INTERNET TELEPHONY OVER WLANS

Eleftherios Dimitriou and Patrik Sörqvist, Global IP Sound

Abstract

This paper will address the issues that most seriously impede the use of IP telephony over WLANs, including spectrum congestion and interference. Other issues, such as the handling of VoIP security and the coexistence of RTP and TCP, will also be considered. Examples of the use of VoIP in different WLAN scenarios will be presented, and a solution that incorporates powerful error recovery and jitter buffering proposed.

INTRODUCTION

Voice over IP (VoIP) is becoming increasingly popular as the technology matures and companies discover the benefits of having a converged network for their data and voice applications. At the same time, wireless local area networks (WLANs) are being adopted to implement the "anywhere, anytime" concept in various places where people tend to utilize much of their time. Offices, airports and popular coffee houses are examples of such places. However, even though VoIP calls are possible today in a WLAN environment, a number of obstacles exist that negatively affect the use, and acceptance, of VoIP. These obstacles are primarily related to the decreased voice quality that can result from the fact that the characteristics of wire-line and wireless networks differ, as well as from congestion and the limited ability of a WLAN to support multimedia communications.

The Internet was originally designed to carry data-based traffic. For that purpose, "best-effort" service is often sufficient, since the network is used for delivery only, and it is up to the end applications to ensure correct delivery of the data packets. Retransmission is generally used when there is a problem with packet delivery. But as the Internet develops into a network carrying extensive multimedia traffic, problems appear. Real-time communications are especially time-sensitive, and the nature of wireless communications puts further constraints on real-time traffic. For example, the free, or unlicensed, nature of the widely used 2.4 GHz spectrum causes the already limited resources of a WLAN to be shared with other technologies, which can lead to significant levels of interference. Real-time communications are time-sensitive which makes retransmissions futile since voice quality can be severely degraded with increasing delay and jitter.

The benefits of VoIP outnumber the drawbacks, and that is reflected in the rapid deployment that it has experienced. New service providers are entering the market, offering an alternative to regular telephony, and VoIP solutions may revolutionize the way people interact, with WLANs offering a cheap and flexible way to increase the reach of the Internet. Video telephony calls and application sharing are just two usage scenarios for the future, and despite security concerns and confusion among standards, voice over WLANs is here to stay.

AN OVERVIEW OF VOICE OVER IP (VOIP)

VoIP started out as a cheap way of making voice calls over the Internet. Today, standards have been developed that enable Internet telephony calls to take place, as well as calls between regular PSTN (public switched telephone network) phones and an Internet voice application. Soon VoIP will be ready for mainstream deployment, and one of the main reasons for this success is, of course, cost-related. The cost for a company that needs to buy a PSTN PBX is significantly higher than buying VoIP-enabling technology that not only enables telephony, but also integrates voice and data into a single network. The operational and management costs are reduced, and the company stands ready for future ways of communication, such as video calls.

VoIP Network Structure

Delivering VoIP functionality requires a number of VoIP-enabled devices to be connected to the Internet. The picture below presents a simplified overview of the various components.

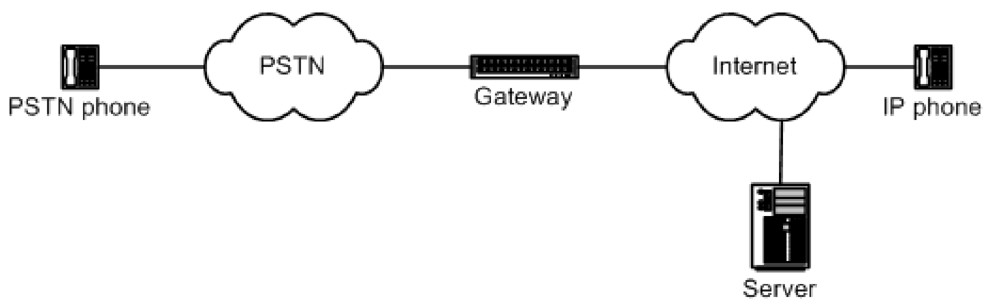

Figure 1: A VoIP network overview.

A VoIP call is initiated in one of the VoIP-enabled end-points. It can be a VoIP client that sets up a call to another user, or it can be a gateway that is trying to reach a VoIP client on behalf of a PSTN phone. VoIP connections between gateways are also possible as a way for two PSTN phones to communicate. VoIP calls can therefore either be between two VoIP clients, a VoIP client and PSTN or cellular phone, or between two PSTN/cellular-phones using the Internet for transport.

The gateway acts as a translator between networks with different signaling protocols. In the PSTN, SS7 is used to make phone calls, but VoIP can use H.323, SIP, or MEGACO, and it is the gateway that enables a proper call setup between two clients on distinct signaling networks. Furthermore, the VoIP equipment translates the information to the correct format. PSTN uses G.711 to encode the voice, while VoIP offers a vast variety of codecs for voice through a flexible call setup protocol. It is the VoIP equipment that, based on the information it gathers from the call setup procedure, decides to which codec the incoming voice should be transcoded. Gateways also exist between the PSTN and cellular network (for example, GSM networks), since the signaling and codec used differ.

To assist in the call setup procedure, a VoIP system needs a variety of servers that are responsible for different tasks. Registering the current location in the form of an IP address for each VoIP client can be the task of one server that is also the server from which VoIP clients request information about the location of other servers.

VoIP Call Setup Protocols

Today, there are two main protocols used for signaling in VoIP: H.323 and the Session Initiation Protocol (SIP). H.323 was the first protocol for commercial use, while SIP is the protocol with the largest momentum today. H.323 and SIP with assistant protocols provide analogous functionality. Another VoIP protocol is MEGACO, which is used for communication between gateways.

SIP defines packet exchange procedures for setting up, modifying and tearing down multimedia sessions. SIP was developed by the Internet Engineering Task Force (IETF) to provide a simple, scalable and easy-to-implement protocol from an IP perspective. It is based on other client-server protocols, such as HTTP and SMTP. SIP only deals with setting up a media session, and depends on other protocols to handle, for example, exchanges of capabilities.

Sessions are set up by having the client calling side send an Invite containing the session information to the server called side. The called side can deny, forward or accept the request. Once accepted, a media session is initiated until either side ends the call.

H.323 was developed by the International Telecommunications Union (ITU). It consists of an umbrella of protocols for dealing with everything related to setting up media connections, which has resulted in a complex protocol. Later versions have been scaled down to meet faster response times and provide less complexity.

Voice quality

The Internet has proven not to be the best carrier of real-time applications. Originally designed for data applications, it offers only one class of service: “best-effort.” The network forwards packets to their destinations without any guarantees. In PSTN, on the other hand, resources are reserved between the two end-points throughout the network to guarantee timely delivery of the voice stream. This leads us to some of the problems facing VoIP in the Internet today: delay, jitter and packet loss.

Figure 2: Illustration of packets traversing the network. Packet 3 was dropped, whilst the other packets arrived with variable delay.

Delay in the case of a real-time voice conversation is the time between the instant a person speaks a word or sentence and the instant the person on the other end hears it.

A number of tasks are performed on that information, all of which take time. The voice is packetized and processed at the sender and then transmitted over the network. On the network, a packet experiences propagation, queuing and other delays until it reaches the receiver, which needs to unpack, process and play out the voice data contained in the packet. The sum of all delays in one direction is not to exceed 150ms [ITU G.114] in order to have a normal conversation between two persons, as the risk of interrupting each other increases with higher delays.

Most delays are constant or vary very little, but as packets traverse the network, some points can be heavily loaded, which leads to a variation in the queuing delay. The variation of the total delay is called the jitter. This is a problem that does not exist in the PSTN because the voice stream has a dedicated line, and voice samples are transmitted continuously to the receiver, where they are play out. In VoIP, a number of voice samples are packetized and sent together in a packet. These packets arrive at uneven intervals, so there will be times when the receiver might not have anything to play out, which causes degradation in voice quality. This is dealt with by having jitter buffers, which save a number of packets at the receiving end before they are processed and played out. By doing this correctly, a continuous stream of voice can be played out without interruptions.

There are occasions when a packet arrives at a network node and the queue is full, causing the packet to be dropped. The network relies on higher protocols to notice this and retransmits that packet, as is done in the Transmission Control Protocol (TCP) that is used in the Internet. Packet loss can also occur due to overflow, CRC-errors, and configuration errors. The User Datagram Protocol (UDP) is used for packet transport in VoIP, since retransmission is not a solution due to the delay constraints involved. Dealing with the problem of lost packets is instead built-in at a higher level, in the voice codecs and the error concealment unit. By making the codecs robust to packet loss, better voice quality can be achieved independent of the network conditions.

There are various standardization efforts that have tried to offer means for providing quality of service (QoS) in the Internet. RSVP, MPLS and DiffServ are some IETF standards that have evolved. Most of these need changes in the current network nodes, which are costly and cumbersome. Also, it is not clear how these technologies will be used and what type of service can be offered or guaranteed. The flexibility achieved by running voice and data on the same network is significantly reduced by using such protocols as RSVP.

AN OVERVIEW OF WIRELESS LOCAL AREA NETWORKS (WLANS)

The WLAN is most likely going to be part of any wireless or mobile system, since it offers a cheap and easy way of providing lots of bandwidth wire-free. The impact can be seen already as offices, coffee shops, airports and other “hotspots” are having WLANs deployed to offer Internet access. People at home have also discovered how easy it is to get rid of all the cables and get the benefit of having Internet in all the rooms of their homes. By just connecting an access point to the Internet, one is ready to go wire-free, and this comes at a cheap price.

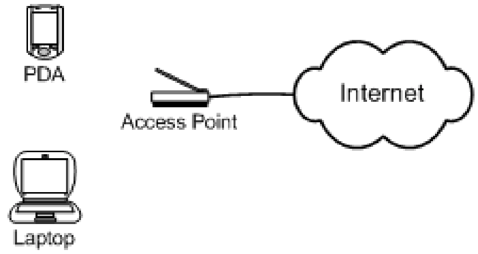

Figure 3: A simple WLAN setup.

It is easy to get lost in the standardization of WLANs. The IEEE has developed the most popular standards in their 802.11 family, and today the IEEE 802.11b standard has more than 80 percent of the WLAN market. It operates in the free, or unlicensed, 2.4 GHz frequency range, which makes it cheap. However, because the number of available channels is limited, it shares this spectrum with other wireless technologies, such as Bluetooth, cordless phones, microwave ovens and, of course, other WLANs.

An alternative solution is the IEEE 802.11a standard, which operates in the 5 GHz frequency band, where the frequency is not shared with a lot of other technologies. 802.11a offers higher bandwidth capabilities compared to 802.11b, but the price for 802.11a components is significantly higher, and the range of the access points is smaller, which makes it a costly alternative.

The success of 802.11b and the need of higher bandwidth have led to the development of 802.11g, which is backwards-compatible with 802.11b. It also operates in the 2.4 GHz frequency band and offers higher bandwidth (54 Mbps compared to a maximum of 11 Mbps for 802.11b), and offers more channels, thereby decreasing the interference challenges.

As the standard that is most widely spread, the success of 802.11b has also exposed its limitations, and interference from other technologies in the 2.4 GHz spectrum is but one of them. For example, an 802.11b access point cannot support the number of VoIP flows that correspond to the available bandwidth of 11 Mbps, even under perfect link conditions. Round-trip delay, jitter and packet loss increase to impractically high levels as the number of flows increases. Furthermore, since many access points do not support QoS, some flows are hurt more than others.

Congestion is another problem with 802.11b WLANs. The belief that there is an 11 Mbps bandwidth available can be misleading, and depending on the access point, the number of active users and the location of the user, the available bandwidth may actually be lower than 1 Mbps. All locations in a WLAN spot do not have the same link and signal quality, which also affects the VoIP connection. For example, moving around an office can impair the link quality as much as other wireless technologies (e.g., Bluetooth and wireless phones) and microwaves. However, since a WLAN is introduced to add mobility, one needs to manage when link quality deteriorates and recover from it when link quality improves. In the case of VoIP this can be done with better sound-processing solutions.

Solving the problem of handling security over WLANs is vital to successful deployment of the technology. Even WEP (Wired Equivalent Privacy), which supports different levels of encryption on the data transmitted over the wireless link, does not always result in secure communications. WEP only offers a simple way to stop eavesdropping and protection against external users, but does not really provide good security. Therefore, companies rely on separating the WLAN from the rest of the network using such tools as Bluesocket and other firewalls. With all of these security methodologies, users have to gain access to the network by registering and going through an authentication process. However, even though there are security concerns, WLANs are still being used and more and more people are adopting this technology.

Quality of service (QoS) is also an issue with WLANs. This is of importance, but as with the Internet, is a complex question, and solving QoS in WLANs first requires QoS to work in the Internet. The IEEE is working on two add-ons to the current standards to deal with security concerns (802.11i) and with QoS (802.11e).

EXAMPLES OF VOIP USED IN A WLAN

In a scenario that is common in small enterprises to increase mobility and productivity, an access point was installed on a 100 Mbps Ethernet local area network. This access point used the 802.11b standard with 64-bit WEP enabled for all traffic, and was similar to the ones implemented in various hotspots.

Previously, all communications between VoIP clients was over the Ethernet. In general, delay, jitter and packet loss levels were low over the Ethernet, even though occasional spikes did occur. In our tests, we were interested in comparing these numbers with the ones experienced by clients in a wire-free environment, and the tests were designed to examine the different network conditions during regular use of VoIP in a WLAN.

To achieve that, two computers were used transmitting continuous streams of data to each other. The following scenarios were investigated:

1. VoIP traffic during excellent RF signaling conditions

   With this test the maximum number of simultaneous VoIP calls was investigated for excellent link conditions.

2. VoIP traffic mixed with TCP

   The impact of TCP traffic in a WLAN environment on VoIP traffic was examined.

3. VoIP client moving around

   Since a WLAN adds mobility it is interesting to see how the network conditions change as a user moves around.

4. Effects of other technologies

   The problem of interference was tested with a simple scenario using a microwave.

All computers used for this test are Pentium III laptops, 600MHz, 256MB RAM, running Windows 2000 with standard 802.11b WLAN cards. The program used to generate and analyze TCP and UDP connections is the Global IP Sound IP WorkBench. All the results were verified by setting up VoIP calls and loading the network with traffic while having an ongoing VoIP conversation.

VoIP traffic during excellent RF signal conditions

An IP/UDP/RTP connection emulates a VoIP call using G.711 that is 80 kbps (16 bytes header + 64 bytes payload) and a TCP connection tries to keep a bandwidth of 100 kbps and packet size of 576 bytes. Furthermore, the tests were also conducted by emulating G.729 traffic (16 bytes header + 8 bytes payload) and similar congestion levels were experienced, which let us assume that the limitation might lie on the number of packets an access point can process

The graphs below present the delay, jitter and packet loss levels for different numbers of VoIP connections.

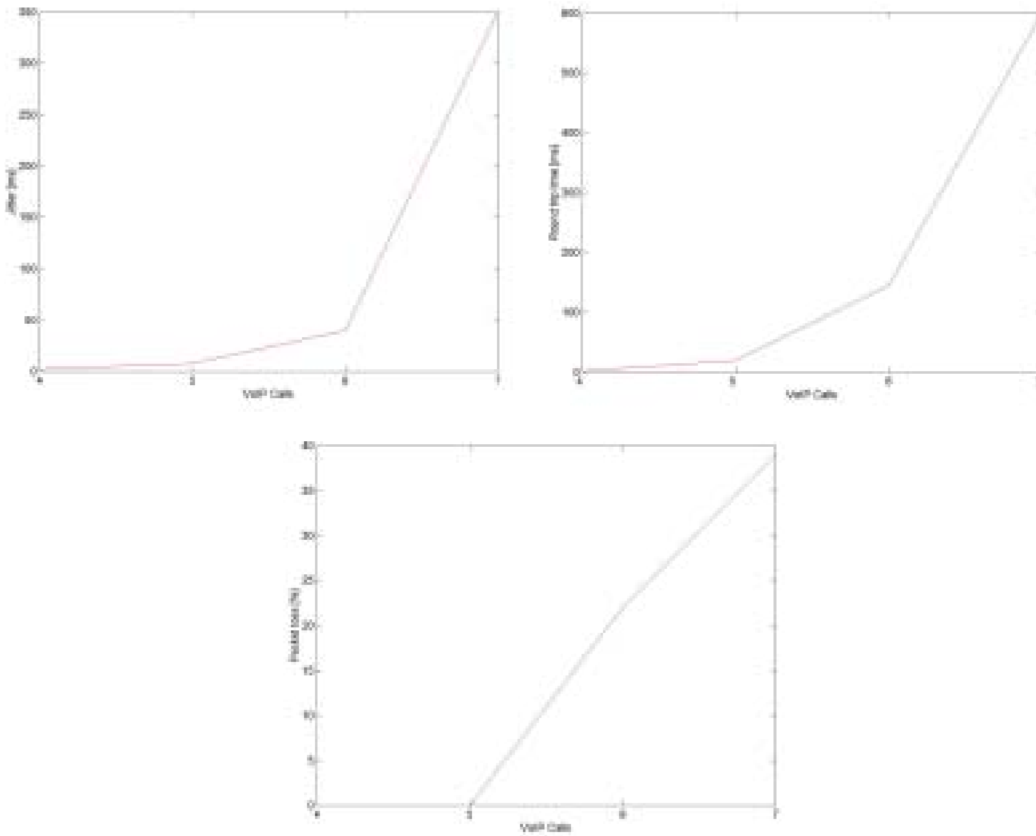

Figure 4: The graphs illustrate the delay, jitter and packet loss levels for 4-7 VoIP calls corresponding to 640-1120 kbps.

It can be seen that after 5 simultaneous VoIP calls, the quality deteriorates significantly even for the perfect link condition. If the link condition is not excellent, the number of simultaneous calls are 3 for good, and not even 1 for bad, link conditions. If one of the users is in an area with a good link condition, that number is lower since the overall achieved end-to-end link condition is worse. Note that when the link quality is bad, a normal conversation cannot be held.

VoIP traffic mixed with TCP

Since no QoS mechanism is in place to protect the real-time VoIP traffic, it is of interest to study the effect of TCP traffic on the network. For that purpose, one VoIP call is established and the TCP traffic is increased. The pictures below show the effect of the TCP traffic on the VoIP call. Bear in mind that with the Global IP Sound IP WorkBench, the TCP connections are aggressive and, if the predefined bandwidth level cannot be sustained, they restart.

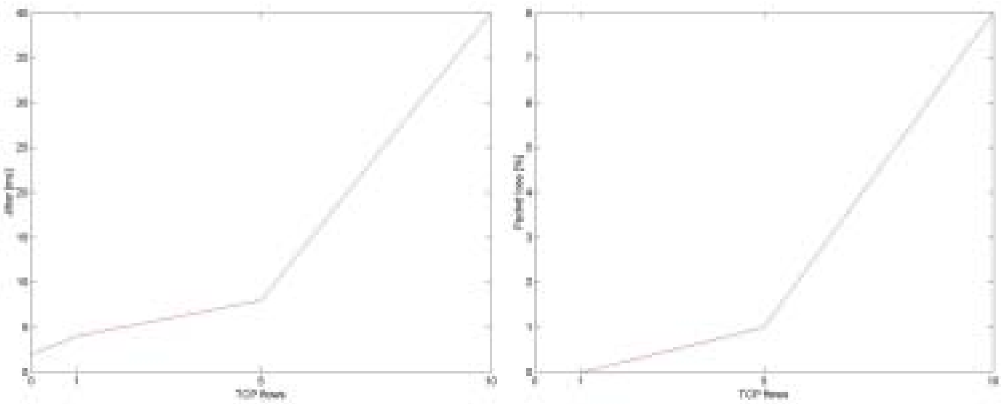

Figure 5: The graphs present the jitter and packet loss levels for the VoIP connection, while 0-10 TCP flows corresponding to 0-1Mbps are active.

It can be seen that the TCP traffic does not pose a serious threat to VoIP at this point. Some degradation of quality can be seen, but it is not one that lasts a long period of time. The effect of VoIP traffic on TCP is not investigated in this paper. TCP will back off and not get any packets through e.g. slow surfing.

VoIP client moving around

Another interesting test involves the effect of a VoIP client moving around in the office because mobility is a large part of the purpose of a WLAN. What we have in mind is a PDA with a built-in WLAN card that can be used as a cordless phone. Carrying a laptop around the office might happen occasionally as well. The pictures below show the distance that was covered, at a pace of around 1meter/second, and the jitter variation during the walk.

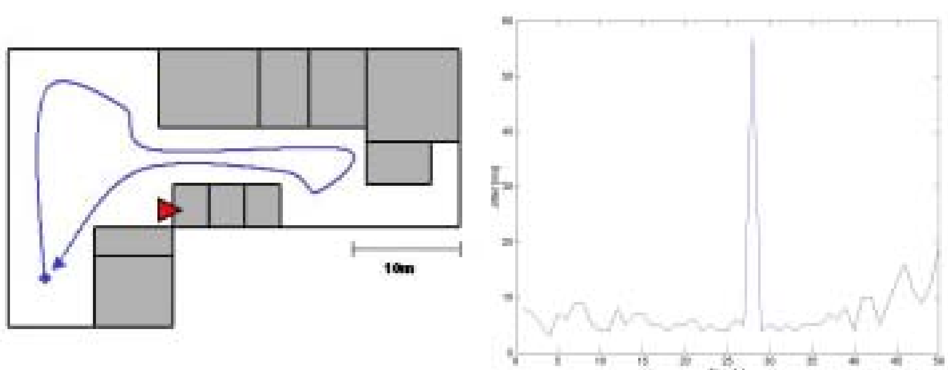

Figure 6: The left picture is a model of the office. The gray areas are offices and the red triangle is the access point location. To the right, one can see the jitter graph, in which each second corresponds to approximately a meter distance.

By moving around, the jitter level that is received can clearly vary. The factor that affects the output the most seems to be the link quality. A slight effect can be seen as a result of the movement, but moving from a good to a bad spot makes the biggest difference. Therefore, recovering from spots with bad link quality is important.

Effects of other technologies

In this test, the computer was put close to a microwave oven. A VoIP call was initiated and after 30 seconds the microwave was turned on. The picture below shows the effect of that on the experienced jitter, but packet loss levels of around 5 percent were experienced as well.

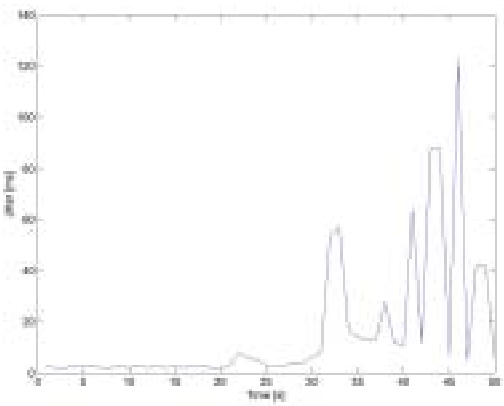

Figure 7: Jitter caused by starting a microwave after 30seconds into a VoIP connection.

The interference caused by other technologies is clearly visible. It can also be heard. The positive fact of having a free spectrum can also be negative as more and more technologies compete. It is also worth noticing is that when a lot of WLANs are close together, for example, at a tradeshow, the interference tends to increase as the number of available channels is filled up.

CONCLUSIONS

By conducting a series of simple tests in a standard network combining wired and wireless components, it is clear that there are a number of problems facing VoIP in the wireless environment.

The problem of not having enough bandwidth can be solved with new technologies, and interference can be avoided by using another frequency. However, not all current WLANs are going to be upgraded, and changing frequencies is expensive, and alternative frequencies have not proven to be as popular as 2.4 GHz.

Voice applications in WLANs need to recover from very bad link conditions caused by interference or a user moving out of range of the access point. If the user moves out of range and never returns, the call will ultimately have to be disconnected. However, if the user realizes the problem and moves to a spot with good signal quality, it is desirable that the voice application be robust enough to manage the short disruption. This requires a very good jitter buffer that will be able to adapt rapidly to the changing conditions. The same applies to the error concealment unit that is providing the voice estimate when no packets are available; when it comes to high packet loss levels, robust codecs are needed.

One attractive answer is to solve the problem in the speech processing design by embedding powerful sound processing software, such as Global IP Sound's NetEq and Enhanced G.711. By making the voice application more robust, voice quality can be sustained and VoIP can be used in a satisfactory way even in congested networks. This makes for a network-independent solution that can be used in all WLANs to enhance voice quality.

By using NetEq and Enhanced G.711, developed by Global IP Sound, to conduct the voice-quality tests, the number of supported calls during excellent link conditions increased 33 percent. Instead of 5, it was shown that 6 simultaneous calls could be supported with no degradation in voice quality. This is a significant increase since it increases the possible users by around 30 percent[*].

---

[*] Using Engset calculator with blocking probability 0.1, given 16 users for 5 lines and 21 users with 6.