Epidemic Enhanced Cellular Networks

Veeraruna Kavitha[1], Eitan Altman[2], Sreenath Ramanath[3]
[1]IEOR, IIT Bombay, Mumbai, India; [2]INRIA, Sophia Antipolis, France;
[3]Lekha Wireless Solutions, Bangalore, India;

***Abstract*—When the base station density is sparse, users are often out of the coverage area of the cellular networks. In such scenarios, the users can rely on fellow users (willing to relay) to deliver delay tolerant information to a base station. Further, the users (and relays) can observe independent environments, during their traverse, if they are moving at considerable speeds. This provides them an opportunity to (independent) search for relays and or base stations at regular intervals of time. However, the user should beacon (transmit short pulses and wait for response), when it desires to be detected by the available relays in its neighbourhood. We derive the performance of such a system, accounting for the power utilized for beaconing, and obtain the beaconing policies that maximize the success (user/any previously contacted relay comes in contact with one of the base stations) probability of message delivery. The base stations and relays (at any given instance of time) are randomly distributed according to a Poisson Point process.**

We formulate the problem as a Markov Decision Process (MDP) to derive optimal policies depending on the system state (closed-loop). We show that the value function satisfies certain monotonicity properties and the optimal policy exhibits a certain switch-off property. We obtain an approximate solution for the continuous control and an exact solution for the ON-OFF (two action) control. Our investigations show that the closed-loop ON-OFF policies perform (almost) as good as the closed-loop continuous policies for all practically viable test cases. We further investigate open-loop policies for ON-OFF control, the policies that can be used when the system state is not known.

Keywords: Markov Decision Processes; Open and Closed-loop policies; Poisson Point process; Delay Tolerant Networks;

I. Introduction

Urban areas offer wireless infrastructure to satisfy unprecedented data rates and user experience. However the base station (BS) density is sparse in less populated areas (rural areas, highways, etc.), where emphasis is on providing basic services. But the users might still require enhanced services, and some of the applications (file transfers, chats, emails, etc.), can tolerate delays. Our emphasis in this work is to improve the (uplink) services offered to the users of a sparse network, when they can tolerate delays. Towards this we propose an epidemic enhanced cellular network: users can transfer messages to fellow (visible) users, if the latter is willing to become a relay. When one such relay encounters a BS on its path, it transfers the message to the later. *We call such users as delay tolerant or DT users. Thus we propose an epidemic aided cellular network, which is very different from traditional DTNs: the purpose here is to transfer the message to any one of the base stations located on a Poisson grid.*

We consider a cellular network where BSs are distributed according to a Poison Point process (PPP) and assume that the time required for transferring message from any BS to the final destination is negligible. Time required to transfer the same message to a BS (any one of them) depends on the contacts made, which are highly random. If a BS is within the range of a DT user, it can pass the message to BS directly which in turn will route the message through its network to the destination instantaneously. Otherwise, the user has to pass the message through alternate routes and our work mainly focuses on this.

A question that arises when considering delay tolerant information transfer via cellular networks is, *whether the existing mechanisms can aid such transfers?* For example, Device to Device (D2D) overlays, introduced in LTE Networks, allow static users to act as relays and enable information transfer for out of coverage users (see [3]). *We propose to extend these kind of protocols to enable mobile relays, which can convey information to one of the encountered BSs (in future)* .

Any BS regularly transmits signals to broadcast information, which is also useful in detecting its presence. This is a common practice in cellular networks. *If a user has to be detected by a relay (without the aid of BSs), it has to beacon to make its presence felt.* This expends power. Power is a scarce resource, which has a great impact on the user and the network performance. Thus an efficient power control is required. We assume *negligible power is used for actual information transfer (occurs few times and to nearby relays), most of the power is used for regular signalling in the form of beacons.* We thus study the control of beacon power.

The main objective is to derive optimal policies describing the power utilized at various beaconing instances, using MDP based (closed-loop) approach. Open-loop policies (not depend upon state) are also considered. We discuss quantitatively the additional advantage obtained by epidemic aided approach. Numerical analysis and simulations are carried out comparing closed-loop continuous, closed-loop ON-OFF (two control) and open-loop ON-OFF policies. *The ON-OFF policies perform almost similar to the continuous policies for all practically viable cases. Thus it is sufficient to use much simpler ON-OFF policies. For many examples open-loop policies also suffice* (mostly when the power constraint is high).

Related research: There has been a large body of research on the control of delay tolerant networks (DTNs). The main objective in the control of DTNs is to use a limited energy/power and then to minimize delays of successful transfers, and or to maximize the system throughput. While our work reminds the work on DTNs, there is a fundamental difference between the two. In DTNs one considers either the time till a single predefined destination (peer-to-peer) is reached or the time till a given set of destinations (broadcast) is reached (e.g., [4], [5], [6], [7]). *We consider, in contrast, the time till the reception by one of the many base stations. We consider a transmission to be complete when (atleast) one of the destinations (BSs here) receive the information. This is an important scenario which is not studied before.*

Majority of the work mentioned above, considers information transfer with no-relevance to cellular networks. A related piece of work considers offloading of delay tolerant traffic of the cellular network by other means (e.g., [1]), including the epidemic based opportunistic forwarding (like our case). Here the focus is on controlled offloading, which also considers fall-back to the cellular back-bone (if required) to ensure minimal QoS guarantees (see recent survey [1] and references there in). While our work focuses on similar epidemic based transfers, but, in the absence of cellular coverage.

In another recent survey ([2] and references there in), the authors discuss mobility as an alternate communication channel when the users are not in the coverage area and when the information is delay tolerant. This scenario is similar to our case, however their focus is on piggybacking relatively large amount of data on storage devices mounted on mobile entities that move on a well defined and well known traces (e.g., ferry based wireless LANs). In contrast we consider data traffic of a specific user moving in an unknown (cellular) and sparse network and derive aid from mobile relays who are also unaware of the network topology (e.g., the existence/non-existence of base stations in their future trajectories).

II. Epidemic enhanced Network and assumptions

If the BS is not in the range of a (high speed) moving user, and if the user is tolerant towards delays in its message transfer, it would search for a relay. If a relay is found, the message is transferred to the detected relay. The relay, if it comes across a BS before the remaining deadline, will transfer the message to the BS. And this continues with more relays.

DTN protocol: DTNs operate with different transfer protocols. A relay transfers message to one of the BSs, but not to other relays in a two-hop protocol. Alternatively DTNs can use full epidemics, i.e., the relays can transfer the message to another relay. Success probability (message delivery) increases with full epidemics, but, power consumption is high and there might be flooding of messages. *We use a two-hop protocol.*

Beaconing: The DT user is travelling across a major street/high way at considerable speeds. It continuously attempts to transfer the message directly to the nearest BS, and this is an uncontrolled part. Apart from this, information or message transfer to a *nearby* relay is attempted at regular intervals of time. Here we call this attempt (transmits short duration pulses and waits for a response) as beaconing. At every beacon instance, the DT user tries to find a relay and transfers its message if it has found one. The range of detection, at any beacon instance, depends on the transmission power. The region of detection (by BS/user/relay) is assumed to be circular with radius equal to the range of detection. If one uses more transmit power, the radius of the detectable region increases. However higher the power per beacon, lesser the number of beaconing chances (as there would be a cost/constraint on the

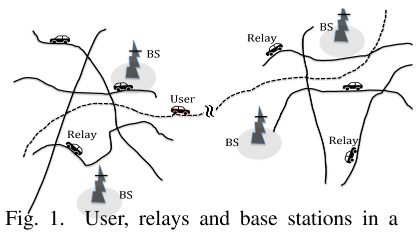

Fig. 1. User, relays and base stations in a sparse network and at two distant locations

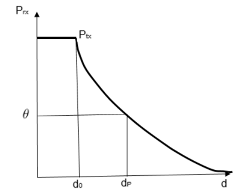

Fig. 2. Detectable distance d_P

power used). In this context, we obtain optimal beaconing policies, basically the sequence of transmission powers and the number of beaconing attempts. We derive closed-loop (MDP based) as well as open-loop policies.

We consider scenarios in which the BSs are sparsely distributed and *the chances of a direct message transfer is negligible*[1]. *Our initial and main focus is on the events related to epidemic aided transfer.* However if a BS is found, the user need not continue with the beaconing efforts. This can also have an impact on the 'optimal' beaconing pattern. We consider such scenarios briefly in Section VII.

More assumptions: We consider mobile relays[2]. At any snapshot of time (beacon instance) and at any point in the network, we assume that the relay positions are governed by a Poisson Point Process (PPP). See Figure 1. For example, if the relays are initially distributed according to a PPP and are moving independently and in all (uniform) random directions, their positions after any time period are again governed by a PPP. Note that the PPP is translation invariant. Further at any point of time, the user if detects a relay, can find the latter moving in a random direction which is uniformly distributed between 0-360 degrees. In other words, at any point in the two dimensional space, one can find a street passing through the neighbourhood (if there is one), in a direction which is uniformly distributed.

One can use DTN concept only if the user can see different (and independent) environments at various beacon instances. *The user is also moving at considerable (and random) speed, such that it observers independent environments at beacon intervals. At any beacon instance, the user transfers the message to at most one relay.*

The BSs are obviously not power deficient, transmit information at regular intervals of time, because of which any nearby user can detect it. In fact in cellular networks this happens automatically, once the user enters the vicinity of a BS. *However for a communication between two users (one as a relay), one of them (the user interested in a favor) should transmit beacons to ensure that the other detects it.*

III. Performance

We derive some important performance measures, which would be the ingredients for deriving the optimal policies. We begin with the discussions related to the received power and

[1]Chances of a relay (or user) detecting a BS can be negligible, but that by one of the many (contacted and independent) relays can be significant.

[2]Our analysis is applicable for highway kind of scenarios which are characterized by: a) low base station density; b) all the entities are mobile, i.e., one rarely finds a stationary element.

the detectable distance. Power received by a relay depends on the distance from the transmitting user d, power transmitted p (e.g. [8]),

$$P_{rx}(p,d) = \begin{cases} P_{tx} & \text{if } d < d_0 \\ P_{tx}(\frac{d}{d_0})^{-\beta} & \text{else}, \end{cases} \quad (1)$$

where d_0 is the lossless distance (usually around 10 meters) and β is the propagation co-efficient (see Figure 2). We are considering short range communications and sparse connectivity (thus no interference), and hence this free space model will suffice. The path-loss factor β depends on the wireless medium and usually ranges between 2 to 4.

A. Detectable distance (d_p):

Relay detects the user only if the received power is greater than a given threshold θ. We define the *detectable distance* as the maximum distance until which a relay can detect the user. Clearly if power transmitted itself is less than θ, no relay can detect the user and hence the detectable distance is 0. On the other hand if $p > \theta$, then the mobile is detected at a distance d, if the received power P_{rx} at that point is greater than or equal to θ. Thus from (1), the maximum distance up to which a user can be detected, with transmit power p, equals:

$$d_p \quad = \quad d_0 \left(\frac{p}{\theta}\right)^{1/\beta} 1_{\{p \geq \theta\}}. \quad (2)$$

B. Beaconing policy and relay detection

We consider a T-time slot problem, wherein one needs to decide the beaconing powers for each time slot. We refer a sequence of powers $\boldsymbol{\pi} = (P_1, \cdots, P_{T-1})$ as a beaconing policy, where P_t is the power used for beaconing at the beginning of t-time slot. As seen from (2) the range of detection depends on the power spent for detection. As the transmit power increases, the range (area of detection) increases resulting in a better chance of detecting a relay. The mobiles usually operate under power budget. So, if more power is used per beaconing attempt, then we may be able to beacon only fewer times. Thus one requires a good beaconing policy. These power choices can depend on the appropriate system state and hence we consider an MDP based approach to derive optimal policies in the next section. The open-loop policies are in later sections.

Relay Detection: The relays are Poisson distributed in the network with parameter λ_r. Thus the probability of finding k relays in the region A is Poisson distributed with parameter $\lambda_r|A|$ and hence equals:

$$e^{-\lambda_r|A|}\frac{(\lambda_r|A|)^k}{k!},$$

where $|A|$ represents the area in two dimension and length in one dimension. When mobile beacons using power P, the area covered equals πd_P^2, where d_P is the detectable distance given by the equation (2). Hence no relay is detected by the user if there are zero relays in a disc of radius d_P. Thus the probability of failure to detect a relay (in two dimensions) in time slot t, with policy $\boldsymbol{\pi}$ equals (see (2)):

$$P^{\boldsymbol{\pi}}(U_{noR}^t) = e^{-\lambda_r \pi d_{P_t}^2} = e^{-cP_t^{2/\beta}} \text{with } c := \lambda_r d_0^2 \pi / \theta^{2/\beta}, \quad (3)$$

where P^{π} is the probability under policy $\pi = (P_1, \cdots, P_{T-1})$ and U_{noR}^t is the event that the user has detected zero relays at the beginning of the time slot t.

C. Overall Failure Probability

Whenever the DT user is within a radius r_b (reference radius up to which the BS can be detected by the users) from a BS, it can directly transfer the information to the BS and the BS will route it through its network and deliver it to the destination. As already mentioned, such chances are rare, and we initially neglect these chances. In later sections, we consider the influence of detecting a BS directly. Note that r_b is the same distance within which any relay can also detect a BS, however since our aim is to contact considerable number of relays, the chances of one of the relays contacting one of the BSs can be significant (based on the beacon policy) and our initial focus is precisely on these events.

The user is moving at considerable speeds and hence we assume it to observe independent environments at the beginning of different time slots. The relays are also moving at considerable speeds but in different directions and hence the events related to the relays (those contacted) of different time slots will also be independent. We use the *independence of these events to derive the probability of failure of message transfer* before deadline, in-spite of the relay contacts.

The DT user will try to communicate with one of the relays, in case one of them is within the contact radius created by the transmitted beacons. Upon a successful contact and subsequent message transfer in the t^{th} beacon attempt, the recipient relay will attempt to pass on the information to a BS during the leftover time. Thus the message is not transferred within the deadline if the following events occur (neglecting direct transfer chances): a)

1) User has detected zero relays in all the beacon chances, i.e., the event $\cap_t U_{noR}^t$;
2) User has detected a relay at some time slot t, but all such relays could not transfer the message within the left over (deadline) time, the later event represented by R_{noBS}^t.

Thus the probability of failure for the delivery of message for any beaconing policy $\boldsymbol{\pi}$ is given by:

$$P_f = P^{\boldsymbol{\pi}}\left(\bigcap_{t=1}^T \left[U_{noR}^t \bigcup \ \left\{(U_{noR}^t)^c \cap R_{noBS}^t\right\}\right]\right). \quad (4)$$

D. Relay failure

Let $\gamma_t = P(R_{noBS}^t)$ *be the probability that a relay fails to transfer the message to one of the BSs in the remaining time before deadline,* $t_r^s = (T-s)D$*, with* D *the duration of one time slot.* Note that *these probabilities do not depend on the policy* $\boldsymbol{\pi}$, but depends only on the mobility profile of the relays and the density of BSs. *One can assume these constants are given in general.* Below we compute these probabilities for one example scenario. One can compute such probabilities even when the relays are randomly wandering as modelled by a Brownian motion or when the mobility pattern of different relays can be of different nature.

Computation of $\{\gamma_t\}_t$ *for uniform relay velocities:* As relay velocity is random, we condition on the same to obtain relay message delivery failure probabilities $\{\gamma_t\}_t$. We assume that the velocity of the relay is uniformly distributed between

$V_{max} = \overline{V}$ and $V_{min} = \underline{V}$. By conditioning on velocity of relay, we have (with $t_r^t := (T-t)D$):

$$\gamma_t = \int_{\underline{V}}^{\overline{V}} e^{-\lambda(r_b v t_r^t)} f_V(v) dv = \int_{\underline{V}}^{\overline{V}} e^{-\lambda(r_b v t_r^t)} \frac{dv}{\overline{V} - \underline{V}}$$
$$= \frac{1}{\overline{V} - \underline{V}} \int_{\underline{V}}^{\overline{V}} e^{-\lambda r_b v t_r^t} dv = \frac{e^{-\lambda r_b \underline{V} t_r^t} - e^{-\lambda r_b \overline{V} t_r^t}}{(\overline{V} - \underline{V})\lambda r_b t_r^t}.$$

In the time (t_r^t) remaining before the deadline, the relay covers a cylindrical area of length vt_r^t and breadth r_b plus a semicircular area at the end of radius r_b minus a similar area[3] in the beginning, if it travels at velocity v and hence the first equality in the above set of equations. It fails to transfer, if no BS is found in the above described area.

IV. Closed-loop (MDP) Policies

If one can estimate/track some important status information (e.g., the time remaining till deadline, the power remaining at the decision time slot, chances of successful message delivery because of previous events, etc.) related to the system, one can design a good beaconing policy. The successful message delivery can, for example, depend upon the number of successful relay contacts till that period and the chances of these successful relays reaching the BS within the remaining deadline etc. In case we were already successful in contacting sufficiently large number of relays, it may be optimal to preserve power. On the other hand, if the user had only succeeded in contacting few/zero relays it might be optimal to beacon few more times. It might be optimal to beacon with different power levels in different situations. The aim is to derive optimal policies, which depend upon an appropriate state of the system and we use the framework of Markov Decision Processes (MDPs) for this purpose. We consider a T- time horizon problem, i.e., the deadline is divided into T time slots, each of duration D. The precise details of MDP formulation are as follows:

State: Let P_{f_t} represent the probability of failure of message delivery within deadline, when all the relays contacted during the first $t-1$ time slots fail. If a relay is contacted in the time slot t then similar probability for the next time slot decreases by a factor γ_t, i.e., $P_{f_{t+1}} = \gamma_t P_{f_t}$, else $P_{f_{t+1}} = P_{f_t}$. To summarize, P_{f_t} is the anticipated probability of failure because of all the events up to time t. The mobile would know the relays contacted in the previous time slots and P_{f_t} is obtained using these events.

Action: Recall a relay is detected with probability $1 - e^{-ca^{2/\beta}}$ (see (3)), if power a is used for beaconing at the beginning of t-th time slot. This influences further evaluation and hence is the variable to be controlled. We consider a compact action space, $\mathcal{A} \subset \{0\} \cup [\theta, \phi]$ for some $\phi < \infty$.

Transition Probabilities: Upon beaconing in time slot t, i.e., if the action were to beacon at non-zero power, there is a possibility to contact a relay (which happens with probability $1 - e^{-ca^{2/\beta}}$) and then $P_{f_{t+1}}$ improves. Thus the controlled transition probabilities can be summarized as below:

$$P(P_{f_{t+1}} = p'_f / P_{f_t} = p_f, A_t = a)$$
$$= \begin{cases} e^{-ca^{2/\beta}} & \text{when} \quad p'_f = p_f \\ 1 - e^{-ca^{2/\beta}} & \text{when} \quad p'_f = p_f \gamma_t \\ 0 & \text{else.} \end{cases}$$

Reward/Costs: These depend upon two factors: i) The factor proportional to beaconing powers utilized in all the $(T-1)$ time slots; and ii) The probability of successful message delivery or equivalently the eventual probability of failure (see (4)). We thus have/consider the following running cost and terminal costs:

$$r_t(P_{f_t}, A_t) = \alpha A_t^{2/\beta} \text{ and } g_T(P_{f_T}) = P_{f_T},$$

where α represents the trade-off between the two factors and $A_t \in \mathcal{A}$ is the power used in time slot t. Let $\boldsymbol{\pi} = [a_1, a_2, ...a_{T-1}]$, represent a typical deterministic policy[4], i.e., for any (t, p_f), $a_t(p_f) = a$ for some $a \in \mathcal{A}$ and equals the the action chosen if $P_{f_t} = p_f$.

Let the initial state be, $P_{f_1} = p_{f_1}$. Usually one can $p_{f_1} = 1$, but *one can also accommodate the chances of direct transfer (to BS) through this initial condition. But if in case the user manages to find a BS before deadline, it would obviously be optimal to stop further beaconing and optimal policies designed based on this stopping would in general be different and are considered in Section VII.* Thus we are interested in the following optimization ($E^{\boldsymbol{\pi}}_{p_{f_1}}$ is conditional expectation, conditioned on $P_{f_1} = p_{f_1}$ and $\boldsymbol{\pi}$):

$$u(p_{f1}) := \min_{\boldsymbol{\pi}} J(p_{f_1}, \boldsymbol{\pi}) \text{ where}$$
$$J(p_{f_1}, \boldsymbol{\pi}) = E^{\boldsymbol{\pi}}_{p_{f_1}} \left[\sum_{t=1}^{T-1} r_t(P_{f_t}, A_t) + g_T(P_{f_T}) \right]. \quad (5)$$

The well known dynamic programming (DP) equations ([9]) provide the value function and optimal control:

$$u_T(p_f) = p_f \text{ and for any } t < T-1 \quad (6)$$
$$u_t(p_f) = \inf_{a \in \mathcal{A}} h_t(p_f, x_a) \text{ where with } x_a := a^{2/\beta} \quad (7)$$
$$h_t(p_f, x) := \alpha x + e^{-cx} u_{t+1}(p_f) + (1 - e^{-cx}) u_{t+1}(p_f \gamma_t).$$

If there is an minimizer a^* in equation (7) for any (t, p_f) then it forms the optimal control ([9]), i.e., $a_t^*(p_f) = a^*$.

A. Some structural results

Below we obtain some structural properties of the optimal policies and the value function. We have the following results whose proofs are in Appendix A.

Lemma 1: **(Monotonicity)** The value function is monotone in state and time, i.e., for any $p_f \geq p'_f$ and any $t \leq T$

$$u_{t-1}(p_f) \leq u_t(p_f) \text{ and } u_t(p_f) \geq u_t(p'_f). \qquad \blacksquare$$

The above property is used to derive the following important result and other results.

[3]The user contacts a relay only because the later has not detected a BS and hence the first circular area is not considered.

[4]These type of policies are sufficient as explained in [9].

Lemma 2: **(Switch off Threshold)** If for some (t, p_f), the optimal control $a_t^*(p_f) = 0$ then $a_{t+1}^*(p_f) = 0$. ■

Thus by Lemma 2 there exists a switch off threshold (t_{off}) for any state p_f in the following sense:

$$\begin{array}{llll} a_t^*(p_f) & = 0 & \text{for any} & t \geq t_{off}(p_f) \text{ and} \\ a_t^*(p_f) & \geq \theta & \text{for} & t < t_{off}(p_f), \text{ where,} \\ t_{off}(p_f) & := & \min\{t : a_t^*(p_f) = 0\}. & \end{array}$$

Basically given a state p_f there exists a time slot number t_{off}, beyond which it is optimal not to beacon.

We now derive optimal policies and towards this we consider two sub-cases. We assume that continuous power control is possible in the first sub-case considered in the next sub-section. i.e., we assume action space $\mathcal{A} = 0 \cup [\theta, \phi]$, where ϕ is the maximum power that can be used. Later we consider ON-OFF control where $\mathcal{A} = \{0, \phi\}$.

B. Continuous control and approximate solution

When θ is close to 0 one can approximate (domain of optimization) the DP equations (6)-(7) and rewrite them as:

$$u_t(p_f) \approx \min_{0 \leq a \leq \phi} h_t(p_f, x_a)$$
$$= \min_{0 \leq a \leq \phi} \left\{ \alpha x + e^{-cx} \left(u_{t+1}(p_f) - u_{t+1}(p_f \gamma_t) \right) + u_{t+1}(p_f \gamma_t) \right\}.$$

If domain of optimization were not bounded, by equating the first derivative to zero and showing that the second derivative is negative (by Lemma 1, $u_{t+1}(p_f) - u_{t+1}(p_f \gamma_t) \geq 0$) the optimal x_t^* (or equivalently a_t^*) would have been (recursively):

$$e^{-c\bar{x}_t^*(p_f)} = \frac{\alpha}{c\left(u_{t+1}(p_f) - u_{t+1}(p_f\gamma_t)\right)} \text{ or}$$
$$\bar{x}_t^*(p_f) = \frac{1}{c} \log\left(\frac{c\left(u_{t+1}(p_f) - u_{t+1}(p_f\gamma_t)\right)}{\alpha} \right). \quad (8)$$

However when one considers optimization over $a \in [0, \phi]$ by convexity of the problem, the optimizer and the value function would be:

$$x_t^*(p_f) = \min\left\{ \phi^{\beta/2}, \max\left\{0, \bar{x}_t^*(p_f)\right\} \right\}, \; a_t^*(p_f) = (x_t^*(p_f))^{2/\beta},$$
$$u_t(p_f) = h_t(p_f, x_t^*(p_f)). \quad (9)$$

Using the above one can recursively compute the required solution. *The above solution is exact, if for all (t, p_f), the above $a_t^*(p_f) \in \{0\} \cup [\theta, \phi]$.*

Switch off Threshold: By Lemma 2 there exists a switch off threshold for any p_f. We compute the same under the approximation used in this sub-section. Since $u_T(p_f) = p_f$, we have after changing the domain to $\{a \geq 0\}$:

$$u_{T-1}(p_f) \approx \min_{a \geq 0} \left\{ \alpha x + e^{-cx} p_f (1 - \nu_{T-1}) + p_f \nu_{T-1} \right\}$$

If $cp_f(1-\gamma_{T-1}) < \alpha$, then $x_{T-1}^*(p_f) = 0$ and then $u_{T-1}^*(p_f) = p_f$. And now if $cp_f(1 - \gamma_{T-2}) < \alpha$ then $x_{T-2}^*(p_f) = 0$ and then $u_{T-2}^*(p_f) = p_f$. One can continue to obtain switch off threshold (Lemma 2):

$$t_{off}(p_f) = \max\{t \leq T-1 : cp_f(1-\gamma_t) > \alpha\},$$

and for all $t > t_{off}$ including $t = t_{off} + 1$ we have:

$$x_t^*(p_f) = 0 \text{ and then } u_t^*(p_f) = p_f.$$

C. Exact optimal policy with ON-OFF control

We now specialize to a case with $\mathcal{A} = \{0, \phi\}$ where $\phi \geq \theta$. For this case we have exact characterization of the optimal policy and the same is provided below with proof in Appendix:

Theorem 1: For any $t \leq T - 1$

$$a_t^*(p_f) = \begin{cases} \phi & \text{if } p_f > \psi_t \\ 0 & \text{else} \end{cases} \quad \text{and}$$
$$u_t^*(p_f) = p_f \text{ if } p_f \leq \psi_t.$$

where $\{\psi_t\}_{t \leq T-1}$ are increasing constants as given below:

$$\psi_t := \frac{\alpha \phi^{2/\beta}}{\bar{\gamma}_t}, \text{ with } \bar{\gamma}_t = (1 - e^{-cx_\phi})(1 - \gamma_t). \quad ■ \quad (10)$$

Numerical Comparison

One obviously expects that continuous control (approximate solution (9)) can provide superior performance in comparison with ON-OFF control (Theorem 1), as the domain of optimization is smaller in the later case. One might still consider ON-OFF control, as it would be a more practically viable solution. However we observe, in Tables I-II, that the loss by considering ON-OFF control is negligible in all practically viable test cases. In this sub-section we discuss only the closed-loop policies (last two columns of the tables) while others (columns) are discussed in Section VI.

Firstly the performance of closed-loop policies with continuous and ON-OFF control is almost the same when we consider examples that are practically interesting (first two and first three rows respectively in Tables I and II). Continuous action space improves significantly in comparison with ON-OFF (closed-loop), when we have lot of power (the last rows of the two tables with $\phi = 20000$ and $\phi = 5000$ respectively). This amount of power in fact implies that the user detectability radius (by relay) is much larger than the BS detectability radius. This definitely is not an interesting case.

We conducted experiments with more case studies and the observations are exactly similar for all cases. The reason for this behaviour is that, *most of the optimizers in DP equations (convex objective function over (almost) convex compact set) are at a boundary. Thus we conclude that, ON-OFF control is sufficient, and continue further with only these controls.*

V. Open-loop policies

The closed-loop policies can perform better, however it might be complicated to keep track of the status of the system. The open-loop policies are useful in such a context. Here we again discuss optimal beaconing policies, however now the power utilized in any time slot depends only upon the time slot and not on any system information. By independence of events (relay detection in different time slots, different relays successfully transferring the message to one of the BSs etc) as discussed already, the probability of failure for message delivery for any open-loop beaconing policy $\boldsymbol{\pi}$ from (4) equals:

$$P_f = \prod_{t=1}^{T} \left[P(U_{noR}^t) + (1 - P(U_{noR}^t)) P\left(R_{noBS}^t\right) \right]$$
$$= \prod_{t=1}^{T} \left[e^{-cP_t{}^{2/\beta}} + (1 - e^{-cP_t{}^{2/\beta}}) \gamma_t \right], \quad (11)$$

$\frac{\phi}{1000}$	Open Loop ON-OFF Soft			Open Loop ON-OFF Hard			Closed Loop ON-OFF			Closed Loop Continuous		
	P_{f_T}	$E[P]$	Total cost	P_{f_T}	$E[P]$	Total cost	P_{f_T}	$E[P]$	Total cost	P_{f_T}	$E[P]$	Total cost
0.1	0.48	208	0.69	0.53	167	0.70	0.50	167	0.66	0.50	166	0.66
0.5	0.29	313	0.60	0.36	252	0.61	0.27	252	0.52	0.27	251	0.52
20	0.38	287	0.67	0.33	348	0.68	0.26	348	0.60	0.21	273	0.48

TABLE I

$\theta = 1e^{-04}$, $\lambda_r = 2e^{-6}$, $\lambda_b = 2e^{-7}$, $d_0 = 10$, $\alpha = .001$, $\beta = 3.5$, $\overline{V} = 35$, $\underline{V} = 2$, $r_b = 1000$, D (TIME SLOT DURATION) = 50, T = 20

$\frac{\phi}{1000}$	Open Loop ON-OFF Soft			Open Loop ON-OFF Hard			Closed Loop ON-OFF			Closed Loop Continuous		
	P_{f_T}	$E[P]$	Total cost	P_{f_T}	$E[P]$	Total cost	P_{f_T}	$E[P]$	Total cost	P_{f_T}	$E[P]$	Total cost
0.1	0.50	125	0.75	0.53	110	0.75	0.50	110	0.72	0.50	109	0.72
0.2	0.43	145	0.72	0.45	135	0.72	0.41	135	0.68	0.41	134	0.68
0.5	0.36	174	0.70	0.37	167	0.70	0.31	167	0.65	0.33	156	0.64
5	0.53	130	0.79	0.47	164	0.80	0.41	164	0.74	0.31	164	0.63

TABLE II

$\theta = 1.0e^{-04}$, $\lambda_r = 2e^{-06}$, $\lambda_b = 2e^{-07}$, $d_0 = 10$, $\alpha = 0.002$, $\beta = 3.5$, $\overline{V} = 35$, $\underline{V} = 2$, $r_b = 1000$, $r_b = 1000$, D = 20 AND T = 15

As already discussed, one needs to consider an optimal trade-off between the power utilized for beaconing and the message successful delivery chances, while designing the policies. We consider two optimization problems: 1) with a hard constraint on a term proportional to the powers spent in various time slots; 2) which optimizes a joint cost of the two factors. Using (11) we define the two problems as below:

Problem 1: $\min_{\boldsymbol{\pi}} \prod_{t=1}^{T-1} \left(e^{-cP_t{}^{2/\beta}} + (1 - e^{-cP_t{}^{2/\beta}})\gamma_t \right),$
s.t. $\sum_{t=1}^{T-1} P_t^{2/\beta} \leq B$ (a hard power constraint)

Problem 2: $\min_{\boldsymbol{\pi}} \prod_{t=1}^{T-1} \left(e^{-cP_t{}^{2/\beta}} + (1 - e^{-cP_t{}^{2/\beta}})\gamma_t \right)$
$+\alpha \sum_{t=1}^{T-1} P_t^{2/\beta}$

We would like to compare the open-loop policies with closed-loop policies. Towards this we consider the hard constraint problem, **Problem 1**: one can set the bound B as the power utilized by the optimal closed-loop policy, and then compare the failure probability performance of open and closed-loop policies when both use the same power.

ON-OFF control: As concluded in previous section, we consider only ON-OFF control, i.e., one can either beacon using power $\phi > \theta$ or remain silent, i.e., $P_t \in \{0, \phi\}$ for each t. It is well known that randomized policies are optimal for constrained problems (**Problem 1**). Thus we redefine policy $\boldsymbol{\pi}$ as the sequence of probabilities $\boldsymbol{\pi} = [p_1, \cdots, p_{T-1}]$, one for each time slot, where p_t is the probability with which we beacon (with non-zero power ϕ) in time slot t. Define $q := (1 - e^{-c\phi^{2/\beta}})$ to rewrite the problems as:

Problem 1: $\min_{\boldsymbol{\pi}} \prod_{t=1}^{T-1} \Big((1 - qp_t) + p_t q \gamma_t \Big),$
s.t. $\sum_{t=1}^{T-1} p_t \phi^{2/\beta} \leq B$ (a hard power constraint)

Problem 2: $\min_{\boldsymbol{\pi}} \prod_{t=1}^{T-1} \Big((1 - qp_t) + p_t q \gamma_t \Big)$
$+\alpha \sum_{t=1}^{T-1} p_t \phi^{2/\beta}$

Solution for Problem 2: By monotonicity (γ_t increases with t) it is easy to see that equivalently we need to find an integer n^* and a (randomized/probability) $p^* \in [0,1]$ which minimizes

$$\prod_{t=1}^{n-1} \Big(1 - q(1 - \gamma_t) \Big) (1 - qp(1 - \gamma_n)) + \alpha(n + p)\phi^{2/\beta}$$
$$= \prod_{t=1}^{n-1} \Big(1 - q(1 - \gamma_t) \Big) + \alpha n \phi^{2/\beta} + p\xi_n \text{ with}$$
$$\xi_n := \alpha\phi^{2/\beta} - q(1 - \gamma_n) \prod_{t=1}^{n-1} \Big(1 - q(1 - \gamma_t) \Big).$$

The above is a piecewise linear function with respect to p, with slopes ξ_n increasing with n. Thus clearly $p^* = 0$ and

$$n^* = \Big((T-1) 1_{\{\xi_{T-1} < 0\}} + \min\{n : \xi_n > 0\} \, 1_{\{\xi_{T-1} > 0\}} \Big) 1_{\{\xi_1 > 0\}}, \quad (12)$$

the first time slope becomes positive ((if at all) and there exists one such n^* only if the slope were positive in the end. If all the slopes are positive it is optimal to never beacon (i.e., $n^* = 0$).

Solution for Problem 1: One can find this solution in a similar way and it is easy to see that:

$$n^* = \left\lfloor \frac{B}{\phi} \right\rfloor \text{ and } p^* = \frac{B - n^*\phi}{\phi}. \quad (13)$$

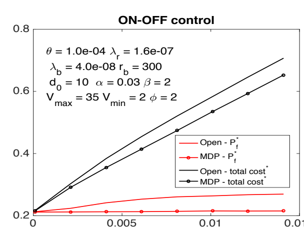

Fig. 3. CL versus OL Hard constraint policies: Optimal P_f versus α

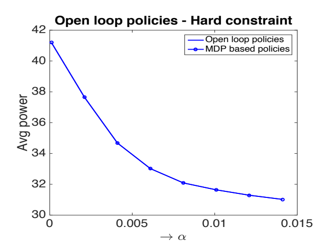

Fig. 4. CL versus OL Hard constraint policies: Optimal Power versus α

VI. NUMERICAL RESULTS

We first continue with examples of Tables I-II. The second major column reports the results under open-loop policy with soft constraint, i.e., Problem 2, with same α as that used in closed-loop policies (solution (12)). We are setting the bound in the hard open-loop problem equal to the optimal power utilized by the closed-loop (ON-OFF) policy to compare open and closed-loop policies and the results are in third major column (solution (13)). We notice that in all cases, without fail, closed-loop policies outperform the open-loop policies. There is significant improvement in some cases (especially the last rows, i.e., ones with large ϕ or power). Interestingly the total cost, $E[P_{f_T} + \alpha \sum_i P_t^{2/\beta}]$, under hard as well as the soft constraint problem is almost the same (when total cost is computed for hard problem also using the same α) for all examples. However the hard problem is better w.r.t. the power utilized while the soft problem provides better failure probability, $P^*_{f_T}$ in most cases.

We also notice that message delivery chances improve significantly with epidemic aided protocol. From the tables, it is seen that $P^*_{f_T}$ can be as low as 0.27.

We consider another set of examples in Figures 3 - 4 and make almost similar observations as in the Tables. We further notice that open-loop policies perform almost as good as the closed-loop polices for small α. We consider only hard constraint problem here. Open-loop soft constraint problems are considered in Figures (5)-(6), are compared with closed-loop policies and the observations are exactly similar.

In Figures (5)-(6), we further consider two different values of T (the number of beacons). The blue curves in both the figures correspond to $T = 22$ case while the black ones correspond to $T = 11$ case, with the duration of the time slots in the former case equal to half that in the latter case. Thus the total time duration is maintained the same, but the number of beaconing instances is varied. We notice that more number of beaconing choices improves the performance.

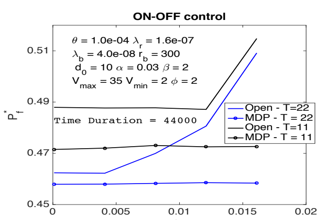

Fig. 5. CL versus OL (Soft constraint) policies: Optimal P_f vs α

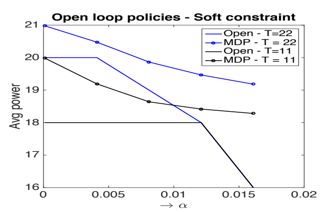

Fig. 6. CL versus OL (Soft constraint) policies: Optimal Power versus α

One can observe one more important aspect from these figures. When one utilizes the same power for both the T, the bigger T provides better P^*_f:

a) from Figure 5, with closed-loop policies, the failure probability $P^*_f \approx 0.455$ for $T = 22$ and $P^*_f \approx 0.475$ for $T = 11$ for almost all α;

b) from Figure 6 the power consumed with $\alpha = 0.0001$ and $T = 11$ is almost the same as the power consumed with $\alpha = 0.0081$ and $T = 22$; and

c) one can make similar observations with open-loop policies. However we should recall here that one can't increase the beaconing instances to a large extent as then the independence (independence environments at various beaconing instances) assumption is lost and then 'diversity' like benefits are lost.

We recall from all the examples that the performance improvement by CL policies is not very significant for small value of trade-off factor α. It is obviously much simpler to implement an open-loop policy. One just needs to remember (n^*, p^*) provided by solution (12) or (13). Thus we have another important suggestion: *one can use much simpler open-loop and ON-OFF policies when power constraint is high.*

VII. When a BS can be found directly

We now consider a brief discussion on the scenario in which the user can directly detect a BS in every time slot. Say this happens with probability ρ in each time slot independent of other time slots and the other events. This is like an optimal stopping finite horizon problem. If the user detects a BS itself, it would obviously stop searching for relays using beaconing. To consider this aspect, we modify the state process to include an extra state named Δ, to indicate absorption. That is, the state at time t denoted by Y_t equals: i) $p_f \in [0, 1]$, the probability of failure in spite of all relay contacts till time $t-1$, if the user has not yet contacted a BS directly; or equals ii) Δ if the user has already contacted a BS. Thus the DP equations change as below:

$$
\begin{aligned}
u_T(y) &= y 1_{\{y \neq \Delta\}} \text{ and for any } t < T \\
u_t(\Delta) &= 0 \text{ and with } x_a := a^{2/\beta} \\
u_t(y) &= \inf_{a \in \mathcal{A}} h_t^{bs}(y, x_a) \text{ for any } y \neq \Delta, \quad \text{where} \\
h_t^{bs}(y, x) &:= \rho * 0 \\
&\quad + (1-\rho)\left(\alpha x + e^{-cx} u_{t+1}(y) + (1 - e^{-cx}) u_{t+1}(y\gamma_t)\right).
\end{aligned}
$$

ON-OFF Control: With ON-OFF control, we have exactly the same optimal policy as in Theorem 1 with the only change being in the thresholds. The new thresholds are (additionally $(1-\rho)^{T-t}$ is included in the denominator):

$$
\psi_t^{bs} = \frac{\alpha x_\phi}{(1-\rho)^{T-t}(1 - e^{-cx_\phi})(1 - \gamma_t)}.
$$

The above expressions are correct as long as ψ_t^{bs} are monotone. Other wise, we need to make some corrections, this and further analysis is a part of future work.

Conclusions

We discussed the idea of augmenting cellular networks with epidemic inspired transfer of (delay tolerant) messages via fellow users in regions of sparse BS density. In previous works, authors consider such epidemic aided transfer, either when a message has to be transmitted from a single source to a single destination (peer to peer) or when a message has to be broadcast from a single source to multiple destinations. The proposed epidemic aided cellular network requires transfer between a single source and one of the base stations, which are distributed according to a (stationary) Poisson Point process.

We obtained beaconing policies for delay tolerant information transfer. We investigated policies that maximize the probability of successful message delivery, while accounting for the power utilized for beaconing purposes. We formulated the problem as a Markov Decision Process to derive optimal policies that depend on system state (closed-loop). We showed that the value function satisfies certain monotonicity properties, and the optimal policy exhibits a certain switch off property. We obtained approximate solution for continuous control and an exact one for ON-OFF (two action) control.

Our investigations show that the closed-loop ON-OFF policies perform on par with the closed-loop continuous policies, for all practically viable test cases. Thus it is sufficient to use ON-OFF policies. We further investigated open-loop policies for ON-OFF control, the policies that can be used when the system state is not known. Detailed numerical analysis and simulations are carried out comparing various policies. The numerical analysis further demonstrated that for many cases the open-loop policies perform as good as the closed-loop policies (this is mostly true when the power constraint is high). However, in many other scenarios there is a good improvement with the closed-loop policies.

We have closed-form expressions for ON-OFF policies (and these are shown to be sufficient). Thus one can obtain optimal policies for any given configuration and then can chose an optimal configuration. For example, one can chose optimal number of beacons or optimal power utilized.

Towards the end, we briefly discussed a scenario in which the user stops beaconing immediately after it (directly) finds a base station. The optimal (ON-OFF) policy for this case has similar structure as before. We have closed-form expressions for the optimal policy under some assumptions.

REFERENCES

[1] Rebecchi, Filippo, et al. "Data offloading techniques in cellular networks: A survey." IEEE Communications Surveys & Tutorials 17.2 (2015): 580-603.

[2] Baron, Benjamin, et al. "Mobility as an Alternative Communication Channel: A Survey." arXiv preprint arXiv:1802.04587 (2018).

[3] Study on further enhancements to LTE Device to Device (D2D), User Equipment (UE) to Network Relays for Internet of Things (IoT) and Wearables, 3GPP TR 36.746, V15.0.0 (2017).

[4] A Ali, M Panda, T Chahed, E Altman, "Improving the transport performance in delay tolerant networks by random linear network coding and global acknowledgements", Ad Hoc Networks, 2013.

[5] E. Altman, L. Sassatelli and F. De Pellegrini, "Dynamic Control of Coding for Progressive Packet Arrivals in DTNs", IEEE Trans. on Wireless Comm., 13(10), 2014.

[6] C. Singh, E. Altman, A. Kumar, R. Sundaresan, "Optimal Forwarding in Delay-Tolerant Networks With Multiple Destinations", IEEE Trans. on Wireless Comm., 12(2), 2013.

[7] A. Ali, T. Chahed, E. Altman, M. Panda and L. Sassatelli, "A new proposal for reliable unicast and multicast transport in Delay Tolerant Networks", IEEE/ACM Trans. on Networking, 21(2), 2013.

[8] A. Kumar, V. Kavitha, and N. Hemachandra, "Power Constrained DTNs: Risk MDP-LP Approach," WiOpt 2015, pp 154-160,

[9] M. L. Puterman, "Markov decision processes: discrete stochastic dynamic programming". John Wiley & Sons, 2014.

APPENDIX A: MDP RELATED PROOFS

Proof of Lemma 1: From the DP equation (7) note that $h_t(p_f, 0) = u_{t+1}(p_f)$ and hence that monotonicity in time is obvious. We prove the monotonicity in state through backward induction. For any $p_f > p'_f$, clearly,

$$u_T(p_f) = p_f > p'_f = u_T(p'_f).$$

Assume the result is true for any $t+1 \le T$, i.e., assume:

$$u_{t+1}(p_f) \ge u_{t+1}(p'_f) \text{ whenever } p_f \ge p'_f.$$

Then for $p_f \ge p'_f$ and with $x := a^{2/\beta}$,

$$\begin{aligned} u_t(p_f) &= \min_a \left\{\alpha x + e^{-cx} u_{t+1}(p_f) + (1-e^{-cx}) u_{t+1}(p_f \gamma_t)\right\} \\ &\ge \min_a \left\{\alpha x + e^{-cx} u_{t+1}(p'_f) + (1-e^{-cx}) u_{t+1}(p'_f \gamma_t)\right\} \\ &= u_t(p'_f). \ \blacksquare \end{aligned}$$

Proof of Lemma 2: Fix a p_f and define $h_t(x) := h_t(p_f, x)$, for any $t < T$, for shorter notation. By hypothesis $u_t(p_f) = h_t(0) = u_{t+1}(p_f)$ and hence for any $a \ge \theta$,

$$u_{t+1}(p_f) < h_t(a^{2/\beta}). \tag{14}$$

Now for any $a \ge \theta$ with $x := a^{2/\beta}$ and using monotone properties of Lemma 1:

$$\begin{aligned} h_t(x) &= \alpha x + e^{-cx} u_{t+1}(p_f) + (1-e^{-cx}) u_{t+1}(p_f \gamma_t) \\ &\le \alpha x + e^{-cx} u_{t+2}(p_f) + (1-e^{-cx}) u_{t+2}(p_f \gamma_t) \\ &\le \alpha x + e^{-cx} u_{t+2}(p_f) + (1-e^{-cx}) u_{t+2}(p_f \gamma_{t+1}) \\ &= h_{t+1}(x). \end{aligned}$$

Further using (14), for any $a \ge \theta$,

$$h_{t+1}(a^{2/\beta}) \ge h_t(a^{2/\beta}) > u_{t+1}(p_f),$$

which in turn implies that $a^*_{t+1}(p_f) = 0$. $\blacksquare$

Proof of Theorem 1: It is easy to compute the optimal control at $t = T-1$ as below. From the DP equations (6)-(7):

$$\begin{aligned} u_{T-1}(p_f) &= \min\{p_f, \ \ h_{T-1}(p_f, x_\phi)\} \\ &= p_f + \min\left\{0, \ \ \alpha x_\phi - p_f \bar{\gamma}_{T-1}\right\}, \end{aligned}$$

where $\bar{\gamma}_{T-k} := (1-e^{-cx_\phi})(1-\gamma_{T-k})$ for any k. Hence

$$a^*_{T-1}(p_f) = \phi 1_{\{p_f > \psi_{T-1}\}}, \ \ \psi_{T-1} := \frac{\alpha x_\phi}{\bar{\gamma}_{T-1}} \text{ and}$$

$$u_{T-1}(p_f) = \begin{cases} \alpha x_\phi + p_f(1-\bar{\gamma}_{T-1}) & \text{if } p_f > \psi_{T-1} \\ p_f & \text{else.} \end{cases}$$

When $t = T-2$ and if $p_f < \psi_{T-1}$, we have $u_{T-1}(p_f) = p_f$ as well as $u_{T-1}(p_f\gamma_{T-2}) = p_f\gamma_{T-2}$. Hence again from DP equations (7) for any $p_f < \Psi_{T-1}$:

$$u_{T-2}(p_f) = p_f + \min\left\{0, \ \ \alpha x_\phi - p_f \bar{\gamma}_{T-2}\right\} \text{ and so}$$

$$a^*_{T-2}(p_f) = \phi 1_{\{p_f < \psi_{T-2}\}} \text{ with } \psi_{T-2} := \frac{\alpha x_\phi}{\bar{\gamma}_{T-2}}.$$

Note here that $\gamma_{T-1} > \gamma_{T-2}$ and hence that $\psi_{T-1} > \psi_{T-2}$. In other words

$$a^*_{T-2}(p_f) = \begin{cases} \phi & \text{if } \psi_{T-2} < p_f < \psi_{T-1} \\ 0 & \text{if } p_f < \psi_{T-2} \end{cases} \text{ and,}$$

$$u^*_{T-2}(p_f) = \begin{cases} \alpha x_\phi + p_f(1-\bar{\gamma}_{T-2}) & \text{if } \psi_{T-2} < p_f < \psi_{T-1} \\ p_f & \text{if } p_f < \psi_{T-2}. \end{cases}$$

For any $p_f > \psi_{T-1}$ we have $a^*_{T-1}(p_f) > 0$, and, hence by Lemma 2 and because of two controls, we have for all such p_f: $a^*_{T-2}(p_f) > 0$ and so $a^*_{T-2}(p_f) = \phi$.

In all, $a^*_{T-2}(p_f) = \begin{cases} \phi & \text{if } p_f > \psi_{T-2} \\ 0 & \text{else} \end{cases}$ and

$$u^*_{T-2}(p_f) = p_f \text{ if } p_f \le \psi_{T-2}.$$

By backward induction (using similar logic) and first considering

$$p_f < \psi_{t+1} := \frac{\alpha x_\phi}{\bar{\gamma}_{t+1}}, \ \ \text{with } \bar{\gamma}_t = (1-e^{-cx_\phi})(1-\gamma_t) \tag{15}$$

and then using Lemma 2 one can complete the proof. $\blacksquare$