

UPGRADE is the European Journal for the Informatics Professional, published bimonthly at <http://www.upgrade-cepis.org/>

Publisher

UPGRADE is published on behalf of CEPIS (Council of European Professional Informatics Societies, <http://www.cepis.org/>) by Novática (<http://www.ati.es/novatica/>), journal of the Spanish CEPIS society ATI (*Asociación de Técnicos de Informática*, <http://www.ati.es/>)

UPGRADE monographs are also published in Spanish (full version printed; summary, abstracts and some articles online) by Novática

UPGRADE was created in October 2000 by CEPIS and was first published by Novática and *INFORMATIK/INFORMATIQUE*, bimonthly journal of SVI/FSI (Swiss Federation of Professional Informatics Societies, <http://www.svifsi.ch/>)

UPGRADE is the anchor point for UPENET (UPGRADE European NETwork), the network of CEPIS member societies' publications, that currently includes the following ones:

- *Informatik-Spektrum*, journal published by Springer Verlag on behalf of the CEPIS societies GI, Germany, and SI, Switzerland
- *ITNOW*, magazine published by Oxford University Press on behalf of the British CEPIS society BCS
- *Mondo Digitale*, digital journal from the Italian CEPIS society AICA
- *Novática*, journal from the Spanish CEPIS society ATI
- *OCG Journal*, journal from the Austrian CEPIS society OCG
- *Pliroforiki*, journal from the Cyprus CEPIS society CCS
- *Pro Dialog*, journal from the Polish CEPIS society PTI-PIPS

Editorial Team

Chief Editor: Llorenç Pagés-Casas, Spain, pages@ati.es

Associate Editors:

François Louis Nicolet, Switzerland, nicolet@acm.org

Roberto Carniel, Italy, carniel@dtg.uniud.it

Zakaria Maamar, Arab Emirates, Zakaria.Maamar@zu.ac.ae

Soraya Kouadri Mostéfaoui, Switzerland,

soraya.kouadrimostefaoui@gmail.com

Rafael Fernández Calvo, Spain, rfoalvo@ati.es

Editorial Board

Prof. Wolfried Stucky, CEPIS Former President

Prof. Nello Scarabottolo, CEPIS Vice President

Fernando Pierra Gómez and Llorenç Pagés-Casas, ATI (Spain)

François Louis Nicolet, SI (Switzerland)

Roberto Carniel, ALSI – Tecnoteca (Italy)

UPENET Advisory Board

Hermann Engesser (Informatik-Spektrum, Germany and Switzerland)

Brian Runciman (ITNOW, United Kingdom)

Franco Filippazzi (Mondo Digitale, Italy)

Llorenç Pagés-Casas (Novática, Spain)

Veith Risak (OCG Journal, Austria)

Panicos Masouras (Pliroforiki, Cyprus)

Andrzej Marciniak (Pro Dialog, Poland)

Rafael Fernández Calvo (Coordination)

English Language Editors: Mike Andersson, David Cash, Arthur Cook, Tracey Darch, Laura Davies, Nick Dunn, Rodney Fennemore, Hilary Green, Roger Harris, Jim Holder, Pat Moody, Brian Robson

Cover page designed by Concha Arias Pérez

"Da Vinci Impossible Machine" / © ATI 2007

Layout Design: François Louis Nicolet

Composition: Jorge Llácer-Gil de Ramales

Editorial correspondence: Llorenç Pagés-Casas pages@ati.es

Advertising correspondence: novatica@ati.es

UPGRADE Newsletter available at

<http://www.upgrade-cepis.org/pages/editinfo.html#newsletter>

Copyright

© Novática 2007 (for the monograph)

© CEPIS 2007 (for the sections UPENET and CEPIS News)

All rights reserved under otherwise stated. Abstracting is permitted with credit to the source. For copying, reprint, or republication permission, contact the Editorial Team

The opinions expressed by the authors are their exclusive responsibility

ISSN 1684-5285

Monograph of next issue (August 2007)

"Ambient Intelligence"

(The full schedule of UPGRADE is available at our website)



The European Journal for the Informatics Professional
<http://www.upgrade-cepis.org>

Vol. VIII, issue No. 3, June 2007

Monograph: ICT Certifications for Informatics Professionals (published jointly with Novática*)

Guest Editors: *Luis Fernández-Sanz, María-José García-García, and Peter Weiß*

- 2 Presentation. Certification Systems for Professionalism in Information Technology — *Luis Fernández-Sanz, María-José García-García, and Peter Weiß*
- 5 E-Skills Competences in Europe: A Partnership to Deliver Sustainable Value — *Hugo Lueders*
- 11 The International IT Professional Practice Programme — *Charles Hughes and Colin Thompson*
- 19 The Demand Context: the IT Professional Labour Market — *Matthew Dixon*
- 31 Certifications for Software Professionals Helps to Assure Safety, Reliability in Vital Systems — *Jefferson Welch*
- 36 Survey of ICT Certification Systems for ICT Professionals in Europe — *Roman Povalej and Peter Weiß*
- 46 Perception of Education and Personal Accreditations by IT Professionals in Spain — *María-José García-García and Luis Fernández-Sanz*
- 56 ontoReadiness: A meta-Ontology for Readiness Certification and Career Portability — *Michael L. Brown, Karsten O. Lundquist, Shirley Williams, and Keith Baker*

UPENET (UPGRADE European NETwork)

- 62 From Novática (ATI, Spain)
Internet
Bioinspired Algorithms for Internet Route Optimization — *José-Luis Gahete-Díaz and Fernando Gómez-González*

CEPIS NEWS

- 71 CEPIS Projects: Most Recent Progresses — *François-Phillippe Draguet*

* This monograph will be also published in Spanish (full version printed; summary, abstracts, and some articles online) by Novática, journal of the Spanish CEPIS society ATI (*Asociación de Técnicos de Informática*) at <http://www.ati.es/novatica/>.



Internet

Bioinspired Algorithms for Internet Route Optimization

José-Luis Gahete-Díaz and Fernando Gómez-González

© 2007 Novática

This paper was first published, in Spanish, by **Novática** (issue no. 184, November-December 2006, pp. 56–62). Novática, <<http://www.ati.es/novatica>>, a founding member of UPENET, is a bimonthly journal published, in Spanish, by the Spanish CEPIS society ATI (Asociación de Técnicos de Informática – Association of Computer Professionals).

In this paper we present a new genetic algorithm that we have called the Genetic Deer Algorithm. This algorithm is based on the social and reproductive behaviour of the deer and is particularly useful for addressing the problem of Internet route optimization. We also show the most salient features of the algorithm in terms of convergence. We end by describing its applicability in terms of the monitoring of traffic between autonomous systems on the Internet.

Keywords: Autonomous Systems, Genetic Algorithms, Internet, Optimization, Routing.

1 Introduction

It is well known that the Internet is a worldwide network made up of millions of computers of varying types using a range of platforms, all interconnected by communication equipment whose main purpose is to locate, select, and exchange information from its place of origin to wherever it may have been requested, using the same set of communications protocols known as TCP/IP [23].

In spite of its complexity, from a macroscopic point of view the Internet can be described hierarchically as a system made up of over 16,000 Autonomous Systems (hereinafter AS) [14] (a set of IP networks interconnected by routers making up an administrative entity) which interact in order to coordinate the delivery of IP traffic.

All these ASs form a dense mesh of interconnections giving rise to a routing problem as there may be up to 65,000 source and destination ASs constantly carrying information. As we

have mentioned, one of the biggest problems of this arrangement is how to successfully locate and deliver the information between source and destination; in other words, the routing problem.

The inter-domain protocol Border Gateway Protocol (hereinafter BGP)

[20] was developed specifically for use in TCP/IP networks and has become the standard external (or inter-domain) routing protocol. The purpose of BGP is to ensure the reachability of IP packets between source and destination, irrespective of route optimization. It is not, therefore, the BGP's job to look

Authors

José-Luis Gahete-Díaz holds an Engineering Doctorate in Computing from the *Escuela Técnica Superior de Ingeniería ICAI* of the *Universidad Pontificia Comillas*, Madrid, Spain, where he lectures in the Dept. of Computer Systems. In 1999 he was appointed Director of Studies of Technical Engineering in Business Information Technology. Since March 2001 he has been Director of Studies of Computer Engineering (ICAI). His main fields of interest are artificial intelligence, communications, and security. He has also worked in CIEMAT (Centre for Energy, Environmental and Technological Research) and the ONCE (National Organization for the Blind). <jlghahete@dsi.ica.upcomillas.es>.

Fernando Gómez-González holds a Doctorate in Industrial Engineering and an MBA from the *Universidad Pontificia Comillas*, Madrid, Spain. He was recognized by the Spanish Royal Academy of Doctors in 1998. He lectures on Software and the Creation of Technology-Based Firms. He is a tutor in Spin-off Business Plans for a cluster of universities of the Region of Madrid. He created the *Oficina Emprendedores* (Entrepreneurial Office) and the *Vivero de Empresas* (Enterprise Nursery) of the *Universidad Pontificia Comillas*. Prior to that has been the IT Manager in a number of engineering companies and has taken part in 54 projects on major management systems, remote detection, image processing, real time data acquisition, and innovation. <fgomez@upcomillas.es>.

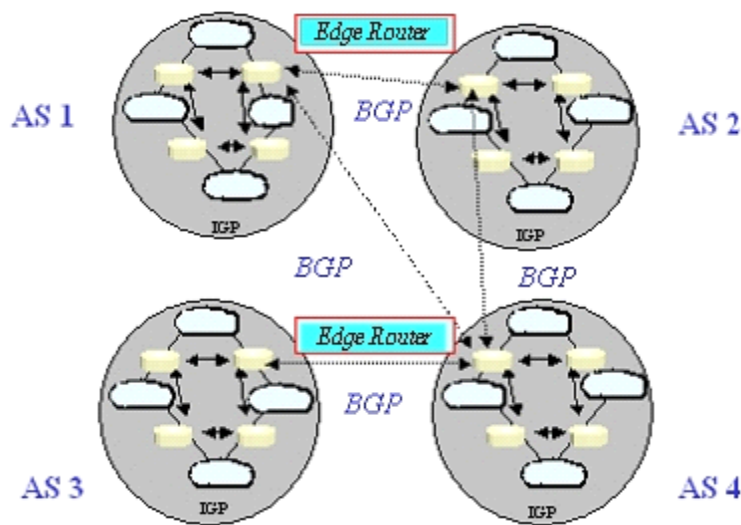


Figure 1: Simplified Diagram of the Internet.

for the global optimal route. This is due, on the one hand, to the restrictions imposed on routing policies chosen by the administrator and, on the other, to the fact that BGP is a distance vector protocol and, more specifically, a path vector protocol (as opposed to a link state protocol) [3].

Figure 1 shows a diagram of how the Internet is interconnected. The grey circles represent ASs; i.e. administrative entities. As we can see in the figure, the AS exchange routing information using BGP. Each AS is made up of various sub-networks connected between routers which will use the intra-domain protocol chosen by the administrator.

As we have already mentioned, the BGP protocol allows AS administrators to set routing policies at the edge routers (see Figure 1). These administrative policies allow decisions to be taken regarding route selection, advertisement, and receiving.

The scenario which our research will focus on is inter-domain routing; that is to say, routing between different ASs which are interconnected by

means of either dedicated links or by public access points to the Internet.

Thus, in Section 2 we describe a new genetic algorithm based on the social and reproductive behaviour of the deer which we have called the Genetic Deer Algorithm (hereinafter GDA). In the same section we formulate the problem we are addressing in mathematical terms and show how our GDA can be used to address that problem. In Section 3 we set out the most salient results provided by the application of our GDA. Finally, in Section 4 we present the most significant conclusions to be drawn from our work and propose the use of the GDA as a fundamental part of the structure of an Internet monitoring entity requiring the convergence speed of the GDA to function optimally. To close we discuss the possibility of extending this work to other problems.

2 Genetic Deer Algorithm

The model we present is based on the social and reproductive behaviour of the deer, which the evolution of the species has rendered extremely selective. It should also be noted that recent research work developed by Ahn and Ramakrishna [1] uses the tournament selection method, which suggests that improvements in Genetic Algorithms (hereinafter GA) should lead to

variations on this theme, as both [3] and [11] demonstrate with the effectiveness of their methods.

The *social GA* developed in the above mentioned reference [11] demonstrated that if the selection criterion is very restrictive and the population obtained by conventional crossovers and mutations, with individuals generated randomly (which he called immigrants) a good solution could be obtained in the first generations. Based on this behaviour, we developed a new improved model in which we included the elimination of local minimums by incorporating perturbation elements (we will be looking at this aspect later in the article).

The problem we address in this work is that of route optimization in the Internet, which means that the GDA needs to know the topology of the Internet in order to be able to determine routes between two Autonomous Systems (AS). To do this the GDA will use information generated by the Route Views (hereinafter RV) project, <<http://www.routeviews.com/>>, from the "Advanced Network Technology Center" of the University of Oregon, USA. The RV project captures multiple views of the global routing table which provides the RV routers with complete routing information by interacting with various ASs distributed around the world. It is therefore a powerful tool intended capture real-time information about the global routing system from the perspective of different backbones¹ and locations around the Internet. Basically this data provides complete information about the routes actually used to reach a destination node from a source node, handling over eight million routes in the process.

Focusing in on the subject of this article, our GDA consists of artificially reproducing the life of the deer to obtain better routes between ASs than the ones obtained by the algorithms currently in use. The model is based on the social and reproductive behaviour of the deer which the evolution of the species has rendered extremely selective. This routing algorithm should use the same metric as BGP; i.e. minimum number of hops.

¹ A high capacity line to which other lower capacity lines are connected by means of connection points called nodes.

2.1 Mathematical Formulation of the Problem

Using data provided by the Route Views project in the form of data files (global Internet routing tables), captured by means of the collaboration and exchange of routing tables from a large number ASs, we obtain the network topology we need for our model. What we are looking at is basically a graph theory problem [2], given that we are dealing with a communications network made up of ASs (nodes) and physical connections (arcs). Therefore the model we propose in this paper is as follows:

Let M be the set of all nodes, let $V(M)$ be the set of all the arcs linking those nodes; i.e. a subset of the Cartesian product $M \times M$:

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

Thus we have a network or graph $G \equiv (M, V(M))$ from which to obtain all possible paths between any two points, which we shall assume to be finite, with cardinality m , and which we shall identify using the sequential numbers 1 to m . A necessary condition is that the network must be strongly connected. Node i is said to be adjacent to another, j , when at least one of the two arcs (i, j) and (j, i) exists. Two arcs (i, j) and (h, k) are adjacent if $i = k, j = h, i = h$ or $j = k$.

An arc or an edge may also be labelled as having a weight or cost C_{ij} , in which case the graph would be called a graph with weights or a weighted graph. In our case cost C_{ij} will equal 1 since we wish to know the reachability between two ASs.

Our task was to find a solution to the routing problem which consisted of starting out from a source node, visiting another set of nodes, and reaching the destination node, in such a way that the distance travelled was as short as possible. It is assumed that each node is visited only once.

Given a graph G , a path between the vertices m_1 and m_{n+1} is a sequence of adjacent arcs whose first source is m_1 and whose final destination is m_{n+1} , that is, a sequence of nodes m_i and arcs $v_i, m_p, v_p, m_2, v_2, \dots, m_n, v_n, \dots, m_{n+1}$ whose source is the vertex m_1 and whose destination is the vertex m_{n+1}

and which contains n arcs $v_i = \{m_i, m_{i+1}\}$, where $1 \leq i \leq n$. The length of a path is n , i.e., the total number of arcs.

The adjacency matrix of G corresponding to the above mentioned vertices is a Boolean square matrix of dimension $n \times n$, whose element a_{ij} is such that:

$$a_{ij} = \begin{cases} 1, & \text{if } (i,j) \in V(M) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

For any two given nodes, h and k , the problem to solve consists of finding the shortest path between the source at h and the destination at k . Thus, this is the optimization problem:

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

in which the decision variable

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

subject to

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

(an arc emanates from vertex h , but no arc terminates there)

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

(an arc terminates at vertex k , but no arc emanates from there)

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

(for every path vertex, whatever "goes in" must be the same as what "comes out")

$$\sum_{i=1}^n x_{ip} \leq 1 \quad , \quad \forall p \neq h, p \neq k \quad (7)$$

(there are vertices that do not form part of the route)

$$\forall i, j \quad x_{ij} \leq a_{ij} \quad (8)$$

element i,j of the adjacency matrix (if there is no arc x_{ij} between i and j it must necessarily be null; if $a_{ij} = 1$ it may or may not be interesting to use that arc).

In the case of a weighted network in which the arc (i,j) has associated cost c_{ij} , the function to minimize is

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

In the problem considered here, $c_{ij} = 1$.

The set of constraints ensures that each node appears only once in the route, that the route is continuous, and that all the nodes of the network are connected and none is left isolated.

As we can see, this model is linear and has an exact solution, but as we have explained it is an NP-hard problem [10] with very long computer run times which means that its exact solution is not feasible. A problem is NP (non-polynomial) if there is a non-deterministic algorithm whose run time is a polynomial expression in the size of the input; in other words, run times will soar with the slightest increase in the number of variables. These problems are classified as "easy" or tractable and "hard" or intractable depending on whether there is a polynomial time algorithm in the former case, or a superpolynomial algorithm in the latter.

This is why heuristic methods tend to be used to solve NP problems as they allow the solution space to be explored so as to find an optimal solution which, while it may not be the best, is valid value obtained in a reasonable time. It is normal for the various classical methods that have been developed to solve this type of problem to handle one solution at a time. However, as computing technology continues to improve, the more modern solutions work with more than one solution at a time. This widens the solution space that can be explored and, if the right solution convergence criteria are applied, makes it easier to obtain a better optimal value when choosing an algorithm to solve this problem.

2.2 Description of the Genetic Deer Algorithm

We use variable length chromosomes in which the genes represent the nodes making up the route between the previously established source-destination pair. The crossover interchanges

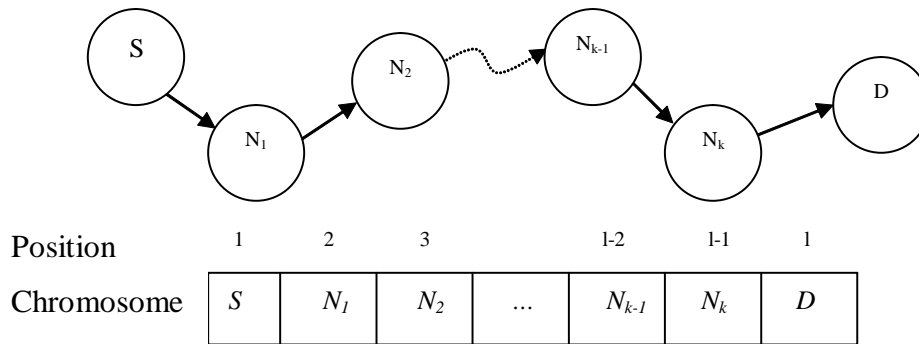


Figure 2: Example of Routing and how it is Encoded.

partial routes from the parent chromosomes; the crossover point in both chromosomes is the same node identifier, thereby ensuring that, a priori, the new routes are good. Nevertheless, a function was designed enabling routes containing some anomalies resulting from the crossover (routes with repeated nodes, etc.) to be "fixed".

The algorithm that we developed is a bioinspired algorithm which aims to imitate the way of life of the deer. This type of algorithm is developed by faithfully reproducing behaviour in nature and provides an innovative alternative approach to problem solving. This new trend in the world of research into artificial intelligence [4][5][12] marks a new direction in this type of developments. All the actual laws of nature (crossover probabilities, survival, population, mutation, etc.) are incorporated into the model.

The GDA works by simulating the basic characteristics of the life of the deer. That is to say, just as the deer strives to evolve in each successive generation towards individuals that are more perfect, stronger, and better equipped for survival, so our aim is to cause the solutions of the routing problem to evolve towards an optimal solution.

As we have already said, this problem consists of minimizing the number of arcs on the route between any AS pair. The GA proposed to solve that problem can be broken down into the following stages:

1. Representation of the problem.

The GDA's chromosome is made up of a sequence of positive whole numbers

representing the node identifiers (ASs) through which a route passes. Thus, each position of the chromosome represents the order in which a node is visited in the route. The first position will always be for the source node, while the last position will be reserved for the destination node. The chromosome encodes the problem including all node identifiers from source to destination on the basis of the information on network topology provided by BGP. In Figure 2 we show a generic example.

2. Initial population. The initial population is generated randomly with the encoding method explained above. Heuristic initialization was not used as it tends to explore only a small fraction of the solution space and never finds a global optimal solution due to the lack of diversity in the population [15]. The gene of the first position encodes the source node, while the gene in the second position is selected at random from the set of nodes connected with the source. Each time a node is selected, the chromosome is analysed to prevent loops from occurring. If this should happen, there is a function responsible for "repairing" the route. This process is repeated until the destination node is reached.

The initial population in the GDA comprises three generations of l chromosomes each, based on the Internet topology obtained from the RV Project. Three generations of possible solutions to the routing problem are created, since the females will be fertile from the second year on, and they evolve in future generations with the aim of

achieving a better solution. In each generation, the GDA will discriminate as to whether the solution generated corresponds to a male or a female, in order to provide a faithful representation of the life of the deer. The proportion of sexes at birth is 1:1, i.e. the same number of males are born as females. Each generation represents a year in the life of the deer. We go on to evaluate the solutions corresponding to this population by applying the objective function.

3. Selection of the dominant male.

The dominant male (i.e. the chromosome with the lowest cost) among all the males belonging to generations 3-8 (generations in which the stags are ready to mate) is determined using tournament selection as the selection operator. In this method the best male chromosome of a defined set of generations "fights" against every male chromosome in that set (generations 3-8) in turn, the size of the tournament being 2, as in nature. The best chromosome, known as the "dominant male", will be the parent of the following generation. In the initial population the dominant male will be in generation 3.

4. Tournament. Before covering all the fertile females between generations 3 and 8, the dominant male has to fight against all the other males. As in real life, a stag (a chromosome with a worse evaluation) may beat the dominant male in a fight; this possibility is the first mechanism preventing the GDA from becoming trapped in a local minimum. Another mechanism to avoid local minimums is the appear-

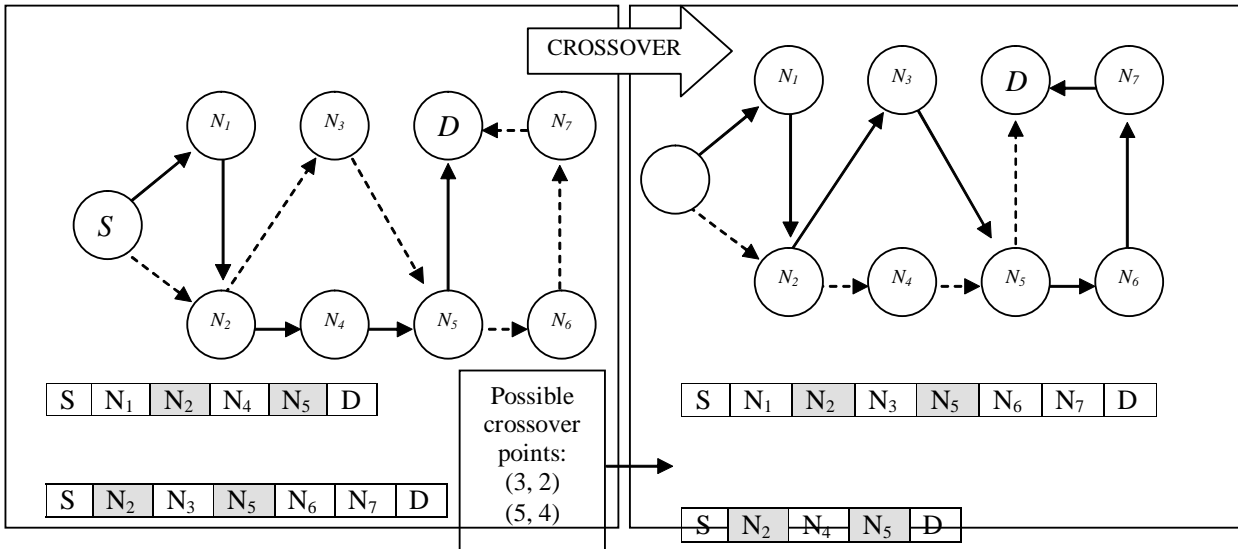


Figure 3: Example of Crossover.

ance of a killer deer. As in real life, this circumstance is a remote probability in our GDA, thereby preventing possible local minimums. This deer, which we call the killer, becomes the dominant male. Another perturbation is introduced into the selection process consisting of reproducing the phenomenon of interlocked antlers which will lead to the death of both stags. In this case we need to find a new dominant male between generations 3 and 8 and then repeat the process. The killer deer corresponds to a conventional mutation in which it may be important to escape from the basin of attraction of a local minimum because the killer may have very different genetic material. However, interlocked antlers would correspond to a milder mutation given that we simply go to the next best candidate. The probability of these phenomena occurring is actually very low. For the purpose of our problem this is a good thing as a great effort may be required to escape from a local minimum and there is no guarantee of escaping from the basin of attraction of the local minimum. This process would therefore have a prejudicial effect on run times.

5. Reproduction (crossover). The crossover operator used in this thesis is known as the "n-point crossover". It allows partial sub-routes to be interchanged between the two selected chromosomes such that each resulting offspring represents a single route. This

type of crossover produces offspring with dominant traits. Logically, in addition to the source and destination, the two chromosomes selected for the crossover must have at least one node in common, though they do not necessarily have to be in the same position in both chromosomes; in other words, the crossover does not depend on the position of the nodes in the chromosomes. Figure 3 [1] shows an example of the crossover process.

The crossover process starts by locating the crossover points in both the dominant male and the female. Next, the chromosome from which the sub-route is going to inherit (from the source to the first crossover point) is chosen at random. The process is then repeated for each successive pair of crossover points until the destination is reached. Sub-routes belonging to either the dominant male or the female can be inherited alternately.

As we have mentioned earlier and as can be seen in the figure above, the crossover points of the two chromosomes may be different in each chromosome. Loops may appear after the crossover operation. To prevent this from happening, a subroutine was developed (recovery function) which eliminates any loops that may occur from the route at a very low computational cost. Figure 4 [1] shows this process.

The dominant male will cover all the female chromosomes at reproduc-

tive age (generations 3 to 8). Thus a new generation will be created which will become generation number 1. The first generation (which already existed) becomes the second and so on and so forth. In each pairing there may be a maximum of 2 offspring and a minimum of 1. The crossover probability (P_c) is set at the beginning. Based on the research work of [12][16] and [21], given that the actual crossover probability is 0.9, this P_c will be used in the executions of the case study.

A random number is generated, $Random_crossover$, of the interval [0.1]. If $Random_crossover \leq P_c$ there is crossover. In this case all that remains is to determine the number of offspring.

To do this another random number is generated, $Random_offspring$, of the interval [0.1] with two possible outcomes:

1. That $Random_offspring < 0.95$. In this case only one offspring will be obtained.
2. That $Random_offspring \geq 0.95$. In this case two offspring will be obtained.

6. Convergence. In order to check the convergence of the solution, we compare the best solutions of generations k and $k-1$, with the following possible outcomes:

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

If there is divergence, the best solution of generation k will take the value of $k-1$

7. Evolution. Steps 3 to 5 are repeated until we arrive at the predetermined number of generations. It has been ascertained that the ideal number of generations is 20 since the behaviour of the GDA is very selective and obtains good results quickly.

8. Stopping criterion. As we stated in the previous paragraph, the GDA will stop once the predetermined number of generations has been reached. Another criterion used is that of obtaining the same optimal solution in three successive generations. This criterion is valid because the problem to be solved is a discrete variable

(number of hops) whose maximum value is around 20 hops. In this problem, given that the idea is to reach a solution in a reasonable length of time, the execution of the GDA can be stopped once a good value has been obtained even if this value is a local minimum. However, case studies are always run for the full 20 generations to see how they evolve.

The GDA can obtain a good solution in a very short time. This is because the dominant male is the best individual of the whole population and transmits its genes through successive generations enabling convergence to be achieved quickly, even though it may be a local minimum. In the GDAs reproduction operation the dominant

male transmits genetic material to successive generations which will subsequently be passed on to the females of all generations capable of reproduction. This means that, as this genetic material evolves, it becomes increasingly more coincident. Thus, the GDA provides us with schemas that will be increasingly longer and of a higher order. This means that with each generation there will be fewer representatives of different solutions and, at the very least, a local optimum will be obtained quickly.

In order to exit a local optimum and explore the solution space more fully, we can use the perturbations – interlocked antlers and killer deer – to provide an escape away from the set of

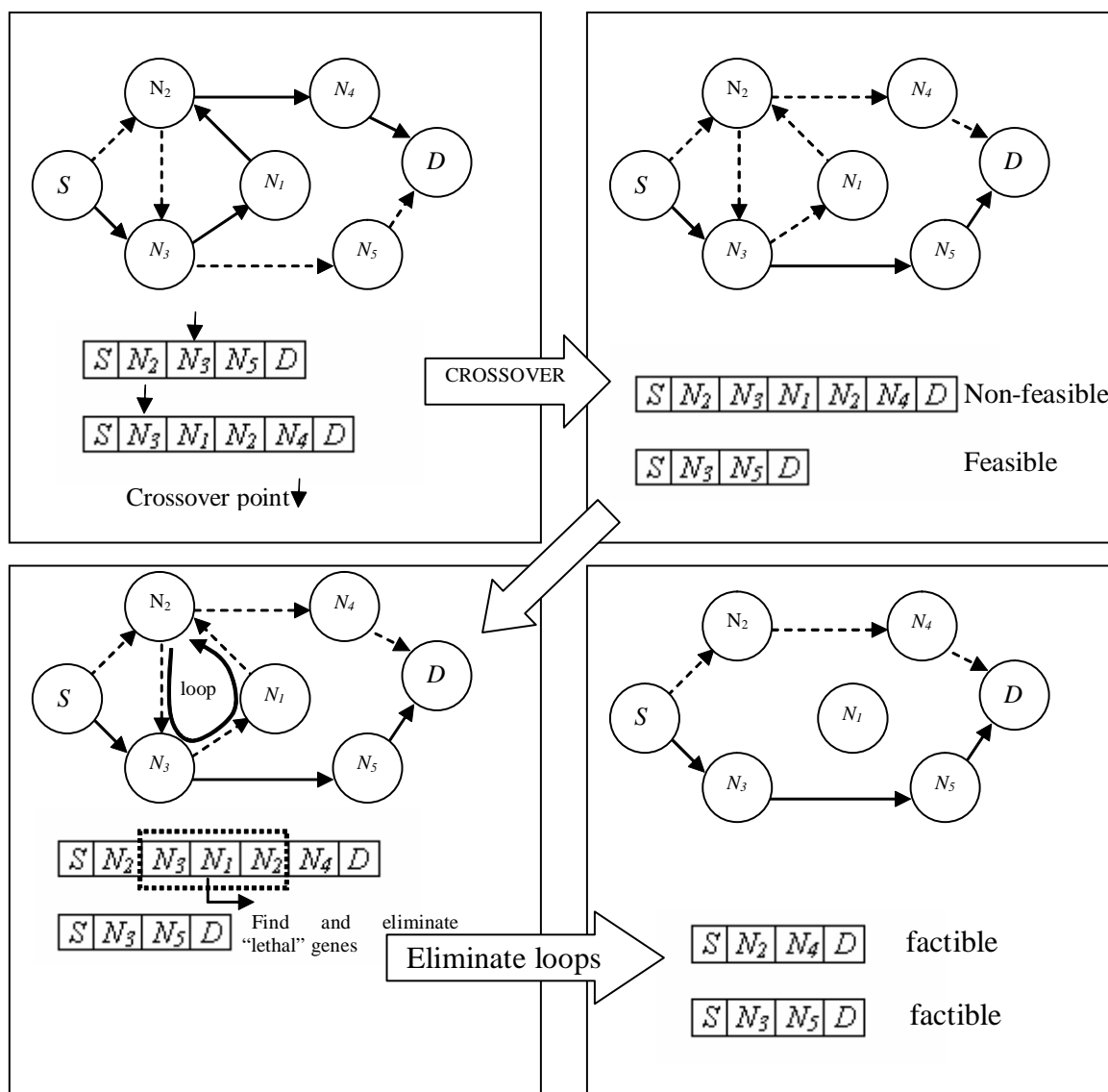
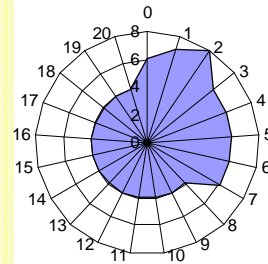


Figure 4: Example of the Function that Fixes Invalid Crossover Chromosomes.

GEN	Crom.		Pc	generaciones		Ruta															
	Mejor	Peor	0.90	Nº machos	Nº hembras	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	
20	4	37		107	93	97	2	458	753												
19	4	35		91	109	97	2	458	753												
18	4	35		96	104	97	2	458	753												
17	4	37		88	112	97	2	458	753												
16	4	35		102	98	97	2	458	753												
15	4	40		101	99	97	2	458	753												
14	4	37		87	113	97	2	458	753												
13	4	44		93	107	97	2	458	753												
12	4	31		82	118	97	2	458	753												
11	4	39		102	98	97	2	458	753												
10	4	44		104	96	97	2	458	753												
9	4	45		84	116	97	2	458	753												
8	4	39		93	81	97	2	458	753												
7	6	37		58	65	97	5	551	550	133	753										
6	6	44		41	40	97	36	2406	2650	765	753										
5	6	39		30	31	97	5	91	2	458	753										
4	6	33		16	25	97	52	46	14	765	753										
3	6	39		17	10	97	52	46	14	765	753										
0	6	34		8	12	97	52	46	14	765	753										
1	7	28		10	10	97	5	3264	6	265	133	753									
2	8	39		6	14	97	36	190	222	684	520	25	753								



Legend: Crom.= Chromosomes; Generaciones = Generations; Mejor = Better; Peor = Worse; Nº Machos = Number of Males; Nº Hembras = Number of Females; Ruta = Route.

Figure 5: Execution of the GDA - Route 97-753.

local solutions towards the global optimum. These perturbations introduce the random factor provided by mutations in conventional GAs, but it should be noted that, a priori, the expected behaviour from each type of perturbation is different. That is to say, the interlocking of antlers causes the death of both combatants and it is highly likely that the candidate for dominant male, which we need to find in all active generations, will have many traits in common with the dead dominant male, thereby reducing the probability of exploring new solutions.

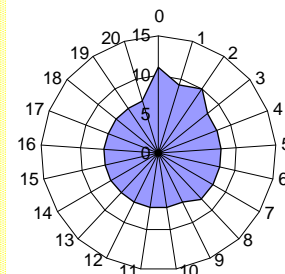
Whereas the appearance of a killer deer means that this will be the one to mate and, therefore, the probability of a set of solutions in progress being repeated is more remote, unless we are already in the basin of attraction of the global optimum.

3 Results and Functioning of the GDA

In this section we set out the results obtained after running our GDA a number of times. The convergence of the method is seen to be empirically proven.

As shown in the tables in Figures 5 and 6, the number of chromosomes of the initial population for each execution is 20, the number of generations to be created is 20, and the crossover probability is 0.9. Each row corresponds to one generation (see column 1) and shows the best and worst result (columns 2 and 3), the number of ASs making up the route, the number of stags and female deer (columns 4 and 5) created, and the best route from a source node to a destination node. Also included is a radar graph (right hand side of Figures 5 and 6) showing GDA

GEN	Crom.		Pc	Generaciones		Ruta															
	Mejor	Peor	0.90	Nº machos	Nº hembras	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	
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	2660	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						



Same legend as in Figure 5.

Figure 6: Execution of the GDA - Route 1907-1509.

Crom.	Pc	Generaciones									
20	0.90	20									
Origen	Destino	Número de saltos mejor ruta		Generación mejor ruta		Ruta					
		7	2652	6	5	7	724	1734	36	321	322
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		

Legend: Crom.= Chromosomes; Generaciones = Generations; Origen = Origin; Destino = Destination; Número de saltos mejor ruta = Number of hops in best route; Generación mejor ruta = Generation in best route; Ruta = Route.

Table 1: GDA Results Summary.

convergence in which the outer numbers 0-20 refer to the generation number and the inner numbers refer to the hopcount of the best route.

As we can see, in most cases the GDA obtains a good solution from the sixth generation on, given that in the first three generations there is no crossover of genetic material and therefore it would correspond to generation zero of a conventional GA.

Table 1 shows a summary of results obtained from a set of executions among which are the two set out above.

4 Conclusions

In this paper we have presented a new bioinspired algorithm (GDA) which has proved to be an effective tool for generating optimal routes in the Internet. This algorithm is based on the dynamics of communities of deer. The convergence of the algorithm is fast thanks to the inclusion perturbation elements (killer deer and interlocking antlers) and due to the fact that much of the genetic material of the dominant male is passed on.

The references [8] and [9] present a model for a monitoring entity, called Observatorio, the purpose of which is to advise the ASs about changes in their routing policies in order to improve Internet traffic. The performance of this task requires a fast and powerful simulation tool that, after multiple executions, is capable of reaching conclu-

sions that will improve global routing. We present our GDA as the ideal algorithm with which to perform this task.

The authors are working on the application of this algorithm to Quality of Service (QoS) and Traffic Engineering related problems in communications networks. It should be stressed that the GDA can be adapted to perform calculations in weighted topologies (link cost other than 1), which makes it ideal for this particular problem in which link weights are associated with various QoS parameters. The problem can also be extended to directed networks.

Future research work may be aimed at solving optimization problems in the world of industry. Some of these problems, such as the problem of the commercial traveller with windows of time and multiple routes, or the problem of optimal task sequencing in a job shop, are modelled with networks whose set of nodes has a very high cardinality. In these models the GDA is the ideal algorithm due to its speed of convergence on the one hand, and its versatility for modelling weighted multidimensional systems on the other. Some cases may call for the implementation of hybrid algorithms tailored to the problem at hand.

Translation by Steve Turpin

References

[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, Berlin, 2000.
- [3] U. Black. IP Routing Protocols, Prentice Hall, New York, 2000.
- [4] P. Bentley. DIGITAL BILOGY. 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.
- [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/ACM Transactions on Networking 9(6): pp 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 of the Conferencia Ibero-Americana WWW/Internet 2005, Lisbon.
- [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 of the Congreso Español de Informática (CEDI), Granada (Spain), 2005.
- [10] M.R. Garey, D.S. Johnson. Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman, San Francisco, CA, 1979.

- [11] F. Gómez. Sistema de Gestión y Control de Redes Medio ambientales. Doctoral thesis ICAI, 1997.
- [12] J.J. Grefenstette. Optimization of control parameters for genetic algorithms. IEEE Transactions on Systems, Man and Cybernetics SMC-16(1), pp 122-128, 1986.
- [13] J. Han et al. Data Mining: Concepts and Techniques. Morgan-Kaufmann, San Mateo, CA, 2000.
- [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 Selection. 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 2178.
- [19] University of Oregon. Route Views project, <<http://routeviews.org>>, 1997.
- [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.
- [22] M. Shipper. Machine Nature. The Coming Age of Bio-Inspired Computing. McGraw-Hill, 2002.
- [23] T. Socolofsky, C. Kale. 1991. A TCP/IP Tutorial. Requests For Comments 1180, 1991.