Opdracht 2: Het verspreide virus

1 Opdracht

Wiske plaatst een aantal computers in een netwerk zodat twee verbonden computers hun opgeslagen informatie makkelijk met elkaar kunnen delen. Wanneer twee computers met elkaar verbinden, wordt telkens hun volledige informatie met elkaar uitgewisseld. Krimson heeft echter een virus ontwikkeld en is kunnen binnendringen in het netwerk van Wiske. Het virus wordt slechts actief wanneer elke computer in het netwerk alle informatie van alle computers kent. Krimson heeft controle over het netwerk en kan naar hartelust computers met elkaar in verbinding stellen. Hij kan echter telkens maar twee computers per keer met elkaar verbinden. Hoe kan hij er met een minimaal aantal connecties voor zorgen dat het virus actief wordt?

1. Bepaal dit minimum voor een netwerk met 4 computers en bewijs dat dit het minimum is.

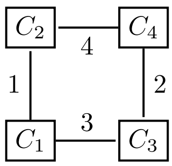
2. Bepaal dit minimum voor een netwerk met 5 computers en bewijs dat dit het minimum is.

3. Bepaal dit minimum voor een netwerk met 10 computers en bewijs dat dit het minimum is.

2 Oplossing

2.1 $n = 4$

Voor $n = 4$ kunnen we onderstaande constructie tekenen om te zien dat 4 verbindingen voldoende zijn. Het is ook duidelijk dat er minstens 3 verbindingen nodig zijn om alle computers met elkaar te verbinden. Wanneer we echter de mogelijkheden met 3 verbindingen overlopen, zien we dat hier niet alle computers alle informatie verkrijgen. Dit bewijst dat 4 verbindingen ook het minimum is.



2.2 $n \geq 5$

Het is makkelijk in te zien dat voor $n \geq 5$ computers, in totaal $2n - 4$ verbindingen voldoende zijn om het virus te activeren. We duiden 4 computers aan als de hoofdcomputers. Eerst verbinden we elk van de overgebleven $n - 4$ computers met een van de hoofdcomputers. Daarna laten we in 4 verbindingen alle info verspreiden onder deze hoofdcomputers. Tenslotte koppelen we alle informatie terug naar de $n - 4$ andere computers. In totaal zijn dit dus $2(n - 4) + 4 = 2n - 4$ verbindingen.

Aantonen dat dit aantal ook effectief het kleinste aantal verbindingen is waarna het virus actief kan worden, is echter moeilijker om algemeen aan te tonen. Daarom bekijken we dit geval per geval.

We bewijzen eerst een algemeen Lemma.

Lemma. *Stel dat er voor m computers een constructie bestaat met k verbindingen die het virus activeren en dat er bovendien een computer bestaat die na een aantal stappen zijn eigen informatie terug ontvangt. Dan bestaat er ook voor $m - 1$ computers een constructie met $k - 2$ verbindingen die het virus activeert.*

Proof. Stel dat computer C_1 zijn eigen info terug onvangt, dan heeft deze informatie een weg als volgt afgelegd:

$$C_1 \to C_2 \to ... \to C_i \to C_1.$$

In onze constructie kunnen we nu C_1 weglaten samen met de 2 verbindingen $C_1 \leftrightarrow C_2$ en $C_1 \leftrightarrow C_i$. Alle andere verbindingen met C_1 herleggen we naar C_2 of C_i afhankelijk van de volgorde wanneer ze in de oorspronkelijke constructie gelegd werden. Was de verbinding nodig om informatie het netwerk in te sturen, dan verleggen we naar C_2; was de verbinding nodig om informatie via C_1 te ontvangen, dan verleggen we naar C_i. Het is makkelijk na te gaan dat elke overgebleven computer nog steeds alle informatie van de andere computers ontvangt. Op deze manier hebben we een constructie met $k - 2$ verbindingen gevonden voor $m - 1$ computers. $\square$

2.2.1 $n = 5$

We moeten aantonen dat er geen constructie bestaat met $2n - 5 = 5$ verbindingen waarna het virus actief wordt.

Stel dat er toch zulk een constructie bestaat. Dan volgt uit het Lemma dat er geen enkele computer zijn eigen informatie na een aantal stappen terug kan ontvangen. Was dit immers wel het geval, dan zou er een constructie met 3 verbindingen voor 4 computers bestaan, en we weten reeds dat dit onmogelijk is.

Beschouw nu een computer C en tel het aantal verbindingen dat nodig is om zijn informatie te versturen naar alle andere computers. Dit zijn er minstens 4. Ook zijn er minstens 4 verbindingen nodig om alle informatie van de overige 4 computers te bezorgen aan C. Dit geeft ons in totaal 8 verbindingen, maar het is mogelijk dat we sommige verbindingen 2 keer hebben geteld. Een verbinding die we 2 keer tellen, is een verbinding die informatie van C verspreidt en tegelijkertijd informatie richting C doorgeeft. Wanneer C niet rechtstreeks in deze verbinding zou zitten, zou C zijn eigen informatie op deze manier terug ontvangen, wat in tegenspraak is met het Lemma.

Daarom hebben we minstens $8 - v(C)$ verbindingen nodig voor computer C, waarbij $v(C)$ het aantal verbindingen is dat rechtstreeks vanuit C vertrekt.

We weten dat $8 - v(C) \leq 5$, het totaal aantal verbindingen, zodat $v(C) \geq 3$. Dit geldt voor elke computer C, dus we hebben minstens $\frac{5v(C)}{2}$ verbindingen in totaal. Dit aantal is groter dan 5, een contradictie.

2.2.2 $5 \leq n \leq 9$

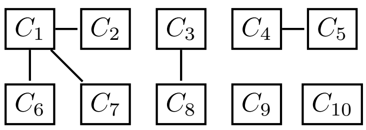
We merken op dat voor $n \leq 9$ het bewijs dat we gaven voor $n = 5$ geldig blijft aangezien voor $n \leq 9$ geldt dat $\frac{3n}{2} > 2n - 5$. Dus voor $5 \leq n \leq 9$ is het mininum aantal verbindingen nodig om het virus te activeren gelijk aan $2n - 4$.

2.2.3 $n = 10$

Voor $n = 10$ hebben we een ander argument nodig aangezien $\frac{3 \cdot 10}{2} = 15 = 2n - 5$.

Stel dat er toch een constructie bestaat met 15 verbindingen. Door het Lemma is elke verbinding ofwel de allereerste voor beide computers, ofwel voor geen van beiden de allereerste. Immers, als een computer A als allereerste verbinding B contacteert die reeds informatie kreeg van een andere computer C, dan weten we dat C op dat moment zeker niet beschikt over de informatie van A. Vanaf het moment dat de informatie van A in het netwerk zit (i.e. vanaf de eerste verbinding van A), zal dus de informatie van A steeds samen met de informatie van C doorgegeven worden. Op een zeker ogenblik moet de informatie van A aan C doorgegeven worden. Op dat ogenblik zal C zijn eigen informatie terugkrijgen, in tegenspraak met het Lemma (dat als we het toepassen leidt tot een tegenspraak voor $n = 9$ computers waar minstens 14 verbindingen nodig zijn). Analoog is elke verbinding ofwel de allerlaatste voor beide computers, ofwel voor geen van beiden. Dit begrijpen we als we even het tegengestelde onderstellen: een computer A maakt nog een verbinding nadat hij in verbinding was met B en dat ook de laatste verbinding van B was. Vermits B daarna niet meer verbindt, moet B in het bezit zijn van alle informatie. Maar dan moet ook A alle informatie bezitten van zodra de verbinding met B is afgelopen. Als er nadien nog een verbinding van pakweg C met A optreedt, zal C zijn eigen informatie opnieuw krijgen van A, weer in tegenspraak met het Lemma. Elke computer heeft een allereerste en een allerlaatste verbinding en elke verbinding verbindt exact 2 computers. Een verbinding $C_i \leftrightarrow C_j$ kan ook nooit de allereerste en tegelijk de allerlaatste zijn, omdat deze computers dan niet (eventueel onrechtsreeks) met de andere verbonden worden. Tellen we al deze allereerste en allerlaatste verbindingen op, dan bekomen we dus $\frac{10+10}{2} = 10$.

Wanneer we de 5 tussentijdse (=overblijvende) verbindingen tekenen voor de 10 computers, krijgen we 5 groepjes van onderling verbonden computers die niet met de andere groepjes verbonden zijn. Bijvoorbeeld:



De enige verbindingen die we niet tekenen zijn de allereerste en allerlaatste verbindingen. Beschouw computer C. Vooraleer de allerlaatste verbindingen worden gelegd, kan C zijn informatie dus enkel verspreiden in zijn eigen groepje of in het groepje van de computer met wie de allereerste verbinding van C werd gemaakt. Bovendien kan computer C na de allereerste verbindingen enkel geïnformeerd worden door zijn eigen groepje of door het groepje van zijn allerlaatste verbinding. Dit betekent dat er 2 groepjes niet bijdragen in het tussentijds verspreiden van de informatie van C, noch in het informeren van C. Met andere woorden, voor computers in deze 2 afgezonderde groepjes wordt informatie met C enkel uitgewisseld via allereerste en allerlaatste verbindingen, geen tussentijdse.

Tellen we opnieuw (zoals in het bewijs voor $n = 5$) het aantal verbindingen dat nodig is om de informatie van C te versturen naar alle andere computers en om alle informatie te bezorgen aan C. Dan krijgen we $9 + 9 - v(C)$ verbindingen, waarbij $v(C)$ het aantal verbindingen is dat rechtstreeks vanuit C vertrekt.

Stel $w(C)$ het aantal tussentijdse verbindingen die niet bijdragen in de informatiekettingen vanuit C, noch in de kettingen naar C. Dit zijn verbindingen die we nog niet geteld hebben, dus we bekomen minstens $18 - v(C) + w(C) \leq 15$ verbindingen. Dus $v(C) \geq 3 + w(C) \geq 3$.

Bijgevolg is $v(C) \geq 3$ en moet dus elke computer 3 verbindingen hebben waarbij 1 allereerste, 1 allerlaatste en dus minstens 1 tussentijdse. Omdat er 2 groepjes waren die niet bijdroegen in het tussentijds verspreiden van en naar C, en elk groepje dus minstens 1 tussentijdse verbinding moet hebben, is $w(C) \geq 2$. Dus $v(C) \geq 5$ en we tellen in totaal $\frac{10v(C)}{2} \geq 25 > 15$ verbindingen, een contradictie.