**Dan KOMOSNY**[1], **Jiri BALEJ**[1], **Hira SATHU**[2], **Ranjana SHUKLA**[2]
**Pavel DOLEZEL**[1], **Milan SIMEK**[1]

Department of Telecommunications, Brno University of Technology, Brno, Czech Republic (1), Department of Computing, Unitec Institute of Technology, Auckland, New Zealand (2)

# Cable Length Based Geolocalisation

***Abstract.*** *This paper discusses a new IP Geolocalisation technique after introducing other contemporary techniques, starting with the earliest known GPS systems for geolocation. The work addresses a key-feature of any constraint-based geolocalisation method, which is finding the maximum linear distance of a target with the unknown position to a set of landmarks with the known geographical position. We define the maximum linear distance using the physical cable length in the interconnecting networks. The proposed method CLBG (Cable Length Based Geolocalisation) first estimates the physical cable length between the target node and the landmark and this in turn is used to compute the linear distance. The computed linear distance is consequently used for the target node localization using multilateration. The CLBG methods outperforms other similar techniques when compared in a both national and global IP networks.*

**Streszczenie.** Niniejsza praca omawia nowe techniki geolokalizacji IP na bazie innych technik, począwszy od najwcześniejszych znanych systemów GPS dla geolokalizacji. Praca dotyczy kluczowej cechy wszystkich metod geolokalizacji, którą jest znalezienie maksymalnej odległości liniowej pomiędzy celem o położeniu nieznanym a zestawem punktów orientacyjnych o położeniu znanym. Maksymalną odległość liniową określamy za pomocą długości fizycznej kabla w sieciach połączonych. Proponowana metoda CLBG (Cable Length Based Geolocation) szacuje długość fiyczną kabla między węzłem docelowym i punktem orientacyjnym. To z kolei wykorzystuje się do obliczenia odległości liniowej. Następnie obliczona odległość liniowa użyta jest do lokalizacji węzła docelowego za pomocą multilateracji. Metody CLBG przewyższają inne podobne techniki geolokalizacji zarówno w sieciach IP krajowych jak i globalnych.

**Keywords:** IP, Geographic, Localisation, Multilateration, Latency, Cable
**Słowa kluczowe:** IP, geograficzny, lokalizacja, multilateracja, opóźnienie, kabel

## Introduction

The benefits of locating one's own position and those of others has applications in the military, business and the community at large. To estimate physical position of a target, one needs to know either the distances or angles to objects with a known position. Multilateration is used when working with distances, while multitriangulation is used when working with angles. Multilateration is a common methodology used by most systems. Over time this extended to popular generalized use for determining a location like a street address rather than the geographic latitude and longitude. Most of these systems utilize mapping displays. Of all the geolocation systems, the most popular Global Positioning System (GPS) ant other are discussed below.

### Geolocalisation using GPS and other systems

GPS system is a space-based global positioning system that gives a user fairly accurate location and time information covering most weather conditions, as long as a line of sight path is present to 4 or more GPS satellites. These GPS services are freely available to any person with a GPS receiver device. Originally the US Department of Defence initiated the GPS system with 24 satellites. Increase in the satellite number has lead to increased location coverage with greater accuracy. Other GPS systems are the Russian GLObal NAvigation Satellite System (GLONASS) used by the Russian military till 2007. Galileo and the Compass navigation systems are the other European and the Chinese initiatives. GPS may not work well indoors or in types of terrain which do not permit line of sight path between satellites and the GPS enabled device. GPS feature is already bundled with majority of smart mobile phones.

For general wireless technologies, direction finding equipment along with triangulation of the bearings were used, for geolocation. For locating mobile host Place Lab [21], Cricket [20] and RADAR [2] could be used with 802.11 based access points and GSM beacons with known locations. These systems have limited coverage (coverage area of access points and cell site coverage). To increase the coverage, it would require wide-spread and dense deployment of access points or GSM hardware, with known geographical locations. Use of IP Geolocalisation software for locating smartphones already exist. TruePosition LOCINT [22] helps not only locating smartphones in active but also in passive mode (not making a call).

The most common of geolocation applications is its mapping capability for reaching a given location, while giving the user his/her own position. This can be used in devices that may also be space borne. However, license restrictions do exist for receivers using this above a given height from the earth's surface to preclude usage of GPS for military applications by other than licensed Defence users. In addition to military applications, there are various civil applications of this technology - persons entrusted to reacting and reaching to emergencies within the shortest possible time in their areas of responsibility. These vary from search & rescue, medical, police and various other commercial applications including the public utility services.

In spite of the many benefits of geolocation using the above discussed technologies, they do have a few constraints. The most significant of these are the non availability of a GPS receiver in computing devices to help in their location. In addition there is also the inability of a GPS enabled device to communicate from locations that do not have line of sight communication with any or sufficient number of GPS satellites. Other mentioned systems have limited coverage and, therefore, they cannot be used globally. The IP Geolocalisation technology discussed in the next subsection and beyond, tries to overcome the above constraints.

### IP Geolocalisation Background

The wide prevalence of IP computing nodes is what makes IP Geolocalisation technology interesting for further research. The nature of these nodes may vary widely. These could be a small web enabled mobile phone, PDA/smart phone, wireless/desktop PC, server or even a main frame computer. IP Geolocalisation is the technology that helps to identify the real world geographical location of a computer node by use of the Internet Protocol (IP) address since it is unique and because most of these devices may not have GPS devices embedded in them. In addition, most of the time, these devices may not have communication access to GPS satellites and they are unsuitable for use by RF direction finding technologies. However, the near universal use of Internet Protocol by these devices makes them suitable for use by IP Geolocalisation techniques discussed in this paper. Large number of companies like Akamai, Digital Envoy, MaxMind, Quova maintain databases that map IP addresses to geographical locations [16]. However, reliability of information is also dependent on the database's being up-to-date and this issue needs to be kept in mind. This would be dependent on the concerned organisation maintaining the database and some guarantees being in place for an appropriate record for each IP address in the database.

**Application areas and the benefits of IP Geolocalisation**

As with any research, it is essential to know the application areas and the benefits that accrue from the research. In case of Geolocalisation in IP networks this potential is significant. This section covers a few of the application areas and the benefits that are realised through IP Geolocalisation. Among the most popular in the commercial area is targeted advertising. An example is a host website identifying the geographical location of a visiting client IP node and placing of advertisements that are relevant to the geographical location of the client. Automated redirection of clients to nearby servers is another area that minimizes internet traffic on the trunks and optimizes user experience too. Locating of nearby social events (movies, concerts, sports etc.) is a useful application for both commercial and community service reasons. These could cover the choice of language, local weather forecast and a range of other locality dependent web content. Display of web content, appropriate to the client's geographical region is the general trend in these applications. Likewise, information or knowledge of the origin of chat/message is being sought by responsible and well informed users of computers. On the other hand, Gueye et al. [8] write in their paper that privacy issues may arise due to development in IP Geolocalisation technology. Geographic Location/Privacy (geopriv) [3], a part of the IETF, is developing policies to control the exchange of geolocation information with privacy in mind.

Security applications are yet another area of interest. Current changes in information technology have seen some initiatives that go towards improved use of IT infrastructure. Few examples of such initiatives are cloud computing, e-commerce, and e-learning. However, these initiatives do bring with them issues such as cyber crime, copyright violation and/or requirement of a need to establish the area (person or organisation) where the crime is committed. In most of the cases existence of international law would help to prosecute the guilty party provided one knows the location of the IP device user. Reducing of credit card and identity theft fraud is a positive contribution of IP Geolocalisation.

IP Geolocalisation software is already being used to restrict specific information being distributed only to certain regions/countries. Censorship of data being downloaded by a client, depending upon the geographical location of the client node can be adopted to enforce regional censorship standards.

While many applications benefit from IP Geolocalisation, there are parties who subvert the IP Geolocalisation process. Few examples of these could be where a client located in an area prohibited from receiving a specific content, is returning incorrect results of its geographical location. Another example could be where a cloud service provider may provide false information for the location of its servers that are in fact not located as covered in the service level agreement but where it is cheaper to host them but not as secure from the clients point of view. Between the network topology aware and delay based IP Geolocalisation techniques, the later techniques fare better against subversion of the location information since they provide lesser time to the parties subverting the IP Geolocalisation process [11].

**Paper contribution and organization**

The paper address accuracy of IP Geolocalisation methods based on latency measurement. We focus on constraint-based method, SOI (Speed of Internet), where a constraint is estimated. Usually the upper bound (the maximum possible distance between nodes) is estimated. When the maximum distances between the target with an unknown location and landmarks with a known geographic location are known, a multilateration algorithm is used for finding an area where the target is located. Our work deals with the maximum distance estimation between any two nodes in the Internet. The maximum distance is usually estimated by the RTT (Round Trip Time) measurement between a target and the respective landmarks. The RTT measurements can be performed from the landmarks or from the target node.

The contribution of the paper lies in identifying and involving a new element in the maximum distance estimation using RTT measurement. From the measured latency, the physical length of cables between nodes can be calculated; provided we know the speed of signal propagation and the delay caused by intermediate devices on the path. The speed of signal propagation was investigated in previous works (propagation speed can not be under 0.5 ms per 100 km). For our research, we used CESNET2 network which connects academic institutions and universities in the Czech Republic and global research network PlanetLab. Hence from the CESNET2 network documentation, we derived the ratio of physical cable length to linear distance between Czech cities. We measured the latency between nodes in the cities connected by this network, and by using the known physical length of cables in this network, we also derived latency per hop on the path. We compared our defined latency per hop with previous works to verify its credibility. Knowing the number of hops and latency each hop causes, we can calculate the cable length between any pair of nodes in any network. Consequently, with the known ratio of cable length to linear distance, we are able to estimate the maximum possible distance between any two nodes.

The results of this study were verified in the global research network PlanetLab. Advantage of accessing the PlanetLab nodes spread over the whole of Europe, was taken to implement the proposed method, which we call CLBG (Cable Length Based Geolocalisation). In order to compare the proposed CLBG method with others, we implemented three other approaches to IP Geolocalisation. The results show that our method outperforms other selected methods in terms of more accurate geographical position estimation. Finally, as an example of use, we described a scenario of finding a position of a target using the CLBG method with an output showing a geographical position on a map.

The balance of this paper has been organized to cover the material as follows: Section 2 covers the related works including the methods for IP Geolocalisation along with a review of the literature. Section 3 covers the investigation of delay sources in IP networks. Section 4 covers the findings and the result of the research. In section 5, we compare the CLBG method with other IP Geolocalisation methods and discuss the results. Section 6 goes on to conclude this paper, recommending further research that could be undertaken in the area of IP Geolocalisation.

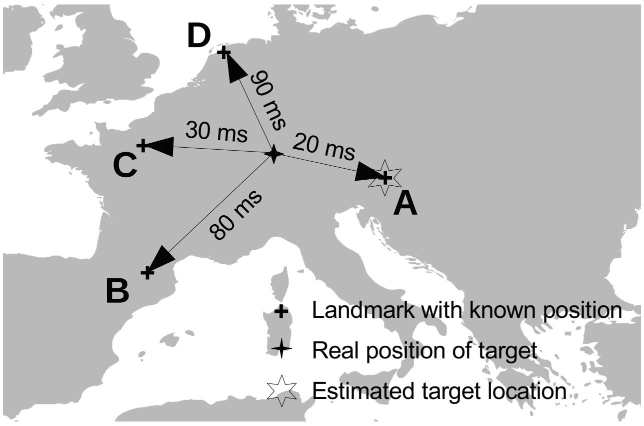

Fig. 1. Example of Shortest Ping method

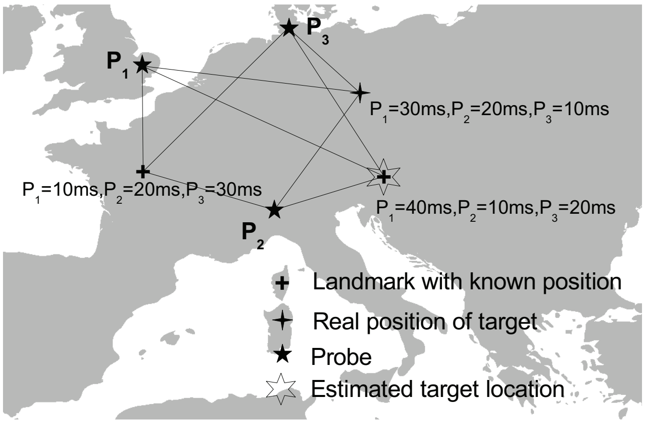

Fig. 2. Example of GeoPing method; euclidean distance is used to identify delay vectors similarities

## Related work

In this section authors have discussed different methodologies for IP Geolocalisation as covered below. They have sorted the methodologies in order from simple to more complex. However, this ordering is rather loose, mainly in the context of methods mentioned at the end of the list. In addition the method complexity does not say anything about its performance. All the methods in this section work with communication latency measurement. We do not mention particular methods based on public or private geolocalisation databases or methods based on the use of the DNS system. We focus on latency based methods since the contribution of this paper falls within this area.

Geolocalisation method Shortest Ping was described in [17] as the simplest delay-based technique. Shortest Ping requires initiating delay probes from each of the landmarks with a known position to the target with an unknown position. This technique maps each target to the landmark's position that is closest to it by measuring RTT (Fig 1).

Padmanabhan and Subramanian [17] presented a collection of three techniques for deducing geographic location from IP addresses of an Internet Host. Geo Track, using the DNS names of the target host to infer the location. GeoCluster uses the combination of rough/estimated host-to-location mapping and BGP prefix information to locate the host. GeoPing uses delay measurements from different locations to compute the coordinates of the target. It estimates the position of a target by assigning it to the most representative landmark with the known position using latency vector similarity. In this methodology the assumption is made that landmarks/targets that are close to each other have similar network delays with respect to other landmarks. The target is mapped to the landmark with the closest delay vector profile. Euclidean distance between delay vectors is used to calculate the similarities. Delay vectors are built by probing targets and landmarks from a set of probes which are capable of delay measurement. Fig 2 shows the geolocalisation process. For the sake of simplicity, the number of landmarks in the figure is lower than the number of probes. The accuracy of GeoPing is restricted by the distance of the target to the nearest landmark.

Bassett et al. [11] has described Speed of Internet (SOI) geolocation delay-based technique. This technique combines distance constraints from multiple landmarks to arrive at a final estimate. The authors consider the fact, that information travels on transmission lines (fibers) at speed 2/3 the speed of light in vacuum (1 ms per 100 km). This creates an upper limit of distance of the communication nodes (constrain). SOI uses a single latency to distance conversion factor across all landmarks. It employs a multilateration method to get the target position.

The CBG method [8] relates delay to distance dynamically. Each landmark measures the delay from itself to all other landmarks periodically to self-calibrate the relationship between latency and distance.

National Security Agency (NSA) Patent uses the time latency of communication in the network to determine the geographical location of the target [10]. A "Network Latency Topology" map is prepared by using latency from various nodes to known nodes. For the target to be geolocated, the minimum latency from it to a node is measured. These measurements are then co-related with the "Network Latency Topology" map to determine the specific geolocation of the target.

Laki et al. [14] describe geolocation technique called Spotter, which uses a probabilistic approach. In this technique analysis of relationship between geographical distance and network delay is performed. The results of the analysis show that the distribution of spatial distances for a given delay follows a universal distribution. It is also independent of the landmark's position from where the measurement was performed. For determining internal model, Spotter uses separate calibration data for each landmark. All the data collected is used for universal delay distance model.

Arif et al. [1] in their paper write about GeoWeight algorithm. With this technique the maximum and minimum distance bounds are divided into different weighted regions. GeoWeight algorithm computes a weight for an intersecting region as the sum of weights of overlapping regions enclosed in the intersection. The location of the target is chosen as the centroid of the intersection region having the highest computed weight. This algorithm assigns weights to sub-regions within the larger region. This is able to constrain the target location to a smaller region.

Eriksson et al. [7] use a methodology that considers a set of measurements from a set of known monitors to a target. This is then used to classify the location of that target based on the most probable geographic region given probability densities learned from a training set. The authors have used the Naive Bayes framework that has low computational complexity and allows additional information to be incorporated to improve process.

Bassett et al. have suggested Topology-Based Geolocation (TBG). This uses delay based measurement by taking into account the logical network topology [11]. In the TBG approach, first traceroute utility is used to measure the delay. Later authors have used algorithms that could combine delay with network topology to reduce error due to the multiple network paths over the Internet. The measurements are converted to a set of constraints for the unknown locations. Then the target node and the routers enroute are geolocated in a way that best satisfies the constraints.

Wong et al. [25] in their Octant approach discuss how Internet hosts are localised based purely on network measurements. The method used consists of error minimising constraint satisfaction. This leads to estimated region in which the target resides. These methods create a sort of cylinder around each landmark. The target is then located inside this cylinder.

### Latency in IP networks

Generally, the communication latency is the time needed for a data packet to be sent by a node, travel and be received by another node. Communication latency is an important property for delay-sensitise applications in IP networks such us Voice and Video over IP [23] [13].

Usually round trip time (RTT) is measured since it avoids the need for time synchronization between nodes. RTT is the time for sending a packet to the remote node, its processing by the remote node, and sending the reply by the remote node to the originating node. Provided that the ongoing route from the source to the destination is the same as the backward route, we can define the one-way delay between the two nodes as half of the measured RTT (packet processing time at the remove node is omitted). However, routing asymmetry in the Internet can significantly affect one-way delay measurement using RTT. Routing asymmetry can be caused by many factors, such as traffic engineering. On the other hand, routing asymmetry does not necessary causes a difference between delays on the forward and reverse paths. The asymmetries of the ongoing and backward paths in the Internet have been studied in several papers. Authors in [15] showed that about 80 % of route pairs can be considered symmetric. For transatlantic connections, 93 % of the router pairs are symmetric. Authors of [9] found that routing asymmetry can be more significant. They involved academic networks and commercial networks in their study. They also considered routing asymmetry at AS (Autonomous System) level and router level. The result is that routing asymmetry is presented in both networks. Academic networks seems to be more symmetric than commercial networks. Routing asymmetry at the AS level is 14-65 %. 10 % of the routes in commecrial networks has asymetry larger then 0.5 and 10 % of the routes in academic networks has asymetry larger then 0.1. Another issue dealing with RTT measurement when used for IP Geolocalisation is that communication delay breaches triangle inequity [24].

RTT consists of particular latencies originated at the elements on the path. The particular latencies originated on the communication path have different impact on the overall delay. Some of them have a constant value and some of them vary a lot depending on actual network conditions. Latency can be divided into two parts – deterministic and stochastic. The deterministic delay has a constant size and it can be theoretically calculated. It is the minimal time required for transmission of information on a particular path. In contrast, the stochastic delay is a random value and it is influenced by network conditions, e.g. actual load of intermediate network devices.

The intermediate device latency consists of delay in the input queue, delay required for forwarding a data unit from input to output and the time in the output queue. The latency in the input queue is affected by the throughput of incoming line and the size of data unit. More accurately, it is the time between receiving the first and last bit of message. For switches, which work in cut-through or fragment free mode, it is the time necessary for receiving the first 14 or 64 B. The time consumed in output queue depends on the number of data units directed on the same port. Forwarding the information from input to output includes the processing data units through the relevant OSI layers used for communication by a particular device. Common latency (without load) for devices working on OSI layer 2 is 1-10 ms, for devices working on OSI layer 3 the latency is about 10-100 ms. If a device overloads, the stochastic part of latency may be much greater than the deterministic part. The overall communication latency is therefore mainly affected by the intermediate devices.

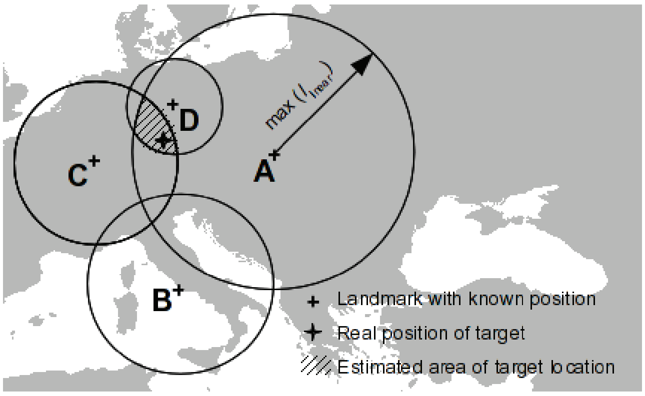

Fig. 3. Example of target location using maximum distance constraint; intersection of circles A, C, D gives estimated area of the target position

### Proposal of distance estimation method for constraint-based IP Geolocalisation

#### Problem statement

The current methods for IP Geolocalisation use different techniques for distance estimation between a pair of nodes. As described before, some constraint-based methods work with the maximum possible distance from landmarks with a known geographical location to a target with an unknown location. The two most representative method for the maximum possible distance estimation between a pair of nodes is SOI. It is simple and performs well when compared to similar methods using the maximum possible distance constrain [11]. SOI uses the speed of information that can travel via optical cables, which is the most common medium in transport lines. Its propagation speed is approximately 194 895 km/s, i.e. 0.65 times the speed of light [18]. The metallic cable has signal propagation speed about 0.75 times the speed of light.

The localization process using the maximum distance from landmarks with known location to a target is depicted in Fig 3. In this case, four landmarks are used. We label the maximum linear distance from a landmark as max ($l_{\text{linear}}$). The great-circles given by the maximum possible distance around landmarks A, C and D estimate an area where the target can be located. The great-circle around landmark B shows the case where the conversion of latency to maximum distance failed. This brings us to the problem description. With an underestimated maximum distance from a landmark, the localization process of a target can fail. With an overestimated maximum distance from a landmark, the localization accuracy decreases since a larger area of the target location is possible. The aim is to find a solution where the estimated area of target position is small enough to achieve a good location accuracy but we still get and estimated location area to have a possibility to locate the target at all (if only a few or none circles intersects).

Our idea is to develop a system that would initially detect the physical network topology and use it in the process of maximum distance estimation using latency measurement. The proposed method is used to estimate the physical cable length and this in turn is used to compute the linear distance of the target node from a source communicating node. For the experiments related to this research, we used a Czech national academic network CESNET2, which is described in the following section.

### CESNET2 national academical network and experiment background

Cesnet is an association of universities in the Czech Republic and the Czech Academy of Sciences. Its purpose is to develop and maintain the academic backbone network called CESNET2. Due to the strong research background of this network, it is well documented including information about its topology and the physical layout of transmission cables. For our experiment purposes, we found out the physical length of cables between the Czech university cities. The network connects the university campuses with only high speed data lines. The network topology is composed of small rings passing through a few cities. The backbone is formed by optical fibres using DWDM (Dense Wavelength Division Multiplexing) and IP/MPLS (Internet Protocol /Multiprotocol Label Switching). The domain names (or IP addresses) of the servers connected to CESNET2 network were found on the web pages of organizations in the Cesnet association (chosen by the organization's home city). If any address did not reply to an ICMP request, we found address of another server in the organization. We created a set of 35 hosts connected to CESNET2 network.

### Communication latency per hop

Our attempt was to estimate the ratio of physical cable length to linear distance in CESNET2 network. For that purpose, we also had to know the latency per hop on the source to destination path. We tried to leave out the stochastic part of delay, so we used the minimal measured RTT. The result then presented the value of latency for one L3 device with minimal load. This value was also influenced by the latency caused by devices on the lower layers (L2 and L1), but these equipments usually have very low latency, in comparison with L3 devices. The measurement was done from two nodes located in Prague and Brno, Czech Republic. RTT was measured, using the ping tool, which uses two ICMP (Internet Control Message Protocol) messages – Echo request and Echo reply. The measurement was done according to recommendation RFC 2544. It was repeated for five times (on different days and hours), the Echo request message was sent every second for 120 seconds to each address. Delay per one hop dhop (one intermediate L3 device) can be identified by

$$d_{\text{hop}} = d_{\text{path}} - \frac{l_{\text{phy}} \times \frac{1}{c_{\text{ew}}}}{n-1}. \tag{1}$$

In order to know the delay between nodes (cities) $d_{\text{path}}$, we divided in half the measured RTT values. We omitted the routing asymmetry and breach of triangle inequity. The physical length of cable ($l_{\text{phy}}$) was found in the CESNET2 documentation. The constant speed of propagation on a kilometre $c_{\text{ew}}$ is 5.13 ns/km (optical fibre). Using traceroute we got the number of L3 intermediate devices $n$. The physical cable length in source and destination network was neglected. The packet processing time at remote host was also ignored.

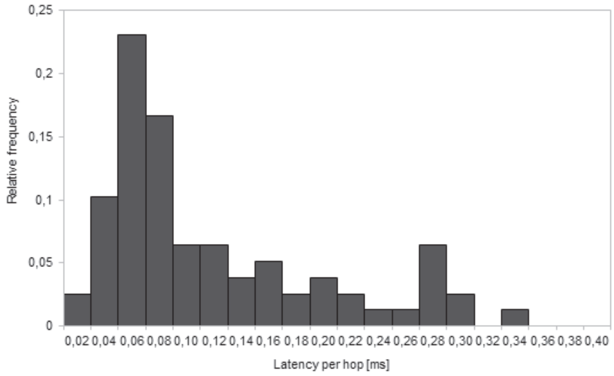

Fig. 4. Distribution of delay per hop , each bar is 0,02 ms wide

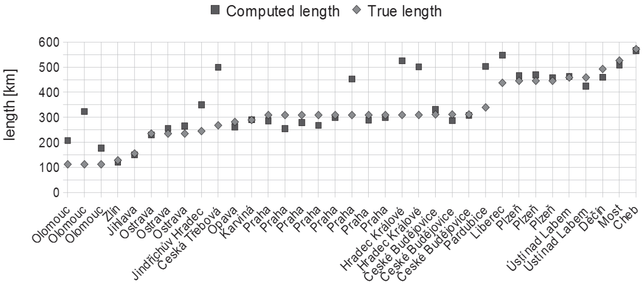

Fig. 5. Comparison of computed and physical cable length

The histogram in Fig 4 shows the distribution of calculated delays per hop. It can be summarized that $d_{\text{hop}}$ is lower than 100 µs in 60 % cases. The minimum for 87 % samples is 40 µs. The 40 µs equals 8 km distance. The average value for all calculated delays per hop is 122 µs, more significant is the median which is 76.5 µs. Result presented in [5] is similar (mean value 101 µs).

The comparison of physical and computed distances between selected cities in the Czech republic is plotted in Fig 5. The computed length $l_{\text{comp}}$ was calculated using

$$l_{\text{comp}} = d_{\text{path}} - (n-1) \times d_{\text{hop}} \times c_{\text{ew}}. \tag{2}$$

The physical distances were obtained from the CESNET2 network documentation. In majority cases, the computed length is by and large equal to the physical cable length keeping in view the the large distances involved. For simplicity, we will suppose that the computed cable length is equal to the physical cable length in the rest of the paper.

### Cable length

As stated before, the idea is to incorporate the estimated physical length of network cables in maximum distance estimation between nodes using latency measurement. For this purpose, we identified the ratio of physical cable length to linear distance $c_{\text{path}}$. The calculated ratio was on average 1.99, the median was 1.92 and the minimal ratio achieved was 1.56. Assuming that the computed cable length is equal to physical cable length, we can conclude that the linear geographic distance is approximately half of the length of the computed cable. Fig 6 shows a line that represents the relation of the linear distance and the physical cable length.

The maximum linear distance $l_{\text{linear}}$ between nodes is identified by

$$l_{\text{linear}} = l_{\text{comp}} \times c_{\text{path}}. \tag{3}$$

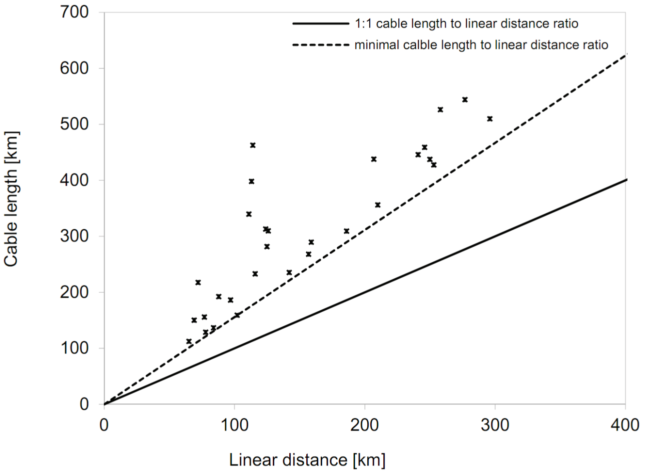

Fig. 6. Relation between straight distance and length of cable

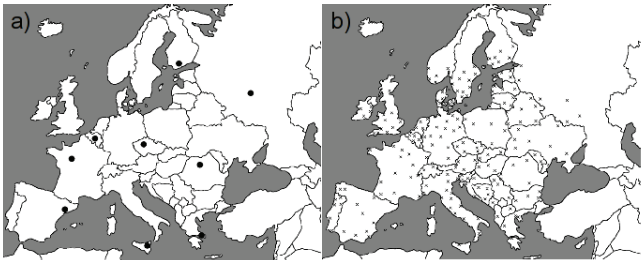

Fig. 7. Placement of nodes used in evaluation of GeoPing - a) probes, b) landmarks; adopted from [6]

The ratio $c_{path}$ used is 1.56 which is the minimal value found. Latency value per hop $d_{hop}$ needed for identifying the computed cable length $l_{comp}$ using equation (2) is the minimal delay for at least 87 % samples (0,04 μs).

**Accuracy evaluation**

This section covers accuracy evaluation of the proposed CLBG method in a cross-border environment. The basic evaluation in a national environment, in a Czech national network CESNET2, is presented in [12]. We compared the CLBG method with SOI, GeoPing and Shortest Ping.

For the purpose of evaluation, we used global research network PlanetLab [19], [4]. It consists of approximately 1000 independent nodes that are placed in various parts of the world (around 500 sites).

We created several datasets consisting of European nodes. Nodes, which were required to be accessible in terms of running scripts (i.e. probes capable of delay measurement), belonged to PlanetLab. Other nodes, acting as passive landmarks, were PlanetLab or ordinary servers. Fig 7 shows a dataset used for evaluation of the GeoPing method [6].

Fig 8 shows the cumulative distribution function of error distance estimation. The CLBG method gives better results in the whole range of distance estimation error. We can conclude that about 90 % of the performed measurements estimated target position within the 310 km distance from its real geographical position. 50 % of the estimations fell in 190 km distance error. 20 % of estimations produces error within 80 km. SOI gives similar, but larger distance errors (20 % of estimations produced error within 130 km, 50 % of samples fell within 210 km error and 90 % of samples fell within 500 km error).

The accuracy evaluation summary is shown in Table 1 which shows the average and the median distance estimation error. Furthermore, it shows the average area radius and the estimated area surface of target location for constraint-based methods, SOI and CLBG.

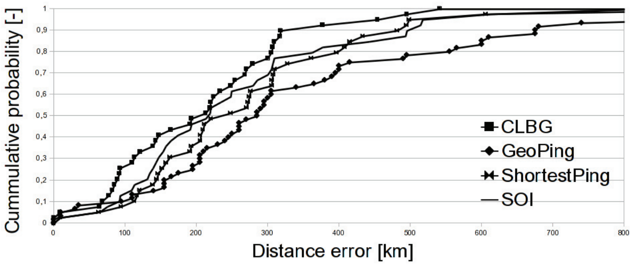

Fig. 8. Cumulative distribution function of geolocation distance error for CLBG, GeoPing, Shortest Ping and SOI

**Conclusion**

This research was focused on IP Geolocalisation techniques based on communication latency to geographical distance conversion. In the paper, we propose an improvement on the constrain-based IP Geolocalisation technique SOI. We address the key-element, which is finding the maximum distance from a target to a given landmark. We define the maximum possible distance by identifying and including a new element, the physical cable length. The physical cable length is estimated by investigating the latency caused by intermediate devices and the speed of signal propagation in optical cables. The physical cable length to linear distance ratio was identified by using the documentation of Czech national network, CESNET2. The estimated maximum distance is then used for finding the area where a target is located using multilateration.

As further research, we would like to compare the effectiveness of constraint-based IP Geolocalisation methods with public geolocalisation databases and with methods based on the use of the DNS system. In addition, we would like to extend our study to cover IP nodes using IPv6 addressing, where the static geolocalisation databases will face a problem of loosely structured IP address space allocation unlike the IPv4 addressing scheme. Furthermore, we would like to investigate the influence of triangle latency inequity on the accuracy of the IP Geolocalisation methods based on latency measurement.

Acknowledgement: The described research was performed in laboratories supported by the SIX project; the registration number CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation. This work was supported by the Czech Ministry of Industry and Trade - project No. FR-TI2/571.

## BIBLIOGRAPHY

[1] Arif, M.J., Karunasekera, S., Kulkarni, S.: Geoweight: internet host geolocation based on a probability model for latency measurements. In: Proceedings of the Thirty-Third Australasian Conferenc on Computer Science - Volume 102, ACSC '10, pp. 89–98. Australian Computer Society, Inc., Darlinghurst, Australia, Australia (2010). URL http://dl.acm.org/citation.cfm?id=1862199.1862209

[2] Bahl, P., Padmanabhan, V.N.: Radar: an in-building rf-based user location and tracking system. In: Proc. IEEE Nineteenth Annual Joint Conf. of the IEEE Computer and Communications Societies INFOCOM 2000, vol. 2, pp. 775–784 (2000). DOI 10.1109/INFCOM.2000.832252

[3] Barnes, R., Lepinski, M., Cooper, A., Morris, J., Tschofenig, H., Schulzrinne, H.: An architecture for location and location privacy in internet applications. RFC 6280, Internet Engineering Task Force (2011). URL http://www.ietf.org/rfc/rfc6280.txt

[4] Bavier, A., Bowman, M., Chun, B., Culler, D., Karlin, S., Muir, S., Peterson, L., Roscoe, T., Spalink, T., Wawrzoniak, M.: Operating system support for planetary-scale network services.

Table 1. Accuracy comparison of IP Geolocalisation methods

| Method | Average error [km] | Median error [km] | Average area radius [km] | Area [km$^2$ ] |
|---|---|---|---|---|
| CLBG | 208 | 213 | 409 | 746923 |
| SOI | 259 | 219 | 472 | 886141 |
| GeoPing | 373 | 285 | - | - |
| Shortest Ping | 280 | 248 | - | - |

In: Proceedings of the 1st conference on Symposium on Networked Systems Design and Implementation - Volume 1, pp. 19–19. USENIX Association, Berkeley, CA, USA (2004). URL http://dl.acm.org/citation.cfm?id=1251175.1251194

[5] Constantinescu, D., Popescu, A.: Modeling of one-way transit time in ip routers. In: Telecommunications, 2006. AICT-ICIW '06. International Conference on Internet and Web Applications and Services/Advanced International Conference on, p. 16 (2006). DOI 10.1109/AICT-ICIW.2006.132

[6] DOLEZEL, P.: Current methods for geolocation of internet hosts. Master's thesis, Brno University of Technology (2001). Thesis. Brno University of Technology. Supervised by Jiri Balej

[7] Eriksson, B., Barford, P., Sommers, J., Nowak, R.: A learning-based approach for ip geolocation. In: Proceedings of the 11th international conference on Passive and active measurement, PAM'10, pp. 171–180. Springer-Verlag, Berlin, Heidelberg (2010). URL http://dl.acm.org/citation.cfm?id=1889324.1889342

[8] Gueye, B., Ziviani, A., Crovella, M., Fdida, S.: Constraint-based geolocation of internet hosts. IEEE/ACM Trans. Netw. **14**, 1219–1232 (2006). DOI http://dx.doi.org/10.1109/TNET.2006.886332. URL http://dx.doi.org/10.1109/TNET.2006.886332

[9] He, Y., Faloutsos, M., Krishnamurthy, S., Huffaker, B.: On routing asymmetry in the internet. In: Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE, vol. 2, p. 6 pp. (2005). DOI 10.1109/GLOCOM.2005.1577769

[10] Huffman, S., Reifer, H.: Method for geolocating logical network addresses (2000). The United States of America as represented by the Director, National Security Agency. United States Patent 752898

[11] Katz-Bassett, E., John, J.P., Krishnamurthy, A., Wetherall, D., Anderson, T., Chawathe, Y.: Towards ip geolocation using delay and topology measurements. In: Proceedings of the 6th ACM SIGCOMM conference on Internet measurement, IMC '06, pp. 71–84. ACM, New York, NY, USA (2006). DOI http://doi.acm.org/10.1145/1177080.1177090. URL http://doi.acm.org/10.1145/1177080.1177090

[12] Komosny, D., Balej, J., Sathu, H., Shukla, R.: Impact of intermediate device latency on ip geolocalisation. In: Proceedings of the 2011 International Conference on Telecommunication Systems (ICTSM 2011). American Telecommunications Systems Management Association (2011)

[13] Kovac, A., Halas, M., Voznak, M.: Impact of jitter on speech quality estimation in modified e-model. In: Proceedings of 12th International Conference on Research in Telecommunication Technologies. Technical University of Ostrava (2010)

[14] Laki, S., Matray, P., Haga, P., Sebok, T., Csabai, I., Vattay, G.: Spotter: A model based active geolocation service. In: INFOCOM, 2011 Proceedings IEEE, pp. 3173 –3181 (2011). DOI 10.1109/INFCOM.2011.5935165

[15] Matray, P., Haga, P., Laki, S., Csabai, I., Vattay, G.: On the network geography of the internet. In: INFOCOM, 2011 Proceedings IEEE, pp. 126 –130 (2011). DOI 10.1109/INFCOM.2011.5934909

[16] Muir, J.A., Oorschot, P.C.: Internet geolocation and evasion. Tech. rep., School of Computer Science, Carleton University (2006)

[17] Padmanabhan, V.N., Subramanian, L.: An investigation of geographic mapping techniques for internet hosts. SIGCOMM Comput. Commun. Rev. **31**, 173–185 (2001). DOI http://doi.acm.org/10.1145/964723.383073. URL http://doi.acm.org/10.1145/964723.383073

[18] Percacci, R., Vespignani, A.: Scale-free behavior of the internet global performance. The European Physical Journal B - Condensed Matter and Complex Systems (2003)

[19] Peterson, L., Bavier, A., Fiuczynski, M.E., Muir, S.: Experiences building planetlab. In: In Proceedings of the 7th USENIX Symp. on Operating Systems Design and Implementation (OSDI (2006)

[20] Priyantha, N.B., Chakraborty, A., Balakrishnan, H.: The cricket location-support system. In: Proceedings of the 6th annual international conference on Mobile computing and networking, MobiCom '00, pp. 32–43. ACM, New York, NY, USA (2000). DOI http://doi.acm.org/10.1145/345910.345917. URL http://doi.acm.org/10.1145/345910.345917

[21] Smith, I., Tabert, J., Wild, T., Lamarca, A., Lamarca, A., Chawathe, Y., Chawathe, Y., Consolvo, S., Consolvo, S., Hightower, J., Hightower, J., Scott, J., Scott, J., Sohn, T., Sohn, T., Howard, J., Howard, J., Hughes, J., Hughes, J., Potter, F., Potter, F., Powledge, P., Powledge, P., Borriello, G., Borriello, G., Schilit, B., Schilit, B.: Place lab: Device positioning using radio beacons in the wild. In: In Proceedings of the Third International Conference on Pervasive Computing, pp. 116–133. Springer (2005)

[22] TruePosition, I.: Trueposition guide to location technologies. White paper (2009)

[23] Voznak, M., Tomes, M., Vaclavikova, Z., Halas, M.: E-model improvement for speech quality evaluation including codecs tandeming. In: Proceedings of the 9th WSEAS international conference on Data networks, communications, computers, DNCOCO'10, pp. 119–124. World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA (2010). URL http://dl.acm.org/citation.cfm?id=1948805.1948828

[24] Wang, G., Zhang, B., Ng, T.S.E.: Towards network triangle inequality violation aware distributed systems. In: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement, IMC '07, pp. 175–188. ACM, New York, NY, USA (2007). DOI http://doi.acm.org/10.1145/1298306.1298331. URL http://doi.acm.org/10.1145/1298306.1298331

[25] Wong, B., Stoyanov, I., Sirer, E.G.: Octant: a comprehensive framework for the geolocalization of internet hosts. In: Proceedings of the 4th USENIX conference on Networked systems design &#38; implementation, NSDI'07, pp. 23–23. USENIX Association, Berkeley, CA, USA (2007). URL http://dl.acm.org/citation.cfm?id=1973430.1973453

**Authors:** Dan Komosny, Jiri Balej, Pavel Dolezel, Department of Telecommunications, Brno University of Technology, Purkynova 118, 612 00 Brno, Czech Republic email: komosny@feec.vutbr.cz, jiri.balej@phd.feec.vutbr.cz, xdolez06@stud.feec.vutbr.cz

Hira Sathu, Ranjana Shukla, Department of Computing, Faculty of Creative Industries and Business, Unitec New Zealand, Private Bag 92025 Auckland, New Zealand email: hsathu@unitec.ac.nz, rshukla@unitec.ac.nz