

Taller01

Taller de Sistemas Operativos
Escuela de Ingeniería Informática

Sebastián Lillo Núñez

sebastian.lillon@alumnos.uv.cl

Resumen. En este trabajo se realiza un análisis de estadística descriptiva frente a los datos entregados por una simulación que muestra variada información sobre las personas que se encuentran en Iquique, ya sean residentes o visitantes. Se plantea una metodología para establecer luego un diseño para la implementación en el lenguaje GNU Bash, se realizan distintas consultas según solicite la tarea específica para obtener resultados y luego exponerlos en un archivo de texto plano. De esta manera se concluye que el realizar un diseño previo es un beneficio para luego realizar la implementación de manera estructurada y ordenada, respetando siempre una directriz de trabajo.

1. Introducción

Nuestro país, es reconocido como uno de los países más sísmicos de la Tierra junto a México, Japón, Nepal y Nueva Zelanda; estos eventos telúricos son los causantes de los tsunamis, definidos como una serie de ondas oceánicas generadas por un disturbio impulsivo en el océano, o en un pequeño y conectado cuerpo de agua. Definido de este modo, el término incluye ondas generadas por desplazamientos abruptos del fondo oceánico, causados por terremotos, deslizamientos de tierra submarinos o de la línea de la costa, erupciones volcánicas y explosiones [1]. El terremoto de mayor intensidad registrado se localizó en el sur de Chile, en la región de la Araucanía, en las cercanías de Traiguén siendo ese lugar el epicentro del sismo, con una magnitud Richter de 9.5, lo que provocó una ola superior a los 10 metros de altura [2].

El objetivo de este informe es realizar un análisis sobre los datos recolectados de una serie de simulaciones de evacuación que se realizaron en la ciudad de Iquique para la prevención de un Tsunami para poder generar una estadística descriptiva de los datos. Para lograr esto se realizará un script en Bash para trabajar los datos que entregaron las simulaciones. Un script es un programa que contiene una lista de instrucciones que son ejecutadas por un intérprete y en esta instancia el lenguaje de comandos a utilizar es GNU Bash. Es por esto que el documento cuenta con una sección de procedimiento donde se darán a conocer los datos utilizados, metodología empleada y diseño de la implementación para realizar el taller, una sección para los resultados de la sección anteriormente y finalmente una sección de las conclusiones obtenidas en el desarrollo del taller.

2. Procedimiento

2.1 Descripción de los datos

Los datos utilizados provienen de lo entregado para la realización de este taller, los cuales están formados por 3 modelos de personas: Residentes, Visitantes Tipo 1 y Visitantes Tipo II. En la tabla 1 se muestra la descripción de cada modelo de persona.

Tipo de Persona	Descripción
Residente	Persona que vive en la ciudad y conoce su zona segura.

Visitante Tipo I	Visitante que durante la evacuación logra determinar su zona segura.
Visitante Tipo II	Visitante que durante la evacuación logra no determinar su zona segura.

Tabla 1 muestra descripción de los grupos de personas simuladas

A su vez los datos se organizan en 3 archivos por cada simulación realizada de la siguiente manera:

Campo	Descripción	Ejemplo
numExperiment	Número del experimento.	0 (Corresponde al código 000)
Tsim	Tiempo de simulación.	3600: segundos de tiempo real que se simuló.
CalibrationTime	Tiempo de calibración del simulador.	100: segundo de tiempo de real que se destina a la calibración inicial de las personas.
Residents	Cantidad de Residentes simulados.	69000
Visitors	Cantidad de Visitantes simulados.	6000
timeExecMakeAgents	Tiempo real que demora en crear en memoria las personas simuladas.	28252: milisegundos
timeExecCal	Tiempo real que demora la calibración de las personas.	33157: milisegundos
timeExecSim	Tiempo real que toma la simulación.	182800: milisegundo
maxMemory	Costo espacial del simulador.	288236: Kbyte
agentsMem	Memoria utilizada por las estructuras de datos relacionados con las personas.	33648: Kbyte

Tabla 2 muestra descripción del archivo 1

Campo	Descripción	Ejemplo
numExperiment	Número del experimento.	0 (Corresponde al código 000)
id	Identificador de la persona simulada.	15
model https://reuna.zoom.us/j/875439493	Identificar del modelo de la persona. 0: Residente 1: Visitante tipo I 2: Visitante tipo II	0
groupAge	Identificador del grupo etario de la persona. 0: G1, 1: G2, 2: G3, 3: G4	2
safeZone	Identificador de la zona segura de la persona. Z1, Z2, Z3, Z4, Z5 Identificadores de zonas seguras. NA: la persona no tiene zona segura asignada.	Z1
distanceToTargetPos	Distancia a la que quedó la persona de su objetivo inicial.	13.871407 metros
responseTime	Número aleatorio que representa cuánto tiempo se demoró la persona en tomar la decisión de evacuar, desde que dieron la orden de evacuar. Si es -1, entonces la persona nunca pudo evacuar.	209.411742 segundos
evacTime	Tiempo que la persona se demoró en llegar a la zona de evacuación. Si es 0, entonces la persona nunca llegó a la zona de evacuación.	2300 segundos

Tabla 3 muestra descripción del archivo 2

Campo	Descripción	Ejemplo
numExperiment	Número del experimento.	0 (Corresponde al código 000)
timeStamp	Tiempo de la medición.	60
usePhone	Cantidad de persona que utilizaron el teléfono móvil en el tiempo especificado.	4381

Tabla 4 muestra descripción del archivo 3

2.2 Método

La metodología a emplear está dada por los siguientes puntos: 1) Lectura de los datos en los archivos por cada simulación; 2) Procesamiento de los datos; 3) Estadística descriptiva; 4) Agrupar resultados por simulación; 5) Presentación de resultados.

2.2 Diseño

La disposición de los archivos se encuentra segmentada en directorios con una etiqueta que representa el número de simulación como muestra la figura 1.

```
sebastian@tallerso:~/TSSOO-taller01/U1/taller01$ cd simulation-results/
sebastian@tallerso:~/TSSOO-taller01/U1/taller01/simulation-results$ ls
000 001 002 003 004 005 006 007 008 009 010
```

Fig. 1 Muestra cómo se disponen los directorios con los datos en los archivos respectivos.

Cada directorio posee 3 archivos de texto plano, llamados executionSummary-NNN.txt, summary-NNN.txt y usePhone-NNN.txt, y los campos están separados el símbolo “:” los cuales contienen los datos que son la fuente de análisis para este taller. Por tanto, se necesita de un ingreso de parámetros para ejecutar el script en la línea de comandos indicando en qué directorio debe inspeccionar los elementos. Se plantean 3 requerimientos en forma de tareas para analizar los datos y obtener resultados referentes a una solicitud en particular. Además, se utilizarán archivos intermedios de existencia volátil para poder organizar y operar los datos leídos en los archivos.

Para poder obtener los resultados de la estadística descriptiva debemos de leer los datos de los archivos de forma individual y por cada uno de ellos se deben realizar los cálculos pertinentes para obtener medidas como la media, valor máximo y mínimo, siendo almacenados en un archivo nuevo con su nombre respectivo, los cuales son: metrics.txt, evacuation.txt, usePhone.txt; como muestra la figura 2.

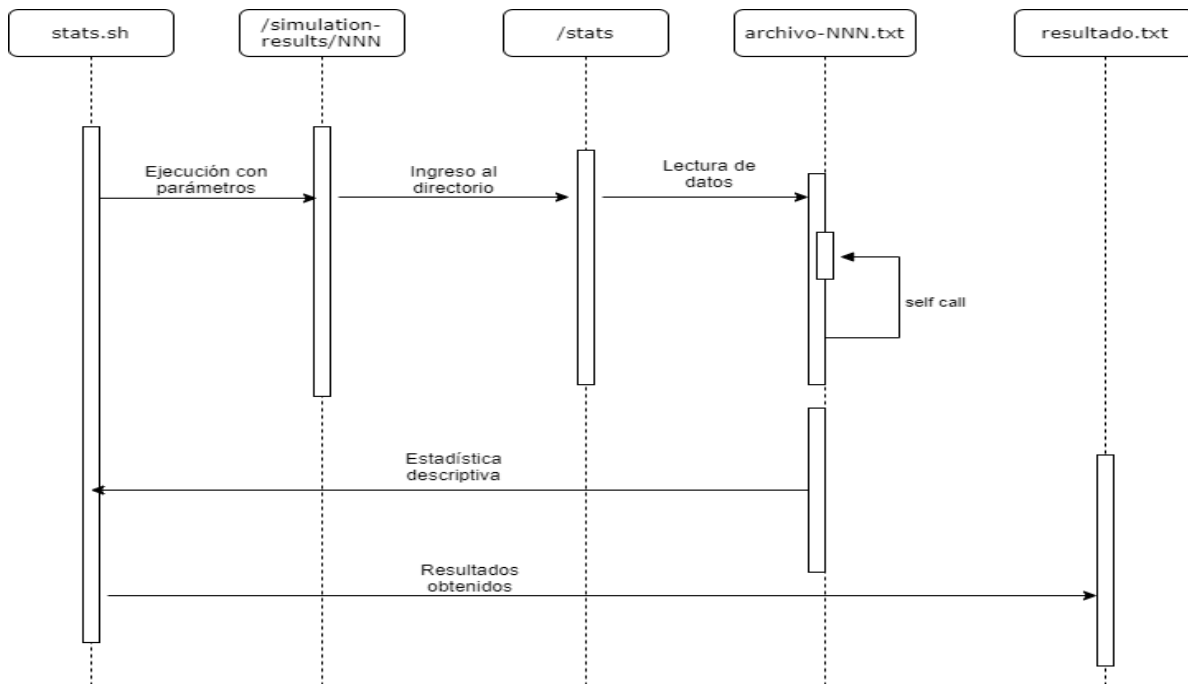


Fig. 2 Diagrama de secuencia de ejecución del script

Así mismo de forma específica cada tarea solicita datos particulares para analizar y obtener resultados, las tareas solicitan la media, valor mínimo y máximo para los siguientes campos:

- Tarea 1: Tiempo de simulación total, memoria utilizada por el simulador.
- Tarea 2: Todas las personas simuladas, solo residentes, solo visitantes tipo I, solo residentes separados por grupo etario, solo visitante tipo I separados por grupo etario.
- Tarea 3: Uso del teléfono en el instante de tiempo específico.

Las figuras siguientes muestran cómo se aplican 3 componentes en el script para satisfacer las solicitudes de las tareas.

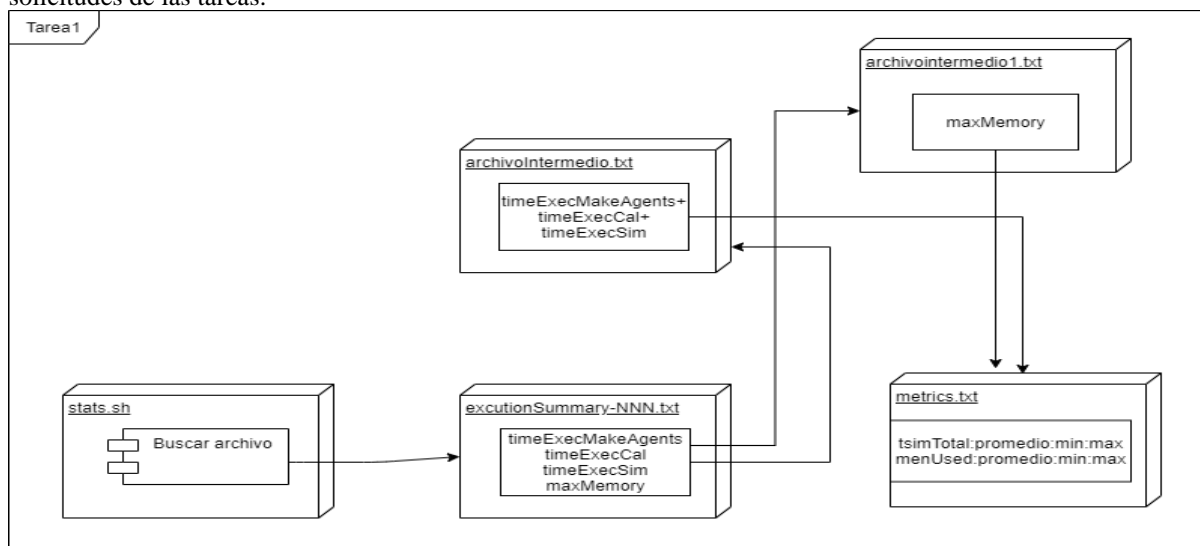


Fig. 3 Diagrama de bloques para tarea 1

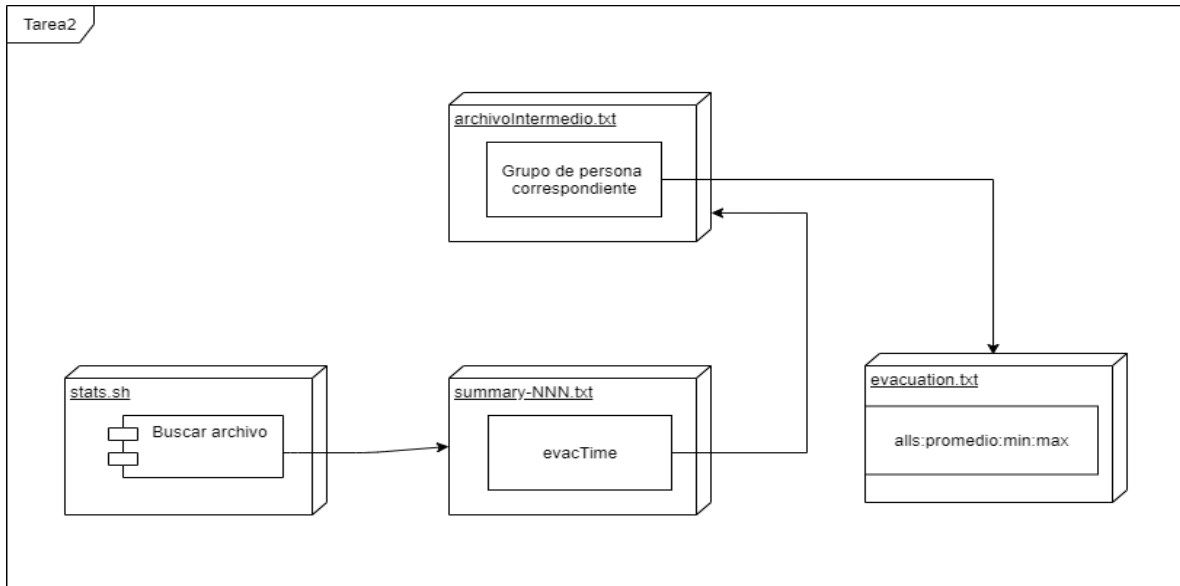


Fig. 4 Diagrama de bloques para tarea 2

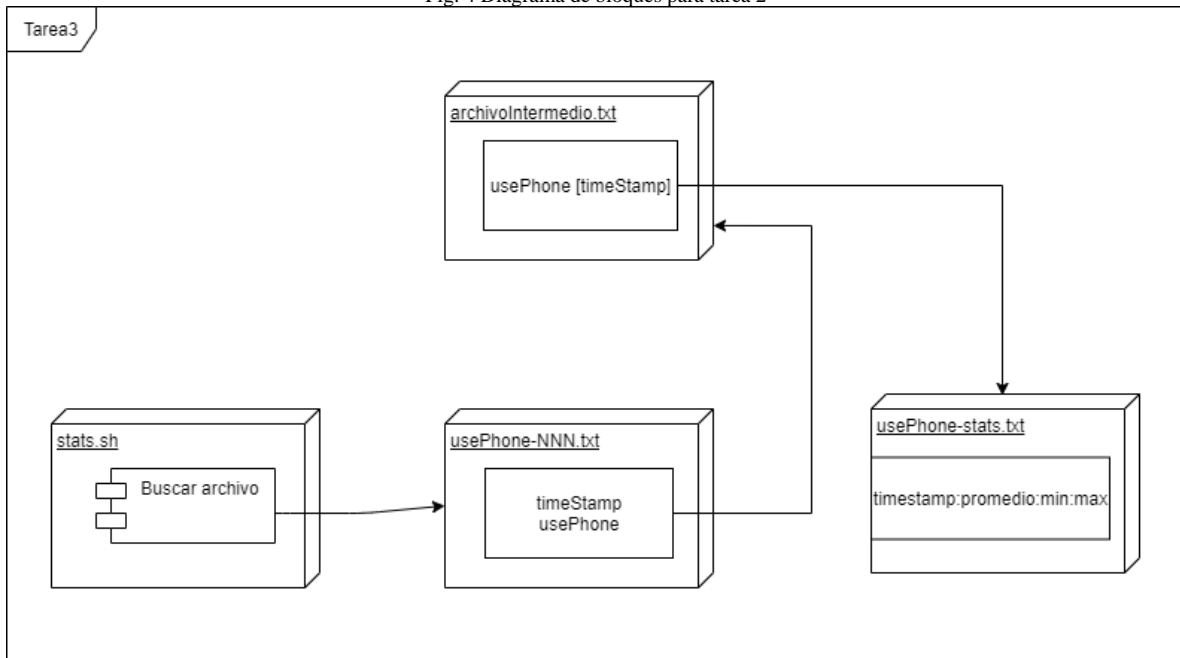


Fig. 5 Diagrama de bloques para tarea 3

En la parte final del script se eliminarán todos los archivos intermedios.

3 Resultados

A partir del diseño previo se obtuvieron los resultados a grandes rasgos, tal como lo muestra la figura 6, para poder obtener los datos específicos de cada campo en los datos de la simulación se utilizó el comando *cat* de Bash para poder recorrer archivo por archivo en el directorio, así como la función *awk* que permite encapsular los campos de manera individual. Con esto los resultados específicos fueron la estadística descriptiva en cada archivo fuente, con esto podemos ver que tanto variaron los valores a partir de la media de cada campo lo que

nos muestra que la variación se comportan como la campana de Gauss [3], por tanto el comportamiento de la simulación demuestra un grado de eficiencia ya que se cumple en su mayoría el objetivo de la evacuación.

```
sebastian@tallerso:~/TSSOO-taller01/Taller_1$ ls
evacuation.txt  metrics.txt  README.md  simulation-results  stats.sh  usePhone-stats.txt
```

Fig. 6 Muestra los 3 archivos que contienen la estadística descriptiva.

De manera específica se puede ver los resultados de que cada operación en la figura 7.

```
tsimTotal:promedio:min:max
492511:246256:244209:248302
memUsed:promedio:min:max
579264:289632:288236:291028
sebastian@tallerso:~/TSSOO-taller01/Taller_1$ more evacuation.txt
resident-G0:promedio:min:max
resident-G1:promedio:min:max
resident-G2:promedio:min:max
resident-G3:promedio:min:max
visitors-G0:promedio:min:max
visitors-G1:promedio:min:max
visitors-G2:promedio:min:max
visitors-G3:promedio:min:max
2530:632.5:0:1300
2530:632.5:0:1300
7180:897.5:0:3620
7180:897.5:0:3620
14590:1215.83:0:3840
14590:1215.83:0:3840
14590:911.875:0:3840
14590:911.875:0:3840
alls:promedio:min:max
14590:7295:7280:7310
residents:promedio:min:max
12470:6235:6170:6300
visitorsI:promedio:min:max
2120:1060:980:1140
sebastian@tallerso:~/TSSOO-taller01/Taller_1$ more usePhone-stats.txt
timestamp:promedio:min:max
1:4404:4475:4333
2:4425.5:4468:4383
3:4352.5:4444:4261
```

Fig7. Muestra cada resultado de cada tarea.

Es importante mencionar que para la generación de estos resultados se usó una fuente de datos de menor tamaño para evitar tiempos de espera prolongados.

4 Conclusiones

En este informe podemos concluir que el diseño como etapa anterior o paralelo a la implementación es una buena manera de guiar el desarrollo de la solución para lo que se está solicitando, permitiendo seguir un orden a la hora de crear código y mantenerlo hasta el final. Además se presentó una forma de calcular las estadísticas

descriptiva con Bash de una lista de fuentes de datos, entregando los resultados tras los cálculos que corresponden, lo cual fue eficiente en términos de procesamiento debido a la poca abstracción del lenguaje, permitiendo un manejo directo de las variables y datos para poder procesarlos.

5. Referencias

[1] Servicio Hidrográfico y oceanográfico de la Armada Available:
<http://www.snamchile.cl/index.php?p=gemTsunami&enl=4>

[2] Armando Cisternas El país más sísmico del mundo Séptima Serie, N°1, pp.8

[3] Wikipedia, La enciclopedia libre Available:
https://es.wikipedia.org/wiki/Distribuci%C3%B3n_normal