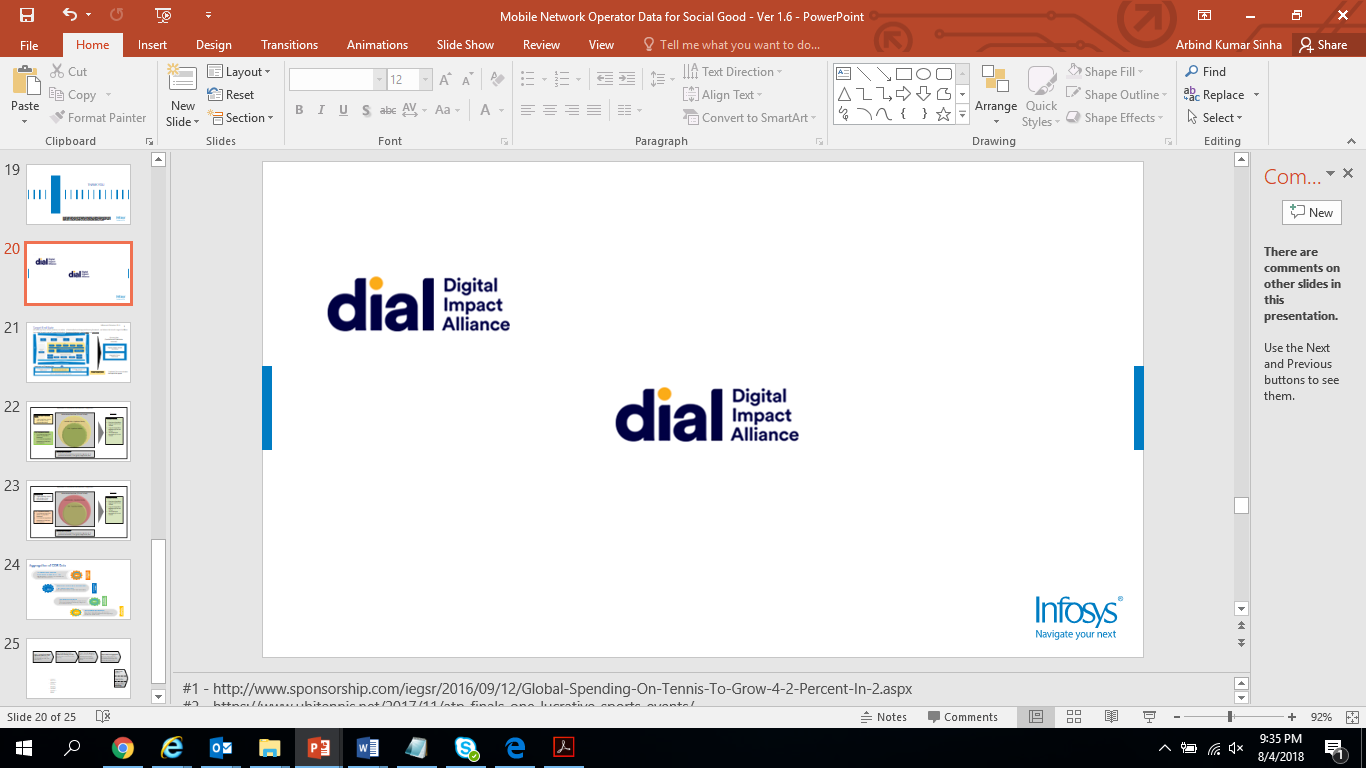
**Malawi Use Case**

***Approach - Optimization of Health Posts***

**Version – 1 0**







Document Information and Revision History-

| **Version** | **Date** | **Author** | **Reviewer** | **Remarks** |
| --- | --- | --- | --- | --- |
| 0 1 | 14/01/2019 | Infosys | Infosys | Draft version |
| 1 0 | 16/01/2019 | Infosys | Infosys | 1st Baseline Version |
|  |  |  |  |  |

**Table of Contents**

[1.0 Introduction 4](#_Toc535256420)

[2.0 Approach 4](#_Toc535256421)

[2.1 Problem Statement 4](#_Toc535256422)

[2.2 Assumptions 4](#_Toc535256423)

[2.3 Optimization Approach 5](#_Toc535256424)

[2.4 Datasets used 7](#_Toc535256425)

[2.5 Mathematical Formulation 8](#_Toc535256426)

[3.0 Reference 9](#_Toc535256427)

# 1.0 Introduction

This purpose of this document is to formulate an optimization problem for the allocation of 900 health posts within the next five years across all districts in Malawi.

This document will cover the approach, assumptions made, datasets required, mathematical formulation of the objective function and the constraints / limitations involved as a part of optimization of 900 health posts.

# Approach

## Problem Statement

The Malawi Ministry of Health, in collaboration with the Digital Impact Alliance, is planning to roll out 900 health posts over next five years across all 28 districts of Malawi. These include both upgrades to existing facilities and the construction of new buildings to expand access, particularly in rural and remote areas, with an emphasis on the provision of primary health care.

As part of its draft Capital Investment Plan, the government of Malawi has drawn up proposed allocations of new facilities using the following 4 criteria:

1. Catchment Population
2. Distance to Nearest Existing Health Facilities
3. Facility Accessibility (high, medium, low)
4. Preferred year for work to take place as expressed by the District Health Monitoring Team

This report proposes to complement the above approach by incorporating population projections, as-well as short term and long-term population movement patterns, using these to optimize the placement of future health posts.

To approach this problem statement, certain assumptions are made as highlighted below.

Assumptions and approach might be revisited based upon availability of new inputs on problem statement.

## Assumptions

* All proposed new health posts are intended to serve a fixed catchment size ‘C’
* Policy preferences are those outlined in the investment plan
* Populations clusters can be assigned to no more than one health-post
* Population clusters always opt for the nearest health post.
* Movement or shifting of existing health facilities are not allowed.
* Travel distance is considered a valid proxy of the cost and time to travel when optimizing the location of health facilities.
* Capacity of a health post is considered based on the maximum number of patients that can be served based on its capacity constraint.
* There is a cap on # of health posts that can be built in a year.
* As in the investment plan, the cost to construct health posts across locations is assumed uniform across districts ($36,573), unless otherwise specified
* Population characteristics and disease burden are similar across clusters, pending the incorporation of additional data on health outcomes.

## Optimization Approach

In this approach, we may assume this problem statement to be a location-allocation problem – with minimizing the uncovered population given binding constraints on the number of facilities. To proceed with this approach, we chalk out the steps as discussed below.

Note: Step 1 is done at TA level.

Step1: Obtain the population densities from phase 1 and project it in the following way for a year:

* Calibrate population density estimate with population density by World Pop:

k =

where k is the calibration factor, P is the world pop population and Ṗ is estimated population

Then αnew = αk which will be the new alpha

Original Power equation:

New Power equation with calibrated α:

where is population density and is call density for a particular TA

* Calculate growth rate in call density at TA level from the 2016 - 2017 CDR data. This growth rate will be used to project call density for next year (at TA level).

Growth rate =

Year(i) call density=Year(i-1) call density (1+ Growth rate)

Here i stands for year 2018 to 2023

* The above projected call density will be input for new power equation to calculate population density for next year.
* Now compute population from population density using:

Estimated Population count = Population density \* Area (in sq km)

* This population count will be adjusted according to the long term population movement (computed in phase 1) using:

Estimated Population count \* (1 + % change in average Netflow at TA level)

Step 2: Get Uncovered area and uncovered population

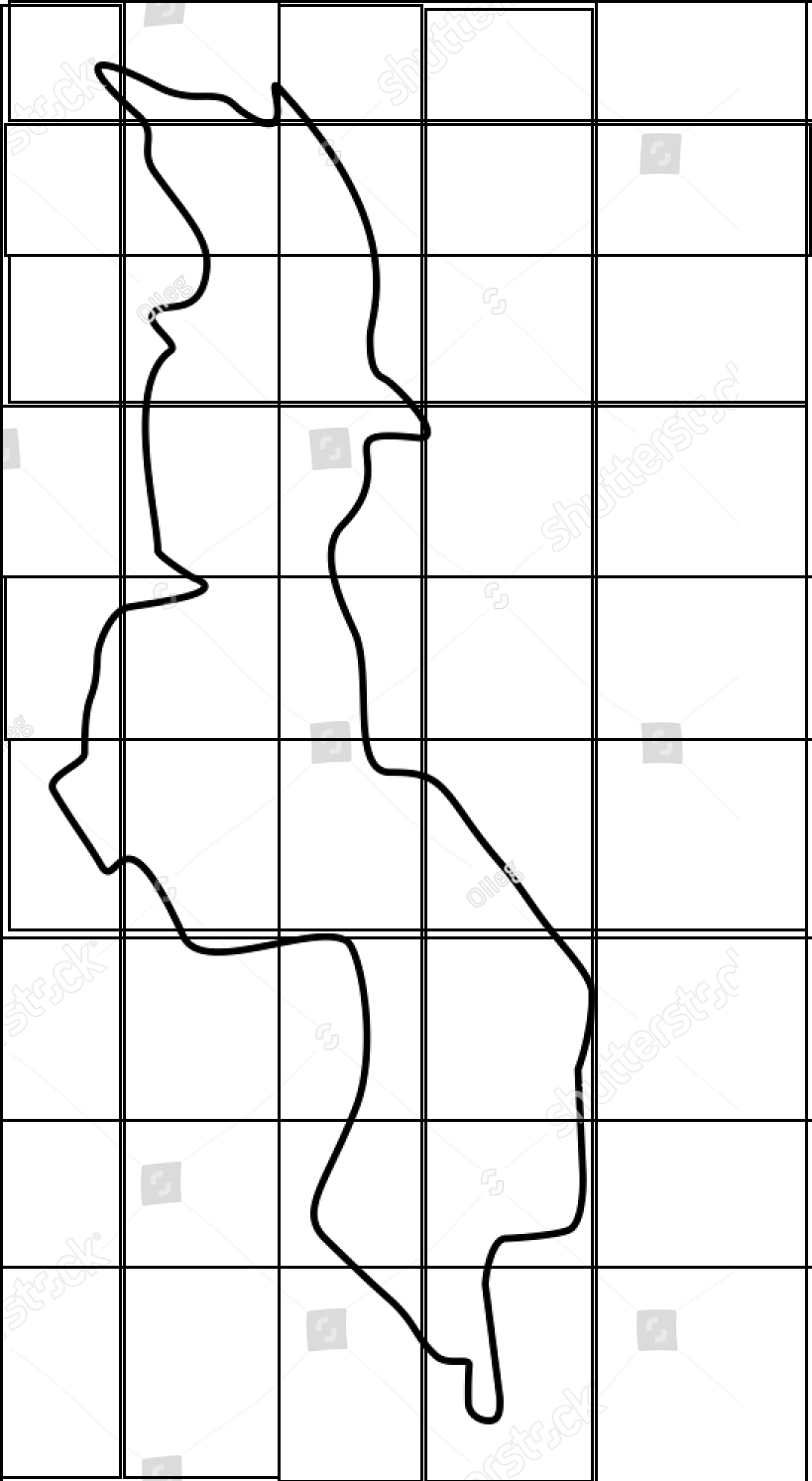
* We have ‘UNICEF catchment shape file’ - UNICEF conducted an extensive survey of every operating clinic in the country and calculated the relevant catchment area. This catchment area constitutes the distance such that a patient would have to walk no more than 5km to reach a clinic. In its analysis, it accounted for the road networks, topography and potential for flooding in calculating the 5km catchment area for each clinic.

(“Thanks to UNICEF Malawi country office for sharing”)

* By using this shape file we will get uncovered area.
* **For first year we will be using this given shape file to get uncovered area and uncovered population. But after that for next four years, for new health post locations which we have predicted for each previous year, we will draw circle of 5 km radius and then find uncovered population.**

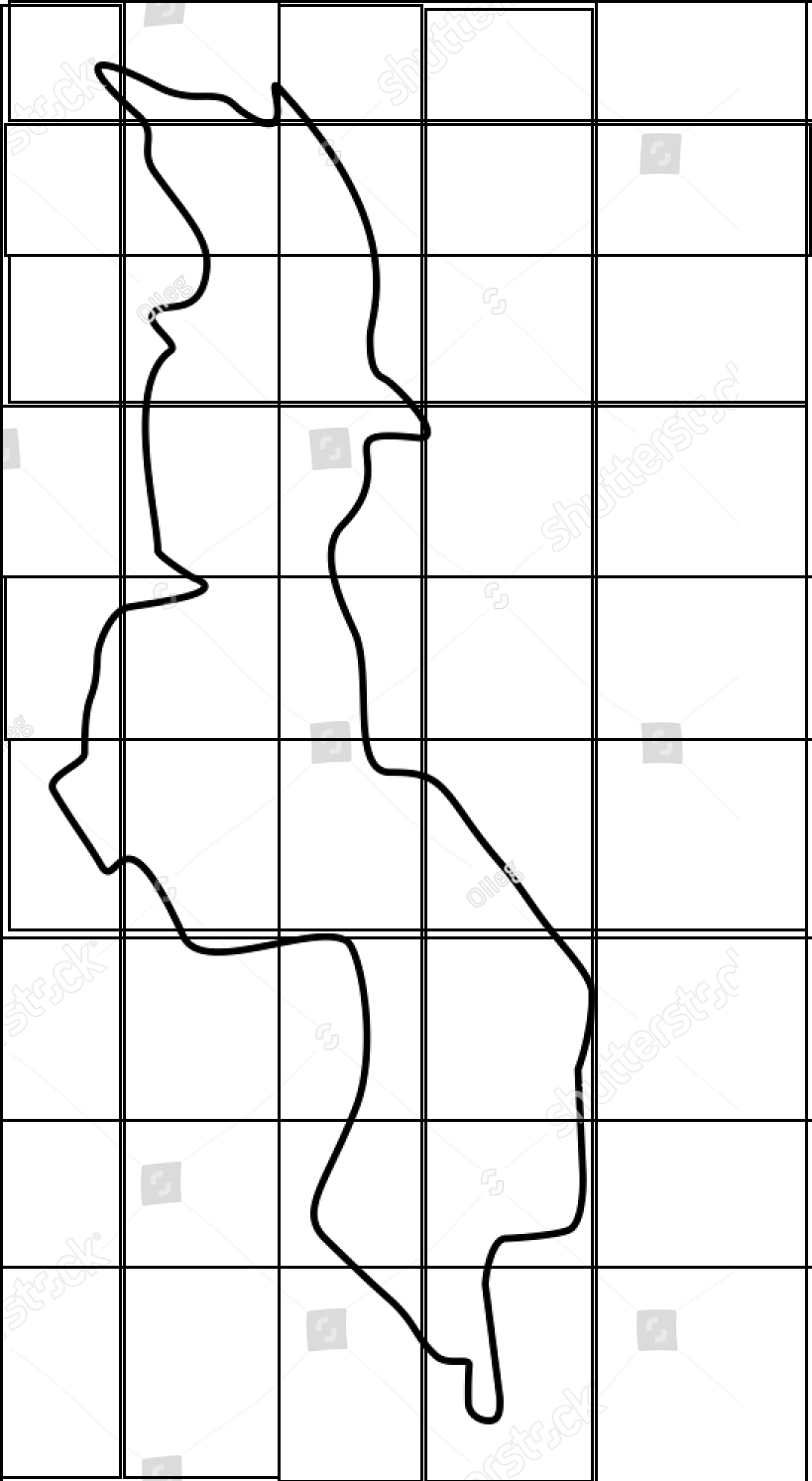
Step 3: Analogous estimation of population

* Overlay grid (eg:1 sqkm) on worldpop raster file and calculate the % of population in each grid.
* This % is then multiplied with estimated population from Step 1 to get distribution of estimated population.



Step 4: Formation of Clusters

* Identify centroid of each grid.
* Cluster grids (by measuring the shortest path between centroids) till you reach catchment population (Cj). This gives one cluster.
* Continue the above till each grid is a part of one cluster.



Step 5: Finding demand points

* These clusters give **uncovered population** (Popu).
* Centroids of these clusters is the **demand points**.

Step 6: Select all the existing health posts as required and new probable health post locations as candidates for **Facilities**. These new probable health post locations (GVH-Tower) will be identified using short term population movement at TA level.

Step 7: Using **demand points**, facilities and impedance (a distance beyond which a patient cannot travel to the health post, i.e. 5 km) obtain the location of new health posts for first year.

Now repeat the above steps from 1 to 6 for next four years. If a potential candidate (health post location) for a given year turns out to be a favorable location, then for the next year the health post becomes a required candidate.

## Datasets used

* List of existing health facilities with location information.
* Facility burden (# of patients per facility) for each facility
* Population density (from phase 1) along with short term (intra-day) and long-term movements (computed) for each administrative unit for each year
* Shape file of UNICEF to know the uncovered population.

(“Thanks to UNICEF Malawi country office for sharing….”)

## Mathematical Formulation

*Objective function***:** Min **Popu - ∑∑ wj \* ∑ Pjjt \* ‘Cj’**

*Constraints:*

∑∑Pjt ≤900 (900 is the maximum number of health posts that can be deployed across all administrative units over the 5 years)

∑Pjt ≤ H (Threshold “H” is the maximum number of health facilities that can be deployed across all administrative units in the year “t”).

A population centroid “i” in administrative unit “j” must be mapped maximum to 1 existing health facility.

∑ Fixjt (summed over “x" existing facilities for a given “i” at adm unit “j” for year “t”) ≤ 1

For a given existing facility “x” in administrative unit “j” must be mapped maximum to 1 population centroid.

∑ Fixjt (summed over for “i" population centroids for a given “x” at adm unit “j” for year “t”) ≤ 1

If a centroid “i” in administrative unit “j” is getting mapped to a particular existing facility “x” and similarly if the existing health facility is getting mapped to the same population centroid “i” then candidate health post (proposed) should not be posted for that centroid.

Pijt=0 ; If the above case is not true,

else Pijt=1

Here, we need to consider deploying a health post. Among the proposed health post candidates, select the location which is closest to the centroid “i”.

Pjt= ∑Pijt (sum over all the centroids “i” for administrative unit “j” for the year “t”)

*Where:*

Pjt: Health post “p” in administrative unit “j” in the year “t”.

Xjt: Existing health post “x” in administrative unit “j” in the year “t”.

Dixjt: distance “d” from existing health post “x” to population cluster centroid “i” in administrative unit “j” in the year “t”.

Dipjt: distance “d” from candidate (proposed) health post “x” to population cluster centroid “i” in administrative unit “j” in the year “t”.

Fixjt: is equal to 1 for the pair of centroid “i” and existing facility “x” which has shortest distance-

“Dixjt” ≤ D (impedance distance). If no pair are between “D” then zero.

Fipjt: is equal to 1 for the pair of centroid “i” and probable candidate facility “p” which has shortest distance- “Dipjt” ≤ D (impedance distance). If no pair are between “D” then zero.

Pijt= 1, if there is requirement for new facility. Else 0.

Calculating Weights at administrative unit level

**wj** = administrative unit level weights to reflect population movement patterns

While calculating Weights we need to consider disease burden on each TA.

Following steps to be followed to calculate weights-

* For each health facility, calculate the average number of people under each disease across all the months in 2016
* Calculate number of people under each disease at TA level by aggregating the above calculated average across health posts
* Calculate the % of people under each disease at TA level using 2016 year’s estimated population for each TA
* Calculate disease burden by taking weighted sum on all the % of people across each disease at TA level
* Calculate percentile rank for above calculated value across TA, which will be used as weights for the optimization model

Since there might be multiple centroids without a health post allocated to it, priorities might be declared based on the above set of information which we will be considering in calculating Weights.

# Reference

**(for basic approach):** Stummer, C., Doerner, K., Focke, A. et al. Health Care Management Science (2004) 7: 63. https://doi.org/10.1023/B:HCMS.0000005399.23600.69