Yang R, Yang ZL, Zheng HP. Some indices of alphabet overlap graph. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 27(4): 897–902 July 2012. DOI 10.1007/s11390-012-1261-9

Some Indices of Alphabet Overlap Graph

Rong Yang[1] (杨　荣), Zhao-Lan Yang[2] (杨兆兰), and He-Ping Zhang[1] (张和平)

[1]*School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, China*

[2]*Normal School, Gansu Lianhe University, Lanzhou 730000, China*

E-mail: {976496552, 359274655}@qq.com; zhanghp@lzu.edu.cn

Received August 26, 2011; revised March 12, 2012.

Abstract The undirected de Bruijn graph is often used as the model of communication network for its useful properties, such as short diameter, small maximum vertex degree. In this paper, we consider the alphabet overlap graph $G(k,d,s)$: the vertex set $V=\{v|v=(v_1\dots v_k);\ v_i\in\{1,2,\dots,d\},\ i=1,2,\dots,k\}$; they are distinct and two vertices $u=(u_1\dots u_k)$ and $v=(v_1\dots v_k)$ are adjacent if and only if $u_{s+i}=v_i$ or $v_{s+i}=u_i$ $(i=1,2,\dots,k-s)$. In particular, when $s=1$, $G(k,d,s)$ is just an undirected de Bruijn graph. First, we give a formula to calculate the vertex degree of $G(k,d,s)$. Then, we use the corollary of Menger's theorem to prove that the connectivity of $G(k,d,s)$ is $2d^s-2d^{2s-k}$ for $s\geqslant k/2$.

Keywords undirected de Bruijn graph, alphabet overlap graph, vertex degree, connectivity

1 Introduction

We first introduce some basic concepts about graph theorem. Let $G=(V,E)$ be a graph with vertex set V and edge set E. For any $u,\ v\in V$, v is called the neighbor of u if $uv\in E$. The set of all neighbors of u is the neighborhood of u and denoted by $N(u)$. The cardinality of $N(u)$ is its degree, denoted by $d_G(u)$. If all the vertices of G have the same degree k, then G is k-regular. Two or more paths are independent if none of them contains an inner vertex of another. G is called k-connected if $|\ V\ |>k$ and $G-X$ is also connected for any set $X\subseteq V$ with $|X|<k$. The greatest integer k such that G is k-connected is the connectivity of G, denoted by $\kappa(G)$. Other definitions not given can be found in [1].

For given integers $k\geqslant 2$ and $d\geqslant 1$, directed de Bruijn graph $dB(k,d)$ is a directed graph with d^k vertices labeled by a word of length k over a certain alphabet with cardinality d: there is an arc from a vertex v labeled by $(v_1\dots v_2)$ to a vertex w labeled by $(w_1\dots w_k)$ if and only if $v_i=w_{i-1}$ for $i=2,\dots,k$. Undirected de Bruijn graph $UB(k,d)$ is obtained from $dB(k,d)$ by replacing each arc by an undirected edge and eliminating loops and multi-edges. Alphabet overlap digraph $DG(k,d,s)$ is a directed graph with the same vertex set as that of $dB(k,d)$: $V=\{v|v=(v_1\dots v_k),\ v_i\in\{1,2,\dots,d\}\ (1\leqslant i\leqslant k)\}$; There is an arc from a vertex $u=(u_1\dots u_k)$ to $v=(v_1\dots v_k)$ if and only if $u_{s+i}=v_i$ $(i=1,2,\dots,k-s)$. Alphabet overlap graph $G(k,d,s)$ is obtained from $DG(k,d,s)$ by replacing each arc by an undirected edge and eliminating loops and multi-edges. Fig.1 gives an illustration of $G(3,2,1)$.

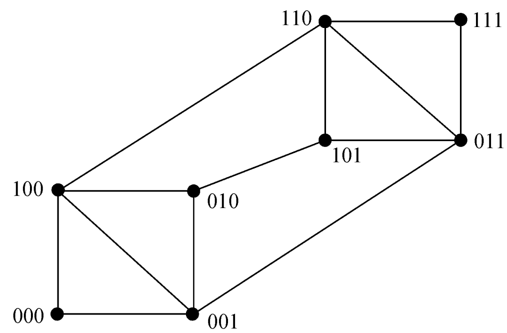


Fig.1. $G(3,\ 2,\ 1)$.

$DG(k,d,s)$ can be viewed as a generalization of $dB(k,d)$, similarly, $G(k,d,s)$ can be viewed as a generalization of $UB(k,d)$.

$G(k,d,s)$ is defined by Knisley *et al.*[2] They showed that $G(k,d,s)$ is Hamiltonian, obtained the chromatic number $\chi(G(k,d,s))=d^{k-2t}+d^t$ when $s\geqslant k/2$, and $\chi(G(k,d,s))\leqslant 1+d^{k-s}$ when $s<k/2$[2].

There are many applications for $G(k,d,s)$. For $DG(k,d,s)$, when $s=1$, it is $dB(k,d)$. It has been utilized to coding theory. When $d=4$, $dB(k,d)$ is DNA graph, which has been utilized to align protein sequences or align DNA fragments[3]. There are also many applications in communication network design theory. Designing a network that can simultaneously

Short Paper

execute k steps is equivalent to designing a k-chromatic graph. The chromatic number of the debruijn graph network was investigated in [4-5].

Knisley *et al.*[6] showed that every α-overlap graph is 3-colorable for any d if k is sufficiently large, and they determined the bound on the chromatic number of α-overlap graphs if d is much large than k.

In [7], the authors denoted $DG(k,d,s)$ by α-overlap graph. They introduced DNA graph which can be viewed as a vertex induced subgraph of $dB(k,4)$. Wang *et al.*[8] generalized the definition of DNA graph as a vertex induced subgraph of alphabet overlap digraph $DG(k,4,s)$. Recently, Li *et al.*[10] considered whether $DG(k,d,i)$ can be embedded in $DG(k,d,j)$ for given integers $1 \leqslant i < j < k$.

In this paper, we investigate some parameters of $G(k,d,s)$. In Section 2, we give a formula to calculate the the vertex degree of $G(k,d,s)$. In Section 3, we deal with the connectivity of $G(k,d,s)$. Using the corollary of Menger's theorem[1], we show that the connectivity of $G(k,d,s)$ is $2d^s - 2d^{2s-k}$ when $s \geqslant k/2$.

For convenience, we call a sequence $v_1 \ldots v_k$ as a *string* and each element v_i as a *letter*. The *length* of a string is the number of letters in it. We shall draw the reader's attention to the cases $d = 2, d = 4, d = 20$, and $d = 26$ as providing concrete applications to binary words, DNA sequences, protein sequences, and words from the English language respectively. For any $1 \leqslant i \leqslant j \leqslant k$, $v_i \ldots v_j$ is called a *substring* of $v_1 \ldots v_k$, denoted by $v_{i\ldots j}$. Moreover, $v_{1\ldots j}$ is called *prefix* and $v_{i\ldots k}$ is called *suffix*. Therefore, there is an edge between vertices $u = (u_1 \ldots u_k)$ and $v = (v_1 \ldots v_k)$ $(u \neq v)$ if and only if the suffix of u with the length of $k-s$ is the same as the prefix of v with length $k-s$ or the suffix with length $k-s$ of v is the same as the prefix with the length of $k-s$ of u. When $d = 1$, there is only one vertex in $G(k,d,s)$; when $s = k$, $G(k,d,s)$ is complete graph; when $s = 0$, it means that two vertices are adjacent if and only if they are completely the same. For any two vertices of $G(k,d,s)$ are different, so it is null. In general, we assume that $d \geqslant 2$, $k \geqslant 2$ and $1 \leqslant s \leqslant k-1$.

2 Vertex Degree $d_G(v)$

In this section, we study the vertex degree of $G(k,d,s)$. Before doing this, let us present some useful definitions.

Let $v = (v_1 \ldots v_k)$ be a vertex of $G(k,d,s)$ and p is a non-negative integer with $p < k$. We call p a *period* of v if for any $1 \leqslant i \leqslant k-p$, $v_i = v_{p+i}$[10]. It is easy to see that if p is a period of v, then any multiple of p smaller than k is also a period of v. For example, 4 is a period of vertex $v = (abcdabcdabcd \ldots abcdab)$, 8 and 12 are also periods of v.

Let $u = (u_1 \ldots u_k)$ and $v = (v_1 \ldots v_k)$ be two vertices of $G(k,d,s)$ (possibly $u = v$). u is called a *right neighbor* of v if $(u_{s+1} \ldots u_k) = (v_1 \ldots v_{k-s})$ and a *left neighbor* of v if $(v_{s+1} \ldots v_k) = (u_1 \ldots u_{k-s})$[11]. The sets of all right neighbors and all left neighbors of v are denoted by $RN(v)$ and $LN(v)$, respectively. Note that, it is possible that $v \in RN(v)$ or $v \in LN(v)$. Clearly, the neighborhood of v is $N_G(v) = RN(v) \bigcup LN(v) - \{v\}$.

For convenience, we use symbol t to replace $k-s$ and divide a vertex $u = (u_1 \ldots u_k)$ into several sections. For example, when $1 \leqslant t \leqslant \frac{k}{2}$, we denote u by abc, where a and c are strings with the same length of t and b is a string with the length of $k-2t$. It is possible that b is an empty string if $k = 2t$ $(s = t = k/2)$. When $\frac{k}{2} < t < k$, we denote u by $(u_1 \ldots u_s u_{s+1} \ldots u_t u_{t+1} \ldots u_k)$. We use x^* to denote an arbitrary string.

Theorem 1. *For a vertex v of $G(k,d,s)$, $d_G(v)$ is given as follows*:

1) *If $1 \leqslant t \leqslant \frac{k}{2}$, let $v = abc$. Then we have*

$$d_G(v) = \begin{cases} 2d^s - d^{2s-k} - 1, & \text{if } a = c, \\ 2d^s - d^{2s-k}, & \text{otherwise.} \end{cases}$$

2) *If $\frac{k}{2} < t < k$, then*

$$d_G(v) = \begin{cases} 2d^s - 2, & \text{if } s \text{ is a period of } v, \\ 2d^s - 1, & \text{if } 2s \text{ is a period of } v, \\ & s \text{ is not}, \\ 2d^s, & \text{otherwise.} \end{cases}$$

Proof. 1) If $1 \leqslant t \leqslant \frac{k}{2}$, let $v = abc$. Then

$LN(v) = \{(cy^*) |\ y^* \text{ has the length of } s\}$,

$RN(v) = \{(x^*a) |\ x^* \text{ has the length of } s\}$,

$RN(v) \cap LN(v) = \{(cz^*a) |\ z^* \text{ with the length of } k-2t\}$.

Clearly, we have that $|RN(v)| = |LN(v)| = d^s$ and $|RN(v) \cap LN(v)| = d^{k-2t} = d^{2s-k}$. Then $|RN(v) \cup LN(v)| = |RN(v)| + |LN(v)| - |RN(v) \cap LN(v)| = 2d^s - d^{2s-k}$.

If $a = c$, then $v \in RN(v) \cup LN(v)$. So $d_G(v) = |N_G(v)| = |RN(v) \cup LN(v)| - 1 = 2d^s - d^{2s-k} - 1$. Otherwise, $v \notin RN(v) \cup LN(v)$, and $d_G(v) = |N_G(v)| = |RN(v) \cup LN(v)| = 2d^s - d^{2s-k}$.

2) If $\frac{k}{2} < t < k$, let $v = (v_1 \ldots v_s v_{s+1} \ldots v_t v_{t+1} \ldots v_k)$. Then

$RN(v) = \{(x^* v_1 \ldots v_t) |\ x^* \text{ has the length of } s\}$,

$LN(v) = \{(v_{s+1} \ldots v_k y^*) |\ y^* \text{ has the length of } s\}$. We have that $|RN(v)| = |LN(v)| = d^s$.

In the following, we will consider the intersection of $RN(v)$ and $LN(v)$. We can see, if u is an element of $RN(v) \cap LN(v)$, then the t-prefix of u is $v_{s+1} \ldots v_k$ and the t-suffix of u is $v_1 \ldots v_t$. Since $2t > k$, $RN(v) \cap$

$LN(v) \neq \varnothing$ if and only if $(v_{2s+1} \ldots v_k) = (v_1 \ldots v_{k-2s})$, i.e., $2s$ is a period of v. Hence, we classify the problem into the following three cases.

Case 1. s is a period of v.

Since $2s < k$, $2s$ is also a period of v. The following two equations hold

$$(v_{2s+1} \ldots v_k) = (v_1 \ldots v_{k-2s}), \quad (1)$$

$$(v_{s+1} \ldots v_{2s} v_{2s+1} \ldots v_k) = (v_1 \ldots v_s v_{s+1} \ldots v_t). \quad (2)$$

The two equations imply that:

$$\begin{aligned} & RN(v) \cap LN(v) \\ &= \{(v_{s+1} \ldots v_{2s} v_{2s+1} \ldots v_k v_{k-2s+1} \ldots v_t)\} \\ &= \{(v_{s+1} \ldots v_{2s} v_1 \ldots v_{k-2s} v_{k-2s+1} \ldots v_t)\} \\ &= \{(v_1 \ldots v_s v_{s+1} \ldots v_{k-s} v_{k-s+1} \ldots v_k)\} \\ &= \{(v_1 \ldots v_s v_{s+1} \ldots v_t v_{t+1} \ldots v_k)\} = \{v\}. \end{aligned}$$

Hence,

$$\begin{aligned} d_G(v) &= |RN(v) \cup LN(v) - \{v\}| \\ &= |RN(v)| + |LN(v)| - 2 = 2d^s - 2. \end{aligned}$$

Case 2. $2s$ is a period of v, but s is not.

Then (1) still holds, but (2) does not. Neither $RN(v)$ nor $LN(v)$ contains v. This implies that:

$$\begin{aligned} & RN(v) \cap LN(v) \\ &= \{(v_{s+1} \ldots v_{2s} v_{2s+1} \ldots v_k v_{k-2s+1} \ldots v_t)\} \neq \{v\}. \end{aligned}$$

Hence,

$$d_G(v) = |RN(v)| + |LN(v)| - 1 = 2d^s - 1.$$

Case 3. $2s$ is not a period of v.

Then both (1) and (2) do not hold. This implies that $RN(v) \cap LN(v) = \varnothing$. Hence, $d_G(v) = |RN(v)| + |LN(v)| = 2d^s$. □

3 Connectivity of $G(k, d, s)$

We mainly use the corollary of Menger's theorem[1] to get the the connectivity of $G(k, d, s)$ $(s \geqslant k/2)$. The following is the corollary.

Theorem 2[1]. *If $ab \notin E$, then the minimum number of vertices $\neq a, b$ separating a from b in G is equal to the maximum number of independent a-b paths in G.*

Let us define $G_0(V, E)$ as a subgraph of $G(k, d, s)$ $(s \geqslant \frac{k}{2})$ induced by the vertex set $V = \{a^* v^* b^* \in V(G(k, d, s)) | a^* \neq b^*\}$. About the graph $G_0(V, E)$, we have the following lemmas:

Lemma 1. *$G_0(V, E)$ is $(2d^s - 3d^{2s-k})$-regular.*

Proof. Let $v = abc$ be any vertex of $G_0(V, E)$. We define two sets: $A = \{cv^*c | v^*$ has the length of $k - 2t\}$ and $B = \{aw^*a | w^*$ has the length of $k - 2t\}$. Clearly $|A| = |B| = d^{2s-k}$. Furthermore, by the definition, we have $N_G(v) = N_{G_0}(v) \cup A \cup B$. Since $a \neq c$, then $N_{G_0}(v) \cap A = \varnothing$, $N_{G_0}(v) \cap B = \varnothing$, and $A \cap B = \varnothing$. Therefore

$$|N_G(v)| = |N_{G_0}(v)| + |A| + |B|.$$

According to Theorem 1, we have

$$2d^s - d^{2s-k} = |N_{G_0}(v)| + 2d^{2s-k}.$$

Hence

$$d_{G_0}(v) = |N_{G_0}(v)| = 2d^s - 3d^{2s-k}. \qquad \square$$

Basing on this, we have $\kappa(G_0(V, E)) \leqslant \delta(G_0) = 2d^s - 3d^{2s-k}$. $\delta(G)$ is the minimum degree of G_0.

Lemma 2. *The connectivity of $G_0(V, E)$ is $2d^s - 3d^{2s-k}$.*

Proof. Since $G_0(V, E)$ is not a complete graph and a separating set must separate two nonadjacent vertices, so we just think about two nonadjacent vertices. In addition, according to the corollary of Menger's theorem[1], the minimum number of vertices separating two nonadjacent vertices x, y is equal to the maximum number of independent x-y paths. Now, we will show that we can find at least $2d^s - 3d^{2s-k}$ independent paths for any two nonadjacent vertices x, y. It means that $\kappa(G_0(V, E)) \geqslant 2d^s - 3d^{2s-k}$.

Let $x = avb$, $y = mwn$ be two nonadjacent vertices of $G_0(V, E)$. Then from the definitions of $G_0(V, E)$, we have $a \neq b$, $m \neq n$, $a \neq n$, $b \neq m$. We will distinguish the following three cases.

Case 1. $a = m, b = n$.

Then x and y have the completely same neighbor set $\{(bw^*z^*)\} \cup \{(e^*w^*a)\}$, $z^* \neq b$, $e^* \neq a$ with the length of $k - s$, and $\{(bw^*z^*)\} \cap \{(e^*f^*a)\} = bw^*a$. We have the order of the neighbor set is $2(d^s - d^{2s-k}) - d^{2s-k} = 2d^s - 3d^{2s-k}$. Therefore there are $2d^s - 3d^{2s-k}$ independent paths between x and y. Every independent x-y path is through only one inner vertex which is the common neighbor of x and y.

Case 2. $a \neq m, b = n$.

We can find three different types of independent x-y paths as follows:

Type 1: avb-bw^*z^*-mwn, where $z^* \neq b$, w^*, z^* are the arbitrary strings with the length of $k - 2t, t$ respectively. Clearly avb, bw^*z^*, mwn are pairwise distinct. The vertex bw^*z^* has $d^s - d^{k-2t}$ choices, we have obtained $d^s - d^{k-2t}$ independent paths.

Type 2: avb-mw^*a-aw^*m-mwn, where avb, mw^*a, aw^*m, mwn, and bw^*z^* are pairwise distinct. Since w^* has d^{k-2t} choices, the number of this type independent x-y paths is d^{k-2t}.

Type 3: avb-l^*w^*a-aw^*l^*-l^*w^*m-mwn, where $l^* \neq a$, m, b, and w^*, l^* are the arbitrary strings with the length of $k-2t, t$ respectively. It is sufficient to show that all of the inner vertices in these paths are pairwise distinct and different from the vertices used by the former paths.

$$l^*w^*a \neq avb, aw^*l^*, l^*w^*m, mwn, bw^*z^*, mw^*a, aw^*m,$$
$$aw^*l^* \neq avb, l^*w^*m, mwn, bw^*z^*, mw^*a, aw^*m,$$
$$l^*w^*m \neq avb, mwn, bw^*z^*, mw^*a, aw^*m.$$

The string l^*w^* has at least $d^s - 3d^{k-2t}$ choices, so there are at least $d^s - 3d^{k-2t}$ independent x-y paths.

Therefore, we have found $2d^s - 3d^{2s-k}$ independent x-y paths.

For the case of $a = m$, $b \neq n$, there exists independent paths analogous as above. So we just give three types of independent x-y paths, but no longer compare the inner vertices. the method is like Case 2.

Type 1: avb-z^*w^*a-mwn, $z^* \neq a$.

Type 2: avb-bw^*n-nw^*b-mwn.

Type 3: avb-bw^*l^*-l^*w^*b-nw^*l^*-mwn, $l^* \neq a$, n, b.

The number of them is $d^s - d^{k-2t} + d^{k-2t} + d^s - 3d^{k-2t} = 2d^s - 3d^{k-2t} = 2d^s - 3d^{2s-k}$.

Case 3. $a \neq m$, $b \neq n$.

We can find four different types of independent x-y path as follows.

Type 1: avb-bw^*m-mwn, mwn-nw^*a-avb, where w^* with the length of $k-2t$. Clearly we have

$$bw^*m \neq avb, mwn, nw^*a,$$
$$nw^*a \neq avb, mwn.$$

They are independent x-y paths. The number of them is $2d^{k-2t}$.

Type 2: avb-bw^*z^*-z^*w^*m-mwn, where $z^* \neq b$, m, and w^*, z^* are the strings with the length of $k-2t$, t respectively. According to the choice of the vertex, we have

$$bw^*z^* \neq avb, mwn, z^*w^*m, bw^*m, nw^*a,$$
$$z^*w^*m \neq avb, mwn, bw^*m, nw^*a.$$

So there are at least $d^s - 2d^{k-2t}$ choices of z^*w^*, we can receive $d^s - 2d^{k-2t}$ independent x-y paths.

Type 3: mwn-nw^*b-aw^*n-mw^*a-avb, where w^* is the arbitrary string of the length of $k-2t$. All of the inner vertices are pairwise distinct. The following is the proof.

$$nw^*b \neq avb, mwn, aw^*n, mw^*a, bw^*m, nw^*a, bw^*z^*, z^*w^*m,$$
$$aw^*n \neq avb, mwn, mw^*a, bw^*m, nw^*a, bw^*z^*, z^*w^*m,$$
$$mw^*a \neq avb, mwn, bw^*m, nw^*a, bw^*z^*, z^*w^*m.$$

The number of them is d^{k-2t}.

Type 4: mwn-nw^*l^*-l^*w^*a-avb, where $l^* \neq a, b, m, n$, and w^*, l^* are the strings with length $k-2t, t$ respectively. It is easy to see that these x-y paths are independent. The quantity of them is $d^s - 4d^{k-2t}$.

$$nw^*l^* \neq avb, mwn, l^*w^*a, nw^*b, aw^*n, mw^*a, bw^*z^*, z^*w^*m, bw^*m, nw^*a,$$
$$l^*w^*a \neq avb, mwn, nw^*b, aw^*n, mw^*a, bw^*z^*, z^*w^*m, bw^*m, nw^*a.$$

The number of the four types x-y independent path we obtained is

$$2d^{k-2t} + d^s - 2d^{k-2t} + d^{k-2t} + d^s - 4d^{k-2t} = 2d^s - 3d^{2s-k}.$$

So for any nonadjacent two vertices x, y, we always find $2d^s - 3d^{2s-k}$ independent x-y paths.

According to the corollary of Menger's theorem[1], we have $\kappa(G_0(V,E)) \geqslant 2d^s - 3d^{2s-k}$. By Lemma 1, we have $\kappa(G_0(V,E)) = 2d^s - 3d^{2s-k}$. □

Theorem 3. *If* $s \geqslant \frac{k}{2}$, *then the connectivity of* $G(k,d,s)$ *is* $2d^s - 2d^{2s-k}$.

Proof. Firstly, we show that $\kappa(G(k,d,s)) \leqslant 2d^s - 2d^{2s-k}$.

For a given string a with length t, let us denote an induced subgraph of $G(k,d,s)$ by G_a which has the vertex set

$$V_a = \{av^*a | v^* \text{ is a string with the length of } k-2t\}.$$

Then G_a is a complete graph $k_{d^{k-2t}}$. For $G(k,d,s)$, the vertex-set has the following partition

$$V(G) = \bigcup_{a^* \text{with length of } t} V_{a^*} \bigcup V(G_0).$$

For any given strings $a \neq b$, there is no edge between G_a and G_b, therefore the neighbor set of V_a in G is just the neighbor set of V_a in G_0. By the definition of G_a, all of vertices of G_a have the completely same neighbor set in G_0 which constructs a separating set of $G(k,d,s)$, denoted by $N_{G_0}(V_a)$. It is sufficient to show

$$N_{G_0}(V_a) = \{av^*z^* | z^* \neq a\} \cup \{z^*v^*a | z^* \neq a\}.$$

Clearly $|N_{G_0}(V_a)| = 2(d^s - d^{k-2t}) = 2d^s - 2d^{2s-k}$.

We have $\kappa(G(k,d,s)) \leqslant 2d^s - 2d^{2s-k}$.

Secondly, we will show that we can find at least $2d^s - 2d^{2s-k}$ independent paths for any two nonadjacent vertices x, y. It means that $\kappa(G(k,d,s)) \geqslant 2d^s - 2d^{2s-k}$.

Let $x \neq y$ be two nonadjacent vertices of $G(k,d,s)$. Then there exists three cases for x and y.

Case 1. x, $y \in G - G_0$.

We denote x, y by ava, mwm respectively, $a \neq m$. We can indicate the following x-y paths represented in three independent types.

Type 1: ava-aw^*m-mwm, ava-mw^*a-mwm, where w^* with the length of $k-2t$. Since $a \neq m$, all of the inner vertices are different. We get $2d^{k-2t}$ independent x-y paths.

Type 2: ava-aw^*z^*-z^*w^*m-mwm, where $z^* \neq a$, m, and z^*, w^* is with the length of t, $k-2t$ respectively. There are at least $d^s - 2d^{k-2t}$ choices of z^*w^*, and we can receive $d^s - 2d^{k-2t}$ independent x-y paths. The following is the proof:

$$aw^*z^* \neq ava, mwm, z^*w^*m, aw^*m, mw^*a,$$
$$z^*w^*m \neq ava, mwm, aw^*m, mw^*a.$$

Type 3: ava-z^*w^*a-mw^*z^*-mwm, where $z^* \neq a$, m, and z^*, w^* is with the length of t, $k-2t$ respectively. We know that the inner vertices are all distinct. The following is the proof.

$$z^*w^*a \neq ava, mwm, mw^*z^*, aw^*z^*, z^*w^*m, aw^*m, mw^*a,$$
$$mw^*z^* \neq ava, mwm, aw^*z^*, z^*w^*m, aw^*m, mw^*a.$$

There are $d^s - 2d^{k-2t}$ independent x-y paths.

The number of all x-y independent paths we have found in these three types is $2d^{k-2t} + d^s - 2d^{k-2t} + d^s - 2d^{k-2t} = 2d^s - 2d^{k-2t} = 2d^s - 2d^{2s-k}$.

Case 2. $x \in G - G_0$, $y \in G_0$.

We denote x, y by ava, mwn respectively, $a \neq m$, $a \neq n$. We can also find $2d^s - 2d^{k-2t}$ independent x-y paths. They are presented in four types.

Type 1: ava-aw^*m-mwn, ava-nw^*a-mwn, where w^* is with the length of $k-2t$. Clearly they are independent x-y paths with the number of $2d^{k-2t}$.

Type 2: ava-aw^*z^*-z^*w^*m-mwn, where $z^* \neq a$, m, and z^*, m^* is with the length of t, $k-2t$ respectively. Since

$$aw^*z^* \neq ava, mwn, z^*w^*m, aw^*m, nw^*a,$$
$$z^*w^*m \neq ava, mwn, aw^*m, nw^*a.$$

They are independent x-y paths. The number is $d^s - 2d^{k-2t}$.

Type 3: ava-l^*w^*a-nw^*l^*-mwn, where $l^* \neq a$, m, n, and l^*, w^* is with the length of t, $k-2t$ respectively. We have the inner vertices which are all distinct. The following are proof.

$$l^*w^*a \neq ava, mwn, nw^*l^*, aw^*z^*, z^*w^*m, aw^*m, nw^*a;$$
$$nw^*l^* \neq ava, mwn, aw^*z^*, z^*w^*m, aw^*m, nw^*a.$$

We have obtained $d^s - 3d^{k-2t}$ independent x-y paths.

Type 4: ava-mw^*a-mw^*m-mwn, where w^* is with the length of $k-2t$. The inner vertices are all distinct.

$$mw^*a \neq ava, mwn, mw^*m, l^*w^*a, nw^*l^*,$$
$$mw^*m \neq aw^*z^*, z^*w^*m, aw^*m, nw^*a.$$

Since $mw^*m \in G - G_0$, the other inner vertices belong to G_0, mw^*m are different from all other vertices. The number of x-y independent paths presented in this type is d^{k-2t}.

The number of independent x-y paths presented in the four types is

$$2d^{k-2t} + d^s - 2d^{k-2t} + d^s - 3d^{k-2t} + d^{k-2t}$$
$$= 2d^s - 2d^{k-2t} = 2d^s - 2d^{2s-k}.$$

Case 3. $x, y \in G_0$.

Let $x = avb, y = mwn$, $a \neq b$, $m \neq n$, $b \neq m$, $a \neq n$. From Lemma 1 we have that there are at least $2d^s - 3d^{k-2t}$ independent paths between x and y. The following are the other independent paths.

If $a \neq m$ and $b \neq n$, we have that

$$mwn\text{-}mw^*m\text{-}mw^*b\text{-}bw^*b\text{-}avb$$

are independent x-y paths, where w^* is with the length of $k-2t$. Furthermore, since $mw^*m, bw^*b \in G - G_0$, they are different from all of the vertices used by Case 3 of Lemma 1. For mw^*b, it is not used by Case 3 of Lemma 1 either. The number is d^{k-2t}.

Otherwise, either $a = m$ or $b = n$, then we have that avb-aw^*a-mwn, or avb-bw^*b-mwn is independent x-y path.

Because aw^*a, $bw^*b \in G - G_0$, they are different from all of the vertices used by Case 1 and Case 2 of Lemma 1. The number of the paths is d^{k-2t}.

Therefore we find out at least $2d^s - 2d^{2s-k}$ independent paths linking x and y. According to the corollary of Menger's theorem, we have $\kappa(G(k, d, s)) \geqslant 2d^s - 2d^{2s-k}$.

We have $\kappa(G(k, d, s)) = 2d^s - 2d^{2s-k}$. □

4 Conclusions

In the paper, we prove that the connectivity of $G(k, d, s)$ is $2d^s - 2d^{2s-k}$ for $s \geqslant k/2$. The connectivity of $G(k, d, s)$ when $\frac{k}{2} < t < k$ is worth studying, but it is difficult to solve with the same method in the paper. We will try to resolve it in succeeding work.

References

[1] Diestel R. Graph Theory (2nd edition), Springer-Verlag, 2000.

[2] Godbole A, Knisley D, Norwood R. Some properties of alphabet overlap graphs, http://arxiv.org/abs/math/0510094.

[3] Patwardhan R, Tang H, Kim S, Dalkilie M. An approximate de Bruijn graph approach to multiple local alignment and

mitif discovery on protein sequences. In *Proc. the 1st International Conference on Data Mining and Bioinformatics*, Sept. 2006, pp.158-169.

[4] Lin H. 1-Fair alternator designs for the de Bruijn network [Master's Thesis]. National Sun Yat-sen University, 2006.

[5] Abed S, Mokhtari Y, Ait-Mohamed O, Tahar S. NuMDG: A new tool for multiway decision graphs construction. *Journal of Computer Science and Technology,* Jan. 2011, 26(1): 139-152.

[6] Debra K, Yared N, Attila P. On the chromatic number of α-overlap graphs. *Journal of Combinatorial Mathematics and Computational Computing*, 2010, 73: 3-13.

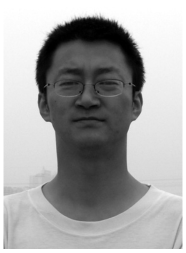
[7] Blasewicz J, Hertz A, Kobler D, de Werra D. On some properties of DNA graphs. *Discrete Appl. Math.,* 1999, 98(1/2): 1-19.

[8] Wang S, Yuan J. DNA computing of directed line-graphs. *Communications Mathematical and in Computer Chemistry*, 2006, 56(3): 479-484.

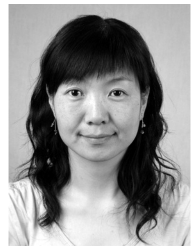
[9] Li X Y, Zhang H P. Embedding on alphabet overlap digraphs. *J. Math. Chem.,* 2010, 47(1): 62-71.

[10] Rivals E, Rahmann S. Combinatorics of periods in strings. *J. Combin. Theory, Series A,* 2003, 104(1): 95-113.

[11] Liu C H, Zhang K M. Super connectivity and restricted connectivity of undirected de Bruijn graph. *Acta Mathematicae Applicatae Sinica.*, 2002, 25(1): 29-35.



Rong Yang is a Ph.D. candidate of the School of Mathematics and Statistics, Lanzhou University. His current research interests include dynamical systems and multimedia communications in wireless networks.



Zhao-Lan Yang received the M.S. degree in operations research and cybernetics from the School of Mathematics and Statistics, Lanzhou University in 2007. And then she works in Gansu Lianhe University. Currently, her research interests include graph theory and its algorithms and application in network.

He-Ping Zhang received his Ph.D. degree of applied mathematics from Sichuan University in 1994. He was promoted as a professor in 1999. Currently, he is the dean of the School of Mathematics and Statistics of Lanzhou University, a tutor of Ph.D. candidates, the director of operations research and cybernetics, the chairman of Mathematical Association of Gansu Province, and the syndic of Combinatorial Mathematics and Graph Theory Association of China. His main research areas include perfect matching theory of graph and chemical graph theory. He has 30 papers indexed by SCI at least.