

Novática, revista fundada en 1975 y decana de la prensa informática española, es el órgano oficial de expresión y formación continua de **ATI** (Asociación de expresion y formación continua de ATT (Asociación de Técnicos de Informática), organización que edita también la revista **REICIS** (Revista Española de Innovación, Calidad e Ingeniería del Software). **Novática** edita asimismo **UP**GRADE, revista digital de **CEPIS** (Council of European Professional Informatics Societies), en lengua inglesa, y es miembro fundador de **UP**ENET (**UP**GRADE *European NETwork*).

> <http://www.ati.es/novatica/>
>

ATI es miembro fundador de CEPIS (Council of European Professional Informatics Societies) y es representante de España en IIFIP (International Federation for Information Processing); tiene un acuerdo de colaboración con ACM (Association for Computing Machinery), como acuerdos de vinculación o colaboración con AdaSpain, AI2, ASTIC, RITSI e Hispalinux, junto a la que participa en Profinova.

Cansolo Editorial Antoni Carbonell Nogueras, Juan ManuelCueva Lovelle, Juan Antonio Esteban Iratar, Erancisco López Crespo, Celestino Martin Atonso, Josep Molas i Bertrán, Olga Pallás Codina, Fernando Plera Comez (Presidente del Consejo), Ramón Pujglaner Trepal, Miguel Sarines Grinó, Asunción Yfutre Herran.

Coordinación Editorial Llorenc Pagés Casas < pages@ati.es > Compésición y autodición Jorge Lider Gil de Ramales Traduccions

Iraulocciones
Grupo de Lengua e Informática de ATI http://www.ati.es/gt/lengua-informatica/, Dpto.
de Sistemas Informáticos - Escuela Superior Politécnica - Universidad Europea de Madrid
Administración

Administración Tomás Brunete, María José Fernández, Enric Camarero, Felicidad López

Secciones Técnicas - Coordinadores Acceso y recuperación de la información

Gumersindo Barcia Ariribas (MAP), - Gumersindo, garcia@map, es>
Arquitactizas
Enrique F. Torres Moreno (Universidad de Zaragoza), - enrique.torres@unizar.es>
Jordi Tubella Morgadas (DAC-UPC), - (jordi@ae.upc.es>
Auritania guitio | folitio, - marinatourino@marinatourino.com>
Marina guitio | folitio, - marinatourino@marinatourino.com>
Marina guitio | folitio, - marinatourino.com>
Marina Palaro Barcia-Suello (ASIA), - manuel@palao.com>
Marina Palaro Barcia-Suello (Collazos (Fac. Derecho de Donostia, UPV), - (hernando@legaltek.net>
Elena Davara Fernández de Marcos (Davara & Davara), - edavara@davara.com>
Esseba Eva Guniversitaria de la hisformatica
Joaquin Expeleta Mateo (CPS-UZAR), - ezpeleta@posta.unizar.es>
Cristobal Parge Flores (DSP-UOM), - coprafigacióp.ucm.es>
Alonso Alvarco Carcia (TID), - «aag@ild.es>
Diego Gacher Page (Universitada truorea de Madrid), - gachet@uem.es>
Basal Baiga Sole (Cgo Deminic Imst & Wung), - (pan.haige@ail.es>
Informatica y Filosochia.

angel Olivas Varela (Escuela Superior de Informática, UCLM)

satisfica profite a company de la compa

niversitat Jaume I de Castellón), <chover@lsi.uji.es> o (Eurographics, sección española), <rvivo@dsic.upv.es>

Roberto Vivo Herrando (Eurographics, sección española), < rivivo@ Ingunterá ela 34/mento (Eurographics, sección española), < rivivo@ Ingunterá ela 34/mento (Eurographics, sección española), < rivivo@ Javier Dolado Cosin (D.SI-UPV), < dolado@si.ehu.es> Luis Fernandez Sanz (PRIS-EI-UEM), < luitem@dpris.esi.uem.es> Mento Ment

Carmen Ugarte García (IBM), <cugarte@ati.es>
tajes informátices
Marín López (Ilnív Cort- ...

Andrés Marin López (Univ. Carlos III), <amarin@it.uc3m.es>
J. Angel Velázquez Itúrbide (ESCET-URJC), <a.velazquez@escet.urjc.es>
Xayer Gámar Guiconad (III-

Xavier Gomez Guinovart (Univ. de Vigo), <a velazquez@escet.uric.es>
Xavier Gomez Guinovart (Univ. de Vigo), <a velazquez@escet.uric.es>
Manuel Palomar (Univ. de Alicante), <mpalomarc@dist.ua.es>
Manuel Palomar (Univ. de Alicante), <mpalomarc@dist.ua.es>
Manuel Palomar (Univ. de Alicante), <mpalomarc@dist.ua.es>
Pederico G. Mon (Trotti, (RITSI) <mpalomarc@dist.ua.es>
Protaclée. Informatica.
Rataler Fernander Calvor (ATV. us.) Processes Informática
Rafael Fernández Catvo (ATI), <firalvo@atl.es>
Miguel Sarries (6finó (Ayto, de Barcelona), <marriera (marriera) (atl.es>
Rafael Fernández (atl.es) (atl.es) (atl.es>
Rafael y sarriclas felenáldez
Jose Lus Marzo Lazaro (Univ. de Girona), <joseluis.marzo@udg.es>
Jose Solé Pareta (DAC-UPC), <pareta@ac.upc.es>
Jose Solé Pareta (DAC-UPC), <pareta@ac.upc.es>

Josep Sole Pareta (DAC-UPC), cpareta@ac.upc.es>
Seguriad
Josep Sole Pareta (DAC-UPC), cpareta@ac.upc.es>
Josep Sole Pareta (Univ. do Deusto), cparetio@eside.deusto.es>
Josep Sole Pareta (Univ. do Deusto), cparetio@eside.deusto.es>
Josep Sole Pareta (Univ. do Deusto), cpareta (Univ. do Deusto), centro (

Didac Lopez Brilli (Inviersitat de Girona). «didac lopez@ati.es> Francisco Javier Cantais Sánchez (Indra Sistemas). «Jicantiais@gmail.com> Tdc y Turtsame Andrés Aguayo Maldonado, Antonio Guevara Plaza (Univ. de Málaga) «(aguayo, guevara)@ilce.uma.es>

Las opiniones expresadas por los autores son responsabilidad exclusiva de losmismos. *Noválica* permite la reproducción, sirá animo de lucró, de artículos, a menos que lo impida la modalidad de © o copyraft, elegida por el autor, debiendose en todo caso citar su procedencia y enviar a *Noválica* un elemplar de la publicación.

Coordinación Editorial, Redacción Central y Redacción ATI Madrid

dilla 66, 3°, dcha., 28006 Madrid n.914029391; fax.913093685 < novatica@ati.es> mpasición, Edición y Redacción ATI Valenci del Reian de Valencia 23, 46005, Valencia

inn/rax 963330392 <serreval@ati.es>
dmilattración y Redacción ATI Cataluña
la Laletana 46, ppal. 1ª, 08018 Barcelona
lin.93412525; tax 934127713 <secregen@ati.es>
vdacción ATI Andalucia
ago Nowton Gr. E.C. C. "

saac Newton, s/n, Ed. Sadiel, sla Cartija 41092 Sevilla, Tilin/lax 954460779 <secreand@atl.es> lotacciba ATI Aragón agasca 9.3-8, 50006 Zazanos

n./lax 970233161 < secteala@ati.es > q**daccién Ati Asturias-Cantabria** <gp-astucant@ati.es > **daccién Ati Castilia-La Mancha** <gp-cimancha@ati.es > **recrincién y Vantas**

y ventas ati.es/novatica/interes.html>, o en ATI Cataluña o ATI Madrid

Padilla 66, 3°, dcha., 28006 Madrid Tlnf.914029391; fax.913093685 < novatica@ati.es>

Derra S.A., Juan de Austria 66, 08005 Barcelona.

Dapásito legat: 315,154-1975 — ISSN: 0211-2124; CODEN NOVAEC

Partada: "Ala Turing & Friends" (variaciones sobre una foto tomada de www.luning.org).

RFCalvo / © Rabael Fernandez Calvo 2001

Bisedia: Fernando Agresta / © ATI 2006

Nº 184, noviembre-diciembre 2006, año XXXII

sumario

aditorial			
editorial La enseñanza de la Informática en España	>	>	02
Un estándar, dos estándares Llorenç Pagés Casas	>	>	02
noticias IFIP Actividades del IFIP TC6 Technical Committee on Communication Network Ramón Puigjaner Trepat	(S >	>	03
monografía Formato de Documento Abierto (ODF) (En colaboración con UPGRADE)			
Editores invitados: Jesús Tramullas Saz, Piedad Garrido Picazo, Marco Fioretti Presentación: OpenDocument, estándar para documentos digitales Jesús Tramullas Saz, Piedad Garrido Picazo	>	>	04
Abierto desde el diseño: el Formato de Documento Abierto para aplicaciones ofimáticas			06
Erwin Tenhumberg, Donald Harbison, Rob Weir ¿Es OpenDocument un estándar abierto?: iSí!			13
David A. Wheeler			
Trampas ocultas en OpenDocument y efectos secundarios en el software libre y de código abierto Marco Fioretti	>	>	19
ISO-26300 (OpenDocument) vs. MS-Office Open XML	>	>	22
Alberto Barrionuevo García Interoperabilidad: ¿se impondrá el verdadero formato universal de fichero Sam Hiser, Gary Edwards	os? >	> 1	28
ODF: el Formato de Documento emergente a elección de los gobiernos Marino Marcich	>	> ;	36
Promoción del uso de los formatos abiertos de documentos por los Progra			
IDA e IDABC Miguel A. Amutio Gómez	>	> ;	39
Una historia resumida de los estándares abiertos en Dinamarca John Gøtze	>	>	42
Formatos estándares abiertos y software libre en la Administración Pública de Extremadura Luis Millán Vázquez de Miguel	>	> 1	44
secciones técnicas			
Enseñanza Universitaria de la Informática Acciones y reacciones en el camino de la mejora docente universitaria Alfonso Blesa Gascón, Pablo Bueso Franc, Carlos Catalán Cantero, Raquel Lacuesta Gilaberte, Mariano Ubé Sanjuán	>	> 1	46
Informática Gráfica Programación de Aplicaciones Gráficas con OpenGL y Java			51
Óscar Belmonte Fernández			
Redes y servicios telemáticos Algoritmo bioinspirado para la optimización de rutas en Internet José Luis Gahete Díaz, Fernando Gómez González	>	> :	56
Referencias autorizadas	>	>	63
sociedad de la información			
Futuros emprendedores Step by Step: Mens sana in corpore sano Miguel Angel Ramos Barroso, Javier Cantón Ferrero, Javier Fernández Rodríguez, Juan María Laó Ramos	>	>	70
Novática interactiva Competencia entre estándares, ¿va a ser posible su coexistencia? Foro de Debate	>	>	74
Programar es crear Polígonos en malla (CUPCAM 2006, problema A, enunciado) Dolores Lodares González	>	>	75
asuntos interiores			
Coordinación editorial / Fe de erratas / Programación de Novática Normas para autores / Socios Institucionales			76 77

Monografía del próximo número: "Buscadores en la Web"

José Luis Gahete Díaz, Fernando Gómez González Escuela Técnica Superior de Ingeniería (ICAI) Universidad P. Comillas, Madrid

Algoritmo bioinspirado para la optimización de rutas en Internet

{jlgahete, fgomez}@dsi.icai.upcomillas.es

1. Introducción

Internet es una red mundial formada por millones de ordenadores de diversos tipos y plataformas, conectados entre sí por medio de equipos de comunicación cuya función principal es la de localizar, seleccionar, e intercambiar información desde el lugar en donde se encuentra hasta aquel donde hava sido solicitada, utilizando un mismo conjunto de protocolos de comunicaciones llamado TCP/IP [23].

A pesar de su complejidad, Internet se puede describir de manera jerárquica desde un punto de vista macroscópico, en el que se puede ver como un sistema formado por más de 16.000 Sistemas Autónomos (en adelante AS) [14] (conjunto de redes IP interconectadas por routers y que conforman una entidad administrativa) que interactúan para coordinar la entrega del tráfico IP.

Este conjunto de AS's da lugar a una tupida malla de interconexiones originando un problema de encaminamiento en el puede haber hasta 65.000 AS's de origen y destino por los que circula constantemente información. Como acabamos de mencionar, uno de los problemas más importantes de este esquema es la correcta localización y entrega de la información entre origen y destino; en otras palabras: el problema del encaminamiento.

El protocolo interdominio Border Gateway Protocol (en adelante BGP) [20] se desarrolló explícitamente para ser utilizado en redes TCP/IP, convirtiéndose en el protocolo de encaminamiento exterior (o interdominio) estándar en Internet. El objetivo del BGP es garantizar la alcanzabilidad de los paquetes IP entre origen y destino, sin tener en cuenta la optimización de rutas. En consecuencia, no es tarea de BGP la búsqueda de una ruta óptima global. Esto se debe, por una parte, a las restricciones impuestas en las políticas de encaminamiento elegidas por el administrador y, por otra parte, al carácter de protocolo de vector distancia de BGP, llamado vector-camino (en contraposición con protocolos de estado de enlace) [3].

En la figura 1 se muestra el esquema de interconexiones en Internet. Los círculos grises representan AS's, es decir, unidades administrativas. Como puede verse en la figura, los AS's intercambian información de encaminamiento mediante BGP. AsimisResumen: en este trabajo presentamos un novedoso algoritmo genético al que hemos denominado Algoritmo Genético del Ciervo. Este algoritmo está inspirado en el comportamiento social y reproductor de los ciervos, siendo particularmente idóneo para la optimización de rutas en el problema del encaminamiento en Internet. Asimismo, mostramos las características más importantes del algoritmo en términos de convergencia. Finalmente exponemos su aplicabilidad en el contexto de la supervisión del tráfico entre Sistemas Autónomos en Internet.

Palabras clave: algoritmos genéticos, encaminamiento. Internet, optimización, sistemas autónomos.

mo, cada AS está formado por diferentes subredes conectadas entre routers que utilizarán el protocolo intradominio elegido por el administrador.

Como ya hemos mencionado, el protocolo BGP permite a los administradores de los AS's configurar manualmente las políticas de encaminamiento en los routers borde (ver figura 1). Estas políticas administrativas permiten tomar decisiones relativas a la selección, anuncio y recepción de las rutas.

El escenario en el que vamos a centrar nuestra investigación es el encaminamiento interdominio, es decir entre AS's diferentes, los cuales se interconectan mediante enlaces dedicados o puntos de accesos públicos a la

Así, en la sección 2 se describe un algoritmo genético novedoso al que hemos denominado Algoritmo Genético del Ciervo (en adelante AGC), inspirado en el comportamiento social y reproductor de los ciervos. En la misma sección se formula matemáticamente el problema que abordamos y la adaptación del AGC a dicho problema. En la sección 3 se recogen los resultados más destacables del funcionamiento del AGC. Finalmente, en la **sección 4** se presenta las conclusiones más importantes de este trabajo y se apunta a la utilización del AGC como parte fundamental de la estructura de un ente supervisor de Internet, que necesita de la rapidez de convergencia del AGC para su óptimo funcionamiento. Por último se discuten posibles extensiones de este trabajo a otros problemas

2. Algoritmo Genético del Ciervo

El modelo que presentamos está basado en el comportamiento social y reproductor de los ciervos, que dentro de la evolución de las especies es tremendamente selectivo. Asimismo, hay que destacar que los últimos trabajos desarrollados por Ahn y Ramakrishna [1] utilizan el torneo como método de selección, lo que sugiere que las mejoras de los algoritmos genéticos (en adelante GA's) apuntan a variantes en este sentido, como demuestran tanto [11] como [1] con la eficacia de sus métodos.

El GA social desarrollado en la referencia [11], demostraba que si el criterio de selección es muy restrictivo y, por otro lado se completaba la población obtenida por cruces y mutaciones convencionales, con individuos generados aleatoriamente a los que denominó inmigrantes, se obtenía una buena solución en las primeras generaciones. Basándonos en este comportamiento, en este trabajo se desarrolla un nuevo modelo mejorado en el que se incluye la eliminación de mínimos locales, gracias a la incorporación de elementos perturbadores que se tratarán más adelante.

El problema que abordamos es el de la optimización del encaminamiento en Internet, por lo que el AGC necesita conocer la topología de Internet para determinar las rutas entre dos sistemas autónomos (AS's). Para ello, utilizará la información generada por el proyecto Route Views (en adelante RV) http://www.routeviews.com/ del "Advanced Network Technology Center" de la Universidad de Oregón. El proyecto RV obtiene múltiples vistas de la tabla de encaminamiento global que permitirá a los routers de RV disponer de una completa información de encaminamiento gracias a la interacción con distintos AS's distribuidos por todo el mundo. Es por tanto una potente herramienta destinada a obtener información en tiempo real acerca de los sistemas de encaminamiento globales, desde la perspectiva de diferentes backbones1 y localizaciones alrededor de Internet. Básicamente estos datos dan la información completa sobre las rutas empleadas, en la realidad, para alcanzar desde un nodo origen un nodo destino, manejando para ello mas de ocho millones de rutas.

Redes y servicios telemáticos secciones técnicas

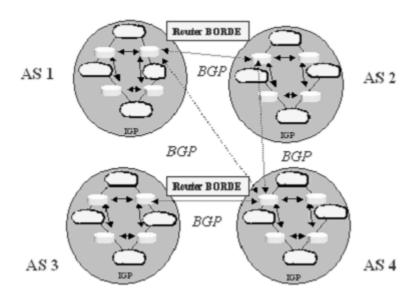


Figura 1. Esquema simplificado de Internet.

Centrándonos en el objeto de este trabajo, el AGC consiste en reproducir artificialmente la vida de los ciervos para obtener rutas mejores entre AS's que las obtenidas por los algoritmos que están al uso. El modelo está basado, por tanto, en el comportamiento social y reproductor de los ciervos que, dentro de la evolución de las especies, es muy selectivo. Este algoritmo de encaminamiento debe utilizar la misma métrica que BGP, es decir, el mínimo número de saltos.

2.1 Formulación matemática del pro-

A partir de los datos aportados por el proyecto RouteViews, en forma de ficheros de datos (tablas de encaminamiento globales de Internet) y obtenidos mediante la colaboración e intercambio de tablas de rutas de un número elevado de AS's, se obtiene la topología de la red necesaria para nuestro modelo. El problema que afrontamos tiene su fundamento en la teoría de grafos [2], dado que partimos de una red de comunicaciones formada por AS's (nodos) y conexiones físicas (arcos). Así, el modelo que se propone en este trabajo es el siguiente:

Sea M el conjunto de todos los nodos, sea V(M) el conjunto formado por todos los arcos que enlazan estos nodos, es decir, un subconjunto del producto cartesiano MxM:

$$V(M) \subset M \times M = \{(i,j)/i, j \in \{1,...,m\}\}\$$
 (1)

Por tanto, disponemos de una red o grafo $G \equiv (M, V(M))$ como soporte para obtener todos los caminos posibles entre dos puntos cualesquiera que supondremos finito, con cardinalidad m, y a los que identificaremos con los números sucesivos de 1 a m. Se impone, como condición, que la red sea

fuertemente conexa. Se dice que el nodo i es adyacente a otro j cuando existe al menos uno de los dos arcos (i, j) y (h, k). Dos arcos(i, j) y (h, k) son advacentes si i = k, j = k $h, i = h \circ j = k$

La etiqueta de un arco o de una arista también recibe el nombre de peso o coste C_{ii} , en cuyo caso el grafo recibiría el nombre de grafo con pesos o grafo valorado. En nuestro caso el coste C_{ii} será igual a 1 ya que se busca la alcanzabilidad entre dos AS's.

El problema que abordamos trata de dar solución al problema del encaminamiento que consiste en visitar, partiendo de un nodo origen, otro conjunto de nodos hasta llegar al nodo destino, de tal forma que la distancia recorrida sea mínima. Se supone que cada nodo se visita una sola vez.

Dado un grafo G, un camino entre los vértices m_1 y m_{n+1} y es una secuencia de arcos adyacentes cuyo primer origen es m_1 y cuyo último extremo es m_{n+1} , esto es, una secuencia de nodos m_1 y arcos v_i , m_1 , v_1 , m_2 , v_2 , ..., m_n , v_m ,..., m_{n+1} , cuyo origen es el vértice m_1 y cuyo destino es el vértice m_{n+1} y que contiene n arcos $v_i = \{m_i, m_{i+1}\}$, siendo $1 \le i$ $\leq n$. La longitud de un camino es n, es decir, el total de arcos.

La matriz de adyacencia de G, respecto a los vértices anteriores, es una matriz booleana X cuadrada, de dimensión $n \times n$, cuyo elemento a_{ii} es tal que:

$$a_{ij} = \begin{cases} 1, & si(i,j) \in V(M) \\ 0, & en \ otro \ caso \end{cases}$$
 (2)

Para dos nodos dados, h y k, el problema a resolver consiste en encontrar un camino de longitud mínima con origen en hy final en k. Es decir, se trata del problema de optimización

Min
$$\sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij}$$
 (3)

donde la variable de decisión

$$x_{ii} \in \{0,1\} \forall i, j$$

sujeto a:

$$\sum_{i=1}^{n} x_{h,j} = 1 \quad ; \quad \sum_{i=1}^{n} x_{i,h} = 0 \quad (4)$$

(del vértice h sale un arco, pero no llega ninguno)

$$\sum_{i=1}^{n} x_{i,k} = 1 \quad ; \quad \sum_{j=1}^{n} x_{k,j} = 0 \quad (5)$$

(al vértice k llega un arco, pero no de él no sale ninguno)

$$\sum_{i=1}^{n} x_{ip} = \sum_{i=1}^{n} x_{pj} , \forall p \neq h, p \neq k$$
 (6)

(en todo vértice del camino lo que "entra" ha de ser igual a lo que "sale")

$$\sum_{i=1}^{n} x_{ip} \le 1 \quad , \forall p \ne h, p \ne k$$

(hay vértices que no forman parte de la ruta)

$$\forall i, j \quad x_{ij} \leq a_{ij}$$

elemento i, j de la matriz de adyacencia (si entre i y j no hay un arco x_{ii} forzosamente es nula; si

$$a_{ii}=1$$

podrá interesar o no utilizar dicho arco).

Si se trata de una red valorada donde el arco (i, j) tiene asociado el coste c_{ii} la función a minimizar es

$$\sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij}$$

En el problema planteado $c_{ii} = 1$.

El conjunto de restricciones asegura que cada nodo aparezca una sola vez en la ruta, la continuidad de la ruta y que todos los nodos de la red están conectados no quedando ninguno aislado. Este modelo, como se puede observar, es lineal y tiene una solución exacta, pero como se ha explicado es un problema "NP- Hard" [10] que tiene unos tiempos de ejecución en ordenador muy elevados, por lo que es inviable su resolución

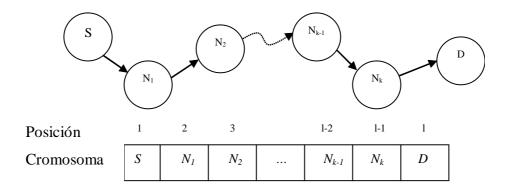


Figura 2. Ejemplo de encaminamiento y su codificación.

por métodos exactos. Un problema es NP (no polinomial) si existe un algoritmo nodeterminista, cuvo tiempo de ejecución es una expresión polinomial en el tamaño de la entrada, es decir, se disparan los tiempos de ejecución al aumentar muy poco el número de variables. Estos problemas se clasifican en "fáciles" o tratables y "difíciles" o intratables en función de si existe un algoritmo polinomial de tiempo para el primer caso; o un algoritmo superpolinomial en el segun-

Es por ello que los problemas NP se suelen resolver por medio de métodos heurísticos que permitan explorar el ámbito de las soluciones y obtener un óptimo, que si bien puede no ser la mejor, si es un valor válido obtenido en un tiempo razonable. Es habitual que los distintos métodos clásicos desarrollados para resolver este tipo de problemas manejen una solución cada vez. Por otro lado, los más modernos, y dado el avance en tecnologías de la computación, trabajan con más soluciones a la vez por lo que la exploración del ámbito de soluciones es más abundante y, por tanto, aplicando los criterios adecuados de convergencia de soluciones, es más fácil mejorar en la obtención del valor óptimo a la hora de seleccionar un algoritmo para resolver este problema.

2.2 Descripción del Algoritmo Genético del Ciervo

Empleamos cromosomas de longitud variable, donde los genes representan los nodos que componen la ruta entre el par origendestino, previamente fijado. El cruce intercambia rutas parciales de los cromosomas padres; el punto de cruce en ambos cromosomas es el mismo identificador de nodo por lo que aseguramos que las nuevas rutas son, a priori, buenas. No obstante, se ha diseñado una función que permite "arreglar" rutas con algunas anomalías procedentes del cruce (rutas con nodos repetidos, etc.)

El algoritmo desarrollado es un algoritmo bio-inspirado en el que se trata de imitar la forma de vida de los ciervos. Este tipo de

algoritmos se han de desarrollar reproduciendo el comportamiento exacto en la naturaleza, constituyendo una alternativa innovadora a la hora de resolver problemas. Esta tendencia en el mundo de la investigación de la inteligencia artificial [4] [5] [22] marca las nuevas tendencias en este tipo de desarrollos. Es decir, todas las características reales de la naturaleza (probabilidades de cruce, supervivencia, población, mutación, etc.) se trasladan al modelo.

El AGC consiste en la simulación de las características fundamentales de la vida de los ciervos. Es decir, lo mismo que el ciervo trata de evolucionar en las sucesivas generaciones hacia individuos más perfectos, fuertes y mejor dotados para sobrevivir, nuestro propósito es hacer evolucionar las soluciones del problema de encaminamiento hacia una solución óptima. Como se ha dicho, este problema consiste en minimizar el número de arcos de la ruta entre cualquier par de AS's.

Para ello el GA propuesto se desarrolla a lo largo de las siguientes etapas:

- 1. Representación del problema. El cromosoma del AGC se compone de una secuencia de números enteros positivos que representan los identificadores de los nodos (AS's) por los que pasa una ruta. Por tanto, cada posición del cromosoma representa el orden en que es visitado un nodo en la ruta. La primera posición siempre será para el nodo origen, mientras que la última estará reservada para el nodo destino. El cromosoma codifica el problema incluyendo los identificadores de los nodos desde el origen hasta el destino basándose en la información de la topología de la red proporcionada por BGP. En la figura 2 mostramos un ejemplo genérico.
- 2. Población inicial. La población inicial es generada aleatoriamente con el método de codificación explicado anteriormente. No se ha utilizado una inicialización heurística ya que ésta suele explorar sólo una pequeña parte del espacio de soluciones y nunca encuentra

una solución óptima global debido a la falta de diversidad en la población [15]. El gen de la primera posición codifica el nodo origen, mientras que el gen de la segunda posición es seleccionado al azar del conjunto de nodos conectados con el origen. Cada vez que se selecciona un nodo, se analiza el cromosoma para evitar la aparición de bucles. Si éste es el caso, hay una función encargada de "reparar" la ruta. Este proceso se repite hasta que se alcanza el nodo destino.

La población inicial en el AGC está compuesta por tres generaciones de l cromosomas cada una, a partir de la topología de Internet obtenida del Proyecto RV. Se crean tres generaciones de posibles soluciones al problema del encaminamiento, ya que las hembras serán fértiles a partir del segundo año, y evolucionan en futuras generaciones con el objetivo de conseguir una mejor solución. En cada generación, el AGC discriminará si la solución generada corresponde a un macho o a una hembra, con el fin de representar fielmente la vida de los ciervos. La proporción de sexos al nacer es 1:1, es decir vienen al mundo tantos machos como hembras. Cada generación representa un año de vida de los ciervos. A continuación se evalúan las soluciones correspondientes a esta población aplicando la función objetivo.

3. Selección del macho dominante. Se determina cual es el macho dominante de entre todos los machos pertenecientes a las generaciones 3-8 (generaciones en las que los ciervos machos está en disposición de procrear), es decir aquel cromosoma que tenga como coste el menor. Para ello, se ha utilizado como operador de selección el torneo, en virtud del cual el mejor cromosoma macho de un conjunto definido de generaciones "lucha" contra todos y cada uno de los cromosomas machos de dicho conjunto (generaciones 3-8), siendo el tamaño del torneo 2, al igual que en la naturaleza. El mejor cromosoma, denominado "macho dominante" será el padre de la siguiente generación. En la población inicial el macho dominante se encontrará en la generación 3.

Redes y servicios telemáticos secciones técnicas

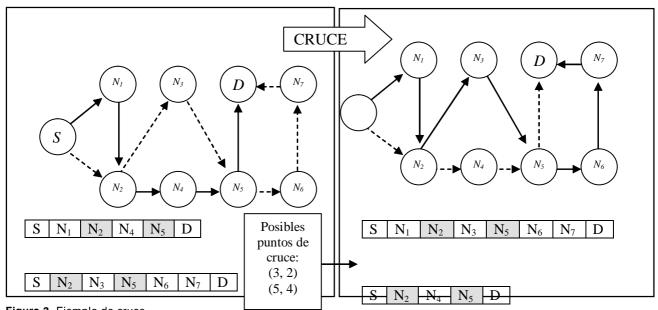


Figura 3. Ejemplo de cruce.

4. Torneo. El macho dominante, antes de cubrir a todas las hembras fértiles entre la generación 3 y la 8, debe luchar contra todos los demás machos. Como en la vida real, puede darse el caso de que un ciervo macho (un cromosoma con peor evaluación) gane en la lucha al dominante, siendo esta posibilidad un primer mecanismo que evita que el AGC quede atrapado en un mínimo local. Otro proceso que permite huir de los mínimos locales es la aparición de un ciervo asesino, circunstancia que, al igual que en la realidad, se da con una probabilidad remota, hecho que es reflejado en el AGC, evitando así posibles mínimos locales. Este ciervo, denominado asesino, pasa a ser el dominante. Por otra parte, se introduce otro proceso perturbador en esta selección que consiste en reproducir el entrecruzamiento de cornamentas, en cuyo caso morirán los dos. Por tanto, se debe encontrar de nuevo un macho dominante entre las generaciones 3 y 8 y repetir el proceso. El ciervo asesino se corresponde a una mutación convencional en la que el salto que se da de la cuenca de atracción de un mínimo local puede ser importante porque el asesino puede tener un material genético muy dispar.

En cambio, el entrecruce de cornamentas correspondería a una mutación más suave dado que se busca el siguiente mejor candidato. La probabilidad de que estos hechos ocurran en la realidad es muy baja. Para nuestro problema viene bien este hecho dado que el esfuerzo que se necesita para salir de un mínimo local puede ser grande y no hay garantías de la salida de la cuenca de atracción del mínimo local. Este proceso por tanto penalizaría los tiempos de ejecución.

5. Reproducción (cruce). El operador cruce utilizado en esta tesis es el denominado cruce "en n-puntos" o cruce "n-puntual".

Permite intercambiar subrutas parciales entre los dos cromosomas seleccionados, de manera que cada hijo resultante represente solo una ruta. Este tipo de cruce produce hijos con rasgos dominantes. Lógicamente, los dos cromosomas seleccionados para el cruce deben tener al menos un nodo en común, además del origen y destino, no siendo necesario que se encuentren en la misma posición en ambos cromosomas; es decir el cruce no depende de la posición que ocupan los nodos en los cromosomas. En la figura 3 [1] se muestra un ejemplo del proceso de cruce.

El proceso de cruce comienza localizando los puntos de cruce, tanto en el macho dominante como en la hembra. A continuación se elige al azar el cromosoma del cual se va a heredar la subruta (desde el origen al primer punto de cruce). A partir de este momento se repite el proceso entre los puntos de cruce intermedios hasta que se alcance el destino. Se pueden ir heredando, alternativamente, subrutas pertenecientes unas veces al macho dominante y, otras a la hembra.

Los puntos de cruce de los dos cromosomas, como se ha comentado anteriormente y puede verse en la **figura 3**, pueden ser diferentes en cada cromosoma. Es posible que aparezcan bucles tras la operación de cruce. Para evitar esta posibilidad se ha desarrollado una subrutina (función de recuperación) que elimina de la ruta los eventuales bucles, con un coste computacional realmente bajo. **La figura 4** [1] muestra este proceso.

El macho dominante cubrirá a todos los cromosomas hembras en edad de procrear (generación de la 3 a la 8). Se creará así una nueva generación que pasará a ser la generación número 1. La primera generación (que ya existía) pasa a ser la segunda y así sucesi-

vamente. En cada apareamiento podrá haber como máximo 2 hijos y como mínimo 1. La probabilidad de cruce Pc es fijada al comienzo. Dado que la probabilidad real de cruce es 0,9, y apoyándonos en los trabajos de investigación de [12] [16] y [21], en las ejecuciones del caso práctico se utilizará esta probabilidad. Se genera un número aleatorio Aleatorio_cruce del intervalo [0,1]. Si Aleatorio_cruce Pc hay cruce. En este caso queda por determinar el número de hijos. Para ello se genera otro aleatorio, Aleatorio_hijos, del intervalo [0,1] pudiendo darse dos casos:

- 1. Que *Aleatorio_hijos* < 0.95. En este caso se obtendrá un solo hijo.
- 2. Que *Aleatorio_hijos* ≥ 0.95. En este caso se obtendrán dos hijos.
- **6. Convergencia.** Para comprobar la convergencia de la solución, comparamos las mejores soluciones de las generaciones ky k-1, pudiendo ocurrir:

$$\frac{S_{k-1}}{S_k} \begin{cases} > 1 & \text{Converge} \\ < 1 & \text{Diverge} \end{cases}$$

Si diverge la mejor solución de la generación k tomará el valor de la k-1.

- 7. Evolución. Los pasos 3 al 5 se repetirán hasta alcanzar el número de generaciones predeterminado. Se ha comprobado que ese número de generaciones es 20 dado que el comportamiento del AGC es muy selectivo y obtiene buenos resultados rápidamente.
- 8. Criterio de parada. Tal y como se señala en el párrafo anterior el AGC se detendrá cuando alcance el número de generaciones que se haya preestablecido. Otro criterio utilizado es el de conseguir una misma solución óptima en tres generaciones sucesivas. Este criterio viene sugerido porque el problema a resolver es una variable discreta (nú-

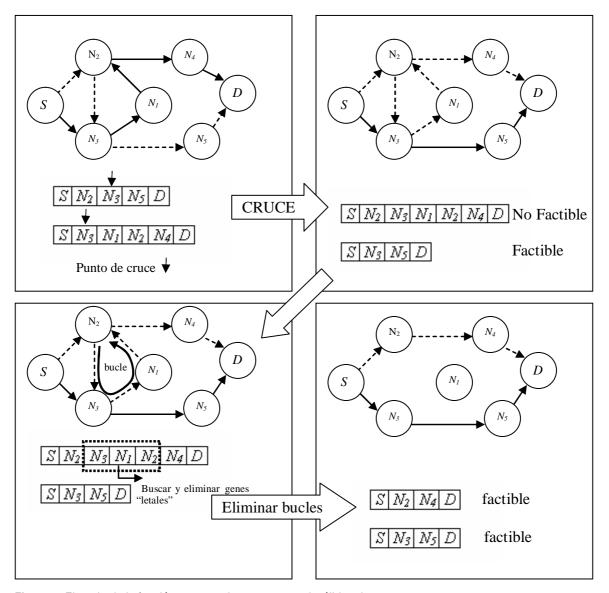


Figura 4. Ejemplo de la función que arregla cromosomas inválidos de cruce.

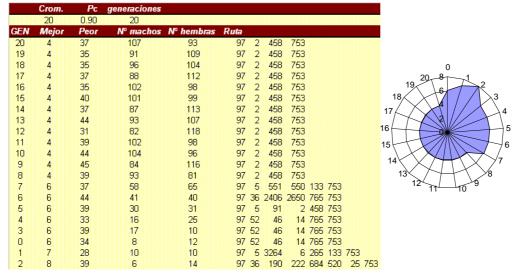


Figura 5. Ejecución del AGC Ruta 1907-1509.

Redes y servicios telemáticos secciones técnicas

	Crom.	Pc	Generaciones												
	20	0.90	20												
GEN	Mejor	Peor	Nº machos	Nº hembras	Ruta										
20	7	40	105	95	1907	615	303	11	140	26	1509				
19	7	37	87	113	1907	615	303	11	140	26	1509				
18	7	29	106	94	1907	615	303	11	140	26	1509				
17	7	45	108	92	1907	615	303	11	140	26	1509				
16	7	34	96	104	1907	615	303	11	140	26	1509				
15	7	33	97	103	1907	615	303	11	140	26	1509				
14	7	37	97	103	1907	615	303	11	140	26	1509				
13	7	36	87	113	1907	615	303	11	140	26	1509				
12	7	34	99	101	1907	615	303	11	140	26	1509				
11	7	45	101	99	1907	615	303	11	140	26	1509				
10	7	38	104	94	1907	615	303	11	140	26	1509				
9	7	45	91	103	1907	615	303	11	140	26	1509				
8	8	40	71	67	1907	615	2738	303	11	140	26	1509			
7	8	38	55	49	1907	615	2738	303	11	140	26	1509			
6	8	47	39	33	1907	615	2738	303	11	140	26	1509			
5	8	42	26	26	1907	615	361	17	485	1833	26	1509			
4	8	27	16	18	1907	615	2738	303	11	140	26	1509			
3	- 8	43	13	11	1907	615	2738	303	11	140	26	1509			
0	11	39	12	8	1907	615	2439	9	736	2860	27	25	483	26	1509
1	9	42	9	11	1907	615	2932	303	10	2	597	26	1509		
2	10	41	8	12	1907	615	616	2467	2	3075	12	484	26	1509	

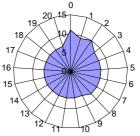


Figura 6. Ejecución del AGC Ruta 97-753.

mero de saltos) cuyo valor máximo está en torno a 20 saltos. En este problema, y dado que se trata de obtener una solución en un tiempo razonable, se puede detener la ejecución al obtener un buen valor aunque este sea un mínimo local. No obstante en las ejecuciones de los casos prácticos siempre se muestra las generaciones indicadas para ver su evolución.

El AGC se caracteriza por ser muy rápido en la obtención de una buena solución. Esto se debe a que el macho dominante es el mejor individuo de toda la población y va transmitiendo su material genético generación tras generación permitiendo converger rápidamente aunque sea a un mínimo local. Es decir, en la operación de reproducción del AGC el macho dominante está transmitiendo material genético a las sucesivas generaciones y en las que, además, se cruzará posteriormente con las hembras de todas las generaciones con capacidad de reproducción.

Esto implica que a medida que se va evolucionando este material genético es cada vez más coincidente. Por tanto, lo que se está obteniendo son esquemas que cada vez serán de longitud y orden mayores. Esto supone que cada vez habrá menos representantes de soluciones distintas y, al menos, se obtendrá rápidamente un óptimo local.

Para salir de ese óptimo local y explorar nuevas zonas de soluciones, las perturbaciones que se emplean, entrecruce de cornamentas y ciervo asesino, pueden permitir dar ese salto del grupo de soluciones locales para ir hacia el óptimo global.

Estas perturbaciones introducen ese factor aleatorio propio de las mutaciones en los GA's convencionales, pero cabe destacar que el comportamiento esperado a priori de cada uno de ellos es diferente. Es decir, el entrecruce de cornamentas lleva a la muerte de ambos púgiles y es más que probable que el candidato a macho dominante, el cual habrá que localizar en todas las generaciones activas, tenga muchas características comunes con el macho dominante muerto, con lo que la probabilidad de explorar nuevas soluciones es menor.

En cambio la aparición del ciervo asesino significa que éste será el que se apareará y, por tanto, la probabilidad a repetir el coniunto de soluciones en curso es más remota: excepto en el caso en el que se encuentre ya en la cuenca de atracción del óptimo global.

3. Resultados y funcionamiento del AGC

En este apartado, se exponen los resultados obtenidos después de realizar un conjunto de ejecuciones del AGC, donde se comprueba de forma empírica su convergencia. Como puede observarse en la tablas incluidas en las **figuras 5** y **6**, para cada ejecución se fija el número de cromosomas de la población inicial en 20, el número de generaciones a crear en 20 y la probabilidad de cruce en 0,9. Cada fila se corresponde con una generación (véase columna 1), mostrándose el mejor y peor resultado (columnas 2 y 3), en número de AS's que componen la ruta, el número de ciervos machos y ciervos hembras (columnas 4 y 5) creados, así como la ruta mejor desde un nodo origen a un nodo destino. Además se incluve un gráfico radial explicativo de la convergencia del AGC (parte derecha de las figuras 5 y 6), donde los números exteriores 0-20 hacen referencia al número de la generación y los números interiores al número de saltos de la mejor ruta

Como se puede observar el AGC obtiene una buena solución, en la mayoría de los casos, a partir de la sexta generación teniendo en cuenta que en las tres primeras no se produce cruce de material genético y por tanto correspondería a la generación cero de un GA convencional. En la tabla 1 se muestra un resumen de resultados obtenidos en un conjunto de ejecuciones entre las que se incluyen la expuestas anteriormente.

4. Conclusiones

En este trabajo hemos presentado un nuevo algoritmo bioinspirado (AGC) que demuestra ser una herramienta eficaz para la generación de rutas óptimas en Internet. Este algoritmo, se basa en la dinámica de las comunidades de ciervos. La convergencia del algoritmo es rápida, gracias a la incorporación de elementos perturbadores (ciervo asesino y cruce de cornamentas), así como la herencia de gran parte del material genético del macho dominante.

En las referencias [8] y [9] se presenta un modelo de ente supervisor, denominado

Crom.	Pc	Generaciones									
20	0.90	20									
		Número de saltos	Generación								
Origen	Destino	mejor ruta	mejor ruta		Ruta						
7	2652	6	5	7	724	1734	36	321	322	2652	
2397	1549	5	6	2397	1066	133	134	9	1549		
7	1321	5	- 6	7	1028	6	1361	4957	1321		
903	4887	4	6	903	724	7	1145	4887			
819	2052	3	6	819	12	824	2052				
97	753	3	- 8	97	2	458	753				
2775	3751	4	- 8	2725	300	2	577	3751			
513	391	3	- 8	513	2	1084	391				
819	2426	3	8	819	6	1285	2426				
1907	1509	6	9	1907	615	303	11	140	26	1509	
2647	3999	6	10	2647	322	321	36	1589	15	3999	
2775	1133	4	11	2725	300	12	442	1133			
7	2318	6	11	7	11	36	775	134	1321	2318	
97	1875	4	11	97	2	458	1253	1875			
2821	1133	5	12	2821	1036	1035	2	442	1133		
741	2426	5	12	741	14	36	2571	1285	2426		

Tabla 1. Resumen de resultados del AGC.

Observatorio, cuya función consiste en asesorar a los AS's sobre cambios en sus políticas de encaminamiento con el fin de mejorar el tráfico de Internet. Para la realización de esta labor se necesita una rápida y potente herramienta de simulación que obtenga, tras múltiples ejecuciones, unas conclusiones de mejora del encaminamiento global. El AGC se presenta como el algoritmo idóneo para la realización de esta función.

Los autores están trabajando en la aplicación de este algoritmo a problemas relacionados con la Calidad de Servicio (QoS) y la ingeniería de Tráfico en redes de comunicaciones. Hay que resaltar, que el AGC se puede adaptar para el cálculo en topologías ponderadas (coste del enlace distinto de 1), lo cual le hace idóneo para este problema concreto en el que los pesos de los enlaces están asociados a diversos parámetros de QoS. Asimismo se puede generalizar el problema a redes dirigidas.

Futuros trabajos de investigación pueden ir encaminados a resolver problemas de optimización de aplicación industrial. Algunos de estos problemas como, por ejemplo, el problema del viajante con ventanas de tiempo y múltiples rutas, o el problema de la secuenciación óptima de tareas en un taller (job shop), se modelan con redes cuyo conjunto de nodos tienen una cardinalidad muy elevada. En estos modelos el AGC se presenta como un algoritmo idóneo por su rapidez de convergencia, por una parte, y por su versatilidad para modelar sistemas multidimensionales ponderados, por otra. En algunos casos, se podrían requerir la implementación de algoritmos híbridos que se ajustasen al problema a tratar.

Referencias

[1] C.W. Ahn, R.S. Ramakrishna. A genetic algorithm for shortest path routing problem and the sizing of populations. Dept. of Inf. & Commun., Kwangju Inst. of Sci. & Technol., South Korea. In Proceedings Evolutionary Computation, IEEE Transactions, Vol 6, pp. 566-579, 2003.

[2] R. Balakrishnan, K. Ranganathan. A textbook of graph theory, Springer-Verlag, Berlín, 2000. ISBN:

[3] U. Black. IP Routing Protocols, Prentice Hall, New York, 2000. ISBN: 0130142484

[4] P. Bentley. DIGITAL BIOLOGY. How Nature is Transforming our Technology. Headline, 2001.

[5] M. Bonabeau, M. Dorigo, G. Theraulaz. Swarm Intelligence. From Nature to Artificial Systems. Oxford University Press, 1999. ISBN: 0195131584

[6] I. Foster et al. The Anatomy of the GRID: Enabling Scalable Virtual Organizations. Intl. J. Supercomputer Applications, 2001.

[7] L. Gao. On inferring autonomous system relationships in the internet. In Proceedings IEEE/ ACMTransactions on Networking 9(6):733-745, 2001.

[8] J.L. Gahete Díaz, F. Gómez González, A. García San Luis, M. Castro Ponce. Sistema global de asesoramiento basado en la generación automática de rutas óptimas alternativas para la optimización del encaminamiento entre Sistemas Autónomos de Internet. In Proceedings Conferencia Ibero-Americana WWW/Internet 2005, Lisboa.

[9] J.L. Gahete Díaz, F. Gómez González, A. García San Luis, M. Castro Ponce. Observatorio de Internet: Modelo de supervisión, optimización y mejora global del encaminamiento de datos entre Sistemas Autónomos. In Proceedings Congreso Español de Informática (CEDI), Granada (España), 2005.

[10] M.R. Garey, D.S. Jonson. Computers and Intractability: A Guide to the Theory of NPCompleteness. Freeman, San Francisco, CA, 1979. ISBN: 0716710447.

[11] F. Gómez. Sistema de Gestión y Control de Redes Medioambientales. Tesis Doctoral ICAI, 1997. [12] J.J. Grefenstette. Optimization of control parameters for genetic algorithms. IEEE Transactions on Systems, Man and Cybernetics SMC-16(1), pp.

[13] J. Han, M. Kamber. Data Mining: Concepts and Techniques. Morgan-Kaufmann, San Mateo, CA, 2000. ISBN: 1-55860-489-8.

[14] J. Hawkinson, T. Bates. Guidelines for creation, selection, and registration of an Autonomous System (AS). Requests For Comments 1930, 1996.

[15] X. Hue. Genetic algorithms for optimization: Background and applications. Edinburgh Parallel Computing Centre, Univ. Edinburgh, Edinburgh, Scotland, Version 1.0, 1997.

[16] J. Koza. Genetic Programming. On the Programming of Computers by Means of Natural Selección. MIT Press, Cambridge, MA., 1994.

[17] G. Malkin. RIP Version 2. Request For Comments 2453, 1998

[18] J. Moy. OSPF Version 2. Request For Comments

[19] University of Oregon. Proyecto Route Views, 1997 http://routeviews.org

[20] Y. Rekhter, T. Li. A Border Gateway Protocol 4 (BGP-4), Request For Comments 1771, 1995.

[21] J.D. Schaffer, R.A. Caruana, L.J. Eshelman, R. Das. A study of control parameters affecting online performance of genetic algorithms for function optimization. In **Schaffer**, pp. 51-60. Morgan Kaufman Publishers, San Mateo, CA, 1989. ISBN:1-55860-

[22] M. Shipper. Machine Nature. The Coming Age of Bio-Inspired Computing. McGraw-Hill, 2002. ISBN: 0070582467

[23] T. Socolofsky, C. Kale. A TCP/IP Tutorial. Requests For Comments 1180, 1991.



Nota

¹ Línea de gran capacidad a la que se conectan otras líneas de menor capacidad a través de puntos de conexión llamados nodos. La traducción literal es "columna vertebral" o "espina dorsal".



I Concurso Universitario de Software Libre

Juegos, distribuciones, utilidades, VoIP, P2P, simuladores, 3D, sonido y mucho más.

93 proyectos, 135 participantes.



Síguelos en http://concurso-softwarelibre.us.es/planet



















