Reducing Call Routing Cost for Femtocells

Yi-Bing Lin, *Fellow, IEEE,* Chai-Hien Gan, *Member, IEEE,* and Ching-Feng Liang

Abstract—Femtocell is an effective solution to improve indoor coverage for cellular networks, where short-range and low-power Base Stations (BSs) called Femto BSs are deployed in small areas (e.g., buildings). We propose the Femto Private Branch Exchange (FPBX) to concentrate the traffic of the Femto BSs in a specific area, such as a campus or an enterprise with central or distributed locations. Through the FPBX, a normal cellular call between two Femto users can be replaced by a low-cost extension call. Therefore, the call routing cost can be significantly reduced. In this approach, the existing mobile network nodes are not modified. We analytically analyze the call routing performance for the FPBX approach, and develop simulation experiments to validate against the analytic model. Our study indicates that the FPBX approach can effectively reduces the call setup costs and the voice trunk costs among Femto users by slightly increasing the normal cellular call setup costs.

Index Terms—Call setup, femtocell, femto base station (femto BS), femto private branch exchange (FPBX).

I. INTRODUCTION

CELLULAR networks have been widely deployed to provide telephone services for mobile users. A cellular network such as Universal Mobile Telecommunications System (UMTS) consists of the core network (Figure 1 (1)) and the Radio Access Network (RAN; Figure 1 (2)) [1], [2]. The core network includes Mobile Switching Centers (MSCs; Figure 1 (3)) and mobility databases (Figure 1 (4)). An MSC is responsible for routing the cellular calls, and the mobility database maintains the location information for mobile users. The RAN consists of Radio Network Controllers (RNCs; Figure 1 (5)) and Base Stations (BSs; Figure 1 (6)). An RNC is connected to an MSC in the core network. A cellular customer uses a Mobile Station (MS; Figure 1 (7)) to communicate with a BS. In this architecture, a cellular call from MS1 to MS2 is set up through path BS1 - RNC - Core - RNC - BS2. Routing in the core network involves the MSCs and the mobility databases, and the details can be found in [2].
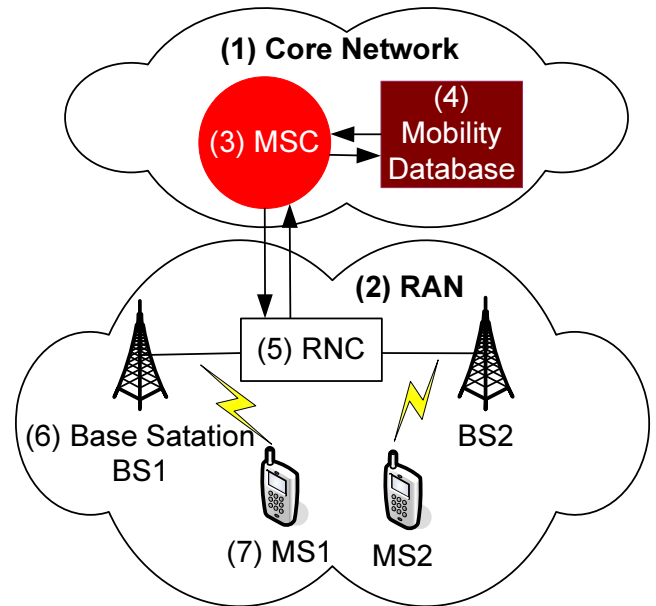
Fig. 1. Cellular network architecture.

Several studies [3]–[5] indicate that more than 50 percent of calls occur indoors, and the operators need to provide better indoor coverage. However, to deploy more BSs for indoor coverage may significantly increase the cost of the mobile network infrastructure. To resolve this issue, the Femtocell technology has been considered as an alternative solution to improve the indoor coverage with low-cost [4], [6]. Femto BSs (or Home Node Bs) are short-range, low-cost, and low-power wireless access points that operate in the same licensed spectrum as the standard cellular BSs (called Macro BSs in this paper). The Femto BSs extend the Macro BSs' coverage for a limited number of authorized users, which connect standard MSs to a mobile operator's network using residential broadband access [7], such as Digital Subscriber Line (DSL) or cable broadband connections. A Femto BS typically serves 4-8 MSs simultaneously.

In standard 3GPP Femtocell architecture, all mobile calls are handled by the MSC in the core network, where the Femto GW acts as a relay to deliver the traffics between the core network and Femto BSs. In other words, call setup for a Femto user (i.e., the MS covered by a Femto BS) is exact the same as a Macro user (the MS covered by a Macro BS) [2]. Even though two call parties reside in the same campus or enterprise area, the call signaling and voice traffic are always routed to the core network, which incurs extra telephony cost. Therefore, in a campus or an enterprise, Internet Protocol Private Branch Exchange (IP PBX) [8] is usually deployed to provide extension calls (i.e., the calls without involving the
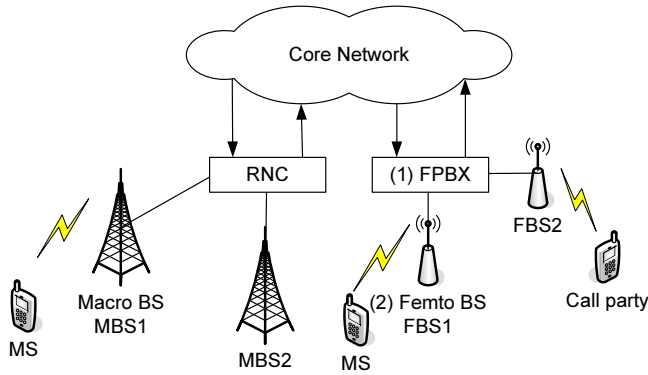
Fig. 2. Femto private branch exchange architecture.

Fig. 3. Call routing paths when location flag is correct (both voice and signaling take the same paths).

core network) by using landline phones to avoid the telephony cost. The problem of directly combining the 3GPP Femto GW and the standard IP PBX solution is that both mobile phone number and landline phone number are required for a user A. When user B in the same campus attempts to make an extension call to user A, B must dials A's landline extension number. If B calls A's mobile phone number through IP PBX, the call always goes to the core network (even if A is in the campus).

To address the above issue, we propose a Femto Private Branch Exchange (FPBX) architecture, which not only combines the Femto GW and the IP PBX, but also implements the call control and mobility management modules to handle call routing and location management of Femto users. With our approach, a Femto user can use the standard mobile phone (e.g., 3G handset) to make and receive extension calls, which does not involve extra landline phone number or the core network routing. Our solution is compatible with the existing mobile telecommunications systems and can be inserted inside them. In this paper, we describe the call setup procedure for the FPBX, and conduct analytic analysis and develop simulation to evaluate the performance of the FPBX.

II. FEMTO PRIVATE BRANCH EXCHANGE ARCHITECTURE

Figure 2 depicts the Femto Private Branch Exchange (FPBX) architecture. In this architecture, the Core, the RNC, the MBSs, and the MSs are standard equipments, which need not to be modified. Since Femto BSs are small base stations and many of them will be installed in a specific area (e.g., a campus or an enterprise), it is more efficient to first concentrate the traffic of these Femto BSs to a gateway. In our solution, the FPBX (Figure 2 (1)) is a "transparent plug-in" gateway that concentrates the traffic of Femto BSs (Figure 2 (2)) installed in a specific service area (e.g., an enterprise which is not limited to a geographic area), and acts as a relay between these Femto BSs and the mobile core network. Femto BSs are connected to the FPBX through an IP network (e.g., managed enterprise intranet). For the discussion purpose, we call an MS a Femto user when it is covered by a Femto BS. The MS is called a Macro user when it moves from the Femto BS coverage to the Macro BS coverage.

In this paper, we define an extension call as the call between two Femto users served by Femto BSs that connect to the FPBX. Besides the traffic concentration, a major advantage
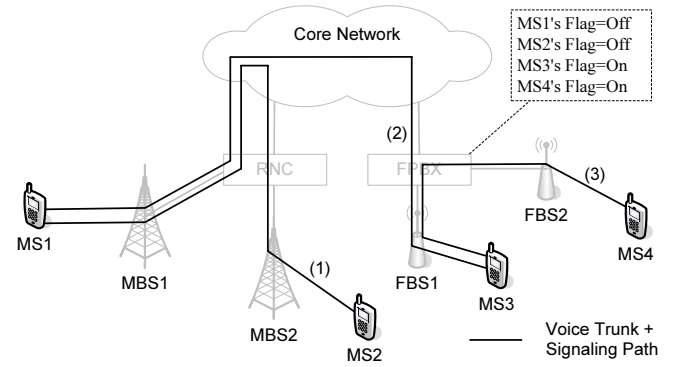
of the FPBX (which is the purpose of this paper) is that it can support the extension calls, where the function is similar to that of a Private Branch Exchange (PBX) installed in an enterprise. For a normal cellular call, the setup message and voice traffic are passed through several core network nodes, such as Mobile Switching Center (MSC), Gateway MSC (GMSC), Visiting Location Register (VLR), and Home Location Register (HLR) [2]. These nodes are carrier-grade equipments, and it is clear that such routing costs are very expensive. By using the FPBX solution, these costs can be avoided for the extension calls.

The FPBX interacts with a Femto BS through the 3GPP Iuh protocol [6], where the FPBX plays the role as the Femto Gateway (or Home Node B Gateway in 3GPP) [6]. Similar, the FPBX interacts with the Femto Gateway through the Iuh protocol, where the FPBX plays the role as a Femto BS. In the higher level procedures such as Mobility Management (MM) and Call Control (CC), the FPBX acts as a VLR to record the MS's location, and acts as an MSC to switch the call setup for an extension call. The reader is referrd to [2] for the details of VLR and MSC operations. Same protocols are used to implement the FPBX in the 3G CS domain. In 3G PS domain or 4G, the calls are most likely carried out by SIP which can be easily used to implement the FPBX, too. In our FPBX deployment, we have modified the Hughes Telecommunications 3G protocol stacks to implement the aforementioned procedures.

To support the extension call service, the FPBX maintains a location flag for each of the MSs who are members of this extension call application, and routes the extension call between two Femto users. When an MS's location flag is "ON", the FPBX considers the MS as a Femto user. When the flag is "OFF", the MS is considered as a Macro user. This flag is updated during call activities of the MS as described below. Figures 3 and 4 depict the call routing between an MS and the other call party. Without loss of generality, we investigate the incoming and outgoing call setup costs of the MS. Specifically, we consider the signaling or call setup path (we will use these two terms alternatively) used to deliver the call setup control messages, and the voice truck path used to deliver the conversation content. The signaling path and the voice trunk path may or may not the same.

Case I (Macro/Macro). The MS is a Macro user covered

(a) A Re-routing Cellular Call Setup

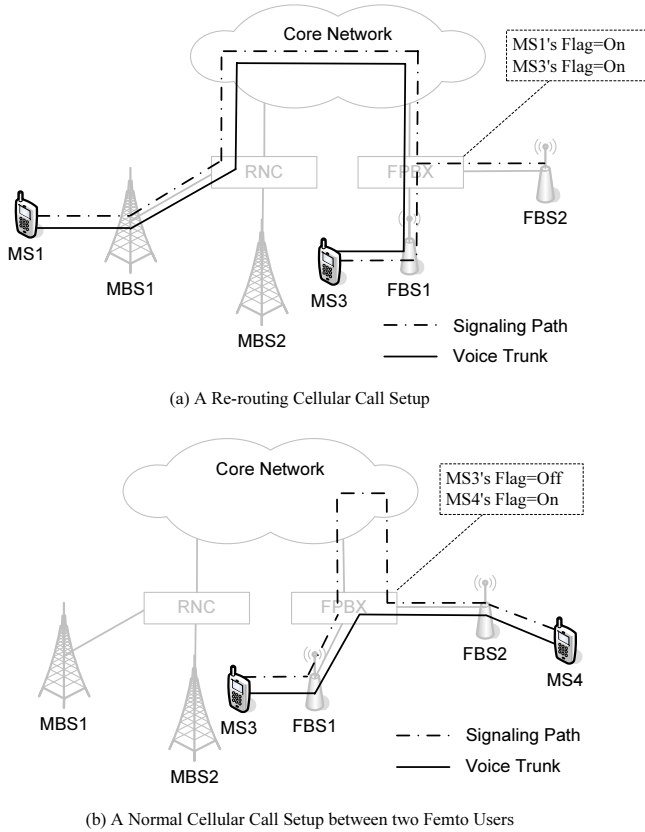(b) A Normal Cellular Call Setup between two Femto Users

Fig. 4. Call routing paths when location flag is obsolete.

by a Macro BS MBS1, and the other call party is a Macro user covered by another Macro BS MBS2. For both incoming and outgoing calls of the MS, the signaling path is MBS1 - RNC - Core - RNC - MBS2 (a normal cellular call setup; see Figure 3 (1), where the MS is MS1 and the other call party is MS2). In this case, the FPBX is not involved. The voice trunk path is the same as the signaling path. In this paper, when we say "same paths" for normal cellular calls, it means that the "core network" is visited. We note that the routing details within the core network may be different [9], which will not affect the analysis in this paper.

Case II (Femto/Macro). The MS is a Femto user covered by a Femto BS FBS1 and the other call party is a Macro user covered by a Macro BS MBS1. For both incoming and outgoing calls of the MS, the signaling path is FBS1 - FPBX - Core - RNC - MBS1 (a normal cellular call setup; see Figure 3 (2), where the MS is MS3 and the other call party is MS1). When the call is set up to the FPBX (either from FBS1 or the core network), the FPBX sets the MS's location flag to "ON". The voice trunk path is the same as the signaling path.

Case III (Macro/Femto; outgoing call). The MS is covered by a Macro BS MBS1, and the other call party is covered by a Femto BS FBS1. The MS makes an outgoing call to the call party. The signaling path is MBS1 - RNC - Core - FPBX - FBS1 (a normal

cellular call setup; see Figure 3 (2), where the MS is MS1 and the other call party is MS3). When the call is routed to the FPBX, and the MS's location flag is set to "OFF". The voice trunk path is the same as the signaling path.

Case IV (Macro/Femto; incoming call). The MS is covered by a Macro BS MBS1, and the other call party is covered by a Femto BS FBS1. The MS receives an incoming call from the call party. When the call routed from FBS1 to the FPBX, the MS's location flag is checked. There are two possibilities:

(a) If the MS's location flag is "OFF", the FPBX sets up the call to the RNC. The signaling path is FBS1 - FPBX - Core - RNC - MBS1 (a normal cellular call setup; see Figure 3 (2), where the MS is MS1 and the other call party is MS3). The voice trunk path is the same as the signaling path.

(b) If the MS's location flag is "ON" (i.e., the flag information is obsolete), the FPBX asks all Femto BSs to page the MS and fails. The MS's location flag is set to "OFF", and the call is routed to the RNC (see Figure 4 (a), where the MS is MS1 and the other call party is MS3). The signaling path is FBS1 - FPBX - all FBSs - FPBX - Core - RNC - MBS1 (a re-routed cellular call setup), and the voice trunk is FBS1 - FPBX - Core - RNC - MBS1 (a normal cellular voice trunk path).

Case V (Femto/Femto; outgoing call). The MS is covered by a Femto BS FBS1, and the other call party is covered by another Femto BS FBS2. The MS makes an outgoing call to the call party. The call is first routed to the FPBX, and the MS's location flag is set to "ON". Then the call is routed to the call party through path FBS1 - FPBX - FBS2 (an outgoing extension call setup if the location flags for the other call party is valid; see Figure 3 (3), where the MS is MS3 and the other call party is MS4). The voice trunk path is the same as the signaling path.

Case VI (Femto/Femto; incoming call). The MS is covered by a Femto BS FBS1, and the other call party is covered by another Femto BS FBS2. The MS receives an incoming call from the call party.

(a) If the MS's location flag is "ON", the FPBX asks all Femto BSs to page the MS and the call is connected. The signaling path is FBS2 - FPBX -FBS1 (an incoming extension call setup; see Figure 3 (3), where the MS is MS3 and the other call party is MS4). The voice trunk path is the same as the signaling path.

(b) If the MS's location flag is "OFF" (i.e., the flag information is obsolete), the FPBX sets up the call to the RNC. The signaling path is FBS2 - FPBX - Core - FPBX - FBS1 (a normal cellular call setup; see Figure 4

(b), where the MS is MS3 and the other call party is MS4). Note that the voice trunk can be shortened by FPBX through path FBS2 - FPBX - FBS1. When the call is routed from the Core to the FPBX, the MS's flag is set to "ON". The voice trunk path is the same as an extension Femto call.

Note that if the Macro and the Femto BSs belong to different UMTS location areas, then a standard UMTS location update is performed when the MS moves from a Macro BS coverage to a Femto BS coverage [2], [10]. In this operation, the FPBX passes the 3GPP location update message to the mobility database in the core network. At the same time, the FPBX can also turn on the MS's location flag. The situation is similar to Case V described above. Also note that handoffs may occur between two Femto BSs. Such Femto BS handoffs are conducted in the FPBX following the standard 3GPP procedure, where the FPBX serves as serving MSC. Due to page limit, we refer the reader to [2] for the details.

We note that FPBX can be developed based on the existing PBX solution (we actually implemented FPBX on ITRI's IP PBX platform [11]). For the same call traffic, the capacity requirement of FPBX is the same as that of PBX. The cost for providing the above FPBX feature is solely software efforts, which includes the Iuh protocol software [6] and the extension call routing software described above. With this small software cost, vendors can sell their PBX products with enhanced FPBX feature with much better price. Therefore, the business justification for FPBX is clear.

III. ANALYTIC MODEL

In the FPBX architecture, when the MS enters/leaves the Femto BSs, no explicit registration and de-registration are performed to update the location flag at the FPBX. Instead, the registration operations are implicitly performed at the call activities in Cases II, V and VI (b). The de-registration operations are implicitly performed at the call activities in Cases III and IV (b). This approach is called implicit location update (i.e., the location flag is implicitly update during call setup), which does not need to modify the existing UMTS nodes, and therefore allows the FPBX to be conveniently plugged in the existing cellular network. However, in Cases IV (b) and VI (b), the location flag is obsolete, which results in extra routing with higher call setup costs (i.e., re-routing in Case IV (b) or normal cellular call setup in Case VI (b)). After an extra routing, the location flag is then set to the correct status.

When both the MS and the other call party are Femto users, we hope that an incoming call to the MS is an extension call within the FPBX; i.e., Case VI (a) occurs with a low call setup cost c_F. On the other hand, when the MS is a Macro user, and the other call party is a Femto user, the incoming call to the MS should be a normal cellular call setup; i.e., Case IV (a) occurs with the cost c_M. Clearly, $c_M > c_F$. Since the location flag of the MS may not be correctly set, Case IV (b) may occur with cost $c_F + c_M$, and Case VI (b) may occur with cost c_M.

Assume that an MS stays in the Macro BS coverage for a period t_M, then it enters the Femto BS coverage for another period t_F, and then it moves to the Macro BS coverage again. Let t_M and t_F be random variables with the density functions $f_M(t_M)$ and $f_F(t_F)$, and the Laplace transforms $f_M^*(s)$ and $f_F^*(s)$, respectively. The (t_M, t_F) pair can be modeled as renewal periods [12]. Define $\theta_F(\mu, k)$ as the probability that there are k calls for the MS during a t_F period, where the call arrivals are a Poisson process with rate μ. Similarly, let $\theta_M(\mu, k)$ be the probability that there are k calls for the MS during a t_M period, where the call arrivals are a Poisson process with rate μ. Then

$$
\begin{aligned}
\theta_F(\mu, k) &= \int_{t_F=0}^{\infty} \left[\frac{(\mu t_F)^k}{k!} \right] e^{-\mu t_F} f_F(t_F) dt_F \\
&= \left[\frac{(-\mu)^k}{k!} \right] \times \left[\frac{d^k f_F^*(s)}{d^k s} \right] \bigg|_{s=\mu}
\end{aligned}
$$

and similarly,

$$
\theta_M(\mu, k) = \left[\frac{(-\mu)^k}{k!} \right] \times \left[\frac{d^k f_M^*(s)}{d^k s} \right] \bigg|_{s=\mu} \tag{1}
$$

Suppose that all incoming calls from Femto users to the MS (Cases IV and VI) are a Poisson process with rate λ, and all other call activities (including incoming and outgoing calls of the MS) in Cases II, III and V are a Poisson process with rate γ. Note that call activities in Case I do not involve the FPBX (and the MS's location flag is not updated), which will not be considered from now on.

After the MS enters the Macro BS coverage, the first call from a Femto user occurs at time t_1 before the MS moves to the Femto BS coverage. This means that the location flag is last updated at time $t_2 < t_1$, where the time point t_2 is in a t_F period with the probability $1 - \theta_F(\lambda + \gamma, 0)$; i.e., a call setup in either Cases II or V (with rate γ) or Case VI (with rate λ) turns on the location flag. Furthermore, following the first t_F period, if there are exactly i (t_M, t_F) pairs in $[t_2, t_1]$, then location flag update occurs in these i periods with probability $[\theta_M(\lambda, 0)\theta_F(\lambda + \gamma, 0)]^i$ (for $i \geq 0$), where $\theta_M(\lambda, 0)$ is the probability that when the MS is in a t_M period, no call setup occurs in Case IV (with rate λ), and the MS's location flag is not turned off.

Let α be the probability that the MS's location flag is "ON" at time t_1 (Case IV (b) occurs). From the above discussion, α is expressed as

$$
\begin{aligned}
\alpha &= \sum_{i=0}^{\infty} [1 - \theta_F(\lambda + \gamma, 0)] \times [\theta_M(\lambda, 0)\theta_F(\lambda + \gamma, 0)]^i \\
&= \frac{1 - \theta_F(\lambda + \gamma, 0)}{1 - \theta_M(\lambda, 0)\theta_F(\lambda + \gamma, 0)}
\end{aligned} \tag{2}
$$

When the MS is a Macro user and the other call party is a Femto user, let C_M be the expected cost of a call from the other call party to the MS including the re-routing overhead. Then

$$
\begin{aligned}
C_M &= (1 - \alpha)c_M + \alpha \\
&\times \left\{ \sum_{k=1}^{\infty} \left\{ \frac{(c_F + c_M) + (k-1)c_M}{k[1 - \theta_M(\lambda, 0)]} \right\} \theta_M(\lambda, k) \right\}
\end{aligned} \tag{3}
$$

In the right-hand side of (3), the first term says that if the MS's location flag is "OFF" when the MS enters the Macro

BS coverage (with probability $1-\alpha$), then every call to the MS is a normal cellular call with cost c_M. The second term says that if the MS's location flag is "ON" at the beginning of a t_M period (with probability α) the extra re-routing overhead c_F is shared by k calls from Femto users occurring in this t_M period (with probability $\theta_M(\lambda, k)$); that is, among these k calls, the first call is re-routed with cost $c_F + c_M$ (Case VI (b)), and each of the subsequent $k-1$ calls is a normal cellular call with cost c_M (Case VI (a)). This term is calculated under the condition that $k > 0$, and therefore is scaled by $1 - \theta_M(\lambda, 0)$. Equation (3) is simplified as

$$C_M = c_M + \left[\frac{\alpha c_F}{1 - \theta_M(\lambda, 0)} \right] \times \left[\sum_{k=1}^{\infty} \frac{\theta_M(\lambda, k)}{k} \right] \quad (4)$$

Clearly, the second term in (4) is the extra cost introduced by implicit de-registration.

Suppose that the MS enters the Femto BS coverage at time t_3. Let β be the probability that the MS's location flag is "ON" at time t_3. Similar to the derivation for α, β can be expressed as

$$\beta = \alpha \theta_M(\lambda, 0) \quad (5)$$

When both the MS and the other call party are Femto users, let C_F be the expected cost of a call from the call party to the MS. Then

$$
\begin{aligned}
C_F &= \beta c_F + (1 - \beta) \left(\frac{\gamma}{\lambda + \gamma} \right) c_F \\
&+ (1 - \beta) \left(\frac{\lambda}{\lambda + \gamma} \right) \\
&\times \left\{ \sum_{k=1}^{\infty} \left[\frac{c_M + (k-1)c_F}{k} \right] \theta_F(\lambda, k) \right\}
\end{aligned}
\quad (6)
$$

The terms in the right-hand side of (6) are explained as follows: The first term says that the MS's location flag is "ON" at the beginning of the t_F period (with probability β), then every call from a Femto user to the MS is an extension call with cost c_F. The second term says that the MS's location flag is "OFF" at the beginning of the t_F period (with probability $1 - \beta$), and the flag is turned "ON" (because of the call activities of Cases II and V) before the first call from a Femto user to the MS (with probability $\gamma/(\lambda + \gamma)$). Therefore, every call from a Femto user to the MS is an extension call with cost c_F. The third term says that if the MS's location flag is "OFF" when the first call from a Femto user to the MS arrives (with probability $(1 - \beta)\lambda/(\lambda + \gamma)$), then this call is a normal cellular call with cost c_M (Case VI (b)), and each of the subsequent calls from the Femto users is an extension call with cost c_F. Equation (6) is simplified to

$$C_F = c_F + \left\{ \frac{(1 - \beta)\lambda(c_M - c_F)}{(\lambda + \gamma)} \right\} \left[\sum_{k=1}^{\infty} \frac{\theta_F(\lambda, k)}{k} \right] \quad (7)$$

Suppose that we modify the software of the MS such that when it moves from a Macro BS to a Femto BS, a registration is explicitly performed to turn on the MS's location flag at the FPBX, and when the MS moves from a Femto BS to a Macro BS, a de-registration is explicitly performed to turn off the flag. Let the cost for location update be c_L. From (4) and (7),

implicit location update outperforms implicit location update if

$$c_L > \alpha c_F \quad (8)$$

and

$$c_L > \left[\frac{(1 - \beta)\lambda(c_M - c_F)}{(\lambda + \gamma)} \right] [1 - \theta_F(\lambda, 0)] \quad (9)$$

By including the cost of accessing the mobility database, we may assume that $c_L \approx c_M$. Then from (8), we have $\alpha > \frac{c_M}{c_F}$, which is always true (because $0 \leq \alpha \leq 1$ and $c_M > c_F$). From (9), we have $(1 - \beta)\left(\frac{\lambda}{\lambda + \gamma} \right)[1 - \theta_F(\lambda, 0)] < \frac{c_M}{c_M - c_F}$ which is also always true (because $0 \leq (1 - \beta) \leq 1$, $0 \leq \left(\frac{\lambda}{\lambda + \gamma} \right) \leq 1$, $0 \leq [1 - \theta_F(\lambda, 0)] \leq 1$, and $c_M > c_M - c_F$). Therefore, implicit location update always outperforms explicit location update.

IV. PERFORMANCE EVALUATION

We have developed both analytic and simulation models to evaluate the FPBX solution. Our simulation model follows the event-driven approach for mobile network studies [13], [14]. Both analysis and simulation results are consistent, and all discrepancies are within 3% (see Figures 6-9). This section discusses the FPBX performance based on the simulation experiments.

A. Input parameters and Output Measures

We summarize the input parameters used for both analytic and simulation models as follows.

- λ: the rate of the arrival calls from Femto users to the MS. The arrivals are assumed to have a Poisson distribution.
- γ: the rate of the incoming calls of non-Femto users to the MS and the outgoing calls of the MS (in Cases II, III, and V). The arrivals are assumed to have a Poisson distribution.
- t_M: the period that an MS stays in the Macro BS coverage, which has a general distribution with mean $1/\eta_M$ and variance v_M.
- t_F: the period that an MS stays in the Femto BS coverage, which has a general distribution with mean $1/\eta_F$ and variance v_F.
- c_M: the signaling cost of a normal cellular call to the MS.
- c_F: the signaling cost of an extension call to the MS.

For the illustration purpose, we use Gamma distributions for t_M and t_F in our study. It is proved that any experimental data can be fit by a mixture of Gamma distributions [15], and is often selected in mobile network studies [13], [16]. We also consider other distributions (including Weibull and Lognormal) [17] for simulation experiments. The results are similar and are not present in this paper.

If t_F is a Gamma random variable with the density function $f_F(t_F) = \left[\frac{\delta(\delta t_F)^{\phi - 1}}{\Gamma(\phi)} \right] e^{-\delta t_F}$ (where $\Gamma(\cdot)$ is the Gamma function, and ϕ is a non-negative real number), then its mean is $\frac{1}{\eta_F} = \frac{\phi}{\delta}$ and the variance is $v_F = \frac{\phi}{\delta^2} = \frac{1}{\phi \eta_F^2}$. The Laplace transform for the t_F random variable is

$$f_F^*(s) = \left(\frac{\phi \eta_F}{\phi \eta_F + s} \right)^{\phi} \quad (10)$$

Similarity, let $f_M(t_M) = \left[\frac{\delta(\delta t_M)^{\phi-1}}{\Gamma(\phi)} \right] e^{-\delta t_M}$, where the mean is $\frac{1}{\eta_M} = \frac{\phi}{\delta}$ and the variance is $v_M = \frac{1}{\phi \eta_M^2}$. We have

$$f_M^*(s) = \left(\frac{\phi \eta_M}{\phi \eta_M + s} \right)^{\phi} \qquad (11)$$

From (1), (10), and (11), we have

$$
\begin{aligned}
\theta_F(\mu, k) &= \left[\frac{(-\mu)^k}{k!} \right] \times \left[\frac{(-1)^k \Pi_{i=1}^{k-1}(\phi + i)}{\phi^{k-1} \eta_F^k} \left(\frac{\phi \eta_F}{\phi \eta_F + \mu} \right)^{\phi+k} \right] \\
&= \left[\frac{\mu^k \Pi_{i=1}^{k-1}(\phi + i)}{k! \phi^{k-1} \eta_F^k} \right] \left(\frac{\phi \eta_F}{\phi \eta_F + \mu} \right)^{\phi+k}
\end{aligned}
$$

and

$$\theta_M(\mu, k) = \left[\frac{\mu^k \Pi_{i=1}^{k-1}(\phi + i)}{k! \phi^{k-1} \eta_M^k} \right] \left(\frac{\phi \eta_M}{\phi \eta_M + \mu} \right)^{\phi+k} \qquad (12)$$

The output measurements in our study include

- α: the probability that the MS's location flag is "ON" when the MS enters the Macro coverage, which can be obtained by (2) and (12). For $k = 0$, $\theta_F(\lambda + \gamma, k) = \left(\frac{\phi \eta_F}{\phi \eta_F + \lambda + \gamma} \right)^{\phi}$ and $\theta_M(\lambda, k) = \left(\frac{\phi \eta_M}{\phi \eta_M + \lambda} \right)^{\phi}$. Therefore, we have

$$\alpha = \frac{1 - \left(\frac{\phi \eta_F}{\phi \eta_F + \lambda + \gamma} \right)^{\phi}}{1 - \left[\left(\frac{\phi \eta_F}{\phi \eta_F + \lambda + \gamma} \right) \left(\frac{\phi \eta_M}{\phi \eta_M + \lambda} \right) \right]^{\phi}} \qquad (13)$$

- β: the probability that the MS's location flag is "ON" when the MS enters the Femto coverage. Form (5) and (13),

$$\beta = \left\{ \frac{1 - \left(\frac{\phi \eta_F}{\phi \eta_F + \lambda + \gamma} \right)^{\phi}}{1 - \left[\left(\frac{\phi \eta_F}{\phi \eta_F + \lambda + \gamma} \right) \left(\frac{\phi \eta_M}{\phi \eta_M + \lambda} \right) \right]^{\phi}} \right\} \left(\frac{\phi \eta_M}{\phi \eta_M + \lambda} \right)^{\phi} \qquad (14)$$

- C_M: the expected cost of an incoming call to the MS, where the MS is a Macro user and the call party is a Femto user, which can be obtained from (4), (12), and (13).
- C_F: the expected cost of an incoming call to the MS, where both the MS and the call party are Femto users, which can be obtained from (7), (12), and (14).

B. Numerical Results

Without loss of generality, we assume that λ, γ, and η_M are normalized by η_F, and c_M is normalized by c_F. The impacts of the input parameters on the output measures are elaborated below.

Effect on α and β: Figure 5 plots α and β as functions of the call arrival rate (λ) from Femto users, where $\gamma = 40\eta_F$. We assume that t_F (t_M) has the Exponential distribution with mean $1/\eta_F$ ($1/\eta_M$).

As λ increases, it is more likely that there are call arrivals at the last t_F period, which will set the location flag of the MS to "ON". In this case, when the MS enters the Macro coverage, the location flag is inconsistent with the actual MS status. Therefore, we observe that α increases as λ increases.

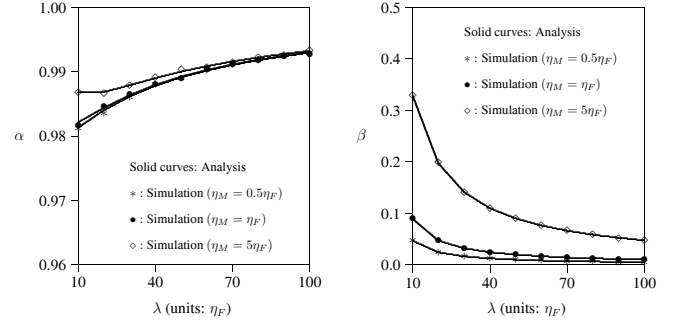When the MS enters the Femto coverage, the location flag of the MS is likely to be consistent with

Fig. 5. Effects on α and β ($\gamma = 40\eta_F$; $\eta_M = \eta_F$).
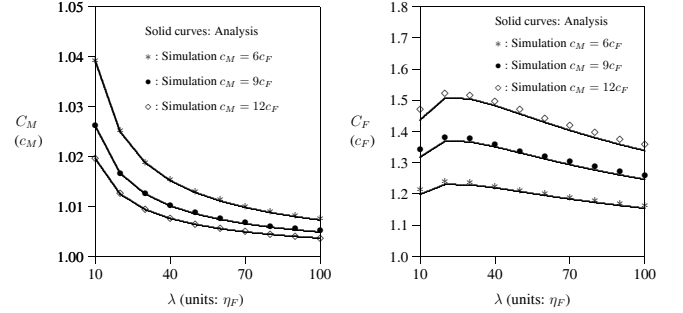
Fig. 6. Effects of λ ($\gamma = 40\eta_F$, $v_M = 1/\eta_M^2$, $v_F = 1/\eta_F^2$, and $\eta_M = \eta_F$).

the MS status if there is no call arrival at the last t_M period. Therefore, β decreases as λ increases.

Since α is mainly affected by the last t_F period rather than the t_M period, Figure 5 shows that α insignificantly increases as η_M increases. On the other hand, β is mainly affected by the last t_M period, and is significantly increased as η_M increases.

Effect of λ: Figure 6 shows the effects of λ on C_M and C_F, where $\gamma = 40\eta_F$, $v_M = 1/\eta_M^2$, $v_F = 1/\eta_F^2$, and $\eta_M = \eta_F$. During a t_M period, the extra re-routing overhead for an MS (if it is actually incurred) is shared by all calls from Femto users. As λ increases (the expected number of call arrivals increases), the shared cost for each call decreases, and thus C_M decreases.

When λ is small (e.g., $\lambda < 20\eta_F$), it is more likely that the first call arrival in a t_F period is an incoming non-Femto call or an outgoing call from the MS, which will correctly set the location flag of the MS (see the β effect when λ is small in Figure 5), and no extra cost for call arrivals will occur in this period. Therefore, C_F increases as λ increases. When λ is very large (e.g., $\lambda > 30\eta_F$), it is more likely that the first call arrival to the MS is from a Femto user when the location flag is "OFF" (see the β effect when λ is large in Figure 5), which results in extra call routing cost. This extra cost is equally shared by all calls from Femto users during the t_F period, and thus C_F decreases as λ increases.

Effect of γ: Figure 7 plots C_M and C_F curves against γ, where $\lambda = 40\eta_F$, $v_M = 1/\eta_M^2$, $v_F = 1/\eta_F^2$, and $\eta_M = \eta_F$. During a t_M period, the calls from non-
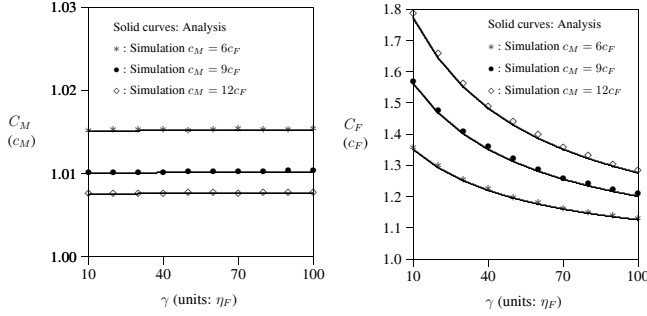
Fig. 7.   Effects of γ ($\lambda = 40\eta_F$, $v_M = 1/\eta_M^2$, $v_F = 1/\eta_F^2$, and $\eta_M = \eta_F$).

Fig. 9.   Effects of v_M ($\lambda = 20\eta_F$, $\gamma = 20\eta_F$, $\eta_M = \eta_F$, and $v_F = 1/\eta_F^2$).
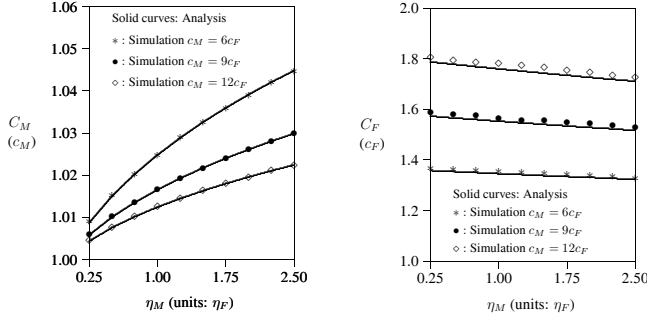

Fig. 8.   Effects of η_M ($\lambda = 20\eta_F$; $\gamma = 20\eta_F$).

Femto users to the MS (or the outgoing calls from the MS) will not change the status of the location flag. Therefore, C_M is not affected by the change of γ.

As γ increases during a t_F period, it is more likely that the first call arrival is a non-Femto user to the MS (or the outgoing call of the MS), which will correctly set the status of the location flag. Therefore, C_F decreases as γ increases.

Effect of η_M: Figure 8 illustrates the effects of η_M on C_M and C_F, where $\lambda = 20\eta_F$, $\gamma = 20\eta_F$, $v_M = 1/\eta_M^2$ and $v_F = 1/\eta_F^2$. As η_M increases (i.e., the expected t_M period decreases), the number of call arrivals in this t_M period decreases. Since the extra cost during the t_M period is shared by all calls from Femto users, a small t_M (a large η_M) implies a large C_M.

When the MS enters the Femto coverage, the status of its location flag is correct (i.e., "ON") if there is no call arrival during the previous t_M periods. Therefore, as η_M increases (the expected t_M period is shorter), it is more likely that no call arrives during the t_M period, and C_F decreases.

Effect of c_M/c_F: Figures 6-9 plot the C_M and C_F as the functions of c_M/c_F. For a t_M period, the extra cost (if incurred) for call rerouting is c_F. Increasing c_M is equivalent to decreasing c_F, which results in the decrease of C_M (see (4)). On the other hand, the extra cost for call routing during the t_F period is $c_M - c_F$ (when an extension call is mistakenly treated as a normal call). Therefore, increasing c_M will increase C_F.

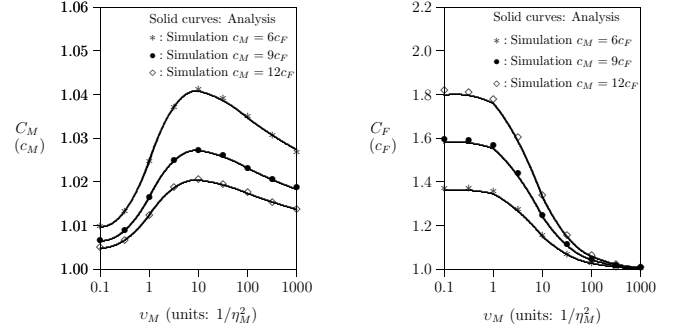As λ increases, it is more likely that the location

flag of the MS is changed by previous periods, which increases the extra cost during the t_F period. Therefore, C_R and C_D increases.

Effect of Variance v_M: Figure 9 plots the effect of the variance of the t_M distribution on C_M and C_F, where $\lambda = 20\eta_F$, $\gamma = 20\eta_F$, $\eta_M = \eta_F$, and $v_F = 1/\eta_F^2$. When v_M is very small ($v_M < 1/\eta_M^2$), if v_M increases, more shorter and longer t_M periods are observed, and the number of shorter t_M periods is more than that of longer t_M periods. In these shorter t_M periods, larger C_M values are observed (see the large η_M cases shown in Figure 8). Therefore, C_M increases as v_M increases when v_M is small (i.e., $v_M < 10/\eta_M^2$). When v_M is large ($v_M > 10/\eta_M^2$), if v_M increases, very short and very long t_M periods are observed. It is much more likely that the calls are not found in the very short t_M, and only arrive at the long t_M periods. In this case, the extra call setup cost is shared by large number of call arrivals. Therefore, C_M decreases as v_M increases. Figure 9 shows that C_F decreases as v_M increases. When v_M increases, more shorter t_M periods are observed, and it is more likely that the status of the location flag is correct when the MS enters the Femto coverage, which reduces the C_F cost.

The effect of v_F is similar to v_M and the details are omitted. Figures 6-9 show that in most cases, the FPBX incurs 0-4% of extra signalling costs for normal calls. The signalling cost can be significantly reduced by 75-90% for Femto-to-Femto calls, and the voice trunks from the BSs to the core network can be totally eliminated. Note that the voice trunks used in the FPBX solution and the standard UMTS network are the same for normal cellular calls (see Figure 3 (2) in Section II).

V. CONCLUDING REMARKS

This paper proposed a Femto Private Branch Exchange (FPBX) architecture. In this solution, the traffic of the Femto Base Stations (Femto BSs) installed in small areas are first concentrated to the FPBX. Then the FPBX is responsible for forwarding the traffic to the core network. In this architecture, an expensive normal call between two Femto users can be replaced by a low-cost extension call. In our approach, no explicit registration and de-registration are performed when the MS enters/leaves the Femto coverage. Therefore, the status

of the location flag may be incorrect, which incurs extra cost for call routing to the MS. However, our performance study indicates that the costs incurred due to implicit location update are not significant, and the FPBX can effectively reduces the call setup costs among Femto users by slightly increasing the normal cellular call setup costs. Beside the benefit of reducing signaling cost, the FPBX approach also totally eliminates the voice trunks to the core network in Femto-to-Femto calls.

As a final remark, the FPBX is a "plug-in" solution, which is transparent to the MSs and the mobile core network. This solution does not need to modify the protocols of the existing mobile telecommunications systems (i.e., MS, Femto BS, and the core network).

REFERENCES

[1] H. Kaaranen, A. Ahtiainen, L. Laitinen, S. Naghian, V. Niemi, *UMTS Networks: Architecture, Mobility and Services*, 2nd edition. John Wiley & Sons, 2005.
[2] Y.-B. Lin and A.-C. Pang, *Wireless and Mobile All-IP Networks*. John Wiley & Sons, 2005.
[3] J.-G. A. V. Chandrasekhar and A. Gatherer, "Femtocell networks: a survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59-67, 2008.
[4] Femto Forum, "Femtocells," 2010. [Online]. Available: http://www.femtoforum.org/.
[5] ABI Research, "Femtocell Market Challenges and Opportunities," 2007, Cellular-Based Fixed Mobile Convergence for Consumers, SMEs, and Enterprises.
[6] 3GPP, "UTRAN Architecture for 3G Home NodeB; Stage 2," 2009, 3G TS 25.467 (Version 8.4.0).
[7] Broadband Forum, "Broadband," 2009. [Online]. Available: http://www.broadband-forum.org/.
[8] K. S. Chava and J. How, "Integration of open source and enterprise IP PBXs," in *Proc. 3rd International Conf. TridentCom 2007*, May 2007, pp. 1-6.
[9] Y.-B. Lin, "Eliminating tromboning effect of international mobile call setup," accepted and to appear in *IEEE Trans. Wireless Commun.*.
[10] S.-R. Yang, Y.-C. Lin, and Y.-B. Lin, "Performance of mobile telecommunications network with overlapping location area configuration," *IEEE Trans. Veh. Technol.*, vol. 2, no. 57, pp. 1285-1292, Mar. 2008.
[11] ITRI, "Embedded IP PBX: complete IP PBX solution, industrial technology research institute," 2009. [Online]. Available: http://www.itri.org.tw/chi/tech-transfer/04.asp?RootNodeId=040&NodeId=041&id=2781.
[12] S. M. Ross, *Stochastic Processes*. John Wiley & Sons, 1996.
[13] S.-R. Yang and W.-T. Chen, "SIP multicast-based mobile quality-of-service support over heterogeneous IP multimedia subsystems," accepted and to appear in *IEEE Trans. Mobile Comput.* [Online]. Available: http://www.cs.nthu.edu.tw/~sryang/paper/SMMQSHIMS.pdf.
[14] P. Lin, W.-R. Lai, and C.-H. Gan, "Modeling opportunity driven multiple access in UMTS," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1699-1677, Sep. 2004.
[15] F. P. Kelly, *Reversibility and Stochastic Networks*. John Wiley & Sons, 1979.
[16] Y. Fang, "Modeling and performance analysis for wireless mobile networks: a new analytical approach," *IEEE/ACM Trans. Networking*, vol. 13, no. 5, pp. 989-1002, Oct. 2005.
[17] A.-M. Law and W.-D. Kelton, *Simulation Modeling and Analysis*, 3rd edition. McGraw-Hill, 2000.

Yi-Bing Lin (M'95-SM'95-F'03) is Chair Professor and the Dean, College of Computer Science, National Chiao Tung University, National Chiao Tung University. His current research interests include wireless communications and mobile computing. Dr. Lin has published over 220 journal articles and more than 200 conference papers. Lin is the co-author of the books *Wireless and Mobile Network Architecture* (co-author with Imrich Chlamtac; published by John Wiley & Sons), *Wireless and Mobile All-IP Networks* (co-author with Ai-Chun Pang; published by John Wiley & Sons), and *Charging for Mobile All-IP Telecommunications* (co-author with Sok-Ian Sou; published by John Wiley & Sons). Lin is an IEEE Fellow, an ACM Fellow, an AAAS Fellow, and an IET(IEE) Fellow.

Chai-Hien Gan (M'05) received his BS degree in computer science from Tamkang University in 1994, Taipei County, Taiwan, and both his MS. and Ph.D. degrees in computer science and information engineering from National Taiwan University, Taipei, Taiwan, in 1996 and 2005, respectively. From March 2005 to July 2007, he was a Research Assistant Professor in Department of Computer Science, National Chiao Tung University, Taiwan. Since July 2007, he has been a Researcher in the Information and Communications Research Labs, Industrial Technology Research Institute (ICL/ITRI), Taiwan. His current research interests include wireless and mobile computing, personal communications services, IP multimedia subsystem, and wireless Internet.

Ching-Feng Liang received M.S. degree in electronic engineering from National Taiwan University of Science & Technology (NTUST) in 1993 and joined the Information & Communication Laboratory (ICL) of Industrial Technology Research Institute (ITRI) as an engineer. Liang has led more than 10 projects of Taiwan Ministry of Economic Affairs (MoEA) to study and develop the technologies of mobile network and services including GPRS/3G core network, WLAN/Cellular interworking and number portability service. Liang is currently the engineering manager of the Telecommunication Network Technology Division of ICL/ITRI and a Ph.D. student of the Department of Computer Science, National Chiao-Tung University (NCTU). His current research interests include wireless and mobile computing systems.