Computing functions over wireless networks

P. R. Kumar

Dept. of Electrical and Computer Engineering, and
Coordinated Science Lab
University of Illinois, Urbana-Champaign

Email: prkumar@illinois.edu
Web: http://decision.csl.illinois.edu/~prkumar

How to process information in the network in sensor networks?

(Or how to do data fusion over a sensor network?

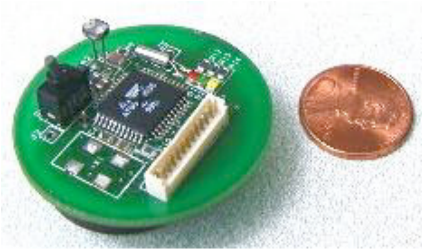
Or how to compute a function of data over a sensor network?

Or how to perform in-network information processing in a sensor network?)

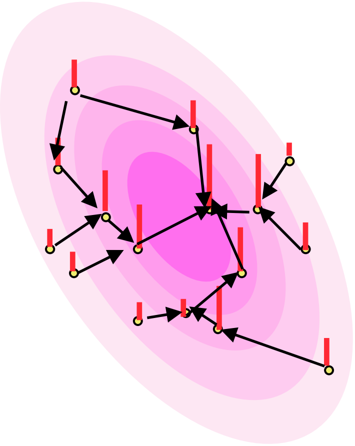
Outline

Sensor networks

- Examples of Tasks
  - Environment monitoring
    - » Determine the Average temperature: $(x_1 + x_2 + \ldots + x_n)/n$
  - Alarm networks
    - » Determine the Max temperature: Max x_i



- Sensor networks are not just data networks with sensor measurements replacing files
  - They are application specific
  - Nodes need not just relay packets
    - » They can discard, combine, process packets
    - » Combination of computing and communication



- More generally: Consider a symmetric function $F(x_1, x_2, \ldots, x_n)$
  - E.g., Average, Mode, Median, Percentile, Max
  - Determined by Histogram or "Types"
- How should information be processed *in the network* to compute such functions?
  - This can also be regarded as *network coding for sensor networks*

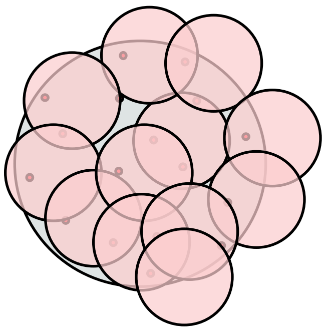
Model of problem:
Protocol model,
Non-information theoretic
(Giridhar & K '05)

Model of problem

- Multi-hop model for wireless communication (Giridhar & K '05)
  - Collocated network
    - » All nodes within range of all
  - Multi-hop random network (Penrose 1997, Gupta & K '98)
    - » Critical common range for connectivity of random graph

$$r(n) = \sqrt{\frac{\log n + c_n}{n}}$$

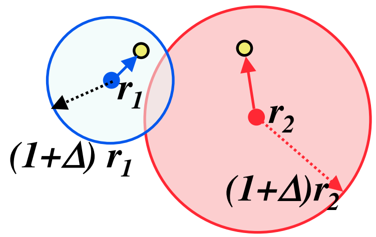
    - » where $c_n \to +\infty$. (Take $r(n) = \sqrt{\frac{2 \log n}{n}}$)



- At time t, sensor i takes a measurement $x_i(t) \in \{1, 2, \ldots, D\}$
- Fusion node needs to calculate $F(x_1, \ldots, x_n)$ exactly
- Non-information theoretic formulation

Model of problem

- Protocol Model for wireless communication
  - Receiver should be outside other transmitters' interference footprints

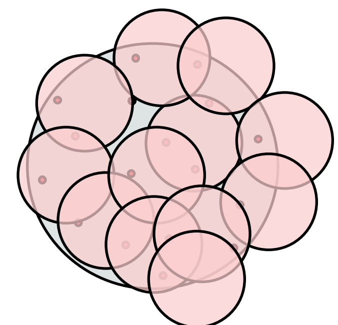


  - Communicate at rate W bits/sec (Take W =1 bit/sec wlog)

- At time t, sensor i takes a measurement $x_i(t) \in \{1, 2, \ldots, D\}$
  - No probability distribution on $x_i(t)$'s
- Fusion node needs to calculate $F(x_1, \ldots, x_n)$ exactly

- Non-information theoretic formulation

- Two types of networks
  - Collocated network
    - Large range so all nodes can hear each other
  - Random network



    - n nodes randomly distributed
    - Need range at no less than

$$r(n) \geq \sqrt{\frac{\log n}{n}}$$

for network to be connected

Definition of Computational Rate $R_{\max}(n)$

- Block coding allowed

N measurements of node 1: $x_1 \rightarrow$ $x_1(1), x_1(2), \ldots, x_1(N)$

N measurements of node 2 : $x_2 \rightarrow$ $x_2(1), x_2(2), \ldots, x_2(N)$

N measurements of node n : $x_n \rightarrow$ $x_n(1), x_n(2), \ldots, x_n(N)$

$F(x_1(1), \ldots, x_n(1)) \ldots\ldots F(x_1(N), \ldots, x_n(N))$

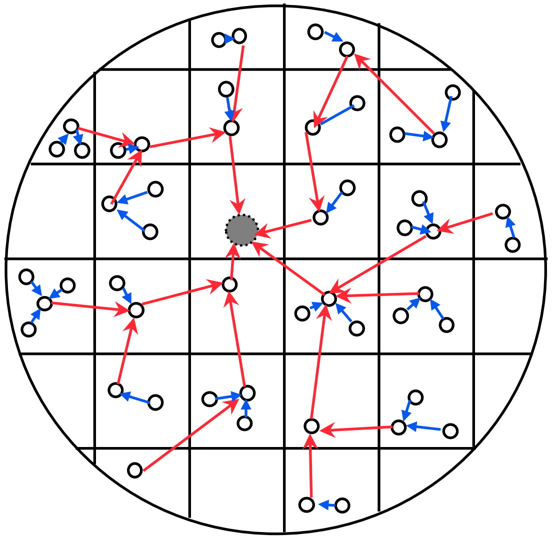
- – If all N functions computed in time T
- – Then Computational Rate $R = \frac{N}{T}$
- – <u>Best Rate over all Strategies S and block lengths N</u>: $R_{\max}(n) = \sup_{S,N} \frac{N}{T^S(N)}$

(Giridhar & K '05)

Sample of results: Average vs. Max

The *Average* versus *Max*

- Theorem (Giridhar & K '05): The rate at which the *Average* can be harvested is $\Theta(\frac{1}{\log n})$
  - Strategy
    - » Tessellate
    - » Fuse locally
    - » Compute along a rooted tree of cells



- Theorem (Giridhar & K '05): The rate at which the *Max* can be harvested is

$$\Theta\left(\frac{1}{\log \log n}\right)$$

  - Strategy: Take advantage of Block Coding
  - » First node announces times of max values: (1 1 1)
  - » Second node announces times of additional max values: (1 1)
  - » Third node announces of yet more max values: (1)

Summary: Order of difficulty of computations

$(1/n)$

Collocated network:
Average, Mode, Type

Data downloading

$(1/\log n)$

Random Multi-hop network:
Average, Mode, Type

Collocated network:
Max

$(1/\log\log n)$

Random Multi-hop network
Max

(Giridhar & K '05)

More details of results (Giridhar & K ‘05)

Results: A classification of functions

- Divisible functions
  - Amenable to divide and conquer
  - $R_{\max}(n) = \Theta(\frac{1}{\log |\mathcal{R}(F_n)|})$ if $\deg(G_n) = O(\log |\Re(F_n)|)$

- Symmetric functions

  $F_n(x) = F_n(\pi(x))$ for every permutation π
  - Data centric paradigm: Identity of node is not important, only its value

- *Type*-sensitive functions
  - Hard to compute
  - $R_{\max}(n) = \Theta(\frac{1}{n})$ in collocated case, and $R_{\max}(n) = \Theta(\frac{1}{\log n})$ random case

- *Type*-threshold functions
  - $R_{\max}(n) = \Theta(\frac{1}{\log n})$ collocated case, and $R_{\max} = \Theta(\frac{1}{\log \log n})$ random case

(Giridhar & K '05)

Examples (Giridhar & K '05)

- Data download problem: $F_n(x_1,\ldots,x_n) = (x_1,\ldots,x_n)$:
  - In collocated or random networks: $R_{\max}(n) = \Theta(\frac{1}{n})$
- Histogram of frequencies or "Type": $F_n(x_1,\ldots,x_n) = (z_1, z_2,\ldots, z_D)$
  - Collocated case: $R_{\max}(n) = \Theta(\frac{1}{n})$ Random networks: $R_{\max}(n) = \Theta(\frac{1}{\log n})$
  - Special case: Any symmetric function $F(x_1, x_2, \ldots, x_n)$
- Mean, Mode, Median, Majority:
  - Collocated case: $R_{\max}(n) = \Theta(\frac{1}{n})$ Random case: $R_{\max}(n) = O(\frac{1}{\log n})$
- Max, Min, Range, Occurrence of a value:
  - Collocated case: $R_{\max}(n) = \Theta(\frac{1}{\log n})$ Random case: $R_{\max} = \Theta(\frac{1}{\log \log n})$

Definition of Rate $R_{\max}(n)$

- Block coding allowed

N **measurements of node 1:** $x_1 \rightarrow x_1(1), x_1(2), \ldots, x_1(N)$

N **measurements of node 2 :** $x_2 \rightarrow x_2(1), x_2(2), \ldots, x_2(N)$

N **measurements of node** n **:** $x_n \rightarrow x_n(1), x_n(2), \ldots, x_n(N)$

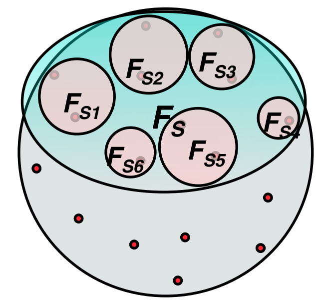
$F(x_1(1), \ldots, x_n(1)) \ldots\ldots F(x_1(N), \ldots, x_n(N))$

Compute all N functions

- If computed in time T
- Then Rate $R = \frac{N}{T}$
- Best Rate over all Strategies S and block lengths N: $R_{\max}(n) = \sup_{S,N} \frac{N}{T^S(N)}$
- Bound on $R_{\max}$: $R_{\max}(n) \leq \frac{W}{\log |\mathcal{R}(F_n)|}$

(Giridhar & K '05)

Divisible functions

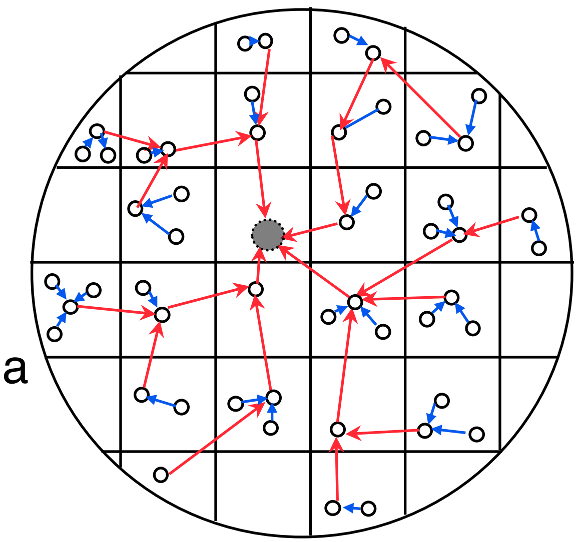


- Divisible functions:
  - There exists $F_S(x_i: i \in S)$ for every subset $S \subseteq \{1,2,\ldots,n\}$
  - With $|\Re(F_S)| \leq |\Re(F_n)|$
  - $F_S(x_S) = g_S(F_{S_1}(x_{S_1}), F_{S_2}(x_{S_2}), \ldots, F_{S_m}(x_{S_m}))$ for partition $\{S_1,\ldots, S_m\}$ of S

- Theorem: $R_{\max}(n) = \Theta(\frac{1}{\log|\mathcal{R}(F_n)|})$ if $\deg(G_n) = O(\log |\Re(F_n)|)$ (Giridhar & K '05)

- Special cases
  - Data Downloading: $\deg(G_n) \leq O(\log |\Re(F_n)|) = O(\log D^n) = O(n)$
  - So $R_{\max}(n) = \Theta(\frac{1}{n})$
  - Histogram: $|\mathcal{R}(F_n)| = \binom{n+D-1}{D}$ and $(\frac{n}{D})^D < \binom{n+D-1}{D} < (n+1)^D$
  - So $\deg(G_n) = \Theta(\log n) = \Theta(\log|\mathcal{R}(F_n)|)$ for Random networks
  - Hence $R_{\max}(n) = \Theta(\frac{1}{\log n})$ for Random networks (Giridhar & K '05)

Proof of $R_{\max}(n) = \Theta(\frac{1}{\log |\mathcal{R}(F_n)|})$ for Divisible Functions



- Tessellate into square cells of area $r^2/2$
- Neighboring occupied cells can communicate with each other
- Form a tree rooted at fusion center out of occupied connected cells
- Choose a relay node in each occupied cell and a parent in the next cell towards the root
- Locally compute and pass on along tree to root
  - Collect data from deg(G_n) nodes within cell
  - Collect functional value of log|$\Re(F_n)$| bits from bounded number of child cells
  - Pass on functional value of log|$\Re(F_n)$| bits to parent cell
- All operations can be performed $\Theta(\log |\mathcal{R}(F_n)|)$ time
- $R_{\max}(n) = \Theta(\frac{1}{\log |\mathcal{R}(F_n)|})$
- Constructive strategy

Symmetric functions

- Symmetric functions depend only on type $z = (z_1, z_2, \ldots, z_D)$
  - where $z_i =$ Number of occurences of i in $\{x_1, x_2, \ldots, x_n\}$
  - $F_n(x_1, x_2, \ldots, x_n) = \bar{F}_n(z_1, z_2, \ldots, z_D)$

- Type-sensitive functions
  - There is a $0 < c < 1$ such that a fraction c of values $(x_1, x_2, \ldots, x_{cn})$ is never enough to pin down the value of the function F_n
  - Examples: Mean, Median, Mode, Majority

- Type-threshold functions
  - Only want to know whether each z_i exceeds a threshold z_i^*
  - There is a threshold vector $z^* = (z_1^*, z_2^*, \ldots, z_D^*)$ such that
    $\bar{F}_n(z) = \bar{F}_n(z \wedge z^*)$ for all n
  - Examples: Max, Min, Range, Occurrence of value

(Giridhar & K '05)

Collision-free strategies in collocated case

- Every node knows when to transmit based on what it hears on channel
- The content of the packet it transmits depends on what it heard, as well as its own information

- Node g_1 transmits packet $P_1(x_{g1})$
- Node $g_2(P_1(x_{g1}))$ transmits packet $P_2(P_1(x_{g1}), x_{g2})$
- Node $g_3(P_1(x_{g1}), P_2(P_1(x_{g1}), x_{g2}))$ transmits packet $P_3(P_1(x_{g1}), P_2\ (x_{g2}), x_{g3})$
- ....

- Note: We are not allowing information transmission to occur through collisions

(Giridhar & K ‘05)

$R_{\max}(n) = \Theta(\frac{1}{n})$ for Type-sensitive functions in collocated case

- Wlog suppose D=2
- Initially, x_{g1} is in the set S^0_{g1} with cardinality $|S^0_{g1}| = 2^N$
- After first transmission, x_{g1} can be in one of two sets depending on whether it transmits 0 or 1
  - Let the transmission correspond to the larger set, call it be S^1_{g1}
  - $|S^1_{g1}| \geq 1/2\ |S^0_{g1}|$
- After t-th transmission of node k, let x_k lie in S^t_k with $|S^t_k| \geq 1/2\ |S^{t-1}_k|$
- So at the end, uncertainty set is: $|S_1 \times S_2 \times \ldots \times S_n| \geq 2^{nN-T}$

- Thus at least nN-T places in the nN values (x_1, x_2 ,…, x_n) are undetermined
- However to compute F_n (x(1),x(2),…,x(N)), at least cnN values are needed
- So nN-$T \leq$ (1-c)nN
- So $T \geq cnN$
- Hence $R = \frac{N}{T} \leq O(\frac{1}{cn})$

- Thus $R_{\max}(n) = O(\frac{1}{n})$ for collocated case

(Giridhar & K '05)

$R_{\max}(n) = \Theta(\frac{1}{\log n})$ for Type-threshold functions in collocated case

- Consider "Max" function (argument can be generalized)
- Threshold vector = (1,1,…,1)
- Lower Bound
  - Take block length $N = \ell n > n$
  - Node 1 transmits its locations of the N_1 1's in $(x_1(1), x_1(2), \ldots, x_1(N))$
  - Node 2 transmits the N_2 "new" 1's in its list $(x_2(1), x_2(2), \ldots, x_2(N))$
  - Node 3 transmits the N_3 "new" 1's in its list $(x_3(1), x_3(2), \ldots, x_3(N))$

  - To describe N_i takes log N bits
  - To describe the locations of N_i 1's requires $\log \binom{N-\sum_{j<i} N_j}{N_j} bits$
  - So $T = n \log N + \sum_i \log \binom{N-\sum_{j<i} N_j}{N_j}$

  - Maximized when N_i = N/n. Use $\binom{N}{\frac{N}{n} \ldots \frac{N}{n}} = \prod_i \binom{(n-i+1)\ell}{\ell} < \binom{n\ell}{\ell}^n < (\frac{ne\ell}{\ell})^{n\ell} = (ne)^{n\ell}$

  - Thus: $R = \frac{N}{T} = \Omega(\frac{1}{\log n})$

(Giridhar & K '05)

Upper bound $R_{\max}(n) = O(\frac{1}{\log n})$ in collocated case ....

- Take $N > 2n$. Consider

$x_1(1), x_1(2), \ldots, x_1(N)$ ← **Exactly $N/2n$ 1's in x_1**

$x_2(1), x_2(2), \ldots, x_2(N)$ ← **Exactly $N/2n$ 1's in x_2**

$x_n(1), x_n(2), \ldots, x_n(N)$ ← **Exactly $N/2n$ 1's in x_n**

At most one 1 .....**At most one 1**

- Claim: Each such x produces a unique set of transmissions $P_1, P_2, \ldots, P_T$
  - Suppose not. Then there are two: x and y which produce same transmissions
  - They differ in some $x_k \neq y_k$
  - Then $\{x_1, x_2, \ldots, x_{k-1}, y_k, x_{k+1}, \ldots, x_n\}$ also produces same transmissions since node k hears the same under x_k or y_k and so reacts the same
  - But this has different "Max" values from x
  - Thus "Max" functions are not determined from transmissions

(Giridhar & K '05)

Finishing the proof for the “Max” function

- Number of such vectors x = $\prod_{1\le i\le n} \binom{N-(i-1)\frac{N}{2n}}{\frac{N}{2n}} > \binom{\frac{N}{2}}{\frac{N}{2n}}^n > (n-1)^N$
- So $2^T >$ $(n-1)^N$
- So $T >$ N log(n-1)
- So $R = \frac{N}{T} \le \frac{1}{\log(n-1)}$
- Hence $R_{\max}(n) = O(\frac{1}{\log n})$
- This proves $R_{\max}(n) = \Theta(\frac{1}{\log n})$

(Giridhar & K ‘05)

Generalizing to any Type-threshold function for collocated case

- Feasibility of $R_{\max} = \Omega(\frac{1}{\log n})$
  - Node i sends only list of 1's for values which threshold has not been attained

- Upper bound of $R_{\max}(n) = O(\frac{1}{\log n})$
  - There exist a threshold vector such that

$$\bar{F}_n(z_1 - 1, z_2, \ldots, z_D) \neq \bar{F}_n(z_1, z_2, \ldots, z_D) \text{ for all } n \geq \sum_i z_i$$

  - Now consider an x which has
    - » z_1-1 vectors of the form (1,1,…,1)
    - » z_2 vectors of the form (2,2,…,2)
    - » z_3 vectors of the form (2,2,…,2)
    - » …
  - Remaining $n \geq (\sum z_i) - 1$ have 1's or 2's only
  - Now problem is reduced to a "Max"

(Giridhar & K '05)

Random networks: Type sensitive networks $R_{\max}(n) = \Theta(\frac{1}{\log n})$

- Theorem (Giridhar & K '05)
  - Type-sensitive functions: $R_{\max}(n) = \Theta(\frac{1}{\log n})$

- Proof
  - Tessellate unit area domain into squares of area $A = (\Delta r)^2/2$
  - Transmissions are local within square
  - Assume "Genie" communicates all messages instantaneously to all nodes
  - We know at least cnN transmissions are needed
  - At least one square has greater than $cnNA$ receptions
  - However only one node can receive at a time in a square
  - So $T \geq cnNA = cnN(\Delta r)^2/2$
  - But $r \geq \sqrt{\frac{\log n}{n}}$ for connectivity
  - So $T \geq c'N \log n$

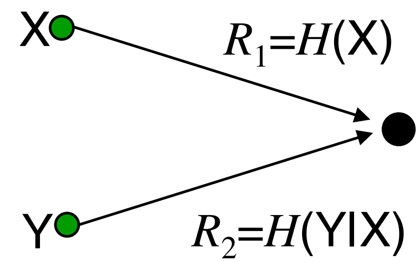
Type-threshold functions

- Theorem (Giridhar & K ’05)
  - Type-threshold functions: $R_{\max} = \Theta(\frac{1}{\log\log n})$
- Proof
  - Consider “Max” for simplicity
  - Tessellate unit area domain into squares of area $A = (\Delta r)^2/2$
  - Some square has greater than nA nodes
  - Suppose all nodes outside this square have value 0
  - Then we need to compute “Max” of nA nodes
  - We need $cN \log (nA)$ transmissions
  - Only one node can receive at any given time
  - So $T \geq cN \log (nA)$
  - But $r \geq \sqrt{\frac{\log n}{n}}$ s needed for connectivity
  - So $A \geq \frac{\log n}{n}$
  - So $T \geq N\, \Omega(\log \log n)$
  - So $R_{\max} = O(\frac{1}{\log\log n})$
  - Achievability can be proved by using tree gathering

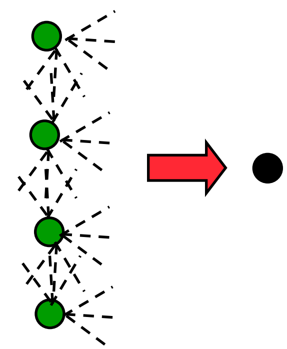
Some information theoretic results for sensor networks

Complexities of function computation over wireless networks

- Slepian-Wolf Theorem ('73)
  - Total information fusion over wires from correlated sources



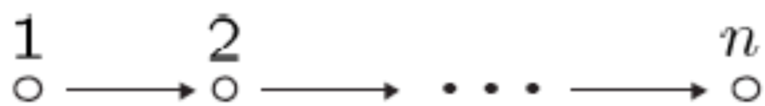
- Several other complexities in sensor networks
- Wireless nodes
  - There are no independent links: Sources share channel
  - Multiple access problem
- Also, sensors can communicate with each other and thus cooperate



- Source-channel separation does not hold
- Only a function needs to be computed, not all information
- Little is known in a pure information theoretic setting

Multiple relay network (Xie and K '05)

- Consider a network $p(y_2(t), y_3(t), \ldots, y_n(t) \mid x_1(t), x_2(t), \ldots, x_{n-1}(t))$



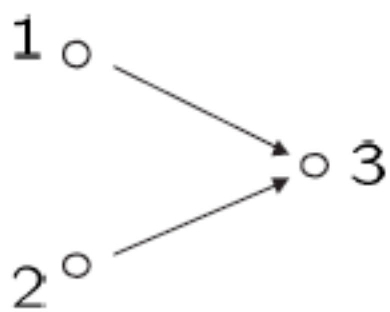
- A feasible rate

$$R < \max_{p(x_1,\ldots,x_{n-1})} \min_{2 \le k \le n} I(X_1, \ldots, X_{k-1}; Y_k | X_k, \ldots, X_{n-1})$$

- Generalization of Cover and El Gamal '79

Multiple access channel (Ahlswede '71, Liao '72)

- Consider the multiple access channel $p\left(y_3(t)|x_1(t), x_2(t)\right)$



- The capacity region

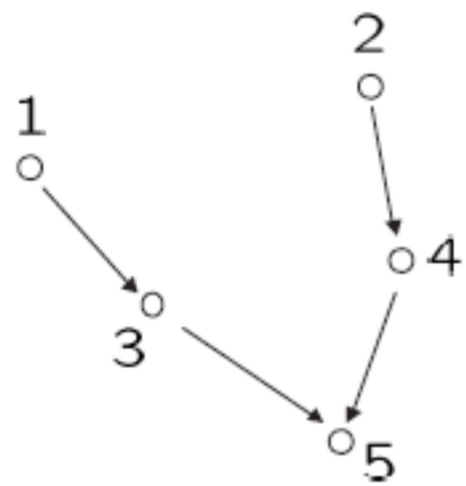
$$R_1 < I(X_1; Y_3|X_2)$$
$$R_2 < I(X_2; Y_3|X_1)$$
$$R_1 + R_2 < I(X_1, X_2; Y_3)$$

for any $p(x_1)p(x_2)$

A multiple source multiple relay network (Xie and K '07)

- Consider the network $p\left(y_3(t), y_4(t), y_5(t) \middle| x_1(t), x_2(t), x_3(t), x_4(t)\right)$

- A feasible rate vector is



$$R_1 < I(X_1; Y_3 | X_3)$$
$$R_2 < I(X_2; Y_4 | X_4)$$

and

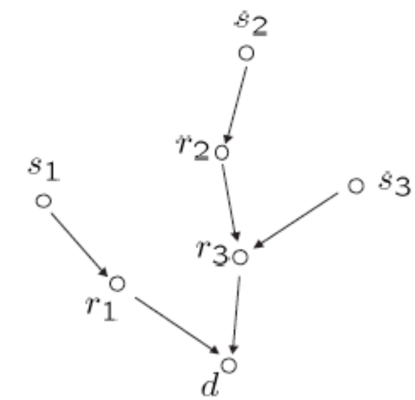
$$\begin{cases} R_1 < I(X_1, X_3; Y_5 | X_2, X_4) \\ R_2 < I(X_2, X_4; Y_5 | X_1, X_3) \\ R_1 + R_2 < I(X_1, X_3, X_2, X_4; Y_5) \end{cases}$$

for some joint distribution $p(x_1, x_3)p(x_2, x_4)$.

A feasible rate for a sensor network with a sink

- Consider the network $p\left(y_1(t),\ldots,y_n(t)\,|\,x_1(t),\ldots,x_n(t)\right)$

- Feasible rate region for acyclic choice of routes



$$R^{(1)} < I(U^{(1)}_{s_1}, U^{(1)}_{r_1}; Y_d | U^{(1)}_d, U^{(2)}_{s_2}, U^{(2)}_{r_2}, U^{(2)}_{r_3}, U^{(2)}_d, U^{(3)}_{s_3}, U^{(3)}_{r_3}, U^{(3)}_d)$$

$$R^{(2)} < I(U^{(2)}_{s_2}, U^{(2)}_{r_2}, U^{(2)}_{r_3}; Y_d | U^{(2)}_d, U^{(1)}_{s_1}, U^{(1)}_{r_1}, U^{(1)}_d, U^{(3)}_{s_3}, U^{(3)}_{r_3}, U^{(3)}_d)$$

$$R^{(3)} < I(U^{(3)}_{s_3}, U^{(3)}_{r_3}; Y_d | U^{(3)}_d, U^{(1)}_{s_1}, U^{(1)}_{r_1}, U^{(1)}_d, U^{(2)}_{s_2}, U^{(2)}_{r_2}, U^{(2)}_{r_3}, U^{(2)}_d)$$

$$R^{(1)} + R^{(2)} < I(U^{(1)}_{s_1}, U^{(1)}_{r_1}, U^{(2)}_{s_2}, U^{(2)}_{r_2}, U^{(2)}_{r_3}; Y_d | U^{(1)}_d, U^{(2)}_d, U^{(3)}_{s_3}, U^{(3)}_{r_3}, U^{(3)}_d)$$

$$R^{(1)} + R^{(3)} < I(U^{(1)}_{s_1}, U^{(1)}_{r_1}, U^{(3)}_{s_3}, U^{(3)}_{r_3}; Y_d | U^{(1)}_d, U^{(3)}_d, U^{(2)}_{s_2}, U^{(2)}_{r_2}, U^{(2)}_{r_3}, U^{(2)}_d)$$

$$R^{(2)} + R^{(3)} < I(U^{(2)}_{s_2}, U^{(2)}_{r_2}, U^{(2)}_{r_3}, U^{(3)}_{s_3}, U^{(3)}_{r_3}; Y_d | U^{(2)}_d, U^{(3)}_d, U^{(1)}_{s_1}, U^{(1)}_{r_1}, U^{(1)}_d)$$

$$R^{(1)} + R^{(2)} + R^{(3)} < I(U^{(1)}_{s_1}, U^{(1)}_{r_1}, U^{(2)}_{s_2}, U^{(2)}_{r_2}, U^{(2)}_{r_3}, U^{(3)}_{s_3}, U^{(3)}_{r_3}; Y_d | U^{(1)}_d, U^{(2)}_d, U^{(3)}_d)$$

- Maximize over $p\,(u^{(1)}_{s_1}, u^{(1)}_{r_1}, u^{(1)}_d)\,p\,(u^{(2)}_{s_2}, u^{(2)}_{r_2}, u^{(2)}_{r_3}, u^{(2)}_d)\,p\,(u^{(3)}_{s_3}, u^{(3)}_{r_3}, u^{(3)}_d),$

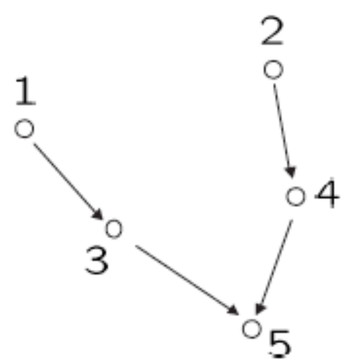
- Based on combination of backward decoding for relay channel and multiple access channel

(Xie and K '07)

Exact capacity region for some Gaussian sensor networks with phase fading

- Kramer, Gastpar and Gupta '03 have determined exact capacity region for relay channel with phase fading for some geometries

- **Theorem** (Xie and Kumar '07)
  - Phase fading unknown to transmitter
  - Node 5 far away from other nodes



- The capacity region is

For node 3,

$$R^{(1)} < \log\left(1 + \frac{P_1/d_{13}^{\alpha}}{N_3 + P_2/d_{23}^{\alpha} + P_4/d_{43}^{\alpha}}\right), \qquad (22)$$

for node 4,

$$R^{(2)} < \log\left(1 + \frac{P_2/d_{24}^{\alpha}}{N_4 + P_1/d_{14}^{\alpha} + P_3/d_{34}^{\alpha}}\right), \qquad (23)$$

and for node 5,

$$\begin{cases} R^{(1)} < \log\left(1 + \frac{P_1/d_{15}^{\alpha} + P_3/d_{35}^{\alpha}}{N_5}\right), & (24) \\ R^{(2)} < \log\left(1 + \frac{P_2/d_{25}^{\alpha} + P_4/d_{45}^{\alpha}}{N_5}\right), & (25) \\ R^{(1)} + R^{(2)} < & (26) \\ \quad \log\left(1 + \frac{P_1/d_{15}^{\alpha} + P_2/d_{25}^{\alpha} + P_3/d_{35}^{\alpha} + P_4/d_{45}^{\alpha}}{N_5}\right), & \end{cases}$$

References-1

- M.D. Penrose, “The longest edge of the random minimal spanning tree,” The Annals of Probability, vol. 7, no. 2, pp. 340-361, 1997.
- Piyush Gupta and P. R. Kumar, “Critical Power for Asymptotic Connectivity in Wireless Networks,” in Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming. Edited by W. M. McEneany, G. Yin, and Q. Zhang, Birkhauser, Boston, MA, pp. 547–566, 1998. ISBN 0-8176-4078-9
- Arvind Giridhar and P. R. Kumar, “Computing and Communicating Functions Over Sensor Networks,” *IEEE Journal on Selected Areas in Communications, pp. 755–764, vol. 23, no. 4, April 2005.*
- Arvind Giridhar and P. R. Kumar, “Towards a Theory of In-Network Computation in Wireless Sensor Networks” *IEEE Communications Magazine, vol. 44, no. 4, pp. 98–107, April 2006.*
- Arvind Giridhar and P. R. Kumar, “In-Network Information Processing in Wireless Sensor Networks.” In *Wireless Sensor Networks: Signal Processing and Communications Perspectives, A. Swami, Q. Zhao, Y.-W. Hong and L. Tong, Editors. pp. 43–67, John Wiley, Chichester, England, 2007.*

References-2

- Slepian D and Wolf J 1973 Noiseless coding of correlated information sources. *IEEE Transactions on Information Theory,* 19(4), 471–480.
- T. Cover and A. El Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inform. Theory, vol. 25, pp. 572-584, 1979.*
- R. Ahlswede, “Multi-way communication channels,” in *Proceedings of the 2nd Int. Symp. Inform. Theory (Tsahkadsor, Armenian S.S.R.),* (Prague), pp. 23-52, Publishing House of the Hungarian Academy of Sciences, 1971.
- H. Liao, *Multiple access channels. PhD thesis, Department of Electrical* Engineering, University of Hawaii, Honolulu, HA, 1972.
- G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inform. Theory,* vol. 51, pp. 3037-3063, September 2005.

References-3

- Liang-Liang Xie and P. R. Kumar, “An Achievable Rate for the Multiple-Level Relay Channel,” *IEEE Transactions on Information Theory, vol. 51, no. 4, pp. 1348–1358, April 2005.*
- Liang-Liang Xie and P. R. Kumar, “Multisource, multidestination, multirelay wireless networks,” *IEEE Transactions on Information Theory, Special issue on Models, Theory and Codes for Relaying and Cooperation in Communication Networks, vol. 53, no. 10, pp. 3586–3595, October 2007.*
- Liang-Liang Xie and P. R. Kumar, “Information-Theoretic Studies of Wireless Sensor Networks.” In *Handbook on Array Processing and Sensor Networks, Simon Haykin and K. J. Ray Liu, Editors. IEEE-Wiley, 2009.*