URSI GASS 2021, Rome, Italy, 28 August - 4 September 2021

Improving Age-of-Information in Distributed Vehicle Tracking

Albin Severinson[(1,2)], Eirik Rosnes[(1)], and Alexandre Graell i Amat[(3,1)]
(1) Simula UiB, Bergen, Norway
(2) Department of Informatics, University of Bergen, Bergen, Norway
(3) Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden

Abstract

We consider the problem of tracking vehicles in a distributed setting, which is important for, e.g., autonomous driving and collision avoidance systems. These applications rely on receiving timely updates and require aggregating sensor data from many sources to improve accuracy. We consider a cloud-assisted scheme that utilizes replication to alleviate the straggler problem, i.e., the problem of random delays in distributed systems. We derive the *age-of-information* (AoI) of estimate updates and show that replication significantly improves the AoI. Furthermore, we derive the probability that the error of the position estimate exceeds some threshold for a given AoI.

1 Introduction

Accurately tracking position is critical for many applications in intelligent transportation systems. In many cases these applications rely on receiving timely updates to operate safely, as is the case for, e.g., autonomous driving and collision avoidance systems [1], and require merging sensor data from multiple sources to improve accuracy [2]. In several recent works timeliness is measured by the probability distribution of the *age-of-information* (AoI) [3], defined as the difference between the current time, t, and the largest generation time of a received update, $U(t)$, i.e., the AoI is $t-U(t)$.

Several distributed tracking systems have been proposed in the literature (see, e.g., [2, 4] and references therein). These systems may both allow several nodes in a network to reach consensus on an estimate and improve accuracy by merging sensor data from multiple sources. One way is to rely on servers in the cloud for aggregation. However, the AoI may be very high in such systems, leading to low tracking accuracy [5]. One of the reasons is the problem of *straggling servers*, i.e., servers that experience transient delays due to, e.g., network congestion. For example, the 99-th percentile response time of individual servers may be several times higher than the average [6].

In [5], the authors introduced a coded distributed tracking scheme, which alleviates the straggler problem by introducing redundancy. The scheme in [5] relies on maximum distance separable codes and replication to recover from transient delays. It is shown that this can significantly improve tracking accuracy compared to the corresponding uncoded scheme.

In this paper, we consider the problem of tracking the position of multiple vehicles over time in a distributed setting, where each vehicle may observe a subset of the other vehicles. In particular, we consider a special case of the distributed tracking scheme of [5], using replication only, applied to this problem. For this scenario we answer the following two questions:

1. What is the probability distribution of the AoI of updates available to the vehicles?
2. What is the probability distribution of the position estimate error for a given AoI?

We answer the first question by deriving the probability of missing a given number of consecutive discrete update steps as a function of the update rate, the replication factor, and a parameter β characterizing the length of the transient delays (Theorem 1). We answer the second question by deriving the cumulative distribution function (CDF) of the estimate error of the position as a function of the estimate covariance matrix (Theorem 2), which can be computed numerically for a given AoI.

2 System Model

We consider the problem of tracking a set of N_{v} vehicles $\mathcal{V} = \{v_1, \ldots, v_{N_{\mathsf{v}}}\}$ in a distributed setting. As in [2], we model the state of each vehicle by a length-4 vector composed of its position and speed in the longitudinal and latitudinal directions. We model the overall system as a stochastic process that evolves over time according to

$$\boldsymbol{x}_t = \boldsymbol{F}\boldsymbol{x}_{t-1} + \boldsymbol{q}_t,$$

where $\boldsymbol{x}_t$ is the concatenation of the state vectors of all vehicles, i.e., $\boldsymbol{x}_t$ is of length $4N_{\mathsf{v}}$, $\boldsymbol{F}$ is the matrix representing the state transition model, and $\boldsymbol{q}_t$ is a noise

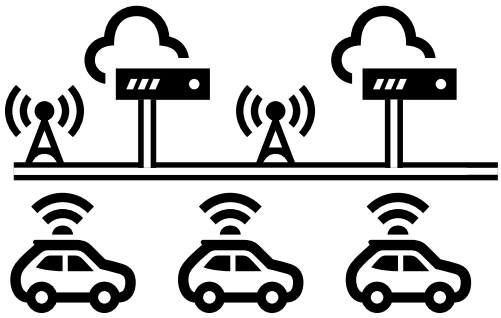

Figure 1. Cloud-assisted tracking of $N_{\mathsf{v}} = 3$ vehicles. Each vehicle uploads sensor data to $N_{\mathsf{w}} = 2$ workers responsible for aggregation and estimation.

vector drawn from a zero-mean Gaussian distribution with covariance matrix $\boldsymbol{Q}$. The estimate is updated in discrete time increments of length Δt. We measure accuracy by the absolute distance between the estimated and true position of each vehicle (see Section 3.2).

At each time step t, each of the N_{v} vehicles obtain a noisy partial observation of its own state and that of a subset of the other vehicles. The observation made by vehicle v at time t is represented by the vector

$$\boldsymbol{z}_t^{(v)} = \boldsymbol{H}^{(v)}\boldsymbol{x}_t + \boldsymbol{r}_t^{(v)},$$

where $\boldsymbol{H}^{(v)}$ is a matrix of size $h^{(v)} \times 4N_{\mathsf{v}}$ representing the observation model of vehicle v and $\boldsymbol{r}_t^{(v)}$ is a noise vector drawn from a zero-mean Gaussian distribution with covariance matrix $\boldsymbol{R}^{(v)}$. We assume that the matrices $\boldsymbol{F}$, $\boldsymbol{Q}$, $\boldsymbol{H}^{(v)}$, and $\boldsymbol{R}^{(v)}$ are known.

2.1 Distributed Tracking

To improve accuracy the observations made by the N_{v} vehicles are sent to a central party, composed of N_{w} severs, referred to as workers, $\mathcal{W} = \{w_1, \ldots, w_{N_{\mathsf{w}}}\}$, responsible for computing an aggregated estimate of $\boldsymbol{x}_t$, denoted by $\hat{\boldsymbol{x}}_t$. Each vehicle v sends the observation $\boldsymbol{z}_t^{(v)}$ to the central party at the start of each time step and the updated estimate $\hat{\boldsymbol{x}}_t$ is sent back to the vehicles at the end of each time step, where it is used, e.g., to generate collision warning messages for obstacles out of line of sight. We depict the system in Fig. 1.

2.2 Probabilistic Runtime Model

We assume that workers become unavailable for a random time after completing a computing task, which is captured by the exponential random variable V with probability density function (PDF) and CDF [7]

$$f_V(v) = \begin{cases} \frac{1}{\beta}\mathrm{e}^{-\frac{v}{\beta}} & v \geq 0 \\ 0 & v < 0 \end{cases} \quad \text{and} \quad F_V(v) = 1 - \mathrm{e}^{\frac{-v}{\beta}},$$

respectively, where β is used to scale the tail of the distribution, which accounts for transient disturbances that are at the root of the straggler problem. We refer to β as the straggling parameter.

For j independent and identically distributed random variables $V_1, \ldots, V_j$, denote by $V_{i:j}$ the i-th order statistic, i.e., the random variable associated with the i-th largest value out of $V_1, \ldots, V_j$. When $V_1, \ldots, V_j$ are exponential random variables $V_{i:j}$ is a Gamma random variable [8]. We denote by $f_{V_{i:j}}(v_{i:j})$ and $F_{V_{i:j}}(v_{i:j})$ the PDF and CDF of $V_{i:j}$, respectively.

2.3 Kalman Filter

Denote by $\tilde{\boldsymbol{x}}_t$ the prediction of the state at time step t based on the state estimate $\hat{\boldsymbol{x}}_{t-1}$ at time step $t-1$ and the state transition matrix $\boldsymbol{F}$, i.e., $\tilde{\boldsymbol{x}}_t = \boldsymbol{F}\hat{\boldsymbol{x}}_{t-1}$, and by $\tilde{\boldsymbol{P}}_t = \boldsymbol{F}\boldsymbol{P}_{t-1}\boldsymbol{F}^{\mathsf{T}} + \boldsymbol{Q}$ the covariance matrix of the error $\tilde{\boldsymbol{x}}_t - \boldsymbol{x}_t$, where $(\cdot)^{\mathsf{T}}$ denotes matrix transposition and $\boldsymbol{P}_{t-1}$ is the covariance matrix of the error $\hat{\boldsymbol{x}}_{t-1} - \boldsymbol{x}_{t-1}$ at time step $t-1$. The Kalman filter is an algorithm for combining the predicted state $\tilde{\boldsymbol{x}}_t$ with an observation $\boldsymbol{z}_t^{(v)} = \boldsymbol{H}^{(v)}\boldsymbol{x}_t + \boldsymbol{r}_t^{(v)}$ to produce an updated state estimate $\hat{\boldsymbol{x}}'_t$ with minimum mean squared error [9]. Let $\tilde{\boldsymbol{y}}_t^{(v)} = \boldsymbol{z}_t^{(v)} - \boldsymbol{H}^{(v)}\tilde{\boldsymbol{x}}_t$ and denote by $\boldsymbol{S}_t^{(v)} = \boldsymbol{R} + \boldsymbol{H}^{(v)}\tilde{\boldsymbol{P}}_t\left(\boldsymbol{H}^{(v)}\right)^{\mathsf{T}}$ its covariance matrix. Then, the updated state estimate is $\hat{\boldsymbol{x}}'_t = \tilde{\boldsymbol{x}}_t + \boldsymbol{K}_t^{(v)}\tilde{\boldsymbol{y}}_t^{(v)}$, where $\boldsymbol{K}_t^{(v)} = \tilde{\boldsymbol{P}}_t\left(\boldsymbol{H}^{(v)}\right)^{\mathsf{T}}\left(\boldsymbol{S}_t^{(v)}\right)^{-1}$ is the Kalman gain that determines how the observation should influence the updated estimate. The covariance matrix of the error $\hat{\boldsymbol{x}}'_t - \boldsymbol{x}_t$ is $\boldsymbol{P}'_t = \left(\boldsymbol{I} - \boldsymbol{K}_t^{(v)}\boldsymbol{H}^{(v)}\right)\tilde{\boldsymbol{P}}_t$, where $\boldsymbol{I}$ is the identity matrix. If more than one observation is available, the estimate can be improved by setting $\tilde{\boldsymbol{x}}_t \leftarrow \hat{\boldsymbol{x}}'_t$ and $\tilde{\boldsymbol{P}}_t \leftarrow \boldsymbol{P}'_t$ and repeating this procedure. After repeating this procedure for all observations, the final estimate $\hat{\boldsymbol{x}}_t$ and its corresponding error covariance matrix $\boldsymbol{P}_t$ is obtained.

3 Resilient Distributed Tracking

In this section, we introduce and analyze the distributed tracking scheme, which is a special case of the scheme in [5]. The key idea is to replicate all observations over the N_{w} workers so that all workers have access to all observations at the start of each time step, which makes the system resilient against stragglers. Each worker computes the estimate $\hat{\boldsymbol{x}}_t$ by first computing the prediction $\tilde{\boldsymbol{x}}_t$ and then combining it with the received observations $\{\boldsymbol{z}_t^{(v)} : v \in \mathcal{V}\}$ one by one as explained in Section 2.3. The first worker to compute $\tilde{\boldsymbol{x}}_t$ shares it with the vehicles and the other workers. Hence, if at least one worker becomes available within a given time step all vehicles and workers have access to an updated estimate at the start of the next time step. If no workers become available during time step t, each vehicle computes $\tilde{\boldsymbol{x}}_t$ locally and we let $\hat{\boldsymbol{x}}_t = \tilde{\boldsymbol{x}}_t$.

The next worker to become available may apply the Kalman filter several times consecutively to catch up.

In the following, we give the probability of the vehicles receiving no update for i consecutive time steps and the probability distribution of the position estimate error.

3.1 Analysis of the AoI

Because the updated estimate is shared with the vehicles at the end of each time step, the AoI at the start of time step t is Δt if at least one worker became available during time step $t-1$. On the other hand, if no workers have become available for i consecutive time steps the AoI is $(i+1)\Delta t$. Here, we give the probability that no workers become available for i consecutive time steps.

Theorem 1. *Let $G_{\Delta t}$ be the random variable associated with the number of unique workers that become available in the time interval from t to $t+\Delta t$, for some t. Then, $\Pr(G_{\Delta t} = N_{\mathsf{w}}) = F_{V_{N_{\mathsf{w}}:N_{\mathsf{w}}}}(t)$ and, for $i < N_{\mathsf{w}}$,*

$$\Pr(G_{\Delta t} = i) = \int_0^{\Delta t} f_{V_{i:N_{\mathsf{w}}}}(v_{i:N_{\mathsf{w}}}) \left(1 - F_{V_{1:N_{\mathsf{w}}-i}}(t - v_{i:N_{\mathsf{w}}})\right) \mathrm{d}v_{i:N_{\mathsf{w}}},$$

where $f_{V_{i:N_{\mathsf{w}}}}$ ($F_{V_{1:N_{\mathsf{w}}-i}}$) is the PDF (CDF) of $V_{i:N_{\mathsf{w}}}$ ($V_{1:N_{\mathsf{w}}-i}$) (see Section 2.2).

The probability of no workers becoming available for i consecutive time steps is $\Pr(G_{\Delta t} = 0)^i$.

3.2 Analysis of the Position Error

We consider the absolute error of the position estimate. Denote by $E^{(v)}_{\mathsf{x},t}$ and $E^{(v)}_{\mathsf{y},t}$ the random variables associated with the error of the estimate of the position in the longitudinal and latitudinal directions, respectively, for vehicle v at time t. Furthermore, denote by $\boldsymbol{P}^{(v)}_{\mathsf{x,y},t}$ the covariance matrix of the random vector $[E^{(v)}_{\mathsf{x},t}\ E^{(v)}_{\mathsf{y},t}]^{\mathsf{T}}$. The covariance matrix can be computed numerically from the Kalman filter equations in Section 2.3. With some abuse of notation, we drop subscript t. Now, let

$$E^{(v)}_{\mathsf{d}} = \sqrt{\left(E^{(v)}_{\mathsf{x}}\right)^2 + \left(E^{(v)}_{\mathsf{y}}\right)^2}$$

be the random variable associated with the absolute distance between the estimated and true position of vehicle v. We have the following results.

Lemma 1. *The probability distribution of $\left(E^{(v)}_{\mathsf{d}}\right)^2$ is equal to that of the sum $\lambda_1 C_1 + \lambda_2 C_2$, where C_1 and C_2 are independent Chi-squared random variables, both with one degree of freedom, and λ_1 and λ_2 are the eigenvalues of $\boldsymbol{P}^{(v)}_{\mathsf{x,y}}$.*

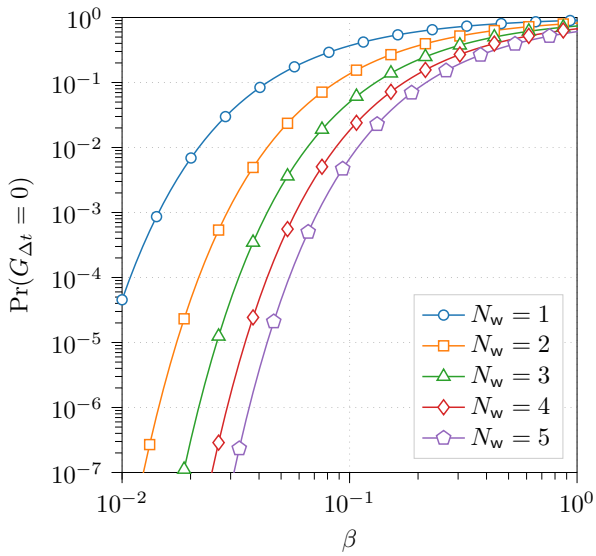

Figure 2. The probability of missing an update as a function of the straggling parameter β when $\Delta t = 0.1$s and $1 \leq N_{\mathsf{w}} \leq 5$.

Theorem 2. *The CDF of $E^{(v)}_{\mathsf{d}}$ is*

$$F_{E^{(v)}_{\mathsf{d}}}\left(e^{(v)}_{\mathsf{d}}\right) = \Pr\left(E^{(v)}_{\mathsf{d}} \leq e^{(v)}_{\mathsf{d}}\right) = \int_0^{\left(e^{(v)}_{\mathsf{d}}\right)^2} f_{C_1}\left(\frac{c_1}{\lambda_1}\right) F_{C_2}\left(\frac{\left(e^{(v)}_{\mathsf{d}}\right)^2 - c_1}{\lambda_2}\right) \mathrm{d}c_1,$$

where f_{C_1} (F_{C_2}) is the PDF (CDF) of C_1 (C_2).

4 Numerical Results

In this section, we plot the probability of missing an update, i.e., $\Pr(G_{\Delta t} = 0)$ (see Theorem 1), and the complementary CDF of $E^{(v)}_{\mathsf{d}}$ (see Theorem 2) for a particular vehicle tracking scenario.

In Fig. 2, we plot $\Pr(G_t = 0)$ as a function of the straggling parameter β when $\Delta t = 0.1$ seconds (s) for $1 \leq N_{\mathsf{w}} \leq 5$ workers. Increasing the number of workers, and thus the replication factor, significantly reduces the chance of missing an update. For example, when $\beta = 0.1$ a replication factor of two ($N_{\mathsf{w}} = 2$) decreases the probability of missing an update from about 0.37 for the scheme with no replication ($N_{\mathsf{w}} = 1$) to about 0.13, thus reducing the AoI. More generally, increasing the replication factor by one lowers the probability of missing an update by about two thirds when $\beta = 0.1$.

In Fig. 3, we consider the vehicle tracking scenario of [2, 5], where at each time step each vehicle observes its absolute position (via, e.g., GPS) and the relative position of a subset of the other vehicles (via, e.g., RADAR or LIDAR). Because relative position can

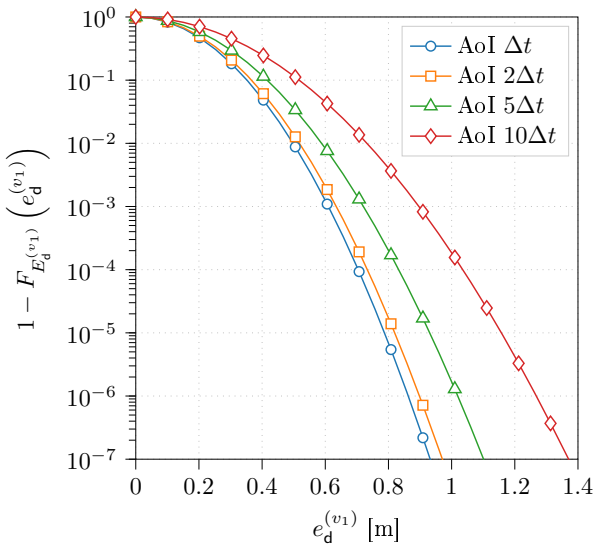

Figure 3. The complementary CDF of the position error $E_{\mathsf{d}}^{(v_1)}$ for vehicle v_1 when there are $N_{\mathsf{v}} = 10$ vehicles, each of which can observe 3 other vehicles (see [5] for details), and $\Delta t = 0.1$s.

be measured much more accurately than absolute position (typically centimeter-level compared to meter-level accuracy), combining observations of relative and absolute positions from multiple sources can significantly improve accuracy. We consider a scenario with $N_{\mathsf{v}} = 10$ vehicles, each of which can observe three of the other vehicles. Matrices $\boldsymbol{F}$, $\boldsymbol{H}^{(v)}$, $\boldsymbol{Q}$, and $\boldsymbol{R}^{(v)}$ are generated as in [5].

We compute the steady state covariance matrix $\boldsymbol{P}_\infty = \lim_{t\to\infty} \boldsymbol{P}_t$ numerically by repeatedly applying the Kalman filter equations (see Section 2.3) until convergence. Note that $\boldsymbol{P}_\infty$ is a function of the statistical properties of the system and is independent of the actual observations $\boldsymbol{z}_t^{(v)}$. Next, for $i \geq 1$, denote by

$$\boldsymbol{P}_{\infty,i} = \begin{cases} \boldsymbol{P}_\infty & i = 1 \\ \boldsymbol{F}\boldsymbol{P}_{\infty,i-1}\boldsymbol{F}^\mathsf{T} + \boldsymbol{Q} & i > 1 \end{cases}$$

the covariance matrix associated with an AoI of $i\Delta t$. We extract the covariance matrix of the position estimate for vehicle v_1 from $\boldsymbol{P}_{\infty,i}$. Finally, we compute the complementary CDF of the absolute position error for vehicle v_1, i.e., $1 - F_{E_{\mathsf{d}}^{(v_1)}}\left(e_{\mathsf{d}}^{(v_1)}\right)$, which we plot in Fig. 3 for $\Delta t = 0.1$s.

The AoI may significantly affect the accuracy of the estimate, especially at the tail. For example, the mean error, $\mathbb{E}\left(E_{\mathsf{d}}^{(v_1)}\right)$, for an AoI of $\Delta t = 0.1$s is 0.19 meters (m), whereas the 90-th and 99-th percentile error is about 0.35m and 0.50m, respectively. For an AoI of $10\Delta t$ the mean, 90-th, and 99-th percentile error is about 50% higher at 0.28m, 0.52m, and 0.73m, respectively, compared to for an AoI of Δt.

5 Conclusion

We considered a distributed vehicle tracking scheme that utilizes replication to alleviate the straggler problem. For this scheme we characterized the AoI of updates available to the vehicles, which depends on the number of workers and a parameter (β) that captures the amount of time that workers become unavailable after completing a task. Replication can significantly reduce the chance of missing an update, thus improving the AoI and accuracy. For example, when $\beta = 0.1$, replication lowers the chance of missing an update from about 0.37 for no replication to about 0.13 when using one replica. Finally, we derived the probability that the error of the estimate of the position exceeds some threshold for a given AoI.

References

[1] 5G Automotive Association, "C-V2X use cases methodology, examples and service level requirements," White Paper, Jun. 2019.

[2] G. Soatti, M. Nicoli, N. Garcia, B. Denis, R. Raulefs, and H. Wymeersch, "Implicit cooperative positioning in vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 12, pp. 3964–3980, Dec. 2018.

[3] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *Proc. IEEE Int. Symp. Inf. Theory*, Cambridge, MA, 2012.

[4] M. S. Mahmoud and H. M. Khalid, "Distributed Kalman filtering: a bibliographic review," *IET Control Theory Appl.*, vol. 7, no. 4, pp. 483–501, Mar. 2013.

[5] A. Severinson, E. Rosnes, and A. Graell i Amat, "Coded distributed tracking," in *Proc. IEEE Global Commun. Conf.*, Waikoloa, HI, Dec. 2019.

[6] G. Liang and U. C. Kozat, "TOFEC: Achieving optimal throughput-delay trade-off of cloud storage using erasure codes," in *Proc. IEEE Conf. Computer Commun.*, Toronto, ON, Canada, Apr. 2014, pp. 826–834.

[7] K. Lee, M. Lam, R. Pedarsani, D. Papailiopoulos, and K. Ramchandran, "Speeding up distributed machine learning using codes," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1514–1529, Mar. 2018.

[8] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja, *A First Course in Order Statistics*, 2nd ed. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2008.

[9] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, Mar. 1960.