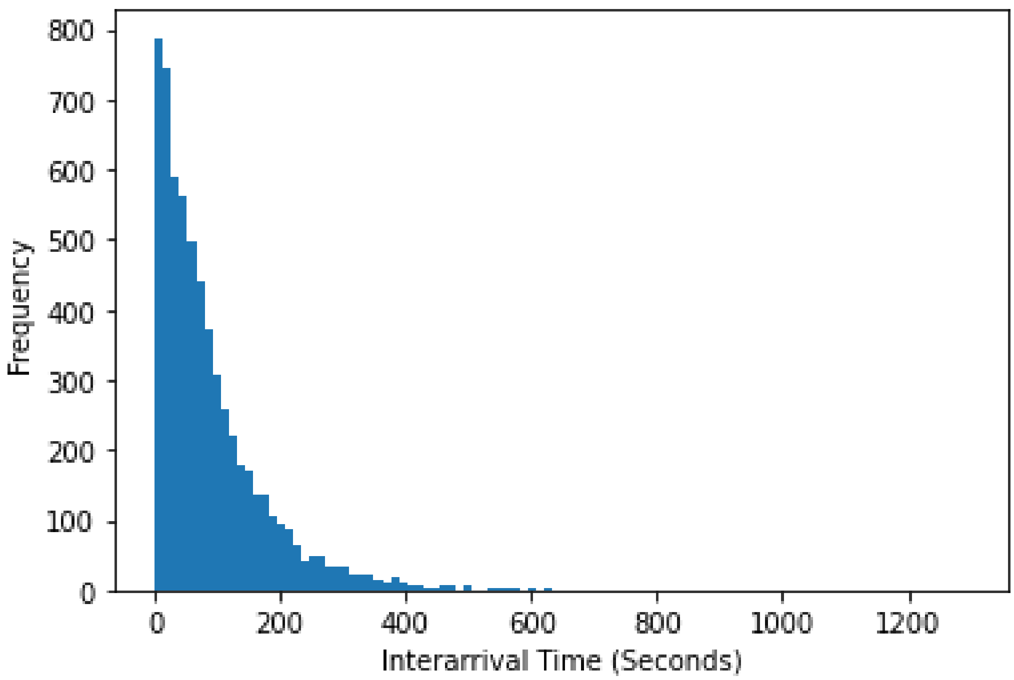

Figure 7.1: Histogram of Inter-arrival times of Call Recieved Event, (Detroit Police:10-10:59PM)

Chapter 8

Estimating Distribution Parameters and Assessing Goodness-of-Fit

I now proceed to fit the exponential distribution to the 911 event/dataset/hour samples by using the exponential distribution's maximum-likelihood-estimator to compute the parameter values for each sample. The goodness of these fits will be assessed using the Chi-Squared test and compared with fits from the Weibull distribution.

8.1 Maximum Likelihood Estimator for Exponential Distribution Parameter

A common method for estimating parameters of a distribution is using *maximum-likelihood estimators* (MLE). MLEs are commonly used for learning parameters from data and have several desirable properties that make them especially useful for statistically evaluating goodness-of-fit [49]. I now derive the MLE for the exponential distribution. Since the exponential distribution is continuous and has only one parameter (λ), the likelihood function can be given as

$$L(\lambda) = f_\lambda(X_1) f_\lambda(X_2) f_\lambda(X_3) .... f_\lambda(X_n) = \lambda^n [e^{-\lambda(X_1+X_2+X_3+...+X_n)}]$$

Where $(X_1, X_2, X_3, ..., X_n)$ are the already observed, independent and identically distributed (iid.) samples of inter-arrival times and

$$f_\lambda(X) = \lambda e^{-\lambda X}$$

or the theoretical density function of the exponential distribution. Solving for the maximum of this function, I take the derivative of the log function (since the argmax of the log function and actual function will be the same) with respect to λ.

$$\begin{aligned} \frac{d}{d\lambda} log(\lambda^n [e^{-\lambda(X_1+X_2+X_3+...+X_n)}]) \\ = \frac{d}{d\lambda} nlog(\lambda) - \lambda(X_1 + X_2 + X_3 + ... + X_n) \\ = n\frac{1}{\lambda} - (X_1 + X_2 + X_3 + ... + X_n) \end{aligned}$$

Setting the derivative equal to zero to solve for the extreme point

$$\begin{aligned} 0 = n\frac{1}{\lambda} - (X_1 + X_2 + X_3 + ... + X_n) \\ (X_1 + X_2 + X_3 + ... + X_n) = n\frac{1}{\lambda} \\ \lambda(X_1 + X_2 + X_3 + ... + X_n) = n \\ \lambda = \frac{n}{(X_1 + X_2 + X_3 + ... + X_n)} \lambda = \frac{1}{\bar{X}} \end{aligned}$$

where $\bar{X}$ is the sample mean.

Now that the estimator is derived, the 911 event/dataset/hour sample groups can be fit and those fits assessed.

8.2 Graphical Goodness-of-Fit Methods

An important technique used to justify whether the fitted distribution appropriately represents that sample data is through graphical procedures. I use two graphical procedures: histogram overplots and quantile-quantile (QQ) plots to assess the fitted exponential distributions to the data samples.

8.2.1 Histogram Overplots

Histogram overplots graphically show how well the expected frequencies generated by the fitted distribution correspond to the histogram bin counts. The expected frequencies were computed by

$$f([x_1, x_2]) = N * (cdf(x_2) - cdf(x_1))$$

where $[x_1, x_2]$ are the respective edges of the bins created by the histogram, *cdf* is the cumulative distribution function for exponential distribution and N is the total number of samples.

Just like in the section 7.3, I use a fixed bin size of 100. The vast majority of histograms, indicate that the fitted distributions did indeed match well with the data samples. See for example the histogram overplot in figure 8.1 demonstrating the fitted exponential distribution on the histogram for an hourly partition of the "First Unit On-Scene Event" from the Cincinannti Police dataset.

8.2.2 Q-Q Plots

Quantile-quantile (QQ) plots are another graphical tool for assessing the accuracy of fitted distributions to samples. QQ plots are created by plotting the values of the sample quantiles against the values of the theoretical quantiles. Theoretical quantile values for a quantile q given an exponential distribution can be defined as x_q where:

$$F(x_q) = q$$

and

$$x_q = F^{-1}(q)$$

where $F(x)$ denotes the CDF of the exponential distribution while $F^{-1}(x)$ is inverse CDF or percent-point function of the exponential distribution. The most commonly identified quantile value is the median or $x_.5$.

QQ plots are helpful for identifying goodness-of-fit because if the fitted theoretical distribution is indeed the same as the true underlying distribution, then the corresponding quantile values will be comparable and the QQ plot will be linear meaning that a $y = x$ line should contain most of the points on the plot [49]. Another benefit of QQ plots is that they will amplify differences between the tails of the sample and fitted distributions. This is important particularly in this analysis because of the large skewness values for some of the events indicated in table 7.2. Although exponential distributions have a skewness value of 2, skewness values > 2 could be detailed by differences between the tails of the sample and fitted distributions.

The "qqplot" function from the "statsmodels" python package was used to create plots for each data sample. Once again as seen in figure 8.1, most of the generated QQ plots indicated the fitted exponential had good fit to the samples. This can be seen by how many of the points follow the fitted regression line. However, it should be noted that in some cases, the QQ plot was less than ideal. As hypothesized, some QQ plots indicated somewhat substantial differences in the tails of the fitted and sample distributions. See for example, figure 8.6, where although the histogram overplot looks somewhat appropriate, the QQ plot shows large deviations between the sample and theoretical quantiles while approaching the tails. As it pertains to the 911 events, the pattern of poorer fits at the tails of the distributions could have an explanation. The exponential distribution is considered *light-tailed*, meaning that the distribution's density function, $pdf(x)$ will approach zero quickly as $x \to \infty$. Although the sample distributions do exhibit many exponential properties, many of these distributions have a few outliers in regards to large inter-arrival times. This can be seen in the histograms of figures 8.1 and 8.6. Although most of the inter-arrival times correspond to the fitted exponential distribution, there are observed inter-arrival times where the probability of that observation, according to the fitted distribution is essentially 0. This likely explains the departure from the fitted regression line in the QQ plots as the quantiles approach the tails of the distributions.

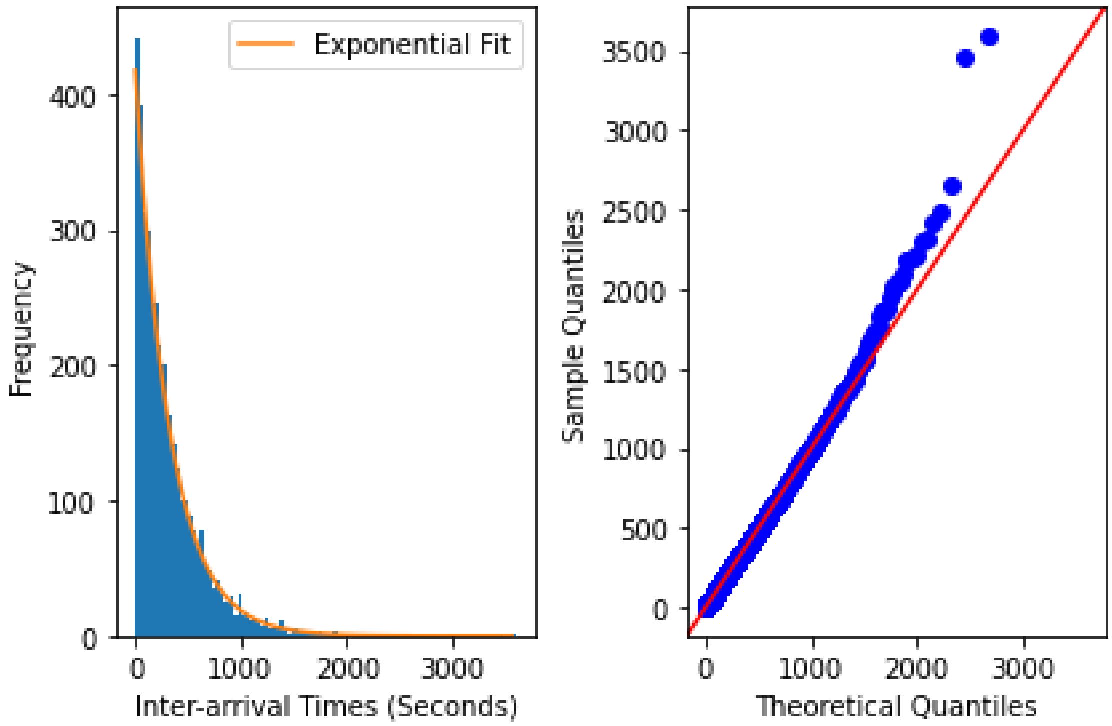

Figure 8.1: Histogram Overplot and Q-Q Plot First Unit On-Scene Event Inter-Arrivals (Cincinnati Police: 5-5:59AM)

8.3 Chi-Squared Goodness-of-Fit

The chi-square test is the oldest goodness-of-fit hypothesis test dating back to the work of Pearson in 1900 [49]. The chi-square test provides a more formal way for evaluating the fit of a theoretical distribution on data than the previous graphical approaches used. Chi-square poses a null hypothesis H_0 that assumes there is no significant difference between the expected values of the fitted probability distribution and the observed values. Although there are many goodness-of-fit tests available and chi-square does have limitations, I chose to use it because of its flexibility in assessing other theoretical distribution's fits in comparison to the exponential fit.

Formally, the chi-square test works as follows as described in [49]:

1. Divide range of fitted distribution into k adjacent intervals $[a_0, a_1), [a_1, a_2), ..., [a_{k-1}, a_k)$.

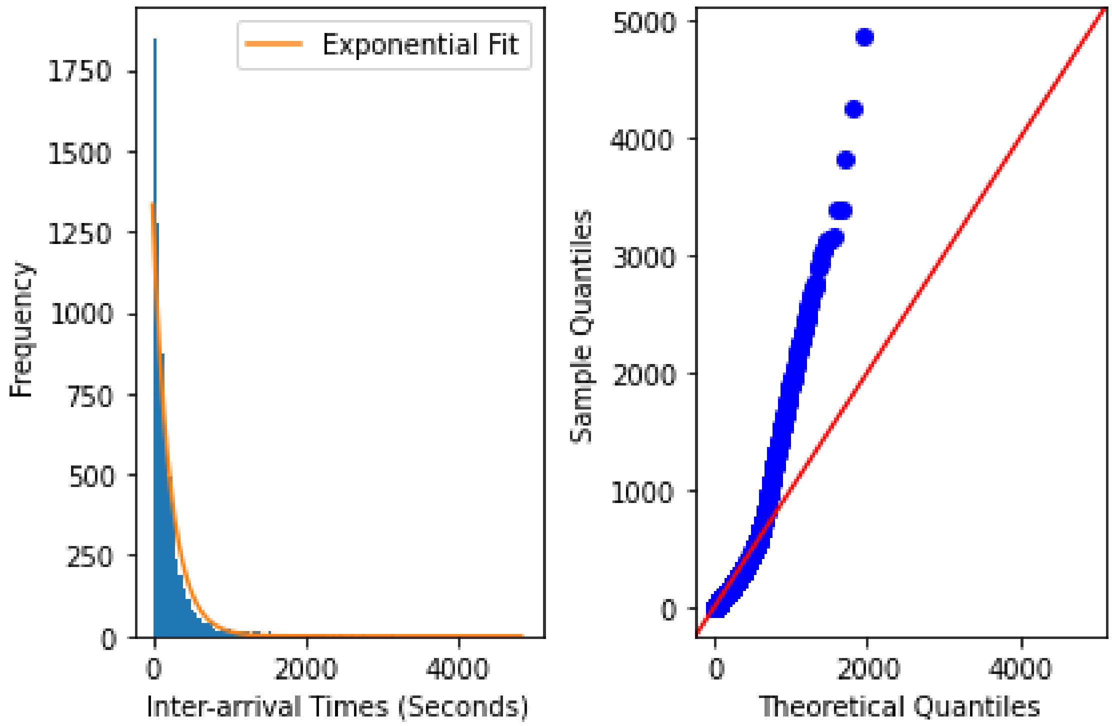

Figure 8.2: Histogram Overplot and Q-Q Plot First Unit On-Scene Event Inter-Arrivals (New Orleans Police: 7-7:59AM)

2. Count the number of X_i's in the j^{th} interval $[a_{j-1}, a_j)$. Denote this as N_j. In this analysis N_j is the number of observations where the inter-arrival time of a 911 event was between the range $[a_{j-1}, a_j)$.

3. Compute the expected proportion p_j which is the probability of observing a value within the interval $[a_{j-1}, a_j)$ according the the CDF of the fitted distribution. Here, p_j is given by:

$$p_j = \int_{a_{j-1}}^{a_j} cdf(x)dx$$

4. Compute the chi-square test statistic χ^2 where:

$$\chi^2 = \sum_{j=1}^{k} \frac{(N_j - np_j)^2}{np_j}$$

8.3.1 Calculating Equiprobable Bins

It turns out that one of the challenges with chi-square is its sensitivity to the binning procedure. It is recommended that in lieu of a definite procedure, that equi-probable bins be created where $np_j \geq 5$ [49]. I used this approach with $k = 20$ intervals such that $p_j = \frac{1}{k} = .05$. Since all of my sample sizes were > 100 this met the criteria of $np_j \geq 5$. The bin edges a_j were computed using:

$$a_j = cdf^{-1}(\frac{j}{20})$$

for $j = 1, 2, ..., 19$ where $cdf^{-1}(x)$ is the inverse cumulative distribution function of the fitted distribution.

8.3.2 Fit Assessment

Given that the test statistic is an error measure similar to sum-square-error, it is obvious that if the fit is appropriate χ^2 will be small and vice-versa if the fit is poor. I use the methodology outlined in [49] to identify the criteria for rejecting H_0. That is given m parameters estimated from the data for the fitted distribution, I reject H_0 if $\chi^2 > \chi^2_{k-1,\alpha}$ where $k = 20$. I evaluate two different alpha values $\alpha = .10$ and $\alpha = .05$, respectively.

The results of the chi-square tests for the fitted exponential distributions were surprising. As tables 8.1 and 8.2 show, the chi-square test rejected the majority of the hour samples. This indicated that for many of the samples across 911 events, there were significant statistical differences between the fitted exponential and sample distribution. To further investigate, I looked at the worst-performing samples in terms of the chi-square statistic. After reviewing the different worst-performing samples from all the datasets, I concluded that the chi-square's sensitivity to large sample sizes played a large role in the null-hypothesis rejection for many of the sample groups. For example, in figure 8.3 although the fitted exponential model appears to represent the data fairly well, the chi-square statistic computed for this sample was $\chi^2 > 1552$, a very high test statistic. The New York Fire dataset was

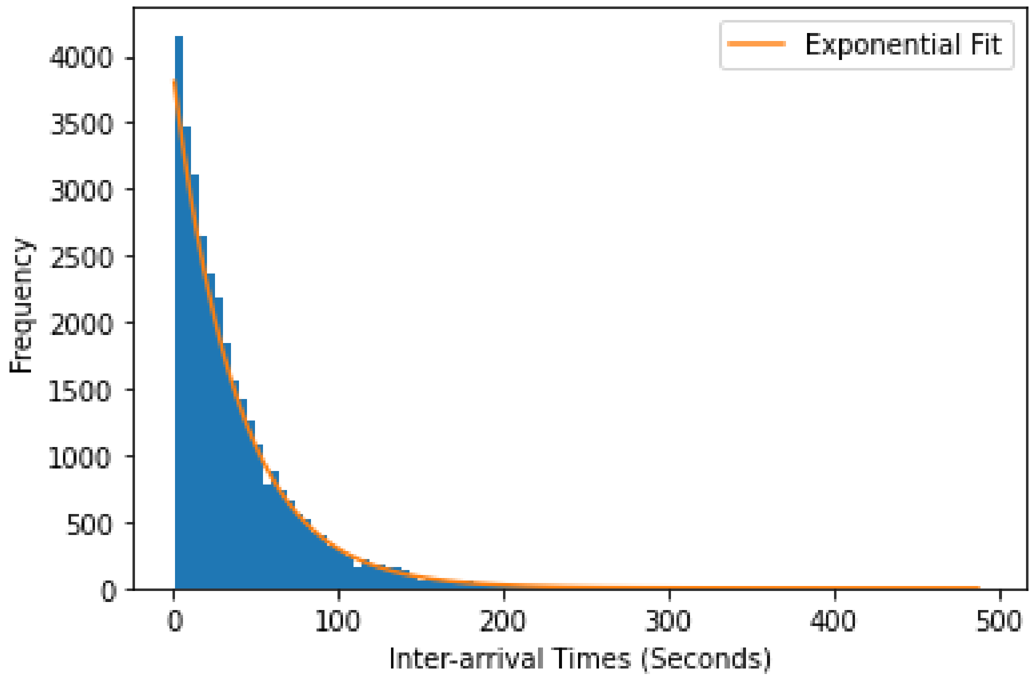

Figure 8.3: Histogram Overplot First Unit Assigned Event Inter-Arrivals (New York Fire: 7-7:59PM)

one of the largest datasets in the study and therefore contained a larger sample sizes for the hourly partitions. As stated in [49], with a large sample size, goodness-of-fit hypothesis tests will "almost always reject H_0" since even small differences between the fitted distribution and sample distributions will be detected for large sample sizes. This is supported by observing in table 8.1 where the least amount of rejections occur. The San Francisco Fire and Richmond Police dataset both perform fairly well across all 911 events. Correspondingly, these data-sets also have much lower sample sizes compared with the New York Fire dataset.

Table 8.1: Number of H_0 Rejections Chi-Square Test ($\alpha = .05$) Exponential Fit

Dataset	Call Received	Entered to CAD	First Unit Assigned	First Unit Enroute	First Unit Onscene	First Unit To Hospital	First Unit At Hospital	Incident Closed	**Dataset Average**
New Orleans Police		24		23	20			24	22.75
Detroit Police	11	6		14	15			12	11.6
Richmond Police		20		17	18			13	17
Boulder Fire Rescue	8		12	12	10			12	10.8
New York Fire		24	24	24	24			24	24
Cincinnati Police		24		23	20			24	22.75
New York EMS		24	24	24	24	24	24	24	24
San Francisco Fire	6	3	24	22	4	1	2	13	9.34
Temple ALS	1		2	3	2				2
Event Average	6.5	17.86	17.20	18	15.22222	12.5	13	18.25	

Table 8.2: Number of H_0 Rejections Chi-Square Test ($\alpha = .10$) Exponential Fit

Dataset	Call Recieved	Entered to CAD	First Unit Assigned	First Unit Enroute	First Unit Onscene	First Unit To Hospital	First Unit At Hospital	Incident Closed	**Dataset Average**
New Orleans Police		24		23	20			24	22.75
Detroit Police	14	14		16	18			13	15
Richmond Police		21		18	20			17	19
Boulder Fire Rescue	12		14	15	11			14	13.2
New York Fire		24	24	24	24			24	24
Cincinnati Police		24		24	21			24	23.25
New York EMS		24	24	24	24	24	24	24	24
San Francisco Fire	10	7	24	24	7	1	3	13	11.13
Tempe ALS	1		4	4	3				3
Event Average	9.25	19.71	18	19.11	16.44	12.5	13.5	19.13	

8.4 F-Test for Exponentiality

Due to the chi-square test's sensitivity with large sample values, I turned to an F-test specifically designed for determining whether data indeed was sampled from an exponential distribution. This test is taken from [80] and the relies on the following distributional relationships (see [1] for elaboration on these relationships) :

1. The sum of k independent exponential random variables with rate parameter λ, is an Erlang distribution with shape and rate parameters (k, λ) such that

$$\text{If } X_i \sim Exp(\lambda)$$
$$\text{then,}$$
$$\sum_{i=1}^{k} X_i \sim Erl(k, \lambda)$$

2. The chi-squared distribution is a special case of the Erlang distribution such that

$$\text{If } X \sim Erl(k, \lambda)$$
$$\text{then,}$$
$$2\lambda X \sim \chi^2(2k)$$

3. The ratio of two independent chi-square random variables with respective degrees of freedom N_1 and N_2 is F-Distributed such that:

$$\text{If } X_1 \sim \chi^2(k_1),\ X_2 \sim \chi^2(k_2) \text{ and } Y = \frac{X_1}{k_1} \div \frac{X_2}{k_2}$$
$$\text{then,}$$
$$Y \sim F(k_1, k_2)$$

The proposed F-test proceeds by forming two randomly ordered groups from the set of n inter-arrival times, t_i. The group sizes are $r \approx \frac{n}{2}$ and $n - r$ respectively. The F statistic is calculated by:

$$F = \frac{\sum_{i=1}^{r} t_i / r}{\sum_{i=r+1}^{n} t_i / (n-r)}$$

Using the above-mentioned distribution relationships, if the inter-arrival times are indeed exponentially distributed, then F should be an F-distributed random variable with $\sim F(2r, 2(n-r))$. Using a two-tailed F-test approach, the rejection criteria can be defined as:

$$F < F_{(1-(\alpha/2)),2r,2(n-r)} \text{ OR } F > F_{\alpha/2,2r,2(n-r)}$$

Performing this test on each sample group reinforced my theory regarding the high rejection rate of the chi-square test. As seen in table 8.3, the number of null-hypothesis rejection drops dramatically. Even for datasets where all the chi-square tests were rejected (e.g. New York Fire/EMS) the F-test for exponential behavior clearly indicates the Type-I error in the chi-square test. From the results of the F-test, it is appropriate to conclude that an exponential distribution is an acceptable model for the overwhelming number of dataset/event/hour sample groups.

Table 8.3: Number of Rejections (Out of 24 Hour Partitions) of H_0 for F-Test, ($\alpha = .05$)

Dataset	Call Received	Entered to CAD	First Unit Assigned	First Unit Enroute	First Unit Onscene	First Unit To Hosptial	First Unit At Hospital	Incident Closed	**Dataset Average**
New Orleans Police		1		6	2			7	4
Detroit Police	2	1		4	4			1	2.4
Richmond Police		4		1	1			2	2
Boulder Fire Rescue	0		3	0	2			4	1.8
New York Fire		0	1	3	3			2	1.8
Cincinnati Police		2		2	0			1	1.25
New York EMS		1	2	1	0	1	1	1	1
San Francisco Fire	0	1	2	2	4	1	0	2	1.5
Tempe ALS	1		2	1	1				1.25
Event Average	0.75	1.43	2	2.22	1.89	1	0.5	2.5	

8.5 Comparing Fit with Weibull Distribution

Although the limitations of the chi-square test are clearly shown in this analysis, especially with large sample sizes, the test does have the flexibility to compare two fitted distributions against each other. As mentioned previously, I was curious whether fitting a heavier-tailed distribution would result in a lower chi-square statistic. Even though the distribution appears to be exponential from the results of the F-test, comparing against another distribution's fit is helpful. I chose to use the Weibull distribution as a comparison. As mentioned previously, the memory-less (Markovian) property of the theoretical exponential distribution is of great importance in analytical queuing models. In the context of inter-arrival times of 911 events, for example the First Unit Enroute event, suggesting an exponential distribution for the inter-arrival times would be implying that the probability of a first unit beginning to travel to an incident scene (the event occurring) does not change as time between successive events increases. For some 911 events, this might not seem like an intuitive assumption. For example, consider the "Call Received" event. As time t increases since the last 911 call, it would seem likely that the probability of receiving a new 911 call would also increase. This increased probability of an event occurring given that it has been a certain time t *since the last arrival* can be defined using a concept called the increasing failure rate (IFR). Distributions with an IFR may present a better fit than the exponential distirbution, who has a constant failure rate (due to the memory-less property). See [80] for an elaboration on the IFR condition.

The Weibull distribution possess the IFR property. It also is a heavy-tailed distribution, which could be a better candidate distribution than the exponential for the reasons discussed in section 8.2.2. In [6] inter-arrival times for calls to a PSAP in Germany were shown to fit best to a Weibull distribution as well. Although [6] only fit parameters for calls and no other 911 events, this strengthened the argument for comparing the exponential fit with a Weibull fit.

The Weibull distribution parameters are shape parameter $\alpha > 0$ and scale parameter $\beta > 0$. The exponential distribution is actually a special case of the Weibull distribution with

$\alpha = 1$. The distribution function of the Weibull distribution is given by

$$F(x) = 1 - e^{-(x/\beta)^{\alpha}}$$

if $x > 0$ and 0 otherwise.

Unfortunately, Weibull ML estimators do not have a simple closed form solution. Numerical optimization was used via the "Scipy" python package to estimate the parameters for the fitted Weibull distribution [85]. See [78] for further explanation of the ML process for the Weibull distribution.

To compare the two distributions, I used the chi-square test following the same procedure as described in section 8.3.

The results as shown in table 8.4 indicate that the fitted Weibull distribution does indeed result in a lower chi-square statistic for most of the dataset/event/hour samples. In some cases the percent decrease from the exponential chi-square statistic is somewhat negligible, as in the case of Tempe ALS's "Call Received Event". However, some differences are substantial, including multiple $> 40\%$ decreases. These results indicate that even with the success of the F-test results, the Weibull distribution could represent a better fit for inter-arrival times across 911 events. An explanation for this could be the extra parameter of the Weibull distribution. This extra parameter compared with the exponential's single λ parameter could be the difference in better fitting the data. A good example of this can be seen graphically in figure 8.4. The Weibull fit in green clearly demonstrates a much tighter fit to the data. Correspondingly the Weibull fit resulted in a $\sim 67\%$ decrease in chi-square statistic value. Interestingly though, the QQ plot for this sample does not indicate that the Weibull is much better. In fact, when comparing the QQ plots, it appears that the Weibull's heavy-tail is actually over-estimating the quantiles of the sample data. This over-estimation from the QQ-plot's perspective does not appear to significantly impact the chi-square statistic. The chi-square statistic does appear to be impacted by under-estimation however. See for instance figure 8.6 and figure 8.7. Although the Weibull distribution appears in the histogram overplot to fit only slightly better than the exponential, the QQ plot clearly shows how the exponential's light-tail is under-estimating quantiles. The chi-square statistic for the Weibull fit in this case shows a $\sim 49\%$ improvement from the exponential fit.

It is also important to note that both datasets and 911 events had differing levels of successful fits, even with the Weibull distribution. Tables 8.5 and 8.6 show that the Weibull distribution still fails many of the chi-square tests. In addition, table 8.4 makes it clear that not all events or datasets responded better to the Weibull distribution. For example, the Cincinnati Police dataset reported a higher average chi-square statistic for the Weibull fit compared to the exponential fit. Many datasets had only marginal improvements like New Orleans Police ($\sim 4.29\%$) and Detroit Police ($\sim 4.55\%$). 911 events also experienced various levels of improvement ranging from a surprising $\sim 10\%$ increase in chi-square statistic for the "First Unit At Hospital" event, to an $\sim 18\%$ decrease for "Entered to CAD" events. These differences once again show that each event in the 911 process has unique stochastic properties. This makes modeling the entire process at event-level granularity difficult.

The overall improvement that the Weibull fits show points out an important trade-off between simplicity and accuracy. The Weibull distribution is more complex than the exponential for many reasons (e.g. non-memoryless, IFR, complex MLE, etc.). As mentioned before, the exponential distribution is also extremely important for many analytical models, which are inherently less expensive than simulation-based models. It becomes important to consider at what cost can the exponential distribution be assumed an appropriate model. As the F-test results showed, there appears to be statistically valid reasons for using the exponential distribution, however if a simulation model is the desired object, than it would seem using a Weibull distribution would be a more appropriate choice.

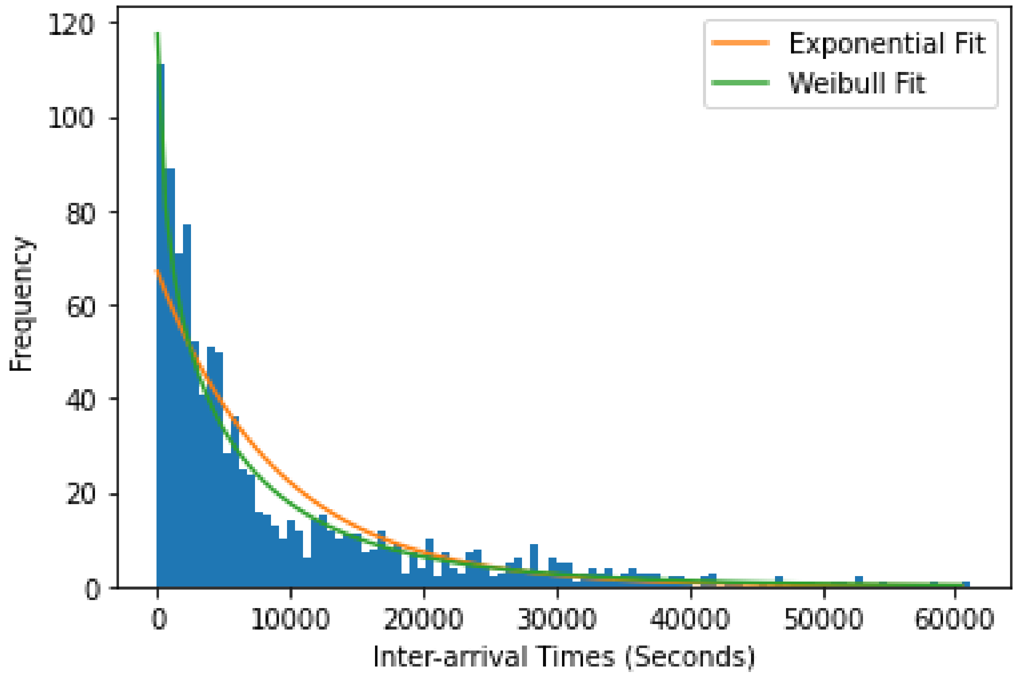

Figure 8.4: Histogram Overplot of First Unit En-Route Event Inter-Arrivals (Richmond Police: 9-9:59AM)

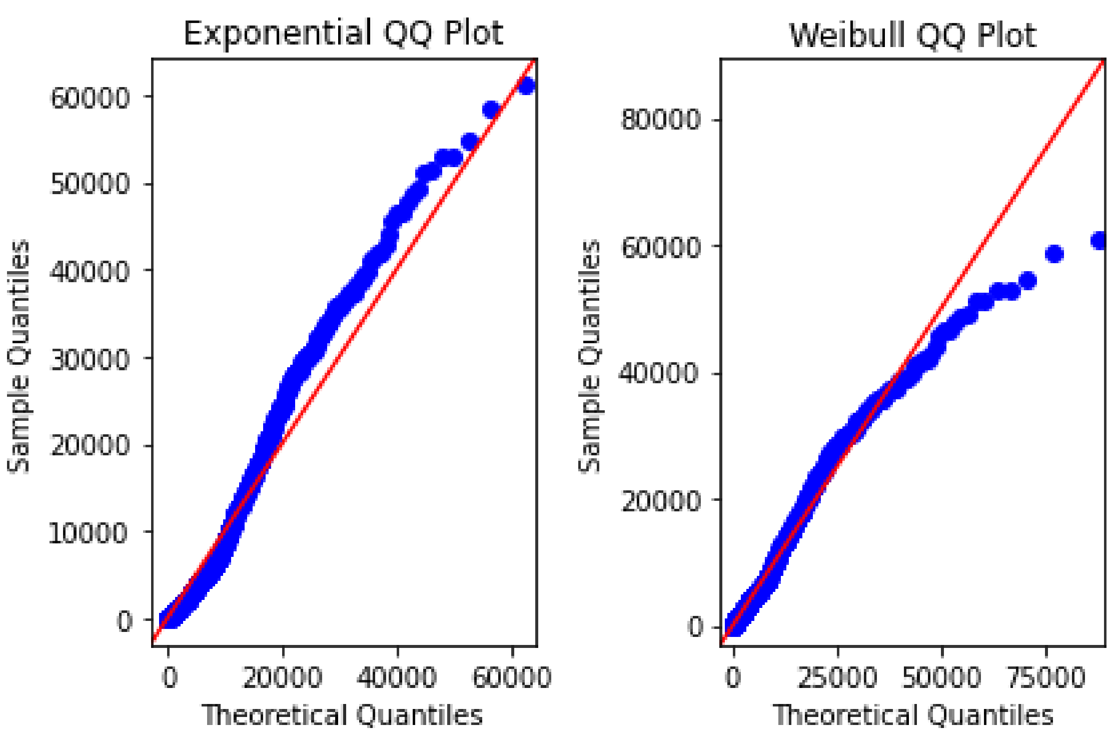

Figure 8.5: QQ Plots (Exponential and Weibull Fits) of First Unit En-Route Event Inter-Arrivals (Richmond Police: 9-9:59AM)

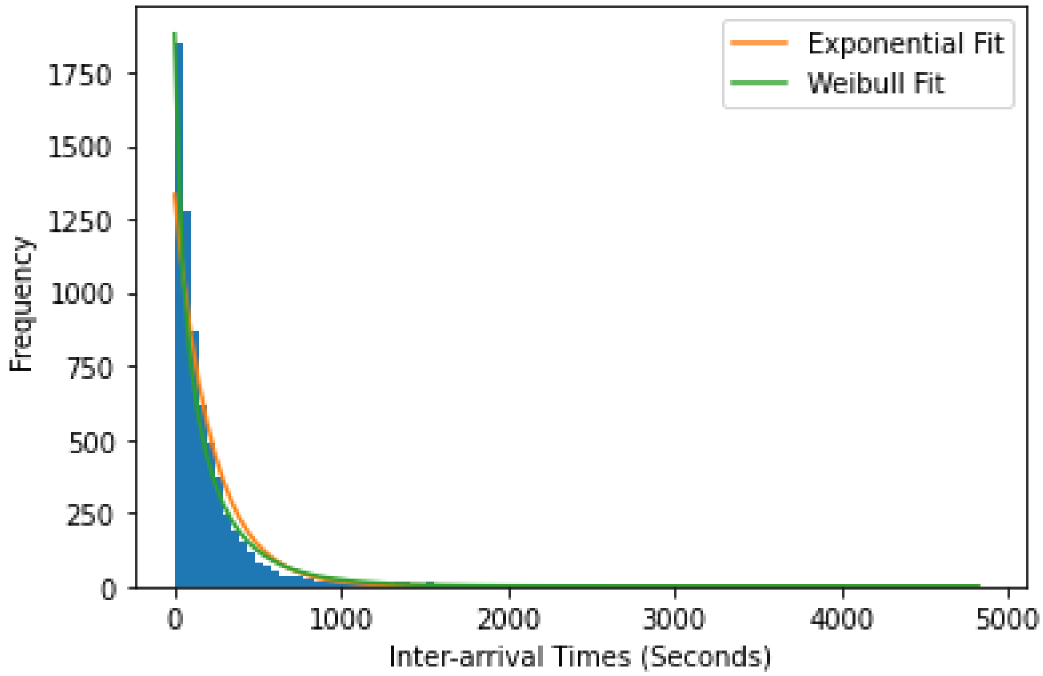

Figure 8.6: Histogram Overplot of First Unit On-Scene Event Inter-Arrivals (New Orleans Police: 7-7:59AM)

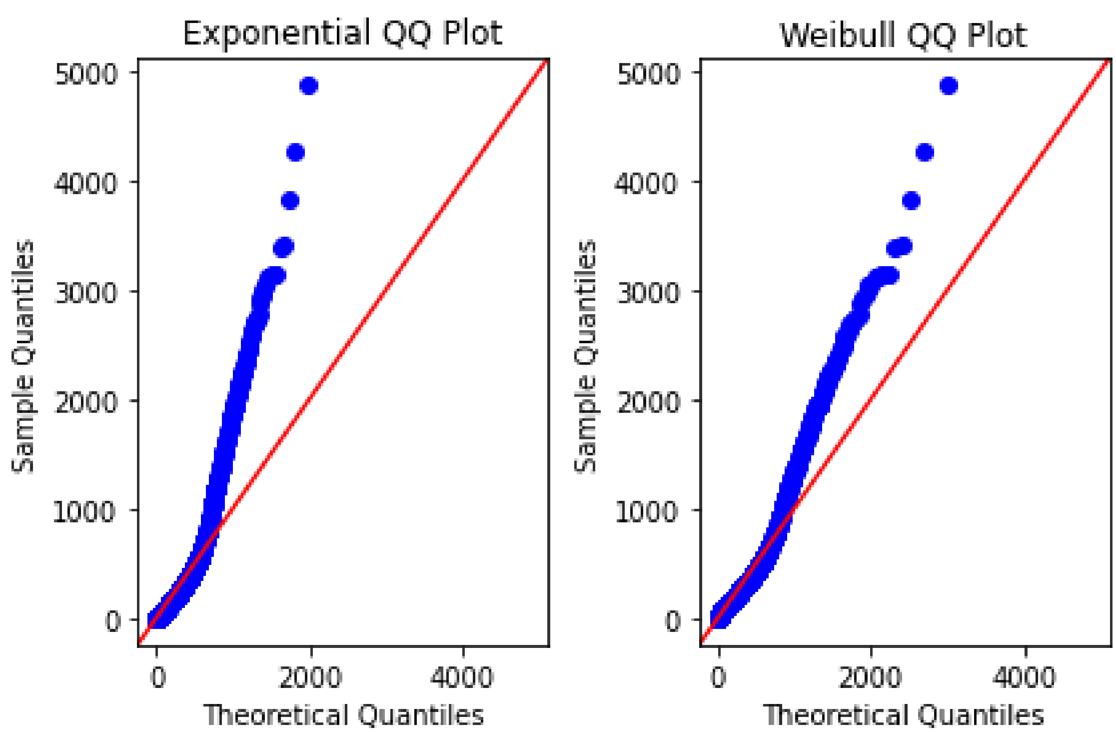

Figure 8.7: QQ Plots (Exponential and Weibull Fits) of First Unit On-Scene Event Inter-Arrivals (New Orleans Police: 7-7:59AM)

Table 8.4: Average Percent Change (%) in Chi-Square Statistic when using Weibull Distribution Fit

Dataset	Call Received	Entered To CAD	First Unit Assigned	First Unit Enroute	First Unit Onscene	First Unit To Hospital	First Unit At Hospital	Incident Closed	**Dataset Average**
New Orleans Police		-3.98%		9.56%	-14.19%			-8.55%	-4.29%
Detroit Police	-11.56%	-4.69%		-5.04%	-2.27%			0.83%	-4.55%
Richmond Police		-40.79%		-35.86%	-28.81%			-31.87%	-34.33%
Boulder Fire Rescue	-39.69%		-33.07%	-30.02%	-31.01%			-39.55%	-34.67%
New York Fire		-34.90%	-30.05%	-22.98%	10.13%			-19.02%	-19.36%
Cincinnati Police		11.18%		-7.05%	2.18%			3.18%	2.37%
New York EMS		-46.08%	-45.68%	-51.30%	-51.75%	25.63%	-3.18%	-44.02%	-30.91%
San Francisco Fire	1.61%	-8.67%	-2.80%	-25.37%	5.59%	-3.77%	-9.65%	-4.39%	-5.93%
Temple ALS	-4.72%		-13.44%	-10.54%	0.01%				-7.17%
Event Average	-13.59%	-18.28%	-25.01%	-19.85%	-12.24%	10.93%	-6.41%	-17.92%	

Table 8.5: Number of H_0 Rejections (Out of 24 Hour Intervals), Chi-Square Test ($\alpha = .05$) Weibull Fit

Dataset	Call Received	Entered to CAD	First Unit Assigned	First Unit Enroute	First Unit Onscene	First Unit To Hospital	First Unit At Hospital	Incident Closed	**Dataset Average**
New Orleans Police		22		23	13			21	19.75
Detroit Police	6	10		10	10			12	9.6
Richmond Police		9		11	10			10	10
Boulder Fire Rescue	4		7	8	8			6	6.6
New York Fire		24	24	24	24			24	24
Cincinnati Police		22		23	22			21	22
New York EMS		24	24	24	24	24	24	24	24
San Francisco Fire	7	3	24	22	5	1	1	10	9.13
Tempe ALS	1		0	1	1				0.75
Event Average	4.5	16.29	15.8	16.22	13	12.5	12.5	16	

Table 8.6: Number of H_0 Rejections (Out of 24 Hour Intervals), Chi-Square Test ($\alpha = .10$) Weibull Fit

Dataset	Call Received	Entered to CAD	First Unit Assigned	First Unit Enroute	First Unit Onscene	First Unit To Hospital	First Unit At Hospital	Incident Closed	**Dataset Average**
New Orleans Police		23		24	18			21	21.5
Detroit Police	13	12		13	13			13	12.8
Richmond Police		11		13	15			15	13.5
Boulder Fire Rescue	6		8	11	12			8	9
New York Fire		24	24	24	24			24	24
Cincinnati Police		22		23	23			21	22.25
New York EMS		24	24	24	24	24	24	24	24
San Francisco Fire	7	5	24	23	6	1	1	12	9.88
Tempe ALS	2		1	2	5				2.5
Event Average	7	17.29	16.2	17.44	15.56	12.5	12.5	17.25	

Chapter 9

Conclusion and Future Work

The 911 response process is an integral part of the Emergency Services Critical Infrastructure Sector in the United States. Understanding the stochastic properties of this system is a crucial part into building an impact-analysis model to determine the impacts of attacks/disasters to the behavior of the 911 system.

This work has been a first effort into modeling the stochastic properties of 911 response process events. A large dataset of 911 events from around the United States was gathered and processed for streamlined analysis. A synthesis effort resulted in the creation of the 911 response process event model. It is my hope that the event model will make it easy to incorporate new data into the modeling effort. Ideally, all that would need to be done is to map the various recorded events in an agency's data to the corresponding event in the model.

Using a variety of statistical techniques, the exponential distribution has been evaluated as a candidate distribution for the inter-arrival times of these events. From the results, it can be concluded that the exponential distribution is indeed an appropriate model for representing the inter-arrival time distributions of all the 911 response process events. However, it should be noted the improvement in fit when using the Weibull distribution. The estimated parameters for each dataset/event/hour sample have also been computed. This will allow for a robust reference to finely tune parameters of future models to agency/location-specific environments.

Identifying the exponential distribution for the inter-arrival times of the 911 events has major implications for the model development process. One of the key implications is that analytical analysis for the 911 process becomes much easier due to the memory-less property of the exponential. A related implication of this finding is that granular components of the 911 process like the PSAP system, can also be modeled individually using analytical tools as well.

9.1 Future Work

The efforts of this research also has paved the way for further research into the stochastic behavior of the 911 response process. A natural next effort would be to investigate the distribution of the time intervals between 911 events. Identifying appropriate distributions for the time a call is being processed at a PSAP and travel-times for the response units would be especially insightful.

Another avenue for research would be to investigate how the various inter-arrival distributions change when assessed by call-type. This research would be particularly interesting because obviously some emergency CFS are more urgent than others (i.e. heart-attack vs lost cat). Aggregating all call-types together, as was done in this study, is helpful for assessing the overall behavior of the system. However, 911 participants and stakeholders may be more interested in assessing the impact of attacks/disasters for certain call-types more than others. Overall, it is my hope that this initial modeling effort will provide some of the foundation for further analysis into the 911 process, specifically for its protection/security. Emergency Services are one of the great public services and its performant functioning is critical for society's well-being. As continued effort is placed into developing models that can accurately represent the system, better policies can be put into place to mitigate future attacks and disasters.

References

[1] Abdelkader Y, Al-Marzouq Z (2010) Probability distribution relationships. *Statistica* 70(1):41–51.

[2] ADI Staff Reporter (2016) Teenager indicted for computer tampering after 911 cyberattack (November 14). *Arizona Daily Independent* URL `https://arizonadailyindependent.com/2016/11/14/teenager-indicted-for-computer-tampering-after-911-cyberattack/`.

[3] Anderson AG (1979) The development of municipal fire departments in the United States. *Journal of Libertarian Studies* 3(3):331–359.

[4] Ani UD, Watson JDM, Nurse JRC, Cook A, Maple C (2019) A review of critical infrastructure protection approaches: Improving security through responsiveness to the dynamic modelling landscape. *CoRR* abs/1904.01551, URL `http://arxiv.org/abs/1904.01551`.

[5] Aradau C (2010) Security That Matters: Critical Infrastructure and Objects of Protection. *Security Dialogue* 41(5):491–514, URL `http://dx.doi.org/10.1177/0967010610382687`.

[6] Aschenbruck N, Fuchs C, Martini P (2009) Traffic characteristics and modeling of emergency calls at the PSAP. *2009 IEEE International Symposium on Modeling, Analysis & Simulation of Computer and Telecommunication Systems*, 1–8 (IEEE).

[7] Barton DC, Stamber KL (2000) An agent-based microsimulation of critical infrastructure systems. *International Energy Foundation's ENERGEX 2000 - 8th International Energy Forum*.

[8] Bennett DS (2018) Police response times to calls for service. *Semantic Scholar* URL `https://pdfs.semanticscholar.org/a215/5b563ee2f45a527b26d3e0fa5557c0b741d9.pdf`.

[9] Bennett KD (2010) *Spatial Analysis of Motor Vehicle Accidents in Johnson City, Tennessee, as Reported to Washington County Emergency Communications District (911)*. Master's thesis, East Tennessee State University.

[10] Bigger JE, Willingham MG, Krimgold F, Mili L, et al. (2009) Consequences of critical infrastructure interdependencies: lessons from the 2004 hurricane season in Florida. *International Journal of Critical Infrastructures* 5(3):199.

[11] Blank F (2020) A hypercube queuing model approach for the location optimization problem of emergency vehicles for large-scale study areas. *International Conference on Dynamics in Logistics*, 321–330 (Springer).

[12] Bush B, Dauelsberg L, Leclaire R, Powell D, Deland S, Samsa M (2005) 3 critical infrastructure protection decision support system (cip/dss) project overview. *23rd International Conference of the System Dynamics Society* (System Dynamics Society).

[13] Casalicchio E, Galli E, Tucci S (2007) Federated agent-based modeling and simulation approach to study interdependencies in it critical infrastructures. *11th IEEE International Symposium on Distributed Simulation and Real-Time Applications (DS-RT'07)*, 182–189.

[14] Cavallini S, d'Alessandro C, Volpe M, Armenia S, Carlini C, Brein E, Assogna P (2014) A system dynamics framework for modeling critical infrastructure resilience. Butts J, Shenoi S, eds., *Critical Infrastructure Protection VIII*, 141–154 (Berlin, Heidelberg: Springer), ISBN 978-3-662-45355-1.

[15] Chaiken JM, Dormont P (1978) A patrol car allocation model: Capabilities and algorithms. *Management Science* 24(12):1291–1300.

[16] Chohlas-Wood A, Merali A, Reed W, Damoulas T (2015) Mining 911 calls in New York City: temporal patterns, detection, and forecasting. *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence* (Association for the Advancement of Artificial Intelligence).

[17] Ciancamerla E, Minichino M, Palmieri S (2013) Modeling cyber attacks on a critical infrastructure scenario. *International Conference on Information, Intelligence, Systems and Applications (IISA)*, 1–6 (IEEE).

[18] City of Seattle (2020) City of Seattle open data portal. URL `https://data.seattle.gov/`.

[19] Connolly C (2005) Thousands of Katrina 911 calls went astray. *The Washington Post* URL `https://www.washingtonpost.com/archive/politics/2005/11/08/thousands-of-katrina-911-calls-went-astray/5dc46197-98cd-4bea-a3d6-7bfa12802430/`.

[20] Cramer D, Brown AA, Hu G (2012) *Predicting 911 Calls Using Spatial Analysis*, 15–26 (Berlin, Heidelberg: Springer), ISBN 978-3-642-23202-2, URL http://dx.doi.org/10.1007/978-3-642-23202-2_2.

[21] Cybersecurity and Infrastructure Security Agency (2020) Critical infrastructure sectors. URL https://www.cisa.gov/critical-infrastructure-sectors.

[22] Dask Development Team (2016) *Dask: Library for dynamic task scheduling*. URL https://dask.org.

[23] Davidson P (2018) With 911 dispatchers in short supply, average response time is slower (September 24). *USA Today* URL https://www.usatoday.com/story/money/2018/09/24/911-centers-struggle-hire-operators-slowing-response/1196541002/.

[24] De Souza RM, Morabito R, Chiyoshi FY, Iannoni AP (2015) Incorporating priorities for waiting customers in the hypercube queuing model with application to an emergency medical service system in Brazil. *European Journal of Operational Research* 242(1):274–285.

[25] Department of Homeland Security (2015) *Emergency Services Sector-Specific Plan* (Department of Homeland Security), URL https://www.cisa.gov/sites/default/files/publications/nipp-ssp-emergency-services-2015-508.pdf.

[26] Department of Homeland Security (2017) DHS risk lexicon. URL https://www.dhs.gov/sites/default/files/publications/18_0116_MGMT_DHS-Lexicon.pdf.

[27] Department of Homeland Security (2017) PSAP 911 service area boundaries. URL https://hifld-geoplatform.opendata.arcgis.com/datasets/psap-911-service-area-boundaries.

[28] Department of Homeland Security (2019) Critical infrastructure security. URL https://www.dhs.gov/topic/critical-infrastructure-security#.

[29] Dill E, Rodgers J, Ibrahim M, Cruzen I, Hendrickson S (2019) Slow 911 response times in SE Como prompt concern (October 2). *The Minnesota Daily* URL https://www.mndaily.com/article/2019/10/n-slow-911-response-times-in-se-como-prompt-concern.

[30] Federal Communication Commision (2018) Emergency communications. URL https://www.fcc.gov/consumers/guides/emergency-communications.

[31] Federal Communications Commission (2020) 911 and e911 services. URL https://www.fcc.gov/general/9-1-1-and-e9-1-1-services.

[32] Federal Communications Commission (2020) Text to 911: What you need to know. URL https://www.fcc.gov/consumers/guides/what-you-need-know-about-text-911.

[33] Goldberg DM, Garvey P, Xia L, Hong S (2019) Modeling local ambulance resource scheduling. *25th Americas Conference on Information Systems, AMCIS 2019, Cancún, Mexico, August 15-17, 2019* (Association for Information Systems), URL http://aisel.aisnet.org/amcis2019/.

[34] Green L (1984) A multiple dispatch queueing model of police patrol operations. *Management Science* 30(6):653–664.

[35] Green LV, Kolesar PJ (2004) Anniversary article: Improving emergency responsiveness with management science. *Management Science* 50(8):1001–1014, URL http://dx.doi.org/10.1287/mnsc.1040.0253.

[36] Grimsman D, Chetty V, Woodbury N, Vaziripour E, Roy S, Zappala D, Warnick SC (2016) A case study of a systematic attack design method for critical infrastructure cyber-physical systems. *2016 American Control Conference (ACC)* 296–301.

[37] Guidotti R, Chmielewski H, Unnikrishnan V, Gardoni P, McAllister T, van de Lindt J (2016) Modeling the resilience of critical infrastructure: the role of network dependencies. *Sustainable and Resilient Infrastructure* 1(3-4):153–168, URL http://dx.doi.org/10.1080/23789689.2016.1254999.

[38] Guri M, Mirsky Y, Elovici Y (2016) 9-1-1 ddos: Threat, analysis and mitigation. *CoRR* abs/1609.02353, URL http://arxiv.org/abs/1609.02353.

[39] Iannoni AP, Morabito R, Saydam C (2011) Optimizing large-scale emergency medical system operations on highways using the hypercube queuing model. *Socio-Economic Planning Sciences* 45(3):105–117.

[40] Industry Council for Emergency Response Technologies (2016) History of 911: And what it means for the future of emergency communications. URL https://www.ng911institute.org/wp-content/uploads/2015/10/iCERT-9EF_Historyof911_WebVersion.pdf.

[41] Ito H, Lee D (2005) Assessing the impact of the September 11 terrorist attacks on U.S. airline demand. *Journal of Economics and Business* 57(1):75 – 95, ISSN 0148-6195, URL http://dx.doi.org/https://doi.org/10.1016/j.jeconbus.2004.06.003.

[42] Jasso H, Fountain T, Baru C, Hodgkiss W, Reich D, Warner K (2006) Spatiotemporal analysis of 9-1-1 call stream data. *Proceedings of the 2006 international conference on Digital government research*, 21–22.

[43] Kean TH, Hamilton L (2004) *National Commission on Terrorist Attacks Upon the United States* (National Commission on Terrorist Attacks Upon the United States), URL https://govinfo.library.unt.edu/911/about/index.htm.

[44] Kendall DG (1953) Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain. *The Annals of Mathematical Statistics* 24(3):338–354, ISSN 00034851, URL http://www.jstor.org/stable/2236285.

[45] Kolesar P, Blum EH (1973) Square root laws for fire engine response distances. *Management Science* 19(12):1368–1378, URL http://dx.doi.org/10.1287/mnsc.19.12.1368.

[46] Kundur D, Feng X, Mashayekh S, Liu S, Zourntos T, Butler-Purry KL (2011) Towards modelling the impact of cyber attacks on a smart grid. *International Journal of Security and Networks* 6(1):2–13.

[47] Larson RC (1974) A hypercube queuing model for facility location and redistricting in urban emergency services. *Computers & Operations Research* 1(1):67–95.

[48] Last Week Tonight with John Oliver (2016) 911 (May 15) URL https://www.youtube.com/watch?v=A-XlyB_QQYs.

[49] Law AM, Kelton WD, Kelton WD (2000) *Simulation Modeling and Analysis*, volume 3 (McGraw-Hill New York).

[50] Layton J (2006) How 9-1-1 works (July 12). *HowStuffWorks.com* URL https://people.howstuffworks.com/9-1-1.htm.

[51] Lewis TG (2019) *Critical infrastructure protection in homeland security: defending a networked nation* (John Wiley & Sons).

[52] Lovejoy B (2016) Teenager plays havoc with multiple 911 call centers trying to collect Apple bug bounty (October 31). *9to5Mac* URL https://9to5mac.com/2016/10/31/ios-exploit-911-call-centers/.

[53] Luiijf E, Nieuwenhuijs A, Klaver M, van Eeten M, Cruz E (2009) Empirical findings on critical infrastructure dependencies in Europe. Setola R, Geretshuber S, eds., *Critical Information Infrastructure Security*, 302–310 (Berlin, Heidelberg: Springer Berlin Heidelberg), ISBN 978-3-642-03552-4.

[54] Luiijf HAM, Nieuwenhuijs AH, Klaver MHA (2008) Critical infrastructure dependencies 1-0-1. *2008 First International Conference on Infrastructure Systems and Services: Building Networks for a Brighter Future (INFRA)*, 1–3.

[55] Manoj BS, Baker AH (2007) Communication challenges in emergency response. *Communications of the ACM* 50(3):51–53.

[56] Markets, Markets (2020) Computer aided dispatch market. URL `https://www.marketsandmarkets.com/ResearchInsight/computer-aided-dispatch-market.asp`.

[57] Mason E (2017) File:9-1-1 system.svg — wikimedia commons, the free media repository. URL `https://commons.wikimedia.org/w/index.php?title=File:9-1-1_System.svg&oldid=249786202`, [Online; accessed 26-April-2020].

[58] McLaughlin M (1994) Automatic vehicle location for law enforcement. *The National Consortium for Justice Information and Statistics, SEARCH Technical Bulletin* (3).

[59] Medhi J (2003) Chapter 2 - queueing systems: General concepts. *Stochastic Models in Queueing Theory (Second Edition)*, 47 – 64 (Burlington: Academic Press), second edition edition, ISBN 978-0-12-487462-6, URL `http://dx.doi.org/https://doi.org/10.1016/B978-012487462-6/50002-3`.

[60] Miller B, Rowe D (2012) A survey SCADA of and critical infrastructure incidents. *Proceedings of the 1st Annual conference on Research in information technology*, 51–56.

[61] National 911 Program (2020) The national 911 program celebrates 50 years of 911. *911.gov* URL `https://www.911.gov/50-years-of-911.html`.

[62] National Emergency Number Assocation (2020) 911-statistics URL `https://www.nena.org/page/911Statistics`.

[63] National Research Council (2003) *The Internet Under Crisis Conditions: Learning from September 11* (National Academies Press).

[64] Neusteter SR, Mapolski M, Khogali M, O'Toole M (2019) The 911 call processing system: A review of the literature as it relates to policing. *The Vera Institute of Justice* .

[65] Newman MEJ (2002) Spread of epidemic disease on networks. *Phys. Rev. E* 66:016128, URL `http://dx.doi.org/10.1103/PhysRevE.66.016128`.

[66] Oliva G, Panzieri S, Setola R (2010) Agent-based input–output interdependency model. *International Journal of Critical Infrastructure Protection* 3(2):76 – 82, ISSN 1874-5482, URL http://dx.doi.org/https://doi.org/10.1016/j.ijcip.2010.05.001.

[67] Ouyang M (2014) Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability Engineering System Safety* 121:43 – 60, ISSN 0951-8320, URL http://dx.doi.org/https://doi.org/10.1016/j.ress.2013.06.040.

[68] Pasqualetti F (2012) *Secure control systems: A control-theoretic approach to cyber-physical security*. Ph.D. thesis, University of California, Santa Barbara.

[69] Perkins KB (1989) Volunteer firefighters in the United States: A descriptive study. *Nonprofit and Voluntary Sector Quarterly* 18(3):269–277.

[70] Potter G (2013) The history of policing in the United States. *Eastern Kentucky University Police Studies Online* URL https://plsonline.eku.edu/insidelook/history-policing-united-states-part-1.

[71] President's Commission on Law Enforcement and Administration of Justice (1967) The challenge of crime in a free society. URL https://www.ncjrs.gov/pdffiles1/nij/42.pdf.

[72] Program N (2017) *911.gov*. URL https://www.911.gov/pdf/National-911-Program-Profile-Database-Progress-Report-2017.pdf.

[73] Program N (2020) Cybersecurity. URL https://www.911.gov/issue_cybersecurity.html.

[74] Sampson R (2002) Misuse and abuse of 911. *Problem-Oriented Guides for Police Problem-Specific Guides Series* URL https://cops.usdoj.gov/RIC/Publications/cops-p024-pub.pdf.

[75] Sandberg H, Amin S, Johansson KH (2015) Cyberphysical security in networked control systems: An introduction to the issue. *IEEE Control Systems Magazine* 35(1):20–23.

[76] Santos JR, Haimes YY, Lian C (2007) A framework for linking cybersecurity metrics to the modeling of macroeconomic interdependencies. *Risk Analysis* 27(5):1283–1297, URL http://dx.doi.org/10.1111/j.1539-6924.2007.00957.x.

[77] Satumtira G, Dueñas-Osorio L (2010) Synthesis of modeling and simulation methods on critical infrastructure interdependencies research. *Sustainable and resilient critical infrastructure systems*, 1–51 (Springer).

[78] Scipy (2020) Statistics (scipy.stats). URL https://docs.scipy.org/doc/scipy/reference/tutorial/stats.html.

[79] Seabold S, Perktold J (2010) statsmodels: Econometric and statistical modeling with python. *9th Python in Science Conference*.

[80] Shortle JF, Thompson JM, Gross D, Harris CM (2018) *Fundamentals of Queueing Theory*, volume 399 (John Wiley & Sons).

[81] Slay J, Miller M (2008) Lessons learned from the Maroochy water breach. Goetz E, Shenoi S, eds., *Critical Infrastructure Protection*, 73–82 (Boston, MA: Springer US), ISBN 978-0-387-75462-8.

[82] Streicher B (2020) Austin first responders too slow to respond to 911 calls, new audit finds (February 27). *KVUE* URL https://www.kvue.com/article/news/investigations/defenders/austin-first-responders-911-calls-too-slow-audit-finds/269-339d15f8-3a90-4cbd-a06d-216bf1663e3b.

[83] The Pandas development team (2020) pandas-dev/pandas: Pandas. URL http://dx.doi.org/10.5281/zenodo.3509134.

[84] Vera Institute of Justice Data Clinic (2020) Announcing a consolidated dataset of 911 calls for five US cities (part 1). URL https://medium.com/dataclinic/announcing-a-consolidated-dataset-of-911-calls-for-five-us-cities-part-1-4320a1a31a88.

[85] Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Jarrod Millman K, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey C, Polat İ, Feng Y, Moore EW, Vand erPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P, Contributors S (2020) SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* 17:261–272, URL http://dx.doi.org/https://doi.org/10.1038/s41592-019-0686-2.

[86] Wallace WA, Mendonça D, Lee E, Mitchell J, Chow J (2001) Managing disruptions to critical interdependent infrastructures in the context of the 2001 World Trade Center attack. *Impacts of and Human Response to the September 11, 2001 Disasters: What Research Tells Us* URL https://www.researchgate.net/publication/241680348_

Managing_Disruptions_to_Critical_Interdependent_Infrastructures_in_the_Context_of_the_2001_World_Trade_Center_Attack.

[87] Wolstenholme EF (1982) System dynamics in perspective. *Journal of the Operational Research Society* 33(6):547–556, URL http://dx.doi.org/10.1057/jors.1982.117.