

Transparent settlement model between mobile network operator and mobile voice over Internet protocol operator

Luzango Pangani Mfupe^{1,2}, Mjumo Mzyece¹, Anish Mathew Kurien¹

¹Department of Electrical Engineering, FSATI, Tshwane University of Technology, Private Bag X680, Pretoria 0001, South Africa

²Meraka Institute, Council for Scientific and Industrial Research (CSIR), PO Box 395, Pretoria 0001, South Africa
E-mail: Lmfupe@csir.co.za

Published in *The Journal of Engineering*; Received on 3rd September 2014; Accepted on 7th October 2014

Abstract: Advances in technology have enabled network-less mobile voice over internet protocol operator (MVoIPO) to offer data services (i.e. voice, text and video) to mobile network operator's (MNO's) subscribers through an application enabled on subscriber's user equipment using MNO's packet-based cellular network infrastructure. However, this raises the problem of how to handle interconnection settlements between the two types of operators, particularly how to deal with users who now have the ability to make 'free' on-net MVoIP calls among themselves within the MNO's network. This study proposes a service level agreement-based transparent settlement model (TSM) to solve this problem. The model is based on concepts of achievement and reward, not violation and punishment. The TSM calculates the MVoIPO's throughput distribution by monitoring the variations of peaks and troughs at the edge of a network. This facilitates the determination of conformance and non-conformance levels to the pre-set throughput thresholds and, subsequently, the issuing of compensation to the MVoIPO by the MNO as a result of generating an economically acceptable volume of data traffic.

1 Introduction

Traditionally, in mobile networks, off-net traffic is considered to be circuit-switched traffic that originates from a mobile network operator (MNO) subscriber, terminated to other networks through a gateway (point of interconnection). The point of interconnection is essentially where call information is collected for the purpose of inter-operator settlements. On-net traffic is considered to be an internal form of circuit-switched traffic that flows between subscribers of the same MNO without passing through a point of interconnection [1]. However, in a mobile voice over internet protocol (MVoIP) context, an MVoIP operator (MVoIPO) such as Viber [2], does not own any network infrastructure but uses software applications to provide services within an MNO's network. In this particular context, the off-net traffic is considered to be packet-switched traffic which originates from an MVoIP user who is subscribed to a packet-based mobile network, terminated by the same mobile network, or other networks, to a non-MVoIP user through a point of interconnection.

In contrast, on-net, or peer-to-peer (P2P), traffic is considered to be packet-switched traffic which has originated from one MVoIP user who is subscribed to a mobile network, terminated at another MVoIP user subscribed in the same or a different mobile network and, vice versa, without passing through a point of interconnection. Additionally, two MVoIP users from the same MVoIPO can communicate in P2P mode regardless of which mobile network they are subscribed to. It is worth noting that the cost of an on-net mobile voice call is considered to be equal to the total sum of the call origination cost and the call termination cost [3]. However, the fact that on-net MVoIP traffic traverses to and from the application-enabled user equipment (UE) within an MNO's network, without necessarily passing through a point of interconnection, raises a fundamental problem of how the primary operator (PO) and the secondary operator (SO) (in this case the MNO and MVoIPO, respectively) can coexist with mutual economic benefits. This problem cannot easily be solved through traditional settlement mechanisms, and sometimes it is the source of conflicts between the two operators. A classic example is a case in the United Kingdom where the PO known as T-mobile decided to disable the

internet telephony feature from its subscribers' handsets blocking them from accessing MVoIP services provided by an SO known as Truphone [4]. Throughout this paper we use the following terms interchangeably; on-net and P2P referring to MVoIP calls made and terminated within an MNO's network. MNO and MVoIPO, respectively, refer to PO and SO.

1.1 Establishment of on-net MVoIP calls within third generation partnership project (3GPP) universal mobile telecommunications system (UMTS) network

Fig. 1 shows how a P2P mobile VoIP call is established between two UEs subscribed to the MNO within a 3GPP UMTS network [5] [The UMTS network architecture is used for illustration purposes only. The model proposed in this paper can be applied in any third generation and beyond packet-based cellular network architecture(s)]. Both application-enabled UEs must also be registered with the MVoIPO via a session initiation protocol (SIP) [6]. This protocol is central to the concept of internet multimedia subsystem envisaged to effectively enable the unbundling of voice services in the data-based next generation mobile networks such as 3G UMTS and the 4G long-term evolution (LTE) [7]. SIP is often utilised to facilitate the end-to-end application layer interaction between the server and UE. Afterwards, either of the UEs can initiate a P2P call to the other by sending a call request (SIP signalling message) that may contain information such as codec type and session type via the proxy server. The SIP proxy server detached from the mobile network infrastructure (i.e. can be hosted in the cloud) routes the call request across the UMTS terrestrial radio access network (UTRAN), via the gateway general packet radio service (GPRS) support node (GGSN) and the Serving GPRS SN (SGSN), to the corresponding UE. Subsequently, the requested UE might send a reply (SIP signalling message) via the same route. Subsequently, an end-to-end data tunnel is established at the UMTS user-plane. Packet data carrying the actual voice conversation between the two calling parties (UE 'a' and UE 'b') will be tunnelled back and forth across the UTRAN, the GGSN and the SGSN.

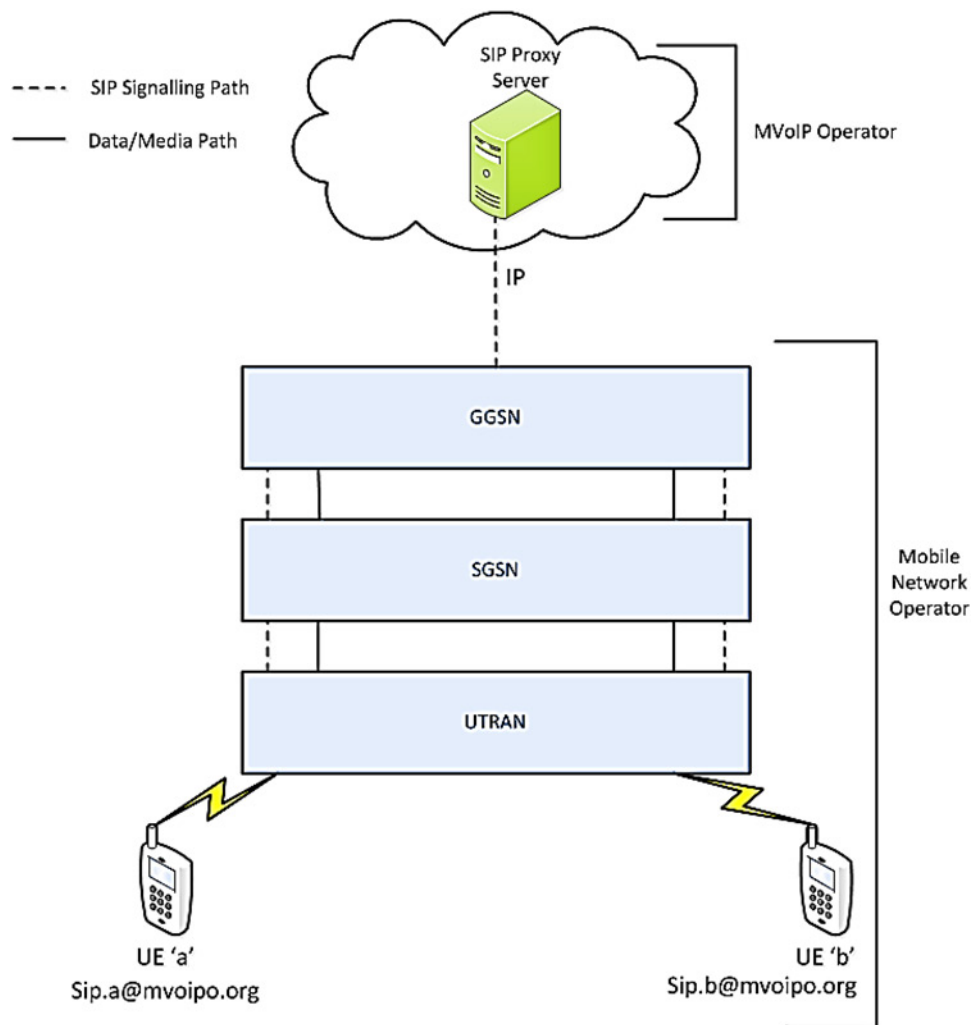


Fig. 1 Illustration of on-net MVoIP call flow within UMTS network

1.2 Contributions of this paper

Decoupling of transport and service has introduced new types of network-less operators at the 'edge' of the network (particularly in the application layer) that can offer best-effort complementary or competing services to the ones offered by the network owner [6]. A recent report [8] has indicated that mobile data traffic grew by 81% in 2013 and forecasted that this trend will surpass 15.9 Exabyte by 2018 which is 11-fold over 2013. Application-enabled MVoIP traffic contributed a significant amount of data in this growth [8]. This paradigm shift comes with profound challenges to the existing interconnection arrangements because the network owner has little or no visibility to the types of services offered by the network users [9]. It is noteworthy that most research on MVoIP tends to mainly focus on how a network user should be charged by the network owner for using network resources (i.e. methods for compositing, aggregating, correlating and description of charging records) or how to structure a bilateral cost sharing arrangement between network user and network owner. This is not directly relevant to how the on-net MVoIP traffic should be settled between the MNO and MVoIPO. Furthermore, we note in [10] that a network user faces consequences from the network owner, for over-utilisation of network resources (exceeding the pre-set service level agreement (SLA) thresholds), where it is rewarded for under-utilisation of network resources. To the best of our knowledge there is not yet a practical model that adequately considers the interconnection settlement mechanism that addresses the coexistence between MNO and

network-less MVoIPO, particularly the key issue of how to handle P2P MVoIP calls originated and terminated within an MNO's network. We argue that a network-less operator offering services over internet protocol (IP) only, such as an MVoIPO, should be compensated by the network owner, the MNO, for generating an economically viable amount of extra data traffic. This is because this extra traffic is because of the joint efforts of both operators. More specifically, this paper makes two main contributions:

- Proposing a new inter-operator settlement scheme, the transparent settlement model (TSM), based on flexible compensation, that is, with multiple possible forms of monetary or non-monetary compensation paid by the network owner (the MNO) to the network user (the MVoIPO).
- Replacing the two central concepts of conventional SLA-based schemes, namely violation and punishment (i.e. users who violate the set SLA thresholds are punished by having their services restricted or removed), with two new concepts, achievement and reward. In this new paradigm, an MVoIPO should be compensated for generating an economically acceptable amount of data traffic (i.e. when an MVoIPO goes above the set SLA thresholds, it gets rewarded instead of getting punished). This is based on the fact that the MNO benefits from the internet data bundle charges paid by users of the MVoIP application.

The rest of this paper is arranged as follows. Section 2 highlights related work. Section 3 presents the TSM framework and design.

The simulation experiments, results and analysis are discussed in Section 4. Section 5 concludes this paper.

2 Related work

There is a significant body of literature and approaches dedicated to the settlements schemes of different types of operators in different layers of IP-based networks that cover the technical, economic and regulatory aspects of interconnection [10–12]. Reviews of pricing schemes and emerging charging methods for services offered by disparate IP-based operators are found in [13–16]. Techno-economic analysis of dynamic pricing framework for secondary users accessing LTE networks using dynamic spectrum access techniques [17]. Approaches and concepts that have a more direct implication to this paper include but are not limited to the following:

- ‘Parameter-based static pricing schemes in dynamic environment [10, 16–18]’ provide a mechanism to predefine usage prices with respect to the SLA contract thereby allowing each network user to be charged in a specific class of services. The most interesting method in this category is the cumulus pricing scheme (CPS); it is based on a flat-rate-type contract between the network user and the network owner such that the network user specifies the expected resource usage.

The SLA is monitored periodically by the feedback mechanism (cumulus points) in the pre-set time scales. During these pre-set times if the user (SO) violates the contract by exceeding the pre-set usage thresholds the network owner (PO) will take an action such as warning or contract re-negotiation. The CPS utilises IP traffic characteristics, such as the mean and standard deviation, as a way of solving the feasibility problem of internet tariffing encountered by internet service providers (ISPs) when pricing multi-service networks in the internet. The feasibility problem refers to the trade-off between customers and their interest in predictable and transparent tariffs against ISPs interest in achieving economic efficiency in operating the network and technical efficiency for the accounting operation that is characterised by predefined prices.

- ‘The digital marketplace approach [19]’ reduces the risk of conflicts among stakeholders in the wireless network business value chain. Broadly, the authors argue that the adoption of IP as a transport mechanism of choice might be a source of conflict among stakeholders. For example, a conflict may arise when a mobile user demands certain services which are prohibited by an MNO but can be made available by a third party provider through the application layer. Therefore the authors proposed an open market-based approach that would effectively allow MNOs and third party operators to coexist with mutual benefits.
- ‘Real option pricing approaches [20]’ deal with the issues of co-existence between the POs (MNOs) offering services in their networks and the network-less SOs such as MVoIPOs offering services only above IP over PO’s networks. It is argued that for such coexistence to succeed, two central aspects must be studied: (i) types of differentiated services offered by the two operators and (ii) the objectives and risk behaviours of the two operators (i.e. the objectives of the two operators greatly differs: the MNO focuses on recovering its costs at the minimum risk possible through traditional pricing schemes. Meanwhile the MVoIPO focuses on attracting as many users as possible using innovative pricing schemes at competitive rates.).
- ‘Packet voice interconnections through the IP exchange (IPX) [21]’. IPX is a predecessor of the GPRS roaming exchange. It is a global private network developed by the global system mobile (GSM) association (GSM-A). IPX is a service-aware system that supports end-to-end SLAs and the principles of cascading interconnection payments. The IPX supports three different types of

interconnection models: bilateral transport only, bilateral service transit and multilateral hub service.

IPX paves the way for a global MNOs consolidation in response to the rising stakes of the new breed of third party application service providers such as mobile virtual network operators (MVNOs). IPX empowers MNOs with the means to resist interconnection with third party operators particularly when the services (e.g. rich communication services and applications) to be offered by the latter are a threat or do not conform to the existing billing arrangements. We therefore argue that this approach does not sufficiently address the settlement issue of a third party’s on-net MVoIP calls within MNO’s network.

3 TSM framework

3.1 Modelling and design

This section presents an overview of our proposed four-stage framework mechanism that enables the MNO and the MVoIPO to handle interconnection settlements, particularly the on-net traffic generated by MVoIPO users within the MNO’s network. To function properly, the TSM framework is subdivided further into two distinct cycles: operational and contractual. In the operational cycle, consisting of stages 1, 2 and 3 of the framework, the MVoIPO’s throughput is monitored and measured at the edge of MNOs network against the set SLA thresholds in real-time, and compensation is issued by the MNO to the MVoIPO. The operational cycle works on short-term timescales (typically minutes). In stage 4 of the framework, the economic merit of the SLA is re-evaluated in order to ascertain its viability on medium-to-long-term timescales (e.g. weeks, months).

3.1.1 TSM framework is based on the following underlying assumptions:

- *Transparency*: As the name suggests, it is in the best interest of both operators to disclose all technical and trade information relevant to the contract.
- *Thresholds*: The PO and the SO initially agree that the latter will eventually qualify for compensation after a consistent generation of uplink and downlink packet data throughput in the network owned by the former. The actual generated packet data throughput levels must consistently be equal to or exceed or equal a pre-set threshold of N bandwidth units per second where $N > 0$.
- *Conformance events*: Compensation will be dynamically issued based on the sum of conformance events (i.e. the moments of time when MVoIP throughput is equal to or above the set thresholds). The conformance events will be monitored and statistically collected at the GGSN to the n th percentile of time.
- *Type of compensation*: There are different possible types of compensation for the MVoIPO to pre-select from in addition to money, for example, prioritisation of MVoIP traffic during congested periods [17]. Fig. 2 shows the four-stage TSM framework.

The framework works as follows:

- *Stage 1*: The agreed MVoIP throughput thresholds (e.g. ‘ $Y \geq X$ ’ bandwidth units) are set by both operators, followed by pre-selection of the type of compensation for the MVoIPO.
- *Stage 2*: The monitoring period starts during which the MVoIP throughput is actively measured and monitored at the edge of the MNO’s core network.
- *Stage 3*: The MNO begins to issue compensation to the MVoIPO according to the usage patterns of the throughput (uplink and downlink) generated by users of the MVoIP application.
- *Stage 4*: Revisits the economic model before executing another transparent settlement cycle. The techno-economic model here refers to the game-theoretic decision-making modelling approach

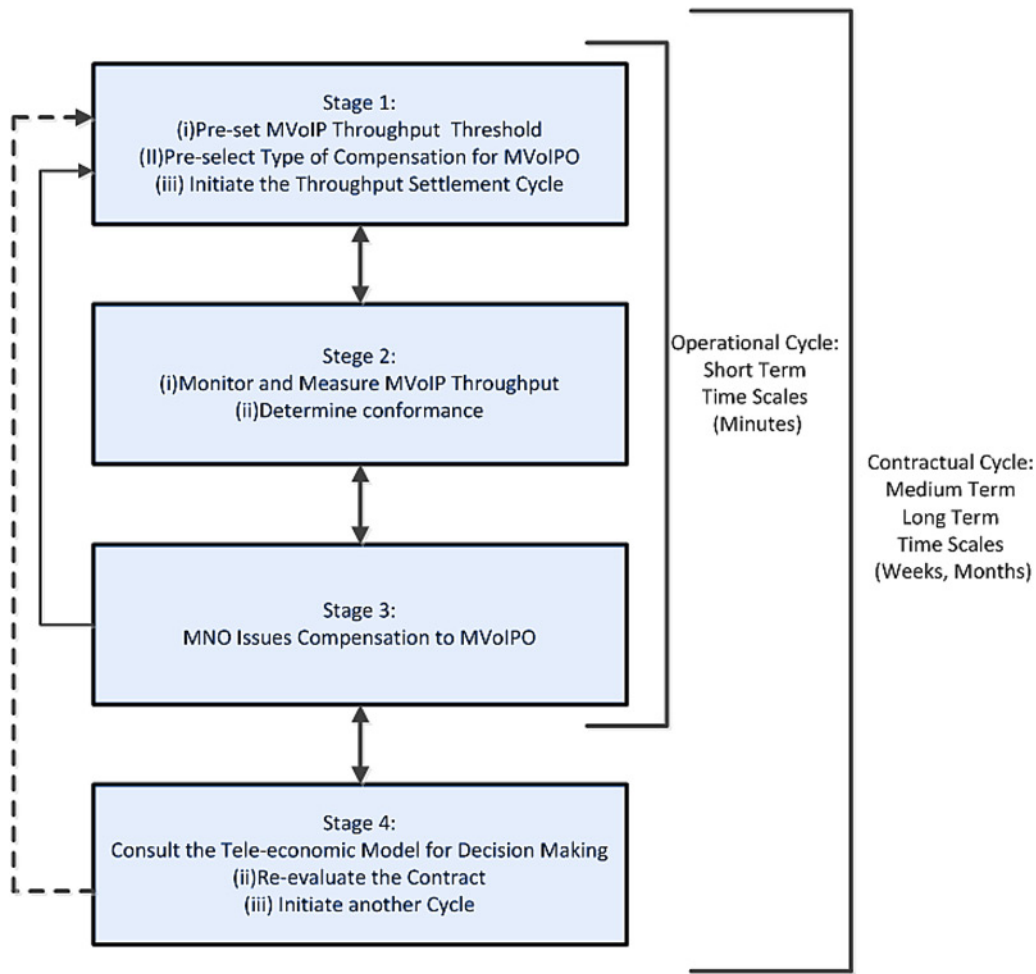


Fig. 2 High-level view of the SLA-based TSM framework

used to govern coexistence between the two types of operators [17, 22, 23]. This approach examines the decisions made by the MNO to invest in excess network capacity as well as the trade-off between the substitutionary effects and positive network externality effects that may be introduced by the MVoIPO.

- *Agreement re-evaluation*: Both operators will review the economic merits of TSM at the end of each contractual cycle on medium-to long-term timescales.

3.2 Monitoring and measuring the MVoIP throughput at the edge of MNO's CN in discrete-time

The discrete-time-based method used in this section allows the continuous fluctuations of the on-net MVoIP traffic to be measured in the predefined fixed-sized windows of time as a stationary process. In [24], it is suggested that when monitoring a continuous time process at a constant measuring distance, the continuous time intervals are transformed into a discrete-time distribution. This observation is true with respect to the Poisson process as it is transformed into a discrete-time process with the correlation between inter-arrival times of duration being zero.

Suppose that the MVoIP throughput measured at the MNO's GGSN is of pure chance traffic (PCT) I type (PCT type I refers to the traffic stream that has a Poisson arrival process and exponential call holding times). Furthermore, suppose that the MVoIP call arrivals are independent identically distributed. The density function of time intervals from each arrival to the first measurement is uniformly distributed and equal to $1/h$ (where h refers to the constant measuring distance in a given time period), whereas the

probability of observing zero measurement during the call holding time is denoted by $p(0)$. If the exponential distribution holding time interval is $F(t) = 1 - e^{-\mu t}$ with a finite mean μ^{-1} , then according to [25] the following discrete distribution will be observed

$$p(0) = 1 - \frac{1}{\mu h} (1 - e^{-\mu h}) \quad (1)$$

$$p(k) = \frac{1}{\mu h} (1 - e^{-\mu h})^2 \cdot e^{-(k-1)\mu h}, \quad k = 1, 2, \dots \quad (2)$$

Therefore the j th derivatives of the probability generating function $Z(z)$ when the value of $z = 1$ is given by [24]

$$Z^{(j)}(1) = \frac{j!}{jh} \cdot \frac{1}{(e^{\mu h} - 1)^{j-1}} \quad (3)$$

In particular, (3) yields the mean m_1 , the variance σ_j^2 and the form factor ε (the form factor refers to the measure of accuracy in the measurements) [25]

$$m_1 = Z^{(1)}(1) = \frac{1}{\mu h} \quad (4)$$

$$\sigma_j^2 = Z^{(2)}(1) + Z^{(1)}(1) - Z^{(1)}(1)^2 \quad (5)$$

$$\varepsilon = \left(\frac{\sigma_j^2}{m_1} \right)^2 = \mu h \cdot \frac{e^{\mu h} + 1}{e^{\mu h} - 1} \geq 2 \quad (6)$$

However, the intensity rate and the variance of the items of call traffic that has a unit holding time ($\mu = 1, = A$) arriving inside the time interval T are yielded by [26]

$$\mu_j = A \quad (7)$$

$$\sigma_j^2 = \frac{A}{T} \cdot \varepsilon \quad (8)$$

Therefore, by inserting the value of the form factor ε obtained in (6) in (8), we can express the variance as

$$\sigma_j^2 = \frac{2A}{T} \quad (9)$$

The related derivations of the variance of PCT I traffic type are found in [27–29].

Suppose that the conformance of MVoIP throughput generated by all hosts and received at the MNO's GGSN needs to be monitored (the conformance and non-conformance of MVoIP throughput is according to the SLA threshold assumptions described earlier). First, we divide the throughput monitoring time between $t=0$ to Z into h steps, such that each step has a size (Z/h) (e.g.

monitoring time lengths in hours). Subsequently, we subdivide each step into small fixed-size throughput-monitoring windows (partitions) in minutes. Secondly, the total throughput at that particular partition needs to be calculated. We calculate the total throughput at the j monitoring window as follows

$$Te_j = \frac{\sum_{t=(z/h)_j}^{(z/h)_{j+1}} ARU}{(Z/H)}, \quad j = 1, 2, 3, \dots, h \quad (10)$$

where Te_j is the total amount of data transmitted in bits per second, A is the MVoIP application's transmitted data from the source to the destination per host, R is the number of repetitions per host per second and U is the number of hosts.

Thirdly, the average of the data measured within the window $(z/h)_j$, $j=1, 2, \dots$ (where j is the monitoring time length in minutes) is determined. Assuming that each j consists of n observations (samples) such that $(z/h)_{j-n+1}, \dots, (z/h)_j$, then we compute their mean by

$$\left(\frac{\bar{z}}{h}\right)_j = \frac{1}{n} \sum_{k=j-n+1}^j \left(\frac{z}{h}\right)_k \quad (11)$$

TSM Algorithm

Input: {Pre-set throughput threshold}

{Pre-set throughput monitoring window time lengths}

{Pre-select the type of compensation to secondary operator}

Output: {Expected economic benefits to secondary operator and primary operator}

Sequence:

1. Transform the continuous Poissonian MVoIP traffic process into a discrete distribution and determine its theoretical mean and variance (Equations (1) to (9));
2. Partition the MVoIP traffic and determine the throughput within the partition (Equation (10));
3. Determine the mean value of the observations within the partition (Equation (11));
4. Determine the normalised variance of the mean of the observations inside the partition (Equation (12));
5. Construct the confidence interval for the mean throughput inside the partition (Equation (13));
6. Issue a conformance event to MVoIPO only if new throughput measurements are equal to or above the pre-set threshold (Equation (14));
7. Issue a pre-selected type of compensation to MVoIPO based on the accumulation of conformance events (Equation (15));
8. Revisit the decision-making model (economic) model in order to re-evaluate the contract before initiating another TSM cycle (this is performed on medium- to long-term timescales e.g., weeks, months).

Fig. 3 Algorithmic form of the TSM framework

from which we derive the normalised variance of the mean as

$$\hat{\sigma}_j^2 = \frac{1}{Te_j^2} \frac{1}{n-1} \sum_{k=j-n+1}^j \left(\left(\frac{z}{h} \right)_k - \left(\frac{\bar{z}}{h} \right)_j \right)^2 \quad (12)$$

That is, the normalised variance is an unbiased estimator of the original distribution such that $\hat{\sigma}_j^2$ has to be close to the theoretical variance derived in (9).

Fourthly, the n th percentile confidence interval estimates are constructed [30] to observe the MVoIP throughput conformance using the Student t -distribution [31] as follows

$$100(1 - \alpha)\% = \left(\left(\frac{\bar{z}}{h} \right)_j \pm \left(\frac{t_{1-\alpha}}{2}, n-1 \right) \hat{\sigma}_j^2 \right) \quad (13)$$

where $((t_{1-\alpha}/2), n-1)$ is the $(1 - (\alpha/2))$ fractal of the Student t -distribution with $n-1$ (degree of freedom), whereas $(1 - \alpha)$ is the probability describing the level of the confidence and n is the number of observations. The values of α are obtained in the t -distribution table [32].

Fifthly, let conf_j denote the conformance event to be gained by the SO in the throughput-monitoring window (such events occur each time when the new throughput measurements found in the confidence interval are found to be within or above the pre-set SLA threshold). We calculate the conformance events by

$$\text{conf}_j \rightarrow \left(\frac{z}{h} \right)_{j+1} \geq N_j \quad (14)$$

where N_j is the pre-set SLA threshold.

Lastly, the accumulation of the conformance events results in the issuing of a pre-selected type of compensation to the SO by the PO such that

$$\phi(\text{comp}) = \sum_{j=1}^n \text{conf}_j \quad (15)$$

where $\phi(\text{comp})$ is the compensation with pre-selected type of choices (e.g. the prioritised MVoIP traffic) as discussed earlier.

In summary, the description of the proposed four-stage TSM framework presented in Fig. 2 can be presented in algorithmic form as shown in Fig. 3.

4 Simulation experiments

This section describes a simulation experiment for the logical inter-connection between the MNO and the MVoIPO at the application layer. The simulation scenarios are constructed in OPNET Modeller version 15.0 [33] by re-using the expert service prediction module in order to analyse the MVoIP traffic growth and the SLA monitoring.

4.1 Simulation parameters

The following inputs to the simulation model have been used during different simulation runs. Table 1 introduces key parameters used in the simulation experiments.

Two simulation rounds, each consisting of five runs, were performed using five different seeds for random number generation (RNG). A traffic growth rate of 10% and a background traffic-scaling factor of 1.5 in best-effort mode were assumed. The UEs used a G.729 voice codec with a rate of 8 kbps and a constant packetisation interval of 20 ms (OFF – silent periods and ON – talk spurts). The call inter-arrival followed the Poisson process with exponential call holding times.

Furthermore, 30 UEs were used to model simultaneous users of the MVoIP application in a single cell with the assumption that

Table 1 Voice application, profile, traffic growth and SLA attributes

Parameters	Values
<i>Voice application and profile attributes</i>	
application start time offset, s	Poisson (2)
application duration	end of profile (the application will end when the profile duration has expired)
application inter-repetition time, s	exponential (2)
application number of repetitions	unlimited
application repetition's pattern	serial (application start time is computed by adding the inter-repetition time to time at which the previous session completed)
profile repetition mode	serial
profile inter-repetition time, s	exponential (3600 s)
profile number of repetitions	unlimited
voice codec	G.729
call signalling protocol	SIP
application type of service	best-effort (0)
UE's P2P communication	enabled
<i>traffic growth and SLA attributes</i>	
simulation rounds	2×5 runs
traffic growth rate type	compound
growth rate	10%
background traffic-scaling factor	1.5
RNG	five seeds
number of UEs	30
MVoIP throughput threshold N_j (uplink, downlink)	$\geq 30\,000$ bps
simulation duration	60 min
throughput-monitoring windows	3 min, 5 min
time lengths	
confidence interval (CI)	98%
monitoring location	GGSN

non-MVoIP users were utilising the remaining cell capacity. In this way, the uplink and downlink throughput thresholds of $\geq 30\,000$ bps were set. It is important to note that the pre-set thresholds were only the expected MVoIPO throughput but not the actual throughput, which was expected to vary during the settlement cycle. Additionally, the time lengths of monitoring windows at the GGSN were set to 3 and 5 min.

4.2 Simulation results

Tables 2 and 3 show the results of the MVoIP throughput (in the uplink and downlink directions) measured during the 3 and 5 min monitoring window time lengths.

Results recorded in Table 2 demonstrate that the MVoIP throughput monitored in the uplink direction conforms to the pre-set threshold ($\geq 30\,000$ bps) for about 98th percentile of the simulation period

Table 2 Mean MVoIP uplink throughput (bps) received at the GGSN during the 3 and 5 min throughput-monitoring windows collected over the simulation period of 1 h

Monitoring window	3 min	5 min
seed 1	35 218 bps	31 907 bps
seed 2	47 027 bps	42 874 bps
seed 3	47 731 bps	44 860 bps
seed 4	47 874 bps	45 271 bps
seed 5	36 017 bps	34 047 bps
mean throughput	42 773 bps	39 792 bps
98% CIs (lower bound–upper bound)	35 900–49 646 bps	32 773–46 811 bps

Table 3 Mean MVoIP downlink throughputs (bps) received at the GGSN during the 3 and 5 min throughput-monitoring windows collected over the simulation period of 1 h

Monitoring window	3 min	5 min
seed 1	24 404 bps	21 664 bps
seed 2	38 718 bps	35 062 bps
seed 3	39 013 bps	35 973 bps
seed 4	39 010 bps	36 470 bps
seed 5-	24 768 bps	22 651 bps
mean throughput	33 183 bps	30 364 bps
98% confidence intervals (lower bound–upper bound)	24 148–33 880 bps	23 687–37 041 bps

of 1 h. Hence, the MVoIPO will receive full compensation from the MNO for the uplink throughput. A statistical graph in Fig. 4 illustrates conformance and non-conformance levels by the MVoIPO traffic at UMTS GTP (GPRS tunnelling protocol) level in the uplink direction received at the GGSN node when the throughput-monitoring window time length was set to 3 min. It is worth noting that the bars in the graph show the statistical average for each monitoring window.

In this scenario, the average value of conformance by the MVoIPO was 35 218 bps; that is, it was above the pre-set threshold

(of $\geq 30\,000$ bps). However, in the first 360 s, the MVoIPO was below the threshold two times.

Results in Table 3 indicate that the MVoIP throughput monitored in the downlink direction partially conforms to the pre-set threshold of $\geq 30\,000$ bps for about 98th percentile of the simulation period of 1 h (i.e. the lower confidence interval bounds fall below the pre-set throughput thresholds, whereas the upper confidence interval bounds fall within the set throughput threshold). In other words, the MVoIPO will be receiving partial compensation from the MNO for the downlink throughput.

Fig. 5 contains a statistical graph illustrating levels of conformance and non-conformance levels by the MVoIPO traffic at UMTS GTP level in the downlink direction received at the GGSN node when the throughput-monitoring window time length was set to 3 min. In this scenario, the average value of non-conformance by the MVoIPO was 24 404 bps. It was below the pre-set throughput threshold (of $\geq 30\,000$ bps) 100% of the time, in which it was below the threshold 20 times.

4.3 Analysis

This section analyses the implications of the results on key factors affecting the successful implementation of the TSM.

Throughput threshold:

- It was observed that the pre-set threshold ($\geq 30\,000$ bps) is roughly about 60% of the MVoIPO's peak throughput. This

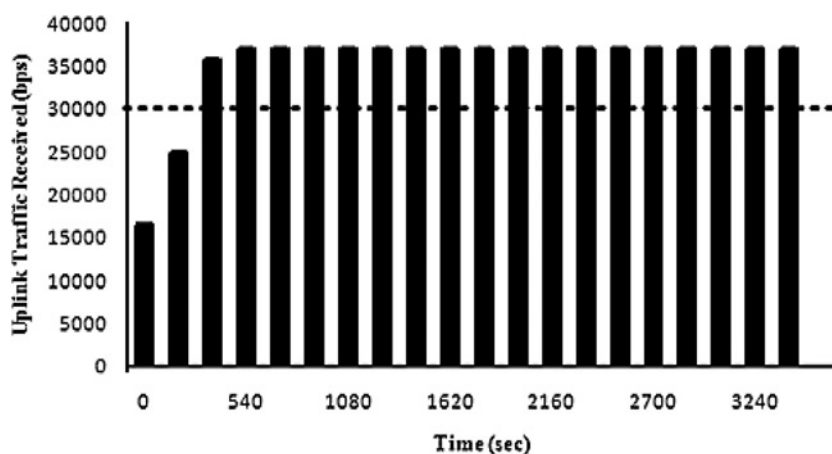


Fig. 4 MVoIPO's uplink throughput received at the MNO's GGSN (Seed = 1; monitoring window time = 3 min; statistics collected over a period of 1 h)

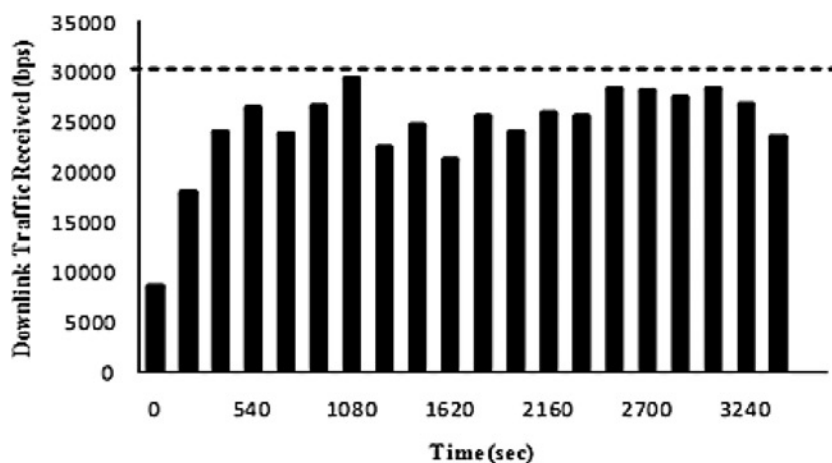


Fig. 5 MVoIP downlink throughput received at the GGSN (Seed = 1; monitoring window time = 3 min; statistics collected over period of 1 h)

raises the question of how to select the optimal threshold values that will offer mutual benefits for both operators (i.e. both the MNO and the MVoIPO). On one hand, setting too large a value might cause the MVoIPO not to benefit from the MNO's incentives despite generating a substantial amount of data traffic.

On the other hand, setting too small a value might enable the MVoIPO to benefit from the MNO's incentives even if it does not generate an economically acceptable amount of data traffic. One possible solution for this crucial issue is to allow a grace period before the formal commencement of the SLA so that both operators can learn about each other's behaviours (i.e. the MNO learns about the MVoIPO's usage patterns, whereas the MVoIPO learns about the MNO's service availability and quality).

Monitoring window time length:

- The next notable factor is the choice of the right monitoring window time lengths. According to [25, 26], the arrival of telephony traffic at intervals of up to 15 min can be treated as a stationary Poisson process. Larger intervals are considered as non-stationary because of the possible intensity variations that may introduce difficulties in taking measurements.

In this paper, smaller throughput-monitoring window time lengths were used to avoid complications related to non-stationary data even though the number of monitored events that fell below the

pre-set throughput threshold was roughly the same in the downlink direction for both the smaller 3 min window and the bigger 5 min window. It is worth observing, however, that an SLA contractual cycle may span several weeks or months, in which case the use of smaller throughput-monitoring window time lengths would result in an increased frequency of monitoring. However, according to [34], raw packet counting is a costly exercise; therefore it is proposed that using bigger throughput-monitoring window time lengths of up to 15 min would be more cost efficient.

Accuracy of the measurements:

- The accuracy of the measurements is largely attributable to the use of the variance of the MVoIP throughput distribution in the confidence interval calculations. The variance, being a second-order statistic, provides much needed granularity of measurements in each monitoring window. It is asserted in [28, 29] that more hidden characteristics of the load are revealed by the use of the second- and higher-order statistics than by the first-order statistics. Fig. 6 compares the sample mean and the variance of the received uplink MVoIP throughput with a 98% confidence interval.

In Fig. 6, the sample mean (in the first graph), a first-order statistic, does not reveal all the hidden behaviour (unevenness characteristics) of the uplink MVoIP throughput. By contrast, the variance (in the second graph), a second-order statistic, reveals both the peaks and the low points of the uplink MVoIP throughput,

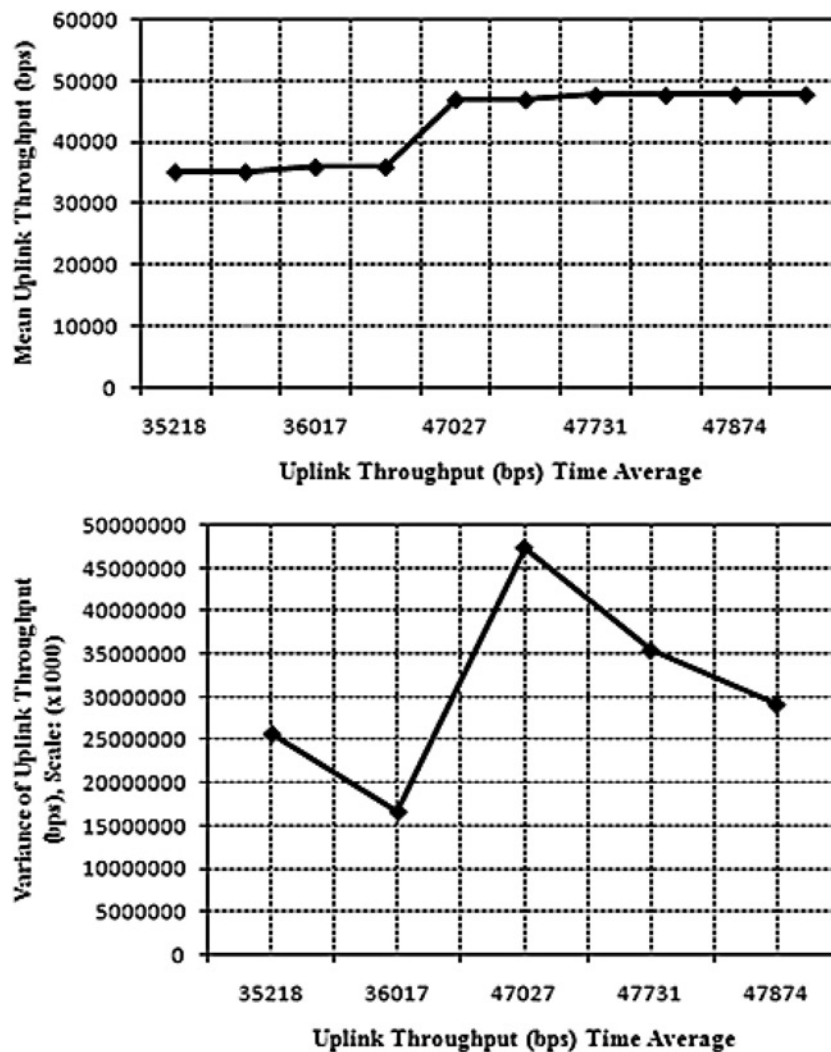


Fig. 6 Mean (upper graph) and the variance (lower graph) at 98% confidence intervals for the aggregated uplink MVoIP throughput measured during the 3 min monitoring windows

which is vitally important for the accuracy of the measurements used in the TSM.

Issuing of compensation to the SO:

- Another important factor relates to the issuing of compensation to the MVoIPO, particularly the exact amount of conformance events required to adequately and accurately determine this matter. We propose that this area should be left to the two operators to negotiate.

5 Conclusion

It is worth noting that despite the on-going explosion of data usage and the need for rollout of next generation high-speed packet-based cellular network infrastructure; many MNOs still find it difficult to gain maximum economic benefits from data traffic. In this paper, we have proposed a settlement model between network owner the MNO and network user the mobile voice internet operator (MVoIPO) for the on-net MVoIP data traffic within the MNO's network. More broadly, our model differs from well-known SLA-based schemes by replacing the two traditional concepts of violation and punishment (i.e. users who violate the set SLA thresholds are punished by having their services restricted or terminated), with two new concepts, achievement and reward. We argued that an MVoIPO should be compensated for generating an economically acceptable amount of data traffic. In other words, MVoIPO should still be incentivised instead of getting punished even when it occasionally goes below the set SLA thresholds. This is due to the fact that MNO will benefit from the internet access charges paid by users of the MVoIP application(s). Our approach can be instrumental in creating innovative business models for POs and SOs to effectively serve the future web 3.0 consumers. Additionally, this approach can aid in establishing 'radio frequency (RF) spectrum usage settlement models' between POs and SOs in the emerging opportunistic networks that are using dynamic spectrum access (DSA) techniques.

6 References

- [1] Kanervisto J.: 'MVNO Pricing Structures in Finland (Publications of Transport and Communications No. 21/2005)', Helsinki, Ministry of Transport and Communications Finland, 2005
- [2] Viber: 'Free calls, text and picture sharing with anyone, anywhere!'. Available at <http://www.viber.com/>, accessed September 2014
- [3] Armstrong M., Wright J.: 'Mobile call termination', *Econ. J.*, 2009, **119**, (538), pp. 270–307
- [4] Judge P.: 'T-mobile blocks rival's VoIP calls', *Techworld*. Available at <http://www.news.techworld.com/mobile-wireless/9191/t-mobile-blocks-voip-rivals-calls/>, accessed January 2012
- [5] 3rd Generation Partnership Project (3GPP): 'Network Architecture (T1TRQ 3GPP 23.002-330)', Standards Committee T1 Telecommunications, version 3.3.0 edition, 2000
- [6] Rosenberg J., Schulzrind H., Camarillo G., *ET AL.*: 'Session Initiation Protocol (SIP) - RFC 3261 (Request for Comments (RFC) No. 3261)', Fremont, Internet Engineering Task Force (IETF), 2002
- [7] Curpen R., Ballan T., Sandu F., Costache C., Cerchez C.: 'Demonstrator for voice communication over LTE'. 2014 Tenth Int. Conf. on Communications (COMM), Bucharest, 29–31 May 2014, pp. 1–4
- [8] Cisco: 'Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018'. Available at http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html, accessed September 2014
- [9] Marcus S., Elixmann D.: 'The future of IP interconnection'. Final Report, European Commission, Brussels, EU-Commission, 2008
- [10] Stiller B., Reichl P., Gerke J., Hasan H., Flury P.: 'Charging and accounting in high-speed networks', *Future Gener. Comput. Syst.*, 2003, **19**, (1), pp. 101–109
- [11] Davoyan R., Effelsberg W.: 'Inter-carrier compensation between providers of different layers: advantages of transmission initiator determination'. Proc. Ninth Int. Conf. on Int. Federation for Information Processing (IFIP), TC6 Networking, Chennai, India, May 2010, pp. 373–381
- [12] Davoyan R., Effelsberg W.: 'A new bilateral arrangement between interconnected providers'. Proc. Sixth Int. Workshop on Internet Charging and QoS Technologies (ICQT): Network Economics for Next Generation Networks, Aachen, Germany, May 2009, pp. 85–96
- [13] Noam E.M.: 'Interconnection practices', in Cave M., Majumdar S., Vogeslang I. (Eds.): 'Handbook of telecommunications economics' (Elsevier, Amsterdam, Netherlands, 2000), vol. 1, pp. 387–409
- [14] Gizelis C., Vergados D.D.: 'A survey of pricing schemes in wireless networks', *IEEE Commun. Surv. Tutor.*, 2011, **13**, pp. 126–145
- [15] Blade V., Ghosal D.: 'Pricing approaches in the design of next-generation wireless networks: a review and a unified proposal', *IEEE Commun. Surv.*, 2007, **9**, (2), pp. 88–101
- [16] Faria F., Nogueira J.M.: 'Exploring context based charging for composite services'. Proc. Int. Conf. Intelligence in Next Generation Networks (ICIN), Berlin, Germany, October 2010, pp. 1–6
- [17] Dixit S., Periyalar S., Yanikomeroglu H.: 'Secondary user access in LTE architecture based on a base-station-centric framework with dynamic pricing', *IEEE Trans. Veh. Technol.*, 2013, **62**, (1), pp. 284–296
- [18] Wang J., Leung V.C.M.: 'Integrating predictable planning and unpredictable dynamics in network utility maximization: an improved quota-based market model'. Proc. Int. IEEE Symp. on Signal Processing an Information Technology, Vancouver, BC, Canada, August 2006, pp. 478–483
- [19] Bush J., Irvine J., Dunlop J.: 'Digital marketplace for tussle in the next generation wireless networks'. Proc. 70th IEEE Int. Conf. on Vehicular Technology (VTC), Anchorage, Alaska, USA, September 2009, pp. 1–5
- [20] D'Acquisto G., Cassara P., Alcuri L.: 'American options based service pricing for virtual operators'. Proc. IEEE Int. Workshops on Network Operations and Management Symp. (NOMS), Salvador, Bahia, Brazil, April 2008, pp. 197–184
- [21] GSM Association: 'Packet Voice Interconnection Service Schedule to AA.80', 2010
- [22] Mfupe L., Mzyece M., Kurien A.: 'A tele-economic approach for MNOs and software based M-VoIP operators to co-exist using game theory'. Proc. Int. Conf. Southern African Telecommunication Networks and Applications (SATNAC), Ezwilini, Swaziland, September 2009
- [23] Ac'ımovic-Raspopović V., Kostić-Ljubisavljević A., Mladenović S.: 'Competitive pricing using Cournot game in the next generation networks'. Proc. Int. Conf. Telecommunications in Modern Satellite Cable and Broadcasting Services (TELKIS), Nis, Serbia, October 2011, pp. 297–300
- [24] Iversen V.B.: 'Teletraffic engineering and network planning' (Technical University of Denmark, Lyngby, Denmark, 2010)
- [25] Iversen V.B.: 'On the accuracy in measurements of time intervals and traffic intensities with applications to teletraffic and simulation'. PhD thesis, IMSOR, Technical University of Denmark, Lyngby, Denmark, 1976
- [26] Palm C.: 'Accuracy of measurements in determining traffic volumes by the scanning method', *Tek. Medd. från K. Telegrafstyrelsen*, 1941, **7**, (9), pp. 97–115
- [27] Glabowski M.: 'Modelling systems with multiservice overflow erlang and engset traffic streams', *Int. J. Adv. Telecommun.*, 2008, **1**, (1), pp. 14–26
- [28] Riordan J.: 'Appendix 1: derivation of moments of overflow traffic', in Wilkinson R.: 'Theories of toll traffic engineering in the USA', *Bell Syst. Tech. J. (BSTJ)*, 1956, **40**, (4), pp. 507–514
- [29] Mandjes M., Saniee I., Stolyar A.: 'Load characterization and anomaly detection for voice over IP traffic', *IEEE Trans. Neural Netw.*, 2005, **16**, (5), pp. 1019–1026
- [30] Jain R.: 'The art of computer systems performance analysis: techniques for experimental designs, measurement, simulation, and modeling' (Wiley-Interscience, 1991, 1st edn.)
- [31] Student: 'The probable error of a mean', *Biometrika*, 1908, **6**, (1), pp. 1–25
- [32] Statsoft: 'Electronic Statistics Textbook'. Available at <http://www.statsoft.com/textbook/distribution-tables/#>, accessed November 2011
- [33] OPNET Technologies: 'OPNET Modeller 15.0'. Available at <http://www.opnet.com>, accessed November 2011
- [34] Cushnie J., Hutchison D., Oliver H.: 'Evolution of charging and billing models for GSM and future mobile internet services (External Publication Report No. HPL-IRI-2000-4)'. Available at <http://www.hpl.hp.com/techreports/2000/HPL-IRI-2000-4.pdf>, accessed December 2011