An approach for network security with graph coloring problem

SEYED MAHMOOD HASHEMI[1] , JINGSHA HE[2], ALIREZA EBRAHIMI BASABI[3]
Hashemi2138@yahoo.com , jhe@bjut.edu.cn, touraj_ebr@yahoo.com
1,2,3: Beijing University of Technology, School of Software Engineering
Beijing Engineering Research Center for IoT Software and Systems, CHINA

Abstract—modeling of network security is useful approach to comprehend the status. In this paper, network is modeled in a graph. Security problem is solved in graph as Graph Coloring Problem (GCP). In GCP, two adjacent nodes must have different colors. Thus GCP provides the security in the network. One objective in GCP is chromatic number and another objective is total price. We present a multi-objective problem for GCP in this paper. Defined problem is solved with multi-objective simulated annealing and multi-objective imperialist competitive algorithm.

Key-Words—Network Security, Graph Coloring, Multi-Objective Optimization

1 Introduction

Security is important issue for network software. According to ISO 7498-2 security includes five elements [1]: 1-authetication 2-authorization 3-integrity 4-confidentiality 5-non-repudiation. Although all stakeholders in software are responsible to security, the major duty of security is done with administrator. First of all, administrator needs to comprehend the vulnerabilities in the system. Actually, searching exhaustive vulnerabilities is impossible practically. Hence, administrator must protect most priority vulnerabilities. Three approaches have attracted from both industry and academia to present priority of vulnerabilities [2].

i. Common vulnerability scoring system (CVSS)
ii. Graphical models (attack tree, attack graph)
iii. Approaches to analyze the interdependency and casual relation between vulnerabilities.

The work in this paper has been supported by National Natural Science Foundation of China (61272500), National High-tech R&D Program (863 Program) (2015AA017204) and Beijing Natural Science Foundation (4142008).

Network can be modeled as a graph $G = (V, E)$. Vertexes, which are shown with V, are nodes and edges, which are represented with E, are communication channels. To provide security, we have two targets: individual channel to communicated nodes. This problem can be modeled by Graph Coloring problem. Assigning must considers the limitations. For example, economic issues must be considered in assignment. Another important limitations are technical limitation, which constraint the communicate channel between nodes.

Graphical models are very useful methods to understand and comprehend the system's circumstances. Among various problems which are studied in graph theory, Graph Coloring Problem (GCP) is compatible with security of network systems. Given a s imple graph, it consists in assigning one color to each vertex such that two vertices linked by an edge get different colors and the total number of colors used is minimized. It is widely used to model several types of real applications such as scheduling, timetabling, frequency allocation, wavelength routing and many more [3]. We can model the network system in graph, such that nodes represent hosts and edges represent connection channels. We ensure to security of system with GCP.

The aim of this paper is presenting Graph Coloring Problem to model the network security and solve it with number of Evolutionary Optimization Algorithms. The rest of this paper is organized as follow: section 2 is assigned to related works. In section 3, we discuss about subjects which are required in this paper. It includes graph theory and multi-objective optimization. Proposed algorithm is presented in section 4. In section 5, we present experimental results and finally, section 6 is conclusion.

2 Related Works

Graph Coloring is potential model for many problems in various fields. In [4], the constrained group decoder is considered. The advantage of this method is lower transmission. Since, there is no need to feedback in this method, traffic of data stream's transmission decreases. This way needs to employs a fixed-rate channel code and modulation scheme. Khattack et al. focuses on cellular networks [5]. They proposed a pattern for networks. In their pattern, cells are arranged in either an infinite linear array or in some two-dimensional pattern, with interference originating only from immediate neighboring cells. Their pattern is very far from real world, because determination of position for cells in real world is difficult. One problem in wireless communication is Interference problem which is similar to security. In Interference problem, two different communicate channel influence to each other. In [6], different scenarios of interference are represented. In [7], graph coloring is used to reduce interference.

3 Preliminaries

In this section, we discuss about some preliminaries that are needed for our proposed algorithm. First of all, number definitions in Graph Theory. Secondly, we declare definition of *Multi-Objective Optimization*. This definition is needed, because our proposed algorithm is according to it.

3.1 Graph Theory

Graph theory is very usable in computer science, because it has ability to model of many problems. Let $G = (V, E)$ be an undirected simple graph and $V(G), E(G)$ are vertexes and edges of a particular graph *G*. the vertexes *v*,*w* are called adjacent if there is an edge $\{v, w\} \in E$ joining them. The Neighborhood of $x \in V$ is $N(x) = \{y \in v | \{x, y\} \in E\}$ and its closed neighborhood is $N(x) \cup \{x\}$ which is denoted by $N[x]$. The cardinality of a set *A* is denoted by $|A|$. The degree of a vertex $x \in V$ is $\delta(x) = |N(x)|$. The size of the neighborhood of x is $\delta\big(N(x)\big) = \sum_{y \in N(x)} \delta(y)$. A sub-graph Δ of *G* is called a clique if two distinct nodes in Δ are always adjacent in *G*. a clique with k nodes is simply called a k-clique. A k-clique in *G* is defined to be a maximum clique if *G* does not contain any (k+*1*)-clique. The graph *G* may have several maximum cliques. Each maximum clique in *G* has same number which is called as clique size of *G* and denoted by $\omega(G)$.

A coloring of a graph is an assignment of colors to its vertexes [8]. Coloring satisfies the following conditions: 1)each node of graph receive exactly one color; 2)adjacent nodes never receive the same color, so they use numbers 1,…,k as colors and the coloring is defined by a map: $f: v \to \{1, \cdots, k\}$. Therefore, $f(v_1) = f(v_2)\,, v_1 \neq v_2$ implies that v_1 and v_2 are not adjacent. Indeed, edges of that graph is colored and also satisfies two conditions: 1)each edge receive exactly one color; 2)if *x*,*y*,*u*,*v* are nodes of a 4-clique in graph, then the edge $\{x, y\}$ and $\{u, v\}$ cannot have same color (*B* type, it is proposed by Bogdan Zavalnij). In other words, if coloring of edges is identified by $g: E \to \{1, \cdots, k\}$ and we have $g(\{x, y\}) = g(\{u, v\})\,, \{x, y\} \cap \{u, v\} = \emptyset$ implies that *x*,*y*,*u*,*v* are not nodes of a 4-clique in graph [9]. Chromatic number is smallest number of colors which can color a specific graph and denoted by $\chi(G)$ [10].

3.2 Multi-Objective Optimization

Suppose there is a function with *n* variables. The goal of optimization is set variables in order to optimum value of function. Optimization algorithms can be divided into two classes according to number of functions: 1)single-objective optimization algorithms. These classes of algorithms are used when there is just one function as target and we want find optimum values for that function. 2)multi-objective optimization algorithms. These classes of algorithms are used when ther are more then one algorithm as target.

Most of the problems in the real world are multi-objective and they require multi-objective optimization for better solutions. Multi-objective optimization is defined as [11]:

$$\begin{aligned} &Optimize \quad \{f_1(X), f_2(X), \ldots, f_k(X)\} \\ &subject \quad g_i(X) \le 0,\ h_j(X) = 0 \\ &i = 1, \ldots, m \quad j = 1, \ldots, p \end{aligned} \tag{1}$$

Where k is number of objective functions, X is the decision vector, m is number of inequality constraints and p is number of equality constraints. Many researchers have tried to find an appropriate approach to solve multi-objective problems [11, 12, 13, 14].

In this paper, two multi-objective optimization algorithms are considered. The first algorithm is *Multi-Objective Simulated Annealing (AMOSA)* [18]. The second method is *Multi-Objective Imperialist Competitive Algorithm (MOICA)*.

Multi-objective optimization is special type of general optimization and characteristic concept of it is *dominance*. Despite of single-objective optimization approaches, which optimum values are wanted, multi-objective optimization algorithms follow non-dominated values.

Formally, a v ector $\vec{u} = (u_1, u_2, \cdots, u_k)$ is said to dominance $\vec{v} = (v_1, v_2, \cdots, v_k)$ (denoted by $\vec{u} \preccurlyeq \vec{v}$ if and only if $\vec{u}$ is partially less than $\vec{v}$, i.e. $\forall i \in \{1, \cdots, k\}: u_i \le v_i \wedge \exists i \in \{1, \cdots, k\}: u_i < v_i$. The set of non-dominance values are goal of multi-objective optimization according to Pareto Optimality.

We say that a v ector of decision variables $X^* \in \mathcal{F}$ is *Pareto Optimal* if there not exist another $X \in \mathcal{F}$ such that $f_i(X) \le f_i(X^*)$ for all $i = 1, \cdots, k$ and $f_j(X) < f_j(X^*)$ for at least one j [11]. In words, this definition says that X^* is Pareto Optimal if there is not any feasible vector of decision variables $X \in \mathcal{F}$ which would decrease some criterion without causing a simultaneous increase at least one other criterion. Unfortunately, this concept almost always gives not a single solution, but rather than a set of solutions called the Pareto Optimal set. The vector X^* corresponding to the solutions included in the Pareto Optimal set are called *non-dominated*. The image of the Pareto Optimal set under the objective functions is called *Pareto Front*. The concept of Pareto Optimal set in mathematical language is defined as (2).

$$P^* \coloneqq \{x \in \mathcal{F} | \nexists x' \in \mathcal{F}\ \vec{f}(x') \preccurlyeq \vec{f}(x)\} \tag{2}$$

In a multi-objective problem and with Pareto Optimal set, the Pareto Front can be defined as follow:

$$PF^* = \{\vec{u} = \vec{f} = (f_1(x), f_2(x), \cdots, f_k(x)) | x \in P^*\} \tag{3}$$

Where k is number of objects.

3.2.1 Multi-Objective Simulated Annealing

Basic concept in Simulated Annealing is evolution of the solution by simulating the decreasing temperature (*tmp*) in a m aterial, where higher the temperature meaning that higher the modification of the solution at a generation. If temperature of a hot material decreases very fast its internal structure may divers and material becomes hard and fragile. Decreasing temperature slowly yields higher homogeneity and less fragile material. Evolution of the solution is carried at specific temperature profiles. At the first iterations a d iverse set of initial solutions for the problem is produced at higher temperature. And, these solutions are evolved while the temperature decreases to get their local optimums. In multi-objective situation, there are non-dominated solutions which must be kept in the archive, as a candidate of optimal solution.

Along the run of AMOSA algorithm, there are two solutions: current-so and new-so [18]. They can have three states compared to each other: i- current-so dominates new-so, ii- current-so and new-so are non-dominated each other and iii- new-so dominates current-so.

If new-so is dominated by current-so, there may be solutions in archive which dominates new-so. New-so is accepted to the archive by the probability:

$$p = \frac{1}{1 + exp\,(\Delta.tmp)} \tag{4}$$

Where Δ is differencing between new-so and other solutions which dominated new-so. If there are A solutions in the archive,

$$\Delta = \frac{\sum_{i=1}^{A} \Delta_i + \Delta}{A+1} \quad (5)$$

Solutions can escape from local-optima and reach to the neighborhood of the global-optima by this probable acceptance.

If new-so is dominated by some solutions in the archive, (5) is modified to:

$$\Delta = \frac{\sum_{i=1}^{A} \Delta_i}{A} \quad (6)$$

When new-so is non-dominated with all members in archive, then new-so is set as current-so and it is added to the archive.

If new-so dominates some solutions in the archive, then new-so is set as current-so and it is added to the archive and solutions in the archive which are dominated by new-so are removed.

If new-so is dominated by by some solutions in the archive, then (4) is changed to:

$$p = \frac{1}{1+exp(-\Delta)} \quad (7)$$

Where Δ is the minimum of the difference between new-so and dominating solutions in the archive. New-so is set as current-so with the probability (4). If new-so is non-dominated by all solutions in the archive it is set as current-so and added to the archive. If new-so dominates some solutions in the archive, it is set as current-so; it is added to the archive; and all dominated solutions are removed from the archive.

3.2.2 Multi-Objective Imperialist Competitive Algorithm

Imperialist Competitive Algorithm (ICA) is a new method belonged to swarm optimization class. ICA has extraordinary ability to solve optimization problems. The objective of this paper is research about ability of Multi-Objective Imperialist Competitive Algorithm (MOICA) to solve multi-objective optimization problems.

ICA is inspired socio-political events, so most names which are used in it can be found in real world. MOICA starts with an initial population (random solutions) named *countries* [15, 16, 17]. Each country is an individual of solution. Population is partitioned to *empires*. The number of countries and empires are defined by algorithm implementer and denoted with $N_{country}$, N_{empire} . Same as r eal world which any country has own fitness; countries of MOICA also have fitness. The cost or fitness of a country represents its *power*. Power of each country is calculated using the value of that country according to an *objective function*, so $cost^i_j$ is power of j^{th} country for i^{th} *objective function*:

$$power_j = \cos^1_j(value_1, \ldots, value_n) + \ldots$$
$$+ \cos^i_j(value_1, \ldots, value_n) \quad (8)$$

In each empire, the country with the highest power is called *imperial*, and all remaining countries are called *colonies*. The total power of the empire is the sum of the powers of all countries in that empire.

Fig. 1, Imperials and colonies

MOICA same as o ther evolutionary optimization algorithms does some jobs in a loop until the termination criterion of the algorithm is satisfied. At each iteration:

- ✓ The power of all countries is changed. Colonies move toward their imperial. This movement is simple model of assimilation policy that was perused by imperialist country. Let d is distance between imperial and colony. x, θ are random numbers with uniform distribution.

$$x \sim U(0, \beta \times d), \theta \sim (-\gamma, \gamma) \quad (9)$$

β, γ are arbitrary numbers that modify the area that colonies randomly search around the imperial. Power of colonies in empire move with

x, θ.

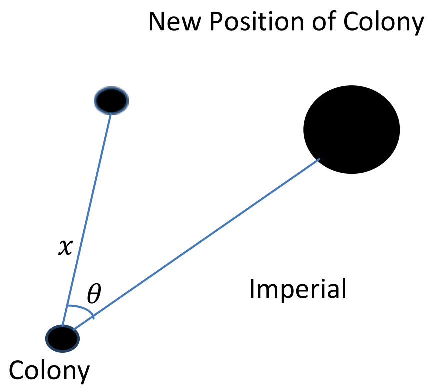

Fig. 2, Movement of colony toward imperial

✓ The power of countries is reevaluated, thus the imperial of any empire may change if one colony enriches power higher than its imperial and can dominate its imperial, substitute with it. if there are number non-dominate country in empire, one of them is selected randomly.

✓ Each imperialist country absorbs colonies of other empires based on power of its empire. This imperialist competition results in the best collection of countries, which corresponds to a solution for a single-objective problem.

Multi-objective ICA requires keeping the non-dominated solutions in a sep arate space, which is called *archive*. There are two important points **(i)** algorithm must keep all non-dominated solutions in an archive to preserve from diversity. **(ii)** The algorithm must avoid deterministic methods to discover a large number of non-dominated solutions. Instead of deterministic movements, using random movements helps the algorithm to escape from *local optimum* and increase the chance of reaching the *global optimum*. Therefore it is better to construct the next population randomly from the archive or from the current population by a probabilistic method.

✓ The algorithm has to decide which country must be kept in archive randomly. All countries are compared with each other (according to (2)) and non-dominated countries are known. If there are k non-dominated countries, the summation of powers for non-dominated countries is given by (10).

$$SumPower = \sum_{i=1}^{k} power_i \qquad (10)$$

All non-dominated countries have probability for entrance to archive and probability for j[th] country is calculated by (11).

$$prob_j = \frac{power_j}{SumPower} \qquad (11)$$

Countries in archive may have 3 situations related to new country:

- The old countries may dominate new country. In this state, new country being fired.
- New country may dominate all or some of the old counties. In this state, dominated countries get out.
- The old countries and new country are non-dominated. In this case, new country is kept in archive.

4 Proposed Algorithm

In this paper, the security problem is solved as Graph Color Problem (GCP). Firstly we need to define the security problem as an optimization problem. Proposed optimization problem must adapt to definition of GCP. In GCP, we look for optimum (minimum) chromatic number, so the objective function of optimization problem is finding the minimum chromatic number for a graph of available nodes. Finding optimized chromatic number is restricted with definition of GCP (two adjacent nodes must have different colors). Thus we reach to following problem:

$$Minimize \quad N, P$$
$$SuchThat \quad \forall x, y \in V \Rightarrow C_x, C_y \in \{1, \cdots, N\},$$
$$if\ x\ is\ adjacent\ to\ y\ \Rightarrow C_x \neq C_y$$

Where V is set of vertexes and C_x means color of node x. P represent the summation price of colors.

The second step is solving the optimization problem. We defined problem with two algorithms which are described above (Multi-Objective Simulated Annealing and Multi-Objective Imperialist Competitive Algorithm).

Proposed algorithm works as following flowchart:

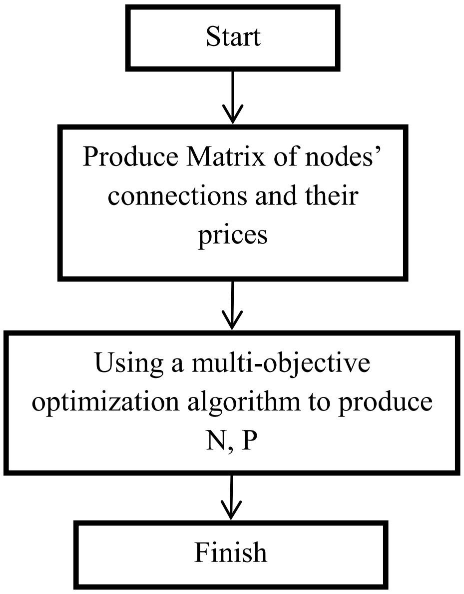

Fig 3. Flowchart of proposed algorithm

5 Experimental Results

For applied the mentioned proposed algorithm and compare between results of three optimization algorithms, we set follow parameters:

$max - iteration = 40, tmp = 150, N_{empires} = 10, N_{countries} = 150$

We applied the proposed algorithm to a network with 10 nodes. Following table represent networks, where numbers present the cost, and ∞ recognizes there is not direct path between two nodes. Note this is directional graph, so $v \rightarrow e$ is not means $e \rightarrow v$.

∞	3	1	5	∞	3	2	4	5	1
3	∞	1	2	1	1	2	2	3	∞
2	3	∞	3	4	2	1	∞	4	2
∞	2	3	∞	3	∞	2	1	3	∞
2	2	3	2	∞	5	3	2	3	∞
3	3	1	1	2	∞	2	2	2	3
∞	1	2	3	2	2	∞	3	2	2
3	5	∞	3	1	2	1	∞	3	2
4	4	3	2	2	3	3	4	∞	3
∞	2	∞	3	1	∞	1	2	3	∞

Table 1. Adjacent matrix for graph

Table 2 presents the final values for N, P with number of iterations which is needed to produce them.

N in AMOSA	P in AMOSA	N in MOICA	P in MOICA	Iterations in AMOSA	Iterations in MOICA
4	4	10	15	8	26

Table 2. Results

Actually both algorithms are stochastic, so the exact value of results may be different in various running. We execute these algorithms more and more and table represents best results within 10 times. Although the values of results are changed in each time, difference between performance of AMOSA and MOICA is constant approximately.

6 Conclusion

Security problem in a network of results is solved in this paper. Network is modeled in a graph and security problem is converted to Graph Coloring Problem (GCP). We define a multi-objective optimization problem for GCP.

References:

[1] International Standard, Information Processing Systems, ISO 7498-2

[2] Shuzhen Wanga, Zonghua Zhang b, Youki Kadobayashi, "Exploring attack graph for cost-benefit security hardening: A probabilistic approach", ELSEVIER, compute rs & s e c u r i t y 3 2 (2 0 1 3) 1 5 8 e1 6 9

[3] Marc Demange, Tınaz Ekim, Bernard Ries, Cerasela Tanasescu, "On some applications of the selective graph coloring problem", ELSEVIER, European Journal of Operational Research 240 (2015) 307–314

[4] Chen Gong, Omar Abu-Ella, Xiaodong Wang, Ali Tajer, "Constrained Group Decoder for Interference Channels", JOURNAL OF COMMUNICATIONS, VOL. 7, NO. 5, MAY 2012

[5] Shahid Khattak, Wolfgang Rave, Gerhard Fettweis, "Distributed IterativeMultiuser Detection through Base Station Cooperation", EURASIP Journal onWireless Communications and Networking Volume 2008, A rticle ID 390489, 2008

[6] M. L. S. N. Swarajya Lakshmi, Niranjan Prasad, "SOURCES OF RADIO LINK INTERFERENCES AND THEIRMITIGATION TECHNIQUES, Golden Research Thoughts, Vol. 2, Issue 7, 2013

[7] Maaly A. Hassan, Andrew Chickadel, " A REVIEW OF INTERFERENCE REDUCTION

IN WIRELESS NETWORKS USING GRAPH COLORING METHODS", International journal on applications of graph theory in wireless ad hoc networks and sensor networks," Vol. 3, No. 1, 2011
[8] Guillemo De Ita Luna, Yuridiana Aleman, Nahum Loya, "Polynomial Strategies for the 3-coloring of a graph", Acta Universitaria, Vol. 22, NE-1, ENC Marzo, 2012
[9] Sándor Szabó, "A Non-Conventional Coloring of the Edges of a Graph", Scientific Research, 119-124, 2012
[10] V. J. Vernold, M. Venkachalam, "A NOTE ON STAR COLORING OF CENTRAL GRAPH OF BIPARTITE
GRAPH AND CORONA GRAPH OF COMPLETE GRAPH WITH PATH AND CYCLE", Transaction of Combinatorics, Vol. 1, No. 1, 2251-8665, 2012
[11] Carlos A. Coello Coello, David A. Van Veldhuizen, Gary B. Lamont, "Evolutionary Algorithms for Solving Multi-Objective Problems", Speringer. 2nd Ed. , 2007.
[12] Cristopher Dimopoulos, "A Review of Evolutionary Multi-objective Optimization, Application in the Area of Production Research", IEEE. Print ISBN 0-7803-8515-2, DOI 10.1109/CEC.2004.1331072. , 2004.
[13] Carlos A. Coello Coello, "An Updated Survey of Evolutionary Multi-Objective Techniques: State of the art and Future Trends", IEEE. Vol.1, Printed ISBN 0-7803-5536-9, DOI 10.1109/CEC.1999.781901. , 1999.
[14] Zdzislaw Kowalczuk, Tomasz Bialaszewski, "Improving Evolutionary Multi-Objective Optimization Using Genders", Springer. Vol. 4029/2006, DOI 10.1007/1178231-42. , 2006.
[15] Arash Khabbazi, Esmaeil Atashpaz, Caro Lucas, "Imperialist Competitive Algorithm for Minimum Bit Error Rate Beamforming", Inte.J. Bio-Inspired Computation, Vol. 1, N os.1/2. , 2009.
[16] Esmaeil Atashpaz, Caro Lucas, "Imperialist Competitive Algorithm: An Algorithm for Optimization Inspired by Imperialistic Competition", IEEE. Congress on Evolutionary Computation. , 2008 .
[17] Caro Lucas, Zahra Nasiri-Gheidari, Farid Tootoonchian, "Application of an Imperialist Competitive Algorithm to the Design of a Linear Induction Motor", ELSEVIER. Energy Conversion and Management. 1407–1411, 2010.
[18] Sanghamitra Bandyopadhyay, Sriparna Saha, Ujjwal Maulik, Kalynmoy Deb, "A Simulated Annealing-Based Multi-Objective Optimization Algorithm: AMOSA", IEEE, Vol. 12, Issue 3, ISSN 1089-778x, 2008