Average Quantized Consensus Building by Gossip Algorithm using 16 Bit Quantization and Efficient Data Transfer Method

Imtiaz Parvez, Nattapong Chotikorn, and Arif Islam Sarwat

Abstract**—**Average consensus building is a critical problem in wireless sensor networks specially with power and bandwidth constraints. This work studies how 16 bit quantization compared to integer approximation of node value, can reduce the mean square error on average consensus building by gossip algorithm using digital (quantized) communication. Moreover, we propose an efficient data transfer method for that the network reaches to consensus and saves significant amount of bandwidth. At the same time, the speed of convergence is estimated.

Keywords**—**Gossip Algorithm, Quantized Consensus, Quantized Communication, Digital Communication, Wireless Sensor Network etc.

I. Introduction

CONSENSUS problem appears in nature and engineering sectors, and has received extensive study. In average consensus building of wireless sensor network, the nodes communicate among themselves, make average and exchange data in the distributive and iterative way. However, most of the work have been done based on the assumption that the nodes exchange real number and the network is ideal (no power and bandwidth constraints).

Let us consider, an undirected graph ξ with set of nodes $v = \{1, 2, ... \mathrm{N}\}$ of N nodes and a real quantity $x_i \in \mathbb{R}$ for every node $i \in v$. In average consensus by gossip algorithm, at each time two randomly chosen nodes (i, j) exchange their values, make average and finally reach to consensus [1]. In some algorithm, a randomly selected node makes average with its two neighbors [2]. Most of the literature assume that the communication channel between nodes allows real number to transfer. But from practical point of view, the sensor network communicates by wireless fashion, and energy and bandwidth constraints limit the capacity of the channel. This suggests that the communication channel rather as digital channel. This clearly forces a quantization on real number that nodes have to send. This issues have been noticed in [3]- [7].

In [7] and [8], the quantization effect was examined from different points of view restricting the attention on integer value, and it introduced the definition of quantized communication which is attained by the vector x if it belongs to the set S defined as

$$S = \{x : \{x\}^N \in \{L, L+1\}, \sum_{i=1}^{N} x_i = T\}$$ with T and L being the sum of initial node and an integer. It was shown that under some constraints and restricted communication, the network reaches to consensus. But the consensus is not clearly strict sense consensus, all nodes may not have exact same value.

However, the use of quantized communication complicates the convergence. This is because of not preserving the initial state. This shortcoming was solved by [6] . To uphold the average at each iteration, the evolution matrix is assumed doubly stochastic at each time.

The quantization method in [4] takes the closest integer value adding some constant and the network is converged into consensus upto the size of the quantization bin for any connected undirected graph. In this study, we propose 16 bit quantization for node value. Since we are reducing resolution size of the quantization, 16 bit represents the node value more accurately compared to integer approximation. Also we propose the optimization of bit transmission. We propose a cycle consists of 5 times bit transmission. We transmit 8 bit MSB (most significant bit) of the 16 bit index through 1st to 4th times and whole 16 bit at the 5th time of a cycle. At the receiving end, from the index, the node value is reconstructed using codebook as like pulse code modulation. For optimizing of bit transmission, the network also converges to average consensus and the mean square error from average is reduced significantly as compared to integer approximation of the node value. At the same time, a significant amount of bandwidth and energy is saved.

This work was supported in part through FIU Graduate School Presidential Fellowship, funded by Graduate School, Florida International University , FL 33174, USA

Imtiaz Parvez is with the Electrical and Computer Engineering Department, Florida International University , FL 33174, USA (e-mail: iparv001@fiu.edu).

Nattapong Chotikorn is with Mahidol University, Thailand.. (e-mail:nattapong1992@hotmail.com).

Arif Islam Sarwat is with the Electrical and Computer Engineering Department, Florida International University, FL 33174, USA (e-mail: asarwat@fiu.edu).

In the section III, the proof of convergence is shown while in the section IV and V, convergence speed and simulation results are provided respectively.

II. STATEMENT

Let us consider, a network of N nodes specified by graph $\xi = (v, \varepsilon)$ where v be the set of vertices (nodes) and ε be the set of edges, which is subset of $\{\{i, j\} : \{i, j\} \in v, i \neq j\}$. The nodes are connected by links. Since the graph is connected, there is a path between any node i and j. Therefore, a path in ξ consists of in a sequence of vertices $(i = i_1, i_2, ...i_r = j)$ such that $\{i_j, i_{j+1}\} \in \xi$ for every $j \in \{1, 2, 3.......r-1\}$. A graph is fully connected or complete if $\varepsilon = \{\{i, j\} : \{i, j\} \in v, i \neq j\}$. Each node is given a time dependent finite value $x_i(t)$ such that $x_i(t) \in \mathbb{R}$. Since the nodes are connected by digital communication links, they cannot extract the real value but the quantized (16 bit) value of exact value. Since ξ is a undirected graph of N nodes, at each iteration, an edge $(i, j) \in v$ is selected randomly with probability $P_{(i,j)}$ such that

$$\sum_{(i,j)\in v} P_{(i,j)} = 1$$

Let, the probability distribution, $P \in \mathbb{R}^{N\times N}$ be, then

$$P_{ij} = P_{ji} = \begin{cases} P_{(i,j)} ....if\ .i \neq j \\ 0..............otherwise \end{cases}$$

P_{ij} and P_{ji} are the probability of selecting edge (i, j) and (j, i) respectively.

The randomly chosen nodes adjourn their states by the below formula:

$$x_i(t+1) = x_i(t) - \alpha q[x_i(t)] + \alpha q[x_j(t)]$$
$$x_i(t+1) = x_j(t) - \alpha q[x_j(t)] + \alpha q[x_i(t)]$$

Where, $q[x(t)] = 16$ bit quantization of $x(t)$ and $\alpha = \frac{1}{2}$.

Now we define, quantized average consensus. A quantized average consensus is a state $\overline{x} = \mathbb{R}^N$ such that $|\overline{x}_i - N^{-1}\sum_{j=1}^{N} x_j(0)| < 1$. If for any time t, there exists a convergence time T_{con} such that $t > T_{con}$, then the network is said converged into consensus and x(t) is a quantized average consensus state.

III. QUANTIZATION METHOD

Let's consider , a scalar $x_i \in \mathbb{R}$ such that bounded by finite interval $[U, -U]$. Further suppose that we want to get quantized message with length of l bits. So, there are 2^l quantization points given by the set $\tau = \{\tau_1, \tau_2, ...\tau_L\}$. This points are uniformly spaced such that $\Delta = (\tau_{j+1} - \tau_j)$ for $j \in \{1, 2, 3.......L-1\}$. It follows that $\Delta = [\frac{2U}{2^l - 1}]$. Now suppose $x_i \in (\tau_{j-1} - \tau_{j+1})$, then $q(x_i) = \tau_j$.

IV. CONVERGENCE

Let us consider, a connected undirected graph with N nodes as $v = \{1, 2........N\}$ and a real quantity $x_i \in \mathbb{R}$ for every node $i \in v$. Let us define, at time t

Minimum node value, $m[t] = \min_i x_i(t)$

Maximum node value, $M[t] = \max_i x_i(t)$

Difference between maximum and minimum node value, $D[t]_{\max} = |M[t] - m[t]|$

Based on the above definition and $D[t]_{\max} \geq \Delta$, we define, the updating rule:

(I) $x_i(t+1) + x_j(t+1) = x_i(t) + x_j(t)$

(II) If the difference between node i and j, $D_{ij}(t) > \Delta$, then $D_{ij}(t+1) < D_{ij}(t)$

(III) If $D_{ij}(t) = \Delta$ (without loss of generality) and $x_i(t) > x_j(t)$, then $x_i(t+1) = x_j(t)$ and $x_j(t+1) = x_i(t)$. This is called swap rule.

At first, we will see the proof of convergence for integer value. Then we will prove that for quantized value of node, the network also converges to consensus.

Since in the first time cycle, the node values are constant and $\sum_{j=1}^{N} x_j(0) = \sum_{j=1}^{N} x_j(t)$ for any time t, we assume below property for the graph ξ,

(1) For any initial value x[0], at any time during the execution of the algorithm, the node value lies in the finite set χ (which may depend on x[0]).

(2) For any state x[t]=x, there exists finite time τ_x such that probability of $x[t+\tau_x]$ given at $x(t)$, $P[x(t+\tau_x) \in S \mid x(t)] > 0$ where S is the set of all vectors which have quantized consensus distribution.

(3) If $x[t] \in S$, then $x[t'] \in S$ for $t' > t$.

Proof: Since χ finite and from property[2], let minimum probability, $\Lambda = \min_{x\in\chi} P[x(t+\tau_x) \in S \mid x(t)] > 0$ and $T = \max_x \tau_x$ is finite.

From property of [3], its follows that

Probability, $P[x(t+T) \notin S \mid x(t) \notin S] \le (1-\Lambda)$ .So

Probability, $P[x(t+T) \notin S \mid x(t) \notin S] \le (1-\Lambda)^{\lfloor t/T \rfloor}$

Its converges to 0 as $t \to \infty$. So the network reaches to consensus for integer node value.

Now, we will prove convergence for quantized node value:

Let's, $D[t]_{\max} > 2\Delta$ and the number of nodes having maximum value $N_{\max}[t] > 1$.

Now let, $\Im(t)$ be the set of nodes which have value $M(t)-2\Delta$ or less and $\psi(t)$ have the value $M(t)$. Since the difference between maximum and minimum node value $D(t) > 0$, $\Im(t)$ will not be an empty set. Let us consider, a path $P = \{v_1, v_2 ....v_p\}$ and $v_1 = M(t) - 2\Delta$ and $v_p = M(t)$. Since ξ is connected, this path exists. Now the path will be shortest if $v_2, v_3 ....v_{p-1}$ have value $M(t)-\Delta$. Since during the communication, each edge has equal positive probability to be selected, in the (p-1) units time following t, the edges are selected sequentially starting from $\{v_1, v_2\}$. At the time t, v_1 and v_2 will swap their value by rule (3). At the last step, v_{p-1} and v_p will swap their value. In quantized gossip algorithm, this update will make both nodes to have value less than $M(t)$. As for example, let us consider quantization resolution is $\Delta = 1, v_1 = 1$ and $v_p = 3$. For the shortest path condition, $v_2, v_3 ....v_{p-1}$ will have value 2. By the swap rule, v_1 and v_2 will have value 2 and 1 respectively. By the same procedure, v_p will have value 2. So the maximum node value is decreased by 1.

So, $P[N_{\max}(t+l_p) < N_{\max}(t)] > 0$ where l_p is the number of edge in the path P.

If $N_{\max}[t] = 1$, for the above reason, there is also positive probability that at $t' = t + l_p$, the node with maximum value $M(t)$ will make average with $M(t)-2\Delta$. So, the maximum value of the node is decreased by at least Δ and therefore $D[t'] < D[t]$ and probability, $P[D[t'] < D[t]] > 0$.

Now consider, the following sequence of times $t_0 = t, t_1, t_2 ....$. For each $i \ge 0$, if $N_{\max}[t_i] > 1$, then we consider t_{i+1} be the first time when there is a positive probability that $N_{\max}[t_{i+1}] > N_{\max}[t_i]$. Now since D is constant, for any integer $k \in \square$, $D[t] < D[t_0]$ or $N_{\max}[t_k] = 1$ with positive probability. For $t_{k+1} \ge t_k$ it follows that $D[t_{k+1}] < D[t_k]$. So proof is completed.

V. Convergence Time

Gossip Algorithm with real value reaches to consensus exponentially fast and for averaging time $\varepsilon(0 < \varepsilon < 1)$, its convergence time T_ε

$T_\varepsilon = \sup_{x(o)} \text{info}\{t : P(\frac{\|x(t) - x_{ave}1\|_2}{\|x(0)\|_2}) \ge \varepsilon) \le \varepsilon\}^1$ has been estimated in [3], [8]and [9]. Thus the averaging time ε is the smallest time it takes for $x(.)$ to get within ε of $x_{ave}1$ with high probability, regardless of the initial value x(0).

For equal probability of selecting any edge

- $T_\varepsilon = \Theta(N)$ for $N \to \infty$ for complete graph
- $T_\varepsilon = \Theta(N^3)$ for $N \to \infty$ for ring graph

Now we see the probabilistic analysis for the bounds of convergence time. Let, $T_1(x)$ is the random variable indicating the time of first non-trivial averaging when $x[0] = x$. For the given graph ξ and the given probability distribution on the edge, lets define $\bar{T}(\xi) = \max_x E[T_1(x)]$, where the maximum expectation is based over all possible initialization x that are not quantized consensus distribution and for which $m \le x_i \le M$ for $i = 1, 2, 3.....N$. Since for all such initializations, the minimum and maximum number of non-trivial averagings are 1 and $\frac{(M-m)N}{8}$ respectively, it follows that the expectation of final convergence time is

$$T_{con} = \bar{T}(\xi) \le \max_{x: m \le x_i \le M} E[T_{con}(x)] \le \frac{(M-m)^2 N \bar{T}(\xi)}{8}$$

The upper bound can be found with high probability as shown in [6]. This is very conservative method for simulation.

Let us define, a function of x(t),

$V[(x)] = x^*(t)\Omega x(t) = (\|x(t) - x_{ave}\|_2)^2$

where, $\Omega = I - N^{-1}11^*$, $x^*(t)$ is the transpose of $x(t)$ and I is the identity matrix.

Now for any node (i, j),

$\sum_{(i,j)} P_{(i,j)}(x_i(t) - x_j(t))^2 = 2x^*(t)(\text{diag}(P1) - P)x(t)$

where, $P_{(i,j)}$ is the probability of selecting edge (i, j) and P is the probability distribution on all edge.

Using the result from [4], the expected value of V at time

$(t+1)$,

$$E[v(x(t+1))] \le (1-\lambda)E[v(x(t))] + \Delta$$

where λ is the smallest eigen value of $(\operatorname{diag}(P1) - P)$ and Δ is the resolution of 16 bit quantization.

For recurrent operation, we can argue that

$$E[v(x(t+1))] \le (1-\lambda)^t E[v(x(t))] + \frac{1-(1-\lambda)^t}{\lambda}\Delta$$

and then $E[v(x(t))] \le (1-\lambda)^t E[v(x(0))] + \frac{1}{\lambda}\Delta$

So, $v(x(t))$ decreases to $(1-\lambda)$ rate initially and saturates to $\frac{\Delta}{\lambda}$. Notice that λ is the function of P and topology of the graph.

For any edge $\{i, j\} \in \varepsilon$, $P = \frac{1}{|\varepsilon|}$ where $|\varepsilon|$ is the cardinality of ξ.

So $\operatorname{diag}(P1) - P = \frac{1}{|\varepsilon|} L_\xi$ where L_ξ is the Laplacian matrix of graph ε.

As shown in [1], for complete graph $\lambda = \frac{2}{N-1}$ and for ring graph $\lambda = \frac{2}{N}(1-\cos(\frac{2\pi}{N}))$ and then $\lambda = \frac{4\pi^2}{N^3} + o(\frac{1}{N^3})$ for $N \to \infty$.

VI. Simulation Result

In most of the literature, the node value is assumed to be the nearest approximate integer value which demonstrates large mean square error from average value and the network is converged to consensus in weaker sense. The quantization method introduced in [1] leads degradation $|q(x) - x| \le \frac{1}{2}$ where, $x \in$ □. But in our algorithm, degradation $|q(x) - x| \le \Delta$. For the smaller step size, in our quantization method, the mean square error is reduced significantly, which is reflected in simulation result.

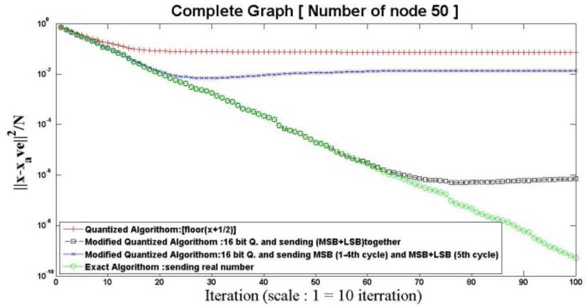

Fig. 1 The mean square error of Complete Graph for N=50 nodes. Remark that with increase of iteration, the mean square error in algorithm sending real number < Algorithm with 16 bit quantization < Algorithm with 16 bit quantization (sending MSB 1-4 cycle and (MSB+LSB) in 5 cycle) < Algorithm sending approximate integer value($\lfloor x + \frac{1}{2} \rfloor$).

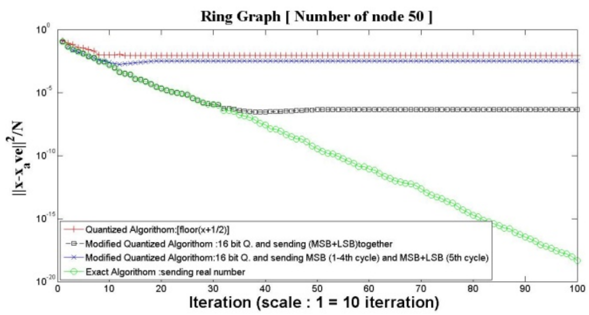

Fig. 2 The mean square error of Ring graph for N=50 nodes. Remark that with increase of iteration, the mean square error in algorithm sending real number < Algorithm with 16 bit quantization < Algorithm with 16 bit quantization (sending MSB 1-4 cycle and (MSB+LSB) in 5 cycle) < Algorithm sending approximate integer value($\lfloor x + \frac{1}{2} \rfloor$).

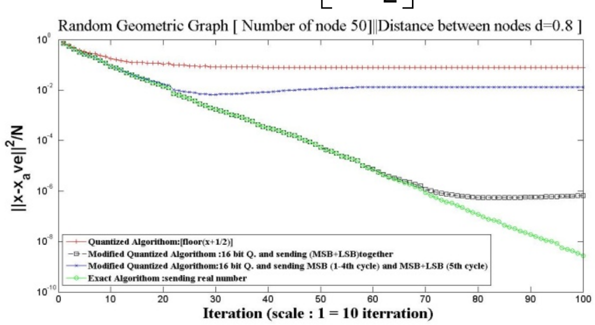

Fig. 3 The mean square error of Random geometric graph for N=50 nodes. Remark that with increase of iteration, the mean square error in algorithm sending real number < Algorithm with 16 bit quantization < Algorithm with 16 bit quantization (sending MSB 1-4 cycle and (MSB+LSB) in 5 cycle) < Algorithm sending approximate integer value($\lfloor x + \frac{1}{2} \rfloor$).

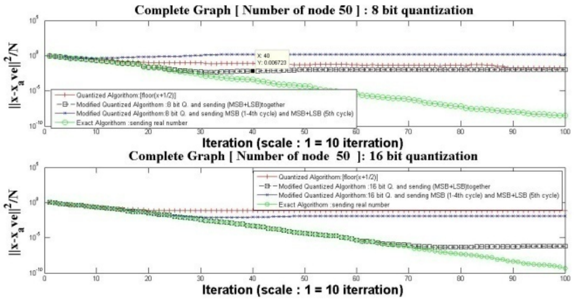

Fig. 4 Effect of bit number on Consensus. Remark that for only 8 bit data communication and based on our definition, the nodes do not converge to consensus

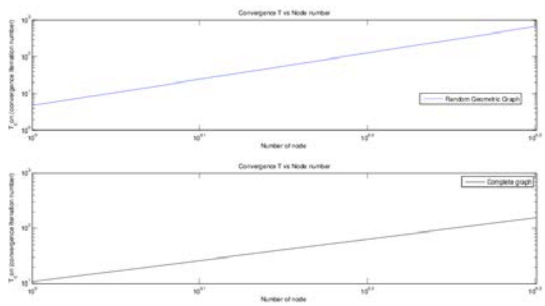

Fig. 5 Convergence Time Vs Node number. Note that with the increase of node number, the convergence time also increases.

From the simulation graphs- Fig. 1 , Fig. 2, Fig 3, we remark that for all method, the networks reach to consensus based on definition of our consensus $\left|\bar{x}_i - \frac{1}{N}\sum_{i}^{N} x_i\right| < 1$. After some iterations, the mean square error for all quantization methods become constant. We also observe that the mean square error for real number is the least and the mean square error for quantization method $\left\lfloor x + \frac{1}{2} \right\rfloor$ is the greatest. At the same time, mean square error for the algorithm with 16 bit (sending MSB in 1-4 cycle & MSB+LSB in 5 cycle) is greater than the algorithm with all through 16 bit quantization but less than the algorithm with quantization method $\left\lfloor x + \frac{1}{2} \right\rfloor$.

We also find that for all through 8 bit data transfer, the network does not reach to consensus. However, for 16 bit convergence is attained which is reflected on Fig. 4. So we need at least 16 bit for consensus building in our proposed data transmission method based on definition of consensus.

With the increase of number of nodes, the convergence time also increases as shown in Fig.5. The relation between convergence time and node is linear (i.e. $T_\varepsilon = \Theta(N)$ for $N \to \infty$).

VII. Bandwidth Save Calculation

In our algorithm, we define a cycle consists of 5 times bit transmission.

We send 8bit MSB of the 16 bit quantized values through 1st to 4th time and whole 16 bit in the 5th time.

In the 1st - 4th time, 50% bandwidth is saved. In the 5th time, 100% bandwidth is utilized.

So, total bandwidth save in the period of a cycle =

$$\frac{0.5 \times 4 + 0}{5} \times 100 = 40\%$$

VIII. Conclusion

From the simulation result, we can conclude that we need at least 16 bit quantization on our data transmission technique to attain consensus. It is also evident that as long as the distance from consensus is much larger than the quantization step, the speed of convergence is almost same as the non-quantized algorithm. Therefore, when we are near the agreement, the granularity effect comes out so that a full understanding of algorithm is based on non-quantization approximation and analysis of integer dynamics.

Since in our algorithm, the optimization of data transfer by sending MSB (most significant bit) and whole 16bit separately, saves 40% bandwidth, it can have great application on large network design specially when nodes are very far from each other, nodes add/drop frequently, and bandwidth and energy constraints is a major issue.

References

[1] P. Frasca, R. Carli, F. Fagnani, and S. Zampieri, "Average consensus by gossipalgorithm using quantized communication,*"47th IEEE Conference on Decision and Control,* 2008.

[2] B. Yang, W. Wu, and G. Zhu., "Distributed averaging in wirelesssensor networks with triple wise gossip algorithm," *TENCONSpring Conference*, 2013 IEEE, vol. 178 - 182, April 2013.

[3] J. Lavaei and R. M. Murray, "Quantized consensus by meansof gossip algorithm," *Automatic Control, IEEE Transactions on*,vol. 57, pp. 19 – 32, January 2012. http://dx.doi.org/10.1109/TAC.2011.2160593

[4] R. Hu, J. Sopena, L. Arantes, P. Sens, and I. Demeure, "Fair comparison of gossip algorithms over large-scale random topologies,"Reliable Distributed Systems (SRDS), 2012 IEEE 31st Symposiumon, vol. 331 - 340, October 2012.

[5] L. Xiao, S. Boyd, and S. Kim, "Distributed average consensuswith least-mean-square deviation," in Journal of Parallel andDistributed Computing, Pasadena, CA 91125-9300, USA, August2006, pp. 67(1):33–46.

[6] R. Carli, F. Fagnani, P. Frasca, T. Taylor, and S. Zampieri,"Average consensus on networks with transmission noise orquantization," In ECC 07, Kos, p. pages 18521857, August 2007.

[7] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," vol. In IPSN 05, LosAngeles, 2005.

[8] A. Kashyap, T. Basar, and R. Srikant, "Quantized consensus,"2007, p. 43:11921203.

[9] A. Boyd, S.andGhosh, B. Prabhakar, and D. Shah, "Randomizedgossip algorithms," *Information Theory, IEEE Transactions on*,pp. 2508 – 2530, 2006.

[10] S. Shang, P. P. H. Cuff, and S. Kulkarni, "An upper bound onthe convergence time for quantized consensus," *INFOCOM, 2013 Proceedings IEEE*, vol. 600 - 604, no. 19 - 32, April 2013.

[11] F. Fagnani and S. Zampieri, "Randomized consensus algorithms over large scale networks," In Proceedings Information Theoryand Application Workshop, 2007.