IEEE - ICICI BME 2013

By Ardian Ulvan

2013 3rd International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME)
Bandung, November 7-8, 2013

Retransmission Issue of SIP session over UDP transport protocol in IP Multimedia Subsystem – IMS

Heru Pranoto [1] and Ardian Ulvan, MSc [2]
Department of Electrical Engineering, State Polytechnic of Medan, Indonesia
[1](Tel: +62618210371; Email: heru.pranoto@gmail.com)
[2]Department of Electrical Engineering, University of Lampung, Bandar Lampung, Indonesia
(Tel: +62721774726; Email: ulvan@unila.ac.id)

Abstract-In this paper, the ability of SIP session over UDP transport protocol retransmission issue in IMS network is investigated and analysed. We described SIP session in IMS with end-to-end SIP signalling flow through the IMS entities such as P-CSCF, I-CSCF, S-CSCF and HSS, including the involvement of related SIP messages. The UDP retransmission issue, e.g., INVITE request and loss rates parameters, were conducted to determine the loss of sessions. It has been found that the gradual increase of session loss occurred after the first attempt of transmission. It is affected by the request rates and loss rates. Therefore, the deployment of RFC3261 TimerB and T1 with value of 32s and 0.5s respectively, will prevent the SIP system from the huge number of retransmission.

Keyword: Wireless Network, IMS, SIP messages, Session Establishment, UDP transport protocol.

I. INTRODUCTION

IP Multimedia Subsystem (IMS) is a new framework, basically specified for mobile networks, for providing Internet Protocol (IP) telecommunication services. It has been introduced by the Third Generation Partnership Project (3GPP) Release 5 as part of the core network evolution from circuit-switching to packet-switching for Universal Mobile Telecommunications System (UMTS) networks [1]. It was refined by subsequent Release 6 and 7 [2] [3]. An IP multimedia framework was later introduced by 3GPP2 as the Multimedia Domain (MMD) for third generation Code Division Multiple Access 2000 (CDMA2000) networks, and finally harmonized with IMS. Real-time services can only be properly supported using the release 6 IMS specifications.

IMS is an international, recognized standard, and now being embraced by other standards bodies including ETSI/TISPAN. The standard supports multiple access types–including GPRS/UMTS, 2G/3G networks, PSTNs, enterprise fixed networks via IP Centrex, residential xDSL or cable modem, WLANs, 3GPP Long Term Evolution (LTE) and LTE-Advanced, and Worldwide Interoperability for Microwave Access (WiMAX).

Session Initiation Protocol (SIP) is used as a signalling protocol in an IMS environment. The SIP protocol [4] provides functionalities such as terminal location, session establishment, session management and participant invocation, including creating, modifying, and terminating sessions with one or more participants. The sessions contain any combination of media such as voice, data, audio, video files, etc., and can be modified at any time by adding new parties or changing the nature of the session. SIP signalling has been designed to be attuned with any transport protocols, typically Transmission Control Protocol (TCP) and User Datagram Protocol (UDP).

This paper focuses on UDP as the transport protocol for SIP signalling. The retransmission issue over UDP and its influence on SIP session performance are respectively determined, investigated and analysed.

The rest of the paper will be as follow: Section II describes IMS architecture. A brief description of SIP and session establishment in registration process is presented in section III. In addition, Section IV addresses the characteristic of UDP protocol and the UDP retransmission issue, respectively. It also includes our results and discussions. The final conclusion is presented in Section V.

II. IMS ARCHITECTURE

In general, the architecture of IMS is divided into three functional layers as depicted in Figure 1 [3] [5]. ***Service layer*** comprises application and content servers to execute basic, supplementary and value-added services for the user. Generic service enablers, as defined in the IMS standard (such as *presence* and *group list management*), are implemented as services in an IP Application Server (AS). The application or service layer contains various types of application servers such as the SIP AS, third-party open service access (OSA) AS, and legacy service control point (SCP) AS. The IMS controls service via the subscriber's home network and those signalling network elements distributed in the service layer and the control layer [6].

The ***control layer*** comprises network control servers for managing call or session set-up, modification and release. The most important of these is the Call Session Control Function (CSCF), also known as a SIP server. This layer contains a full suite of support functions, such as provisioning, charging and operation and management (O&M). Interworking with other operators networks and or other types of networks is handled by border gateways called Session Border Controller (SBC). This layer also includes the Home Subscriber Server (HSS) database, Subscriber Location Function (SLF) database, Policy Decision Function (PDF), and Breakout Gateway Control Function (BGCF).

The ***connectivity layer*** comprises routers and switches, both for the backbone and the access network. The connectivity or access layer is used to transport signalling traffics and media streams. This layer contains switches, router, and media processing entities (MGWs), signalling gateways (SGWs), Media Resource Function Controls (MRFCs), and MRF Processors (MRFPs). Since IMS is designed to be accessed independently, it can connect to different types of existing and emerging access networks as long as they have IP connectivity. Access networks that can connect with the IMS include GSM, GPRS/UMTS, PSTNs, enterprise fixed networks via IP Centrex, residential fixed networks via xDSL or cable modem, WLANs, 3GPP Long Term Evolution (LTE) and LTE-Advanced, and Worldwide Interoperability for Microwave Access (WiMAX).

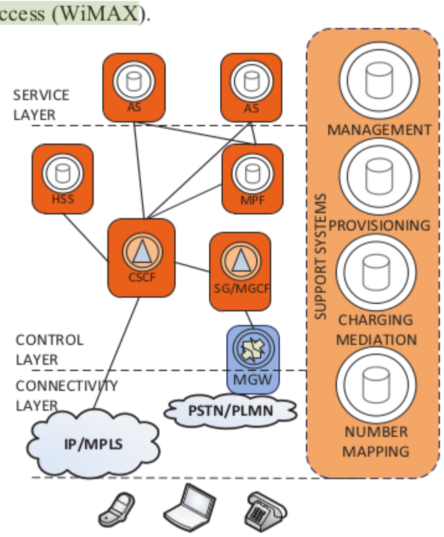

Fig. 1. Layered architecture in IMS

III. Session Establishment in IMS

A. Session Initiation Protocol

Session Initiation Protocol (SIP) is a signalling control protocol at application-layer for creating, modifying, and terminating sessions with one or more participants.

Traditionally, the Internet Protocol (IP) suite provides two core transport layer protocols, i.e., Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). Both have different characteristics when implemented in the communication system. Later, the Internet Engineering Task Force (IETF) through the Signalling Transport (SIGTRAN) working group introduced the Stream Control Transmission Protocol (SCTP) as a transport protocol, serving a similar role to the TCP and UDP.

The SIP is designed to be independent of the underlying transport layer, therefore it should run over TCP, UDP or SCTP. Since each protocols have different transport characteristics, the performance of SIP transmission for particular traffic may also be influenced and vary. The SIP messages can be in the form of request or response from a client to the server or server to the client [4]. Example of the SIP messages is shown in Table 1.

TABLE I
Type of Session Establishment Message

Session establishment message	Type of message
INVITE	Request
100 TRYING	Response
183 SESSION PROGRESS	Response
PRACK	Request
200 OK	Response
UPDATE	Request
180 RINGING	Response
ACK	Request

B. Session Establishment in Registration Process

In this paper, the signalling analysis is just focused in the session establishment registration of an UE in IMS network. The UE is considered to have connected to the network either through the internet or via mobile phone networks.

Figure 2 shows the session establishment flows in case an UE want to register to IMS Network. The UE generates the Request Register and sends to the P-CSCF via the UDP protocol. As no port is indicated in the route header, the request will be sent to the default SIP port (port 5060) [4].

Upon receiving the initial Register request, the P-CSCF acts as the SIP outbound proxy, since the UE is not authenticated yet. The P-CSCF processes the request by forwarding the Register request to the I-CSCF. The P-CSCF removes its own entry from the Route header, hence the Route header will be empty. The only routing-related information left now is the registrar address in the request URI, which point to UE's home network. In order to discover the address of a SIP proxy in home network, the P-CSCF needs to resolve the domain name via the DNS (it is given in the request URI). In addition, the P-CSCF will not put the address of the I-CSCF in the Route header. Nevertheless, the P-CSCF will put the address of the I-CSCF as the destination address in the UDP packet in which the SIP request is transported. Before sending the Register message, the P-CSCF also adds itself to the "Via" header, in order to obtain the response of the request [7].

The I-CSCF receives the Register request and interrogating the S-CSCF that is assigned to serve the UE who is registering. If there is no S-CSCF has been assigned, then the I-CSCF would select one. After putting its own entry in the Via header, the I-CSCF sends the Register request to the S-CSCF, that it either, obtained from the HSS or that is selected by I-CSCF.

Upon receiving the initial Register request, the S-CSCF will request the UE to authenticate himself. This will result another Register request from the UE. This second Register request will include the same registration-related information and will also be route exactly the same way as the initial Register request. If the authentication procedures are successful, the S-CSCF will register the UE. The S-CSCF will also update the information in the HSS to indicate that the UE has been registered now. The HSS will upload UE's user profile to the S-CSCF via the Cx interface. Afterward, the S-CSCF will send back a 200 OK message as the response to the UE. It indicates that the registration procedure has succeeded [8].

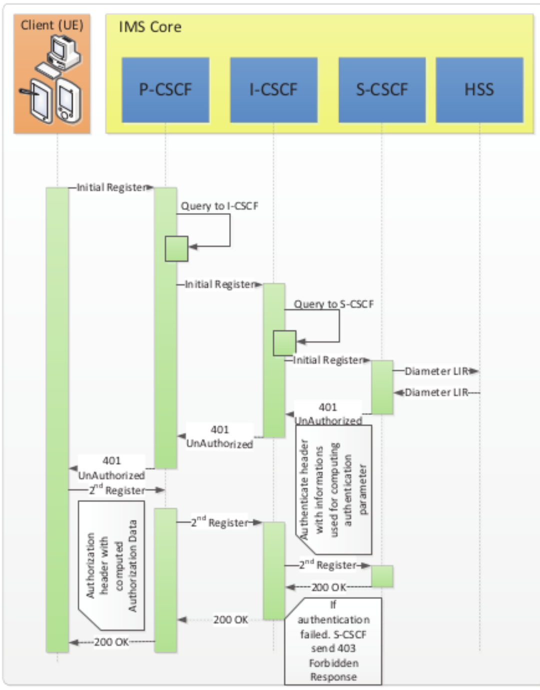

Fig. 2. Session establishment in registration process.

IV. Transport Protocol Retransmission Issue

A. User Datagram Protocol (UDP)

The characteristic, which have been considered in this paper, is the behaviour of UDP when experiencing the loss of packets, and its nature to perform the packet/session retransmission.

UDP is defined to make available a best-effort datagram service mode to an end-system (IP terminal). UDP provides a procedure for application programs to send messages to other programs with a minimum of protocol mechanism. It means the UDP is unreliable hence there is no guarantee for delivery and no protection from duplication [9].

UDP communication accordingly does not incur connection establishment and teardown overheads and there is minimal associated end system state. Based on these characteristics, a very efficient communication but has no inherent congestion control. Another unique characteristic of UDP is no specific requirements acquired when working in many platforms. Applications can send UDP datagrams at the line rate of the link interface, which is often much greater than the available path capacity, and doing so would contribute to congestion along the path, applications therefore need to be designed conscientiously.

When transmitting a UDP datagram, a terminal should complete the appropriate fields in the UDP header and forwards the data along with the header for transmission to the IP network layer. UDP header composes of four fields: Source port, Destination port, UDP length, UDP Checksum.

The pseudo header that conceptually prefixed to the UDP header contains the source address, the destination address, the protocol, and the UDP length (Figure 3). This information gives protection against misrouted datagrams. If the computed checksum is zero, it is transmitted as all ones (the equivalent in one's complement arithmetic). An all zero transmitted checksum value means that the transmitter generated no checksum (for debugging or for higher level protocols that don't care) [9].

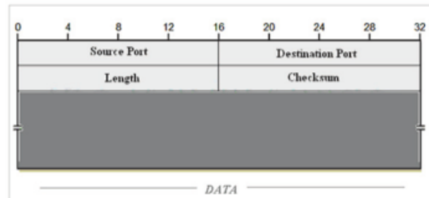

Fig.3. UDP headers.

In IMS, the implementation using UDP transport protocol has a different procedure with TCP, figure 4 depicted when a request is sent via UDP. UDP header indicates the IP address and port number to where all related responses should be routed. When a request is sent out using TCP, the TCP header information is overridden and the response is rerouted back to the sender's address and port number. In contrast, TCP is a connection-oriented transport protocol, so that applying this rule will ensure that there is no need to open additional TCP connections to send responses and requests received via TCP. This cause the routing of SIP messages between P-CSCF and the UE will behave differently [6].

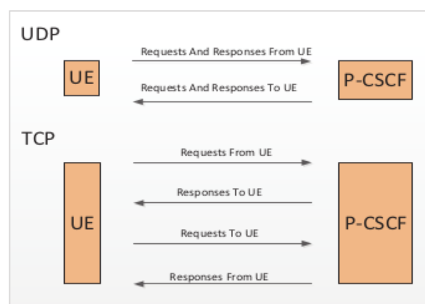

Fig. 4. Requests and Responses UDP and TCP between UE to P-CSCF

B. UDP Retransmission Issue

The UDP [9] uses the connection-less oriented mode which provides a simple transmission model without implicit handshaking, set up channels or data paths between hosts before the communication is established. It is timely delivery therefore the reliability, packet ordering or data integrity cannot be highly expected in this protocol. Its stateless nature make the UDP is compatible with packet broadcast, multitasking and streaming media application such as VoIP.

The performance of SIP transmission for particular traffic may be influenced by the characteristics of the UDP. When working with UDP, SIP conducts an exponential retransmission to enable reliable message delivery. In all presented scenarios, if the SIP sender does not receive a response to a sent request, either due to the loss of the request itself or the response to it, the sender will retransmit the request after a particular time [4], e.g., T1 second. Thus, the timeout value is increased. When the maximum number of retransmission was sent, without any respond, the sender will stop sending the request.

For the case of INVITE requests, the exponential retransmission characteristic is used up to 32 seconds [4], which is called TimerB. The requests are retransmitted at time points T1, 3T1, 7T1, 15T1 and up to TimerB. This can be represented as a series in the form of:

$$(2^1-1)T1, \quad (2^2-1)T1, \quad (2^3-1)T1, \ldots, (2_i^N-1)T1 \qquad (1)$$

Where $(2_i^N-1)T1 = TimerB$.

A loss rate of l is defined as the rate of loss transmitted packets/sessions in the network, and INVITE request rate r is defined as number of requested transmission and retransmission due to the session loss. The average sessions loss per $T1$ seconds can be determined by $r \times l$. On the other hand, the retransmitted sessions would also be suffer from a loss and will have to be retransmitted later. Therefore, the INVITE request session generation rate (I_R) can be determined as shown in Table 2. Detail description on INVITE requests rate and determination of session loss can be found in [10].

TABLE 2
DETERMINED SESSIONS LOST [10]

Ti	I_R	Session loss
0	r	lr
1T1	$r+lr$	$lr+l^2r$
2T1	$r+lr$	$lr+l^2r$
3T1	$r+lr+l^2r$	$lr+l^2r+l^3r$
4T1	$r+lr+l^2r$	$lr+l^2r+l^3r$
5T1	$r+lr+l^2r$	$lr+l^2r+l^3r$
…	…	…
7T1	$r+lr+l^2r+l^3r$	$lr+l^2r+l^3r+l^4r$
8T1	$r+lr+l^2r+l^3r$	$lr+l^2r+l^3r+l^4r$
…	…	…
15T	$r+lr+l^2r+l^3r+l^4r$	$lr+l^2r+l^3r+l^4r+l^5$
…	…	…
31T	$r+lr+l^2r+l^3r+l^4r+l^5$	$lr+l^2r+l^3r+l^4r+l^5r+l^6r$
…	…	…
63T	$r+l^2r+l^3r+l^4r+l^5r+l^6$	$lr+l^2r+l^3r+l^4r+l^5r+l^6r+l^7r$
64T	$r+l^2r+l^3r+l^4r+l^5r+l^6$	$lr+l^2r+l^3r+l^4r+l^5r+l^6r+l^7$

As recommended in [4], $T1$ is set to 0.5 s, whereas the loss rate (l) is assumed to be 5% and 10 % and INVITE request rate (r) were set to 1500, 2000 and 2500 request per second. With all those parameters, the loss of sessions in a particular time can be determined as depicted in Figure 5.

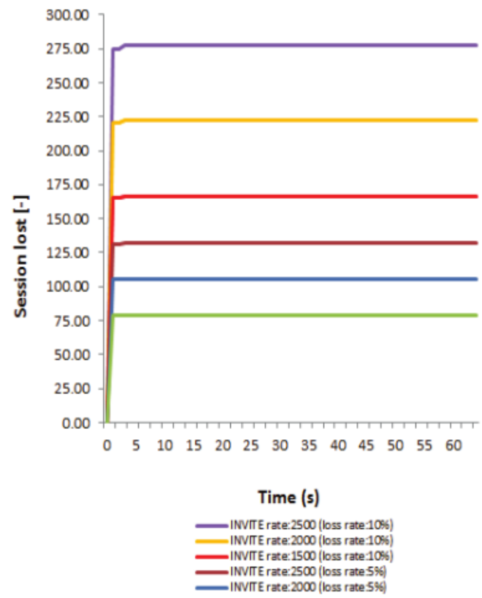

Fig. 5. The characteristic of lost SIP session in UDP transport protocol

As can be seen in Figure 5, number of lost session is influenced by the designed INVITE request rate and loss rate. In the given time, the higher INVITE rate and loss rate will cause the higher session lost during the transmission. With the assigned loss rate, the transmission rate started with 1 session per second, and then increased gradually by the maximum number of additional re-transmission sessions. It is due to the behaviour of UDP that reacted to perform the re-transmission if the respond of the request session is not received when the $T1$ second has been reached. *Timer*B (32 seconds) is applied to deal with this characteristic issue. The deployment of *Timer*B will prevent the SIP system from the huge number of retransmission. In addition, assigning the proper value of $T1$, related to the INVITE and loss rates condition, will also give the solution for a very frequent session re-transmissions.

V. CONCLUSION

In this paper we have described IMS and analysed retransmission SIP over UDP with assumption that the delay occurred due to the nature of UDP unreliable characteristic was true. The performance evaluation based on UDP retransmission, INVITE request and loss rate was conducted to determine the loss of sessions. It has been found that the gradual increase of session loss occurred after the first attempt of transmission. It is affected by the request rates and loss rate. Therefore, the deployment of *Timer*B (e.g., 32 s) will prevent the SIP system from the huge number of retransmission. In addition, assigning the proper value of *T*1 (e.g., 0.5 s), related to the INVITE and loss rates condition, will also give the solution for a very frequent session re-transmissions.

From this study it can be concluded that despite having used widely in IMS system, the UDP transport protocol for SIP session establishment still has some drawbacks due to the unreliability nature of UDP transmission. By providing the given solutions, UDP remains the most prolific protocol for carrying SIP messages. However, taking the other transport protocols such as TCP and SCTP for further investigation and deep analysis is worth to be done.

REFERENCES

[1] ETSI TS 122.228 v5.7.0,"Digital Cellular Telecommunication System (Phase 2+); Universal Mobile Telecommunication System (UMTS); Service Requirements for The Internet Protocol (IP) Multimedia Core Network Subsystem (IMS)", Stage 1, Release 5, 2006.

[2] ETSI TS 122.228 v6.11.0,"Digital Cellular Telecommunication System (Phase 2+); Universal Mobile Telecommunication System (UMTS); Service Requirements for The Internet Protocol (IP) Multimedia Core Network Subsystem (IMS)", Stage 1, Release 6, 2006.

[3] ETSI TS 122.228 v7.6.0,"Digital Cellular Telecommunication System (Phase 2+); Universal Mobile Telecommunication System (UMTS); Service Requirements for The Internet Protocol (IP) Multimedia Core Network Subsystem (IMS)", Stage 1, Release 7, 2007.

[4] Rosenberg, J.,Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and Schooler, E., SIP: Session Initiation Protocol.

[5] RFC 3261, June 2002.Ulvan, A., Ulvan, M.: IP Multimedia Subsystem - IMS: Converged Network Architecture for the Intelligent Interaction of Network Applications and Services. In *Proceedings CD-ROM of Digital Technologies 2007* [CD-ROM]. Žilina: Slovenská elektro technická společnost, 2007, ISBN 978-80-8070-786-6.

[6] Poikselkä, M., Niemi, A., Kartabil, H., Mayer, G.,"The IMS: IP Multimedia Concepts and Services", second edition, Wiley. 2006. ISBN-10: 0470019069.

[7] Kist, A.A., Harris, R.J., "A Simple Model for Calculating SIP Signalling Flows in 3GPP IP Multimedia Subsystems", Springer-Verlag London, UK ©2002

[8] Ulvan, M., Bestak, R., "Delay Performance of Session Establishment Signalling in IP Multimedia Subsystem", in proceeding of 16th International Conference on Systems, Signals and Image Processing (IWSSIP), Chalkida, Greece, June 2009. E-ISBN: 978-1-4244-4530-1. Print ISBN: 978-1-4244-4530-1 INSPEC Accession Number: 11032514.

[9] Postel J, "User Datagram Protocol", RFC 768. IETF. August 1980.Eyers, T., Schulzrinne, H., "Predicting Internet Telephony Call Setup Delay", in IPTel 2000 the first IP telephony workshop. 2000.

[10] Sisalem, D., Liisberg, M., Rebahi, Y., "A Theoretical Model of the Effects of Losses and Delays on the Performance of SIP", in IEEE Global Telecommuncations Conference, 2008 (IEEE GLOBECOM), New Orleans, USA, December 2008. ISSN: 1930-529X. Print ISBN: 978-1-4244-2324-8.

IEEE - ICICI BME 2013

ORIGINALITY REPORT

28%

SIMILARITY INDEX

PRIMARY SOURCES

1	Niebert. "Ambient Networks– The Consequence of Convergence", Ambient Networks, 04/04/2007 Crossref	110 words — 4%
2	www.alcatel.com Internet	78 words — 3%
3	Melvi Ulvan, Robert Bestak. "Chapter 5 Analysis of Session Establishment Signaling Delay in IP Multimedia Subsystem", Springer Science and Business Media LLC, 2009 Crossref	75 words — 2%
4	www.getrfc.ru Internet	60 words — 2%
5	P. F. Rosa. "An IMS Control and Access Layer PR-SCTP Based Network", Fourth International Conference on Networking and Services (icns 2008), 03/2008 Crossref	50 words — 2%
6	Camarillo. "Session Control in the IMS", The 3G IP Multimedia Subsystem (IMS), 06/29/2004 Crossref	46 words — 2%
7	Ajay R. Mishra. "Cellular Technologies for Emerging Markets", Wiley, 2010 Crossref	33 words — 1%
8	rfc768.openrfc.org Internet	29 words — 1%
9	Ali, Sinan Ghassan Abid, Mohd Dani Baba, Mohd Asri Mansor,	

and Labeeb Mohsin Abdullah. "SIP based IMS registration signalling for LTE-based femtocell networks", 2014 IEEE 5th Control and System Graduate Research Colloquium, 2014.
Crossref

26 words — 1 %

10 Yong Fan, Mikko Valkama. "Efficient Control Channel Resource Allocation for VoIP in OFDMA-Based Packet Radio Networks", EURASIP Journal on Wireless Communications and Networking, 2011
Crossref

26 words — 1 %

11 Din, George. "A Performance Test Design Method and its Implementation Patterns for Multi-Services Systems", Technische Universität Berlin, 2009.
Publications

22 words — 1 %

12 mail.ral.ro
Internet

21 words — 1 %

13 www.ericsson.com
Internet

18 words — 1 %

14 www.lib.ncsu.edu
Internet

17 words — 1 %

15 creativecommons.org
Internet

17 words — 1 %

16 Sinda Boussen, Nabil Tabbane, Sami Tabbane. "Performance analysis of SCTP protocol in WiFi network", 2009 First International Conference on Communications and Networking, 2009
Crossref

17 words — 1 %

17 Rogier Noldus. "Introduction to the IMS Network", IMS Application Developer s Handbook, 2011
Crossref

16 words — 1 %

18 Roy, Radhika. "SIP Message Elements", Handbook on Session Initiation Protocol, 2016.
Crossref

16 words — 1 %

19	www.re2ee.org Internet	15 words — < 1%
20	Perez, André. "The IMS Network", Voice Over LTE EPS and IMS Networks, 2013. Crossref	13 words — < 1%
21	www.unb.ca Internet	13 words — < 1%
22	hal.inria.fr Internet	13 words — < 1%
23	L. Veltri, S. Salsano, G. Martiniello. "Wireless LAN-3G Integration: Unified Mechanisms for Secure Authentication based on SIP", 2006 IEEE International Conference on Communications, 2006 Crossref	11 words — < 1%
24	jkobielus.blogspot.com Internet	10 words — < 1%
25	hdl.handle.net Internet	10 words — < 1%
26	Wei-Kuo Chiang, Po-Chiang Kuo. "IMS-Based Centralized Service Continuity", Wireless Personal Communications, 2012 Crossref	9 words — < 1%
27	link.springer.com Internet	9 words — < 1%
28	Ardian Ulvan, Vit Andrlik, Robert Bestak. "The analysis of IEEE802.16e MAC layer overhead and efficiency in PMP topology", 2008 5th IFIP International Conference on Wireless and Optical Communications Networks (WOCN '08), 2008 Crossref	9 words — < 1%
29	www.coursehero.com Internet	8 words — < 1%

30 "Introduction to the IP Multimedia Subsystem (IMS)", Wiley, 2018
Crossref
8 words — < 1%

31 V.C. Joseph, K.K. Lucky, G.N.S. Mohan Rao. "SIP as an enabler for convergence in future wireless communication networks", 2006 IFIP International Conference on Wireless and Optical Communications Networks, 2006
Crossref
8 words — < 1%

32 www.pentasi.net
Internet
8 words — < 1%

33 orbilu.uni.lu
Internet
8 words — < 1%

34 Herlinawati, Umi Murdika, Grienda Elan, Titin Yulianti. "Ultrasonic signal denoising based on wavelet haar decomposition level", 2016 3rd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), 2016
Crossref
7 words — < 1%

35 Ilan Sadeh, Stojan Kitanov, Danco Davcev. "Application of Multiple Description Coding for Adaptive QoS Mechanism for Mobile Cloud Computing", International Journal of Cloud Computing and Services Science (IJ-CLOSER), 2014
Crossref
6 words — < 1%

36 Miran Mosmondor, Lea Skorin-Kapov, Renato Filjar, Maja Matijasevic. "Conveying and Handling Location Information in the IP Multimedia Subsystem", Journal of Communications Software and Systems, 2017
Crossref
6 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE MATCHES	OFF
EXCLUDE BIBLIOGRAPHY	ON		