Bull. Nov. Comp. Center, Comp. Science, 6 (1997), 17–27

Iterative switching networks*

V.N. Kasyanov

In the paper the construction of the so called iterative networks is studied. This class contains rearrangeable N-inputs M-outputs networks carrying m connections with roughly $2(N+M)\log(NM/(N+M))$ contacts, if $m = \min(N, M)$ and with roughly $2(N+M)\log m$ contacts, if $m < \min(N, M)$; these results are the best obtainable by the methods used.

1. Introduction

The switching (N, M, m)-networks, where $n \leq \min\{N, M\}$, is a system for establishing the simultaneous paths from N terminals called inputs to other M terminals called outputs. The paths are established through single-pole single-throw switches called contacts. The contacts may be either "on" ("closed") or "off" ("open"). At any moment in time, at most m paths may be established simultaneously. Such a set of paths will be called a *state*. In any state of the (N, M, m)-network any input may be connected to at most one output and each output can be connected to at most one input.

Rearrangeable and nonblocking networks arise in a variety of communications contexts. Common examples include telephone systems and network architectures for parallel computers.

Rearrangeable (N, M, m)-networks can establish any set of s, $s \leq m$, connections from inputs to outputs. An additional request for connection in a state α, $|\alpha| < m$, however, may require a complete rearrangement of the state α. A request for disconnection, of course, presents no problems.

Nonblocking (N, M, m)-networks like rearrengeable (N, M, m)-networks, can establish any set of s, $s \leq m$, connections from inputs to outputs. In contrast, however, an additional request for connection can be satisfied without disturbing connections and irrespective of which state the history of connections and disconnections has left the network in.

A slightly weaker notion of nonblocking network called also a *strict-sense* nonbloking network is that of a *wide-sense* nonbloking network. A wide-sense nonbloking network does not make guarantee as a strict-sense one. A network is a wide-sense nonbloking network, if there is an algorithm for establishing path in the network, one after another, so that after each path is established, it is still possible to connect any unused input to any

*Supported in part by the Russian Committee on Higher Education.

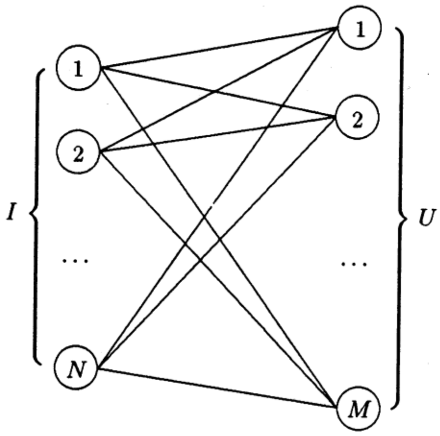

Figure 1. N-inputs M-outputs crossbar

unused output. Still weaker is the notion of rearrangeable network because a rearrangeable network is capable of realizing any m connections of inputs to outputs with node-disjoint paths provided that all the requests for connections to be made are known in advance.

A nonblocking or rearrangeable network is called a *generalized* network, if it has the additional property that each input can be simultaneously connected to an arbitrary set of outputs, provided that every output is connected to just one input.

A network may be represented as a directed graph in which nodes represent terminals and edges represent switches (Figure 1). Any desired connection of inputs and outputs corresponds to subgraph of the graph representing the network. This subgraph includes all edges (or contacts) in the on state. Input i is connected to output j, if there is a path from i to j in the subgraph just described.

Nonblocking and rearrangeable (N, N, N)-networks (or simply N-networks) have a rich and lengthy history. See [13] for an excellent survey and [7] for more comprehensive description of previous results. In 1950, Shannon [14] proved that any rearrangeable or nonblocking N-network must have $\Omega(N \log N)$ contacts. Further work on lower bounds can be found in [2]. In 1953, Clos [6] constructed a strict-sense nonblocking network with $O(N^{1+1/j})$ contacts and depth j, for fixed j. (The degree of the nodes is not bounded.) Bounded-depth nonblocking networks have subsequently been studied extensively [7–9], [12]. In the early 1960s, Beizer [3] and Beneś [4] independently discovered bounded-degree rearrangeable N-network with depth $O(\log N)$ and size $O(N \log N)$, and Waksman [15] gave an elegant algorithm for determining how the nodes should be set in order to realize any particular permutation. Ofman [10] follows with a generalized rearrangeable N-network of size $O(n \log^2 N)$. The existence of a bounded-degree strict-sense nonblocking N-network with size $O(N \log N)$ and depth $O(\log N)$ was first proved by Bassalygo and Pinsker [2].

More recent work has focuses on the construction of generalized nonblocking networks [1, 7, 13]. A generalized rearrangeable N-network with $O(N \log N)$ contacts and a generalized nonblocking N-network with $O(N \log^2 N)$ contacts were studied by Pippenger [13]. Arora, Leighton and Maggs [1] described a nonblocking N-network with $O(N \log N)$ bounded-degree nodes and an algorithm that can satisfy any request for connection or disconnection between an input and an output in $O(\log N)$ bit step, even if many requests are made at once.

For $N \neq M$ and $m = \min(N, M)$ nonblocking (N, M, m)-networks with $O(N \log^2(M+1))$ contacts were constructed by Ofman [11].

In the paper the construction of the so called iterative networks is studied. This class contains rearrangeable N-inputs M-outputs networks carrying m connections with roughly $2(N+M)\log(NM/(N+M))$ contacts, if $m = \min(N, M)$ and with roughly $2(N+M)\log m$ contacts if $m < \min(N, M)$; these results are the best obtainable by the methods used.

These networks may be more useful in the context of real multiprocessor computer systems or telephone systems, where a number of inputs can be unequal to a number of outputs and there are limitations on the number of connections which may be established simultaneously (e.g., it is unlikely that everyone on the East Coast will call someone on the West Coast at the same time).

2. Iterative networks

The networks that we use to obtain these results are constructed in such a way that is described in [8]. We refer to these networks as *iterative* networks. The nonblocking networks of Beneś [5] and Ofman [10] are similar.

Any iterative networks G is either simple or compound. We start by describing the class V^0 of simple iterative networks being nonblocking networks.

The N-inputs M-outputs crossbar (or $N \times M$-crossbar) has a separate contact for connecting each input to each output (see Figure 1).

A simple iterative (N, M, m)-network G, $G \in V^0$, is either the $N \times M$-crossbar, if $m = \min(N, M)$, or a network obtained from the $N \times m$ crossbar and the $m \times M$ crossbar by merging the output l of the first crossbar with the input l of the second one for all $1 \leq l \leq m$, if $m < \min(N, M)$ (Figure 2).

Suppose for some integer $n > 0$ an iterative network $G_{n-1} \in V^{n-1}$ be constructed as $(N_{n-1}, M_{n-1}, m_{n-1})$-network. For any integer $k_n > 0$ the compound iterative (N_n, M_n, m_n)-network $G_n \in V^n$, where $k_n = N_n/N_{n-1}$ $= M_n/M_{n-1} = m_n/m_{n-1}$, can be constructed from G_{n-1} used as an inter-

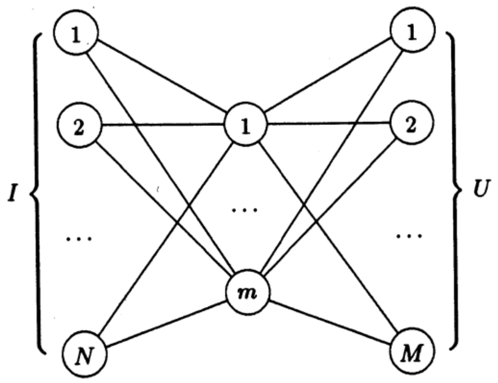

Figure 2. A simple iterative (N, M, m)-network for $m < \min(N, M)$

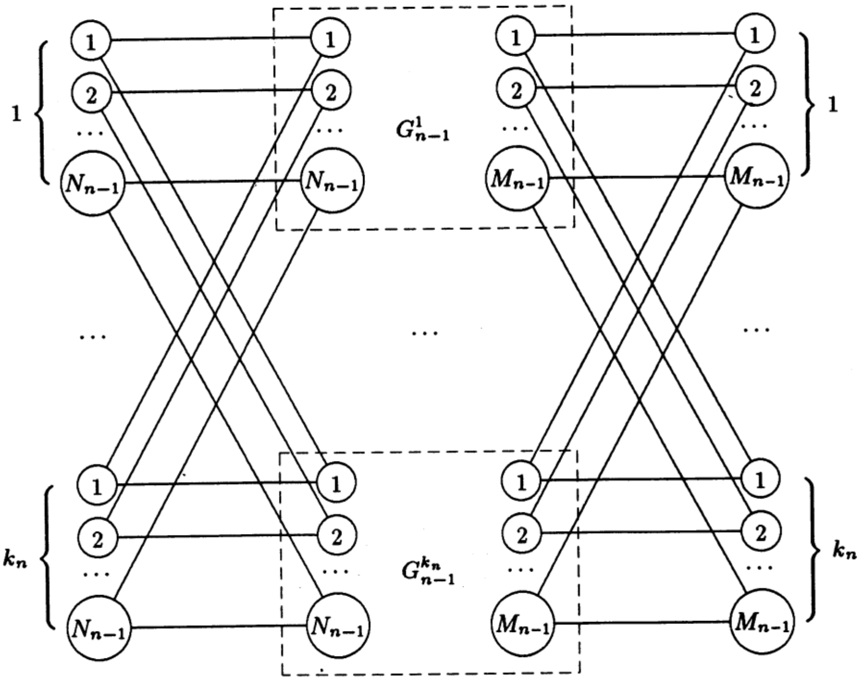

Figure 3. A compound iterative network G_n

nal subnetwork. G_n is formed by gluing together k_n copies of G_{n-1} and $N_{n-1} + M_{n-1}$ copies of $k_n \times k_n$-crossbar in the way represented in Figure 3 (here internal subnetworks are denoted by $G^1_{n-1}, \ldots, G^{k_n}_{n-1}$). The compound iterative network G_n is formed by merging the output l of the input crossbar s with input s of the internal subnetwork l for all $1 \le l \le k_n$ and $1 \le s \le N_{n-1}$ and by merging the input t of the output crossbar r with input r of the internal subnetwork t for all $1 \le t \le k_n$ and $1 \le r \le M_{n-1}$.

We shall say that G_n has the construction coefficients $k_n, k_{n-1}, \ldots, k_1$ where $k_{n-1}, \ldots, k_1$ are the construction coefficients of subnetwork G_{n-1}.

Let L_n denote the size of G_n (or number of switches of G_n). It is easy to see that G_n consists of $L_n = (N+M)\sum_{i=1}^{n} k_i + NM(\prod_{i=1}^{n} k_i)^{-1}$ switches, if $m = \min(N, M)$, and G_n consists of $L_n = (N+M)(\sum_{i=1}^{n} k_i + m(\prod_{i=1}^{n} k_i)^{-1})$ switches if $m < \min(N, M)$.

Theorem 1.

$$\min_{k_1,\ldots,k_n} L_n = \begin{cases} (n+1)(N+M)^{n/(n+1)}(NM)^{1/(n+1)}, & \textit{if } m = \min(N,M), \\ (n+1)(N+M)m^{1/(n+1)}, & \textit{if } m < \min(N,M), \end{cases}$$

and reached when

$$k_1 = \ldots = k_n = \begin{cases} (NM)^{1/(n+1)}(N+M)^{-1/(n+1)}, & \textit{if } m = \min(N,M), \\ m^{1/(n+1)}, & \textit{if } m < \min(N,M). \end{cases}$$

Proof. It is sufficient to show that for any $a > 0$

$$F(a) = \min_{k_1,\ldots,k_n>0}\Big\{\sum_{i=1}^{n} k_i + a\Big(\prod_{i=1}^{n} k_i\Big)^{-1}\Big\} = (n+1)a^{1/(n+1)},$$

and reached when $k_1 = \ldots = k_n = a^{1/(n+1)}$.

Note that for any $a > 0$

$$F(a) = \min_{k_1,\ldots,k_{n+1}>0}\Big\{\sum_{i=1}^{n+1} k_i : \Big(\prod_{i=1}^{n+1} k_i\Big)^{-1} = a\Big\} = (n+1)a^{1/(n+1)},$$

and reached when $k_1 = \ldots = k_n = a^{1/(n+1)}$. But

$$1/(n+1)\Big(\sum_{i=1}^{n+1} k_i\Big) \geq \Big(\prod_{i=1}^{n+1} k_i\Big)^{1/(n+1)},$$

i.e., $F(a) \geq (n+1)a^{1/(n+1)}$.

From the other hand

$$F(a) \leq \sum_{i=1}^{n+1} a^{1/(n+1)} = (n+1)a^{1/(n+1)}. \qquad \square$$

Corollary 1. *If $m = \min(N, M)$, then*

$$\min_{k_1,\ldots,k_n>0} L_{n=\log(NM/(N+M))-1} = 2(N+M)\log(NM/(N+M)).$$

Corollary 2. *If $m < \min(N, M)$, then*

$$\min_{k_1,\ldots,k_n>0} L_{n=\log m-1} = 2(N+M)\log m.$$

3. Establishing connections

Below we will consider a compound iterative (N', M', m')-network G' which is constructed from a rearrangeable (N, M, m)-network G with a coefficient k, i.e., $k = N'/N = M'/M = m'/m$. Let the inputs and outputs of G' be denoted by I' and U' correspondingly. Inputs and outputs of the i-th subnetwork G^i of G', where $1 \leq i \leq k$, will be denoted by I^i and U^i correspondingly. Let $\bigcup_{1 \leq i \leq k} I^i$ and $\bigcup_{1 \leq i \leq k} U^i$ be denoted by $\tilde{I}$ and $\tilde{U}$. So, $|I'| = k|I| = |\tilde{I}|$ and $|U'| = k|U| = |\tilde{U}|$.

In order to describe the technique for reconfigurating G' the term of connection sequence will be considered.

Two nodes from $\tilde{I} \cup \tilde{U}$ will be called *equivalent*, if either they are merging with the outputs of the same input crossbar or they are merging with the inputs of the same output crossbar.

Consider a state α of G'. We define a node to be *busy* in α, if there is a path currently routing through it in α and *idle* otherwise.

Let Y be a set of all inputs and outputs of some subnetworks G^s and G^t of network G'.

A sequence of nodes from Y

$$(p_{i_1}, p_{i_2}, p_{i_3}, p_{i_4}, \ldots, p_{i_r}), \quad r \geq 1,$$

is called a *connection sequence* in α under $\{s, t\}$ with initial node p_{i_1} and terminal node p_{i_r} (denoted by $[\alpha, p_{i_1}, s, t]$), if the following properties hold:

1) p_{i_1} and p_{i_r} are idle nodes,
2) p_{i_j} is a busy node for any j, $1 < j < r$,
3) p_{i_j} and $p_{i_{j+1}}$ are equivalent nodes for any odd j, $1 \leq j < r$,
4) p_{i_j} is connected with $p_{i_{j+1}}$ for any even j, $1 < j < r$.

The connection sequence $x = [\alpha, p, s, t]$ is the *trivial* one, if its length (denote $|x|$) is equal to 2. A trivial connection sequence contains only idle nodes. Let a number of pairs of such busy nodes of a subnetworks G^l, $1 \leq l \leq k_n$, which belong to x and connected in α be denoted by $n_l[x]$.

It is easy to see that the following properties hold.

Proposition 1. *Any connection sequence* $(p_{i_1}, p_{i_2}, p_{i_3}, p_{i_4}, \ldots, p_{i_r}), r \geq 1$, *consists of different nodes, i.e., for all* k, r $p_{i_k} \neq p_{i_r}$, *if* $k \neq r$.

Proposition 2. *For any state* α *and any* s, t *two connection sequences in* α *under* $\{s, t\}$ *either contain no the same node or consist of the same set of nodes.*

Proposition 3. *For any connection sequence x and any node $p \in Y$, if p is either connected with a node $q \in x$ or equivalent to a node $q \in x$, then $p \in x$.*

Proposition 4. *For any state α and any idle nodes p and q, if eiter $p \in I^s$ and $q \in U^t$, or $p \in U^s$ and $q \in I^t$, then $q \notin [\alpha, p, s, t]$.*

Lemma 1. *For any connection sequence $x = [\alpha, p, s, t]$*

$$n_s[x] = \begin{cases} n_t[x] - 1, & \textit{if } |x| = 0 \pmod 4, \\ n_t[x], & \textit{otherwise.} \end{cases}$$

Proof. From Proposition 3 the following two properties hold:

1) the connection sequence x includes the same number of nodes of subnetworks G^s and G^t,
2) all busy nodes of the connection sequence x are decomposed onto pairs of connected nodes.

From the other hand, the connection sequence x includes only two idle nodes: the initial node of x and the terminal node of x. By this, the initial node of x belongs to G^s always, and the terminal node belongs to G^s, if and only if $|x| = 0 \pmod 4$. □

Corollary 3. *If p is an idle node of subnetworks G^s, q is an idle node of subnetwork G^t and $|[\alpha, p, s, t]| = 0 \pmod 4$, then $q \notin [\alpha, p, s, t]$.*

Lemma 2. *If in the given state α a number of connections established trough the subnetworks G^s is less than a number of connections established trough the subnetworks G^t, then there is such an idle output p of G^s that $|[\alpha, p, s, t]| = 0 \pmod 4$.*

Proof. We will find such an idle output p of G^s that $|[\alpha, p, s, t]| = 0 \pmod 4$.

Let A and B be the sets of all busy outputs of subnetworks G^s and G^t, respectively. By condition of Lemma 2 we have that $|A| < |B|$.

So, there are equivalent outputs p and q of the subnetworks G^s and G^t such that $p \notin A$ and $q \in B$. Consider connection sequence $[\alpha, p, s, t]$. If $|[\alpha, p, s, t]| = 0 \pmod 4$, then the needed output p is obtained.

Let $|[\alpha, p, s, t]| \neq 0 \pmod 4$. Then the above considerations can be applied to sets $A' = A \setminus [\alpha, p, s, t]$ and $B' = B \setminus [\alpha, p, s, t]$, since $|A'| < |B'|$ by Proposition 5. From Proposition 2 it follows that in this case also by means of a finite steps the needed output p will be obtained. □

Let α be a state of G'. A *trace* of a path $P \in \alpha$ is the element $(p_1, p_2, p_3, p_4) \in I \times \tilde{I} \times \tilde{U} \times U$ in which every $p_i \in P$ is a busy node in the state $\{P\}$. The set of traces of all paths $P \in \alpha$ will be denoted by $TRACES(\alpha)$.

A state β is called *immediately reached* from a state α, if there is such a communication sequence x in α that for all nodes p_1, p_2, p_3, p_4

$$(p_1, p_2, p_3, p_4) \in TRACES(\beta),$$

if and only if there is such $(q_1, q_2, q_3, q_4) \in TRACES(\alpha)$ that the following properties hold:

1) $p_1 = q_1$ and $p_4 = q_4$,
2) either $p_2 \notin x$, $q_2 \notin x$, $p_2 = q_2$, or $p_2 \in x$, $q_2 \in x$, p_2 and q_2 are equivalent nodes,
3) either $p_3 \notin x$, $q_3 \notin x$, $p_3 = q_3$, or $p_3 \in x$, $q_3 \in x$, p_3 and q_3 are equivalent nodes.

A state β is called *reached* from a state α, if there is such a finite sequence $(\gamma_1 = \alpha, \gamma_2, \ldots, \gamma_r = \beta), r \geq 1$, of states of G' that for any t, $1 < t \leq r$, the state γ_t is immediately reached from the state γ_{t-1}.

For a given subnetwork G^r a number of all paths in α currently routing through G^r will be denoted by $n_r[\alpha]$.

Lemma 3. *For any state α and any input (or output) p of a subnework G^s, $1 \leq s \leq k$, if $|[\alpha, p, s, t]| \neq 0 \pmod 4$ or $n_s[\alpha] < m$, then there is such a state β that β is reached from α by the communication sequence $[\alpha, p, s, t]|$ and*

$$n_t[\beta] = \begin{cases} n_t[\alpha] - 1, & \textit{if } |[\alpha, p, s, t]| = 0 \pmod 4, \\ n_t[\alpha], & \textit{otherwise.} \end{cases}$$

Proof. Lemma immediatly follows from the definition of iterative networks, Lemma 2 and Proposition 3. □

Theorem 2. *G' is a rearrangeable network.*

Proof. Let α be such a state of G' that $|\alpha| < km = m'$, and let $(u, v) \in I \times U$ be a pair of idle terminals, i.e., u and v are unused in α.

Let a set of all subnetworks in which the input being adjancent with p is idle in α be denoted by $N[\alpha]$. $m[\alpha]$ will denote a set of all such subnetworks G^r that $n_r[\alpha] < m$. Let a set of all subnetworks in which the output being adjacent with q is idle in α be denoted by $M[\alpha]$.

It is clear that if

$$N[\alpha] \cap M[\alpha] \cap m[\alpha] \neq \emptyset,$$

then a connection of the input p with the output q can be established without rearrangement of current connections of α.

Let $N[\alpha] \cap M[\alpha] \cap m[\alpha] = \emptyset$. Thus, one of the following four cases arises:

1) $N[\alpha]$, $M[\alpha]$ and $m[\alpha]$ are mutually disjont,
2) $N[\alpha] \cap M[\alpha] \neq \emptyset$,
3) $N[\alpha] \cap m[\alpha] \neq \emptyset$,
4) $M[\alpha] \cap m[\alpha] \neq \emptyset$.

Since case 4 is reduced to case 3 by reorientation of network archs, in order to prove the theorem it is sufficient to consider cases 1–3 and to show that for each case there is such a state β being reached from α for which

$$N[\beta] \cap M[\beta] \cap m[\beta] \neq \emptyset.$$

Case 1. For subnetworks $G^s \in m[\alpha]$ and $G^t \in M[\alpha]$ there is such an output p of G^s that $|[\alpha, p, s, t]| = 0 \pmod 4$ by Lemma 2. From Lemma 3 it follows that there is such a state β that β is reached from α by the connection sequence $[\alpha, p, s, t]$ and $G^t \notin m[\beta]$. In addition, by Corollary 3 the idle output q of G^t being equivalent to p does not belong to $[\alpha, p, s, t]$ and so $G^t \notin M[\beta]$. Thus the state β is reached from α and $G^t \in m[\beta] \cap M[\beta]$, i.e., we have case 4 for the state β being reached from α.

Case 2. From Proposition 4 follows that a consideration presented above and applied to subnetworks $G^s \in m[\alpha]$ and $G^t \in N[\alpha] \cap M[\alpha]$ can be used in order to obtain such a state β that β is reached from α and

$$N[\beta] \cap M[\beta] \cap m[\beta] \neq \emptyset.$$

Case 3. Let $G^s \in m[\alpha] \cap N[\alpha]$, $G^t \in M[\alpha]$ and p be such an idle input of G^t which is adjacent to u. By Lemma 3 there is a state β being reached from α by the connection sequence $[\alpha, p, s, t]$. Since the input q of G^t being adjacent to u is equivalent to p and so belongs to $[\alpha, p, s, t]$, we have that $G^t \in N[\beta]$. From the other hand, the idle output of G^t being adjacent to v does not belong to $[\alpha, p, s, t]$ by Proposition 4 and so $G^t \in N[\beta]$. Thus, by Lemma 3 follows that β is either a needed state, if $|[\alpha, p, s, t]| = 0 \pmod 4$, or β is such a state that $N[\beta] \cap M[\beta] \neq \emptyset$, i.e., we have case 2 for the state β being reached from α, if $|[\alpha, p, s, t]| \neq 0 \pmod 4$. □

Corollary 4. *To establish an additional connection of an idle input with an idle output in a state α, $|\alpha| < m$, it is sufficient to rearrange at most of $3m - 1$ connections of α.*

Theorem 3. *If $k > 1$ and $m > 2$, then G' is not a wide-sense nonblocking network.*

Proof. Let $I(i, j)$ denote the input i of the input crossbar j, and $U(i, j)$ denote the output i of the output crossbar j, for all $1 \leq i \leq k$ and $1 \leq j \leq m$.

Consider the following sequence of pairs input-output: $(I(1,1), U(1,1))$, $(I(2,1), U(2,1)), \ldots, (I(k-1,1), U(k-1,1))$, $(I(k,1), U(1,2))$, $(I(1,2), U(1,3))$, $(I(2,2), U(2,3)), \ldots, (I(k-1,1), U(k-1,3))$. The length of the sequence is equal to $2k-1$, i.e., it is less than $m' = k \times m$.

It is clear that there are only two following kinds of states for connection the pairs considered:

1) such a state α that there are a subnetwork which connects input 1 with output 2 and input 2 with output 3, a subnetwork which connects input 1 with output 1 and $k-2$ subnetworks in every one of which input 1 with output 1 and input 2 with output 3 are connected,
2) such a state β that there are a subnetwork which connects input 1 with output 2 and $k-1$ subnetworks in every one of which input 1 with output 1 and input 2 with output 3 are connected,

It is clear that for the first case without reconnection of the existing state α it is impossible to connect $I(k,2)$ with $U(k,1)$. For the second case without reconnection of the existing state β it is impossible to connect $I(k,2)$ with $U(2,2)$. □

References

[1] S. Arora, F.T. Leighton, and B.M. Maggs, *Path selection in nonblocking network*, SIAM J. Comput., **25**, 1996, 600–625.

[2] L.A. Bassalygo, M.S. Pinsker, *Complexity of optimal non-blocking switching systems without reconnections*, Problems of Inform. Transmission, **9**, No. 1, 1973, 84–87 (in Russian).

[3] V.E. Beizer, *The analysis and synthesis of signal switching networks*, In Proc. Symp. on Math. Theory of Automata, Brooklyn Polytechnic Institute, Brooklyn, NY, 1962, 563–576.

[4] V.E. Beneś, *Optimal rearrangeable multistage connecting networks*, Bell System Technical J., **43**, 1964, 1641–1656.

[5] V.E. Beneś, *Mathematical Theory of Connected Networks and Telephone Traffic*, Academic Press, New York, 1965.

[6] C. Clos, *A study of non-blocking switching networks*, Bell System Technical J., **32**, 1953, 406–424.

[7] P. Feldman, *Wide-sense nonblocking networks*, SIAM J. Discrete Math., **1**, 1988, 158–173.

[8] V.N. Kasyanov, *On rearrangeability of switching networks*, In: Computational mathematics and programming, Computing Center, Novosibirsk, 1975, 123–134 (in Russian).

[9] D. Nassimi and S.S. Ahni, *Parallel permutation and sorting algorithms and a new generated connection network*, J. Assoc. Comput. Math., **29**, 1982, 642–667.

[10] Y.P. Ofman, *An universal automation*, Trans. Moscow. Soc., **14**, 1965, 186–199 (in Russian).

[11] Y.P. Ofman, *On parallel computer*, Problems of Inform. Transmission, **4**, No. 3, 1968, 58–61 (in Russian).

[12] N. Pippenger, *On rearrangeable and nonblocking switching networks*, J. Comput. System. Sci., **17**, 1978, 145–162.

[13] N. Pippenger, *Telephone switching networks*, In: Proc. Symposia in Applied Math., **26**, AMS, Providence, RI, 1982, 101–133.

[14] C.E. Shannon, *Memory requirements in a telephone exchange*, Bell System Technical J., **29**, 1950, 343–349.

[15] A. Waksman, *A permutation network*, J. Assoc. Comput. Mach., **15**, 1968, 159–163.