Multi-Rate Network Coding for Minimum-Cost Multicasting

Jasper Goseling
IRCTR/CWPC, WMC Group
Delft University of Technology
The Netherlands
j.goseling@tudelft.nl

Jos H. Weber
IRCTR/CWPC, WMC Group
Delft University of Technology
The Netherlands
j.h.weber@tudelft.nl

***Abstract*—We consider multicast network codes that allow to tradeoff throughput against cost. We construct a single code that enables the source to control the throughput, always achieving the minimum possible cost per transmitted symbol. Nodes in the network perform linear coding operations that are the same for all achievable throughput–cost pairs. On each of their outgoing edges nodes transmit either symbols that are obtained by these fixed linear combinations or nothing at all.**

I. Introduction

There is a natural *tension between maximizing throughput and minimizing cost*. Consider e.g. the multicasting problem in the butterfly network depicted in Figure 1, with unit capacity edges and for each edge a cost one if it is used to transmit a symbol. Figure 2 shows the canonical network code that achieves the maximum possible throughput of 2, at a cost of $4\frac{1}{2}$ per symbol. Figure 3, on the other hand, shows a solution achieving a cost 4 per symbol at a reduced throughput of 1.

It is well known that for multicasting on networks represented by directed graphs, network coding is beneficial for the throughput [1]. Also, network coding with a cost criterion has been considered [2]. Existing work, however, has focussed either on maximizing throughput or minimizing cost.

In this paper we show that a single network code can be used to *tradeoff throughput against cost*, i.e. that it can be used to operate at any achievable throughput–cost pair, subsequently referred to as an *operating point*. More precisely, we construct a code for which the source and intermediate nodes perform linear coding operations that do not have to be changed if a different operating point is used. The only difference between operating points is in which edges of the network are used. We refer to this type of code as a *multi-rate network code*. Our interest is in multi-rate codes that operate at the minimum possible cost for each of the values of the throughput that are achievable.

In [3] a network code is constructed, that, like our multi-rate code, allows to operate at different rates, while preserving local coding vectors. Cost is, however, not taken into consideration in this work. The code that is constructed uses all edges of the network, making it unsuitable for minimum-cost operation. Also, we consider a fixed set of receivers, whereas in [3] the set of receivers at a specific rate are those nodes in the network with a sufficiently large min-cut.

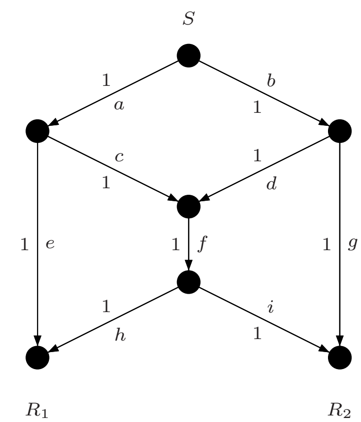

Fig. 1. Butterfly network with source S, receivers R_1 and R_2 and edges $a, b, \ldots, i$ of cost one.

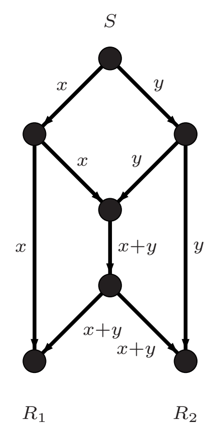

Fig. 2. Multicast of two symbols x and y, i.e. at throughput 2, at a cost per symbol of $9/2 = 4\frac{1}{2}$.

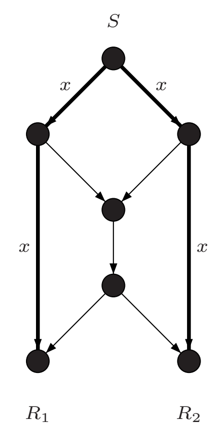

Fig. 3. Multicast of one symbol x, i.e. at throughput 1, at a cost per symbol of $4/1 = 4$.

The problem of constructing a network code for multiple operating points is, however, closely related to the problem of constructing a network code that is robust against a set of edge failure patterns, a problem considered e.g. in [4] and [5]. The similarity to our work is that a single code needs to be constructed that is valid on different subgraphs. The difference is that in [4], [5] the supported rate on all subgraphs is the same, whereas we vary the rate.

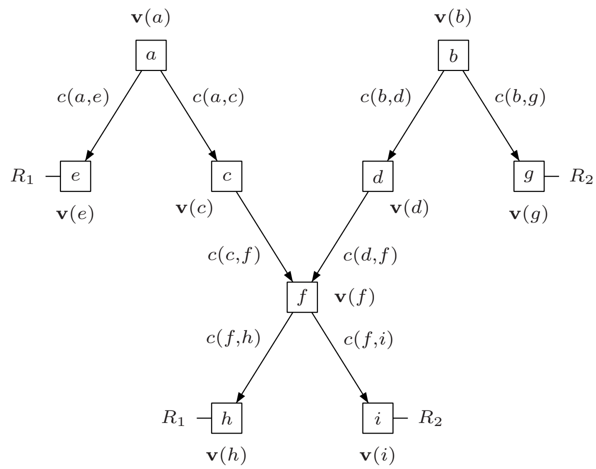

Fig. 4. Line graph of the network from Figure 1, annotated with global coding vectors $\mathbf{v}(e)$, $e \in \mathcal{E}$, and local coding coefficients $c(e, f)$, $(e, f) \in \mathcal{L}$.

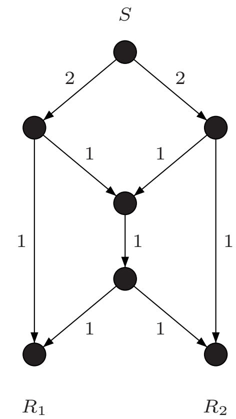

Fig. 5. The network from Figure 1 with a cost assignment on the edges such that the minimum-cost network coding solution has throughput 2.

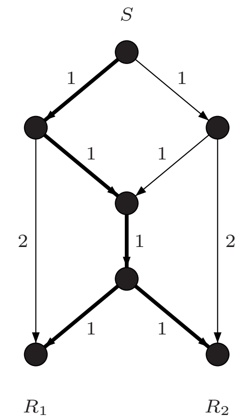

Fig. 6. The network from Figure 1 with different costs on the edges. The thick edges depict a minimum cost solution at throughput 1.

In Section II we introduce our network model and some notation. Section III briefly discusses some issues related to network coding for minimum cost. In Section IV we provide a construction for a minimum-cost multi-rate network code and show that it can be applied on all networks. In Section V we mention some interesting points for future research.

II. Model and Definitions

Our interest is in networks represented by directed acyclic graphs $G(\mathcal{V}, \mathcal{E})$, i.e. with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$, in which there is one source $S \in \mathcal{V}$ that is supposed to multicast its information to a set of receivers $\mathcal{R} \subset \mathcal{V}$. There is a non-negative cost associated with each edge. Throughout this paper we will assume that all edges have unit capacity. However, our results are also valid for networks with edges of arbitrary capacity. We assume that there is no delay on the edges and that information flows through the acyclic network in zero time.

For vertex $V \in \mathcal{V}$, let $\delta^+(V)$ and $\delta^-(V)$ be the sets of edges whose head respectively tail is V. Similarly for edge $e \in \mathcal{E}$, let $\delta^+(e)$ be the set of edges whose head is equal to e's tail, and $\delta^-(e)$ those whose tail is e's head. Let $G(\mathcal{E}, \mathcal{L})$ be the line graph of $G(\mathcal{V}, \mathcal{E})$, i.e.

$$\mathcal{L} = \left\{(e, f) | e \in \mathcal{E}, f \in \delta^-(e)\right\}. \tag{1}$$

Let $\mathbb{F}_q$ be the field of operation and $x_n \in \mathbb{F}_q$, $n = 1, \ldots, N$ the information symbols to be sent, collected in a row vector $\mathbf{x} = [x_1, \ldots, x_N]$. We call N the dimension of the code. Let $\mathbf{v}(e) \in \mathbb{F}_q^N$ be the global coding vector of edge $e \in \mathcal{E}$, i.e. the inner product $\langle \mathbf{v}(e), \mathbf{x} \rangle$ is transmitted over edge e. Associate with all $(e, f) \in \mathcal{L}$ a local coding coefficient $c(e, f) \in \mathbb{F}_q$. The coding operation performed for an edge $f \in \mathcal{E}$ is given by

$$\mathbf{v}(f) = \sum_{e \in \delta^+(f)} c(e, f)\mathbf{v}(e). \tag{2}$$

This is illustrated in Figure 4. Note that a network code will, in general, not utilize all edges of the network. In that case, only a subset of the edges carries symbols and the above definitions apply to the corresponding subgraph.

Throughput is defined as the number of information symbols that are successfully decoded by all receivers. *Cost per symbol* is defined as the sum of the costs of all edges carrying a symbol divided by the throughput.

III. Network coding at minimum cost

For minimum-cost routing it is sufficient to use only a single path to each receiver, i.e. in a network with unit capacity edges there always exists a solution with throughput one achieving minimum cost over all possible rates. If network coding is allowed this is no longer true. Consider for example the network from Figure 1, but with costs on the edges as depicted in Figure 5. It is easy to verify that each solution at throughput 1 has cost at least 6, whereas the canonical network code at throughput 2 has cost $5\frac{1}{2}$ per symbol. In fact, in this network there is no tradeoff between throughput and cost. This example also demonstrates that for networks represented by directed graphs network coding can achieve lower cost per symbol than routing only.

The above might suggest that the construction of a minimum-cost network code is more complex than finding a minimum-cost routing solution, but a minimum-cost network code can be found in polynomial time [2]. The minimum-cost multicast routing problem on the other hand corresponds to the Steiner tree problem, which is NP-complete.

The difference between the construction of an arbitrary network code and one that needs to operate at minimum cost is in the selection of the edges that are used by the code. For an arbitrary code one can select, independently for each receiver, any set of disjoint paths that meet the throughput requirements. If we consider the network in Figure 6 at throughput 1, we see that this is no longer the case for a minimum-cost network code. The edges that need to be used by the code can, however, be found in polynomial time by means of linear programming [2]. After the appropriate edges have been found, standard algorithms can be used to construct the minimum-cost network code, e.g. the algorithm presented in [5].

In the remaining part of this paper we will assume that for each operating point under consideration a minimum-cost set of edges satisfying the throughput requirements is given. Note that these different sets of edges correspond to the different subgraphs that need to be considered when constructing a robust network code [4], [5]. The main difference is that in our case, the supported rate on each subgraph is different.

IV. Multi-rate Network Coding

In this section we will construct a network code for which the source is able to control the throughput. We refer to this type of code as a *multi-rate network code*.

Let $h_1, h_2, \ldots, h_K$, $h_1 > h_2 > \cdots > h_K$ be the desired throughputs for operating points $1, \ldots, K$ respectively. As stated in Section III we assume that for each operating point a set of minimum-cost edges is given. Let $\mathcal{E}_k \subset \mathcal{E}$ be the edges for operating point $k = 1, \ldots, K$, with $\mathcal{L}_k \subset \mathcal{L}$ accordingly. Also, let $\mathcal{E}_1^K = \bigcup_{k=1}^K \mathcal{E}_k$ and $\mathcal{L}_1^K = \bigcup_{k=1}^K \mathcal{L}_k$.

A. Encoding and decoding

The dimension of the code, i.e. the number of symbols in $\mathbf{x}$, needs to be high enough to support throughput h_1, the maximum throughput among the operating points. For the other operating points, not all information symbols in $\mathbf{x}$ will have to be used. In fact, since we operate at minimum cost, the number of coded symbols observed by a receiver will be too low to decode all information symbols.

We will construct a code in which, for each operating point, the source and receivers agree on which information symbols in $\mathbf{x}$ will be used to transmit actual information. We assume that at rate h_k the first h_k symbols are used. The remaining symbols are fixed at zero by the source.

The receivers will observe a number of coded symbols for which the coding vectors span a subspace of $\mathbb{F}_q^N$. If the dimension of this subspace is smaller than N, the observed symbols and their coding vectors by themselves are not sufficient to successfully decode, i.e. to solve the system of linear equation represented by the coding vectors. Knowledge about which information symbols are not used and fixed at zero makes decoding possible, however. The receiver can immediately eliminate the corresponding variables from the system and consider a reduced system of linear equations.

B. Code construction

Our code is defined in terms of global coding vectors $\mathbf{v}^*(s)$, $s \in \left(\delta^-(S) \cap \mathcal{E}_1^K\right)$ and local coding coefficients $c^*(e,f)$, $(e,f) \in \mathcal{L}_1^K$. A node $V \in \mathcal{V}$ in the network is provided with $c^*(e,f)$, for all $e, f \in \mathcal{E}_1^K$ for which $e \in \delta^+(V)$ and $f \in \delta^-(V)$. Also, for all $e \in \delta^-(V)$ and all $k = 1, \ldots, K$, V needs to know if $e \in \mathcal{E}_k$.

We assume that all nodes in the network have knowledge of the operating point chosen by the source. Justification comes from practical considerations. Implementations of network coding that have been suggested, e.g. in [6], transmit in the header of a packet, a coding vector that is used by all symbols in the payload of the packet. We also include a description of the operating point in the header. The overhead in terms of throughput and cost of this is small.

At operating point k, an internal node $V \in \mathcal{V}$ upon receiving coded symbols, for each outgoing edge $f \in \delta^-(V)$, transmits a symbol coded according to (2) if $f \in \mathcal{E}_k$ and does not transmit anything on f otherwise. The receivers consider the (reduced) system of linear equation of dimension h_k as described in Section IV-A and successfully decode.

C. Existence of a valid code

With each of the operating points obtained by the multi-rate code presented above we can associate a traditional network code, consisting of the coding vectors used at that operating point. For $k = 1, \ldots, K$ this code, for $(e,f) \in \mathcal{L}_1^K$, has local coding coefficients

$$c_k(e,f) = \begin{cases} c^*(e,f), & \text{if } (e,f) \in \mathcal{L}_k \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The global coding vectors transmitted by the source are

$$\mathbf{v}_k(s) = \begin{cases} \mathbf{v}^*(s), & \text{if } s \in (\delta^-(S) \cap \mathcal{E}_k) \\ \mathbf{0}, & \text{otherwise,} \end{cases} \tag{4}$$

for $s \in \left(\mathcal{E}_1^K \cap \delta^-(S)\right)$. The remaining coding vectors are defined accordingly, following (2). Finally, the source utilizes the first h_k information symbols.

We need to find $\mathbf{v}^*(e)$ and $c^*(e,f)$ for which all K codes defined according to (3) and (4) are valid, i.e. the vectors $\mathbf{v}_k(e)$ allow all receivers to decode the first h_k information symbols assuming the remaining symbols are kept zero.

Theorem 1. *If $\mathbb{F}_q$ is sufficiently large, there exist valid $\mathbf{v}^*(s)$ and $c^*(e,f)$, $s \in \left(\delta^-(S) \cap \mathcal{E}_1^K\right)$, $(e,f) \in \mathcal{L}_1^K$.*

Proof: Following the algebraic framework [4], introduce for all $(e,f) \in \mathcal{L}_1^K$ an unknown over $\mathbb{F}_q$, i.e. let

$$c(e,f) = \alpha_{e,f}, \quad \forall (e,f) \in \mathcal{L}_1^K. \tag{5}$$

Also introduce unknowns for the coding vectors transmitted by the source, i.e. let

$$\mathbf{v}(s) = \left[\ldots \beta_{s,n} \ldots\right]_{n=1}^N, \quad \forall s \in \left(\delta^-(S) \cap \mathcal{E}_1^K\right). \tag{6}$$

All coding vectors in the network can be expressed in these unknowns. Since the network is acyclic, each entry of each coding vector is a bounded degree polynomial in these unknowns.

We construct matrices $\mathbf{A}_R^k$, $R \in \mathcal{R}$, $k = 1, \ldots, K$ that represent the system of linear equations that needs to be solved by receiver R at operating point k. First, let $\tilde{\mathbf{A}}_R^k$ be the concatenation of all $\mathbf{v}(e)$ for which $e \in \mathcal{E}_1^K$ is observed by

R and used at operating point k, i.e if we label the respective edges as $\{e_1, e_2, \ldots, e_{h_k}\} = \delta^+(R) \cap \mathcal{E}_k$,

$$\tilde{\mathbf{A}}_R^k = \begin{bmatrix} \mathbf{v}(e_1) \\ \mathbf{v}(e_2) \\ \vdots \\ \mathbf{v}(e_{h_k}) \end{bmatrix}. \tag{7}$$

Since at operating point k, the source only uses the first h_k information symbols, a receiver can eliminate the remaining symbols from the system by considering only the first h_k columns of $\tilde{\mathbf{A}}_R^k$. Let these be given by the $h_k \times h_k$ matrix $\hat{\mathbf{A}}_R^k$.

Also, since we need to consider the codes defined by (3) and (4), for all entries in $\hat{\mathbf{A}}_R^k$ put $\alpha_{e,f} = 0$, $\forall (e,f) \notin \mathcal{L}_k$ and $\beta_{s,n} = 0$, $\forall s \in \left(\delta^-(S) \cap \left(\mathcal{E}_1^K \setminus \mathcal{E}_k\right)\right), n = 1, \ldots, N$. Denote the new matrix by $\mathbf{A}_R^k$. Now, the codes defined by (3) and (4) are valid if all $\mathbf{A}_R^k$, $R \in \mathcal{R}$, $k = 1, \ldots, K$ are full rank matrices.

In $G(\mathcal{V}, \mathcal{E}_k)$ the source has, by assumption, min-cut h_k to each receiver. There exists therefore a routing solution, that routes information symbols $1, \ldots, h_k$ over edges in $\mathcal{E}_k$ to receiver R. This routing solution corresponds to values of the unknowns $\alpha_{e,f}$, $(e,f) \in \mathcal{L}_1^K$ and $\beta_{s,n}$, $s \in \left(\delta^-(S) \cap \mathcal{E}_1^K\right)$, $n = 1, \ldots, N$, that make the matrix $\mathbf{A}_R^k$ full rank. The determinant of $\mathbf{A}_R^k$ is, therefore, a non-zero polynomial in these unknowns. Moreover, since the degree of each entry is bounded, the degree of the determinant is bounded. Hence, it has only a finite number of zeros.

Now consider the polynomial that is the product of the determinants of all $\mathbf{A}_R^k$. It has again a finite number of zeros. If the field $\mathbb{F}_q$ is sufficiently large, there exist therefore values for $\alpha_{e,f}$, $(e,f) \in \mathcal{L}_1^K$ and $\beta_{s,n}$, $s \in \left(\delta^-(S) \cap \mathcal{E}_1^K\right)$, $n = 1, \ldots, N$, for which this product polynomial evaluates to a non-zero value. Therefore, for these values of the unknowns all individual determinants also evaluate to non-zero and all $\mathbf{A}_R^k$ have full rank. ■

D. Example

We construct a minimum-cost multi-rate code for the butterfly network given in Figure 1 for the operating points depicted in Figures 2 and 3. We have $K = 2$, $h_1 = 2$, $h_2 = 1$, $\mathcal{E}_1 = \{a, b, c, d, e, f, g, h, i\}$, $\mathcal{L}_1 = \{(a,c), (a,e), (b,d), (b,g), (c,f), (d,f), (f,h), (f,i)\}$, $\mathcal{E}_2 = \{a, b, e, g\}$, $\mathcal{L}_2 = \{(a,e), (b,g)\}$.

One can verify that the canonical network code depicted in Figure 2 does not allow receiver R_2 to decode at operating point 2. There exists, however, the following solution over $\mathbb{F}_2$. Take $c^*(e,f) = 1$, $\forall (e,f) \in \mathcal{L}_1^2$, and

$$\mathbf{v}^*(a) = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \mathbf{v}^*(b) = \begin{bmatrix} 1 & 1 \end{bmatrix}. \tag{8}$$

For operating point 1, the global coding vectors for the other nodes in the network are depicted in Figure 7. At operating point 2 the coding vectors for edges not in $\mathcal{E}_2$ are zero, the others are unaffected.

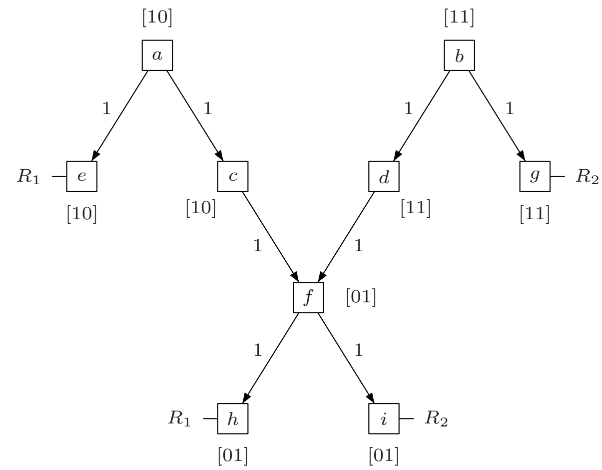

Fig. 7. Line graph of the network of Figure 1 with global and local coding vectors for a multi-rate code for the operating points depicted in Figures 2 and 3.

At operating point 1, receivers R_1 and R_2 obtain the system of linear equations represented by the matrices

$$\begin{bmatrix} \mathbf{v}(e) \\ \mathbf{v}(h) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{v}(g) \\ \mathbf{v}(i) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \tag{9}$$

respectively. Since these are full rank matrices both receivers can decode.

At operating point 2 receiver R_2 obtains only

$$\begin{bmatrix} \mathbf{v}(g) \end{bmatrix} = \begin{bmatrix} 1 & 1 \end{bmatrix}, \tag{10}$$

which corresponds to an underdetermined system, but since the source only uses the first information symbol the receiver can eliminate the second symbol and succesfully decode. Receiver R_1 obtains

$$\begin{bmatrix} \mathbf{v}(e) \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{11}$$

and can directly decode.

V. Discussion

We have shown that network coding can be used to tradeoff throughput against cost. In the code that we construct, nodes need only one set of coding vectors. Based on the operating point, nodes either perform a linear coding operation that is the same for all operating points or do not transmit anything on specified edges.

The idea to use one network code to tradeoff throughput against cost, holds many interesting questions for future research. The example from Section IV-D shows that the canonical network code on the butterfly network can not be used as the basis for a multi-rate code. There does, however, exist a multi-rate code over the field $\mathbb{F}_2$, which is the smallest field required for any code of throughput 2. It will be interesting to see if there always exists a multi-rate code over a field not larger than would be required by any of the operating points individually.

From the proof of Theorem 1 it follows that a multi-rate code can be constructed by considering an appropriate algebraic structure. This, however, requires knowledge of the complete structure of the network. One could also try to find decentralized ways to construct a multi-rate code, using e.g. ideas presented in [7].

VI. Acknowledgements

We would like to thank Christina Fragouli and Suhas Diggavi for many useful discussions.

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[2] D. S. Lun, N. Ratnakar, M. Médard, R. Koetter, D. Karger, T. Ho, E. Ahmed, and F. Zhao, "Minimum-cost multicast over coded packet networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2608–2623, 2006.

[3] S. L. Fong and R. W. Yeung, "Variable-rate linear network coding," in *IEEE Inf. Theory Workshop*, Chengdu, China, Oct. 2006, pp. 409–412.

[4] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, 2003.

[5] S. Jaggi, P. Sanders, P. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhuizen, "Polynomial time algorithms for multicast network code construction," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1973–1982, 2005.

[6] P. Chou, Y. Wu, and K. Jain, "Practical network coding," *Allerton Conference on Communication, Control, and Computing*, 2003.

[7] T. Ho, M. Médard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, 2006.