

## 細胞アレイ解析のための線形計画法を用いた レポーター遺伝子の最適配置

林田 守広<sup>†</sup> 孫 富艶<sup>‡</sup> 油谷 幸代<sup>†</sup> 堀本 勝久<sup>†</sup> 阿久津 達也<sup>†</sup>

<sup>†</sup>京都大学 化学研究所 バイオインフォマティクスセンター  
<sup>‡</sup>産業技術総合研究所 生命情報工学研究センター

**概要** 生きた細胞中の様々な環境下で、タンパク質パスウェイや遺伝子転写制御ネットワークのダイナミクスを観測するために、細胞アレイを用いた生物実験手法が開発されている。この手法を適用するためにはレポーター遺伝子を適切に配置することが有用である。そこで本稿では最も効率的なレポーター遺伝子の配置問題を考え、グラフ理論の立場から定式化を行い、二つの問題を定義する。両方の問題ともにNP困難であり、任意の近似比での多項式時間近似も難しいことを示すが、現実的な大きさの問題に対しては最適解を導く線形計画問題による手法も提案する。ヒール細胞のアポトーシスパスウェイに適用し、結果の生物学的意味について考察する。

## Integer Programming-based Approach to Allocation of Reporter Genes for Cell Array Analysis

Morihiro HAYASHIDA<sup>†</sup>, Fuyan SUN<sup>‡</sup>, Sachiyo ABURATANI<sup>†</sup>, Katsuhisa HORIMOTO<sup>‡</sup>, and Tatsuya AKUTSU<sup>†</sup>

<sup>†</sup>*Bioinformatics Center, Institute for Chemical Research, Kyoto University*

<sup>‡</sup>*Computational Biology Research Center, National Institute of Advanced Industrial Science and Technology*

**Abstract** The biological experimental technique using transfected cell arrays has been developed for observing behaviors of protein pathways and genetic networks under various environments in living cells. It is often useful to allocate several reporter genes for cell array analysis. In this paper, we consider the problem of selecting the most effective set of reporter genes. We propose two graph theoretic formulations of the reporter gene allocation problem, and show that both problems are hard to approximate. We propose integer programming-based methods for solving practical instances of these problems optimally. We apply them to apoptosis pathway maps, and discuss biological significance of the result.

### 1 Introduction

Finding novel target genes for the treatment of diseases is an important topic in drug design and systems biology. By reason of its importance, various approaches have been proposed. Among these, *transfected cell microarrays* (*cell arrays* for short) are regarded as a potentially powerful approach [1, 2, 3, 4]. Cell arrays are complementary technique to DNA microarrays. The most important difference is that each spot in a DNA microarray corresponds to a gene, whereas each spot in a cell array corresponds to a cluster of several tens or hundreds of *living cells*. This property enables us to observe times series data of gene expression in living cells. Furthermore, upon the addition of cells and a lipid transfection reagent, slides printed with cDNA become living microarrays, in which some specific gene is overexpressed. On the other hands, it is also possible to knock out some specific gene by using siRNA [1, 3]. Therefore, we may be able to observe effects of gene overexpression or gene knockdown by using cell arrays. We may also be able to observe effects of external signals on gene expressions in living cells. In order to observe the effects using cell arrays, we may need *reporter genes*, which are designed to measure the expression level of gene or the corresponding product through the magnitude of fluorescence. Over the past decade, a battery

of powerful tools that encompass forward and reverse genetic approaches have been developed to dissect the molecular and cellular processes that regulate disease. In particular, the advent of genetically-encoded fluorescent proteins, together with advances in imaging technology, make it possible to study these biological processes in many dimensions [5]. Importantly, these technologies allow direct visual access to complex events as they happen in their native environment, which provides greater insights into human diseases than ever before [6, 7]. However, the cost (both in labor and money) of introduction of reporter genes to a cell is very high. Thus, we cannot use a lot of reporter genes. Instead, we should allocate several or several tens of reporter genes which are the most efficient for identifying the pathways that are significantly activated or inactivated by means of external signals or environmental changes.

There exist related studies. Ruths et al. recently proposed a framework for computational hypothesis testing in which signaling networks are represented as bipartite directed graphs [8]. In their framework, each network contains two types of nodes: nodes corresponding to molecules and nodes corresponding to reactions. They considered two problems: the constrained downstream problem and the minimum knockdown problem. The latter one is closely related to our problem and is to find a minimal set of nodes removal of which disconnects two given sets of compounds. They defined the minimum knockdown problem as a graph theoretic problem. They proved that the problem is NP-hard and proposed an iterative and randomized heuristic algorithm.

In this paper, we consider graph theoretic formulations of the reporter gene allocation problem. Since there is no consensus mathematical model of genetic networks or signaling pathways, we do not assume any specific models such as Boolean networks and Bayesian networks. Instead, we treat each network as a directed graph, where each edge can have a weight. Then, we formulate the reporter gene allocation problem as problems of selecting a set of nodes that covers as many nodes as possible, or selecting a minimal set of nodes that covers all the nodes in a network, where we say that node  $v$  is covered by node  $u$  if there exists a directed path from  $u$  to  $v$  within a specified length. We prove that these problems are NP-hard. Furthermore, we prove that these problems are hard to approximate. We also show that some connection between these problems and the set cover problem (along with its variant). In order to solve realistic instances, we formulate these problems as integer programs (IPs) and apply a famous IP solver (CPLEX) to solving instances of these IPs. This approach is reasonable because a close relationship between integer programming and the set cover is known [9]. It should be noted that our approach is significantly different from that in [8]: (i) problems and network representations are different from each other, (ii) optimality of the solution is not guaranteed in [8], whereas optimality is guaranteed in our approach.

We perform computational experiments using a real biological network. The set of allocated reporters for a real network is reasonable from a biological viewpoint. Furthermore, the proposed method can find the optimal solutions within a few seconds. These suggest that the proposed approach is practically useful for finding an optimal set of reporter genes.

## 2 Allocation Problems

In this section, we define two optimal allocation problems, P1 and P2. Biological networks such as gene regulatory networks and signaling pathways can be considered as a directed graph  $G = (V, E)$  with a set of nodes  $V = \{v_1, \dots, v_n\}$  and a set of directed edges from  $v_i$  to  $v_j$ ,  $(v_i, v_j) \in E$ . In gene regulatory networks, a node means a gene, and in signaling pathways, a node means a protein. It should be noted that a reporter gene can be used both for measuring gene expression and for measuring abundance of proteins.

We define that a node  $v$  is a *neighboring upstream node* of a node  $v_r$  if there is a directed path within the length of a constant  $L$  from  $v$  to  $v_r$  in  $G$ . In this case, we also say that  $v$  is *covered* by  $v_r$ . For a set of nodes  $R$ , we say that  $v$  is covered by  $R$  if  $v$  is covered by some node in  $R$ . This definition can be justified as follows: if some node  $v$  covered by  $v_r$  is affected by external signals and/or environmental changes, it is highly expected (for small  $L$ ) that  $v_r$  is also affected. That

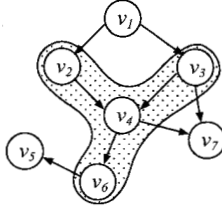


Figure 1: Example of nodes covered by a reporter node when  $L = 2$  in a directed graph  $G = (V, E)$  with  $V = \{v_1, \dots, v_7\}$ . In this case,  $v_2, v_3, v_4$  and  $v_6$  are covered by  $v_6$ .

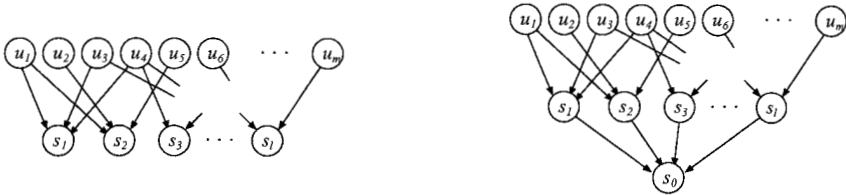


Figure 2: Left: Transformation of an instance  $I = \langle U = \{u_1, \dots, u_m\}, S = \{s_1, \dots, s_l\}, k \rangle$  of the maximum coverage problem to Problem P1. Right: Transformation of  $I = \langle U, S \rangle$  of the set cover problem to Problem P2.

is, we may infer that a subnetwork around  $v_r$  is affected by external signal or environmental change if  $v_r$  is affected, and we want to cover as many parts of the network as possible.

We assume in this paper that  $L$  does not depend on the reporter node and each edge has unit length. This assumption is reasonable because it is difficult to determine  $L$  for each gene or protein and the length of each edge. However, the proposed methods can be modified for a general case in which  $L$  depends on the reporter node and each edge has distinct length (or weight). Figure 1 shows an example of covered nodes by using a reporter when  $L = 2$ .

Problem P1 maximizes the number of covered nodes by using  $K$  reporters, and is defined as follows.

**Definition 1 (Problem P1)** *Given a directed graph  $G = (V, E)$  and two integers  $L$  and  $K$  ( $\leq |V|$ ), find a set  $R \subseteq V$  of cardinality at most  $K$  maximizing the number of nodes covered by  $R$ .*

It should be noted that  $R$  corresponds to a set of reporters. For sufficiently large  $K$ , we can cover all nodes of  $V$  using the solution of Problem P1. In some cases, we may want to cover all the nodes by using a minimum number of reporter nodes. Thus, we also consider the following problem.

**Definition 2 (Problem P2)** *Given a directed graph  $G = (V, E)$  and an integer  $L$ , find a minimum cardinality set  $R \subseteq V$  such that all nodes of  $V$  are covered by  $R$ .*

### 3 Theoretical Results

We show that Problem P1 is MAX SNP-hard, which means that no PTAS exists unless  $P=NP$ . It should be noted that MAX SNP-hardness also implies NP-hardness. For terminology on approximation algorithms, refer to [9].

**Theorem 1** *Problem P1 is MAX SNP-hard.*

*Proof.* We show an  $L$ -reduction from the maximum coverage problem [9, 10], which is known to be MAX SNP-hard [11], to Problem P1. The maximum coverage problem is defined as follows: Given a family of sets  $S$  over  $U$ , and an integer  $k$ , find  $C \subseteq S$  of cardinality at most  $k$  which maximizes the number of covered elements in  $U$ . From an instance  $I = \langle U = \{u_1, \dots, u_m\}, S = \{s_1, \dots, s_l\}, k (\leq l) \rangle$  of the maximum coverage problem, we construct an instance  $I' = \langle G = (V, E), L, K \rangle$  of P1 in the following way (See Figure 2):

$$V = \{u_1, \dots, u_m, s_1, \dots, s_l\}, E = \bigcup_{j=1}^l \bigcup_{u_i \in s_j} \{(u_i, s_j)\}, L = 1, K = k.$$

It should be noted that  $|V| = m + l, |E| = \sum_{j=1}^l |s_j|$ . Thus,  $I'$  can be constructed in polynomial time.

Let  $OPT(I)$  and  $OPT(I')$  be optimal solutions of  $I$  and  $I'$ , respectively. Then,  $OPT(I') = OPT(I) + k$  holds. Without loss of generality, we can assume that  $OPT(I) \geq k$ . Therefore,  $OPT(I') \leq 2OPT(I)$ .

Given any solution  $R \subseteq V$  of  $I'$  with cost (i.e., the number of covered nodes)  $c'$ , we produce a solution  $C$  of  $I$  in polynomial time by letting  $C = R - U$ , where  $R - U = \{r \in R \text{ and } r \notin U\}$ . Then,  $|C| \leq |R| \leq k$ . Let  $c$  be the cost (i.e., the number of covered elements) of  $C$ . Since  $c' \leq c + k$  holds,

$$OPT(I') - c' = OPT(I) + k - c' \geq OPT(I) - c.$$

Therefore, the above reduction is an  $L$ -reduction and thus Problem P1 is MAX SNP-hard.  $\square$

For Problem P2, we can show a much stronger hardness result as follows.

**Theorem 2** *There is no polynomial time algorithm for Problem P2 with approximation ratio less than  $\frac{1-\delta}{4} \log n$  for any constant  $0 < \delta < 1$  unless  $NP \subseteq DTIME(n^{\text{poly} \log(n)})$ .*

*Proof.* We prove the theorem by contradiction. Suppose that there is a polynomial time algorithm for Problem P2 with approximation ratio less than  $\frac{1-\delta}{4} \log n$  for some constant  $0 < \delta < 1$ .

The set cover problem is defined as follows: Given a family of sets  $S$  over  $U$ , find a minimum cardinality set  $C \subseteq S$  such that all elements of  $U$  are covered by  $\bigcup_{s_i \in C} s_i$ . From an instance  $I = \langle U = \{u_1, \dots, u_m\}, S = \{s_1, \dots, s_l\} \rangle$  of the set cover problem, we construct an instance  $I' = \langle G = (V, E), L \rangle$  of P2 in the following way (See Figure 2):

$$V = \{u_1, \dots, u_m, s_1, \dots, s_l, s_0\}, E = \bigcup_{j=1}^l \left( \{(s_j, s_0)\} \cup \bigcup_{u_i \in s_j} \{(u_i, s_j)\} \right), L = 1,$$

where  $s_0$  is a node not in  $S$ .

Let  $OPT(I)$  and  $OPT(I')$  be optimal solutions of  $I$  and  $I'$ , respectively. Then,  $OPT(I') = OPT(I) + 1$  holds.

Given any solution  $R \subseteq V$  of  $I'$  with cost  $c'$  (i.e., the number of selected nodes), we produce a solution  $C$  of  $I$  in polynomial time by letting  $C = (R - U - \{s_0\}) \cup \{s_j \mid u_i \in R - S - \{s_0\}, u_i \in s_j\}$ . Let  $c$  be the cost (i.e., the number of selected elements) of  $C$ . Since  $c = |C| \leq |R| = c'$  holds,

$$\frac{c}{OPT(I)} = \frac{c}{OPT(I') - 1} \leq \frac{c'}{OPT(I') - 1}.$$

For any constant  $0 < \delta < 1$ ,

$$\frac{c'}{OPT(I') - 1} \leq \frac{1}{1 - \delta} \frac{c'}{OPT(I')} < \frac{1}{4} \log n$$

holds for sufficient large  $n = m + l + 1$ . Therefore,

$$\frac{c}{OPT(I)} < \frac{1}{4} \log n.$$

This contradicts to the fact that there is no polynomial time algorithm for the set cover problem with approximation ratio less than  $\frac{1}{4} \log n$  unless  $NP \subseteq DTIME(n^{\text{polylog}(n)})$ . Thus, the theorem is proved.  $\square$

We can also show positive results on approximation ratios using a well-known greedy algorithm for the set cover [9]. For that purpose, we let  $U = V$  and  $S = \{s_v | s_v \text{ is the set of nodes covered by } v \in V\}$ , and simply apply the greedy algorithm. Then, the following propositions are directly obtained from the results on the greedy algorithm [9, 10, 11].

**Proposition 1** *P1 can be approximated within a factor of  $e/(e-1)$ .*

**Proposition 2** *P2 can be approximated within a factor of  $O(\log n)$ .*

## 4 Integer Programming Formulation

In this section, we propose methods to solve Problem P1 and P2 using integer programming. In the previous section, we showed that both Problem P1 and P2 are very hard to find optimal or approximate solutions. However, efficient algorithms such as branch-and-bound methods have been developed for *integer programming*, which is also NP-hard. Therefore, we formulate Problem P1 and P2 as integer programs, and call IP1 and IP2 respectively. In the next section, we show that IP1 and IP2 are solved in practical time through computational experiments.

Problem P1 is formulated as follows.

$$\begin{aligned} \text{(IP1) Maximize } & \sum_{i=1}^n y_i, \\ \text{Subject to } & \\ & y_i \leq \sum_{j \in S_i^L} x_j \text{ for } i = 1, \dots, n, \\ & \sum_{i=1}^n x_i \leq K, \\ & x_i = \{0, 1\}, \\ & y_i = \{0, 1\}, \end{aligned}$$

where  $S_i^L$  is the set of nodes covering  $v_i$ . Thus, for  $j \in S_i^L$ , the length of a directed path from the node  $v_i$  to  $v_j$  is less than or equal to  $L$ .  $x_i = 1$  if  $v_i$  is selected as a reporter, otherwise  $x_i = 0$ .  $y_i = 1$  if  $v_i$  is covered by some reporter, otherwise  $y_i = 0$ . IP1 maximizes the number of covered nodes using at most  $K$  reporter nodes.

Similarly, Problem P2 is formulated as follows.

$$\begin{aligned} \text{(IP2) Minimize } & \sum_{i=1}^n x_i, \\ \text{Subject to } & \\ & \sum_{j \in S_i^L} x_j \geq 1 \text{ for } i = 1, \dots, n, \\ & x_i = \{0, 1\}. \end{aligned}$$

IP2 minimizes the number of reporters such that all nodes are covered. If the parameter  $K$  of IP1 is greater than or equal to the optimal solution of IP2, the optimal solution of IP1 is always  $n$ .

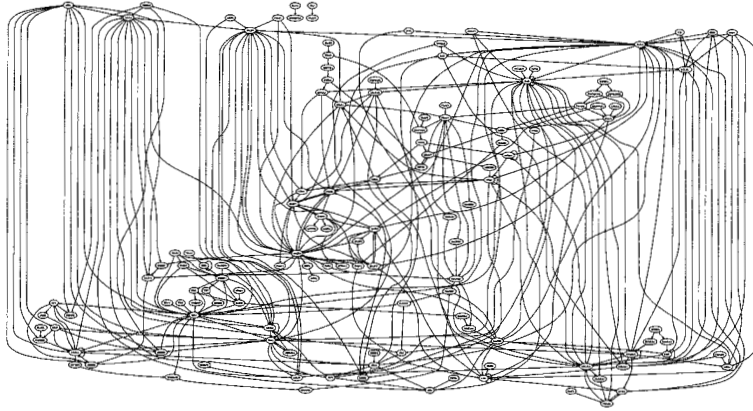


Figure 3: Apoptosis pathway maps in a HeLa cell, which contain 132 proteins and 337 binomial relations.

## 5 Computational Experiments

We applied the proposed methods to apoptosis pathway maps in a HeLa cell (See Figure 3). The maps are composed of major signal pathways of apoptosis, which are initiated by TRAIL (tumour necrosis factor apoptosis inducing ligand) ligation [12]. The maps were constructed by a commercial software, MetaCore (GeneGo Corp., <http://www.genego.com/metacore.php>), in which findings presented in peer-reviewed scientific publications were systematically encoded into an ontology by content and modelling experts, and a molecular network of direct physical, transcriptional and enzymatic interactions was computed from this knowledge base. The maps thus constructed contain 132 proteins and 337 binomial relations.

All of these computational experiments were done on a PC with a Xeon 5160 3GHz CPU and 8GB RAM running under the Linux (version 2.6.19) operating system. We used ILOG CPLEX (version 10.1, <http://www.ilog.com/products/cplex/>) for solving IP1 and IP2, and measured execution time of the optimization function CPXmipopt() for mixed integer programming problems in CPLEX. We must calculate  $S_i^L$  for all  $i$  in order to give integer programming problems to the function. However, the preparation takes at most  $O(n^2)$  time.

Table 1 shows the results on the optimal solution of IP1 and IP2 for each  $L (= 1, \dots, 6, 132)$  and  $K (= 1, \dots, 6)$ . The solution of IP2 for each  $L$  gives the required number of nodes to cover all nodes of  $V$ . For example, 42 reporters are required for  $L = 1$ , and 9 reporters for  $L = 6$ .

In the case that  $L$  is equal to the number of nodes  $n = 132$ , a node  $v_i$  is always covered by another  $v_j$  if there is a directed path from  $v_i$  to  $v_j$ . Since 121 proteins among 132 proteins are covered by protein BAK1 in the case of both  $L = 6$  and  $L = 132$ , we can see that the distance between almost all pairs of proteins in this network is at most 12. Thus, it is considered that the network also has a small-world property [13]. It should be noted that most nodes (126 nodes) are covered by 6 reporters in the case of  $L = 6$ . It is also observed that 104 nodes are covered by 6 reporters even in the case of  $L = 2$ . For  $L = 1, \dots, 3$ , TP53, BCL2 and BAX were selected as the most significant reporters respectively. These proteins are considered as hubs of the network because they have large indegrees and outdegrees. On the other hand, BAK1 is not considered as a hub, but is as an accumulation node of the network, and is selected as a reporter. Moreover, it seems that some of the selected proteins have significant biological meanings as follows. p53, a tumour suppressor gene that responds to DNA-damage, is influential on TRAIL-induced apoptosis by up-regulating TRAIL receptor [14]. Bcl-2 superfamily regulates cell death that is amplified via the mitochondrial pathway [15]. BAX may be related with possible amplification of apoptosis via

Table 1: The optimal solution of IP1 and IP2 for each  $L$  and  $K$  in apoptosis pathway maps, where the numbers of covered nodes and the numbers of the selected reporters are shown for IP1 and IP2, respectively.

$L$	IP1 for each $K$						IP2	Reporter in $K = 1$ (indegree/outdegree)
	1	2	3	4	5	6		
1	20	36	47	56	62	68	42	TP53 (19/5)
2	60	76	85	92	98	104	22	BCL2 (17/4)
3	88	103	110	116	118	120	15	BAX (16/6)
4	109	116	120	122	124	126	12	BAX (16/6)
5	118	121	123	125	127	128	10	BAK1 (6/1)
6	121	123	125	127	128	129	9	BAK1 (6/1)
132	121	123	125	127	128	129	9	BAK1 (6/1)

the intrinsic pathway in response to JNK. The caspase-9 may be essential for border-cell migration in the *Drosophila* ovary [16], and the regulation of cell migration may also point to a roll in the cleavage of several adhesion- and cell motility- related proteins during mammalian apoptosis [17].

On the other hand, our integer programming-based methods can always find optimal solutions if any. For each case, the elapsed time of optimizing IP1 or IP2 was at most 0.023 seconds. These results suggest that our methods are practical.

## 6 Concluding Remarks

We defined two problems P1 and P2 to allocate reporter genes that are effective for observing behaviors of various biological networks. We showed hardness results on approximation of these problems. On the other hand, by means of reduction to the set cover problem, we showed that P1 and P2 can be approximated within a factor of  $e/(e - 1)$  and  $O(\log n)$ , respectively.

We proposed integer programming-based methods IP1 and IP2 for solving practical instances of P1 and P2, respectively. We applied them to apoptosis pathway maps, and found that such proteins as TP53, BCL2 and BAX selected by our methods often correspond to hubs in the network. These proteins are also considered to play important biological roles. The elapsed time of optimizing IP1 or IP2 was very short. It suggests that our methods are practical.

Though we considered directed and unweighted networks in this paper, IP1 and IP2 can be modified for undirected and/or weighted networks. Furthermore, we can add various kinds of constraints to IP1 and IP2 because these are based on integer programming. Such a flexibility would be useful for modifying the proposed methods according to requirements from experimental biologists.

## Acknowledgments

We would like to thank Prof. Yuichi Sugiyama in University of Tokyo for valuable suggestions. This work is partially supported by the Cell Array Project from NEDO, Japan and by a Grant-in-Aid ‘‘Systems Genomics’’ from MEXT, Japan.

## References

- [1] S. N. Bailey, R. Z. Wu and D. M. Sabatini. Applications of transfected cell microarrays in high-throughput drug discovery. *Drug Discovery Today*, 7, S113–S118, 2002.

- [2] K. Kato, K. Umezawa, M. Miyake, J. Miyake and T. Nagamune. Transfection microarrays of nonadherent cells on an oleyl poly (ethylene glycol) ether-modified glass slide. *Biotechniques*, 37, 444–452, 2004.
- [3] T. Yoshikawa, E. Uchimura, M. Kishi, D. P. Funeriu, M. Miyake and J. Miyake. Transfection microarray of human mesenchymal stem cells and on-chip siRNA gene knockdown. *Journal of Controlled Release*, 96, 227–232, 2004.
- [4] J. Ziauddin and D. M. Sabatini. Microarray of cells expressing defined cDNAs. *Nature*, 411, 107–110, 2001.
- [5] A. K. Hadjantonakis, M. E. Dickinson, S. E. Fraser and V. E. Papaioannou. Technicolour transgenics: imaging tools for functional genomics in the mouse. *Nature Reviews Genetics*, 4, 613–625, 2003.
- [6] R. S. Stearman, M. C. Grady, P. Nana-Sinkam, M. Varella-Garcia and M. W. Geraci. Genetic and epigenetic regulation of the human prostacyclin synthase promoter in lung cancer cell lines. *Molecular Cancer Research*, 5, 295–308, 2007.
- [7] M. Golzio, L. Mazzolini, A. Ledoux, A. Paganin, M. Izard, L. Hellaudais, A. Bieth, M. J. Pillaire, C. Cazaux, J. S. Hoffmann, B. Couderc and J. Teissié. In vivo gene silencing in solid tumors by targeted electrically mediated siRNA delivery. *Gene Therapy*, 14, 752–759, 2007.
- [8] D. A. Ruths, L. Nakhleh, M. S. Iyengar, S. A. G. Reddy and P. T. Ram. Hypothesis generation in signaling networks. *Journal of Computational Biology*, 9, 1546–1557, 2006.
- [9] V. V. Vazirani. *Approximation Algorithms*. Springer, 2001.
- [10] D. S. Hochbaum. Approximation algorithms for the set covering and vertex cover problems. *SIAM Journal on Computing*, 11, 555–556, 1982.
- [11] T. Akutsu and F. Bao. Approximating minimum keys and optimal substructure screens. *Lecture Notes in Computer Science 1090 (Proc. COCOON 96)*, 290–299, 1996.
- [12] F. C. Kimberley and G. R. Screaton. Following a TRAIL: Update on a ligand and its five receptors. *Cell Research*, 14, 359–372, 2004.
- [13] D. J. Watts and S. H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393, 440–442, 1998.
- [14] G. S. Wu, T. F. Burns and E. R. McDonald III, W. Jiang, R. Meng, I. D. Krantz, G. Kao, D.-D. Gan, J.-Y. Zhou, R. Muschel, S. R. Hamilton, N. B. Spinner, S. Markowitz, G. Wu and W. S. El-Deiry. KILLER/DR5 is a DNA damage-inducible p53-regulated death receptor gene. *Nature Genetics*, 17, 141–143, 1997.
- [15] M. R. Sprick and H. Walczak. The interplay between the Bcl-2 family and death receptor-mediated apoptosis. *Biochim Biophys Acta*, 1644, 125–132, 2004.
- [16] E. R. Geisbrecht and D. J. Montell. A role for Drosophila IAP1-mediated caspase inhibition in Rac-dependent cell migration. *Cell*, 118, 111–125, 2004.
- [17] U. Fischer, R. U. Janicke and K. Schulze-Osthoff. Many cuts to ruin: a comprehensive update of caspase substrates. *Cell Death Differentiation*, 10, 76–100, 2003.