Uniwersytet Warszawski
Wydział Matematyki, Informatyki i Mechaniki

Wiktor Zuba
Nr albumu: 320501

Efektywne algorytmy generacji obiektów kombinatorycznych???

Praca magisterska
na kierunku INFORMATYKA

Praca wykonana pod kierunkiem
prof. Wojciech Rytter
Instytut Informatyki

??? 2017

Oświadczenie kierującego pracą

Potwierdzam, że niniejsza praca została przygotowana pod moim kierunkiem i kwalifikuje się do przedstawienia jej w postępowaniu o nadanie tytułu zawodowego.

Data | Podpis kierującego pracą

Oświadczenie autora (autorów) pracy

Świadom odpowiedzialności prawnej oświadczam, że niniejsza praca dyplomowa została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem tytułu zawodowego w wyższej uczelni.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Data | Podpis autora (autorów) pracy

Streszczenie

???

Słowa kluczowe

???

Dziedzina pracy (kody wg programu Socrates-Erasmus)

???

Klasyfikacja tematyczna

???

Tytuł pracy w języku angielskim

Effective algorithms of combinatorial objects generation???

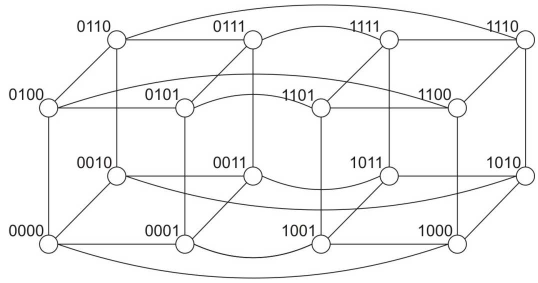
Spis treści

Wprowadzenie

Rozdział 1

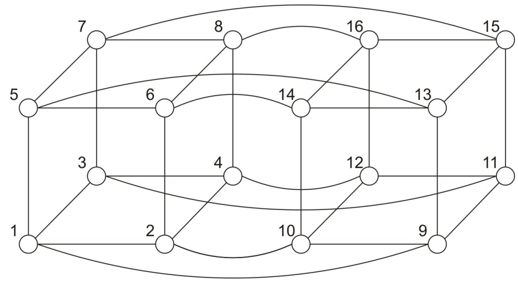
Własności hiperkostki

Definicja 1.0.1. Hiperkostką wymiaru n (Q_n) *nazwiemy graf, w którym każdy wierzchołek odpowiada ciągowi binarnemu długości n, zaś krawędzią połączone są te wierzchołki, których ciągi binarne różnią się na dokładnie jednej pozycji.*
$V(Q_n) = \{(v_0, ..., v_{n-1}) : v_i \in \{0, 1\}\}, E(Q_n) = \{(u, v) : \sum_i |u_i - v_i| = 1\}$

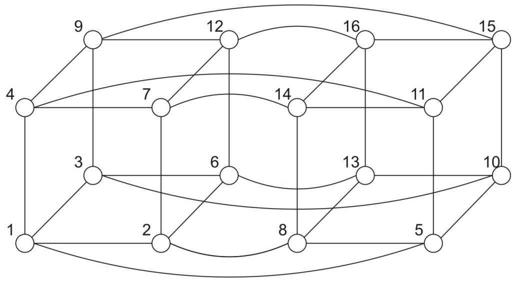


W przypadku pełnej hiperkostki bardzo łatwo jest określić długość najkrótszej ścieżki pomiedzy wierzchołkami – jest ona równa ilości pozycji na których rózni się ciągi tych wierzchołków.
Hiperkostka jest grafem dwudzielnym, w którym jedną składową jest zbiór wierzchołków o ciągach z parzystą liczbą jedynek, zaś drugą tych o ich nieparzystej liczbie.
Cowięcej przy badaniu hiperkostek często dzieli się je na $n+1$ warstw, gdzie dla $i \in \{0, 1, ..., n\}$ i-tą warstwę stanowią te wierzchołki, których ciągi binarne mają dokładnie i jedynek (warstwa zawiera zatem wierzchołki oddalone o i od wierzchołka zerowego ($\overline{0}$)).

Definicja 1.0.2. Numerowaniem klasycznym (naturalnym) *hiperkostki nazwiemy takie numerowanie* $\varphi : V(Q_n) \to \{1, ..., |V(Q_n)|\}$ *jej wierzchołków, że* $\varphi(v) = 1 + \sum_i v_i \cdot 2^i$



Definicja 1.0.3. Numerowaniem warstwowym *hiperkostki nazwiemy jej numerowanie w kolejności przeszukiwania grafu wszerz zaczynając od wierzchołka* $\overline{0}$ *z wybieraniem sąsiadów w kolejności leksykograficznej.*



Uwaga 1. *Jest to takie numerowanie $\varphi : V(Q_n) \to \{1, ..., |V(Q_n)|\}$ jej wierzchołków, że wierzchołki z i-tej warstwy otrzymują numery od $\sum_{j=0}^{i-1} \binom{n}{j} + 1$ do $\sum_{j=0}^{i} \binom{n}{j}$. W obrębie jednej warstwy numery przyznawane są przeciwnie do kolejności leksykograficznej na odwróconych słowach.* $\varphi(v) > \varphi(u) \Leftrightarrow (\sum_{i=0}^{n} v_i > \sum_{i=0}^{n} u_i) \vee ((\sum_{i=0}^{n} v_i = \sum_{i=0}^{n} u_i) \wedge (\sum_{i=0}^{n} 2^{n-i} v_i < \sum_{i=0}^{n} 2^{n-i} u_i))$

Dowód. Indukcyjnie po warstwach.
Dla warstwy 0 oczywiste.
Zakładając, że i-ta warstwa jest ponumerowana w tym porządku weźmy dwa wierzchołki u, v z warstwy $i+1$: $u = (\overline{y_1}, 1, \overline{x}), v = (\overline{y_2}, 0, \overline{x})$.
Jeśli $\overline{y_1}$ zawiera same 0, to $\overline{y_2}$ zawiera dokładnie jedną 1, sąsiedzi tych wierzchołków z poprzedniej warstwy o namniejszych numerach to odpowiednio $(\overline{0}, 0, x), (\overline{0}, 0, x)$, tak więc zostaną ponumerowane jako sąsiedzi tego samego wierzchołka, jednak u otrzyma mniejszy numer jako sąsiad mniejszy leksykograficznie.
Jeśli $\overline{y_1}$ zawiera 1, to $\overline{y_2}$ też, więc sąsiedzi tych wierzchołków z poprzedniej warstwy o namniejszych numerach to odpowiednio $(\overline{y_1'}, 1, \overline{x}), (\overline{y_2'}, 0, \overline{x})$, gdzie $\overline{y_1'}$ i $\overline{y_2'}$, to odpowiednio $\overline{y_1}$ i $\overline{y_2}$ z pierwszymi 1 zamienionymi na 0. Z założenia indukcyjnego sąsiad u ma mniejszy numer niż sąsiad v, więc u ma mniejszy numer niż v. □

1.1. Podstawy kombinatoryczne

$\binom{2n}{n} = \frac{2^{2n}\Gamma(n+\frac{1}{2})}{\sqrt{\pi}\Gamma(n+1)}$ $\quad \binom{2n+1}{n} = \binom{2n+1}{n+1} = \frac{2^{2n+1}\Gamma(n+\frac{3}{2})}{\sqrt{\pi}\Gamma(n+2)}$
$\Gamma(z) = \int_0^{\infty} x^{z-1} e^{-x} dx$ dla $n \in \mathbb{N}$ $\Gamma(n) = (n-1)!$,
ogólniej dla $x \in \mathbb{R}, x > 1$ $\frac{\Gamma(x+1)}{\Gamma(x)} = x$, $\frac{\Gamma(x+\frac{1}{2})}{\Gamma(x)} < \frac{\Gamma(x+1)}{\Gamma(x+\frac{1}{2})} \Rightarrow \sqrt{x - \frac{1}{2}} < \frac{\Gamma(x+\frac{1}{2})}{\Gamma(x)} < \sqrt{x}$

Lemat 2. *Dla* $k \leqslant \lfloor \frac{n+1}{2} \rfloor$ *zachodzi ograniczenie* $\sum_{i=0}^{k-1} \binom{n}{i} \leqslant 2^{n-1} \frac{\binom{n}{k}}{\binom{n}{\lfloor \frac{n}{2} \rfloor}}$

Dowód. (Uogólnienie dowodu z podobnego lematu dla $n = 2m, k < m$ z [1])
Załóżmy najpierw, że $k < \lfloor \frac{n}{2} \rfloor$
Zdefiniujmy $c = \frac{\binom{n}{k}}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} < 1, t = \lfloor \frac{n}{2} \rfloor - k$, $A = \sum_{i=0}^{k-1} \binom{n}{i}, B = \sum_{i=k}^{\lfloor \frac{n}{2} \rfloor - 1} \binom{n}{i}$
$\forall_{1 \leqslant i \leqslant k}$ $\frac{\binom{n}{k-i}}{\binom{n}{\lfloor \frac{n}{2} \rfloor - i}} < \frac{\binom{n}{k-i+1}}{\binom{n}{\lfloor \frac{n}{2} \rfloor - i+1}}$ $(\Leftrightarrow \frac{k-c+1}{n-k+c} < \frac{\lfloor \frac{n}{2} \rfloor - c+1}{\lceil \frac{n}{2} \rceil + c}$,
co wynika z szeregu prostych nierówności $\frac{k-c+1}{n-k+c} \leqslant \frac{k-c+1}{k+c+1} < \frac{\lfloor \frac{n}{2} \rfloor - c+1}{\lfloor \frac{n}{2} \rfloor + c+1} \leqslant \frac{\lfloor \frac{n}{2} \rfloor - c+1}{\lceil \frac{n}{2} \rceil + c})$
Daje nam to ograniczenia $\forall_{1 \leqslant i \leqslant k}$ $\frac{\binom{n}{k-i}}{\binom{n}{\lfloor \frac{n}{2} \rfloor - i}} < c$.
Suma ostatnich t wyrazów szeregu A jest majoryzowana przez $c \cdot B$, wcześniejszych t przez

c razy suma ostatnich t (a więc przez $c^2 \cdot B$). Daje nam to oszacowanie $A < (c + c^2 + c^3 + ... + c^{\lfloor \frac{k}{t} \rfloor}) \cdot B < (c + c^2 + c^3 + ...) \cdot B = \frac{c}{1-c} \cdot B$. Jednocześnie $A + B = \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor - 1} \binom{n}{i} < 2^{n-1}$.
$A = c \cdot A + (1-c) \cdot A = c(A + \frac{1-c}{c} A) < c \cdot (A + B) < c \cdot 2^{n-1}$.
Pozostaje udowodnić przypadki większych k:
Dla $n = 2m, k = m$ $\sum_{i=0}^{m-1} \binom{2m}{i} = 2^{2m-1} - \frac{1}{2}\binom{2m}{m} < 2^{2m-1} = 2^{n-1} \cdot \frac{\binom{n}{k}}{\binom{n}{\lfloor \frac{n}{2} \rfloor}}$
Dla $n = 2m+1, k = m$ $\sum_{i=0}^{m-1} \binom{2m+1}{i} = 2^{2m} - \binom{2m+1}{m} < 2^{2m} = 2^{n-1} \cdot \frac{\binom{n}{k}}{\binom{n}{\lfloor \frac{n}{2} \rfloor}}$
Dla $n = 2m+1, k = m+1$ $\sum_{i=0}^{m} \binom{2m+1}{i} = 2^{2m} = 2^{n-1} \cdot \frac{\binom{n}{k}}{\binom{n}{\lfloor \frac{n}{2} \rfloor}}$ (jedyna nieostra nierówność)

□

1.2. Własność ekspansji

Definicja 1.2.1. *Dla grafu G oraz wierzchołka $v \in V(G)$ definiujemy* sąsiedztwo wierzchołka *jako zbiór wierzchołków połączonych z nim krawędzią: $N(v) = \{u \in V(G) : (u, v) \in E(G)\}$.*

Definicja 1.2.2. *Dla grafu G oraz zbioru wierzchołków $S \subseteq V(G)$ definiujemy* sąsiedztwo zbioru wierzchołków *jako zbiór tych sąsiadów wierzchołków ze zbioru, które same do tego zbioru nie należą: $N(S) = (\bigcup_{v \in S} N(v)) \backslash S$*

Definicja 1.2.3. *Dla grafu G oraz zbioru wierzchołków $S \subseteq V(G)$ definiujemy* wnętrze zbioru wierzchołków *jako zbiór tych wierzchołków z S, których wszyscy sąsiadzi również należą do tego zbioru: $In(S) = \{v \in S : N(v) \subseteq S\}$*

Definicja 1.2.4. *Graf G posiada własność ε–*ekspansji wierzchołkowej*, jeżeli dla każdego zbioru wierzchołków $S \subseteq V(G)$ takiego, że $|S| \leqslant \frac{|V(G)|}{2}$ zachodzi $|N(S)| \geqslant \varepsilon \cdot |S|$*

Lemat 3. *Zbiór pierwszych l wierzchołków hiperkostki według numerowania warstwowego posiada maksymalne wnętrze wśród zbiorów wielkości l.*

Jest to jeden z lematów dowodzonych w pracy [2].

Lemat 4. *Dla hiperkostki do udowodnienia własności ε_n–ekspansji wierzchołkowej wystarczy rozważyć zbiory S postaci $S_k, k \leqslant 2^{n-1}$.*

Dowód. Weźmy dowolne $S \subseteq V(G), l = |S| + |N(S)|$ z Lematu 1.2 wynika, że $\frac{|N(S)|}{|S|} = \frac{|N(S)|+|S|}{|S|} - 1 \geqslant \frac{|S_l|}{|In(S_l)|} - 1 = \frac{|S_l \backslash In(S_l)|}{|In(S_l)|} \geqslant \frac{|N(In(S_l))|}{|In(S_l)|}$. Z definicji S_l wynika, że $In(S_l) = S_k$ dla $k = In(S_l)$.
Pozostaje udowodnić, że wystarczy rozważyć te S_k, że $k \leqslant 2^{n-1}$
Dla $l = |N(S)| + |S| \geqslant (\varepsilon_n + 1) \cdot 2^{n-1}$ mamy $|S| > 2^{n-1}$ lub $|N(S)| \geqslant \varepsilon_n |S|$, wystarczy więc rozważyć przypadek $l < (\varepsilon_n + 1) \cdot 2^{n-1}$.
Dla $n = 2m+1$ weźmy $k = 2^{n-1} = \sum_{i=0}^{m} \binom{2m+1}{i}$, wtedy S_k = pełne $m+1$ pierwszych warstw i $N(S_k)$ = warstwa $m+1$. Przykład ten pokazuje, że $\varepsilon_n \leqslant \frac{\binom{2m+1}{m+1}}{2^{2m}}$, więc $l < 2^{2m} + \binom{2m+1}{m+1} \Rightarrow S_l$ mieści się w piewszych $m+2$ warstwach $\Rightarrow S_k = In(S_l)$ mieści się w pierwszych $m+1$ warstwach $\Rightarrow k \leqslant 2^{2m} = 2^{n-1}$.
Dla $n = 2m$ weźmy $k = 2^{n-1} = \sum_{i=0}^{m-1} \binom{2m}{i} + \frac{1}{2}\binom{2m}{m}$, wtedy S_k = pełne m pierwszych warstw + połowa środkowej. W środkowej warstwie pierwsze $\binom{2m-1}{m-1} = \frac{1}{2}\binom{2m}{m}$ wierzchołków to dokładnie te, których ciągi binarne kończą się na 1. Wtedy też $S_k \cup N(S_k)$ to dokłanie pełne $m+1$ pierwszych warstw plus te wierzchołki z warstwy $m+2$, które kończą się na

$1 \Rightarrow |N(S_k)| = \binom{2m-1}{m} + \binom{2m-1}{m} = \binom{2m-1}{m-1} + \binom{2m-1}{m} = \binom{2m}{m}$. Przykład ten pokazuje, że $\varepsilon_n \leqslant \frac{\binom{2m}{m}}{2^{2m-1}}$, więc $l < 2^{2m-1} + \binom{2m}{m} \Rightarrow S_l$ mieści się w piewszych $m+1$ warstwach plus tych wierzchołkach z warstwy $m+2$, które kończą się na $1 \Rightarrow k \leqslant 2^{n-1}$. □

Wniosek 5. *Hiperkostka wymiaru n nie posiada własności $\frac{2\sqrt{2}}{\sqrt{\pi n}}$–ekspansji wierzchołkowej.*

Dowód. Dla $n = 2m+1$
$\frac{|N(S_{2^{2m}})|}{|S_{2^{2m}}|} = \frac{\binom{2m+1}{m+1}}{2^{2m}} = \frac{2\Gamma(m+\frac{3}{2})}{\sqrt{\pi}\Gamma(m+2)} < \frac{2}{\sqrt{\pi(m+1)}} = \frac{2}{\sqrt{\pi(\frac{n}{2}+\frac{1}{2})}} = \frac{2\sqrt{2}}{\sqrt{\pi(n+1)}} < \frac{2\sqrt{2}}{\sqrt{\pi n}}$.
Dla $n = 2m$ $\frac{|N(S_{2^{2m-1}})|}{|S_{2^{2m-1}}|} = \frac{\binom{2m}{m}}{2^{2m-1}} = \frac{2\Gamma(m+\frac{1}{2})}{\sqrt{\pi}\Gamma(m+1)} < \frac{2}{\sqrt{\pi m}} = \frac{2}{\sqrt{\pi \cdot \frac{n}{2}}} = \frac{2\sqrt{2}}{\sqrt{\pi n}}$. □

Twierdzenie 6. *Hiperkostka Q_n posiada własność $\frac{1}{\sqrt{\pi n}}$–ekspansji wierzchołkowej.*

Dowód. Jeśli $k = \sum_{i=0}^{r} \binom{n}{i}$ (pełne $r+1 \leqslant \lfloor \frac{n}{2} \rfloor + 1$ warstw), to $\frac{|N(S_k)|}{|S_k|} = \frac{\binom{n}{r+1}}{\sum_{i=0}^{r}\binom{n}{i}} > \frac{\binom{n}{r+1}\binom{n}{\lfloor \frac{n}{2} \rfloor}}{2^{n-1}\binom{n}{r+1}} = \frac{\binom{n}{\lfloor \frac{n}{2} \rfloor}}{2^{n-1}} = \frac{2^n\Gamma(\lfloor \frac{n}{2} \rfloor + \frac{1}{2})}{2^{n-1}\sqrt{\pi}\Gamma(\lfloor \frac{n}{2} \rfloor + 1)} = \frac{2\Gamma(\lfloor \frac{n}{2} \rfloor + \frac{1}{2})}{\sqrt{\pi}\Gamma(\lfloor \frac{n}{2} \rfloor + 1)} > \frac{2}{\sqrt{\pi(\lfloor \frac{n}{2} \rfloor + \frac{1}{2})}} \geqslant \frac{2\sqrt{2}}{\sqrt{\pi(n+1)}} \geqslant \frac{2}{\sqrt{\pi n}}$(dla $n \geqslant 2$).
($n = 2m+1, r = m$ rozważone w 4)
Jeśli $k = \sum_{i=0}^{r} \binom{n}{i} + \binom{n-1}{r}$ (pełne $r+1 \leqslant \lfloor \frac{n}{2} \rfloor + 1$ warstw ? TODO + te wierzchołki z warstwy $r+2$, których ciągi binarne kończą się na 1).
$|N(S_k) \cup S_k| = \sum_{i=0}^{r+1} \binom{n}{i} + \binom{n-1}{r+1} \Rightarrow |N(S_k)| = 2 \cdot \binom{n-1}{r+1}$
$\frac{|N(S_k)|}{|S_k|} = \frac{2 \cdot \binom{n-1}{r+1}}{\sum_{i=0}^{r}\binom{n}{i} + \binom{n-1}{r}} > \frac{2 \cdot \binom{n-1}{r+1}\binom{n}{\lfloor \frac{n}{2} \rfloor}}{2^{n-1} \cdot \binom{n}{r+1} + \binom{n-1}{r} \cdot \binom{n}{\lfloor \frac{n}{2} \rfloor}} = \frac{2 \cdot \binom{n-1}{r+1}\binom{n}{\lfloor \frac{n}{2} \rfloor}}{2^{n-1} \cdot (\binom{n-1}{r} + \binom{n-1}{r+1}) + \binom{n-1}{r} \cdot \binom{n}{\lfloor \frac{n}{2} \rfloor}} >$
$\frac{2 \cdot \binom{n-1}{r+1}\binom{n}{\lfloor \frac{n}{2} \rfloor}}{2^{n} \cdot \binom{n-1}{r+1} + \binom{n-1}{r} \cdot \binom{n}{\lfloor \frac{n}{2} \rfloor}} = \left(\frac{2^{n-1}}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} + \frac{\binom{n-1}{r}}{2 \cdot \binom{n-1}{r+1}} \right)^{-1} > \left(\frac{\sqrt{\pi(n+1)}}{2\sqrt{2}} + \frac{1}{2} \right)^{-1} = \frac{2\sqrt{2}}{\sqrt{\pi(n+1)} + \sqrt{2}} > \frac{2}{\sqrt{\pi n}}$(dla $n \geqslant 6$).
W pozostałych przypadkach można otrzymać ograniczenie choć dużo gorsze wiedząc, że dodanie wierzchołka do S zmniejszy $N(S)$ o conajwyżej 1.
Weźmy teraz $\sum_{i=0}^{r} \binom{n}{i} < k < \sum_{i=0}^{r} \binom{n}{i} + \binom{n-1}{r}$
$\frac{|N(S_k)|}{|S_k|} > \frac{\binom{n}{r+1} - \binom{n-1}{r}}{\sum_{i=0}^{r}\binom{n}{i} + \binom{n-1}{r}} = \frac{\binom{n-1}{r+1}}{\sum_{i=0}^{r}\binom{n}{i} + \binom{n-1}{r}} \geqslant \frac{\frac{1}{2}\binom{n}{r+1}}{\sum_{i=0}^{r}\binom{n}{i} + \frac{1}{2}\binom{n}{r+1}} > \left(\frac{\sqrt{\pi(n+1)}}{\sqrt{2}} + 1 \right)^{-1} > \frac{1}{\sqrt{\pi n}}$(dla $n \geqslant 6$).
Analogicznie da $\sum_{i=0}^{r} \binom{n}{i} + \binom{n-1}{r} < k < \sum_{i=0}^{r+1} \binom{n}{i}$
$\frac{|N(S_k)|}{|S_k|} > \frac{2\binom{n-1}{r+1} - \binom{n-1}{r+1}}{\sum_{i=0}^{r+1}\binom{n}{i}} = \frac{\binom{n-1}{r+1}}{\sum_{i=0}^{r+1}\binom{n}{i}} > \frac{\binom{n-1}{r+1}}{2^{n-1}\frac{\binom{n}{r+1}}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} + \binom{n}{r+1}} \geqslant \frac{\binom{n}{r+1}\binom{n}{\lfloor \frac{n}{2} \rfloor}}{(2^{n-1} + \binom{n}{\lfloor \frac{n}{2} \rfloor})\binom{n}{r+1}} = \frac{\binom{n}{\lfloor \frac{n}{2} \rfloor}}{2^{n-1} + \binom{n}{\lfloor \frac{n}{2} \rfloor}} >$

$\left(\frac{\sqrt{\pi(n+1)}}{\sqrt{2}} + 1 \right)^{-1} > \frac{1}{\sqrt{\pi n}}$(dla $n \geqslant 6$).
Dla przypadku $n \leqslant 5$ można ręcznie sprawdzić wszystkie 2^{n-1} przypadków, aby również otrzymać oszacowanie $\frac{1}{\sqrt{\pi n}}$. □

Bibliografia

[1] L. Lovasz, J. Pelikan and K. Vesztergombi. "Discrete Mathematics, Elementary and Beyond." *Undergraduate Texts in Mathematics. New York: Springer, first edition, 2003*

[2] L. H. HARPER, "Optimal Numberings and Isoperimetric Problems on Graphs" *JOURNAL OF COMBINATORIAL THEORY 1, 385-393 (1966)*

[3] Tomas Dvorak, Jirı Fink, Petr Gregor, Vaclav Koubek and Tomasz Radzik, "Efficient connectivity testing of hypercubic networks with faults"