Green Traffic Compression in Wireless Sensor Networks

Kang-Hao Peng*, Kwang-Cheng Chen*, Shao-Lun Huang†, Shao-Chou Hung* and Xinhao Cheng‡
*Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan
†Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA
‡China Research Laboratory, IBM Research, Shanghai, China
r99942078@ntu.edu.tw, ckc@ntu.edu.tw, shaolun@mit.edu, r00942124@ntu.edu.tw and chengxh@cn.ibm.com

***Abstract*—Emerging multi-hop machine-to-machine (M2M) communications that likely support a large number of wireless devices create new challenges for spectrum scarcity and energy efficiency. In parallel to pursuing physical layer transmission efficiency, traffic compression to reduce required wireless transmissions suggests a new paradigm of wireless networks. Utilizing the natures of broadcasting and information collection in wireless sensor or machine networks, cognitive traffic compression can be facilitated by our proposed optimal fusion rules and topology compression algorithm. Therefore, only the necessary and connected sensors/machines in M2M networks are required to transmit, to achieve the desirable distortion of information collection (i.e. detection/estimation error). In other words, given the desirable distortion, the number of sensors to transmit, or equivalently the total energy consumption, serves our purpose of energy efficiency for end-to-end networking. Numerical results show successful compression of total network traffic to significantly enhance networking energy efficiency.**

I. Introduction

Multi-hop machine-to-machine (M2M) communications and wireless sensor networks (WSN) [1], [2] are promising technologies to fulfill human intelligent life. Possible applications include meticulous healthcare services [3], and fictional wireless robotics [4]. Suggested by [1], M2M and WSN share similar features except additional action executors in M2M network, and are both information collection networks.

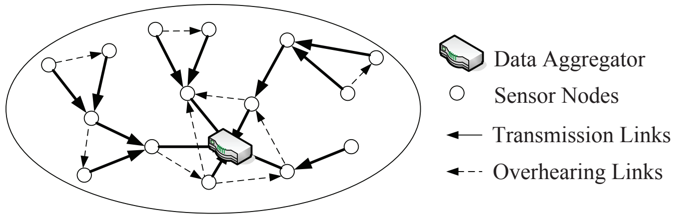
However, to implement such network, the massiveness of M2M devices engenders new technology bottleneck. Sensors, data aggregators (DA), action executors and personal communication devices integrally form a machine swamp that indicates a thirst for new protocols to sustain a leaped number of wireless communication devices. Consequently, as suggested in [4], [5], the major challenges in M2M communications include spectrum scarcity, network scalability, device deployment, device management and energy efficiency.

Approaches to mitigate these challenges are many, but we discover the elementary root-cause is: sensor signals are highly temporal and spatial correlated [6], whereas sensors are too dense [7]. This suggests that big amount of data transportation in a network might not have similarly large amount of information (in terms of entropy). Instead of energy efficient physical layer transmission, it shall be worth compressing total amount of traffic by identifying and dynamically turning on only the necessary transmissions in such network.

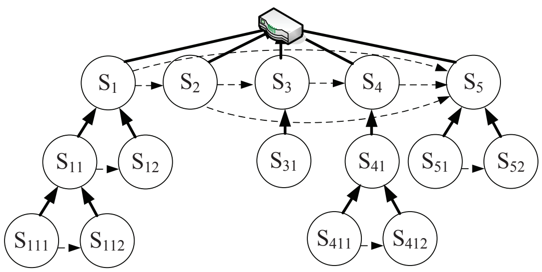
Related traffic compression researches include information selection [8] that measures importance of information and discard those unimportant, Slepian-Wolf based energy-efficient clustering algorithm [9] that reduces data transmission and saves energy, and the utility of network coding [10] that compresses correlated data. Specifically, [11] proposed progressive estimation to fuse observation without overhearing to an optimal estimator at DA, which serves as the baseline of this paper. However, these approaches have no attempts on exploiting signal correlations from broadcast and interaction between sensors to compress traffic.

Notably, signals that are transmitted from other transmitter-receiver pairs are typically deemed as interference, such as cognitive radio (CR) network [12]. Different from such widely accepted thinking, since transmission order can be pre-determined and sensors have low duty cycles, such interference can be fully utilized and received using TDMA in WSN and M2M communication without resource competition. Therefore, we call the interfering signals as overheard signals in WSN and M2M communication. Furthermore, exploiting overheard signals to compress traffic is theoretically assured by distributed source coding with side information [13], [14], where side information can be deemed as overheard signals. [15] study Gaussian interference channel to indicate cooperation (by overhearing) between transmitters/receivers increases network capacity region. [16] indicates that chatting in WSN through chatting channels dramatically improve detection/estimation performance. Specifically, however, [17] indicated no sum-rate gain for broadcast and interaction between agents (sensors) in Gaussian CEO problem for ideal channel, and thus no noise is introduced into sensors' overhearing. The assumption of noisy channels is more practical and, acts as a key to traffic compression performance.

In this paper we develop a novel traffic compression methodology that consists of 2-stage processing. The first stage is the optimal fusion rule for sensors to leverage its overheard signals. The subsequent second stage is a cognitive topology compression algorithm for DA to identify necessary operating devices. Via correlation among signal transmissions, a portion of network devices are no longer necessary to transmit, therefore the purpose of energy efficiency and spectral efficiency is accomplished.



(a) Receiving overheard/interference signals in WSN is feasible.



(b) Network topology of Fig. 1(a). Since sensors are highly probable to overhear nearby sensors, overhearing links are assumed to exists between sensors that have the same topological parent.

Fig. 1: In WSN and M2M communication, overheard signals can be exploited to reduce estimation error.

II. Traffic Compression Methodology

As shown in Fig. 1(a), sensors in a WSN collect information of a physical quantity $\theta \in \mathbb{R}$ from the environment, where θ is a random variable. Subsequently, sensors transmit their signals through a multi-hop tree-structured network, and fuse and relay their signals to a data aggregator (DA) according to a transmission order. Each sensor transmit once, and parent node transmits after all its child nodes complete transmission. By geographically closeness, a sensor may overhear signals from other sensors before its transmission. Finally, the DA that is the sink of the network receives signals from sensors, and estimates θ by $\hat{\theta}$. Assume noises are introduced into sensors' observations and transmissions, we measure the quality of $\hat{\theta}$ by a distortion function $d(\theta, \hat{\theta})$, and we wish the following distortion criterion to be satisfied,

$$E[d(\theta, \hat{\theta})] \leq D, \text{ for a given distortion constraint } D > 0,$$

where $E[.]$ denotes expectation over θ and $\hat{\theta}$. In this paper, we will not construct optimal transmission order and network topology. Instead, we will show that for every given transmission order and topology, the utility of the overheard signals reduces distortion. Therefore an algorithm can be designed to reduce network topology and save energy.

A. Notation for WSN Topology

We denote the topology of a WSN by a directed graph $G_o = (\mathcal{V}, \mathcal{E})$ ("o" represents overhearing). Vertices $\mathcal{V}$ represent all sensors and the DA. Edges $\mathcal{E}$ are disjointly composed of transmission links $\mathcal{E}_t$ and overhearing links $\mathcal{E}_o$. Because G_o is a WSN, the transmission topology $G = (\mathcal{V}, \mathcal{E}_t) \subset G_o$ is assumed to be a tree rooted at the DA.

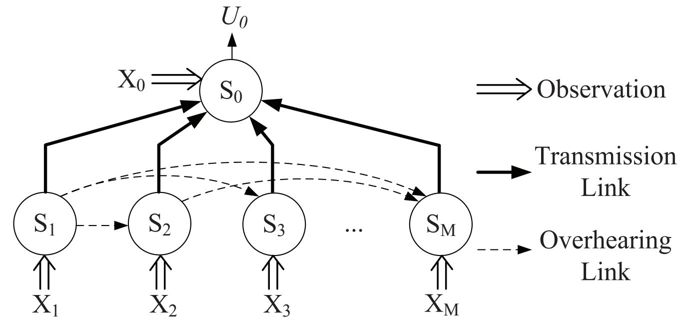


Fig. 2: Data fusion on one-layer network. Since inter-layer overhearing links are discarded, consider Optimal Fusion Rule problem on star-structured network is sufficient.

We assume that due to geographical deployment, overhearing links exist only between sensors that have the same parent in G. This assumption leads us to an easy form of fusion rules.

B. Stochastic Model and Data Fusion

Fig. 2 shows a layer in a tree network. The parent node S_0 has M children $S_1, \cdots, S_M$. Without loss of generality, each child S_k transmits according to the increment of k. The parent S_0 is the last one to transmit. Child S_k observes $X_k = \theta + Z_k$, receives an overheard signal $Y_{i,k} = U_i + Q_{i,k}$ from child S_i that transmits antecedently ($i < k$), and transmits the fused signal $U_k(X_k, \mathbf{Y}_k)$ to S_0, where $\mathbf{Y}_k = [Y_{1,k} \cdots Y_{k-1,k}]^T$ denote all the overheard signals for S_k. If there are no overhearing link exists between S_k and S_i, then S_k treats $Y_{i,k}$ as a dummy variable. Finally, parent S_0 observes $X_0 = \theta + Z_0$, receives $\{R_k = U_k + V_k\}_{k=1}^{M}$ from all its child nodes, and transmits fused signal U_0 to its parent. This fusion process is conducted from bottom to root, and DA fuses $\hat{\theta}$ as the final estimator. For each i, k, the zero mean additive Gaussian noises $\{Z_k\}, \{Q_{i,k}\}$ and $\{V_k\}$ are independent, each has variance $\text{Var}(Z_k) = \sigma_k^2$, $\text{Var}(Q_{i,k}) = s_{i,k}^2$ and $\text{Var}(V_k) = N_k$. Specifically, $\{Z_k\}$ model the spatial variation of sensors to the impact of observations, and $\{Q_{i,k}\}$ and $\{V_k\}$ model the quantization and wireless channel noises [18]. Assume the DA knows the *a priori* distribution of θ as $\theta \sim N(0, \sigma^2)$.

C. Procedures of Traffic Compression Methodology

Traffic compression methodology consists of the following 2-stage processes:

- **Stage 1** (Optimal Fusion Rule) Given a distortion/cost function $d(\theta, \hat{\theta}) > 0$, the topology of the network G_o and transmission order, find every optimal fusion rule U_k for sensor S_k in G_o, such that the Bayes' risk $r(G_o)$ is minimized,

$$\min_{U_k, \forall S_k \in G_o} r(G_o) = E[d(\theta, \hat{\theta})] \tag{1}$$

- **Stage 2** (Topology Compression Algorithm) Subsequently, given an initial topology G_o, a distortion constraint $D > 0$, the transmission order and the optimal fusion rules in stage 1, find a reduced/necessary topology $G_o' = (\mathcal{V}', \mathcal{E}') \subset G_o$ such that $G' = (\mathcal{V}', \mathcal{E}' \cap \mathcal{E}_t)$ is also a tree rooted at the DA (thus G_o' is connected), distortion criterion is satisfied $r(G_o') \le D$, and the necessary sensor number (cardinality of $\mathcal{V}'$) is minimized,

$$\min_{G_o'} |\mathcal{V}'| \text{ s.t. } r(G_o') \le D. \tag{2}$$

III. Optimal Fusion Rule

In this section we first derive optimal fusion rules on one-layer topology, then we will show that if these fusion rules are conducted from leaves to root, the global optimality of $\hat{\theta}$ at the DA is also guaranteed. Through out this paper, we consider squared error as distortion measure, $d(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2$.

A. Fusion Rules on One-Layer Star-Structured Network

As shown in Fig. 2, where $\{S_k\}_{k=1}^M$ are leaf nodes. We assume each sensor uses unbiased linear fusion rule,

$$U_k = \boldsymbol{\alpha}_k^T \mathbf{Y}_k + \beta_k X_k, \tag{3}$$

where $\boldsymbol{\alpha}_k, \mathbf{Y}_k \in \mathbb{R}^{k-1}$, $\beta_k \in \mathbb{R}$. The unbiasedness is assured by regulating the linear coefficients be normalized to unity,

$$\boldsymbol{\alpha}_k^T \mathbf{1}_{k-1} + \beta_k = 1, \text{ where } \mathbf{1}_{k-1} = [1 \ 1 \ \cdots \ 1]^T \in \mathbb{R}^{k-1}. \tag{4}$$

Further, if S_k does not overhear signal from S_i, then regulate the ith element in $\boldsymbol{\alpha}_k$ be zero, $[\boldsymbol{\alpha}_k]_i = 0$, if $(S_i, S_k) \notin \mathcal{E}_o$. As a result, since each U_k is unbiased Gaussian estimator of θ, so does each R_k. This feature enables every parent node S_0 to fuse the minimum variance unbiased estimator (MVUE) $\hat{X}_0$ easily by

$$\hat{X}_0 = \frac{[r_1 \ \cdots \ r_M] \, \Sigma^{-1} \mathbf{1}_M}{\mathbf{1}_M^T \Sigma^{-1} \mathbf{1}_M}, \tag{5}$$

where we assume the conditional covariance matrix $\Sigma := \text{Cov}(R_1, \cdots, R_M | \theta)$ is locally known by S_0. Now, since $\hat{X}_0$ and X_0 are conditionally independent, S_0 fuse them into X_0^* (also an MVUE) by,

$$X_0^* = \frac{\hat{\sigma}_0^2 X_k + \sigma_0^2 \hat{X}_0}{\hat{\sigma}_0^2 + \sigma_0^2}, \text{ where } \hat{\sigma}_0^2 = \text{Var}(\hat{X}_0 | \theta). \tag{6}$$

Remark 1. *If S_0 overhears $\mathbf{Y}_0$, we see that both X_0^* and X_0 are conditionally independent of $\mathbf{Y}_0$. Therefore X_0^* can replace X_0 as a better version of observation by the following (7). In fact, if any S_k has child, we can deem its observation X_k as the replaced version by X_k^*.*

$$\sigma_0^{2*} := \text{Var}(X_0^* | \theta) = \left(\mathbf{1}_M^T \Sigma^{-1} \mathbf{1}_M + \sigma_0^{-2} \right)^{-1} < \sigma_0^2 \tag{7}$$

Remark 2. *DA is also a parent node that has the a priori distribution knowledge of θ, but it does not observe nor overhear. DA fuses the minimum mean square error (MMSE) estimator by $\hat{\theta} = X_0^* = \frac{\sigma_0^2 \hat{X}_0}{\hat{\sigma}_0^2 + \sigma^2}$.*

The next question is, how can each sensor S_k fuse $U_k(\mathbf{Y}_k, X_k)$ so to minimize σ_0^{2*}, or equivalently maximize $\mathbf{1}_M^T \Sigma^{-1} \mathbf{1}_M$ in (7)? This is a more challenging task and we may want to define more notations for mathematical simplicity. Define conditional covariance matrix $\Sigma_1 := \text{Var}(R_1 | \theta)$ and

$$\Sigma_k := \text{Cov}(R_1, \cdots, R_k | \theta) = \begin{pmatrix} \Sigma_{k-1} & \mathbf{b}_k \\ \mathbf{b}_k^T & a_k \end{pmatrix}, \tag{8}$$

where $\mathbf{b}_k := E[(R_k - \theta)[R_1 - \theta \ \cdots \ R_{k-1} - \theta]^T | \theta]$, and $a_k := \text{Var}(R_k | \theta)$ for $k = 2, \cdots, M$. We can derive $\boldsymbol{\alpha}_k$ in (3) to have an one-to-one relationship with $(a_k, \mathbf{b}_k)$ by

$$\begin{aligned} &\mathbf{b}_k = (\Sigma_{k-1} - \text{diag}(N_1, N_2, \cdots, N_{k-1})) \boldsymbol{\alpha}_k, \\ &a_k = \beta_k^2 \sigma_k^2 + N_k + \\ &\quad \boldsymbol{\alpha}_k^T (\Sigma_{k-1} + \text{diag}(s_{1;k}^2 - N_1, \cdots, s_{k-1;k}^2 - N_{k-1})) \boldsymbol{\alpha}_k. \end{aligned} \tag{9}$$

Note that $\Sigma_M = \Sigma$, so the optimal fusion rules for S_k can be derived as an iterative series of optimization problems, by applying matrix block-wise inversion formula,

$$\begin{aligned} \mathbf{1}_k^T \Sigma_k^{-1} \mathbf{1}_k &= \mathbf{1}_{k-1}^T \Sigma_{k-1}^{-1} \mathbf{1}_{k-1} + \frac{(\mathbf{b}_k^T \Sigma_{k-1}^{-1} \mathbf{1}_{k-1} - 1)^2}{a_k - \mathbf{b}_k^T \Sigma_{k-1}^{-1} \mathbf{b}_k} \\ &= (\sigma_1^2 + N_1)^{-1} + \sum_{i=2}^{k-1} \frac{(\mathbf{b}_i^T \Sigma_{i-1}^{-1} \mathbf{1}_{i-1} - 1)^2}{a_i - \mathbf{b}_i^T \Sigma_{i-1}^{-1} \mathbf{b}_i}. \end{aligned} \tag{10}$$

In (10), since all $\{\Sigma_k\}_{k=1}^M$ are positive definite matrices, the denominator $a_i - \mathbf{b}_i^T \Sigma_{i-1}^{-1} \mathbf{b}_i$ in (10) is always positive. Due to the transmission order, Σ_{k-1} is invariant to U_k, so by assuming S_k locally knows Σ_{k-1}, the objective of S_k is to control $\boldsymbol{\alpha}_k$ (thus control a_k and $\mathbf{b}_k$) to maximize the following (11),

$$\max_{a_k, \mathbf{b}_k} \frac{(\mathbf{b}_k^T \Sigma_{k-1}^{-1} \mathbf{1}_{k-1} - 1)^2}{a_k - \mathbf{b}_k^T \Sigma_{k-1}^{-1} \mathbf{b}_k}, \quad \text{given } \Sigma_{k-1} \tag{11}$$

Therefore, combined with (9), $\boldsymbol{\alpha}_k$ can be obtained by solving (11), and β_k is given by (4). Now, let us find out why these fusion rules result in global optimality of $\hat{\theta}$.

B. Local Optimality Implies Global Optimality

For any parent node S_0 and child node S_k, let us consider S_k further has child nodes $S_{k1}, \cdots, S_{kM}$. Since overhearing links only exist within layer, the following Markov chain holds,

$$X_{kj}^* \to X_k^* \to X_0^*, \text{ for any } k, j = 1, \cdots, M. \tag{12}$$

Therefore, the only way the node S_{kj} can affect the performance of X_0^* is through minimizing the distortion of X_k^*, which is exactly what the above fusion rules achieve.

IV. Topology Compression Algorithm

Since the optimal fusion rules (1) are derived, we are ready to develop topology compression algorithm (2). We assume that DA knows all the channel conditions of σ_k^2, N_k and $s_{i;k}^2$ for any S_i, S_k in the whole network G_o. As we will show, overhearing significantly reduces necessary sensors, therefore this is not a strict assumption.

Topology Compression Algorithm
For $k = N : -1 : 0$, (bottom to root)
Calculate each $\sigma^{2*}_{(k)}$ without overhearing;
Set distortion constraint $D = \sigma^{2*}_{(0)}$;
Set unnecessary sensor set $V_{\text{off}} = \phi$;
For $k = N : -1 : 0$, (bottom to root)
Calculate each $\sigma^{2*}_{(k)}$ with overhearing;
For $k = 1 : +1 : N$, (root to bottom)
If $S_{(k)}$ has an ancestor in V_{off}
$V_{\text{off}} = V_{\text{off}} \cup S_{(k)}$;
Set $\sigma^{2*}_{(k)} = \infty$;
Else (Try to turn of $S_{(k)}$)
Store $\{\sigma^{2*}_{(i)}\}_{i=0}^{k}$ into Temp $= \{\sigma^{2*}_{(i)}\}_{i=0}^{k}$;
Turn off $S_{(k)}$ by Set $\sigma^{2*}_{(k)} = \infty$;
For $j = k-1 : -1 : 0$, (bottom to root)
If $S_{(j)} \in V_{\text{off}}$
Set $\sigma^{2*}_{(j)} = \infty$;
Else
Calculate each $\sigma^{2*}_{(j)}$ with overhearing;
If distortion constraint is satisfied ($\sigma^{2*}_{(0)} < D$)
$V_{\text{off}} = V_{\text{off}} \cup S_{(k)}$;
Else (Transmission of $S_{(k)}$ is turned on again)
$\{\sigma^{2*}_{(i)}\}_{i=0}^{k} =$ Temp;

A. Basic Topology Compression: Exhaustive Search

Consider there are N sensors in the network. If DA wishes to identify minimal necessary and connected topology as in (2), then there are $O(2^N)$ possible candidates. However, the objective function $r(G_o^{'})$ in (2) is not convex. Therefore, it is unwise to exhaustively search all $O(2^N)$ topology, since it spends exponential time computational complexity. Instead, it is necessary to propose the following topology compression algorithm that has complexity $O(N^2M^3)$ and performance nearly as good as exhaustive search (as shown in numericals).

B. Topology Compression Algorithm

Without loss of generality, let $S^{(k)}$ denotes the $(N-k)^{th}$ sensor in the transmission order, and $S^{(0)}$ denotes the DA. Therefore, the sensors transmit according to the decrement of (k), and DA is the last one to transmit. Denote $\sigma^{2*}_{(k)}$ as the conditional variance (7) of the fused observation of $S^{(k)}$. The algorithm works as follows: top-down (from DA to leaf nodes) turns off transmission of one sensor at a time and check if the Bayes' risk $r(G_o^{'})$ satisfies the distortion constraint D. If transmission of $S_{(k)}$ is turned off and the distortion constraint D is satisfied, then transmission of $S_{(k)}$ remains turned off and so are all the offsprings of $S_{(k)}$; otherwise the transmission of $S_{(k)}$ is turned on again. This algorithm has computational complexity of order $O(N^2M^3)$, because transmission of all nodes $\{S_k\}_{k=1}^{N}$ are traversed to be turned on or off, each time N-many $\{\sigma_j^{2*}\}_{j=1}^{N}$ are calculated with complexity lower bounded by matrix inversion (in the order of M^3.)

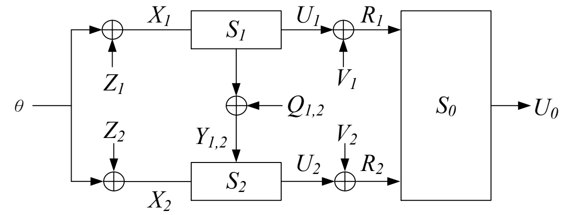


Fig. 3: 2 sensors 1 DA scenario - basic element of binary tree

V. Numerical Results

In this section we provide quantitative numerical results to measure the performance of traffic compression methodology. We focus on sensor energy saving in WSN via the proposed traffic compression methodology.

A. Traffic Compression Methodology for Energy Efficiency

Consider each sensor in a WSN can provide bit-error rate $P_e \leq 10^{-3}$, which corresponds to a signal-to-noise-ratio (SNR) above 7dB in Gaussian channels using binary phase shift keying (BPSK). Therefore, the signal-to-quantization-noise-ratio (SQNR) is dominated by quantization errors. Further, suppose sensors use high resolution quantization (small quantization level), so the noises that are introduced during sensors' transmission can be well modeled by additive white Gaussian noise (AWGN). In addition, since real world noise powers may perturb, we consider the observation and quantization noise powers are all exponential random variables $\sigma_k^2 \sim \exp(1)$, $N_k \sim \exp(0.01)$ and $s_{i,k}^2 \sim \exp(0.01)$, for all i, k. The *a priori* distribution is $\theta \sim N(0, 1)$.

We also assume that sensor's computation and receiving powers are negligible comparing to its transmission power. Each sensor uses a fixed power for transmission, hence the necessary number of sensors is proportional to overall network energy consumption. Consider the network transmission topology be complete binary tree for basic topology construction algorithm, where the basic elements are shown in Fig. 3.

The fusion rule α for S_2 maximizes the following equation,

$$f(\alpha) = \mathbf{1}_2^T \Sigma^{-1} \mathbf{1}_2 = \frac{u_1\alpha^2 - 2u_2\alpha + u_3}{u_4\alpha^2 - 2u_5\alpha + u_6},$$

and $u_1 = \sigma_1^2 + \sigma_2^2 + s_{1,2}^2$, $u_2 = \sigma_1^2 + \sigma_2^2$, $u_3 = \sigma_1^2 + \sigma_2^2 + N_1 + N_2$, $u_4 = \sigma_1^2 s_{1,2}^2 + \sigma_1^2\sigma_2^2 + \sigma_1^2 N_1 + s_{1,2}^2 N_1 + \sigma_2^2 N_1$, $u_5 = \sigma_1^2\sigma_2^2 + N_1\sigma_2^2$, $u_6 = \sigma_1^2\sigma_2^2 + \sigma_1^2 N_1 + \sigma_2^2 N_1 + N_1 N_2$. Note that the denominator $u_4\alpha^2 - 2u_5\alpha + u_6$ is always greater than 0. We can easily calculate the corresponding α as,

$$\alpha = \frac{u_1u_6 - u_3u_4 + \sqrt{(u_3u_4 - u_1u_6)^2 - 4(u_1u_5 - u_3u_5)}}{2(u_1u_5 - u_2u_4)}.$$

Fig. 4 shows the necessary number of sensors versus achieved mean squared error (MSE), with/without overhearing. Under the same level of MSE, we see that necessary sensors are significantly reduced. For instance, in Fig. 4, 35.4% of sensors are unnecessary to achieve an MSE of 2×10^{-2}.

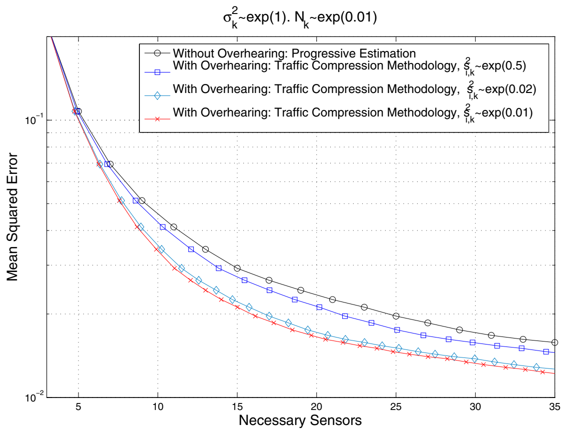


Fig. 4: Traffic compression on complete binary tree. The 2-steps traffic compression methodology significantly reduces necessary sensors to achieve energy efficient communication.

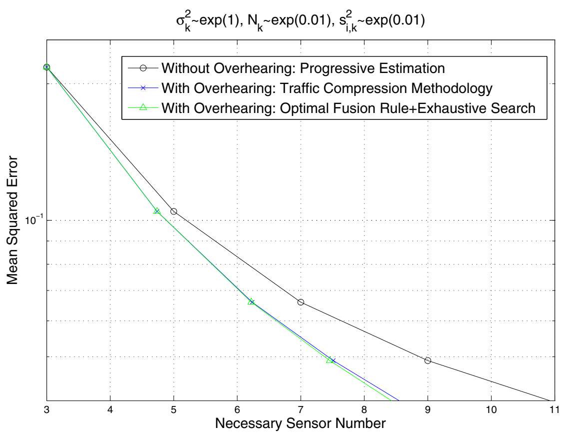


Fig. 5: Traffic compression algorithm performs nearly as optimal as exhaustive search.

As suggested by numerical, the reduction ratio (number of reduced sensors versus number of total sensors) is dominated by the channel conditions near the DA, i.e., the reduction ratio reaches a constant as N approaches infinity. Finally, in situations of larger SQNR ($s_{i,k}^2 \sim \exp(0.02)$ and $s_{i,k}^2 \sim \exp(0.5)$ in Fig. 4), we see that the traffic compression methodology exhibits robustness in that it still reduces topology. The difference for the larger SQNR is the more asymptotic performance to the case without overhearing.

B. Topology Compression Algorithm and Exhaustive Search

We also compare the topology compression algorithm with exhaustive search. In Fig. 5, when sensor numbers are few (below 11), topology compression algorithm performs very similar to exhaustive search. Moreover, leveraging overhearing, we do not need many sensors to reach the desired distortion constraint, therefore the sufficiency of topology compression algorithm is demonstrated.

VI. Conclusion

In large Machine-to-Machine (M2M) network of tremendous wireless devices, spectrum scarcity and energy efficiency are critical problems. One potential solution is to identify redundant signals and cognitively compress traffic in the network. We successfully propose a mechanism that utilizes the broadcast nature of wireless communications and the signal correlations among sensors to cognitively compress traffic. The mechanism consists of optimal fusion rules and topology compression algorithm such that the system can dynamically turn on necessary devices' transmission to achieve a desired estimation quality. This mechanism cognitively and significantly conserves energy and lowers sensors deployment cost to suggest a new design paradigm in M2M communications.

References

[1] S.-Y. Lien, K.-C. Chen, and Y. Lin, "Toward ubiquitous massive accesses in 3gpp machine-to-machine communications," *IEEE Communication Magazine*, vol. 49, no. 4, April 2011.

[2] G. Wu, S. Talwar, K. Johnsson, N. Himayat, and K. D. Johnson, "M2m: From mobile to embedded internet," *IEEE Communication Magazine*, vol. 49, no. 4, April 2011.

[3] M. Seyedi, B. Kibret, D. T. H. Lai, and M. Faulkner, "A survey on intrabody communications for body area network applications," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 8, August 2013.

[4] F.-M. Tseng, C.-H. Lin, and K.-C. Chen, "In-network computations of machine-to-machine communications for wireless robotics," *Wireless Personal Communications*, vol. 70, pp. 1097–1119, June 2013.

[5] K.-C. Chen and S.-Y. Lien, "Machine-to-machine communications: Technologies and challenges," *Ad Hoc Netw.*, p. http://dx.doi.org/10.1016/j.adhoc.2013.03.007, 2013.

[6] M. C. Vuran and I. F. Akyildiz, "Spatial correlation-based collaborative medium access control in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 14, no. 2, April 2006.

[7] I. F. Akyildiz and et el., "A survey on sensor networks," *IEEE Communication Magazine*, vol. 40, no. 8, August 2002.

[8] E. C.-H. Ngai, E. Gelenbe, and G. Humber, "Information-aware traffic reduction for wireless sensor networks," *IEEE 34th Conference on Local Computer Networks*, October 2009.

[9] Z. Huang and J. Zheng, "A slepian-wolf coding based energy-efficient clustering algorithm for data aggregation in wireless sensor networks," *IEEE ICC 2012*, 2012.

[10] K. Rajawat and A. Cano, "Network-compressive coding for wireless sensors with correlated data," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, December 2012.

[11] Y. Hua and Y. Huang, "Progressive estimation and detection," *SenSIP Workshop*, May 2008.

[12] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE journals on selected areas in communications*, vol. 23, pp. 201–220, Februrary 2005.

[13] A. D. WYNER and J. ZIV, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, 1976.

[14] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder-ii: General sources," *IEEE Transactions on Information Theory*, 1978.

[15] I. H. Wang and D. N. C. Tse, "Interference mitigation through limited transmitter cooperation," *IEEE International Symposium on Information Theory*, June 2010.

[16] J. Z. Sun and V. K. Goyal, "Intersensor collaboration in distributed quantization networks," *IEEE Transactions on Communication*, vol. 61, no. 9, September 2013.

[17] V. Prabhakaran, D. Tse, and K. Ramchandran, "Rate region of the quadratic gaussian ceo problem," *IEEE International Symposium on Information Theory*, June 27-July 2 2004.

[18] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Transactions on Information Theory*, vol. 44, no. 6, October 1998.