Lecture 9: Small World II. chapter 6.3

(diameter)

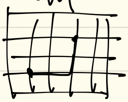
⇒ (Strogatz and Watts) model : (small world structure가 어떻게 arise / 각각 사이에 long shortcut connection ⇒ small world 구조)

This chapter (Kleinberg) — Jon Kleinberg / Robert "

= Milgram의 실험 : routing using localized limited information → short path를 determine

결론) the network contains "structure" that enables each individual could "leverage" "navigability"

(Model) : parameter (β, α)



node의 갯수는 n : $n = m^2$, 즉 $m = \sqrt{n}$

- Grid (same as SW)
- shortcut의 생성 방법 : 각 node마다 β개의 shortcut을 갖는다 → β개의 shortcut을 어떻게 만드느냐?

"이게 핵심 extension"

v에서 w로 가는 shortcut을 생성할 확률 $\propto d(v, w)^{-\alpha}$

$d(v,w) = |v-w|_{L_1}$: (v_1, v_2) (w_1, w_2) grid 상에서 distance (α는 parameter)

$$|v-w|_1 = |v_1 - w_1| + |v_2 - w_2|$$

$$\text{확률} = \frac{|v-w|^{-\alpha}}{\boxed{\sum_{u \neq v} |v-u|^{-\alpha}}}$$

normalization term

거리에 무관해짐 ← $\alpha = 0$ .... $\alpha = \infty$ → 거리를 고려해서

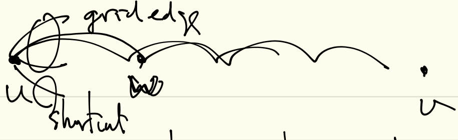
확률 $\propto \dfrac{1}{|v-w|^{\alpha}}$

$\boxed{\alpha = 0, \beta = 1}$ = Strogatz and Watts $(p = 1)$

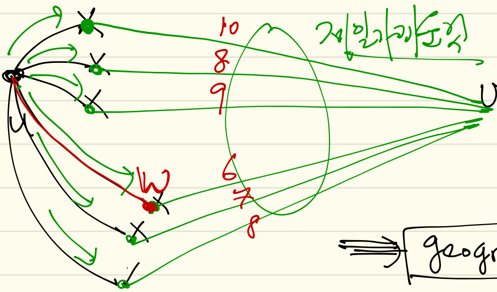
α가 크면 ⇒ 가까운 친구들을 선호

↳ 먼 거리에 있는 node에 shortcut을 덜 만듦

(결론정리)



① **Greedy routing**: A node u, trying to reach a node v forwards the message to w among its grid and shortcut edges that is closest (according to L^1 distance) to the target v.



매우 상식선의 algorithm

"greedy": ~~the distance~~ destination에 가장 가까운 node

⇒ $\boxed{\text{geographic routing}}$

(결과)

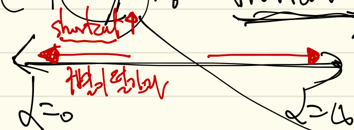
(Θ(√n) 보다 짧은 path가 존재)

(i) $\alpha=2$: shortcut이 존재하고, greedy routing이 실제로 그 path를 찾아서 routing 한다.

(ii) $\alpha<2$: shortcut은 존재하나, greedy routing으로는 short path를 찾지 못한다.

(dest)

(iii) $\alpha>2$: shortcut이 존재하지 않는다.



→ 너무 가까운 곳으로만 shortcut을 만들기 때문에

← 거리에 무관한 shortcut 생성

SW ($\alpha=0$)

shortcut이 길수록 better

너무 가까운 친구들과의 shortcut

$\alpha = 0$ — $\alpha = 2$ — $\alpha = \infty$

destination까지 지름길(X)

그래프의 구별을 덜하는 것이다

→ greedy routing 처럼 "stupid" 평범한 어떤 path가 좋은 path인지 구별X)

구별을 잘하면서도 shortcut이 길게

$\alpha = 2$

② $2 < \alpha$ ① $\alpha = 2$ $\cancel{\alpha > 2}$

6.3.2 $\alpha = 2$

Thm 6.5 Assume $\alpha = 2$. Under the greedy routing scheme, Let $T_{greedy}(u, v)$ be the number of steps used ~~the~~ to reach from u to v. The following holds:

$$\boxed{E(T_{greedy}(u, v)) = O(\log n)^2}$$

(i) 해석: 평균 hops 수: 많아 봐야 $(\log n)^2$

routing에서 필요한 개수는 적어도

(Note) 의미: $\alpha = 2$ (greedy routing이 working) → 수학적 의미대로

직관 → 어떠한 수학적 결과

$E(\text{Diameter}(\text{전체})) = O$

$$\boxed{\max_{u,v} T_{greedy}(u, v) \leq c(\log n)^2 \quad w.h.p}$$

그렇게 결과가 나왔으니까! 직관: verbal statement

→ 질문에 답: 여러 형태가 가능 (= ~~화살표~~)

교수: A가 (될 것 같은데 / 증명 가능...)

학생 추가: 해봤는데 A가 안된다.

(학생 바로: 해봤는데 A는 안되지만 그걸 보니 weak한 B는 된다)

상식적 sense

일치 ||

프로그래머의 sense

안되는 이유

정확히

weak → 새로운 화살표

(가정을 추가 / 조건 restriction)

더 버기

~~수학은 잘하는데 coding은 취약~~

~~coding은 잘하는데 수학은 싫어~~

~~O X~~

→ 엄격해진다. 자신감이 없다

자신감은 있지만 / 엄격해진다.

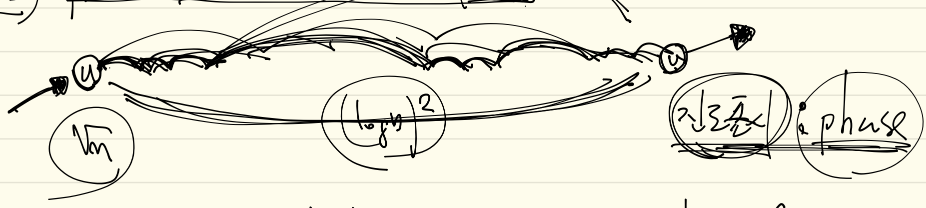
mental robust

긍정적(X), 꼼꼼하게

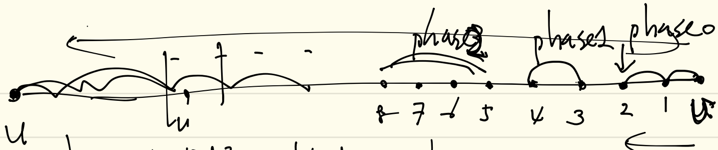
항상 모서리로 뭔가를 찾는다

긍정적

Pf) 1번째로 나갈 때면 어떻게 proof



- $u(t)$: the node to which the greedy algorithm has forwarded the message after t steps
- phase: We say that the algorithm is in phase j if at time t $2^j < |u(t) - v| \le 2^{j+1}$



phase 0
$1 < | \ | \le 2$
phase 1
$2 < \quad \le 4$

phase의 변화를 겪는 과정 : (phase가 많이 줄어도 되고
phase의 변화가 없어도 되고)

Let X_j be a random variable that represents the #. of steps spent in a given phase j

Y : random variable ~ Geometric distribution with some parameter

$$X_j \le Y \Rightarrow E(X_j) \le E(Y) \approx (\log n)^2$$

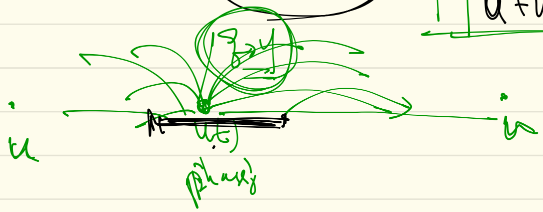
$$E(T_{greedy}(u,v)) = O(\log n)^2$$

사이 2^j의 확률 - ~~Assu~~ Assuming $u(t)$ is in phase j, the probability that a shortcut to some node, say w, leads to a phase $k < j$ is lower bounded

$$f(j) = \min_{u(t):\, 2^j < |u(t) - v| \le 2^{j+1}} \frac{\sum_{w:\, |v-w| \le 2^j} |u(t) - w|^{-2}}{\sum_{u' \ne u(t)} |u(t) - u'|^{-2}}$$

① (분자)
② (분모)

(phase j 주변에서 개선을 이룰 확률의 lower bound)

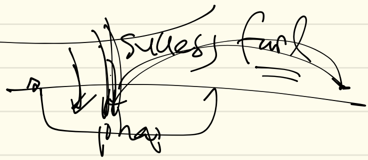


Lemma 9.1 $\quad f(j) \geq \dfrac{1}{144(1+\log 2m)}$ (page 73 의 가운데)

(Pf) ① 분자 $\geq (2^{j+1}+2^{j})^{-2} \sum_{i=1}^{2^j} i$ $\quad$ ② 분모 $\leq \sum_{i=1}^{2m} (4i) i^{-2}$

triangle inequality $\geq \dfrac{1}{36}$ $\quad\quad \leq 4(1+\log 2m)$

공리의 연습 상한식대로 재보강



(Homework) Proof of Lemma 9.1

Y_j = geometric random variable with $1-\dfrac{1}{144(1+\log 2m)}$

↳ # of steps in phase j

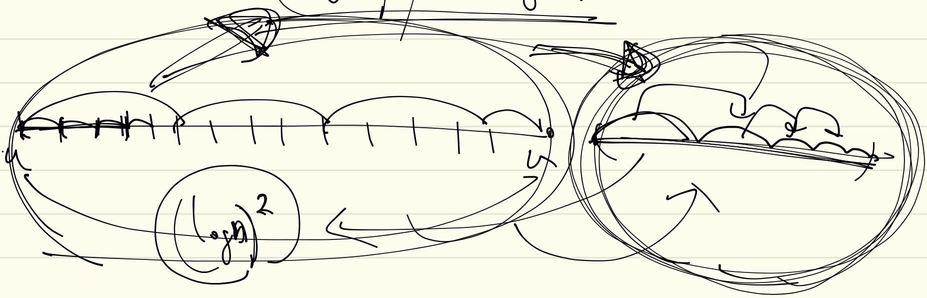
(ii) Y

$E(T_{greedy}(u,v)) \leq$ (phase마다 필요한 step수) · 총 step수

$= 144(1+\log 2m)\cdot \log_2 2m$

$= O(\log m)^2 = O(\log n)^2$

(Question)



② $\alpha<2$ (chapter 6.3.3)

$\boxed{\text{Thm 6.6}}$ Assume $0 \le \alpha < 2$. Then for any decentralized algorithm alg,

"for "most" pairs (u,v)", $E[T_{alg}(u,v)] \ge \delta \left(n^{\frac{2-\alpha}{3}} \right)$, for some constant $\delta > 0$.

polynomial 값) $(\log n)^2$

예시 i) what is decentralized algorithm?

: Routing decision made at step t depends only on knowledge of (i) the nodes $u(0), u(1), u(2), \ldots, u(t)$ visited so far and (ii) coordinates of the destinations of shortcuts generated at these nodes

좌표

⇒ Greedy algorithm belongs to this category.

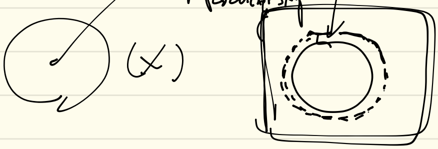
(destination의 좌표)

clever한 alg (x)

(ii) $\alpha<2$: shortcut은 많이 존재하나 "decentralized algorithm" 으로는 그 shortcut을 찾기가 못한다.

(좋은 research topic) $\alpha<2$. 어떤 algorithm이 shortcut을 enjoy 할 수 있을까? 있나?

Routing Alg



(Proof) 잔머리 발휘하여 적어도 xxx 보다 크다. $\boxed{E(X) \ge \ }$

저자의 머리 (Q) $\alpha<2$, greedy algorithm working 할까? subpolynomial

⇒ 6.3.2 의 proof를 $\alpha<2$ 일 때도 try해 봤을 것이다.

Homework: 6.3.2의 증명을 $\alpha<2$ 일 때 사용하려면 어디서 break이 날까?

(Q) Simulation 결과가 → $(\log n)^2$ (X), polynomial

↳ $E(T_{alg}(u,v)) \geq \boxed{}$ polynomial

$\boxed{a \Pr(T_{alg}(u,v) \geq a)} \leq E(T_{alg}(u,v))$ (markov inequality)

모를 경우

편하자

↳ a만 잘 choose 하면

다루기 쉽다.

어떤 step t를 찾는가 s.t $(\boxed{T_{alg}(u,v) \geq t \text{ 일 확률}}) \times t$ = polynomial

u에 출발해서 t step 안에 들어가지 못할 확률 (Q1) 확률(?)의 lower bound

(도달하지 못할)

v에

(증명에서)

a를 (parameterize 해서 tunable) → 확률 전개 → 우리가 원하는 방향으로 (tune 선택 조절)

(Q2) a는 어떻게 choose?

해결되지 않았는가?

우리는 $\Pr(T_{alg}(u,v) > t)$

$= \Pr$(u에서 출발해서 t step 안에 못 들어갈 확률) —— ①

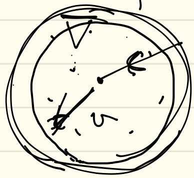
t step 안에 v에 도착하려면, t step 동안에 움직임에서 first t visited locations 들 중에서 ◯ 으로 들어가는 shortcut이 존재.

first t visited locations 중에서 ◯ 로 들어가는 shortcut이 존재하지 않으면 t step안에 v에 도달할 수 없다.

why? grid edge 만으로 가면 1 step에 거리 1씩 전진

① ≥ Pr(first t visited location 에서 ~~shortcut~~ ◯ 로 들어가는 shortcut이 존재하지 않음)

For some c, consider a set $V = \{w : |v-w| \le c\}$



u

Assume $|u-v| > c$ (질문: 만약 $|u-v| \le c$ 나중에 c를 choose) ↳ 나는 고려하지 않을거임

$t = \varepsilon c$, where $0 < \varepsilon < 1$ (나중에 ε로 우리가 choose)

Pr(t step안에 v에 도달하지 못할 확률)

$\geq$ Pr(◯로 들어가는 shortcut이 존재하지 않는 in the first t visited location)

우리의 목표: $Pr(T_{dg}(u,v) > t) \geq$ 예쁜 모양 (ε와 c를 잘 choose하여)

$\boxed{\text{Lemma 9.2}}$: V에 안으로 도달하는 shortcut의 확률 $\leq$ $6c^2 m^{2-r}$

숙제 $\boxed{\text{Homework 9.2}}$ 증명

Lemma 9.3 $\Pr(T_{alg}(u,w) > t) \geq 1 - \rho\varepsilon\left(6C^3 m^{\alpha-2} / 2^{3-\alpha}\right)$

(Proof) At each step 에서. 끝에의 shortcut이 존재. $P(A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_t) \leq \sum P(A_i)$

생각해보기

$\Pr(T_{alg}(u,v) > t) = 1 - \Pr(t \text{ step 안에 도착할 확률})$

t step 안에 v에 도착할 확률

$\geq 1 - \rho t \Pr(\text{shortcut generated from } w \text{ reaches } V)$

From Lemma 9.2 $\geq 1 - \rho\varepsilon C \cdot 6C^2 \cdot m^{\alpha-2} / 2^{3-\alpha}$

(ε와 C로 표현, 결국은 t로 표현) $t = \varepsilon C$

ε와 C를 잘 바꿔서 선택

$E(T_{alg}(u,v)) \geq t \cdot \Pr(T_{alg}(u,v) > t)$

$\geq \varepsilon C \cdot (1 - \rho\varepsilon 6C^3 m^{\alpha-2})$

$\varepsilon = \frac{1}{128}$ style 비슷하게

$C = \left(\frac{m^{(2-\alpha)}}{2^{3-\alpha}}\right)^{\frac{1}{3}}$ $\alpha < 2$

C와 ε는 잘 choose

$\varepsilon \cdot m^{\frac{2-\alpha}{3}}$ for some constant $\rho > 0$

$1 - \rho \cdot \frac{1}{128} \cdot 6 \frac{m^{2-\alpha}}{2^{3-\alpha}} \cdot m^{\alpha-2} = 1 - \frac{6}{128} \frac{6}{2^{3-\alpha}}$

$= 1 - \frac{1}{2^{4-\alpha}} < k$

내가 C^2_2 잘 선택하여 $m^{\text{some order}}$로 만든다.

결과론 · 질문과 결과론 intuition
직관

질문 : network structure와 routing의 방법 연결

Strogatz and Watts

Short small world를 나타내는 structure를 발명하기

decentralized, greedy routing

①개별

간편한 파수령

훌륭히 좋은 연구 할 수 있다.

모든 사실, 결과

무작정 concept

질문하는 순간

[redacted]

straight 한 질문

(그 간단한 궁금증들)

순간

새로운 문제를 define 할 수 있는 질문

질문

세밀화

순환

이론

상호관계