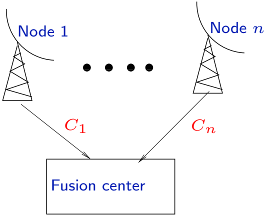
Energy Scaling Laws for Distributed Inference in Random Networks

A. Anandkumar[1] J. E. Yukich[2] Lang Tong[1] A. Swami[3]

[1]School of Elect. & Comput. Engineering, Cornell University, Ithaca, NY.
[2]School of Mathematics, Lehigh University, Bethlehem, Pa.
[3]Army Research Laboratory, Aldephi, MD

Allerton 2008
24^{th} Sept., 2008

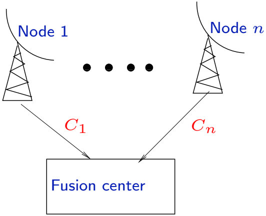
Distributed Statistical Inference

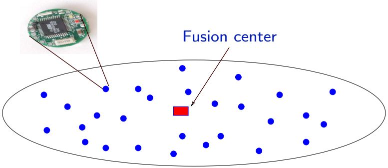


Classical distributed inference

- Many-to-one data fusion
- Rate constraints on fusion links
- Quantization rule at local sensors
- Inference rule at fusion center

Distributed Statistical Inference





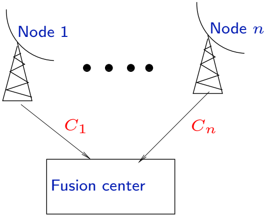
Classical distributed inference

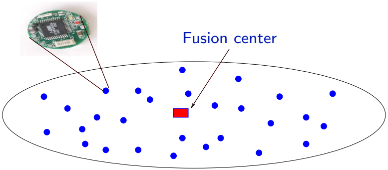
- Many-to-one data fusion
- Rate constraints on fusion links
- Quantization rule at local sensors
- Inference rule at fusion center

Sensor networks for inference

- Multihop data fusion
- Energy constraints
- Transmission and routing policies
- Quantization and inference rules

Distributed Statistical Inference





Classical distributed inference

- Many-to-one data fusion
- Rate constraints on fusion links
- Quantization rule at local sensors
- Inference rule at fusion center

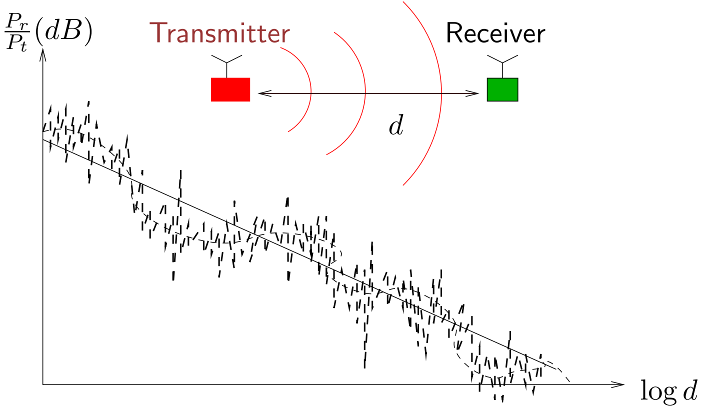
Sensor networks for inference

- Multihop data fusion
- Energy constraints
- Transmission and routing policies
- Quantization and inference rules

Energy Consumption for Distributed Inference

Outline

Propagation model and assumptions

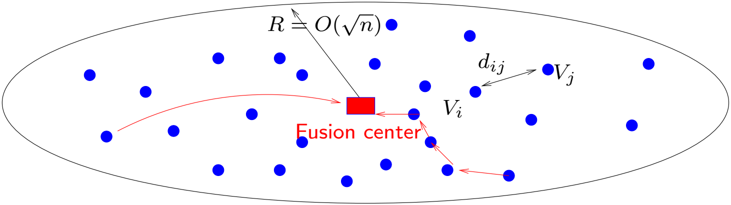


Energy cost per sample: $\mathcal{E} = O(d^\nu)$, $2 \le \nu \le 4$, ν is Path Loss.

Network Model and Assumptions

Network Model

- Network has a fixed node density $\lambda = \frac{n}{\pi R^2}$: $R = O(\sqrt{n})$.
- Sensor locations $V_i \overset{\text{i.i.d.}}{\sim} \sqrt{\frac{n}{\lambda}}\kappa(x), i = 1, \cdots, n$ (e.g., Uniform)
  - $\kappa(x)$ has support on unit square
- Adjustable transmission power for multihop or direct transmission.
- For connectivity, $\max \mathcal{E}_i, i \in \mathbf{V}_n$ grows at least at $O((\sqrt{\log n})^\nu)$.



Inference Model and Assumptions

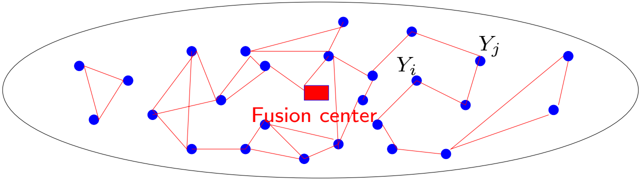
Binary Hypothesis Testing

- Location $\mathbf{V}_n \stackrel{\Delta}{=} (V_1, \cdots, V_n)$ and observations $\mathbf{Y}_n \stackrel{\Delta}{=} (Y_1, \cdots, Y_n)$.
- $\mathcal{H}_k : (\mathbf{Y}_n, \mathbf{V}_n) \sim f_k(\mathbf{y}_n|\mathbf{v}_n)\kappa(\mathbf{v}_n), \quad k = 0, 1$

$\mathbf{Y}_n$: Markov random field with dependency graph $\mathcal{G}_k = (\mathbf{V}_n, E_k)$

$$-\log f_k(\mathbf{y}_n|\mathbf{v}_n; \mathcal{G}_{n,k}) = \sum_{c \in \mathcal{C}_k} \Psi_{k,c}(\mathbf{y}_c)$$

$\mathcal{C}_{n,k}$: collection of maximal cliques, $\Psi_{k,c} > 0$: clique potentials

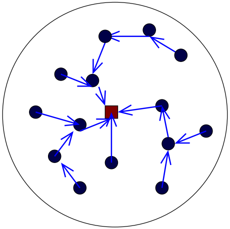


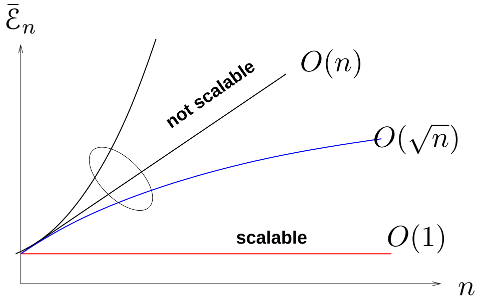
Outline

Energy Scaling Law

Energy Consumption For Inference

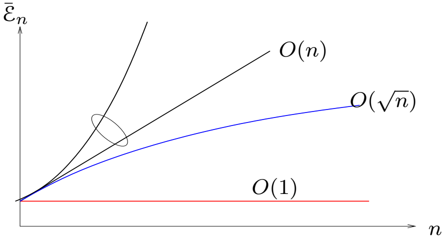
- Require optimal inference at the fusion center
- Examine average energy expenditure: $\bar{\mathcal{E}}_n = \frac{1}{n} \sum_i \mathcal{E}_i$



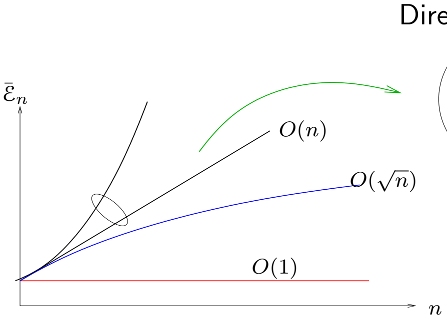


Goal: find a scalable data fusion strategy for optimal inference

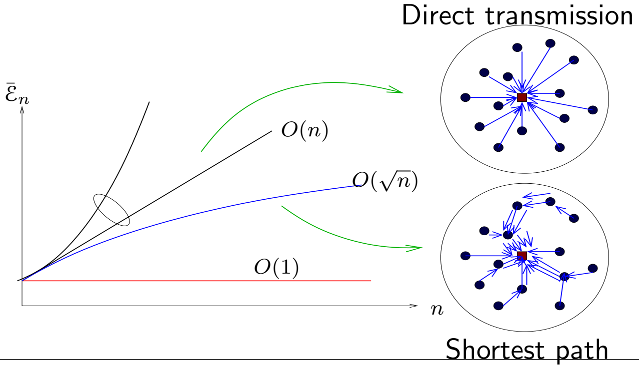
Energy Scaling Law



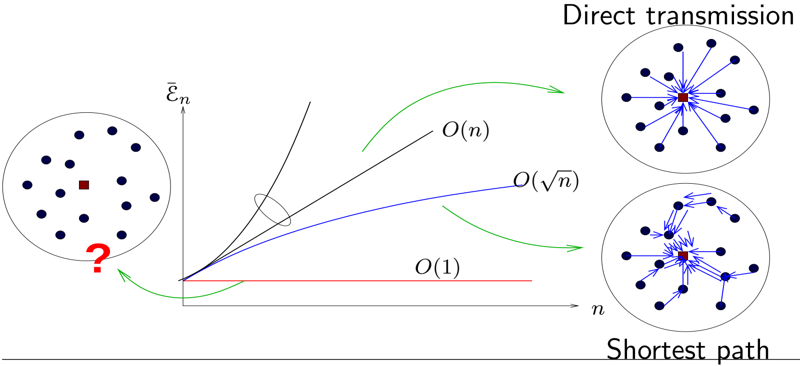
Energy Scaling Law



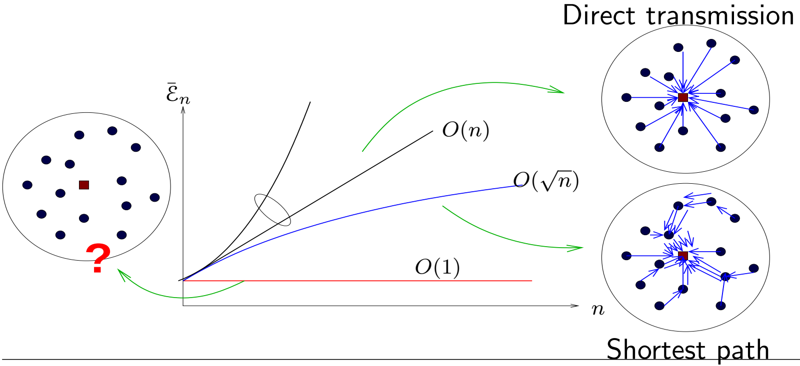
Energy Scaling Law



Energy Scaling Law

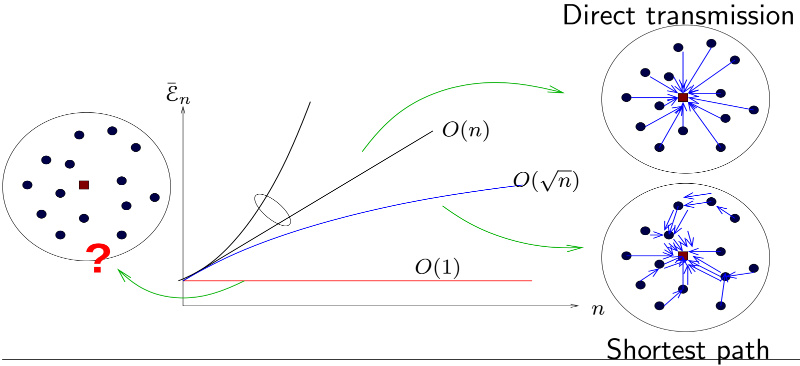


Energy Scaling Law



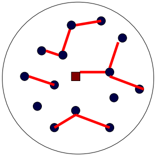
Finite Scaling for Local Spatial Dependencies

Energy Scaling Law

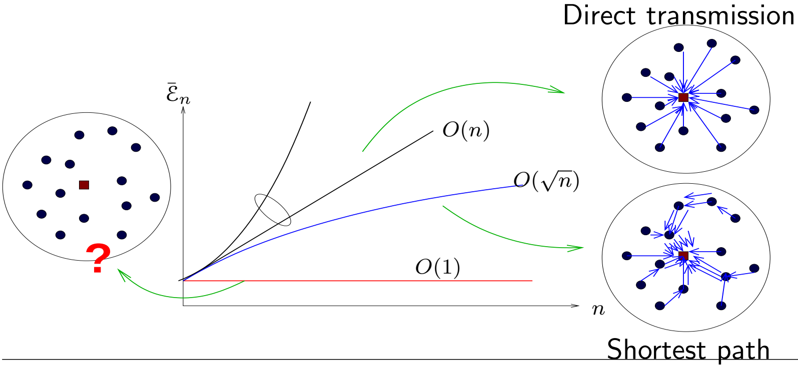


Finite Scaling for Local Spatial Dependencies

Finite Disk Graph: **Yes**

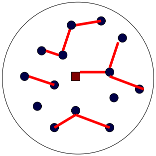


Energy Scaling Law

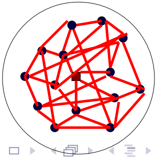


Finite Scaling for Local Spatial Dependencies

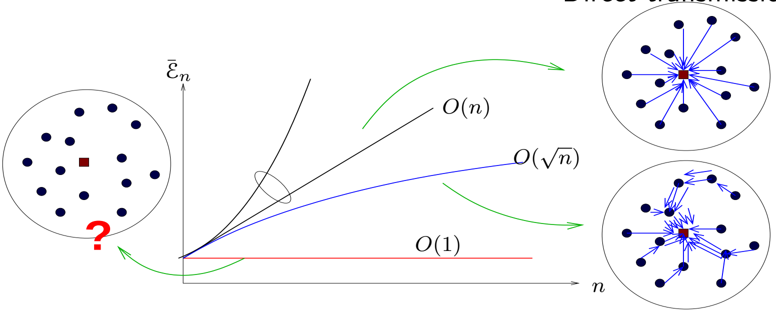
Finite Disk Graph: **Yes**



Complete Dependency: **No**



Energy Scaling Law



Finite Scaling for Local Spatial Dependencies

Finite Disk Graph: **Yes** Complete Dependency: **No**

Summary of Results

Finite Average Energy Scaling

For Stabilizing Dependency Graphs e.g., k-NNG, disk graph

- Construction of a suboptimal fusion scheme DFMRF and its analysis

Summary of Results

Finite Average Energy Scaling

For Stabilizing Dependency Graphs e.g., k-NNG, disk graph

- Construction of a suboptimal fusion scheme DFMRF and its analysis

Additional Results

- Lower Bound on Energy Consumption for Optimal Inference
- Constant Factor Approximation Ratio for DFMRF

Summary of Results

Finite Average Energy Scaling

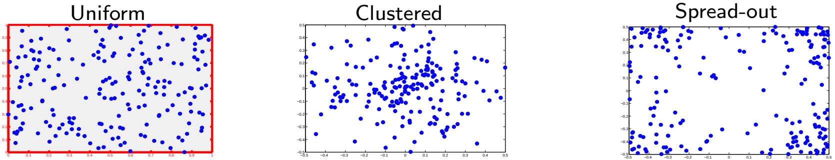
For Stabilizing Dependency Graphs e.g., k-NNG, disk graph

- Construction of a suboptimal fusion scheme DFMRF and its analysis

Additional Results

- Lower Bound on Energy Consumption for Optimal Inference
- Constant Factor Approximation Ratio for DFMRF

Influence of Node Placement Distribution κ



Summary of Results

Finite Average Energy Scaling

For Stabilizing Dependency Graphs e.g., k-NNG, disk graph

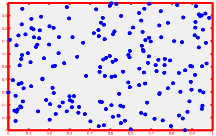
- Construction of a suboptimal fusion scheme DFMRF and its analysis

Additional Results

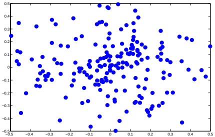
- Lower Bound on Energy Consumption for Optimal Inference
- Constant Factor Approximation Ratio for DFMRF

Influence of Node Placement Distribution κ

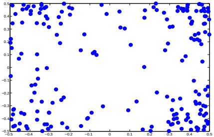
Uniform



Clustered



Spread-out



Optimality of Uniform Placement Over IID Placements

For Scale-Invariant Dependency (k-NNG) and Path loss ≥ 2

Outline

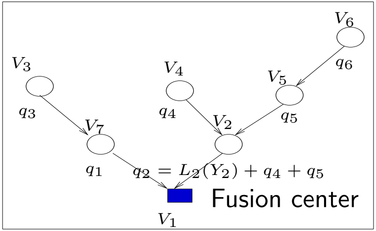
Optimal Fusion: Independent Case

IID Measurements

$\mathcal{H}_k : \mathbf{Y}_n \sim \prod_i f_k(y_i)$

Sufficient Statistic

$$L(\mathbf{y}_n) = \log \frac{f_0(\mathbf{y}_n)}{f_1(\mathbf{y}_n)} = \sum_i L_i(y_i)$$



Optimal data fusion is LLR aggregation over MST

- Each node must transmit at least once
- MST minimizes edge sum for spanning trees: $\min \sum_i |e_i|^\nu$
- Fusion rule: $q_i = \sum_{j \in \mathcal{N}(i)} q_j + L_i(Y_i)$

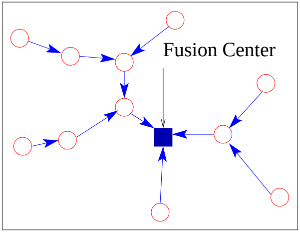
Optimal fusion: energy analysis

Average energy per node

$$\bar{\mathcal{E}}_n = \frac{1}{n} \sum_{e \in \mathsf{MST}} |e|^\nu$$

LLN: Steele'88, Penrose-Yukich'03

$$\frac{1}{n} \sum_{e \in \mathsf{MST}} |e|^\nu \xrightarrow{L^2} \bar{\mathcal{E}}_\infty$$



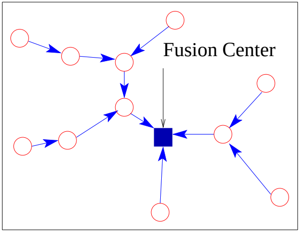
Optimal fusion: energy analysis

Average energy per node

$$\bar{\mathcal{E}}_n = \frac{1}{n} \sum_{e \in \mathsf{MST}} |e|^\nu$$

LLN: Steele'88, Penrose-Yukich'03

$$\frac{1}{n} \sum_{e \in \mathsf{MST}} |e|^\nu \xrightarrow{L^2} \bar{\mathcal{E}}_\infty$$



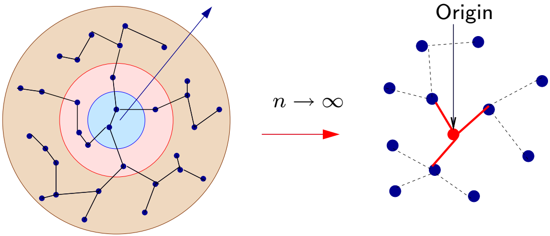
Scaling Constant for κ bounded away from 0 & ∞

$$\bar{\mathcal{E}}_\infty = \zeta(\nu; \mathsf{MST}) \int_{[-\frac{1}{2},\frac{1}{2}]^2} \kappa(x)^{1-\frac{\nu}{2}} dx, \quad \zeta(\nu; \mathsf{MST}) = \frac{1}{2}\mathbb{E}\left[\sum_{\substack{e \in \mathsf{MST}(\mathcal{P}_1 \cup \mathbf{0}) \\ \mathbf{0} \subset e}} |e|^\nu \right]$$

Key idea: global property to local property

Scaling Constant: Law of Large Numbers (Penrose & Yukich '03)

$$\bar{\mathcal{E}}_\infty = \zeta(\nu;\textsf{MST}) \int\limits_{[-\frac{1}{2},\frac{1}{2}]^2} \kappa(x)^{1-\frac{\nu}{2}}dx, \quad \zeta(\nu;\textsf{MST}) = \frac{1}{2}\mathbb{E}\left[\sum_{\substack{e\in\textsf{MST}(\mathcal{P}_1\cup\mathbf{0})\\ \mathbf{0}\subset e}} |e|^\nu\right]$$



Normalized sum of edge weights

Expectation for edges of origin of Poisson process

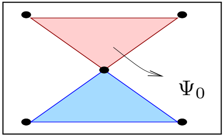
$$\frac{1}{n}\sum_{e\in\textsf{MST}(\mathbf{V}_n)} |e|^\nu \longrightarrow \bar{\mathcal{E}}_\infty$$

Outline

Fusion for Markov random field

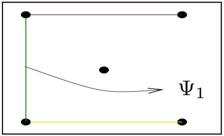
Null Hypothesis

$\mathcal{H}_0 : \{\mathcal{G}_0(\mathbf{V}), \mathcal{C}_0, \Psi_0\}$



Alternative Hypothesis

$\mathcal{H}_1 : \{\mathcal{G}_1(\mathbf{V}), \mathcal{C}_1, \Psi_1\}$



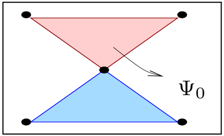
Binary hypothesis on MRF

for $k = 0, 1, \quad \mathcal{H}_k : \mathbf{Y}_n \sim f_k(\mathbf{y_V}; \mathcal{G}_k) = \exp\{-\sum_{c \in \mathcal{C}_k} \Psi_{k,c}(\mathbf{y}_c)\}$

Fusion for Markov random field

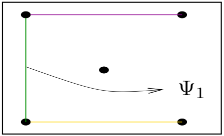
Null Hypothesis

$\mathcal{H}_0 : \{\mathcal{G}_0(\mathbf{V}), \mathcal{C}_0, \Psi_0\}$



Alternative Hypothesis

$\mathcal{H}_1 : \{\mathcal{G}_1(\mathbf{V}), \mathcal{C}_1, \Psi_1\}$



Binary hypothesis on MRF

for $k = 0, 1, \quad \mathcal{H}_k : \mathbf{Y}_n \sim f_k(\mathbf{y_V}; \mathcal{G}_k) = \exp\{-\sum_{c \in \mathcal{C}_k} \Psi_{k,c}(\mathbf{y}_c)\}$

Minimal Sufficient Statistic: $\mathcal{G} \triangleq \mathcal{G}_0 \bigcup \mathcal{G}_1 = (\mathbf{V}, \mathcal{E}_0 \bigcup \mathcal{E}_1)$

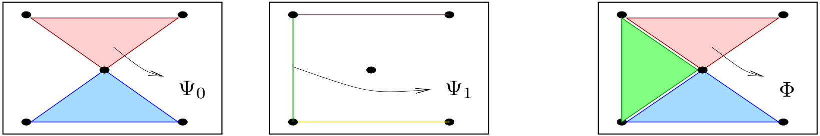
$$L_{\mathcal{G}}(\mathbf{Y_V}) = \log \frac{f_0(\mathbf{Y_V}; \mathcal{G}_0)}{f_1(\mathbf{Y_V}; \mathcal{G}_1)} = \sum_{c \in \mathcal{C}} \phi_c(\mathbf{Y}_c),$$

where $\mathcal{C}$ is the collection of maximal cliques of $\mathcal{G}$

Fusion for Markov random field

Null Hypothesis | **Alternative Hypothesis** | **Effective MRF For LLR**

$\mathcal{H}_0 : \{\mathcal{G}_0(\mathbf{V}), \mathcal{C}_0, \Psi_0\}$ | $\mathcal{H}_1 : \{\mathcal{G}_1(\mathbf{V}), \mathcal{C}_1, \Psi_1\}$ | $\{\mathcal{G}_0(\mathbf{V}) \cup \mathcal{G}_1(\mathbf{V}), \mathcal{C}_0 \cup \mathcal{C}_1, \Phi\}$



Binary hypothesis on MRF

for $k = 0, 1, \quad \mathcal{H}_k : \mathbf{Y}_n \sim f_k(\mathbf{y}_\mathbf{V}; \mathcal{G}_k) = \exp\{-\sum_{c \in \mathcal{C}_k} \Psi_{k,c}(\mathbf{y}_c)\}$

Minimal Sufficient Statistic: $\mathcal{G} \stackrel{\Delta}{=} \mathcal{G}_0 \bigcup \mathcal{G}_1 = (\mathbf{V}, \mathcal{E}_0 \bigcup \mathcal{E}_1)$

$$L_\mathcal{G}(\mathbf{Y}_\mathbf{V}) = \log \frac{f_0(\mathbf{Y}_\mathbf{V}; \mathcal{G}_0)}{f_1(\mathbf{Y}_\mathbf{V}; \mathcal{G}_1)} = \sum_{c \in \mathcal{C}} \phi_c(\mathbf{Y}_c),$$

where $\mathcal{C}$ is the collection of maximal cliques of $\mathcal{G}$

Optimal Fusion Scheme for Inference

Optimization Statement: π^*

- Minimize sum routing costs s.t. $L_{\mathcal{G}}(\mathbf{Y}_n) = \sum_{c \in \mathcal{C}} \Phi_c(\mathbf{Y}_c)$ is delivered
- Steiner tree reduction under local processor assignment: NP-hard

Optimal Fusion Scheme for Inference

Optimization Statement: π^*

- Minimize sum routing costs s.t. $L_{\mathcal{G}}(\mathbf{Y}_n) = \sum_{c \in \mathcal{C}} \Phi_c(\mathbf{Y}_c)$ is delivered
- Steiner tree reduction under local processor assignment: NP-hard

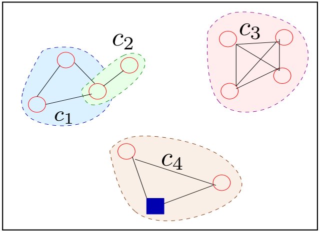
Data Fusion for Markov random field: DFMRF

Optimal Fusion Scheme for Inference

Optimization Statement: π^*

- Minimize sum routing costs s.t. $L_{\mathcal{G}}(\mathbf{Y}_n) = \sum_{c \in \mathcal{C}} \Phi_c(\mathbf{Y}_c)$ is delivered
- Steiner tree reduction under local processor assignment: NP-hard

Data Fusion for Markov random field: DFMRF

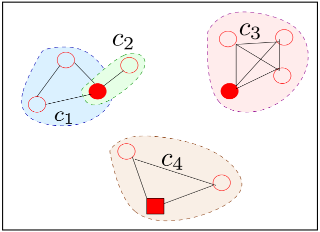


Optimal Fusion Scheme for Inference

Optimization Statement: π^*

- Minimize sum routing costs s.t. $L_{\mathcal{G}}(\mathbf{Y}_n) = \sum_{c \in \mathcal{C}} \Phi_c(\mathbf{Y}_c)$ is delivered
- Steiner tree reduction under local processor assignment: NP-hard

Data Fusion for Markov random field: DFMRF

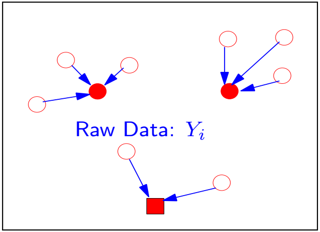


Optimal Fusion Scheme for Inference

Optimization Statement: π^*

- Minimize sum routing costs s.t. $L_{\mathcal{G}}(\mathbf{Y}_n) = \sum_{c \in \mathcal{C}} \Phi_c(\mathbf{Y}_c)$ is delivered
- Steiner tree reduction under local processor assignment: NP-hard

Data Fusion for Markov random field: DFMRF

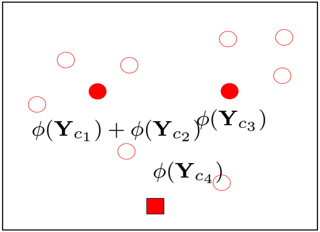


Optimal Fusion Scheme for Inference

Optimization Statement: π^*

- Minimize sum routing costs s.t. $L_{\mathcal{G}}(\mathbf{Y}_n) = \sum_{c \in \mathcal{C}} \Phi_c(\mathbf{Y}_c)$ is delivered
- Steiner tree reduction under local processor assignment: NP-hard

Data Fusion for Markov random field: DFMRF

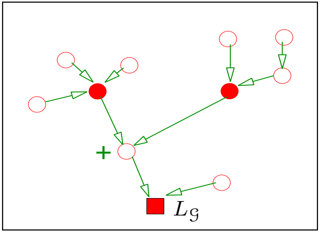


Optimal Fusion Scheme for Inference

Optimization Statement: π^*

- Minimize sum routing costs s.t. $L_{\mathcal{G}}(\mathbf{Y}_n) = \sum_{c \in \mathcal{C}} \Phi_c(\mathbf{Y}_c)$ is delivered
- Steiner tree reduction under local processor assignment: NP-hard

Data Fusion for Markov random field: DFMRF

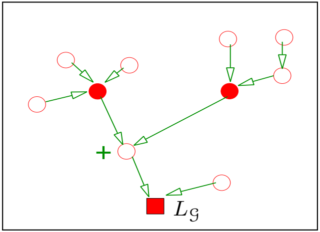


Optimal Fusion Scheme for Inference

Optimization Statement: π^*

- Minimize sum routing costs s.t. $L_{\mathcal{G}}(\mathbf{Y}_n) = \sum_{c \in \mathcal{C}} \Phi_c(\mathbf{Y}_c)$ is delivered
- Steiner tree reduction under local processor assignment: NP-hard

Data Fusion for Markov random field: DFMRF



DFMRF

- Local processor assignment and MST aggregation
- Total energy consumption= Data Forwarding + Aggregation

Fusion on Markov random field: energy scaling law

Assumptions

- Dependency $\mathcal{G}$ translation & scale invariant, stablizing (k-NNG)
- Set of feasible links is a u-energy spanner for finite constant u
  - SP energy no more than u times SP energy on complete graph
- Placement distribution κ is bounded away from 0 and ∞ on $[-\frac{1}{2}, \frac{1}{2}]^2$

Fusion on Markov random field: energy scaling law

Assumptions

- Dependency $\mathcal{G}$ translation & scale invariant, stablizing (k-NNG)
- Set of feasible links is a u-energy spanner for finite constant u
  - SP energy no more than u times SP energy on complete graph
- Placement distribution κ is bounded away from 0 and ∞ on $[-\frac{1}{2}, \frac{1}{2}]^2$

Scaling Result for DFMRF

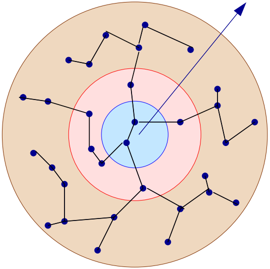
$$\limsup_{n\to\infty} \frac{\mathcal{E}(\text{DFMRF})}{n} \leq \lambda^{-\frac{\nu}{2}} [\underbrace{u\,\zeta(\nu; \mathcal{G})}_{\text{data forward}} + \underbrace{\zeta(\nu;\ \text{MST})}_{\text{MST aggregation}}] \int_{[-\frac{1}{2},\frac{1}{2}]^2} \kappa(x)^{1-\frac{\nu}{2}} dx,$$

$$\zeta(\nu; \mathcal{G}) \triangleq \frac{1}{2}\mathbb{E}\left[\sum_{\substack{e\in\mathcal{G}(\mathcal{P}_1\cup\mathbf{0})\\ \mathbf{0}\subset e}} |e|^\nu\right]$$

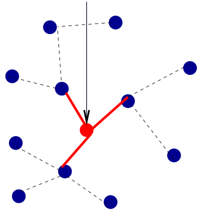
Key ideas

Bound on Energy for Forwarding to Processors Proc

$$\mathcal{E}(\text{Forward}) \leq \sum_{c\in\mathcal{C}(\mathbf{V})}\sum_{i\subset c} \text{SP}(i, \text{Proc}(c)) \leq u \sum_{c\in\mathcal{C}(\mathbf{V})}\sum_{i\subset c} |i, \text{Proc}(c)|^{\nu} \leq u \sum_{e\in\mathcal{G}} |e|^{\nu}$$



$n \to \infty$



Normalized sum of edge weights

Expectation for edges of origin of Poisson process

$$\frac{1}{n}\sum_{e\in\mathcal{G}(\mathbf{V}_n)} |e|^{\nu} \longrightarrow \frac{1}{2}\mathbb{E}\Big[\sum_{\substack{e\in\mathcal{G}(\mathcal{P}_1\cup\mathbf{0})\\ \mathbf{0}\subset e}} |e|^{\nu}\Big] \int_{[-\frac{1}{2},\frac{1}{2}]^2} \kappa(x)^{1-\frac{\nu}{2}}\, dx$$

Illustration of Stabilization

General LLN on Stabilizing Graph $\mathcal{G}$

$$\frac{1}{n} \sum_{e \in \mathcal{G}(\mathbf{V}_n)} |e|^{\nu} \to \text{constant}$$

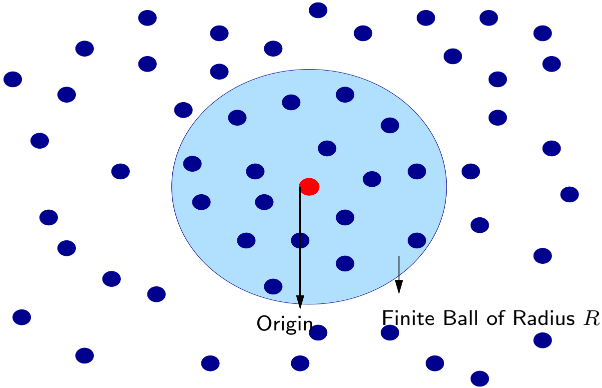


Illustration of Stabilization

General LLN on Stabilizing Graph $\mathcal{G}$

$$\frac{1}{n}\sum_{e\in\mathcal{G}(\mathbf{V}_n)}|e|^{\nu}\rightarrow \text{constant}$$

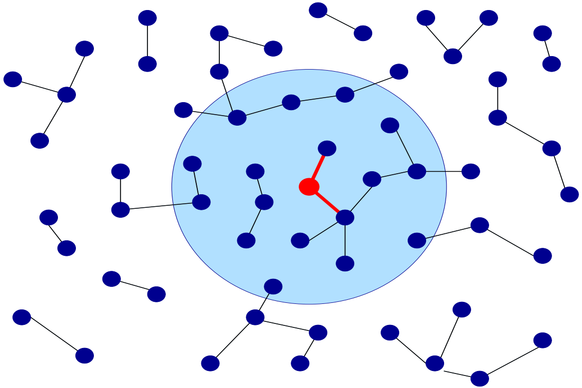


Illustration of Stabilization

General LLN on Stabilizing Graph $\mathcal{G}$

$$\frac{1}{n}\sum_{e\in\mathcal{G}(\mathbf{V}_n)}|e|^\nu \longrightarrow \text{constant}$$

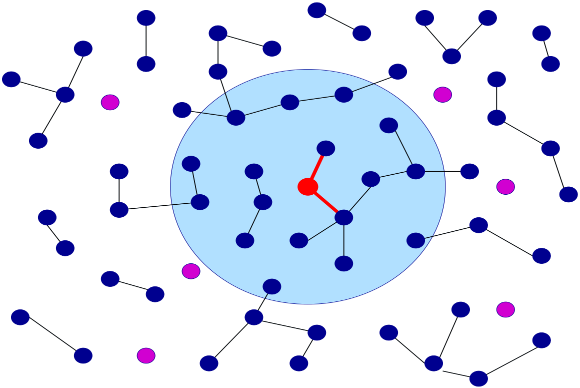


Illustration of Stabilization

General LLN on Stabilizing Graph $\mathcal{G}$

$$\frac{1}{n} \sum_{e \in \mathcal{G}(\mathbf{V}_n)} |e|^\nu \rightarrow \text{constant}$$

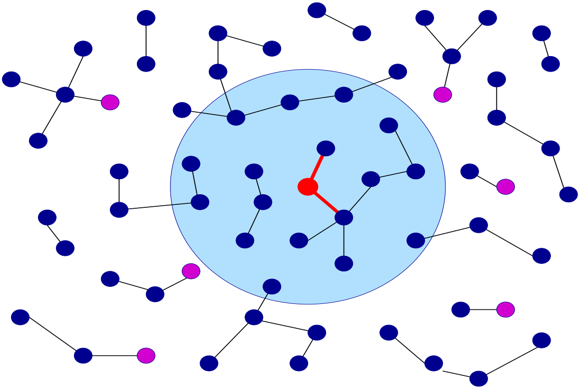


Illustration of Stabilization

General LLN on Stabilizing Graph $\mathcal{G}$

$$\frac{1}{n} \sum_{e \in \mathcal{G}(\mathbf{V}_n)} |e|^{\nu} \rightarrow \text{constant}$$

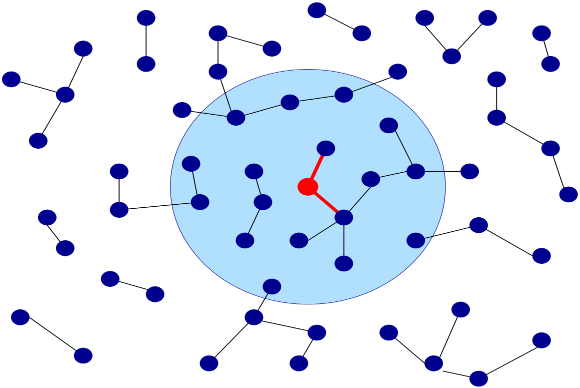


Illustration of Stabilization

General LLN on Stabilizing Graph $\mathcal{G}$

$$\frac{1}{n} \sum_{e \in \mathcal{G}(\mathbf{V}_n)} |e|^{\nu} \to \text{constant}$$

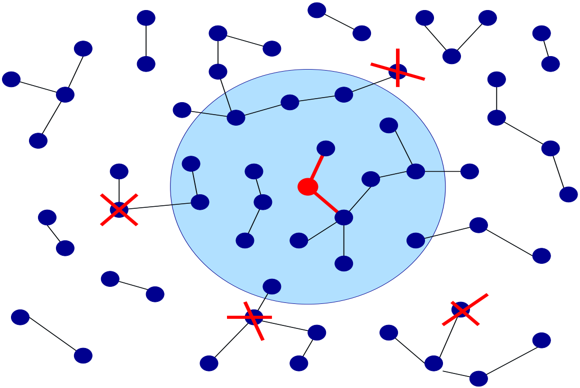
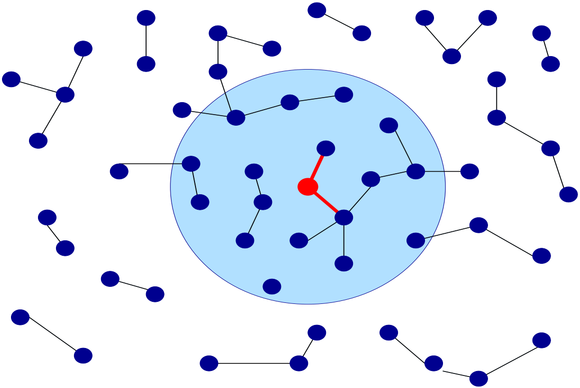


Illustration of Stabilization

General LLN on Stabilizing Graph $\mathcal{G}$

$$\frac{1}{n} \sum_{e \in \mathcal{G}(\mathbf{V}_n)} |e|^{\nu} \rightarrow \text{constant}$$



Scaling Law for Optimal Fusion

Lower and Upper Bounds For Any Network

$$\bar{\mathcal{E}}_n(\mathsf{MST}(\mathbf{V}_n)) \leq \bar{\mathcal{E}}_n(\pi^*(\mathbf{V}_n)) \leq \bar{\mathcal{E}}_n(\mathsf{DFMRF}(\mathbf{V}_n))$$

Bounds For Large Random Networks Under $\mathcal{G}$ Dependency

$$\zeta(\nu; \mathsf{MST}) \leq \lim_{n\to\infty} \frac{\bar{\mathcal{E}}_n(\pi^*(\mathbf{V}_n))}{\lambda^{-\frac{\nu}{2}} \int\limits_{[-\frac{1}{2},\frac{1}{2}]^2} \kappa(x)^{1-\frac{\nu}{2}}\, dx} \leq [u\,\zeta(\nu;\mathcal{G}) + \zeta(\nu;\mathsf{MST})]$$

Finite Average Energy Scaling For Distributed Inference

Approximation ratio of DFMRF for Large Random Networks

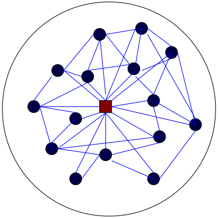
$$\limsup_{n\to\infty} \frac{\mathcal{E}(\mathsf{DFMRF}(\mathbf{V}_n))}{\mathcal{E}(\pi^*(\mathbf{V}_n))} \leq \left(1 + u\frac{\zeta(\nu;\mathcal{G})}{\zeta(\nu;\ \mathsf{MST})}\right)$$

Constant Factor Approximation for DFMRF

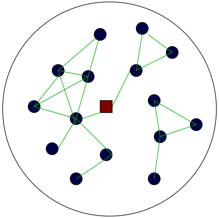
Outline

Summary

Network graph
Feasible Links for Communication

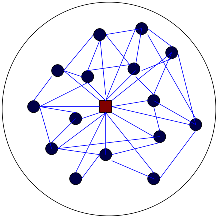


Dependency graph
Correlation Model of Data

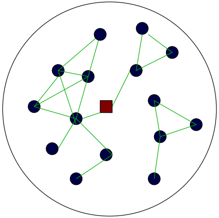


Summary

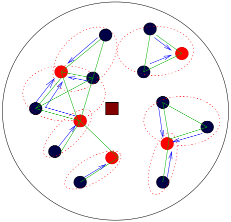
Network graph
Feasible Links for Communication

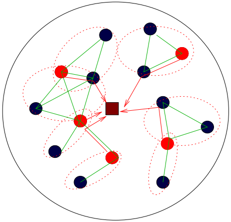


Dependency graph
Correlation Model of Data



DFMRF Scheme





Forwarding on shortest path   Aggregation on MST

Conclusion and future work

Concluding remarks

- Energy consumption is a key design parameter for large wireless sensor networks.
- Sensor location is a new source of randomness in distributed inference
- Asymptotic techniques are useful in overall network design.

Conclusion and future work

Concluding remarks

- Energy consumption is a key design parameter for large wireless sensor networks.
- Sensor location is a new source of randomness in distributed inference
- Asymptotic techniques are useful in overall network design.

Future work

- General behavior of error exponents in MRF
- Impact of sensor distribution on energy-performance tradeoff.

References

http://acsp.ece.cornell.edu/members/anima.html

Aggregation and Error Exponents in Large Random Networks

- A. Anandkumar, J.E. Yukich, L. Tong, and A. Swami "*Energy Scaling Laws for Distributed Inference in Random Networks*," sub. to JSAC, Aug.08. Available on Arxiv
- A. Anandkumar, L. Tong, and A. Swami, "*Detection of Gauss-Markov random fields with nearest-neighbor dependency*," accepted to IEEE Tran. Information Theory, Jul. 08
- A. Anandkumar, L. Tong, and A. Swami, "*Optimal Node Density for Detection in Energy Constrained Random Networks*," accepted to IEEE Tran. Signal Processing

Minimum Cost Data Fusion in Arbitrary Networks

- A. Anandkumar, L. Tong, A. Swami, A. Ephremides, "*Minimum Cost Data Aggregation with Localized Processing for Statistical Inference*," INFOCOM 2008
- A. Anandkumar, A. Ephremides, L. Tong, and A. Swami, "*Minimum Cost Routing with Local Processing for Distributed Statistical Inference*," in handbook on Array Processing and Sensor Networks, (S. Haykin and R. Liu, eds.), 2008.

Cost-Performance Tradeoff in Arbitrary Networks

- A. Anandkumar, L. Tong, A. Swami, and A. Ephremides, "*Cost-Performance Tradeoff in Multi-hop Aggregation for Statistical Inference*," in Proc. IEEE ISIT, July 2008