Final Version Published in Wireless Networks, (2014) 20:639–654.

Weak state versus strong state: an analysis of a probabilistic state mechanism for dynamic networks

Utku Günay Acer · Shivkumar Kalyanaraman · Alhussein A. Abouzeid

Abstract Network protocols coordinate their decision making using information about entities in remote locations. Such information is provided by *state* entries. To remain valid, the state information needs to be refreshed via control messages. When it refers to a dynamic entity, the state has to be refreshed at a high rate to prevent it from becoming stale. In capacity constrained networks, this may deteriorate the overall performance of the network. The concept of *weak state* has been proposed as a remedy to this problem in the context of routing in mobile ad-hoc networks. Weak state is characterized by probabilistic semantics and local refreshes as opposed to strong state that is interpreted as absolute truth. A measure of the probability that the state remains valid, i.e. *confidence*, accompanies the state. The confidence is decayed in time to adapt to dynamism and to capture the uncertainty in the state information. This way, weak state remains valid without explicit state refresh messages. We evaluate the consistency of weak state and strong state using two notions of distortion. Pure distortion measures the average difference between the actual value of the entity and the value that is provided by the remote state. Informed distortion captures both this difference and the effect of confidence value on state consistency. Using a mathematical analysis and simulations, we show that weak state reduces the distortion values when it provides information about highly dynamic entities and/or it is utilized for protocols that is required to incur a low amount of overhead.

Keywords Weak state · Strong state · State maintenance

An earlier version of this paper appeared in Infocom'09 [2].

Electronic supplementary material The online version of this article (doi:10.1007/s11276-013-0618-5) contains supplementary material, which is available to authorized users.

U. G. Acer (✉)
Alcatel-Lucent Bell Laboratories, Antwerp, Belgium
e-mail: utku_gunay.acer@alcatel-lucent.com

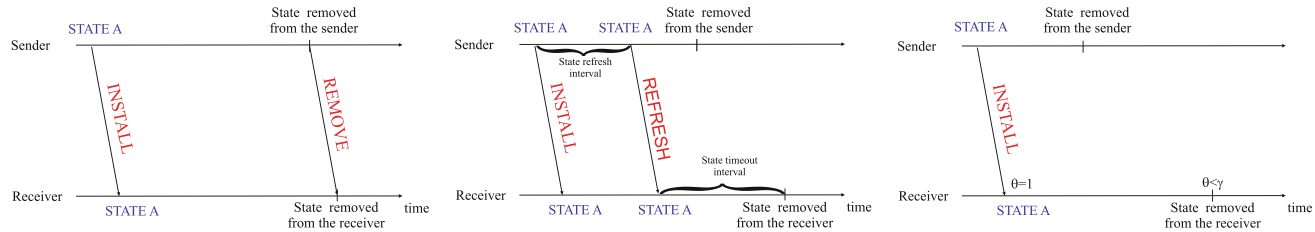
S. Kalyanaraman
IBM Research-Australia, Melbourne, Australia
e-mail: shivkuma@gmail.com

A. A. Abouzeid
Rensselaer Polytechnic Institute, Troy, NY, USA
e-mail: abouzeid@ecse.rpi.edu

1 Introduction

The operation of a majority of network protocols relies on the information about entities or conditions that are actually present in other parts of the network. This information is provided by *state* entries to the protocols. As the state-of-the-art protocols push the limits of dynamism in the information they utilize, it becomes harder to maintain consistent state because the true value of the entity changes often.

Traditionally, the concept of state can be classified into two categories based on the way that the state is signaled to remote nodes: hard and soft state approaches. Hard state remains valid until it is explicitly removed using state teardown messages by the node that installs the state. The installer node refreshes the state at the remote nodes only when the original state is updated. Since the state is removed explicitly, reliable communication is essential. On the other hand, soft state, the term first coined by Clark [5], *times out* unless it is refreshed within some time-out duration. The state installer node periodically issues a refresh message. Once this message is received by the node maintaining the state, the timer corresponding to the state is



(a) Hard State **(b)** Soft State **(c)** Weak State

Fig. 1 A comparison summary of hard state, soft state and weak state approaches. Hard state requires explicit control message to be removed. Soft state times out if it is not refreshed within the timeout interval. Weak state is associated with a confidence value θ, which is a decreasing function of time. When the confidence is below a threshold value γ, it is removed

rescheduled. If the timer expires, the state times out and is removed from the system. Soft state does not require explicit removal messages. As a result, reliable signaling is not required. Refresh message losses can be tolerated since the state is refreshed periodically.

Both hard state and soft state are deterministic and regarded as absolute truth. We say that such information has strong semantics, i.e. it is *strong state*. If strong state provides information about a dynamic entity whose value changes frequently, it is rapidly invalidated and needs to be refreshed often through control messages in order to provide up-to-date information. Refreshing the remote state at a higher rate increases the number of refresh messages roaming around the network. These messages can quickly fill up the transmission queues of the nodes and crowd out the network's entire capacity for data traffic [1].

We have recently proposed the concept of *weak state* in order to address this problem within the context of routing in dynamic mobile ad-hoc networks [3] and delay tolerant networks [4]. Unlike traditional state, weak state is not deterministic; it has probabilistic semantics, i.e. it is interpreted as a probabilistic hint. The state is associated with a *confidence* value that measures the probability that the value provided by the state remains valid. The confidence is locally decayed (i.e. weakened) over time to capture the uncertainty in the information. Once the confidence is below a threshold value, the state is removed from the system. Weakening the state corresponds to aging it and is equivalent to a soft timeout. Hence, weak state is a generalization of soft state. A comparison of hard, soft and weak state is given in Fig. 1. Weak state mechanism design is characterized by two properties:

1. Probabilistic semantics: The information that the state yields is not deterministic. Instead, the validity of the information is subject to a probabilistic confidence value.
2. Local refreshes: The information maintained at remote nodes can be refreshed without explicit control messages from the host of the original state (i.e. sender). The receiver locally refreshes the confidence value in order to capture the uncertainty. In addition, the receiver can estimate the actual value of the state.

The consistency of state information may have serious implications in protocol design. For example in AODV [19], one protocol parameter that governs the consistency of the maintained state information is active route timeout (ART) value. In [22], it has been shown that improper values for ART causes routes to expire before or after they should, which leads to inconsistency. This in turn deteriorates performance in terms of packet delivery ratio and the number of control messages that correct the invalidated routes. Hence, protocols that utilize state entries should employ mechanisms to ensure the consistency of the state. When doing this, overall performance of the network should also be taken into account as well.

In this paper, we compare the consistency of weak state with that of strong state using the concept of distortion. We first consider the scenario where the original value of the state changes at discrete instances and there is no correlation between the new and old values of the state. In this case, the distortion is defined to be the probability that the original value of the state and the perceived value maintained at a remote location are different. When the state value changes continuously, this definition of distortion does not suffice. In such scenarios, we use the notions of *pure* distortion and *informed* distortion. Pure distortion measures the average difference between the actual value of the interested entity and the value of the state entry that provides information about it. With weak state, pure distortion is lower because the protocol can use an estimate of the state value instead of a mere last reported value. In addition, using the confidence value makes the protocol capable of adapting to dynamism. If the confidence in the maintained value is low, the state is likely invalid and the protocol is less likely to use information provided by the state. Thus, the protocol can cope with high pure distortion. In other words, it is superior to have a hint with a measure of accuracy of the state than having invalid deterministic state which may lead to wrong decisions with probability 1. The informed distortion metric captures both the difference between the original value of the entity and the effect of the confidence value.

We have used a variety of stochastic processes to model the value of the actual state value and evaluate the consistency of the weak state and strong state in terms of distortion metrics. For all processes, our results globally show that strong state may cause lower distortion by adjusting the state refresh interval appropriately to small values. As the level of dynamism increases, this approach requires issuing more refresh packets, which is not desirable in networks that are vulnerable to high overhead such as large scale mobile networks. For such networks, weak state is a more appropriate design tool because it reduces the distortion values significantly when the number of refresh messages generated in the network is low. *This however does not mean that there is no trade off*. The maintained information is not perfect and the protocol may choose different ways to interpret and deal with the confidence. For instance in the *weak state routing* (*WSR*) protocol that utilizes a weak state realization for routing in large scale and dynamic networks, a packet is successively biased towards the points yielded by the intermediate nodes that contain increasingly more confident state [3]. WSR increases the delivery ratio and decreases the overhead significantly at the cost of increasing path length. With the dramatically reduced control traffic, longer paths need not imply longer end-to-end delivery latency since the protocol reduces queuing in the intermediate nodes. In this paper, we do not address such practical issues regarding the implementation.

1.1 Related work

The concept of weak state has recently been proposed and utilized for routing in large scale and dynamic mobile ad-hoc networks [3]. The same state model has later been used to perform routing in delay tolerant networks [4]. However, other realizations are possible such as PROPHET [14] and EDBF [11]. We believe a large variety of techniques could be used to realize weak state but the focus of this paper is on the consistency of probabilistic and deterministic state maintenance when the state changes dynamically. We do that by using an abstract model of weak state and strong state. Their actual implementation is beyond the scope of this paper. In a survey of methods for approximate global state for distributed systems, it is mentioned that approximate or "weak" state could be a useful primitive for dynamic networks [10].

Analytical comparisons of hard state, soft state and the hybrid approaches in terms of consistency have been presented by Raman and McCanne [21] and Ji et al. [9]. In both papers, state is modeled with strong semantics and the focus is on the signaling mechanisms. On the other hand, we take state maintenance mechanisms into account and compare states with weak and strong semantics. In terms of robustness, Lui et al. [15] has shown that soft state can be a better design tool in comparison to hard state for protocols that operate under scenarios that might contain denial of service attacks and correlated lossy feedback channel in comparison to hard state.

Manfredi et al. [16] categorizes the state information into control state and data state, where the first maintains information about the network while the latter yields information about the data packets and stored on packets. The authors propose a framework that suggests that in unpredictable scenarios one should opt for data state and when the network is well connected control state should be used. In this paper, our attention is on control state where the network information is maintained in remote nodes and whether the nature of the state should be deterministic or probabilistic.

1.2 Organization of paper

In Sect. 2, we lay out the foundations of our analysis. In Sect. 3, we extend our analysis to a broader range of scenarios. We model the state information using a variety of stochastic processes in Sect. 4 and evaluate the consistency of strong state and weak state in terms of pure and informed distortion. In Sect. 5, we demonstrate that protocol performance may benefit from low distortion and use experiments to evaluate the distortion in more complex scenarios. Finally, we conclude the paper in Sect. 6.

2 Analysis framework

2.1 System and data model

In this paper, we adopt the single hop signaling system used in previous work [9, 21]. The system consists of a node that hosts the entity of interest and another node that maintains the remote state that provides information about the entity to the instance of the protocol. The first sends refresh messages to the latter. We refer to these nodes as the sender and the receiver nodes, respectively. The true

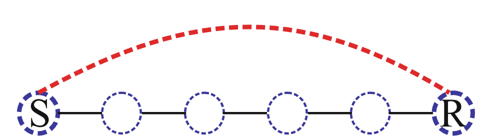


Fig. 2 Single hop signaling system. The sender *S* is connected to the receiver *R* with a logical link. The logical link can consist of a series of physical links which enables end-to-end signaling

value of the entity corresponds to the sender's state or the state at the sender. On the other hand, the receiver's state or the state at the receiver provides information about the entity to the instance of the protocol running at the receiver.

The sender installs the state at the receiver and periodically refreshes it. The sender and the receiver are connected to each other over a logical link that typically consists of a sequence of physical hops (see Fig. 2). INSTALL, REFRESH and REMOVE messages on Fig. 1 are transferred over the logical link. The signaled messages can be lost while being transferred through the logical link. In our analysis, we consider the soft state signaling mechanism. For convenience, we assume that each message is delivered to the receiver independently from other messages. For all messages, the delivery is subject to a homogeneous probability value.

We assume that the sender's state X, is a continuous time stochastic process. The state update corresponds to the changes in the value of the sender's state. The intervals between state updates are independent and exponentially distributed random variables, with parameter λ. The sender periodically refreshes the perceived state value, $\hat{X}$, at the receiver. The interval between the refresh messages is fixed and denoted by T. The value of $\hat{X}$ at any given time is the value of X as reported by the most recently received refresh message. Strong state times out if it is not refreshed within some timeout interval, χ. Weak state, on the other hand, is a tuple $(\hat{X}, \theta)$, where θ is the confidence value, the probability that the state at the receiver remains valid, i.e. its value is still equal to the value of the sender's state. The confidence of the state is decayed in time locally at the receiver because the probability of the state validity decreases with time. At time t, the confidence is $\theta(t) = P(\zeta > t) = \exp(-\lambda t)$ where ζ denotes the time interval in which the state at the sender remains the same. Once the confidence is below some threshold, γ, the receiver removes the state. The protocol designer controls the lowest confidence value below which the state entry has no importance through γ parameter. Note that the protocol designer does not decide on the state timeout value; instead, it is implicitly deducted from the selected γ value. The state times out when the receiver reduces the confidence value locally to γ. We assume that the protocols always work, and the duration for which the state at the sender is maintained is very long in comparison to the average state update interval and state timeout interval; it approaches infinity. The summary of notation used in the paper is given in Table 1.

In this model, the incurred overhead is tied to the state refresh interval, T. The overhead also depends on the path that the refresh messages follow and it includes the messages that are associated with the formation and the maintenance of the logical hop. However, these messages are in the lower layers and do not depend on the signaling parameters and/or whether the state information is interpreted deterministically or probabilistically. Hence, these details can be omitted in the comparison of weak state and strong state.

Table 1 Summary of notation

Notation	Description
X	State value at the sender
$\hat{X}$	State value at the receiver
ζ	Random variable for state update interval
λ	Average state update rate, i.e. $E[\zeta] = 1/\lambda$
χ	State timeout interval
p	Signaling loss rate
T	State refresh interval
θ	Confidence in state information
γ	Confidence threshold for removing weak state

2.2 State consistency

We evaluate the consistency of the protocol state using a distortion measure. The term distortion refers to the changes in the original characteristics, such as frequency and amplitude, of a signal due to an operation such as filtering [17] and is associated with loss of quality in signal. In our case, the quality of the state depends on both X and $\hat{X}$ values and how the maintained value of $\hat{X}$ is utilized by the receiver, i.e. probabilistically or deterministically. It is important to note that distortion is not something that the receiver calculates for a practical purpose in the operation of the protocol. Rather, we introduce it in order to measure the correctness or the consistency of the maintained state.

We define the distortion as the probability that the receiver interprets the state mistakenly, which takes place if the receiver's state is different from the state at the sender and the receiver thinks that it maintains the correct state value. When the values are actually the same but the receiver thinks that it maintains an incorrect value, the state is also misinterpreted since the receiver discards the valid information provided by the state entry.

The state values at the sender and the receiver differ if (1) the state entry at the sender is updated but the receiver's state remains the same, or (2) the receiver falsely removes the state due to a series of refresh messages being lost on the logical link while the sender's state and the receiver's state have equal values. In addition, the values can also be different if the sender removes the state but the receiver still maintains it. Since the lifetime of the sender's state is assumed to be very large and to approach infinity, this case is not a factor in our analysis.

Without loss of generality, let the state at the receiver be most recently refreshed at time $t = 0$, $X(0) = \hat{X}(0)$. Let $d(t)$ denote the instantaneous distortion between the state value at the sender and the receiver at time t. If the state has strong semantics[1], the information is deterministic. If the receiver's state is different from the sender's state, the protocol definitely uses an inaccurate value. In other words, given the value of the state at the receiver at this time instant, the probability for inconsistency at time t is either 0 or 1.

$$d_S(t) = \begin{cases} 0 & \text{if} \quad X(t) = \hat{X}(t) \\ 1 & \text{otherwise.} \end{cases} \tag{1}$$

The state information is probabilistic in weak state. Even when the state at the receiver is different from the state at the sender, the protocol at the receiver is less likely to use the maintained value if the accompanied confidence is low. If the receiver's state does not match with the sender's state, the distortion is characterized by the probability, as inferred by the receiver, that the value it maintains is equal to the sender's state. This probability is given by the confidence of the receiver in the maintained value. Weak state also can cause distortion when state at the sender and the state at the receiver are the same because the receiver infers that the state is invalid with some probability. This probability defines the distortion in this case. The resulting instantaneous distortion for weak state is

$$d_W(t) = \begin{cases} 1 - \theta(t) & \text{if} \quad X(t) = \hat{X}(t) \\ \theta(t) & \text{otherwise.} \end{cases} \tag{2}$$

At time t, probability that there is no update in the sender's state is $P\big(X(t) = \hat{X}(t)\big) = P(\zeta > t) = \exp(-\lambda t)$. Hence, the expected instantaneous distortion at time t for strong state and weak state can be modeled as

$$D_S(t) = \begin{cases} 1 - \exp(-\lambda t) & \text{if} \quad t \leq \chi \\ \exp(-\lambda t) & \text{otherwise} \end{cases} \tag{3}$$

$$D_W(t) = \begin{cases} 2\exp(-\lambda t)(1 - \exp(-\lambda t)) & \text{if} \quad t \leq \chi \\ \exp(-\lambda t) & \text{otherwise.} \end{cases} \tag{4}$$

The second terms in both (3) and (4) are due to false removal of the state at the receiver. If a number of refresh messages are lost, the receiver removes the state it maintains. These terms yield the probability that the value of the removed state is still equal to the value of the sender's state given that the time elapsed since the reception of the last refresh message is larger than χ.

[1] We denote the semantics of the state as a subscript throughout the paper.

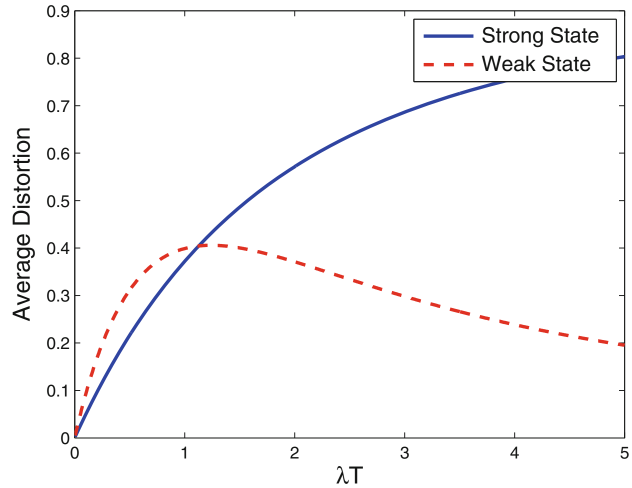


Fig. 3 Distortion against the λT product. A higher λ indicates that the information is more dynamic, and smaller T indicates that the state at the receiver is refreshed more frequently. At a fixed T, the figure shows that strong state causes lower distortion for static systems whereas weak state performs better in dynamic settings. Under a constant λ, strong state causes lower distortion if the state is refreshed at a high rate at the cost of larger number of refresh messages while weak state incurs lower distortion if the state is refreshed by a lower number of refresh messages

The average overall distortion, which is denoted by $\overline{D}$, within the interval between two state refresh messages received by the receiver is

$$\overline{D} = \sum_{i=1}^{\infty} \frac{1}{iT} p^{i-1}(1-p) \int_{0}^{iT} D(\tau)d\tau \tag{5}$$

where i is the number of attempts to deliver a refresh message successfully to the receiver and iT is the time between the two consecutive refresh message receptions.

The average distortion values for both weak state and strong state are functions of the λT product once Eqs. (3) and (4) are integrated into (5). Figure 3 shows how $\overline{D}$ varies with respect to the λT product. Obtaining this figure, we used $\chi = 3T$ for strong state. We have adjusted the γ parameter for weak state so that it has the same timeout value as the strong state for a given T value.[2] The probability of loss for the refresh messages, p, is equal to 0.02.

[2] We note that the distortion also depends on the timeout parameter, χ. Using this parameter, the receiver detects whether a number of consecutive refresh messages are missed, which implies that the sender stopped refreshing messages. Hence, χ should be a function of T. Scaling λ and T does not yield the identical distortion values because χ scales with T but not with λ. However, the resulting difference is always very small (less than 0.001) as the timeout probability is very small. For clarity, we represent the performance with respect to λT. We take the same approach in the following sections as well.

If the state refresh interval is taken constant, Fig. 3 shows how average distortion changes with the dynamism in the state, i.e. λ. For example if it is assumed $T = 1$, then the figure plots the distortion against λ. When λ is small, the state at the sender does not change frequently and the state value at the receiver is likely to remain the same. Given $\hat{X} = X$, the strong state does not cause any distortion whereas weak state can cause some distortion because of the wrong interpretation that the maintained value might be inaccurate. This stems from the probabilistic nature of weak state. Hence, the expected distortion of weak state is slightly higher than strong state when the dynamism is low. As λ increases, the sender's state is more likely to change between two consecutive instants the sender refreshes the receiver, which increases the distortion for strong state. In weak state, the confidence drops with time and the distortion does not increase even though the state value changes rapidly.

Figure 3 similarly shows the relationship between the average distortion and the state refresh interval, T, when the state update rate is constant. In other words, if λ is taken as equal to 1, the figure shows how the distortion changes with respect to T. When T is low, state is refreshed frequently, which decreases the average distortion for strong state. Because weak state can interpret the state information mistakenly even when $\hat{X} = X$, its average distortion is slightly higher in this case. If T is large, the duration for which the sender's state does not match the state at the receiver increases. Weak state can adapt to this by locally decaying the confidence, resulting a lower distortion.

Strong state causes lower distortion if the state refresh interval is set to a very small value. However, this approach costs a large number of refresh messages when the state entry yields information about a dynamically changing entity. If a large state refresh interval is required by the network, weak state is more advantageous since it incurs significantly lower distortion.

3 Extension of analysis: state with correlated space

In some scenarios, knowing that the receiver's state is different from the sender's state may not be sufficient and the amount by which they are apart is also important. For example, in a routing protocol for mobile networks that uses the geographical locations of nodes instead of link states, the distance between the exact location of a destination and its perceived location matters. As the distance grows, the protocol tries to deliver a packet to a point that is very far away from the destination's location and it likely gets harder to rectify this invalid decision elsewhere. In this section, we extend the analysis to capture the difference between the state values at the sender and receiver in addition to the effect of the probabilistic confidence parameter.

Let $\mathbf{x}(t)$ and $\hat{\mathbf{x}}(t)$ denote the values of the state at the sender and the receiver, respectively. Let t_0 and t_1 be two consecutive time instants at which the sender issues refresh messages for the state maintained at the receiver, with $\mathbf{x}(t_0) = x_0$. In strong state, the state information perceived by the receiver remains constant until t_1, i.e.

$$\hat{\mathbf{x}}_S(t) = \mathbf{x}(t_0), \quad t_0 \le t \le t_1. \tag{6}$$

On the other hand, weak state can be refreshed locally using the statistical properties of the process[3]. $\hat{\mathbf{x}}_W(t)$ is the expected value of $\mathbf{x}(t)$, which is the best estimate given $\mathbf{x}(t_0) = x_0$.

$$\hat{\mathbf{x}}_W(t) = E[\mathbf{x}(t)|\mathbf{x}(t_0) = x_0]. \tag{7}$$

Equation (7) captures that weak state is refreshed locally and continuously. The confidence of the state at time t, $\theta(t)$ is characterized by the probability

$$\theta(t) = P(|\mathbf{x}(t) - \hat{\mathbf{x}}_W(t)| \le \rho) \tag{8}$$

where the protocol can tolerate a difference between the state at the sender and the receiver that is at most ρ. Equation (8) yields the probability that maintained state is within this tolerable interval of the sender's state.

Similar to Sect. 2, we start with an intermediate step that introduces the instantaneous distortion, $e(t)$. The instantaneous distortion at time t is the difference between the state values at the sender and the receiver at t in terms of mean square error. Let the mostly recently received refresh message be issued at time t_0 by the sender. Then,

$$e(t) = E[(\mathbf{x}(t) - \hat{\mathbf{x}}(t))^2|\mathbf{x}(t_0) = x_0]. \tag{9}$$

For strong state, $\hat{\mathbf{x}}_S(t) = x_0$.

$$\begin{aligned} e_S(t) &= E[(\mathbf{x}(t) - x_0)^2|\mathbf{x}(t_0) = x_0] \\ &= \mu_t^2 + \sigma_t^2 - 2x_0\mu_t + x_0^2 \\ &= \sigma_t^2 + (\mu_t - x_0)^2 \end{aligned} \tag{10}$$

where μ_t is the expected value of $\mathbf{x}_t$ given $\mathbf{x}(t_0) = x_0$ and σ_t^2 is the variance under the same condition. For weak state, we have $\hat{\mathbf{x}}_W(t) = \mu_t$.

$$\begin{aligned} e_W(t) &= E[(\mathbf{x}(t) - \mu_t)^2|\mathbf{x}(t_0) = x_0] \\ &= \sigma_t^2. \end{aligned} \tag{11}$$

Since $(\mu_t - x_0)^2 \ge 0$ for all t and x_0, we have $e_W \le e_S$. In other words, the instantaneous distortion for weak state is never higher than that of strong state for the same t.

[3] Obtaining the information about these properties is beyond the scope of this paper.

Note that these relations are given if $t \leq t_0 + \chi$. Otherwise, there is a distortion involved due to the possibility of false removal of the state, which is denoted by constant C.

Using instantaneous distortion, we define two notions of distortion: pure distortion and informed distortion. Pure distortion merely measures the difference between the sender's state and the receiver's state. Informed distortion incorporates the effect of confidence and the tolerable interval for state validity into the analysis.

3.1 Pure distortion

Pure distortion quantifies the average separation between the state at the sender and the receiver; i.e., it is the time average of instantaneous distortion.

Without loss of generality, let's assume $t_0 = 0$. The average pure distortion within two consecutively received state refresh messages is

$$\overline{e^P} = \sum_{i=1}^{\infty} \frac{1}{iT} p^{i-1}(1-p)E(iT) \tag{12}$$

where $E(t) = \int_0^t e(\tau)d\tau$.

Since weak state can be refreshed locally, the receiver can use an estimate for the current value of the sender's state rather than merely relying on the value given by the most recently received message, i.e. $e_W(t) \leq e_S(t)$ for all t. Hence, $\overline{e^P_W} \leq \overline{e^P_S}$ for the same T. In other words, we can use a larger T value for weak state to achieve the same average pure distortion.

3.2 Informed distortion

In weak state, local refreshes are particularly useful because they also involve locally weakening the confidence value. This way, weak state captures the uncertainty in state, which is a result of the dynamism. Here, we introduce the concept of *informed distortion* that captures the effect of the confidence together with the difference between the values of the state at the sender and the receiver. Therefore, informed distortion evaluates strong state and weak state better.

Let the value maintained at a receiver node be ϑ. When asked about the state information, the receiver will reply "The actual value of the sender's state is within the interval $[\vartheta - \rho, \vartheta + \rho]$ with probability v." Even if the pure distortion is very large, its effect will be limited if v is low. If weak state mechanism design is used, the protocol adapts appropriately to dynamism by taking the probability v into account, instead of making deterministic and invalid decisions. For example in WSR, intermediate nodes do not always bias the packets they receive even if they contain information about the location of the node. If the confidence in the information is low, the information provided by this state is discarded. This way, the packets are not forwarded to invalid locations and the information maintained at an intermediate node does not deteriorate the performance even though the information at this node causes high pure distortion.

In informed distortion, if the receiver's state is within the interval ρ of the sender's state the strong state does not cause distortion because such a difference can be tolerable for the protocol. For instance in routing for mobile networks, the protocol can successfully deliver the packets when the difference between the perceived location by an intermediate or the source node and the actual location of the destination node is within some threshold. In order to capture this, we incorporate the confidence by weighting the instantaneous distortion with the probability for the receiver to interpret the validity of the state mistakenly at a time instant. The informed distortion is the time average of this product:

$$s(t) = \begin{cases} 0 & \text{if} \quad |\mathbf{x}(t) - \hat{\mathbf{x}}_S(t)| \leq \rho \\ e_S(t) & \text{otherwise} \end{cases}$$

$$w(t) = \begin{cases} (1-\theta(t))e_W(t) & \text{if} \quad |\mathbf{x}(t) - \hat{\mathbf{x}}_W(t)| \leq \rho \\ \theta(t)e_W(t) & \text{otherwise.} \end{cases}$$

The expected values are:

$$e_S^{(I)}(t) = (1 - \Theta(t))e_S(t) \tag{13}$$

$$e_W^{(I)}(t) = 2(1 - \theta(t))\theta(t)e_W(t) \tag{14}$$

where $\Theta(t) = P(|\mathbf{x}(t) - \hat{\mathbf{x}}_S(t)| \leq \rho)$, i.e. the probability that the strong state remains valid since last refresh message. Note that in weak state, even if $|\mathbf{x}(t) - \hat{\mathbf{x}}_W(t)| \leq \rho$, the protocol may interpret $\hat{\mathbf{x}}_W(t)$ as invalid because of the probabilistic confidence value.

In order to derive an expression for the average informed distortion, we substitute (10) and (11) with (13) and (14) in (12), respectively. Also, when $t > \chi$ the distortion incurs with probability $P(|\mathbf{x}(t) - \hat{\mathbf{x}}(t)| \leq \rho)$, which captures the likelihood of the event that the state is still valid even though it is removed.

4 Particular example processes for state information

In this section, we demonstrate the results obtained by applying the analysis in the previous section on several examples of random processes. These results give insights about the quantitative performance differences stemming from adopting strong state and weak state under various dynamic conditions. Although the actual numbers may change, the results have identical characteristics for all these example processes.

4.1 Wiener process

The Wiener Process is a Gaussian process, which is widely used as a mobility model in the literature for mobile networks (e.g. [7]). In this process, two samples obtained at two time instances are jointly Gaussian random variables.

In a Gaussian process, the confidence for a weak state at time t according to (8) is

$$\theta(t) = \mathrm{erf}\left(\frac{\rho}{\sigma_t\sqrt{2}}\right) \tag{15}$$

where erf(.) is the error function.

The Wiener process is characterized by independent, stationary increments. The conditional density is

$$(\mathbf{x}(t_0+t)|(\mathbf{x}(t_0)=x_0)) \sim \mathcal{N}(x_0, \sigma^2 t)$$

for all t_0, $t \geq 0$ and $\sigma_t^2 = \sigma^2 t$. Without loss of generality, we assume $t_0 = 0$. Because $\hat{\mathbf{x}}(t) = x_0$ in both weak state and strong state, the instantaneous distortion equals $\sigma^2 t$ and the pure distortion is identical. Using (12), we have

$$\begin{aligned}\overline{e^P} &= \sum_{i=1}^{\kappa} p^{i-1}(1-p)\frac{1}{2}\sigma^2 iT \\ &\quad + \sum_{i=\kappa+1}^{\infty} \frac{1}{iT} p^{i-1}(1-p)\left[\frac{1}{2}\sigma^2\chi^2 + C(iT-\chi)\right] \\ &\leq \frac{\sigma^2 T(1-p^{\kappa+1})}{2(1-p)} + p^{\kappa} C\end{aligned} \tag{16}$$

where $\kappa = \left\lfloor \frac{\chi}{T} \right\rfloor$. Recall that $e(t) = C$ if $t > \chi$. (16) is true if $C > \frac{1}{2}\sigma^2\chi$ and it is proper to have $C \geq \sigma^2\chi$, which is the instantaneous distortion value at $t = \chi$.

The state information maintained at the receiver can be valid even if the receiver removes the state due to loss of refresh messages over the logical link. For a state entry that has been removed by the receiver, the probability that its value remains equal to the sender's state is associated with Θ for strong state and θ for weak state. In the Wiener Process, we have $\Theta(t) = \theta(t)$, which is given in (15). Hence, the expected informed distortion values are

$$e_S^{(I)}(t) = \begin{cases} (1-\theta(t))\sigma^2 t & t \leq \chi \\ \theta(t)C & t > \chi \end{cases}$$

$$e_W^{(I)}(t) = \begin{cases} 2(1-\theta(t))\theta(t)\sigma^2 t & t \leq \chi \\ \theta(t)C & t > \chi \end{cases}$$

Figure 4 presents the pure and informed distortion for strong and weak state with respect to the $\sigma^2 T$ product as both the displacement and the confidence at time t can be represented in terms of $\sigma^2 t$. We take the message loss probability on the logical hop as $p = 0.02$. The tolerable

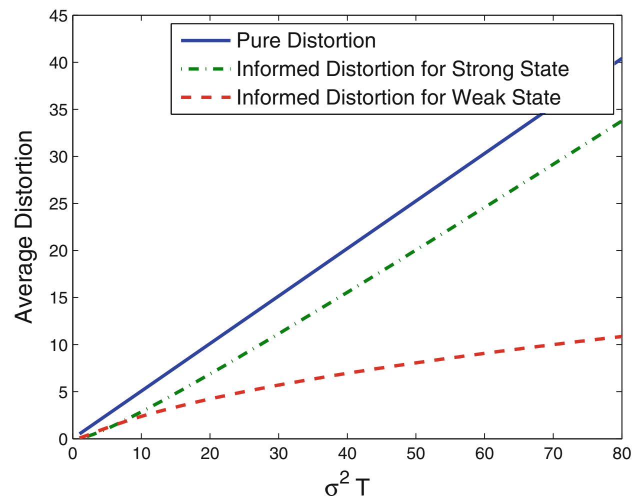


Fig. 4 Pure distortion and informed distortion with respected to the product $\sigma^2 T$. Pure distortion is identical for both strong state and weak state. Because of probabilistic semantics, weak state causes slightly higher informed distortion when the dynamism is low and/or state is refreshed frequently. Otherwise, informed distortion is much lower for weak state

interval in the state value is taken $\rho = 2$. For strong state, $\chi = 3T$. The γ parameter for weak state is adjusted so that weak state and strong state have the same timeout value when other parameters are the same. Because $\hat{\mathbf{x}}_S(t) = \hat{\mathbf{x}}_W(t)$, pure distortion is identical for weak state and strong state. If T is taken constant, the figure shows how the average distortion values change with respect to the variance of the state value, σ^2, which controls the dynamism in the state. When the state variance is low, the receiver's state is likely to remain within the tolerable interval of the sender's state where strong state does not incur any distortion. Though it happens with low probability, weak state can still cause some distortion and the value in (14) is slightly larger than the one in (13). Consequently, the same trend follows in the time average as well. When the dynamism is high, the informed distortion for weak state is significantly lower because the receiver can deduce whether the maintained value is within the tolerable interval of the actual value at the sender.

Figure 4 also shows the effect of the state refresh interval T on the distortion values. When the state refresh interval is small, the receiver's state is frequently refreshed and the probability that it remains within the tolerable interval of sender's state is high. Hence, the informed distortion for strong state is lower than weak state because of the latter's probabilistic nature. If T is large, the receiver's state does not remain within the tolerable interval until the next refresh message is received, with high probability. Weak state can infer this situation because of locally refreshed confidence. When the protocol cannot afford to refresh the state information at a high rate, weak state causes significantly lower informed distortion than strong state.

4.2 Ornstein–Uhlenbeck process

The Ornstein–Uhlenbeck Process is a stationary, Gaussian and Markovian process that is widely used in mobile networking research [13]. $E[\mathbf{x}(t)] = 0$ as $t \to \infty$. The autocorrelation function is of the form $R(\tau) = \sigma^2 e^{-\alpha|\tau|}$. Given $\mathbf{x}(0) = x_0$, the random variable $\mathbf{x}(t)$ is a Gaussian random variable with mean $\mu_t = x_0 e^{-\alpha t}$ and variance $\sigma_t^2 = \sigma^2 (1 - e^{-2\alpha\, t})$ [18] (Chapt. 11-1).

Derived from (10) and (11), the instantaneous distortion values are given by

$$e_S(t) = \begin{cases} \sigma^2(1 - e^{-2\alpha t}) + (\mu_t - x_0)^2 & t \le \chi \\ \sigma^2 & t > \chi \end{cases} \quad (17)$$

$$e_W(t) = \begin{cases} \sigma^2(1 - e^{-2\alpha t}) & t \le \chi \\ \sigma^2 & t > \chi \end{cases} \quad (18)$$

In both cases, the instantaneous distortion in case of false removal at the receiver is $C = \sigma^2$. It corresponds to the variance of a sample obtained from the process without any conditional information since the state is removed the receiver does not maintain any value when $t > \chi$. The expected informed distortion values are

$$e_S^{(I)}(t) = \begin{cases} (1 - \Theta(t))\left(\sigma^2(1 - e^{-2\alpha t}) + (\mu_t - x_0)^2\right) & t \le \chi \\ \Theta(t)\sigma^2 & t > \chi \end{cases}$$

$$e_W^{(I)}(t) = \begin{cases} 2(1 - \theta(t))\theta(t)(\sigma^2(1 - e^{-2\alpha t})) & t \le \chi \\ \theta(t)\sigma^2 & t > \chi \end{cases}$$

where $\Theta(t) = P(|\mathbf{x}(t) - x_0| \le \rho)$ and $\theta(t)$ is given in (15).

The pure distortion values can be obtained by integrating (17) and (18), respectively as in (12). For $t \le \chi$, we have

$$E_S(t) = \sigma^2\left(t + \frac{1}{2\alpha}e^{-2\alpha t} - \frac{1}{2\alpha}\right) + x_0^2\left(t + \frac{2}{\alpha}e^{-\alpha t} - \frac{2}{\alpha} - \frac{1}{2\alpha}e^{-2\alpha t} + \frac{1}{2\alpha}\right) \quad (19)$$

$$E_W(t) = \sigma^2\left(t + \frac{1}{2\alpha}e^{-2\alpha t} - \frac{1}{2\alpha}\right).$$

Note that (19) is obtained given that x_0 is known. However, x_0 is also a stochastic value. Since $E[x_0^2] = \sigma^2$, the cumulative distortion is

$$E_S(t) = 2\sigma^2\left(t + \frac{1}{\alpha}e^{-\alpha t} - \frac{1}{\alpha}\right).$$

Similar to the Wiener process, we divide the time axis into two, before and after χ. The average pure distortion considering the refresh packet losses can be derived as the following:

$$\overline{e^P} = \sum_{i=1}^{\kappa} \frac{1}{iT} p^{i-1}(1-p)E(iT) + \sum_{i=\kappa+1}^{\infty} \frac{1}{iT} p^{i-1}(1-p)\left[E(\chi) + \sigma^2(iT - \chi)\right].$$

For weak state,

$$\overline{e_W^P} \approx \sigma^2\left[1 - \frac{1}{2\alpha T} + \frac{1}{2\alpha T}e^{-2\alpha T}(1-p) \times \frac{1 - (pe^{-2\alpha T})^\kappa}{1 - pe^{-2\alpha T}} + \frac{1}{\kappa T}p^\kappa \frac{1}{2\alpha}e^{-2\alpha\chi}\right]. \quad (20)$$

The similarity in (20) holds because $\frac{u^i}{i} \approx u^i$ when $u \ll 1$ and i is large. Also note that $\frac{u^i}{i} \le u^i$ when i is a positive integer. Since u^i converges if $0 < u < 1$, $\frac{u^i}{i}$ converges as well. For strong state, we have

$$\overline{e_S^P} \approx 2\sigma^2\left[1 - \frac{1}{\alpha T} + \frac{e^{-\alpha T}}{\alpha T}(1-p)\frac{1 - (pe^{-\alpha T})^\kappa}{1 - pe^{-\alpha T}} + 2p^\kappa \frac{1}{\alpha\kappa T}e^{-\alpha\chi}\right]. \quad (21)$$

Unless otherwise noted, the following default parameters are used for distortion evaluation: $\sigma = 1$, $p = 0.02$, and $\rho = 1/\sqrt{2}$. For strong state, $\chi = 3T$. For consistency in the parameters, we have adjusted γ for weak state so that its timeout value is same as that of strong state.

We compare the distortion values for strong state and weak state against the αT product in Fig. 5. Similar to the Wiener process, the evolution of the process at time t can be represented by the value of αt. State correlation parameter α

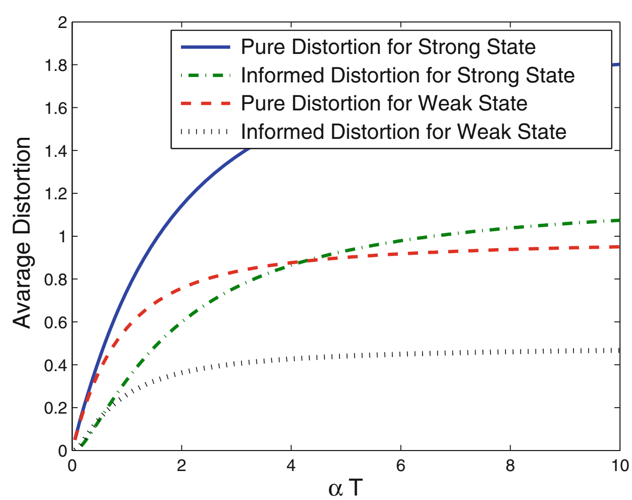


Fig. 5 Average distortion values with respect to the α T product. When the state is dynamic (large α) or state is refreshed at a small rate (large T), weak state reduces the pure distortion by obtaining better estimates about the state value. Utilizing the probabilistic confidence value decreases the informed distortion for weak state

indicates how fast the process deviates from previous values. Under a constant T value, consider two samples obtained from the state $\mathbf{x}$ at times t_1 and t_2. The correlation coefficient between $\mathbf{x}(t_1)$ and $\mathbf{x}(t_2)$ is given by $e^{-\alpha|t_1-t_2|}$. With a larger α, the correlation between these two samples is smaller. This way, α is similar to λ parameter used in Sect. 2, the rate at which the state is updated. As α increases, the correlation between the most recently reported state value and the current value of the state decreases. The strong state approach turns into modeling a random variable with another (almost) uncorrelated random variable that pertains to the same mean and variance. In this case, the instantaneous distortion quickly becomes twice the variance of the random variables. On the other hand, refreshing the state locally limits the instantaneous distortion to variance of the state. Pure and informed distortion values are affected correspondingly. Figure 5 also presents the way the state refresh interval T affects the distortion values when α is taken constant. The figure shows that the distortion values increase with the refresh interval. Because the value of the weak state can be locally refreshed, it always causes lower pure distortion. Weak state causes slightly higher informed distortion when T is small since the receiver can misinterpret an accurate state value as invalid because of the probabilistic confidence value. However, it is much more advantageous for protocols that cannot afford to refresh remote state at a very high rate.

4.3 Poisson process

In a Poisson process with parameter λ, given that $\mathbf{x}(t_0) = x_0$, the value of the process at time $t_0 + t$ equals $x_0 + n_t$ where n_t is a Poisson random variable with parameter λt:

$$P(n_t = k) := \frac{\exp(-\lambda t)(\lambda t)^k}{k!}.$$

The expected value of $\mathbf{x}(t)$ is $x_0 + \lambda t$ whereas its variance equals to λt [18]. This process can be relevant in a traffic monitoring sensor network where the sensor nodes report the number of vehicles they detect [6].

Note that considering only the equity of the state values at the sender and the receiver makes the distortion identical to that in Sect. 2. Here, we are also interested in the difference between these values. From (10) and (11), the instantaneous distortion becomes

$$e_S(t) = \begin{cases} \lambda t + (\lambda t)^2 & t \le \chi \\ C & t > \chi \end{cases} \tag{22}$$

$$e_W(t) = \begin{cases} \lambda t & t \le \chi \\ C & t > \chi \end{cases} \tag{23}$$

given $t_0 = 0$. The corresponding informed distortion values can be obtained in a way similar to the previous process examples. For $t \le \chi$, the time cumulative distortion is

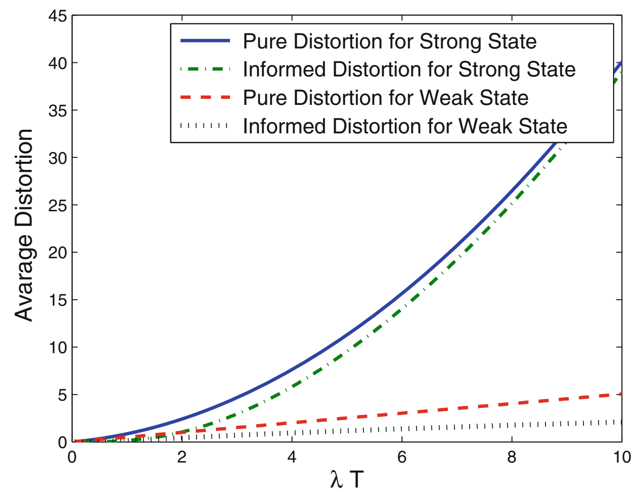


Fig. 6 Distortion values with respect to λT product in Poisson process. Weak state can statistically estimate the current value of the state; hence it causes lower pure distortion. When the sender's state is dynamic (large λ) and/or the receiver's state is refreshed at low rate (large T) due to overhead concerns, weak state causes lower informed distortion

$$E_S(t) = \frac{1}{2}\lambda t^2 + \frac{1}{3}\lambda^2 t^3$$
$$E_W(t) = \frac{1}{2}\lambda t^2.$$

The average pure distortion, considering the refresh packet losses, can be derived as follows

$$\overline{e^P} = \sum_{i=1}^{\kappa} \frac{1}{iT} p^{i-1}(1-p)E(iT) + \sum_{i=\kappa+1}^{\infty} \frac{1}{iT} p^{i-1}(1-p)[E(\chi) + C(iT-\chi)]$$

where κ is defined as before. Informed distortion for strong state and weak state can be obtained by plugging (22) and (23) into (13) and (14), respectively. The following default values are used to obtain the distortion results: $\lambda = 0.1$, $p = 0.02$, $T = 10s$, $\rho = 1$, and $\chi = 3T$ both for weak state and strong state.

Figure 6 presents the distortion values for weak and strong state with respect to the λT product since the value of the process at t, $\theta(t)$ and $\Theta(t)$ are all governed by λt. Because it statistically estimates the current value of the state, weak state always causes lower pure distortion. As in previous cases, the informed distortion for weak state can be slightly higher when the dynamism in the state is low (small λ) or the receiver's state is frequently refreshed (small T). However, when the sender's state changes more dynamically, weak state causes lower informed distortion in comparison to strong state. Also, if the protocol chooses to refresh the remote state at a low rate due to concerns

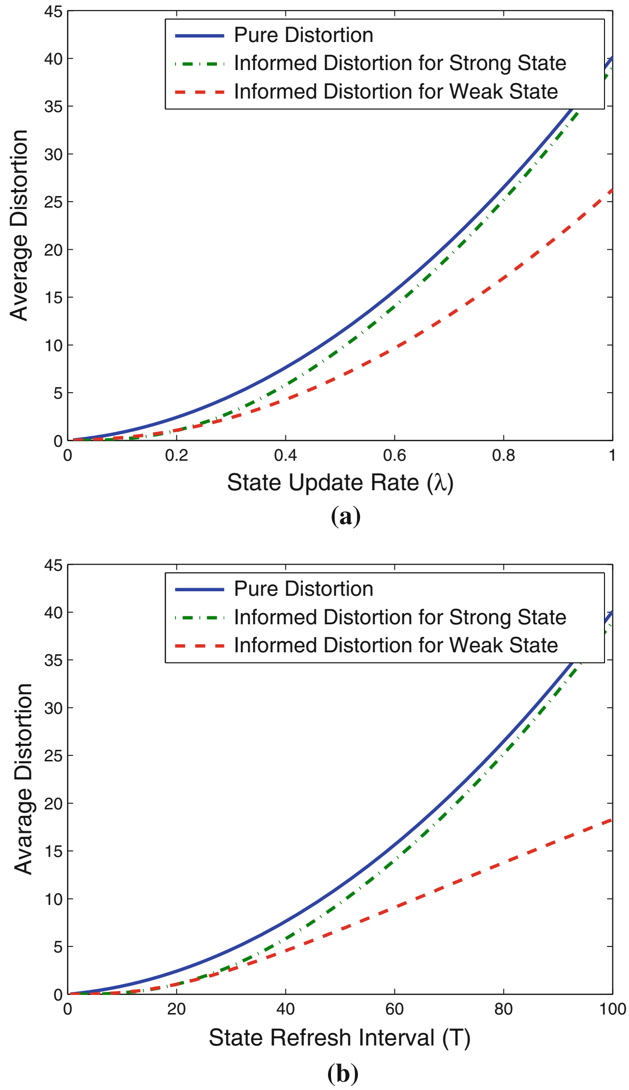


Fig. 7 Distortion metrics with respect to the process parameter λ and state refresh interval T when the receiver cannot use estimates. When the state is dynamic or refreshed at a lower rate, informed distortion is lower for weak state. **a** Distortion versus λ. **b** Distortion versus T

regarding the overhead, weak state reduces the average informed distortion.

One might argue that the advantages of weak state are due to using estimates at the receiver rather than the probabilistic semantics, and estimates may not always be available if the receiver is not aware of the properties of the state at the sender. We address this in Fig. 7 where we assumed that no information about the stochastic process is available and used $\hat{\mathbf{x}}_W(t) = \hat{\mathbf{x}}_S(t) = x_0$. Because of the lack of information about the underlying process, confidence value can no longer represent the exact probability that the state at the receiver is within the tolerable interval of the sender's state. Instead, $\theta(t)$ provides a measure of this probability. We used $\theta(t) = q^t$, where $0 < q < 1$ and adjusted the informed distortion calculation accordingly. Figure 7(a), (b) show that the pure distortion is now identical for weak state and strong state since weak state cannot provide an estimate of the current value of the sender's state. Previous observations regarding the informed distortion still hold although the margin by which the distortion of strong state exceeds that of weak state is smaller than before[4].

4.4 Discussion on assumptions

The analysis in this paper makes several assumptions for convenience. Some of them are related to deducing the paths the refresh messages follow to single logical hops. In reality, the paths of different size potentially lead to different values for parameters such as message loss probability. In addition, the analysis assumes that each message is delivered to the receiver over the logical single hop independently with identical probabilities. In reality, if the communication between the end points break down, multiple messages are likely to fail. Also, refresh messages that are transferred on longer paths are dropped with higher probability. Other assumptions include that there is no end-to-end signaling delay, that the receiver is aware of the underlying process of the state at the sender and that the sender always maintains the state. In the next section, these assumptions are not adopted.

In our analysis, we use a signaling mechanism in which the sender periodically sends refresh messages to the receiver, i.e. soft state. Note that what we propose in this paper is a probabilistic state maintenance mechanism, which is orthogonal to the way the sender signals the state to the receiver. It is possible to use this mechanism together with hard state and soft state with triggering mechanisms (i.e. sender sends a refresh message when the change in the state value is above some threshold), which both reduce the distortion, but such a reduction comes at the cost of extra overhead. Ji et al. [9] already show that soft state uses fewer refresh messages than hard state and soft state with triggering approaches require in order to achieve a target level of distortion. Hence, our claim that weak state achieves the same amount of distortion using much less messages in comparison to strong state holds for all signaling mechanisms.

5 Experimental evaluation

In this section, we conduct simulation experiments of actual networking scenarios that do not attempt to follow

[4] Note that in this case, the state update interval λ has no effect on the confidence value. As a result, scaling λ and T yields different informed distortion values for weak state. Thus, we show the effect of these two parameters in two separate figures.

the assumptions in the previous sections. First, we show that the protocol performance benefits from low distortion. To do this, we use WSR for routing in mobile networks [3]. Then, we evaluate state distortion for weak state and strong state in a number of scenarios using simulations.

In all the simulations, we use the realistic vehicular networking tool TraNs to simulate node mobility [20]. TraNs integrates traffic and network simulators, SUMO and ns2, respectively. SUMO simulates the road traffic for a number of users on a given road map whereas ns2 is a packet level network simulator. In full implementation, there is a bi-directional channel between the two simulators, i.e. the communication between two users can change their mobility. However, we used its stripped-down version for generating realistic mobility traces from SUMO. These traces later on are fed into the ns2. Our scenario consists of 500 nodes that move in a 3,000 $\times$ 3,000 area in Manhattan, NY, USA for 1,000 s.

5.1 Distortion versus protocol performance

In this section, we demonstrate that the protocol performance is tied to the informed distortion metric using a routing protocol for mobile ad-hoc networks, WSR. In WSR, the nodes periodically send information about their location in random directions. Hence, for every sender, every other node may or may not be a receiver. The state information is characterized by a probabilistic bloom filter which also yields the confidence. Once a node learns the location of the destination, it instantiates a weak state with full confidence that corresponds to the destination and the confidence is decayed in time. Once a node receives a packet, it first checks whether it maintains information about the destination. If the confidence in the local information is greater than what is carried on the packet header, the intermediate node biases the direction in which packet is forwarded. Otherwise, it merely relays the packet in a way that it advances towards the direction yielded in the packet header.

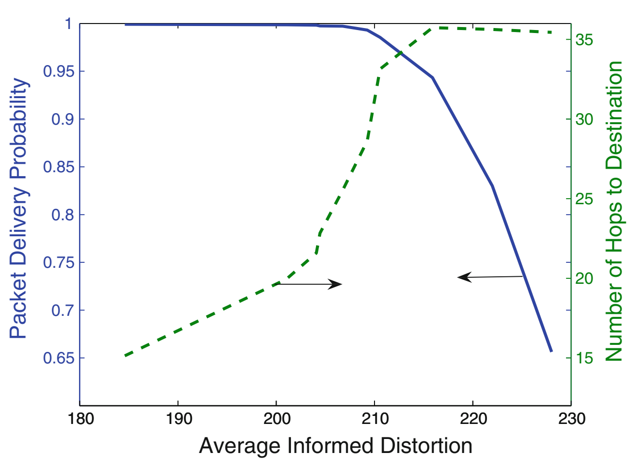


Fig. 8 WSR performance in terms of packet delivery ratio (*solid curve*) and the average number of times a packet is forwarded until the destination (*dashed curve*) against the average informed distortion. The y-axis on the left gives the figures for the delivery ratio and the y-axis on the right shows the values for the number of hops. WSR performs better in both metrics if the informed distortion is low

In Fig. 8, we show how the performance in terms of packet delivery ratio and the average number of hops packets are forwarded until the destination is related to informed distortion. Plotting this figure, each data point corresponds to a separate run rather than an average of multiple runs. The figure shows both performance metrics benefit from low informed distortion. When the informed distortion is low, the packet delivery ratio is high and it takes packets smaller number of hops to reach the destination. On the other hand, as the distortion increases, delivery ratio drops and the successfully delivered packets takes larger number of hops to destination on average.

5.2 Evaluation of distortion

In this section, we consider a scenario in which a base station maintains a state entry for every node that is located within a smaller 2, 000 $\times$ 2, 000 region A, inside the larger area. A sender sends refresh messages as long as it is located in A. If the receiver keeps a state entry for a mobile node that is no longer in A, the state becomes *orphaned*, which also causes distortion. Premature timeout of the state before it is orphaned is also associated with distortion. As in previous sections, weak state and strong state have the same timeout interval given all the other parameters are identical. Nodes use geographical routing to send their refresh messages over a lossy wireless medium characterized by a bit error rate of 0.001. The refresh messages can also be lost due to MAC layer errors and/or lack of paths to the base station.

For weak state, the confidence value decays exponentially with time such that $\theta(t) = q^t$, where t is the time elapsed since the state at the base station is last refreshed and q is the decaying constant. Since there is no information available regarding the properties of node mobility, the receiver cannot adjust the exact probability that the state at the sender remains unchanged. Moreover, local refreshes do not consist of estimating the current locations of the nodes. The perceived entry for weak state is the most recent value reported by the corresponding node, similar to the strong state.

We consider two location tracking applications. In the first scenario, a central base station tracks down which subregion each node is located in. An example for this scenario is presented in KioskNet, where mobile users are associated with kiosks [8]. The association for a user is

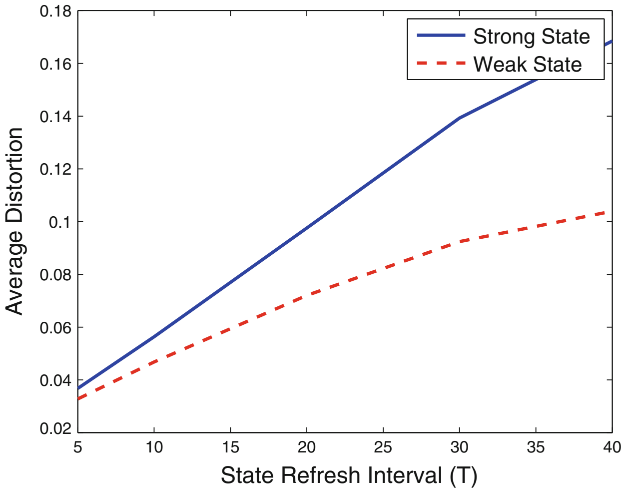


Fig. 9 Distortion values with respect to state refresh interval (T) when the state information yields the region in which nodes are located (*discrete* state) as perceived by the base station. The average distortion for weak state is lower because it involves the probabilistic confidence value

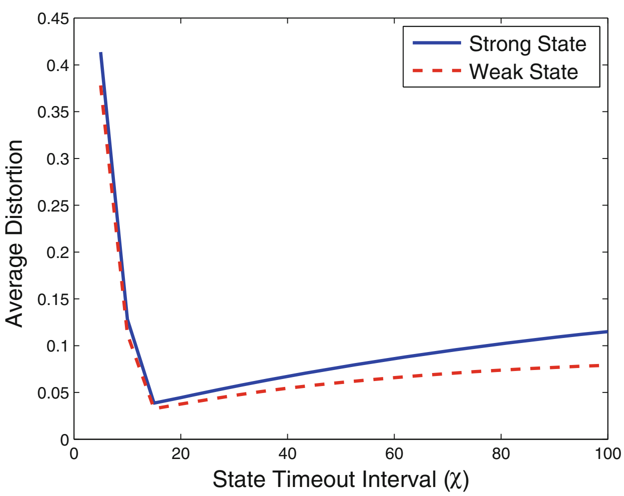


Fig. 10 Distortion with respect to state timeout interval (χ) when a entry state yields the sub-region a node is located in (*discrete* state) as perceived by the base station. Weak state causes lower distortion than strong state for any value of χ

based on its proximity to the kiosks. The packets that are destined to a mobile user is transferred to its corresponding kiosk. We divide A into four sub-regions each with a kiosk; and the state entry that corresponds to a particular node yield the sub-region that the node is located in. In the second set, the base station tracks the geographical coordinates of each node in the network, similar to the location servers in geographical routing. Once a new packet is initialized, the source node queries the location of the destination from the base station and uses this information in forwarding [12]. In the first scenario, the base station is only interested in whether the node is actually located in

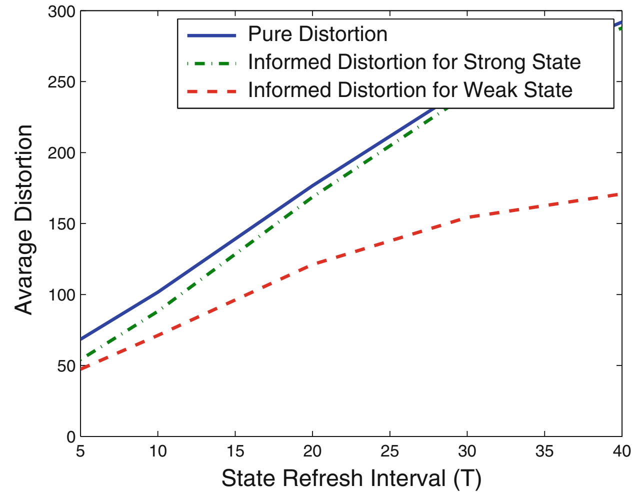


Fig. 11 Distortion with respect to state refresh interval (T) when the state information yields the node location (*continuous* state) as perceived by the base station. Since no estimate is used, pure distortion is the same for either case. When the state is frequently refreshed by its sender, strong state may yields better informed distortion at the cost of increased number of refresh messages. When the state is less frequently refreshed, weak state causes lower informed distortion

the region of its associated kiosk. In the second scenario, the difference between the actual position and the perceived position is important. Hence, the first scenario reflects the analysis in Sect. 2 whereas the second scenario is compatible with the analysis in Sects. 3 and 4.

Figure 9 shows how the distortion values changes with respect to state refresh interval, T, in the first scenario. For strong state, the state at the base station is removed if it is not refreshed for a duration of $3T$. For weak state, the γ parameter is adjusted appropriately to have the same timeout interval. The distortion for weak state is $1 - q^t$ when a sender is located in the region given by the state at the base station. In this case, strong state does not cause any distortion. Neither strong state nor weak state causes distortion when a node is out of region A and the base station does not maintain any state information for that node. If the state entry indicates that the node is in A when it is actually not, the distortion is equal to 1 for strong state and q^t for weak state. These values correspond to the probabilities with which the receiver thinks that the node is inside A when it is not. Figure 9 shows the average distortion with respect to state refresh interval. To obtain this figure, we used $q = 0.99$. Similar to our analysis in Sect. 2, weak state causes lower distortion because it provides probabilistic information to the base station.

Figure 10 shows the average distortion with respect to state timeout interval χ when the state refresh interval is set to $T = 10$. When χ is very small, the state at the receiver is removed very quickly in both cases even though the state maintained at the receiver is still legitimate. This causes

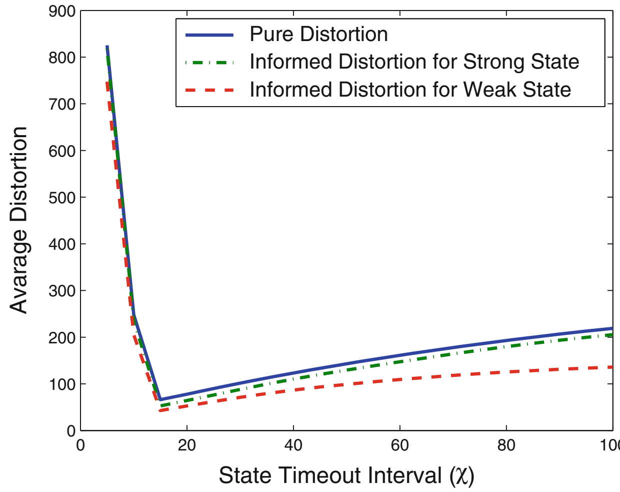


Fig. 12 Distortion with respect to state timeout interval (χ) when the state information yields the node location (*continuous* state) as perceived by the base station. Pure distortion is the same for both weak state and strong state due to lack of estimates. Both mechanisms show similar characteristics against χ, yet weak state causes lower informed distortion because of utilizing probabilistic confidence value

higher distortion in both cases. When the timeout interval is slightly more than the state refresh interval, both strong state and weak state cause their smallest distortion values. For large values of timeout value, the state at the base station is more likely to be orphaned. As a result, the distortion increases with the state timeout interval. However, for any value of χ, the distortion weak state causes is lower than the distortion for strong state.

Similar to the prior case, a node issues refresh messages only if it is located within A in the second scenario. Let $\hat{x}(t)_a$ denote the location entry for node a at time t in the base station. The instantaneous distortion at time t for node a is $|x(t)_a - \hat{x}(t)_a|$, the distance between the actual position and the maintained position. The calculation of pure distortion and informed distortion is similar to that in Sect. 3. In addition, we now also consider the case that the base station might keep a state about a node when the node is outside the region A. In this situation, the instantaneous distortion equals C[5]. We calculate the average distortion values as given in Sect. 3. Obtaining this figure, we set $\chi = 3T$ for both strong state and weak state and $\rho = 50$. In Fig. 11, we compare the distortion values of strong state and weak state against state refresh interval. Since weak state cannot use any prediction, weak state and strong state cause identical pure distortion values. Similar to the previous cases, informed distortion is significantly lower for weak state when large values of T is used.

In Fig. 12, we compare weak state and strong state against the state timeout interval χ when $T = 10$ s. The pure distortion is again identical in both cases since both weak state and strong state use the last reported value. When χ is very small, the base station removes the valid state prematurely when the sender is still in A, which leads to higher distortion. Informed distortion also increases with increasing χ since the receiver takes a longer time to detect the orphaned state. Still, with the probabilistic confidence value, informed distortion is lower for weak state.

Note that state timeout interval for weak state can be also determined by setting a constant γ while varying the q. However, q captures how fast the state at the sender diverges from the previously reported value. Since the dynamism in node mobility remains fixed, we use a constant q value.

6 Conclusion

In this paper, we study the consistency of weak state, which is characterized by local refreshes and probabilistic semantics unlike traditional strong state where the information is regarded as absolute truth. If the state provides information about a dynamic process, strong state quickly becomes invalidated. On the other hand, weak state is more stable and consistent. In order to evaluate the consistency of the state, we use two notions of distortion. Pure distortion quantifies the average difference between the actual state value and value maintained at the remote node. Protocols that rely on weak state also utilize the information about the confidence in the state value. Even if the pure distortion is large, a communication protocol that utilizes weak states can take the necessary steps to improve the performance if the confidence is low. We capture this effect as well as the difference between the state values at the sender and the receiver using the informed distortion metric.

We present an analysis to compare/contrast weak state and strong state in terms of consistency, using pure and informed distortion metrics. Our mathematical analysis and numerical analysis based on simulations show that weak state reduces both pure distortion and informed distortion significantly when it provides information about a dynamic entity. When the state is refreshed at a very high rate, the distortion for strong state is lower. However, this comes at the cost of increased overhead which reduces the data capacity of the network. When the network cannot tolerate a large amount of overhead to refresh state information in remote locations, weak state is a useful alternative to achieve consistency. It is a more appropriate design block than strong state for protocols that need to maintain information about dynamically changing entities but not cause high overhead.

In this paper, we do not propose a particular protocol that interprets the confidence in the state value. Instead, we

[5] In the example scenario, we set C to the maximum distance between two points in the region A, i.e. $C = 2,000\sqrt{2}$

show weak state provides more consistent information about a dynamic entity than the strong state does. In future work, we plan to develop a generic protocol that can use both weak state and strong state and compare the distortion values in the context of that protocol. This will also allow us to shed light on the effect of weak state on other aspects of protocol design and performance. In addition, we will develop mechanisms that will gather information about the state information so that the protocol can calculate the distortion and adaptively change from weak state to strong state or vice versa to operate with more consistent state.

Acknowledgments This material is based upon work supported by the National Science Foundation under Grant Nos. 0627039 and 0322956.

References

1. Abouzeid, A., & Bisnik, N. (2011). Geographic protocol information and capacity deficit in mobile wireless ad hoc networks. *IEEE Transactions on Information Theory, 57*(8), 5133–5150.
2. Acer, U. G., Abouzeid, A., & Kalyanaraman, S. (2009). An evaluation of weak state mechanism design for indirection in dynamic networks. In *INFOCOM*.
3. Acer, U. G., Kalyanaraman, S., & Abouzeid, A. A. (2010). Weak state routing for large-scale dynamic networks. *IEEE/ACM Transactions on Networking, 18*(5), 1450–1463.
4. Acer, U. G., Kalyanaraman, S., & Abouzeid, A. A. (2011) DTN routing using explicit and probabilistic routing table states In Wireless Networks.
5. Clark, D. D. (1988). The design philosophy of the DARPA internet protocols. In *SIGCOMM* (pp. 106–114). ACM
6. Coleri, S., Cheung, S. Y., & Varaiya, P. (2004). Sensor networks for monitoring traffic. In: *Allerton conference on communication, control and computing*.
7. Grossglauser, M., & Vetterli, M. (2006). Locating mobile nodes with EASE: Learning efficient routes from encounter histories alone. *IEEE/ACM Transactions on Networking, 14*(3), 457–469.
8. Guo, S., Falaki, M. H., Oliver, E. A., Rahman, S. U., Seth, A., Zaharia, M. A., et al. (2007). Design and implementation of the kiosknet system. In *International conference on information technologies and development*.
9. Ji, P., Ge, Z., Kurose, J., & Towsley, D. (2007). A comparison of hard-state and soft-state signaling protocols. *IEEE/ACM Transactions on Networking, 15*(2), 281–294.
10. Keshav, S. (2006). Efficient and decentralized computation of approximate global state. *Computer Communication Review, 36*(1), 69–74.
11. Kumar, A., Xu, J., & Zegura, E. W. (2005). Efficient and scalable query routing for unstructured peer-to-peer networks. In *INFOCOM*.
12. Li, J., Jannotti, J., De Couto, D., Karger, D., & Morris, R. (2000). A scalable location service for geographic ad hoc routing. In *MobiCom*.
13. Liang, B., & Haas, Z. J. (2003). Predictive distance-based mobility management for multidimensional pcs networks. *IEEE/ACM Transactions on Networking, 11*, 718–732.
14. Lindgren, A., Doria, A., & Schelen, O. (2004). Probabilistic routing in intermittently connected networks. In *First international workshop on service assurance with partial and intermittent resources (SAPIR)*.
15. Lui, J. C. S., Misra, V., & Rubenstein, D. (2004). On the robustness of soft state protocols. In *International conference on network protocols, ICNP*.
16. Manfredi, V., Crovella, M., & Kurose, J. (2011) Understanding stateful vs stateless communication strategies for ad hoc networks. In *MobiCom* (pp. 313–324).
17. Oppenheim, A. V., Willsky, A. S., & Nawab, S. H. (1996). *Signals & systems* (2nd ed.). Upper Saddle River, NJ, USA: Prentice-Hall, Inc.
18. Papoulis, A., & Pillai, U. S. (2001). Probability, random variables and stochastic processes. McGraw-Hill Science/Engineering/Math, December.
19. Perkins, C. E., & Royer, E. (1999). Ad-hoc on-demand distance vector routing. In *IEEE WMCSA*.
20. Piorkowski, M., Raya, M., Lugo, A., Papadimitratos, P., Grossglauser, M., & Hubaux, J.-P. (2008). TraNS: Realistic joint traffic and network simulator for VANETs. *ACM SIGMOBILE Mobile Computing and Communications Review, 12*(1), 31–33.
21. Raman, S., & McCanne, S. (1999). Model, analysis, and protocol framework for soft state-based communication. *Computer Communication Review, 29*(4), 15–25.
22. Richard, C., Perkins, C., & Westphal, C. (2005). Defining an optimal active route timeout for the AODV routing. In *SECON*.