

**SMART HOME CONTROL BY VOICE USING NEURAL NETWORKS****HOGAR INTELIGENTE POR CONTROL DE VOZ USANDO REDES  
NEURONALES**

**MSc. Esperanza Camargo, Ing. Cesar Coronel Segrera  
Ing. Milton Calderón**

**Universidad Distrital Francisco José de Caldas, Facultad Tecnológica.**

Grupo de Investigación DIGITI.

Trasversal 70B No.73 a 35 Sur, Bogotá, Cundinamarca, Colombia.

Tel: +(57-1) 3239300 Ext 5003 Fax: +(57-1) 3239300.

E-mail: [ecamargoc@udistrital.edu.co](mailto:ecamargoc@udistrital.edu.co)

**Abstract:** This paper presents the design of an intelligent low-cost home-based Google Voice API chained to a number of neural network topology feedforward-back propagation, which provide interaction with the machine in a fun mimicking human conversation and leaving monotonous side orders, the state of art technologies presents all options chosen under open source platform to minimize costs and make the most globalized and knowledge accessible to the reader, also find the evolution of the system itself, the results obtained is the total automation of a apartment of 57 square meters with interactive functions based on voice and you can turn on or off any light housing complex by a vocal request addressed to the device with a fast response of the system, plus a number of useful for everyday life choices.

**Keywords:** Domotics, open source, android, artificial intelligence, embeded systems.

**Resumen:** Este artículo presenta el diseño de un hogar inteligente de bajo costo basado en el *Google Voice* API encadenada a una serie de redes neuronales de topología *feedforward-backpropagation*, las cuales brindan la interacción con la máquina de una manera amena simulando una conversación humana y dejando de lado órdenes monótonas, el estado del arte presenta las opciones tecnologías escogidas todas bajo plataforma *Open Source* para minimizar costos y hacer el conocimiento más globalizado y accesible al lector, también encontrarán la evolución del sistema en sí, los resultados obtenidos es la automatización total de un apartamento de 57 metros cuadrados con funciones interactivas basadas en la voz ya que se puede encender o apagar cualquier luz del complejo habitacional haciendo una solicitud vocal direccionada al dispositivo con una respuesta instantánea del sistema, además de una serie de opciones útiles para la vida cotidiana.

**Palabras clave:** Domótica, open source, android, inteligencia artificial, sistemas embebidos.

## 1. INTRODUCCIÓN

La domótica consiste en dotar de inteligencia a una casa para que se ocupe por si misma de ciertas

tareas, estas tareas pueden ser tan simples como controlar la temperatura de un recinto, o encender una luz, o tan complejas como administrar el consumo energético de la casa.

La estructura básica de un sistema domótico se compone de tres elementos principales, los cuales son los sensores, los actuadores y el controlador (Zamora, 2010).

Los sistemas domóticos se han venido desarrollando desde la década de los noventa en donde los primeros prototipos funcionaban mediante un bus de datos en el cual viajaba la información necesaria para el funcionamiento e interconexión de los dispositivos que hacen posible la administración de variables en el hogar, posteriormente se unificó el flujo de energía con la transmisión de datos a cada uno de los dispositivos. Con el auge de la tecnología inalámbrica ésta comenzó a hacer parte de los sistemas domóticos. Sin embargo, esta nueva tecnología tuvo varios inconvenientes ya que necesitaban una conexión en línea directa, para solucionar el problema, se desarrolló una estructura en malla, con la cual se solucionaron los problemas incorporados por la tecnología inalámbrica (Bromley, Perry y Webb, 2003).

Los motores de voz han tenido un gran auge en los dispositivos móviles en los cuales hacen que la interacción usuario – dispositivo se de una forma más natural, el primer fabricante que incorporó esta tecnología en sus dispositivos fue Apple dicho sistema de interacción se llamó Siri, posteriormente el sistema también fue implementado por Google para su sistema operativo de dispositivos móviles llamado Android.

## 2. ESTADO DEL ARTE

Hoy en día en el mercado existen infinidad de opciones Domóticas para su hogar, las cuales oscilan entre los 200\$ USD y 700\$ USD, pero con un poco de destreza y innovación puedes fabricar un sistema hecho a tu medida y necesidades. Para esto se usará tecnologías Open Source con el fin de minimizar la mayoría de gastos económicos. Un motor de voz es un sistema capaz de reconocer la voz humana y transcribirla a texto para ser utilizada en otras aplicaciones, una aplicación de este estilo es la que desarrollará a través de este artículo. El motor de voz mas potente y abierto al público es el suministrado por Google, este motor de voz funciona bajo sistema operativo Android el cual puedes encontrar en cualquier celular de gama media en nuestro hogar, por este motivo además de otras innumerables como información, tutoriales y facilidad para programar optaremos por este sistema operativo para el desarrollo de la

aplicación para nuestro entorno Domótico. Uno de los dispositivos más usados a nivel electrónico son las placas de desarrollo tales como Arduino, RaspberryPi, Intel Galileo, Cerebot, Chipkit Max32. En este caso se usará el sistema de desarrollo Arduino Mega ADK, el cual posee un Host USB Android, este Host nos brinda comunicación entre la placa Arduino y la Aplicación la cual se realizara en Android uno de los sistemas operativos para móvil más usados en el mundo, de esta manera tu celular será el cerebro. El Arduino Mega ADK es una placa electrónica basada en el microprocesador Atmega2560. Cuenta con una interfaz de host USB para conectar con los teléfonos basados en Android, basado en el maxZ3421e IC. Cuenta con 54 pines digitales de entrada / salida (de los cuales 15 se pueden utilizar como salidas PWM), 16 entradas analógicas, 4 UARTs (puertas seriales), un oscilador de 16MHz, una conexión USB, un conector de alimentación, un header ICSP, y un botón de reinicio (Arduino, 2013).

La estructura de malla utilizada en este proyecto. La topología de red malla es una topología de red en la que cada nodo está conectado a todos los nodos. De esta manera es posible llevar los mensajes de un nodo a otro por distintos caminos. Si la red de malla está completamente conectada, no puede existir absolutamente ninguna interrupción en las comunicaciones. Cada servidor tiene sus propias conexiones con todos los demás servidores. Esta topología, a diferencia de otras (como la topología de árbol y la topología de estrella), no requiere de un servidor o nodo central, con lo que se reduce el mantenimiento (un error en un nodo, sea importante o no, no implica la caída de toda la red).

Las redes de malla son auto ruteables. La red puede funcionar, incluso cuando un nodo desaparece o la conexión falla, ya que el resto de los nodos evitan el paso por ese punto. En consecuencia, la red malla, se transforma en una red muy confiable. Es una opción aplicable a las redes inalámbricas (*wireless*), a las redes cableadas (*wired*) y a la interacción del software de los nodos.

Una red con topología en malla ofrece una redundancia y fiabilidad superiores. Aunque la facilidad de solución de problemas y el aumento de la confiabilidad son ventajas muy interesantes, estas redes resultan caras de instalar, ya que utilizan mucho cableado. Por ello cobran mayor importancia en el uso de redes inalámbricas (por la no necesidad de cableado) a pesar de los inconvenientes propios de las redes sin hilos.

En muchas ocasiones, la topología en malla se utiliza junto con otras topologías para formar una topología híbrida (Castells, 1997).

### 3. METODOLOGÍA

El primer paso requerido es familiarizarse con el API de Google con el cual se realizará el reconocimiento de voz, esta API es muy sencilla de utilizar ya que consta de dos comandos básicos los cuales son “escuchar” y “detenerEscucha”. en el primer comando la API habilita el dispositivo de audio y comienza a capturar información por medio del micrófono, al detener la escucha interrumpe la captura de audio y procede a realizar la transducción de voz a texto. En la figura 1 se puede observar la utilización de la API para decodificar voz (Android, 2014).

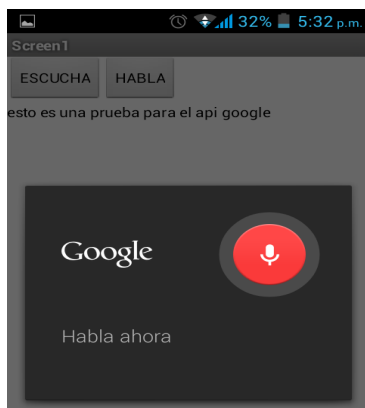


Fig. 1. Interfaz de audio, Captura de pantalla DIGITIDOMOTIC.

El siguiente paso requerido es identificar la acción adecuada al comando de voz determinado mediante el motor de voz, dicha identificación se realizará mediante una red neuronal con topología feedforward-backpropagation (Augusto y Nugent, 2006), la cual se puede observar en la figura 2, en ésta se ejecutará el proceso de decodificación de la orden a ejecutar, el sistema tiene la flexibilidad de detectar la misma orden a ejecutar aun cuando el comando de voz sea diferente, esto se realiza mediante la individualización de las palabras ingresadas en el comando de voz, las cuales serán analizadas una a una y en dicho análisis se descartarán palabras no útiles que hayan sido descartadas en la red neuronal. una vez descartadas las palabras se procede al análisis posicional de cada una de ellas, una vez evaluadas estas variables la red determina el comando a ejecutar la cual será la variable de salida, tal como encender o apagar

una luz, subir o bajar un telón de proyección, abrir o cerrar las cortinas de una habitación.

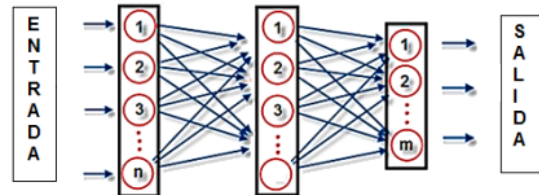


Fig. 2. Topología red neuronal, DIGITIDOMOTIC.

El proceso de individualización de palabras se realiza de la siguiente manera: luego de adquirido el comando de voz, la API de Google se encarga de convertir la voz en texto, este texto es almacenado en una variable conocida como *SpeechVar*, en el caso de la figura 1 el *SpeechVar* correspondiente es: “esto es una prueba para el API GOOGLE”. El procesamiento de esta variable se realiza mediante los espacios entre la oración, en este caso se usa un Array de palabras.

1	esto
2	es
3	una
4	prueba
5	para
6	el
7	api
8	google

Fig. 3. Array de palabras, DIGITIDOMOTIC.

En este Array almacena cada palabra según en el orden que fue pronunciada para el posterior análisis. La entrada a cada una de las neuronas mostradas en la Figura 2 son los componentes de la Array y sus combinaciones. Se realizará la descripción del proceso para el caso de encender una luz: Antes de programar las neuronas se toman todas las palabras que se pueden relacionar a encender una luz, tales como “encender, encendido, enciende, enciéndeme, activa, activado, actívame, active, prender, prendido, prendida.” Estas palabras se asignan a una neurona en la capa 2, la cual esta encargada de identificar cuando alguna de estas palabras se presentan dentro del Array, al identificar alguna la neurona en la capa 2 se encarga de modificar el estado en la capa 3, a lo cual se llama una Variable de estado. Esta variable de estado es la encargada de asignar un estado digital a uno de los elementos del recinto residencial.

Este mismo procedimiento se realiza para las neuronas de identificación de los lugares del hogar tal como: para la habitación se asignará una neurona que identifique las palabras: “Cuarto, Habitación, Pieza, Alcobá, Principal.” Al realizar el mismo proceso anterior la salida en la capa 3 de la neurona será asignada a la Variable Lugar.

Teniendo la Variable Lugar, y La Variable estado se pueden dar órdenes vocales al sistema de forma fluida con esto la red neuronal tomara las decisiones, se puede decir: enciende la luz de la habitación, cuyo resultado al finalizar el proceso será: “Variable Estado =1, y la Variable Lugar = Habitación”.

O se puede decir: “por favor serias tan amable de prender la luz del cuarto”, cuyo resultado al final de la red neuronal será: “Variable Estado =1 Variable Lugar = Habitación”.

Para realizar el sistema más ameno y amable con el usuario se han configurado una serie de respuestas aleatorias a cada una de las instrucciones tales como: (listo, luz de “Variable lugar” “variable estado”, con mucho gusto mi labor es servirte, como ordenes encenderé la luz de la “variable lugar”, ok ya está “variable estado”), además de esto se configuraron sugerencias en el sistema, después de cada orden el sistema sugiere encender o apagar el sistema de audio o subir o bajar las cortinas. Cada una de las variables lumínicas posee un contador de tiempo individual, el cual se activa o se desactiva cuando encendemos o apagamos una luz, conociendo previamente el consumo en kw/H de cada una de las zonas de iluminación del hogar se puede establecer un estimado de consumo a través del tiempo referente a iluminación. Con este consumo fijamos alertas de alto consumo o advertencias antes de encender una luz o visualizar el consumo del mes en cuestión de luminarias.

Otra de las características que posee el sistema es el encendido o apagado de todas las luces del apartamento por medio de los comando de voz “Encender todas las luces” o “Apagar todas las luces” estos comandos pasan por el mismo proceso de redes neuronales explicado anteriormente.

El sistema de Alarma usa los métodos convencionales de detección de intrusos tales como sensores de movimientos, interruptores relay en puertas y ventanas, con la diferencia que mediante el sistema se puede configurar un mensaje de texto, Twitter, Facebook o WhatsApp con el dispositivo

Android para la notificación de violación de seguridad en el apartamento.

En la figura 4 se puede observar la interfaz final realizada en Android para un apartamento habitacional de tamaño regular, este tipo de implementaciones pueden tener un valor que oscila de los USD \$200 hasta los USD \$700 incluyendo un dispositivo móvil de gama media, y el valor de la vivienda puede verse incrementado entre USD \$4000 y USD \$6000 debido a lo futurista e innovador que luce el apartamento con esta implementación.



Fig. 4. Interfaz final del sistema domótico, Captura de pantalla aplicación DIGITIDOMOTIC.

Para la implementación de este sistema se tuvieron en cuenta como sistemas embebidos el Raspberry Pi (Raspberry Pi, 2013) y el Arduino Mega (Arduino, 2013), los cuales cuentan con la característica que son Open Source esto hace que toda la documentación de dichos sistemas sea accesible y además reduce costos.

Tabla 1. Funcionalidades, DIGITIDOMOTIC..

LUGARES	TIPO DE FUNCIONALIDAD	ESTADOS	ACTUADOR	CONTROL POR VOZ
SALA	LUCES	ON/OFF	ESTADO SOLIDO	SI
COCINA	LUCES	ON/OFF	ESTADO SOLIDO	SI
BAÑO	LUCES	ON/OFF	ESTADO SOLIDO	SI
PASILLO	LUCES	ON/OFF	ESTADO SOLIDO	SI
COMEDOR	LUCES	ON/OFF	ESTADO SOLIDO	SI
ALCOBA	LUCES	ON/OFF	ESTADO SOLIDO	SI
LABORATORIO	LUCES	ON/OFF	ESTADO SOLIDO	SI
TEATRO	ENTRETENIMIENTO	ON/OFF	MOTORIZADO	SI
CORTINAS	VENTILACION	ARRIBA/ABAJO	MOTORIZADO	SI

#### 4. RESULTADOS

Se obtuvo la automatización total de la iluminación de un apartamento regular implementando comandos de voz y órdenes táctiles a través de un software desarrollado en Android, este sistema además monitorea el consumo energético y realiza

sugerencias basándose en las condiciones lumínicas previamente configuradas.

Se facilita la supervisión y control del hogar desde un lugar remoto ya que las variables pueden ser observadas y modificadas por internet en tiempo real, además de esto el sistema puede enviarte alertas o mensajes a través de Facebook, Twitter, WhatsApp o mensaje de texto, como apertura de puertas o activación de alarmas.

Se puede obtener un consolidado del consumo por luminarias al día del mes de la consulta, con lo cual se puede estimar el gasto total energético asignado a iluminación, adicionalmente a esto el sistema también puede generar alertas por exceso del consumo, dicho consumo es configurado al inicio del periodo mensual.

Entrando un poco en detalle en el comportamiento del modelo realizado para el apartamento podemos describir que desde tu dispositivo móvil con una orden de voz no establecida siempre y cuando lleve la variable "lugar" como (sala, cocina, alcoba, laboratorio, pasillo) o alguno de sus sinónimos como (habitación, pieza, cuarto), y una variable de "estado" como: (encendido, prendido, activado, apagado, desactivado) se puede activar o desactivar cualquier luz del complejo habitacional, Además en la aplicación de Android se desarrollo una serie de botones con los que se logran los mismos resultados. Estas variables pueden ser monitoreadas o cambiadas desde la red de internet. Esta opción es muy práctica en caso tal que abandonemos nuestros hogares en periodos largos de vacaciones y no queremos levantar sospechas que nuestro hogar se encuentra deshabitado, con esta opción podemos realizar encendido y apagado de luces.

## 5. CONCLUSIONES

Se implementó un control de voz mediante el cual el usuario interactúa con el sistema domótico, sin previas instrucciones con un lenguaje muy interactivo.

Se desarrolló un sistema neural capaz de interpretar y asimilar los comandos de voz generados por el usuario.

Se obtuvo un sistema domótico confiable en el cual la interfaz humano-maquina funciona de una manera interactiva similar a una conversación humana.

## REFERENCIAS

- Zamora, M. Santa, J. y Gómez, A. (2010) "An Integral and networked home automation solution for indoor ambient intelligence". IEEE. Pervasive Computing, Vol. 9.
- Bromley, K. Perry, M. y Webb, G. (2003) "Trends in Smart Home Systems, Connectivity and Services". [www.nextwave.org.uk](http://www.nextwave.org.uk)
- Android. (2013) "Reference Android Speech Package"  
<http://developer.android.com/reference/android/speech/package-summary.html>
- Google (2013) "Google Applications Program Interface Developers Experiences"  
<http://commondatastorage.googleapis.com/io-2013/presentations/4057%20Web%20Speech%20API%20creates%20Interactive%20Experiences%20-%20Google%20I-O%202013%20Codelab.pdf>
- Augusto, J. y Nugent, C. (2006) "Designing Smarthomes: The role of artificial intelligence" Springer Berlin, Vol 1.
- Raspberry Pi Foundation. (2013). "Resources. 01/06/2014, de Raspberry Pi Foundation" <http://www.raspberrypi.org/resources/>
- Arduino. (2013). "Arduino Mega. 01/06/2014, de Arduino" <http://arduino.cc/en/pmwiki.php?n=Main/ArduinoBoardMega>.
- Castells, M. (1997). "La era de la información. Economía, sociedad y cultura" Alianza Editorial. Madrid, Vol. 1.