Published in final edited form as:

IEEE Trans Syst Man Cybern Syst. 2014 December ; 44(12): 1569–1583. doi:10.1109/TSMC.2014.2332137.

RE-PLAN: An Extensible Software Architecture to Facilitate Disaster Response Planning

Martin O'Neill II, Armin R. Mikler, Saratchandra Indrakanti, Chetan Tiwari, and Tamara Jimenez

Center for Computational Epidemiology and Response Analysis University of North Texas, Denton, TX USA

Abstract

Computational tools are needed to make data-driven disaster mitigation planning accessible to planners and policymakers without the need for programming or GIS expertise. To address this problem, we have created modules to facilitate quantitative analyses pertinent to a variety of different disaster scenarios. These modules, which comprise the REsponse PLaN ANalyzer (RE-PLAN) framework, may be used to create tools for specific disaster scenarios that allow planners to harness large amounts of disparate data and execute computational models through a point-and-click interface. Bio-E, a user-friendly tool built using this framework, was designed to develop and analyze the feasibility of ad hoc clinics for treating populations following a biological emergency event. In this article, the design and implementation of the RE-PLAN framework are described, and the functionality of the modules used in the Bio-E biological emergency mitigation tool are demonstrated.

Keywords

biological emergencies; disaster mitigation planning; emergency response; evacuation planning; POD throughput; public health preparedness

1 Introduction

The REsponse PLaN ANalyzer (RE-PLAN) framework was designed to facilitate the construction of computational tools for the analysis and development of disaster mitigation and evacuation plans. Development of the framework centered around creating computational methods to analyze plans for the treatment of large populations following the release of hazardous biological substances [1][2]. A widely accepted approach to this problem includes the implementation (by local governments) of ad hoc medical countermeasure dispensing facilities known as Points of Dispensing (PODs) [3]. These PODs must be strategically placed and designed in order to treat the entire target population within a short mandated time frame. This presents a challenge for high-population localities which must develop plans to reach hundreds of thousands or millions of individuals under time constraints as short as twenty-four to forty-eight hours [3]. Computational tools created with the RE-PLAN Framework can be designed to make data-driven creation and analysis of these plans accessible to public health preparedness practitioners (PHPP) using a graphical interface [4].

Throughout development of the framework, collaborations with PHPP from local counties were maintained in order to develop the requirements for the system. Hardware requirements are based upon the availability of computational hardware to the PHPP of local counties, and software interface requirements are based upon observations of expertise generally possessed by PHPP. System output requirements are based upon federal guidelines, local PHPP needs, and the availability of input data. The RE-PLAN framework was designed to conform to these requirements.

POD placements with respect to the geographic distribution of the population affect many critical aspects of planning such as road network congestion, average distance-to-assigned-POD, expected population distribution among PODs, and resource distribution requirements of the PODs. Road network congestion may prevent portions of the population from reaching the PODs within the time frame, resulting in the under-utilization of mitigation resources at the PODs. Large average distance-to-assigned-POD may contribute to this congestion by forcing individuals to traverse a greater number of roads to reach their assigned PODs. The Centers for Disease Control and Prevention (CDC) has recommended that population and resource distribution among PODs in a plan should be uniform in order to avoid public perceptions of service quality disparities which could cause certain PODs to be overwhelmed [3].

A set of facilities is considered feasible if its operational efficiency defined by [3] is capable of meeting service requirements (e.g. specific time frames for service completion or proportions of populations to be served) without exceeding available resources (e.g. transportation network capacities or limitations of facility infrastructure). The RE-PLAN framework has been constructed to analyze and optimize the feasibility of response plans. This paper highlights the following main architectural components of the RE-PLAN framework and the modules designed to implement them:

- Plan Initializer - Facilitates the naming of plans, choice of study areas, and specification of plan-wide assumptions.
- Plan Designer/Optimizer - Generates service areas of selected facilities. Sets of facilities in existing plans may be analyzed or sets of feasible facilities may be generated with respect to the populations' geographic distributions.
- Logistics calculator - Calculates how the population utilizes the transportation network to travel to the facilities. These calculations facilitate the analysis of conditions on the transportation network resulting from response plan implementation.
- Plan Analysis Tools - Creates graphical representations and maps of processed data for analyses.

A case study is introduced which uses a tool based on the RE-PLAN Framework to illustrate differences between the plan design and optimization methods. A discussion concerning the work described is presented, and the paper is concluded with a description of the limitations of the system.

Computational models of biological emergency events show the importance of a policy of aggressive mass treatment [5][6], and delays in this treatment can lead to increased numbers of casualties [7]. Routing and scheduling for timely delivery of medical countermeasures to treatment facilities have been examined [8], and strategies regarding resource distribution among the facilities have been explored [9]. Nonetheless, creating feasible plans for the distribution of medical countermeasures to large populations remains an open problem [10].

To aid larger cities in planning their biological emergency mitigation procedures, the United States Department of Health and Human Services instituted the Cities Readiness Initiative (CRI) in 2004 [11]. An initial evaluation of CRI indicates that the initiative has improved mass treatment preparedness [12]. Studies have been conducted regarding shortcomings and optimization strategies inside service facilities during a biological emergency [13][14], and tools have been created to study and practice policies and procedures within each facility [15][16]. However, little attention has been paid to how the population will be delivered to facilities for treatment during response plan implementation.

Surveillance systems, such as the BioSense system created by the Centers for Disease Control and Prevention [11], use data from disparate sources to facilitate early detection of biological emergency events. The World Health Organization (WHO) recognizes the potential release of the harmful biochemical substance anthrax as being one of the new health threats of the twenty-first century [17] and endorses the use of public health surveillance systems, referring to them as being, “the cornerstone of public health security.” [17] A wide variety of different data sources, such as over-the-counter drug sales [18][19], internet search query patterns [20][21], and personal web log (blog) data [22], have been explored to detect public health events. However, further exploration of available data sources to allow decision makers to develop and analyze their response plans is needed [13].

Disaster mitigation and evacuation planning must rely on an analysis of available data [23], and plan design must consider both the length of the mandated time frame and the size of the target population [3]. However, the vast amounts and different types of data needed make this analysis intractable without the use of computational tools [4]. A set of modules was designed for data analysis problems relevant to a variety of disaster mitigation and evacuation planning and analysis. These modules, which facilitate the analyses of many different types of geographic, spatio-temporal, and demographic data, comprise the RE-PLAN Response Plan Analysis framework. Collaboration with public health departments has led to the implementation of the Bio-E tool which is based on the RE-PLAN framework. This tool was designed to create and analyze POD-based biological emergency mitigation plans.

A feasible response plan must be able to accomplish its assigned tasks without exceeding the available resources. Facilities hosting ad-hoc clinics during a biological emergency must serve their entire populations within specific time frames. This *facility throughput* problem is affected by the geographic population distribution, facility locations, constraints of the transportation network, limitations of the facilities, and the availability of personnel and supplies. All of these factors can be separated into two groups: problems that may occur at the facilities themselves, and problems that may occur in each facility's service area.

1.1 Problems at the Facilities

The primary question to be answered when analyzing problems that may occur at a facility is, “Can the facility serve the number of people in its assigned population under the given time constraints?” To answer this question, the service area of each facility must be determined. Once the service areas have been determined, the population of each service area may be analyzed to estimate the requirements of each facility. These requirements lead to further questions such as:

- “Can the parking lot at each facility support the number of cars which must enter and exit under the time constraints?”
- “How long will it take for each facility to serve its assigned population?”
- “Based on the assigned population of each facility, are there any special requirements for the facilities?”

These questions must be answered in order to assess plan feasibility.

1.2 Problems in the Service Areas

The population in each service area is unlikely to be uniformly distributed. Further, transportation network infrastructure is likely to be irregularly distributed across the service areas. Therefore, the locations of the facilities in relation to population distribution and transportation network resources must be examined. If the transportation network is incapable of delivering individuals to the facility in a timely manner, resources at the facility may be under-utilized. This motivates the question, “How will the implementation of the facilities in a given plan affect the traffic situation on the transportation network?”

2 RE-PLAN Framework

The RE-PLAN architectural framework consists of the Plan Initializer, Plan Designer/Optimizer, the Logistics Calculator, and the Plan Analysis Tools. These four components communicate with the RE-PLAN Database to facilitate plan creation and analysis. Each component consists of a series of modules which may be redesigned, augmented, or replaced in order to modify its features or underlying methodology. The work flow diagram in Figure 1 depicts an overview of how the modules are used in each of the architectural components to initialize and design plans, calculate logistics, and perform analyses. RE-PLAN was designed to allow different sets of facilities, assumptions, or areas of interest to be explored in the planning process. Therefore, each plan inside the RE-PLAN system is referred to as a *scenario*. Screen captures from the RE-PLAN Bio-E tool developed for POD planning and analysis are included in this article as examples.

2.1 Plan Design

The **Plan Initializer** is used to create new scenarios. It creates a designation and description of the new scenario as well as a regional data set to be used by the new scenario based on data pre-loaded into the database and user input. Once a scenario has been initialized, a map of the study area is displayed, and the scenario is ready to be designed.

The **Plan Designer/Optimizer** facilitates the creation of response plans consisting of sets of facilities and assigned service areas. Figure 1 shows different paths through the modules of the Plan Designer based on different optimization criteria. Each of these paths results in a set of facilities and service which are stored in the RE-PLAN Database.

The Facility Editor has a graphical interface which facilitates the creation, modification and deletion of individual facilities. Entire sets of facilities may also be imported or exported. This module updates the RE-PLAN Database and the interactive map as facilities are modified in a plan. Each facility is represented by a set of attributes that are provided by the planner as depicted in Table 1. Many of these attributes are only stored to assist policymakers during the planning process. The attributes of each facility which affect plan analysis calculations are the facility's location (longitude and latitude), type, status, and number of booths. A facility's type may be used to specify whether a facility is public and should be included in logistical calculations or analysis of a plan. A facility's status reflects whether it is active and whether the location is used in the current response plans. Only facilities which are designated as being *public* and *active* are included in the logistic and analysis calculations. The number of booths of a facility represents the number of individuals which may be treated at this facility concurrently. Once the facilities of a plan have been input, service areas can be created in preparation for logistics calculations.

Plan optimization includes a number of methods for choosing facility locations and calculating service areas. Some planning activities may aim to minimize the spatial distance individuals must travel to reach their assigned facility. Others may focus on the generation of uniform population service areas. Each of these optimization goals utilize specific optimization algorithms which are part of the RE-PLAN Framework. The specifics regarding these algorithms are beyond the scope of this article and may be found in [24].

2.2 Plan Analysis

Once a plan has been initialized and designed, the **Logistics Calculator** prepares it for traffic analysis. For each service area, it models a round-trip flow of individuals along the road network to the facility. An estimated load (in numbers of individuals) on the road network is calculated and saved for a load:capacity analysis. A more detailed description of this procedure is provided in Section 3.3.

The **Plan Analysis Tools** facilitate analysis of response plans at the facility and on the transportation network. The Facility Requirement Analyzer provides detailed information about each facility both graphically and as a written report. This information includes the population distribution among facilities and service throughput estimates and requirements for each facility and its parking lot. The distance distribution of each population to its assigned facility is displayed as a cumulative graph. The Traffic Analysis Toolbox uses data from the RE-PLAN Database calculated by the Logistics Calculator to facilitate exploration of traffic conditions which will impact response plan implementation. Assumptions which affect traffic conditions can be modified. Although these modules are functionally separate, they comprise the set of analysis tools available in the RE-PLAN Framework.

The *Facility Requirement Analyzer* facilitates the analysis of response plan data created by the Plan Designer. Graphical representations of population distribution among the facilities are created, allowing PHPP to easily analyze personnel and supply distribution requirements of the facilities. Further, the population assigned to each facility is combined with user-specified assumptions regarding the number of booths assigned to each facility, the average number of people who will travel to the facility in each car, and the amount of time required to serve each person to create graphical representations for the analysis of each facility. The Facility Requirement Analyzer provides the following five methods for examining plan scenarios:

- A written report for the entire plan and for each individual facility.
- Two different graphical representations of the population distribution among the facilities.
- A chart depicting the time required for each facility to serve its assigned population under the current assumptions with those unable to meet assumed deadlines highlighted in red.
- A chart showing estimates of the situation in each facility's parking lot under different assumptions.
- A graph illustrating the distance the population must travel to reach the facilities.

The *Traffic Analysis Toolbox* combines geographic population data with transportation data in the context of the response plan logistics data. Traffic conditions at specific points may be examined with respect to the physical properties of the transportation network and the load on these points resulting from *business as usual* base traffic, traffic caused by implementation of the response plans, or a combination of both. Traffic conditions at each point are classified into one of six different classes which represent the ratio of load to the maximum physical capacity. Figure 2a shows how these classes are visually represented on a map to facilitate analysis of traffic conditions by personnel without the need for expertise in Geographic Information Systems (GIS) or computer programming.

The Traffic Toolbox graphical interface in Figure 2b allows users to easily change analysis parameters through a point-and-click interface. Traffic resulting from implementation of response plans and base traffic may be toggled *on* or *off* independently. Parameters which affect base traffic projections such as *weekday* versus *weekend* traffic and *time of day* may be adjusted. In total, 192 different base traffic estimations are available for analysis (ninety-six per day at fifteen minute intervals on either a weekday or weekend). Assumptions which affect response plan traffic such as *people per car* and time constraints (in numbers of hours) may also be adjusted. Table 2 provides a list of possible values for the options in the Traffic Toolbox.

2.3 RE-PLAN Database

The RE-PLAN Database stores all information about the response plans. It is also the primary way data are shared among the different architectural components of the RE-PLAN Framework. Three main categories of data are stored in the RE-PLAN Database: data

concerning the area of interest, data which comprises response plans, and other data which facilitates the user experience and collaboration among users.

Geographic data sets for building areas of interest must be loaded into the RE-PLAN Database before plan creation or analysis begins. This includes spatial demographic and population data, spatial transportation network data, and traffic count observation data. These data are available from sources such as the U.S. Census Bureau and regional governments, respectively. If traffic count observation data is unavailable for a specific area, data from another area may be used as an estimate. During initialization, RE-PLAN integrates these data to build each region for the newly created scenarios.

The set of facilities, their service areas, and the load distributed across the transportation network are all stored in the RE-PLAN Database. For each facility, the facility's name, location, address, city, zip code, status, type, and other comments are stored in the database. If the Logistics Calculator has been used on a response plan, the divided service areas and load on the transportation network at the crossing points are also stored in the database. Notes about each plan may also be stored in the database.

The RE-PLAN Database stores data that is used to describe the area of interest, create and save response plans, and provide a personalized user experience. It is designed to facilitate response plan creation and provide a platform to share response plans while minimizing storage requirements. In order to accomplish this, the tables in the database are divided among schemas of two kinds: the public schema and user schemas. Data from both the public schema and a user schema are required for RE-PLAN execution. Public schema tables generally consist of pre-loaded data and a *scenarios* table used to organize and track the state of saved plans. Each user is assigned a user schema under which tables specific to response plans created by that user are stored.

In order to maintain consistency of data in the public schema and in saved response plans, temporary copies of all tables required for execution of RE-PLAN are created during run-time under the current user's schema. All computations are performed on the temporary tables. To save a plan, copies of most temporary tables are made, and entries are added or updated in the *scenarios* table to track the state of saved plans. Temporary tables such as aggregations of data resulting from the user's selection of an area of interest are not stored. Instead, a description of the area of interest is saved in the *scenarios* table which can be used to reconstruct these tables. Figure 3 shows the tables that exist and are saved depending on the state of the scenario.

All temporary tables corresponding to a plan are dropped when the plan is closed. However, if RE-PLAN crashes or loses its connection to the database, temporary tables may be orphaned. When temporary tables are created, a creation time for them is also saved in the *scenarios* table. At run time, RE-PLAN checks this table for temporary tables with time stamps older than a user-specified age. Any tables older than this age are dropped under the assumption that they have been orphaned.

Each saved scenario is owned by a specific user, and the *scenarios* table maintains details about saved scenarios and their owners. When a user accesses another user's saved

scenarios, relevant tables are copied to their own schema. This protects the original user's plans from unwanted modifications while facilitating collaboration among users. Further, state information for each scenario is maintained in the *scenarios* table, allowing scenarios to be saved and reloaded with the correct options and features enabled. For example, if a plan was saved before the Logistics Calculator was executed, when the plan is reloaded, the Traffic Analysis Toolbox should not be available until the Logistics Calculator has been executed.

3 Methodology

The methodology employed by the RE-PLAN Framework was designed to analyze plan feasibility using large amounts of data from disparate sources. These data include population data, transportation network data, and traffic count observation data. Combined with assumptions from local public health officials drawn from an understanding of their region and populations, this methodology is used to create a model which facilitates the analysis of conditions resulting from the implementation of response plans. Further, the model allows public health officials to explore alternate plan scenarios while exploring the data underlying the computational model.

3.1 Facility Selection and Service Area Determination

Two different facility selection and service area determination methods have been developed based on different optimization criteria. The simplest method allows users to select facility locations directly and then calculates service areas such that the distance to the facility in each service area is minimized. An alternate method creates approximately uniform service areas based on selected demographic variables. Once these service areas have been created, suggested facility locations, which minimize the distance to the facility in each service area, are automatically selected within them.

3.2 Service Facility Analysis

Once service areas have been determined for each facility, the population assigned to each of them can be calculated using spatial population data. Each facility's requirements and feasibility can then be calculated with respect to the population it is assigned to serve. The outcome of this analysis may lead planners to redistribute personnel and resources among the facilities or to modify their response plans. Methods for analyzing the set of facilities include examining the distribution of the population among them, calculating the amount of time it will take for each to serve its assigned population, and estimating the traffic situation at each facility's parking lot.

Decisions regarding the distribution of personnel and supplies among the facilities require an analysis of the distribution of the population across the set of facilities. Facilities assigned a greater population must be assigned more personnel and resources. Those assigned drastically smaller or larger populations can be identified, allowing planners to adjust their response plans accordingly. Once the population distribution across the set of facilities is acceptable, and personnel and resource distributions have been determined, analysis can continue.

The estimated service time T_i required for each facility i to serve its assigned population p_i can be calculated by

$$T_i = \frac{p_i s}{w_i}, \quad (1)$$

where s is the amount of time it would take to serve a single individual and w_i is the number of individuals who may be served at facility i concurrently. Any facility whose estimated service time is close to or exceeds the mandated time constraints may be incapable of meeting its service requirements. The estimated service time of a particular facility may be decreased by increasing the number of individuals whom it may serve concurrently, by decreasing the assigned population, or by shortening the amount of time it would take to serve a single individual.

Assuming all individuals travel to their assigned facilities in cars, the total number of cars a_i which will visit each facility i can be estimated by

$$a_i = \frac{p_i}{f}, \quad (2)$$

where f is the average number of people who will travel in each car. This average is estimated by planners and may be based upon such factors as regional demographics or familiarity with the population to be served. The maximum time interval r_i during which cars must enter and exit the parking lot at each facility i in order to meet the facility's estimated service time is

$$r_i = \frac{T_i}{a_i}, \quad (3)$$

Further, the maximum time interval ρ_i during which cars must enter and exit each parking lot in order to serve the population within a time frame of ω is

$$\rho_i = \frac{\omega}{a_i}, \quad (4)$$

3.3 Traffic Analysis

Plan feasibility is contingent upon the populations' ability to reach their assigned facilities and return home. This ability is dependent upon the geographic distribution of capacity on the transportation network with respect to the demand placed on the network as a result of plan implementation. Traffic congestion will result in instances where demand exceeds capacity, slowing the arrival of individuals to their assigned facilities and leaving facility resources under-utilized. Therefore, the ratio of demand to capacity resulting from the implementation of a response plan must be estimated in order to evaluate the feasibility of this response plan.

Capacities of links in the transportation network can be obtained from local planning authorities or estimated using the type of road (e.g. freeway or arterial), speed limits, and

numbers of lanes. Geographic population and transportation network data must be used to model the flow of individuals from their homes to their assigned facilities and back again to generate an estimate of demand at specific points in the network. However, the numbers of road segments in the transportation data and of geographic population data points makes this estimation a very complex process.

A model was developed to approximate the flow of the populations across the transportation network to their assigned PODs. In this model, each service area is divided into rings of proximity around the facility as shown in Figure 4. It can be observed that individuals who live in the outer rings must travel through the inner rings to arrive at the facility in the center ring as shown in Figure 5. Further, individuals must travel a similar but opposite route to return home. Each ring of the service area is divided further into a series of segments around links in the transportation network which connect them as exemplified in Figure 6.

The geographic population data points are used as a source for load estimations across the transportation network. For each service area, the flow of individuals from their origins to their assigned facility is modeled. Figure 7 illustrates how roads where the population crosses from ring to ring can be examined with respect to the assumed time constraints. Population flow estimates for each road are converted to automobile traffic estimates using the assumed average number of individuals per car. Properties of the transportation network are likely to vary. Therefore, the traffic caused by the implementation of the plan at a specific point must be analyzed with respect to the properties of the transportation network at that point (e.g. speed limit, number of lanes, or estimated maximum physical capacity).

It cannot be assumed that traffic resulting from implementation of the response plan will be the only source of load on the transportation network. Therefore, an estimate of normal *business as usual* traffic must be estimated to serve as a base upon which traffic caused by the implementation of the response plan may be added. To accomplish this task, traffic count observation data must be used with other properties of the transportation network to create an estimation of base traffic across the network. Further, *time of day* and *day of week* become important considerations when dealing with base traffic to adequately represent peak and off-peak traffic periods.

4 Implementation

The RE-PLAN framework employs a client-server model. Client-side modules are written in Java for portability. Spatial data display and interaction are implemented using the GeoTools open source Java GIS toolkit. Data analysis charts displayed by the Facility Analyzer are rendered using the JFreeChart open source chart library. A PostgreSQL database with PostGIS is used on the RE-PLAN server to store and manipulate both spatial and non-spatial data. The client-server design facilitates collaboration among users whose client programs connect to the same server, thus allowing users to access each other's mitigation plan scenarios. Further, the client programs have minimal system requirements since most of the complex calculations are executed on the server. As a result, there is little need for hardware upgrades on otherwise up-to-date client machines.

4.1 Extensibility of the RE-PLAN Framework and the Creation of Tools for Emergency Response Planning

The RE-PLAN Framework was designed to facilitate the addition of new features. It maintains a set of state variables which can be accessed to determine which tables should exist in the database at any given time (described in greater detail in Section 2.3) and the designation (name) chosen for the current plan. New features can access the database by building queries using a set of constant string variables which support the framework's naming conventions of database tables. Although the GeoTools open source Java GIS toolkit is used to render the map display, map style builders are included in the framework to simplify the addition of new map layers. This facilitates the integration of new, region-independent features into the RE-PLAN Framework.

Tools for a variety of planning scenarios (e.g. mass prophylaxis planning, evacuation planning, or emergency response planning) can be created using the RE-PLAN Framework. A large set of features (e.g. uniform population partitioning or closest distance partitioning) is available in the framework for inclusion into these tools. However, in order to simplify the software interface, tools should only include features required by the planners who will use them. Therefore, a tool created using the RE-PLAN Framework includes a specific subset of the framework's features chosen to support the tasks most relevant to a specific planning problem.

4.2 RE-PLAN Bio-E POD Design and Analysis Tool

The RE-PLAN Bio-E POD analysis tool allows PHPP to create response plans and analyze their feasibility with respect to a variety of resource constraints. Features from the RE-PLAN Framework included in this tool have been customized around specific issues related to the POD planning and analysis problem. These features can be accessed through the graphical interface of the tool.

Execution of logistics and analysis features is available through the graphical interface of the application. The primary tools offered by the framework to be used on an existing scenario can be classified into plan creation and plan analysis; hence the tools menu is organized into *plan creation* and *plan analysis* categories. Plan creation provides access to functionality to create response plans that consist of facilities and their respective service areas. Plan creation methods currently part of this tool include an equal population partitioning feature [24], a spatial distance partitioning feature based on Voronoi tessellations [25], and an n-choose-k partitioning feature which chooses an optimal subset of facilities from a larger group. Plan analysis features facilitate exploration of factors which can affect plan feasibility including traffic conditions, population distance-to-assigned-facility, and population distribution. As new plan creation and analysis features are developed, access to them will be organized similarly.

The availability of plan creation and analysis features in the framework depends on the state of the system. For example, a scenario's facilities cannot be analyzed if they have not yet been created. Figure 8 illustrates the state-dependence of these features. A class

ScenarioState is implemented in the RE-PLAN Framework to track the state of scenarios and preserve the state-dependence of different features.

Modules in the framework have been designed to be interoperable by incorporating the ability to import and export features using widely used data standards. Sets of facilities may be imported or exported as Comma Separated Values (CSV) files. Many software packages commonly used in public health, disaster management, and city planning are capable of importing and exporting data as CSV files. Therefore, existing data may be imported into tools created with the RE-PLAN Framework, and RE-PLAN data may be exported to other commonly used software packages.

Functionality to export entire plans as standard ESRI shapefiles allows data to be shared and further analyzed by those with GIS expertise. These plans may be published by creating maps in software packages such as ArcGIS or may be used to create online, interactive maps using OpenMap, Google Maps, or Microsoft Bing Maps. Figure 9 shows a plan exported from RE-PLAN Bio-E and loaded into Google Earth for visual exploration. A feature which opens Google Earth and zooms to a location clicked on the RE-PLAN map has been implemented. This enables users to easily move between the Bio-E interface and Google Earth to facilitate interactive data visualization and exploration. Therefore, Bio-E facilitates not only plan analysis, but plan distribution and implementation as well.

5 Case Study

RE-PLAN may be used to build tools to create and analyze response plans for a variety of disaster scenarios. In the event of a biological emergency, local public health personnel must have solid, functional plans to provide medical countermeasures to the entire population of their jurisdiction within a short, mandated time-frame [3]. In this section, the application of the RE-PLAN Bio-E POD Analysis Tool to the biological emergency mitigation planning problem is demonstrated. Using the tool, a sample response plan was created and analyzed for Denton County, Texas.

According to the 2010 Census, Denton County has a population of 662,614. If a forty-eight hour time frame during which the entire population must be served is assumed, then at least 13,805 individuals must be served every hour. A response plan was created for Denton County consisting of twenty-five POD facilities, each with the capability of treating twenty individuals concurrently. Therefore, an average of twenty individuals must be treated at each facility every 130 seconds in order to meet time requirements. Service areas were generated for each facility, and the RE-PLAN analysis features were used to evaluate the feasibility of the response plan.

POD facility locations for the experiments were selected using the equal population partitioning algorithm. The algorithm generates a set of approximately equal-population sized service areas and places a facility at the geographic centroid of each service area. The facility locations generated by the equal partitioning algorithm were further analyzed by computing alternate service areas using the spatial distance partitioning method in a different RE-PLAN scenario. Both the equal population partitioning scenario and the spatial

distance partitioning scenario are analyzed and compared below using the RE-PLAN POD analysis features to study plan feasibility.

A facility analysis was performed on both scenarios using the assumptions of an average of four individuals per car and that a single individual takes 125 seconds to serve. Figures 10 and 11 show the results produced by the RE-PLAN Bio-E facility analyzer for the spatial distance partitioning scenario. Service areas were generated to assign populations to their closest facilities, thus minimizing distances which populations must travel for treatment. This service area generation method resulted in a population distribution among service areas which varied widely. Service time requirements for each facility also varied widely since time requirements are, in part, a function of population size. A blue horizontal line in the Time Requirements chart is placed at forty-eight hours to represent the mandated time frame. Green bars represent facilities capable of meeting the time frame, and red bars represent facilities which will exceed the time frame. Although parameters and assumptions may be modified to reduce time requirements in an effort to meet the time frame, the Parking Lot Maximum Entry/Exit Periods chart can reveal limitations in plans which may require significant plan modifications. These entry and exit periods represent how quickly cars must be able to simultaneously depart and park at the facility parking lot in order to support facility throughput.

In the equal population partitioning scenario, the same facility locations were used, but service areas were generated to equally distribute populations among the facilities. Figures 12 and 13 shows the results produced by the RE-PLAN Bio-E facility analyzer for the equal population partitioning scenario. The population distribution chart shows that a nearly uniform population distribution is achieved. This nearly uniform population distribution makes time requirements lower and parking lot entry/exit periods more uniform, thus making the scenario feasible. However, the distance the population must travel to the POD facility is longer than in the spatial distance partitioning method.

Maps of service areas and traffic analysis crossing points shown in Figure 14 were created assuming an average of four people per car at 5:00 p.m. on a weekday. Expected traffic situations are depicted using five classes on a color ramp from green to red. Green crossing points indicate low levels of expected traffic, while orange and red points indicate points of considerable congestion. Areas of higher expected traffic indicated in these maps can be further investigated using the traffic toolbox to determine whether they result from implementation of the response plan or from the estimations of *business as usual* base traffic. Modifying POD facility locations, adding additional facilities, or recalculating the service areas are strategies which may be employed in an effort to reduce load on high-traffic areas.

6 Discussion

The RE-PLAN Framework serves as a foundation upon which new algorithms for facility location and emergency response may be implemented. The framework provides access to region-specific data sets by aggregating separate regional data collections as needed based on user input. Therefore, new features added to the framework are automatically region

independent. The framework employs naming conventions to standardize how data is stored on the database and manages the creation, saving, and loading of plan *scenarios*. A graphical user interface is provided by the framework which includes an interactive map to allow users to develop and view their plans. These aspects of the RE-PLAN Framework facilitate the quick translation of new analysis and optimization algorithms into features available in the framework. Sets of these features can be used to create tools usable by preparedness planners.

The Bio-E tool created using the RE-PLAN Framework answers important questions regarding the implementation of specific response plans. The tool harnesses large amounts of data to estimate conditions at, and requirements for, each facility during implementation of specific mitigation plans. These conditions and requirements include the projected traffic situation at each facility's parking lot, the amount of time each facility requires to serve its assigned population, and the infrastructure needed by each facility to serve its assigned population. Further, the tool facilitates the analysis of the traffic situation on the transportation network resulting from the implementation of specific response plans.

The RE-PLAN Framework was developed in collaboration with local PHPP. Their suggestions and comments drove developments in the methodology and implementation. Graphical interfaces were incorporated into the framework to allow use without the need for GIS or computer programming expertise. Although large amounts of data are used and may be accessed through the framework, the design of graphical displays focused on specific aspects of the response plan analysis methodology. Therefore, while the data used by the computational methods may be accessed through the graphical interface, it is hidden by default to avoid clutter and confusion.

The two response plans described above were created as separate RE-PLAN Bio-E scenarios using different partitioning tools. Although both plans used the same set of facilities, analysis of the results underscores the differences in the partitioning tools. The choice of which partitioning tool to use is left up to local preparedness planners whose expertise is instrumental in feasible response planning. Partitioning algorithms are incapable of predicting the facility each individual will visit during actual implementation of a response plan. Nonetheless, they assist planners in assigning individuals to facilities with the goals of influencing which facility each individual will visit and facilitating quantitative resource distribution decisions.

A version of the RE-PLAN Bio-E tool has been created for and deployed at a local county public health department. County public health officials have been trained to use this tool for analyzing their mitigation plans. Analysis performed at the county using this tool has led to the continued revision and modification of response plans. Local stakeholders have been trained regarding the modified plans, and preparations to implement these plans have been conducted.

Although the RE-PLAN Framework was originally designed for a specific emergency mitigation planning problem involving the accidental or deliberate release of anthrax, its modules facilitate data-driven analysis pertinent to a variety of different emergency response

scenarios. Existing modules may be modified and new modules may be added to create tools for different mitigation planning problems. For example, the RE-PLAN Framework can be extended for planning resource allocations for mass evacuations or for the distribution of life-sustaining supplies following a natural disaster. Further, since the framework uses the Model-View-Controller design pattern, it can be modified more easily to create a new browser-based interface or cloud-based back-end in order to scale RE-PLAN for use on larger problems and/or by larger audiences.

7 Limitations

As with all computational models, fidelity is limited by the accuracy and availability of underlying data. The RE-PLAN Framework uses several sets of data for which availability or currency may be problems. For each specific region, population data, transportation network data, and traffic count observation data may have to be obtained from different sources. Corresponding data sets for different regions are likely to be organized differently. Therefore, data must be prepared for use by the framework by manipulating it into a specified format.

Geographic population distribution data needed by the RE-PLAN Framework is available in the United States as polygons from the U.S. Census Bureau. Aggregation of population counts at the polygon level obscures the distribution of the population within the given polygons. Therefore, measures of distance involving the population are calculated using the Euclidean distance and the geometric centroid of a population polygon. Use of data from the decennial census may be undesirable as it is updated infrequently. Another product from the Census Bureau with more frequent updates is the American Community Survey (ACS). While ACS data may be more up-to-date, it also has a lower spatial granularity. If more accurate or current population distribution data is available from other sources, the RE-PLAN Framework is capable of incorporating these data into its models. The only limitation regarding which population distribution data may be used is that the framework is only designed to incorporate vector (not raster) data sets. However, raster data can be transformed into vector data using a wide variety of available tools.

Transportation network data may be incomplete or out-of-date for certain areas. Smaller neighborhood roads may be excluded from available data sets. However, these smaller roads do not greatly affect traffic analysis. Although newer links on the transportation network may not be included in available data, local planners who are using the tool will likely be familiar with the roads in their area. Attributes available for each link of the transportation network may differ from location to location. To address this problem, a variety of methods have been developed to use the framework with the available attributes. Nonetheless, acquisition and manipulation of local data remains an unavoidable task for those performing RE-PLAN-based tool deployments in new areas.

Traffic count observation data may be unavailable for the vast majority of links in the transportation network. The RE-PLAN Framework addresses this sparse data problem by classifying the links in the transportation network and assigning links of the same class the same traffic loads per lane. Traffic count observation data may contain inaccuracies due to

the methods used in the collection of this data. If this data is too sparse or entirely unavailable for an area, the RE-PLAN Framework may be trained on data from a similar, but different, area. Nonetheless, it may be difficult to accurately determine whether a specific traffic count represents conditions of high traffic speed and low traffic density, or of low traffic speed and high traffic density [26]. A recent and innovative approach to estimating traffic uses crowdsourcing of anonymous data from mobile phone users to observe current speeds on major roads and highways in real time [27]. If these data become publicly available, they may be used in conjunction with traffic count observation data to better understand traffic patterns on roads which may affect plan feasibility.

The complexity of the response planning problem makes identification of the absolute best response plan for a region prohibitively expensive. Included optimization algorithms focus on maximizing specific aspects of a plan. These algorithms are likely to include a trade-off, where maximizing specific aspects of a plan may lead other aspects of the plan to be less efficient. Nonetheless, these new optimization tools provide response planners with the ability to increase the efficiency of their response plans using their knowledge of local communities.

Estimations regarding traffic conditions and time requirements to serve populations under specific assumptions represent best-case situations where the population departs their homes and travels to their assigned facilities at a uniform rate. During actual implementation of plans, it is unlikely that this uniform departure and arrival pattern will occur. Rather, transportation and facility infrastructure may be overwhelmed during the first few hours of mitigation, ultimately leading to the under-utilization of mitigation resources. Nonetheless, tools constructed using the RE-PLAN Framework represent an opportunity for planners to build their plans in the context of a rich set of data which would otherwise be unavailable.

Models in the RE-PLAN Framework convert numbers of individuals directly into numbers of cars on the road network or in parking lots. The assumption that individuals will be able to travel to and from their assigned facilities using private vehicles may be unrealistic. Additional data regarding the availability of private vehicles within specific populations must be integrated into the analysis to address this issue. Modifications to the framework for modeling public transit infrastructure may be required.

Acknowledgments

The project described was supported by Grant Number NIH 1R01LM011647-01 and Grant Number NIH 1R15LM010804-01 from the National Institutes of Health. An earlier version of this paper has been published in the Proceedings of the International Conference on Bioinformatics and Computational Biology (BIOCOMP'11: July 2011, USA), Editors: Hamid R. Arabnia and Quoc-Nam Tran; ISBN # (for set): 1-60132-172-4.

Biographies

Martin O'Neill II received the M.S. degree in Computer Science from the University of North Texas (UNT) in 2009 and is currently A.B.D. in Computer Science at UNT. Since 2010, he has been a research analyst in the Center for Computational Epidemiology and Response Analysis (CeCERA) at UNT. His research interests include spatial algorithm development, computational response plan analysis, and infectious disease modeling.

Armin R. Mikler received his Diplom in Informatics from Fachhochschule Darmstadt, Germany in 1988. He received a Fulbright Scholarship in 1986 and completed his M.S. and Ph.D. in Computer Science at Iowa State University in 1990 and 1995 respectively. Since 1997, Dr. Mikler is a faculty member in Computer Science at the University of North Texas (UNT) with joint appointment in the Department of Biological Sciences. Dr. Mikler has established and is the director of the Center for Computational Epidemiology and Response Analysis (CeCERA) at UNT and leads the Computational Epidemiology Research Laboratory (CERL).

Saratchandra Indrakanti received the M.S. degree in Computer Science from the University of North Texas (UNT) in 2012 and is currently a Ph.D. student in Computer Science at UNT. He is a member of the Center for Computational Epidemiology and Response Analysis (CeCERA) at UNT. His research includes developing methods and algorithms that facilitate emergency response planning and analysis.

Chetan Tiwari received his Ph.D. in Geography from The University of Iowa, USA in 2008. Dr. Tiwari is a faculty member in the Department of Geography at the University of North Texas (UNT). He is also affiliated with the Center for Computational Epidemiology and Response Analysis (CeCERA) at UNT. His research interests are in the applications of Geographic Information Science and spatial analysis methods for problems in Public Health.

Tamara Jimenez received the Diplom in Informatics from the University of Passau, Germany in 2007 and the Ph.D. in Computer Science from the University of North Texas (UNT), United States in 2010. Dr. Jimenez was a post-doctoral research associate at the Center for Computational Epidemiology and Response Analysis (CeCERA) at UNT and is currently employed at UNT as a Lecturer in Theoretical Computer Science. Her research includes the design of algorithms for the analysis and optimization of bio-emergency response plans, as well as the development of methodology for generation of synthetic geographic regions.

References

1. Schneider T, Mikler AR. RE-PLAN: A Computational Tool for Response Plan Analysis. *International Journal of Functional Informatics and Personalised Medicine*. 2010; 3(no. 2):103–121.
2. Schneider, T.; Mikler, AR.; O'Neill, M, II. Computational Tools for Evaluating Bioemergency Contingency Plans; *Proceedings of the 2009 International Conference on Disaster Management*; 2009.
3. Centers for Disease Control and Prevention. Receiving, Distributing, and Dispensing Strategic National Stockpile Assets: A Guide for Preparedness, Version 10.02. 2006
4. O'Neill, M., II; Mikler, AR.; Schneider, T. BIOCAMP'11. Las Vegas, NV: 2011. An Extensible Software Architecture to Facilitate Disaster Response Planning.
5. Kaplan EH, Craft DL, Wein LM. Emergency response to a smallpox attack: the case for mass vaccination. *Proceedings of the National Academy of Sciences of the United States of America*. Aug.(no. 16):10935–10940.
6. Wein LM, Craft DL, Kaplan EH. Emergency response to an anthrax attack. *Proceedings of the National Academy of Sciences of the United States of America*. Apr.(no. 7):4346–4351.
7. Baccam P, Boechler M. Public health response to an anthrax attack: an evaluation of vaccination policy options. *Biosecurity and bioterrorism : biodefense strategy, practice, and science*. 2007 Mar; 5(no. 1):26–34.
8. Herrmann, JW.; Lu, S.; Schalliol, K. A Routing and Scheduling Approach for Planning Medication Distribution; *Proceedings of the 2009 Industrial Engineering Research Conference*; 2009.
9. Lee, YM. Analyzing Dispensing Plan for Emergency Medical Supplies in the Event of Bioterrorism; *Proceedings of the 2008 Winter Simulation Conference*; 2008. p. 2600-2608.
10. Rotz LD, Hughes JM. Advances in detecting and responding to threats from bioterrorism and emerging infectious disease. *Nature medicine*. 2004 Dec; 10 Suppl(no. 12):S130–S136.
11. U.S. Department of Health & Human Services. FY 2005 Performance and Accountability Report. 2005
12. Willis, HH.; Nelson, C.; Shelton, SR.; Parker, AM.; Zambrano, JA.; Chan, EW.; Wasserman, J.; Jackson, BA. Initial Evaluation of the Cities Readiness Initiative. Santa Monica, CA: 2009.
13. Lee EK, Chen C-H, Pietz F, Benecke B. Modeling and Optimizing the Public-Health Infrastructure for Emergency Response. *Interfaces*. 2009 Oct; 39(no. 5):476–490.
14. Giovachino M, Calhoun T, Carey N, Coleman B, Gonzalez G, Hardeman B, McCue B. Optimizing a District of Columbia Strategic National Stockpile Dispensing Center. *J Public Health Manag Pract*. 2005; 11(no. 4):282–290. [PubMed: 15958925]
15. Monahan C, Franco JA. From Research to Practice : Creating a Virtual POD Using a Collaborative Virtual Environment. 2011

16. Center for the Advancement of Distance Education at the School of Public Health of the University of Illinois at Chicago. The POD Game.
17. World Health Organization. Geneva, Switzerland: 2007. The World Health Report 2007: A Safer Future - Global Public Health in the 21st Century.
18. Chen H, Zeng D, Yan P. Syndromic Surveillance Data Sources and Collection Strategies. *Infectious Disease Informatics*. 2010; 21(no. 1):33–48.
19. Krenzelok E, MacPherson E, Mrvos R. Disease surveillance and nonprescription medication sales can predict increases in poison exposure. *Journal of medical toxicology*. 2008 Mar; 4(no. 1):7–10. [PubMed: 18338303]
20. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature*. 2009 Feb; 457:1012–1014. [PubMed: 19020500]
21. Hulth A, Rydevik G, Linde A. Web queries as a source for syndromic surveillance. *PloS one*. Jan. (no. 2)
22. Corley, CD.; Mikler, AR.; Singh, KP.; Cook, DJ.; Worth, F.; Science, C. Proceedings of the 2009 International Conference on Bioinformatics and Bioengineering (BIOCOMP09). Las Vegas, NV: 2009. Monitoring Influenza Trends through Mining Social Media.
23. Brandeau ML, McCoy JH, Hupert N, Holty J-E, Bravata DM. Recommendations for modeling disaster responses in public health and medicine: a position paper of the society for medical decision making. *Medical decision making : an international journal of the Society for Medical Decision Making*. 2009; 29(no. 4):438–460. [PubMed: 19605887]
24. Jimenez T, Mikler AR, Tiwari C. A Novel Space Partitioning Algorithm to Improve Current Practices in Facility Placement. *IEEE transactions on systems, man, and cybernetics. Part A, Systems and humans : a publication of the IEEE Systems, Man, and Cybernetics Society*. 2012 Sep; 42(no. 5):1194–1205.
25. Aurenhammer F. Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Computing Surveys*. 1991 Sep; 23(no. 3):345–405.
26. U.S. Department of Transportation Federal Highway Administration. The 1985 Highway Capacity Manual - A Summary. Washington D.C.: Office of Traffic Operations; 1986.
27. Barth D. The bright side of sitting in traffic : Crowdsourcing road congestion data. 2009

Figure 1.
RE-PLAN Framework Architectural Components and Work Flow Diagram

Figure 2.

An example of traffic displayed on the map interface is provided in 2a, and a screen capture of the Traffic Toolbox is provided in 2b

	Areas of Interest Selected	PODs Selected	Catchment Areas Created	Logistics Calculated
Roads	T	T	T	T
Population Blocks	T	T	T	T
County Outlines	T	T	T	T
PODs		T S	T S	T S
Catchment Areas			T S	T S
Block-to-POD Assignments			T S	T S
Traffic Rings				T S
Crossing Points				T S

T Temporary copy of table exists **S** Saved copy of table created upon saving of scenario

Figure 3.

Tables saved to the database at different states

Figure 4.
Breaking a service area into rings of proximity to the facility

Figure 5.
How the population travels through the rings of proximity to the facility

Figure 6.

Using only three rings of proximity, the rings are divided into segments to model population flow across roads. Each segment is uniquely identified by two numbers. The first number represents the ring in which it is located, and the second uniquely identifies this segment in that ring.

Figure 7.

Using only three rings of proximity, the segments shown in Figure 6 are used with the road network to create a graph model. In this model, edges represent specific links in the road network where expected traffic may be examined.

Figure 8.
State-dependence diagram and analysis method requirements of the RE-PLAN Framework

Figure 9.
Example of RE-PLAN data displayed in Google Earth

(a) Population Distribution

(b) Time Requirements

Figure 10.

Population Distribution and Service Time Requirements generated using the Spatial Distance Partitioning method

(a) Population Distance to Facility

(b) Parking Lot Maximum Entry/Exit Periods

Figure 11.

Parking Lot Requirements and Population Distance to Facility generated using the Spatial Distance Partitioning method

(a) Population Distribution

(b) Time Requirements

Figure 12.

Population Distribution and Service Time Requirements generated using the Equal Population Partitioning method

(a) Population Distance to Facility

(b) Parking Lot Maximum Entry/Exit Periods

Figure 13.

Parking Lot Requirements and Population Distance to Facility generated using the Equal Population Partitioning method

Figure 14.
Service Areas and Traffic Estimations

Table 1

Facility Attributes

Attribute	Example Value(s)	Client Data Type	Displayed in RE-PLAN Interface
fid	0	int	no
id	1	int	yes
name	"University Union"	String	yes
address	"1155 Union Circle"	String	yes
city	"Denton"	String	yes
zip	"76203-5017"	String	yes
status	"active" or "closed"	Boolean	yes
comments	"Unused areas of this large building need to be secured"	String	yes
numbooths	23	int	yes
type	"public" or "private"	Boolean	yes
latitude	33.210863	double	yes
longitude	-97.146971	double	yes

Table 2

Range of values for parameters of Traffic Toolbox

Parameter	Possible Values
People per car	1–5
Weekday / Weekend	Weekday OR Weekend
Time of day	0:00 – 23:45 at 15-minute intervals
Distribute traffic over	1–240 hours