A Study toward link quality improvement and communication performance enhancement on mobile satellite communication systems

February 2013

Fumiaki Nagase

Doctoral Program in Electrical and Information Engineering

Graduate School of Science and Technology

Niigata University

A Study toward link quality improvement and communication performance enhancement on mobile satellite communication systems

移動体衛星通信システムの回線品質向上および通信性能高速化技術に関する研究

永瀬　文昭

新潟大学大学院自然科学研究科

電気情報工学専攻 情報工学コース

1. Introduction

1.1 Background and motivation

It has been for a long dream of an ideal society traditionally stressed as available to individuals to communicate wherever, whenever and with whomever they please. Recent progress in mobile communication and other network technologies have now been making this dream come true, and mobile communication system will encourage not only people but also computers or other objects to communicate with each other. Of course, service coverage should be continually expanded while service cost is kept low. To achieve the communication system, a mobile multimedia satellite communication for high-speed data transmission will play a key role. This is because no other medium can support communication from anywhere of all over the world.

Pushed by those social demands, mobile communication systems have been spreading rapidly. There were over 127 million mobile subscribers in Japan with around 98% of these using 3G services such as WCDMA (Wideband CDMA) or LTE (Long Term Evolution) into 2012 [1-1]. Mobile communication users demand explosively more weight from phone calls to data service with smart phones, which can dedicate to provide large-capacity multimedia contents including still images and the animation while in a high-speed vehicle of the Shinkansen [1-2][1-3][1-4]. Figure 1-1 shows trends of the mobile communication system. The LTE system [1-5] which is developed from a cellular system that applied an OFDMA/SC-FDMA to realize a higher-speed data transmission for mobile communications of up to one hundred mega-bps is commenced commercialized service in 2010 and expands the service area from urban area. The LTE-advanced system [1-8] has been currently developing as 4G system for up to Giga-bps over wide area. On the other hand, the IEEE 802.11n which is developed from wireless LAN system that applied an OFDM/MIMO technology to realize high-speed data transmission for wireless communications of several hundred mega-bps has completed the standardization process in 2009. The IEEE 802.11ac/ad [1-6][1-7] system that enable super-high-speed data transmission of Giga-bps class have been currently under developing. The wireless LAN system is higher transmission data rate, but can be used under lower mobility of walking and narrower area of several hundred meters compared with cellular system.

Wireless designers constantly seek to improve the transmission data rate, link reliability over wide coverage even in high-speed mobile environments. Mobile multimedia satellite communication with satellite-tracking antennas along with appropriate signaling and receiver techniques can offer a powerful candidate for improving mobile communication performances.

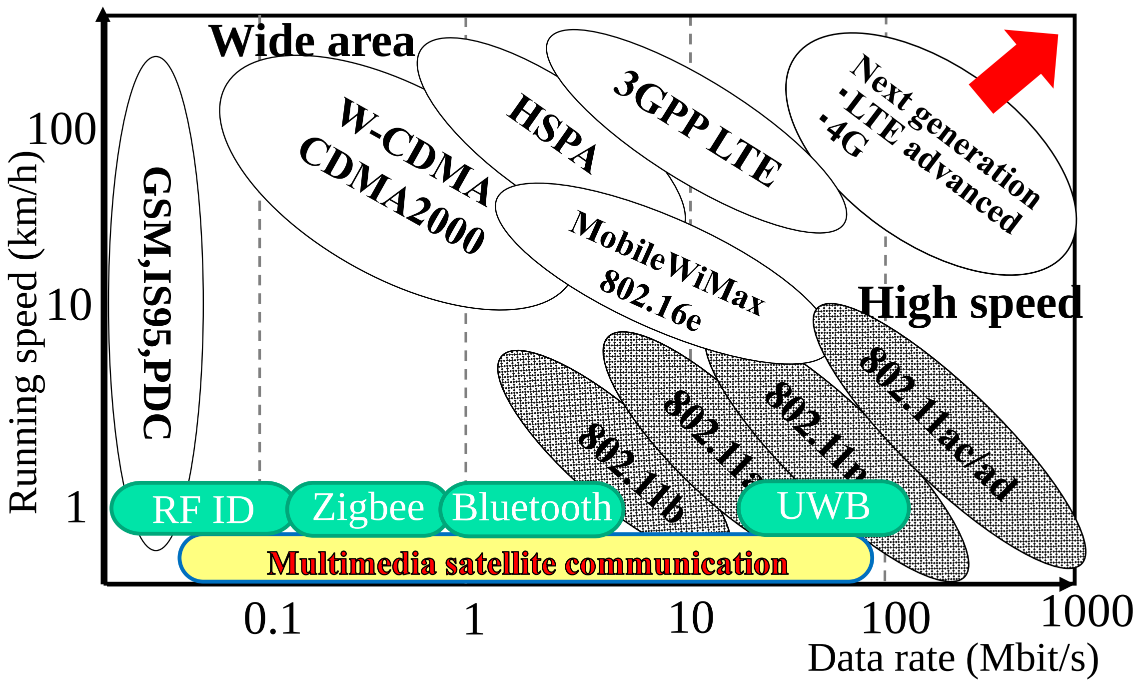


Figure 1-1 Trends of wireless communication systems.

As for the predominant traffic feature of the multimedia communication widely popularizing WWW in the Internet, it is asymmetrical communications between a WWW server and a client. The information including still images and animations are very large capacity data in the forward link (from a WWW server to a client), whereas in the return link (from a WWW server to a client) signals including information such as request signals and the acknowledgement signals of notification confirmation are small capacity data. By taking this feature into account, the multimedia satellite communication system that make use of a hybrid network consisting of high-speed data transmission satellite signals for the return link and a terrestrial network for the return link have been developed and commercialized [1-9]. Figure 1-2 shows the configuration of the system. Employing the DVB (digital video broadcast) standard [1-10] of the satellite link have been able to achieve low-cost reception facilities in the user terminals, and a high-speed data transmission signal [1-11]. These systems dedicate users through a fixed antenna to receive satellite signals for the forward link and terrestrial network for the return link. Since fixed antennas employ high gain and sharp beam antennas, and multi-path signals can be ignored and the received signal level is approximately static in direct-path transmission of satellite communications. Wide area coverage, which is one advantage of the satellite communication, can dedicate to mobile multimedia system immediately as for when and anywhere. The system is expected to usefully provide urgently necessary information to all people even when the terrestrial network systems are damaged under disasters [1-12].

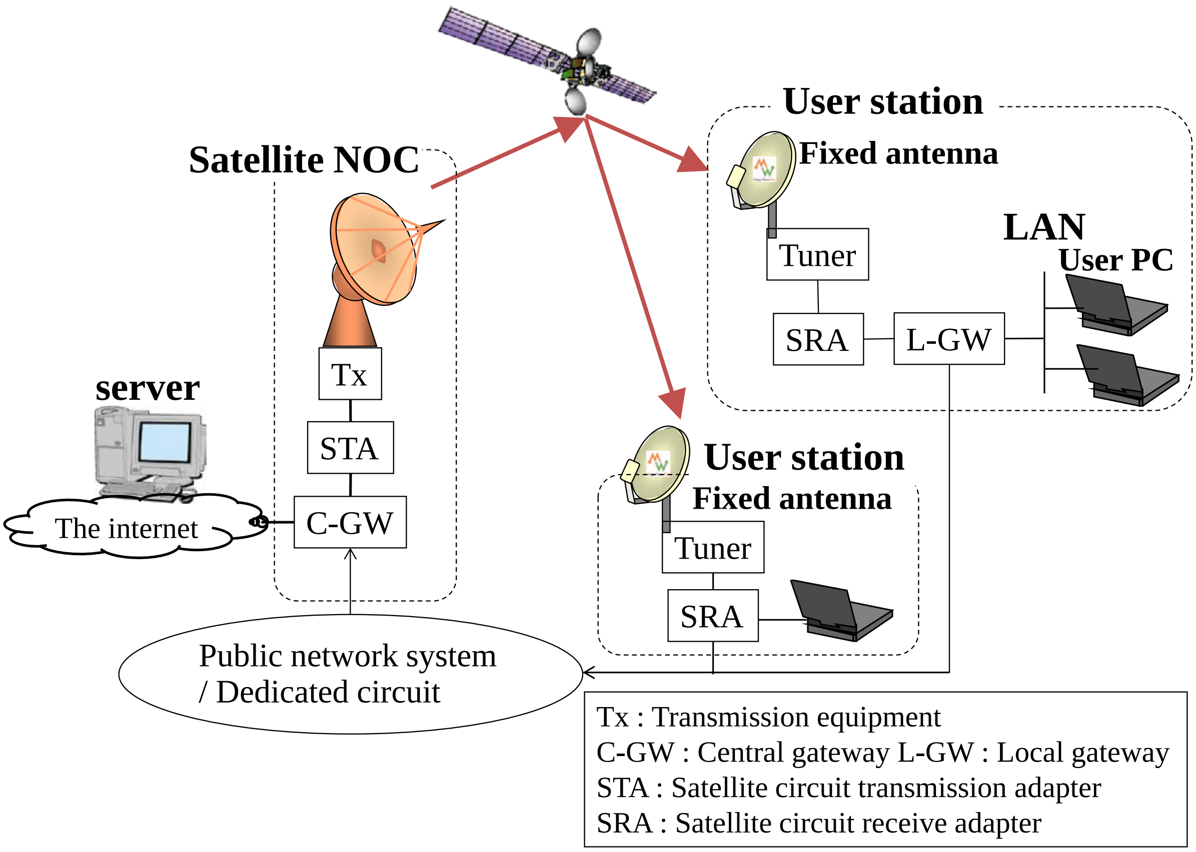


Figure 1-2 Configuration of the multimedia satellite communication system.

1.2 Aims and objectives

In this dissertation, a multimedia satellite communication system to provide mobile multimedia communication for high-speed running vehicles such as buses, trains and the Shinkansen will be proposed. Considering trends of mobile communication systems and the characteristics of the multimedia communication traffic, the proposed system can be dedicated in together with current land mobile systems and high-speed data transmission of the multimedia satellite communication system. Basic configuration of the system and main technique subjects will be clarified. In order to quickly construct the proposed system as objective, the system will be assembled by developing a satellite tracking antenna mounting on those vehicles utilizing the current platform of the multimedia satellite communication system with Ku band [1-9] and exiting land mobile communication systems. Figure 1-3 shows the service concept of the proposed system.

The main object on realizing the system is developing a satellite tracking antenna and middleware techniques. It is required to a developing tracking antenna that tracking a satellite highly precise for securing a same gain as fixed parabola antenna of effective aperture diameter more than about 45cm diameter in mobile environment of running vehicle[1-13]. Some types of Ku band tracking antennas have already been developed for the Satellite News Gathering (SNG) system and for receiving Broadcast Satel-

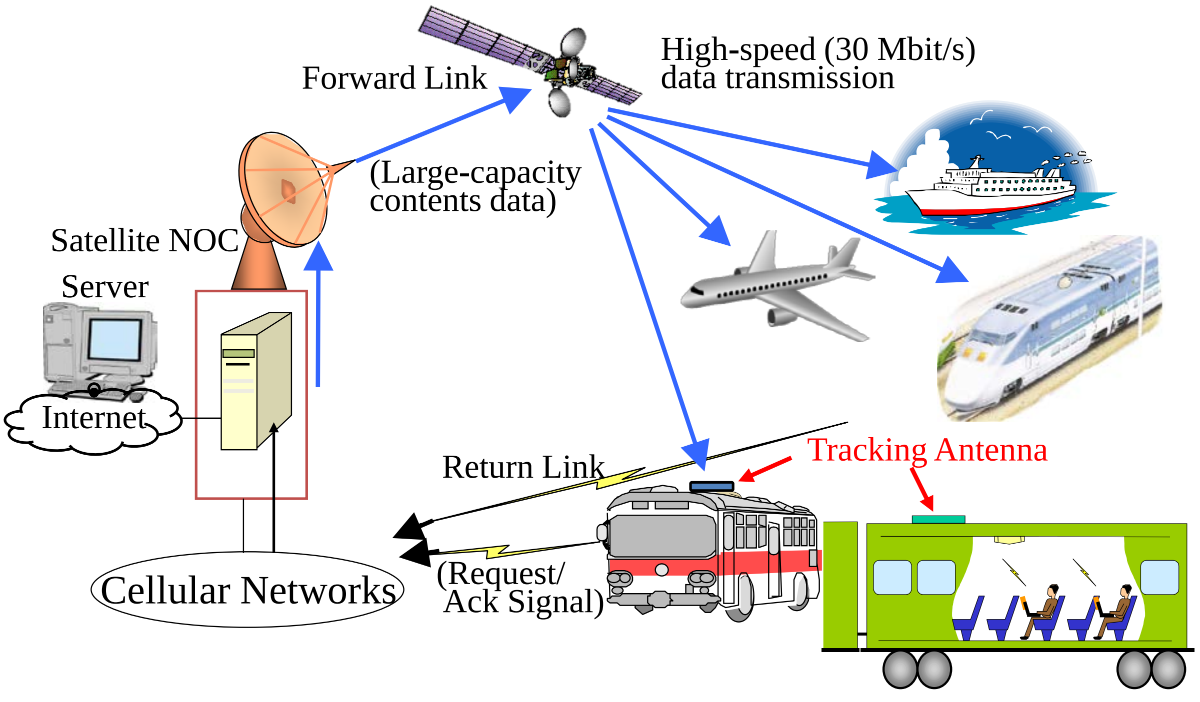


Figure 1-3 Service concept of the proposed system.

lite (BS) signals. Antennas for SNG have already achieved as for a tracking antenna possessing a transmission and reception signals on a vehicle [1-14]. Not only receiving equipment but also transmitting equipment are required, the SNG antenna is very expensive. Furthermore, its applicable region is restricted by its large size; the vehicle must be specially modified to mount it. BS tracking antennas are already commercially available [1-15][1-16]. However, their antenna gain is too small to catch CS signals, which is our target satellite, because the equivalent isotropically radiated power (EIRP) of a BS is about 5 times larger (89dBm) than that of a CS (81.7dBm) [1-13][1-17]. Therefore, the developing antennas are required higher-gain and precisely tracking by the beam narrowing of the gain. In addition, the tracking antenna must distinguish each CS that there is around geostationary orbit longitude every four degrees. Since then, the CS tracking antenna which is an essential key component to realize the system will be described. Improving with the tracking accuracy of the tracking antenna to apply with the system, which enables high-speed data transmission while satisfying requirements of the mounting on the medium/full-size vehicle such as buses and high-speed running vehicles, will be described as the first aim.

The proposed system can come true at low cost, by replacing fixed networks with current mobile communication networks for the return link and replacing a fixed antenna with tracking antenna to receive the satellite signals for the forward link in existing the Ku band multimedia satellite communication systems. In recent years, demands for providing high transmission rate Internet services have been largely emerged for long body vehicles such as trains, buses and airplanes. When applying the system to those vehicles, the satellite signal received at the vehicles is cut off (shadowing) by tunnels, buildings, bridges, and trees. The received satellite signal level is degraded when running direction changing is over the tracking-performance of tracking antennas. Such shadowing and degrading the satellite signal level must occur packet loss [1-18]. Furthermore, the loss of the control packets such as acknowledgement signals in the return link may occur due to handoff (also known as handover) of the mobile communication system [1-19]. Those matters must cause uncomfortable communications due to throughput-deterioration, unexpected disconnection and decreasing effective usage of link bandwidth by retransmission traffic. Development of the middleware to compensate for the comfortable communication is indispensable. Therefore, prevention satellite signal level interception by shadowing and satellite link quality deterioration by tracking-performance of tracking antenna, and improving the quality of the received satellite signal will be described as the second aim.

Then, providing high-speed data transmission of the WWW (World Wide Web) service, that is the most popular and spreading Internet application, with the system will be described. The WWW service can offer various multimedia contents including text files, still images, animations, and sounds with HTTP (Hyper Text Transfer Protocol) for WEB pages. A Web page is composed of several tens or hundreds of web contents, each of which is small of approximately 10-kBytes. The WWW service with HTTP often uses the Transmission Control Protocol (TCP) as the transmission control protocol layer. Owing its reliability of TCP with small start algorithm of the congestion control scheme and small size data transmission, extremely larger latency of about 500-msec in the system than terrestrial network between the clients and the WEB servers, restricts the throughput of HTTP performance regardless of expanding the satellite link bandwidth [1-20]. As an HTTP performance enhancement technology in extremely large latency environments, WEB prefetch schemes [1-21][1-22] using WEB prefetch servers have been devised conventionally. The greatest http performance enhancement effectively doubles is expected without any enhancement schemes when the conventional WEB prefetch server installs at the point of halfway the latency among a user and a WEB server. That means the maximum effective is obtained when the conventional web prefeth server is installed on the satellite. The conventional scheme is not as effective for achieving the maximum performance in the target system, which would incorporate the conventional prefetch server in the vehicle or the ground station. Furthermore, a technology of sending web contents data continuously with user datagram protocol (UDP) as transmission control protocol layer can make high-speed HTTP transmission. Since the UDP characterizes a connectionless communication, retransmission scheme for reliable quality of compensation packet loss is required. However, the transmission rate will reduce with repeated retransmissions over times. Therefore a prefetch scheme with a retransmission scheme with UDP for performance acceleration of the Web services over extremely large latency networks as an application protocol layer in the large latency environment will be described as the third aim.

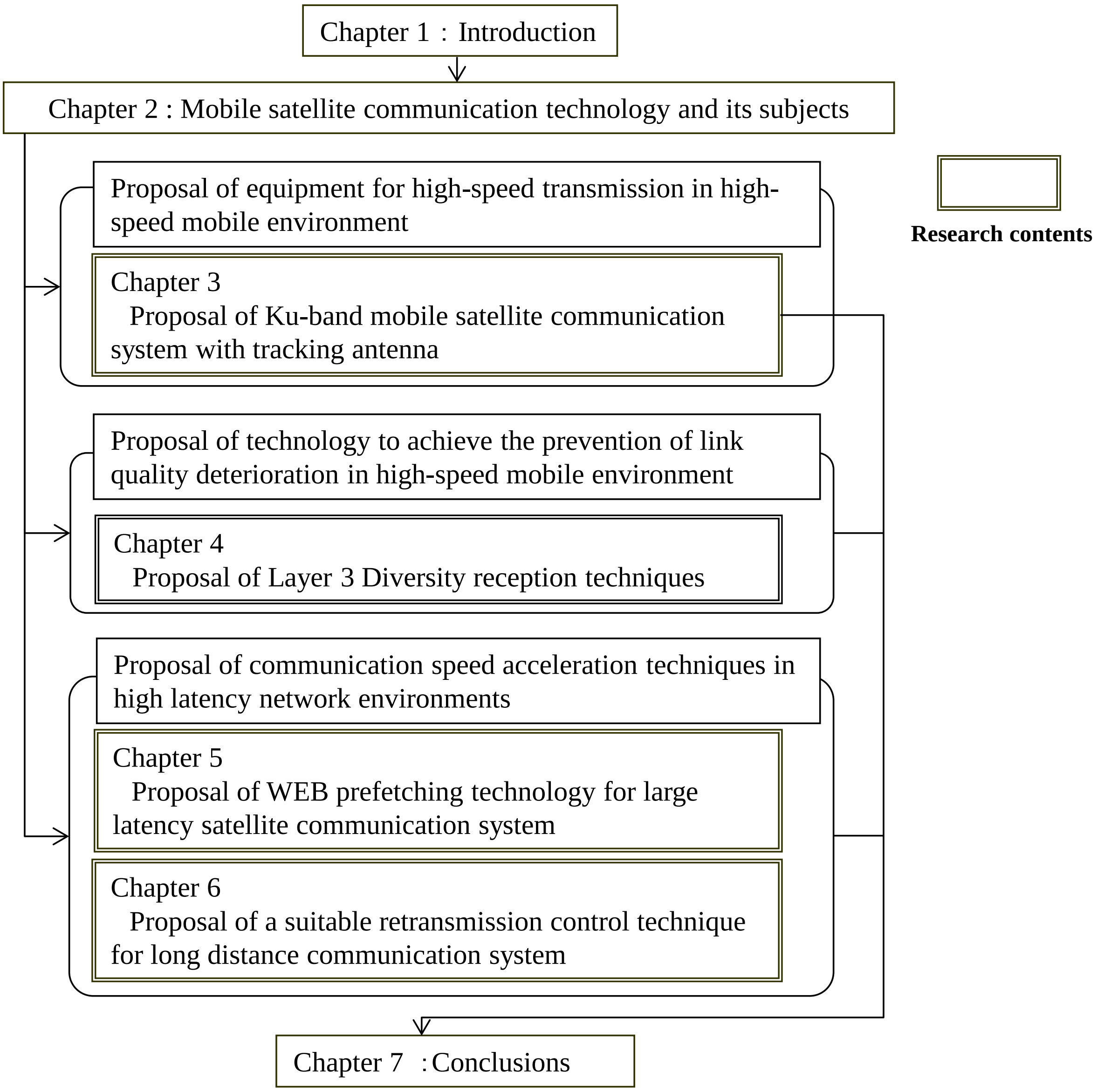


Figure 1-4 Configuration of this dissertation.

1.3 Dissertation structure

The flow of the dissertation is described by a flow chart as shown in Figure 1-4. It clearly shows that the flow from the background and major issues in mobile multimedia satellite communication system, proposal, cost-effective and quickly development until implementation of the system under real environment and test-bed system in the laboratory.

Chapter 2 provides techniques and issues to build a mobile satellite communication system for high-speed running vehicles economically and quickly over wide area. This chapter also shows related matters with the middleware technology essential for realization of the system.

Chapter 3 focuses the requirements for the satellite tracking antenna which is the key component technology to realize the proposed mobile satellite communication system for vehicles such as buses, trains, and airplanes. A low-profile satellite tracing antenna for mounting onto the roof of vehicles is proposed and developed. A practical system that implements the developed tracking antenna on a bus is constructed. It shows tracking accuracy of the tracking antenna and evaluation its network performance through various demonstrations with the practical system by running the bus on roads and highways or by running real trains.

Chapter 4 focuses the requirements for receiving stably the satellite signal in large vehicles like trains for high quality communication and proposes packet selection schemes of layer-3 diversity receiving technology by using more than two tracking antennas or receiving systems with gap-filler system. Making the same satellite signal receiving shadowing environment of running a train in the laboratory, measuring the receiving packet order and verification effectiveness of the network performance are described.

Chapter 5 and chapter 6 describe HTTP performance enhancement schemes and retransmission schemes to use comfortably the WEB service, which is the most popular application in the Internet, for users in long latency network systems.

Chapter 5 describes web prefetch schemes that can enhance the HTTP throughput of the WEB service in large latency network systems. A new web prefetch scheme is proposed and mathematical evaluation is performed. A test-bed machines installed the proposed scheme is developed and constructed the test-bed system in the laboratory is constructed with the same environment of the real system. Network performance test carrying out with the test-bed system is described.

Chapter 6 describes a new retransmission scheme based on a Selective Repeat technique in large latency networks for compensate high-speed data transmission under link conditions of high packet loss ratio. Simulation evaluation of performance effect with the scheme by packet loss ratio is discussed.

Finally, the last chapter draws conclusions on the work, and provides scope and direction for further work.

References for chapter 1

[1-1]"Telecommunications services", 2011 WHITE PAPER Information and Communications in Japan, section 3 chapter 4, part 3, pp.50, Ministry of Internal Affairs and Communications, Japan, 2011.

[1-2]Hiroshi Hojo, "Trends in Wireless Access System Technology toward Expanded Broadband based on Optical and Wireless Systems", NTT technical review, vol.6, no.5, May, 2008.

[1-3]Kita, N., Ito, T., Yokoyama, S., Ming-Chien Tseng, Sagawa, Y., Ogasawara, M., and Nakatsugawa, M., "Experimental study of propagation characteristics for wireless communications in high-speed train cars", Antennas and Propagation, 2009. EuCAP 2009. 3rd European Conference on Date of Conference: 23-27 March 2009.

[1-4]Yukiya Tateishi, Takahiro Baba, Yusuke Suzuki, and Ko Takani, "Broadband Radio Transmission in Railways", JR EAST Technical Review-No.20, pp.34-pp.39, summer 2011.

[1-5]Takehiro Nakamura, Sadayuki abeta, Mikio Iwamura, Tetsushi Abe, Motohiro Tanno, “Overview of LTE-Advanced and Standardization Trends,” NTT docomo technical journal, vol.12, no.2, Sep. 2010.

[1-6]IEEE 802, “IEEE P802.11ac/D0.3”, Institute of Electronic Engineers, April 7, 2011.

[1-7]"IEEE 802.11ad". Standards.ieee.org. Retrieved 2012-07-24.

[1-8]3GPP TR 36.912 V.9.0.0, “Feasibility study for Further Advancements for E-UTRA (LTE-Advanced),” 2009.

[1-9]M. Nakayama, M. Nakagawa, Y. Hashimoto, K. Tanaka and H. Nakashima, “A Satellite Communication System for Interactive Multimedia Networks,” IEICE Trans. On Commun. , Vol. E80-B, No. 1, pp. 103-108, January 1997.

[1-10] Hajime MATSUMURA and Hisakazu KATOH, “Satellite Digital Broadcasting Systems,” Broadcast Technology no.18, Spring 2004.

[1-11] M. Nakayama, K. Araki, M. Kobayashi,and H.Nakashima, ”A Satellite Communication System using ATM Multiplex Scheme for Interactive Multimedia Networks”,48th IAF, Oct. 6-10,1997,Turin,Italy.

[1-12] Yutaka Imaizumi, Takashi Hirose, and Hidekuni Yoshida, “Small Satellite Earth Stations for Disaster Recover Operations,” NTT Technical Review, vol 10, no. 7, July 2012.

[1-13] Hideki Toshinaga, Kiyoshi Kobayashi, Kohei Ohata, Masazumi Ueba and Hideki no, ”Bi-directional Multimedia Satellite Communication System Employing Superposed Transmission,” Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE

[1-14] Toru IWASAKI and Takao MURATA, “Electromagnetic Coupling-Fed Active Slot Array Antenna - Toward an ultra low-profile transmission antena suitable for a portable SNG-“, NHK STRL R&D No.73, May. 2002.

[1-15] Toshiaki Watanabe, Masaru Ogawa, Tomohisa Harada, Kunitoshi Nishikawa, "Mobile Antenna System for Direct Broadcasting Satellite," R&D Review of Toyota CRDL, Vol. 32 No. 1, pp.23-pp.33, March 1997.

[1-16] Y. Minami, S. Aoshima, M. Morita, K. Nishikawa and S. Inamuera, “Development of BS Reception System for Small Car”, Car Technical Journal, Vo.51, No.2, pp.59-pp.63, 1997.

[1-17] Saito, T. Hashimoto, A. Minematsu, F. Katoh, H. Matsumura, H., “Transmission system for satellite ISDB,” Global Telecommunications Conference, 1998. pp.2942 – pp.2947, vol.5, 1998.

[1-18] Fumiaki Nagase、Jin Mitsugi、Masayoshi Nakayama、Masazumi Ueba, "Ku Band Mobile Multimedia Satellite Communication System for Trains", 21st AIAA ICSSC, AIAA-2003-2205, 2003.

[1-19] Yutaka Imaizumi, Satoshi Harada, Fumiaki Nagase, Masayoshi Nakayama, Kohei Ohata and Masazumi Ueba. ”Network performance evaluations for Ku band mobile satellite communications system based on Shinkansen railroad's propagation conditions”, Vehicular Technology Conference, 2005. VTC-2005-Fall. 2005 IEEE 62nd, pp. 890 – pp.894, 2005.

[1-20] Hiroyasu Obata, Kenji Ishida, Junichi Funasaka, and Kitsutaro Amano, “Data Transfer Time by HTTP 1.0/1.1 on Asymmetric Networks Composed of Satellite and Terrestrial Links,” IEICE transaction on communication, E85-B/ 12, pp.2895-pp.2903, 2002.

[1-21] K. Chinen, “An interactive prefetching proxy server for improvement of WWW latency,“ Proc. INET ’97 Conference, 1997.

[1-22] V. N. Padmanabhan and J.C. Mogul, “Using Predictive Prefetching to Improve World Wide Web Latency,” ACM SIGCOMM Computer Communication Review, Vol. 26, No.3, July 1996.

2. Technologies and issues for mobile multimedia satellite communication system

2.1 Introduction

A demand to access the Internet in high-speed running vehicles such as trains, buses, and airplanes is increasing by rapidly spreading of mobile multimedia communication terminals such as smart phones. As well as the multimedia data transmission such as photographs, and animations, it is essential to realize a high-speed data transmission mobile communication system with stable quality corresponding to a variety of application including the WWW service. Mobile communication systems by using satellite are a strong candidate to provide a mobile multimedia communication immediately at low cost over a wide area for users in such environments.

In satellite communication systems, not only link switching technologies but also packet switching technologies such as random access control schemes have been researching and developing conventionally [2-1][2-2]. Satellite communication technologies that can realize for high-speed data transmission of multimedia communications for fixed users have been developed and practically used [2-3][2-4][2-5]. Fundamental technologies for mobile multimedia satellite communication system [2-6], which can realize high-speed data transmission in high-speed running vehicles is firstly described in this chapter. The system uses tracking antenna mounting on a vehicle in substitution for fixed parabola antenna in the user station. Users in vehicles can receive high transmission rate signals via satellite links in mobile environment from servers. High-speed data transmission communication can be provided for users in the vehicles with the proposed mobile multimedia satellite communication system, which incorporates with the current mobile communication systems over wide area immediately at low cost. Since high-speed data transmission requires high C/N of link quality. It means that the antenna of higher tracking accuracy than that of the antenna used in the Inmarsat system [2-7], which is popular installed for ships and planes, is essentially requisite. Therefore, requirements for tracking antenna in mobile environment and issues for realizing stable link quality under the mobile environments in mobile multimedia satellite communication system are described. Finally, communication quality of the Internet-access by the network properties of this system and approaches for issues are explained for providing comfortably the WWW services.

2.2 Fundamental technologies for mobile multimedia satellite communication system

The forward link specifications of the multimedia satellite communications [2-3] are shown in Table 2-1. In order to enable high-speed data transmission, the DVB (Digital Video Broadcasting) standard [2-8], which is widely used for broadcasting satellite system into the satellite link, is adapted. Therefore, the receiver equipment such as antennas or the reception tuners can economically cost lower. Table 2-2 shows the link budget of the multimedia satellite communication system.

Table 2-1 Forward link system specifications.

Item	Specification
Frequency band	Ku band(12.25-12.75GHz)
Modem	QPSK
Transmission Bandwidth	27MHz
Transmission rate	42.192Mbit/s
Information rate	29.162Mbit/s
Error correction	Outer coding: Reed Solomon(188/204) Inner coding: Convolutional encoding Viterbi decoding (3/4)
Framing	MPEG2 transport stream
Multiplexing	ATM cell multiplexing

Table 2-2 The link budget of the multimedia satellite communication system

Modulation scheme	QPSK
Forward error correction	Conventional coding (K=7,R=3/4) Reed Solomon coding (204,188)
Required C/N	5.5 dB
Uplink	
Data rate	29.162 Mbit/s
Clock rate	21.096 MHz
EIRP at CES	91 dBm
Power at satellite antenna	-115.9 dBm
C/N	21.4 dB
C/I	23.2 dB
Downlink	
EIRP (satellite)	81.7 dBm
Power at UES antenna	-123.9 dBm
C/N	12.4 dB
C/I	11.8 dB
Total C/N+I	8.7 dB
Rain attenuation	3.8 dB
Total C/N	6.6 dB
Margin	1.1 dB

Upsizing schemes of satellite onboard antenna, increasing transmission power schemes of satellite onboard equipment and terrestrial equipment, error correction schemes of Reed–Solomon/Viterbi-decoded convolutional coding, it has been realized that small antenna aperture size of about 45cm can be achieved in the user stations for high-speed data transmission communication of about 30Mbit/s.

In this chapter, the requirements for the configuration of the network system and tracking antenna schemes to realize the mobile multimedia satellite communications system are explained.

2.2.1 Configuration of the proposed mobile multimedia satellite communication system

Figure 2-1 shows the configuration of a prototype system constructed for evaluation purposes. In Figure 2-1, the forward link is indicated by a solid line and the return link by a dotted line. As for the mobile multimedia satellite communication system, the network system constitutes as asymmetry routing that a forward link and a return link are not the same routes. The basic configuration of the mobile multimedia satellite communication system combines a current mobile network for the return link and the Ku-band satellite network for the forward link. It comprises a vehicle mounted the developed tracking antenna, an information server center (ISC), a satellite network operation center (NOC), and a communication satellite (CS). The NOC and the CS are now in commercial operation for satellite Internet/Intranet services. The vehicle contains a tracking antenna, user PCs connected to a LAN, a satellite router, and a cellular terminal. The tracking antenna is for receiving satellite signals, and the cellular terminal is connected to the ISC through a cellular network. The ISC has an evaluation server, an access server, and an ATM router. The access server connects the user PCs to the server via a cellular network, and the ATM router forwards data from the server to the NOC via a leased line.

2.2.2Tracking antenna design

High gain causes the beam width narrow for high C/N link quality in high-speed transmission.

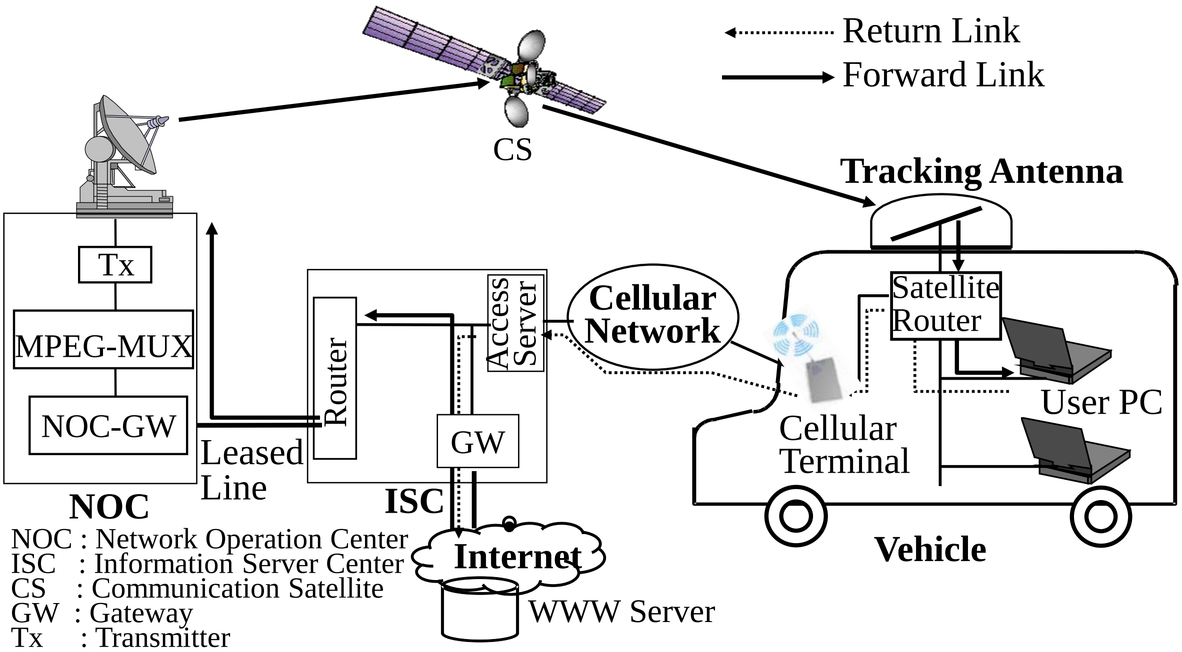


Figure 2-1 Configuration of the high-speed mobile multimedia satellite communication system.

Since then, antennas for consistently tracking the point direction of the main robe to the satellite are essential in the vehicle, where the satellite direction changes with the variation of running direction. The tracking control scheme is classified roughly for the open loop control system and the close loop control system. The point direction of the antenna is calculated from information of inertial navigation systems or gyro systems in the opened loop control system. The point direction is controlled by feedback loop that the variation of received satellite signal level in the closed loop control system. The step truck scheme and the monopulse scheme are mainly used as closed loop control system [2-9] as shown in Figure 2-2. In the step truck scheme, the pointing direction of the antenna is controlled at the received satellite signal level of maximum. In the monopulse scheme, the pointing direction of the antenna is controlled to detect the point error at zero degree with comparing the signal phase difference of the signal received by plural sub-array antennas.

Advantages of steptrack schemes are relatively cheap and simple to implement. Disadvantages of steptrack schemes are stepping results in loss of gain and easily confused by level changes (e.g. as a result

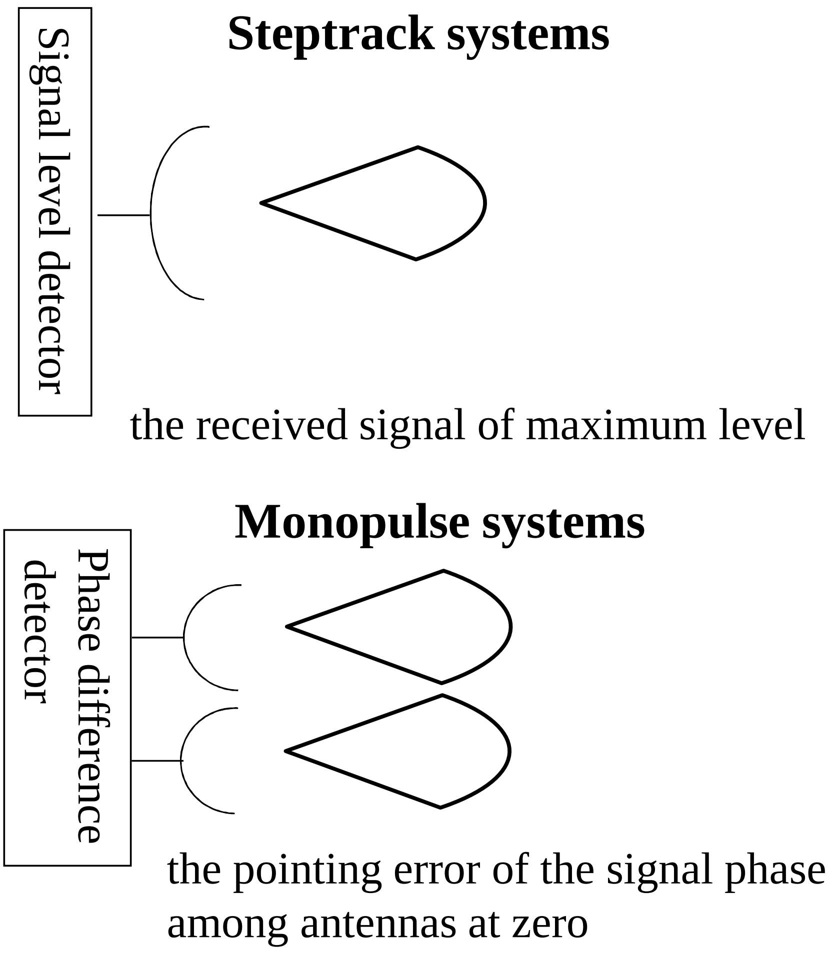


Figure 2-2 Classical closed loop tracing scheme algorithm.

of rain fade). Advantages of monopulse schemes are the best tracking error sensitivity, and high tracking precision by easily realization of array antennas with 4 quadrants sub-arrays. Disadvantages of monopulse schemes are complicated constitution of the feeding circuits and a receiver for the phase difference signal detection is necessary.

Monopulse tracking have been utilizing over alternative tracking schemes, such as conical scan or steptrack, due to the challenging pointing requirements. The alternative scheme generate error information by making small changes in the antenna pointing angle to modulate a received beacon signal at the cost of increasing the overall angular pointing error budget. Another significant advantage of the monopulse system is the wide pull-in range. This allows the tracking system to rapidly re-acquire the satellite should the link be lost e.g., due to shadowing by buildings, dense vegetation or overhead structures.

2.3 Approaches and issues for mobile multimedia satellite communication system

2.3.1 Requirements and countermeasure technologies for issues

In this chapter, approaches and issues to realize mobile multimedia satellite communication system is described. The network configuration is shown in Figure 2-3. Traffic trends of the current broadband of wireless multimedia services are described for high-data speed transmission about the Internet access. Technologies in the system that can provide stable link quality are described. Development of highly precise tracking antenna is essential requisites. Therefore the requirements to the tracing antenna is described and following requirements are clarified to realize stable and high communication quality without packet losses under the high speed mobile environment where shadowing occurred frequently.

(i) It must adapt to signal symbol differences which are caused by the diversity distance and the running direction of the train.

(ii) It must furnish seamless connectivity with gap-fillers

(iii) It must keep the packet order correct.

(iv) A minimum alteration of receivers and transmitting servers

Finally, it is described that Quality of Service in the WWW services comfortably under high latency environment. It is also described about a new retransmission control scheme to realize high transmission efficiency under high packet loss rate and high latency environment. In Chapter 3 through Chapter 6, suggestions and the examinations of the technique to solve the issues mentioned above are pushed forward

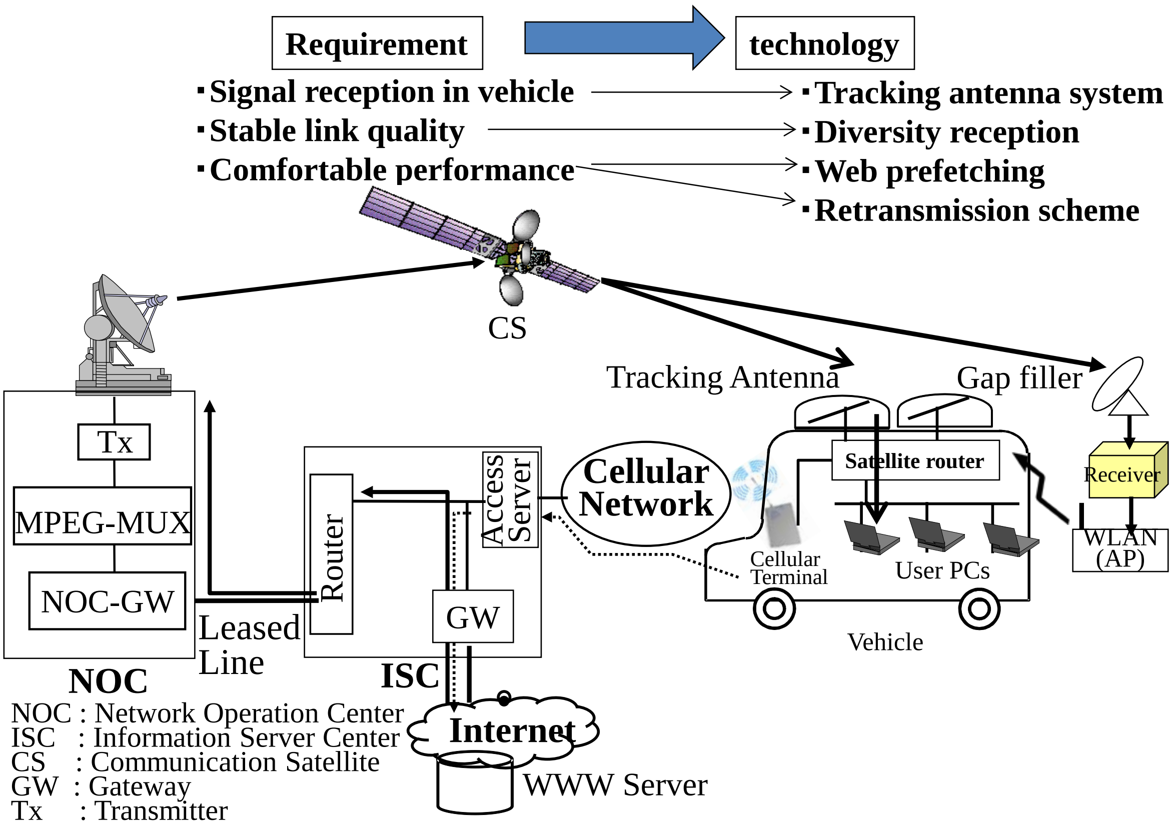


Figure 2-3 The configuration of the mobile multimedia satellite communication system and its issues.

2.3.2The first issue: low profile tracking antenna with high accuracy

The tracking antenna is the key component of the system [2-10][2-11]. Some types of Ku band tracking antennas have already been developed for vehicle systems; examples include antennas developed for the Satellite News Gathering (SNG) system and for receiving Broadcast Satellite (BS) signals [2-12]. A mobile SNG car is shown in Figure 2-4. The SNG antenna, however, is very expensive because of its receiving and transmitting functions. Furthermore, its applicable region is restricted by its large size; the vehicle must be specially modified to mount it. BS tracking antennas are already commercially available. A tracking antenna for BS is shown in Figure 2-5. However, their antenna gain is too small to catch CS signals, which is our target satellite, because the equivalent isotropically radiated power (EIRP) of a BS is larger than that of a CS. The major requirements and specifications for the tracking antenna are shown in Table 2-3. The radio link analysis imposes the requirement of a high antenna gain, including radome loss and accurate tracking control. The height of the radome is also imposed by mounting on vehicles.

Figure 2-4 A mobile SNG car.

Figure 2-5 A TV-car of Keihan railway mounting a tracking antenna for BS.

Table 2-3 Main requirements, specification and features of tracking antenna.

Requirement	Specification & Feature
High gain antenna	> 32.0 dBi in operation (including radome loss and gain reduction caused by tracking error)
Small size (Height dimension)	< 350 mm (Radome height) Microstrip antenna (32 x 16 elements)
Accurate tracking control	<1 dB gain reduction by tracking error

2.3.3The second issue: link quality degradation by shadowing

Satellite signals received in vehicles are blocked of shadowing by tunnels, buildings, bridge, and trees and so on. Shadowing may cause link quality degradation, which deteriorates throughput performances [2-13]. There are many kinds of obstacles along railways that cause shadowing.

Figure 2-6 shows the shadowing environment in a running train. By analyzing different obstacles, the satellite signals under the railroad running environment are blocked periodically in a short time by an overhead wire structure installed every the approximately same distance. We found that the shadowing influence can be reduced by making the diversity distance long. Therefore, diversity reception of long distance with satellite tracking antennas is effective [2-14] if the shadowing is shorter in length than the train itself. Diversity reception with a satellite tracking antenna and gap-filler system is effective if the shadowing is longer than the train itself. Considering from above usage of diversity reception, the requirements for diversity reception can be summarized as follows.

(i) It must adapt to signal symbol differences which is caused by the diversity distance and the running direction of the train: Since shadowing influence can be reduced by making the diversity distance long, long diversity distance is effective in order to avoid the network performance degradation caused by shadowing from terrestrial obstacles. However this may generates a large symbol difference.

(ii) It must furnish seamless connectivity with gap-fillers: Some kind of gap-filler should be used to provide network connectivity in shadowing areas where satellite signals cannot be directly received such as in tunnels, longer than the train. It is desired that the diversity reception be applied to satellite reception

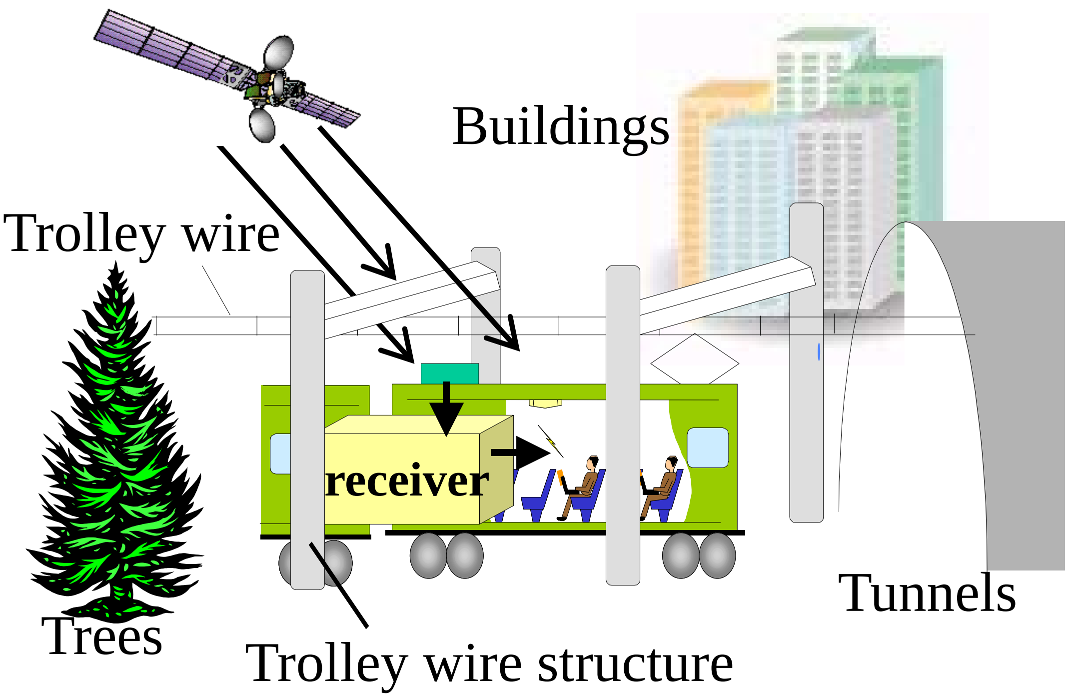


Figure 2-6 Shadowing environment.

and the gap-filler system.

(iii) It must keep the packet order correct: When the train leaves the tunnel, the packet arrival times on the two branches (satellite and wireless LAN, IEEE802.11b with DCF method9) are quite different. These packet arrival time differences can yield packet order scrambling. Failure to keep the original packet order can degrade the quality of the streaming video.

(iv) A minimum alteration of receivers and transmitting servers: Since we want a technology that can be simply added to an existing multimedia satellite communication system, alterations to existing equipment must be minimized.

Table 2-4 summarizes a comparison between diversity reception systems. There are mainly two diversity systems, which are layer 1 diversity and layer 3 diversity. Three conventional techniques are used for diversity reception on layer 1: selection combining, maximal ratio combining, and equal gain combining. The reception branch with the highest power signal is chosen with a switch for selection combining; arbitrary diversity branches are first co-phased and then weighted proportionally to their signal level before summing for the maximal ratio combining; and diversity branches are summed after co-phasing for equal gain combining. These techniques need switch or co-phasing circuits in the satellite receiver. Satellite receivers need equalizers to compensate for the signal symbol difference, whose number varies due to the diversity distance and the running direction of a train, between diversity reception branches. However it is not easy to adapt the dynamic signal symbol difference for equalizers if the diversity distance is long, thus causing a large symbol difference. Therefore, layer 1 diversity usually cannot satisfy requirements (i) and (iv). Furthermore, it is impossible to satisfy requirement (ii) for layer 1 diversity because it selects or combines the same layer signals between branches.

Table 2-4 Comparison between diversity reception systems.

<table>
<tr><th colspan="2">Requirements / Diversity system</th><th>Adaptation to symbol difference</th><th>Seamless connectivity with gap -fillers</th><th>Assuring packet order</th><th>Alteration of existing equipment</th></tr>
<tr><td rowspan="3">Layer1</td><td>Selection combining</td><td>△ (Alteration of receivers and need equalizers)</td><td>×</td><td>△</td><td>×(Need switches in receivers)</td></tr>
<tr><td>Maximal ratio combining</td><td rowspan="2">△ (Alteration of receivers and need equalizers)</td><td rowspan="2">×</td><td rowspan="2">○</td><td rowspan="2">× (Need co -phasing circuits in receivers)</td></tr>
<tr><td>Equal gain combining</td></tr>
<tr><td>Layer3</td><td>Reorder packet by ID No.</td><td>○</td><td>○</td><td>○</td><td>×(Alteration in transmitting servers)</td></tr>
</table>

2.3.4The third issue: HTTP performance for the WWW service in large latency network

The World Wide Web (WWW), also known as Web service, is standardized by the World Wide Web Consortium (W3C). WWW is the most popular service among Internet applications. A Web server provides Web contents composed from hypertexts that include multimedia contents, such as pictures, moving videos, and music. For example, in the case of constitution as shown in Figure 2-7, the WEB page is made from a content of text 1, text 2, text 3, text 4, text 5, picture 1, picture 2, picture 3, picture 4 and picture 5. The figure also shows transmission of these contents for requests from users (clients) to the Web server, and the Web server replies Web contents for the request. A Web service is composed of several tens or hundreds of web contents, each of which is approximately 10 kBytes. The request and the Web contents are transferred by the Hypertext Transfer Protocol (HTTP) [2-16][2-17] which is an application layer protocol.

Wide-spread use of broadband services can enable high transfer speed communication for Web service.

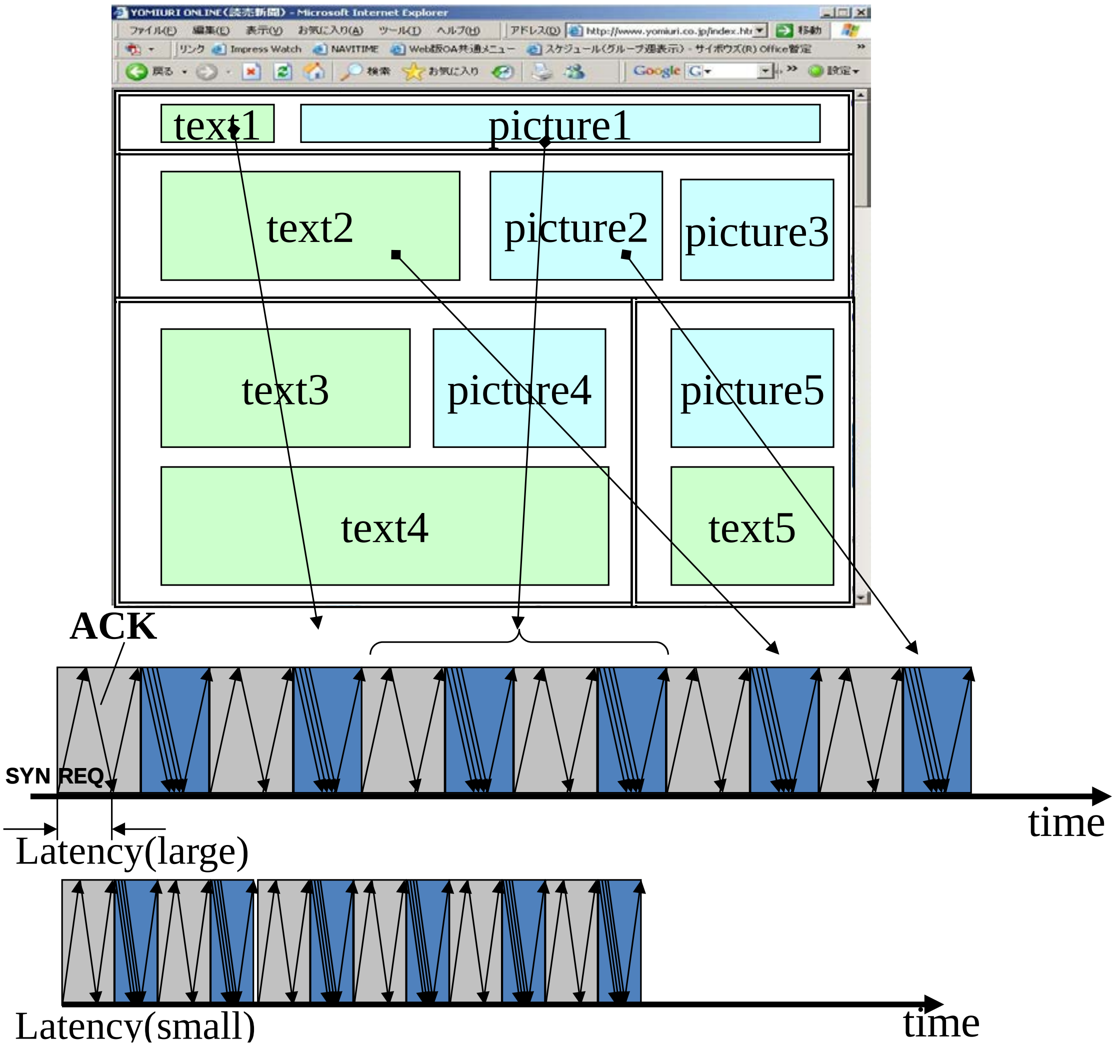


Figure 2-7 an example of web browsing display and its delivery.

Web service with HTTP often uses the Transmission Control Protocol (TCP) as the transmission protocol.

Clients in a vehicle can access the Internet via the satellite link between the vehicle and the ground station. The long propagation distance via the satellite link causes extremely large latency of about 500 ms between the clients in the train and the Web server. In this system for high-speed running vehicles, since clients in the trains connects to WLAN (Wireless Local Area Network) equipment located at in-vehicle, the latency of between clients and WLAN will not be almost influenced. However, the latency between the vehicle and the ground station becomes much larger than usual since the propagation distance is long.

The throughput performance of a Web service with HTTP is restricted in such an extremely large latency environment [2-18][2-19]. To improve these issues, several TCP acceleration schemes [2-20] and protocol conversion schemes [2-21] have been studied to boost Web service performance. These schemes may accelerate throughput by enhancement of TCP, which improves congestion control algorithms. In applications using TCP, extending receive window size of clients can generally improve transfer speed under the condition of no packet losses. Data transfer completes within slow transfer rate by the slow start mechanism [2-22] of TCP when a small size content data in the WWW service. High transfer rate of TCP cannot be expected so enough in such short time transmissions. HTTP throughput is restricted up to about 105 kbps more than 3.6 k-Bytes of receive window size when a data size of web content to be 10 kBytes in the system[2-23]. Therefore, the HTTP performance by window size expansion and the band reinforcement of the satellite line does not improve.

It is pointed that increasing concurrent connection number as one of the HTTP performance enhancement schemes. Growing up of concurrent connection number makes performance deterioration due to the load for the WEB server [2-24]. In addition, concurrent connection number should be restricted in RFC 2616. Since the maximum number of concurrent connections from a single host process connecting via broadband to a single server has been increased to 6 with Internet Explorer 8 [2-25], approximately 500kbps as for the real HTTP throughput is estimating assuming the number of the concurrent connection is around 4. The effectiveness traffic of approximately 1Mbps per person can be found out from the recent trends of the broadband communication [2-26], users do not enough satisfy the HTTP performance with the mobile multimedia satellite communication system. Therefore, under unusual environments or system condition, the extremely large latency has a major influence on the throughput of a Web service, though it doesn’t influence usually.

2.3.5 The fourth issue: retransmission control scheme in large latency networks of high packet loss rate environment

Satellite communication systems of long distance transmission have been widely used in various applications and services. As a resulting in such a long distance transmission, it has been conventionally difficult to provide WEB services with high-speed data transmission owing to network latency. Efforts have been expanding to overcome this difficulty. User datagram protocol (UDP) packet usage is effective for real-time transmission. Consecutive UDP packet transmission enables efficient data transmission [2-27][2-28]. However, the UDP cannot support compensation for packet loss; hence a highly reliable scheme is essential for long-distance UDP wireless transmission.

TCP is a highly reliable communication protocol because of its flow control retransmission function [2-27] and [2-28]. This retransmission function uses a slow start procedure to decrease the transmission speed when packet loss occurs. TCP is highly efficient in multi-traffic flow environments, such as the Internet. However, this study assumes wireless relay by peer-to-peer networks, making TCP flow control unsuitable. Therefore, a retransmission function that is suitable for this condition is required. Hence, automatic repeat-request (ARQ) [2-29][2-30], which can reliably transmit data packets, is suitable for long-distance transmission.

ARQ is classified into three main categories: stop-and-wait (SAW), go-back-N (GBN), and selective-repeat (SR). All these types of ARQs have advantages and disadvantages. SAW is suitable for short-range transmission and environments with high bit errors, such as wireless LAN systems, because of its simplicity. GBN is employed in the TCP/IP employed for the world wide web (WWW). Unlike SAW and GBN, SR is generally used for long-distance wireless communications that have long delays, such as satellite systems. We adopted SR in this study because of our focus on systems having long delays owing to long-distance transmission.

Figure 2-8 shows a conventional SR transmission procedure. Acknowledgment (ACK) or selective ACK (SACK) messages are used in ARQ to detect whether or not a data packet is successfully received at a receiver. As shown in Figure 2-8, the data packets are successfully received in the conventional SR. The

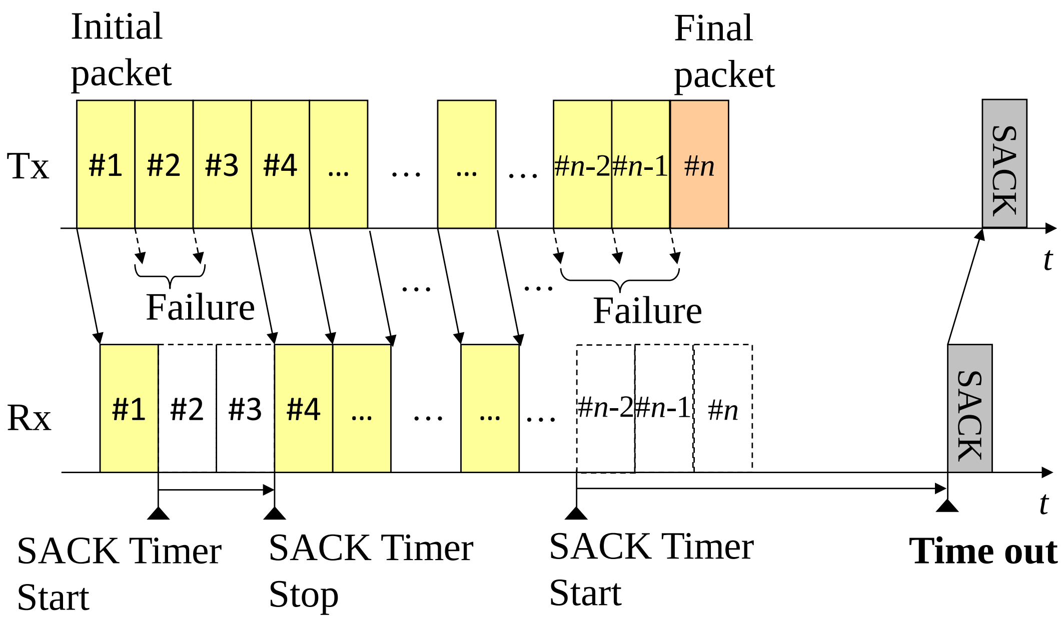


Figure 2-8 Conventional SR scheme.

receiver sends an SACK message, including the successful packet number information, to the transmitter after a certain time lag, called a ”Timeout”, when the success/failure of reception of the last data packet is detected, regardless of other packet losses. Therefore, the transmitter will retransmit only failure packets because the successful data packet numbers are transmitted at the transmitter by the SACK message. The receiver waits for a certain time before the SACK, which is retransmitted if appropriate packets are not received. Retransmission is not employed at the transmitter site unless the SACK is received after the transmission of the last data packet. The receiver must wait for a period corresponding to the Timeout when the final data packet cannot be received. Hence, a SACK cannot be instantaneously sent by the receiver site. This problem results in system delays and degradation of frequency utilization. Increased transmission times owing to retransmission control schemes must be avoided in long latency network systems.

References for chapter 2

[2-1]Hiroshi KAZAMA, and Fumiaki NAGASE,” RANDOM ACCESS CONTROL SCHEME FOR SPREADSPECTRUM FDMA SATELLITE COMMUNICATIONS SYSTEM,” PIMRC ’99, C1-1,1999.

[2-2]YAMASHITA FUMIHIRO, KOBAYASHI KIYOSHI, OHATA KOHEI, and UEBA MASAZUMI, “Seamless Symbol Rate Switching Scheme for Multi-Rate FDMA Modem,” IEICE Trans Commun., VOL.E89-B;NO.2;PAGE.545-555, 2006.

[2-3]M. Nakayama, M. Nakagawa, Y. Hashimoto, K. Tanaka and H. Nakashima, “A Satellite Communication System for Interactive Multimedia Networks,” IEICE Trans. On Commun. , Vol. E80-B, No. 1, pp. 103-108, January 1997.

[2-4]Toshinaga, H. Kobayashi, K. Ohata, K. Ueba, M. and Mizuno, H., ” Bi-directional multimedia satellite communication system employing superposed transmission,” Global Telecommunications Conference, 2001. Vol.4, GLOBECOM '01. IEEE, pp.2769 – pp.2773, 2001.

[2-5]M.H. Hadjitheodosiou, A. Ephremides, and D. Friedman, “Broadband access via satellite,” Computer Networks Volume 31, Issue 4, 25 February 1999, pp.353–pp.378, 1999.

[2-6]H. Tanaka, H. Kazama, M. Nakayama, T. Seki, F. Nagase, and H. Mizuno, “High-Speed Mobile Multimedia Satellite Communication System –Proposal of New Mobile System and its System Configuration-“, Technical Report of IEICE, SAT99-163(2000-02), pp.141-pp.146,2002.(in Japanese)

[2-7]YASUNAGA MASAYUKI, KIMURA YOSHIFUMI, IMATA SATOSHI, and CHIBA EIJI, “Overview of Recent Trends in Inmarsat System,” Technical report of IEICE, VOL.102, NO.548 (SAT2002 134-157), pp.61-pp.66, 2003.(in Japanese)

[2-8]Hajime MATSUMURA and Hisakazu KATOH, "Satellite Digital Broadcasting Systems," Broadcast Technology no.18, Spring 2004.

[2-9]T. Yoshida, N. Honma, H. Tanaka, T. Sueyoshi, T. Tsuchida, and H. Kazama, “Low-profile Auto-tracking Antenna System for Broadband Mobile Satellite Communication,” Ka and Broadband Communications Conference, Naples, Italy K000021, Sep. 2006.

[2-10] Fumiaki Nagase, Hiroshi Tanaka, and Masazumi Ueba, “IP Network Performance in Mobile Environments over Ku Band Satellite Links”, Proceedings of 2001 Korea-Japan Joint Conference on Satellite Communications (JC-SAT2001), pp.77-pp.82, Kyoto, Japan, Oct. 2001.

[2-11] Tomihiko Yoshida, Naoki Honma, Tadahiro Sueyoshi, and Toshihiro Tsuchida, “Development of a low-profile wideband antenna system for mobile satellite communication,” IEICE Tech. Report, SAT 2006-29, pp.67-72, 2006.

[2-12] Y. Minami, S. Aoshima, M. Morita, K. Nishikawa, and S. Inamura, ”Development of BS reception system for small car,” J. Society of Automotive Engineers of Japan, vol. 51, no. 2, pp. 59-63,1997.

[2-13] Fumiaki NAGASE, Hiroshi TANAKA, Jin MITSUGI, and Masazumi UEBA,” IP NETWORK PERFORMANCE OF A Ku BAND MOBILE MULTIMEDIA SATELLTE SYSTEM”,AIAA-paper-2002-2061, 20th AIAA ICSSC, May, 2002.

[2-14] Nagase, F., Mitsugi, J., Nakayama, M., and Ueba, M., “Ku band Mobile Multimedia Satellite Communications System for Trains,” AIAA-paper-2003-2205, ICSSC, 2003.

[2-15] Tomihiko Yoshida, Kiyoshi Kobayashi, Kohei Ohata, and Masazumi Ueba, “On-Frequency Gap-Filler using Adaptive Tracking Null Controller for Next Generation Mobile Satellite Communication Systems,” 2005 IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications”, pp.2257-pp.2261, 2005.

[2-16] T. Berners-Lee, R. Fielding and H. Frystyk, “Hypertext Transfer Protocol, HTTP/1.0.,” RFC 1945 MIT/LCS, UC Irvine, May 1996.

[2-17] R. Fielding, J. Getty P. Leach and T. Berners Lee, “Hypertext Transfer Protocol, HTTP/1.1,” Internet Draft Standard RFC 2616, June 1999.

[2-18] Hiroyasu Obata, Kenji Ishida, Junichi Funasaka, and Kitsutaro Amano, "Data Transfer Time by HTTP 1.0/1.1 on Asymmetric Networks Composed of Satellite and Terrestrial Links," IEICE transaction on communication, E85-B/ 12, pp.2895-pp.2903, 2002.

[2-19] Go Hasegawa, Masayuki Murata and Hideo Miyahara, “Performance Evaluation of HTTP/TCP on Asymmetric Networks,” International Journal of Communication Systems, Volume 12, Number 4, pp.281-296, July-August 1999.

[2-20] Takeshi Yoshino, Yutaka Sugawara, Katsushi Inagami, Junji Tamatsukuri, Mary Inaba, Kei Hiraki, ”Performance Optimization of TCP/IP over 10 Gigabit Ethernet by Precise Instrumentation,” in Proceedings of the 2008 ACM/IEEE conference on Supercomputing (SC08), IEEE Press, pp.1-12, Nov 2008.

[2-21] F. Nagase, H. Tanaka, M. Nakayama, T. Seki, H. Kazama, H. Mizuno," Mobile Multimedia

Satellite Communication System at Ku Band," IEICE trans. on Com. VOL.E84-B;NO.4;PAGE.903-909(2001).

[2-22] Stevens, W.R.,"TCP Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery Algorithms," RFC 2001, Jan. 1997.

[2-23] Fumiaki Nagase, Tetsu Sakata, “A study on HTTP performance acceleration with WEB prefetching in large latency network system,” IEICE Technical Report, RCS2011-47, pp.73-pp.78, Jun. 2011.

[2-24] Steve Souders ,"High Performance Web Sites: Essential Knowledge for Frontend Engineers," O'Reilly Media, Sep. 2007.

[2-25] "AJAX-Connectivity Enhancements in Internet Explorer 8," http://msdn.microsoft.com/en-us/library/cc304129(v=vs.85).aspx.

[2-26] "Aggregation and Preliminary Calculation of Internet Traffic in Japan -Publication of aggregate results as of November 2011-," Press Release-Telecom, Ministry of Internal Affairs and Communications, Japan, March 16, 2012.

[2-27] J. Nakamura, A. Makishima, D. Umehara, and M. Kawai, ”High-speed data transmission performance on Ka-band GEO-synchronous satellite communication link using TCP, “ 2002 Korea-Japan Joint Conference on Satellite Communications, Daejeon, Korea, pp. 33-38, Oct. 2002.

[2-28] K.Ienaga and M.Hiraishi, “A study of Fast File Transfer Protocol on PCs,” IEICE Technical Report, NS2002-102, pp.35-40, Sept. 2002.

[2-29] K.Hasegawa, and I.Sasase, ”A CONTROLLING RETRANSMISSION SIZE H-ARQ SCHME BY BLER BASED ON RESULTS OF CRC DECODING,” The 18th IEEE Annual International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC 2007), 54, Athens, Greece, Sept. 2007.

[2-30] S.Lin, D.Costello Jr., and M.Miller, “Automatic-repeat-request error-control schemes,” IEEE Commun.Mag., Vol.22, No.12, pp.5-17, Dec. 1984.

3. Proposal of Ku-band mobile multimedia satellite communication system with tracking antenna system

3.1 Summary

This paper presents a new satellite communication system that enables high-speed communication in a mobile environment. The system configuration combines a terrestrial mobile network and an existing satellite system, and includes a tracking antenna that was developed to receive 30Mbit/s signals from commercial communication satellites. A prototype system comprising the mobile network, the satellite system and a vehicle in which the tracking antenna is installed was constructed for purposes of evaluation and demonstration. A LAN system was achieved in the experimental vehicle by using the tracking antenna, a satellite router and a Personal Digital Cellular phone. The validity of the proposed system was verified in a driving test for the tracking antenna, system UDP tests, and FTP throughput tests in a mobile environment.

3.2 Introduction

Mobile communication systems have rapidly become widespread in recent years, and the number of subscribers of mobile terminals (Personal Digital Cellular (PDC) and Personal Handyphone System (PHS)) in Japan is now about 60 million [3-1]. One important user demand is for small mobile terminals that provide high-speed communication ability at low cost. This demand reflects the need for users to access large-capacity data and pictures in mobile environments. In response to this demand, faster and better mobile systems have been studied and developed, and the third generation mobile system, the so-called IMT2000, is scheduled to go into commercial operation in 2001.

According to the basic specification for the IMT2000 system determined by the International Telecommunication Union (ITU), the minimum transmission rate is 144 kbit/s in mobile environments such as in automobiles, trains and so on [3-2]. The actual rate is expected to range from 144 to around a couple of hundreds kbit/s. This speed, however, is considered to be insufficient for obtaining large-capacity data or for receiving moving pictures.

Against this background, the use of satellite systems has become popular, especially in the broadcasting field. The principal features of such systems are broad bandwidth and wide coverage area, which make them advantageous for the development of high-speed communication and multicast communication systems. As a matter of course, satellite systems can be easily expanded to mobile systems by using their characteristics as wireless systems. The ARIB (Association of Radio Industries and Businesses), an organization of broadcasting companies and equipment manufacturers, has studied the issues involved in standardizing the digital video broadcasting (DVB) system in Japan. Standardization will enable general users to economically purchase DVB system equipment such as tuners (receivers) and receiving antennas.

The predominant feature of multimedia communications at the current stage is that it is basically asymmetrical communications, in which access link signals such as requests or control signals (i.e., URL address information), or acknowledgment signals from user terminals are small- capacity data, while downlink signals from content servers, etc., including moving pictures, are large-capacity data (see Figure 3-1). By taking this feature into account, satellite multimedia communication systems that make use of a hybrid network consisting of a satellite for downlinks (from server to user terminal) and a terrestrial network (e.g., the Public Switched Telephone Network (PSTN), or the Integrated Services Digital Network (ISDN)) for access links (from user terminal to server) have been developed and commercialized [3-3]. These systems were made cost-effective through their utilization of a high-speed satellite network for downlinks and an economical terrestrial network for access links.

The applicable area of these current multimedia satellite communication systems could be expanded to mobile environments by making use of a mobile network. This could be accomplished by replacing the fixed parabola antenna on the roof of an office building or house, which is the current mode, with a tracking antenna placed on the roof of a vehicle.

This paper presents the concept of a new mobile satellite system, describes the system configuration and the newly developed tracking antenna that is a key component of the system. Antenna performance and system throughput test results are shown and the validity of the system is confirmed.

3.3 System Concept and Configuration

3.3.1System Concept

The basic configuration of the proposed mobile satellite communication system combines a land mobile network for access links and a satellite network for high-speed downlinks. The system concept is expansion of the applicable region of current commercial systems, presently limited to fixed use, to mobile environments. The main features of the proposed system are as follows:

(1) Quick, low-cost system construction is enabled by making full use of existing systems, i.e., land mobile networks and commercial satellite systems including such terrestrial facilities as transmitting earth stations.

(2) The problem of interference to adjacent orbiting satellites is avoided through the system's receiving

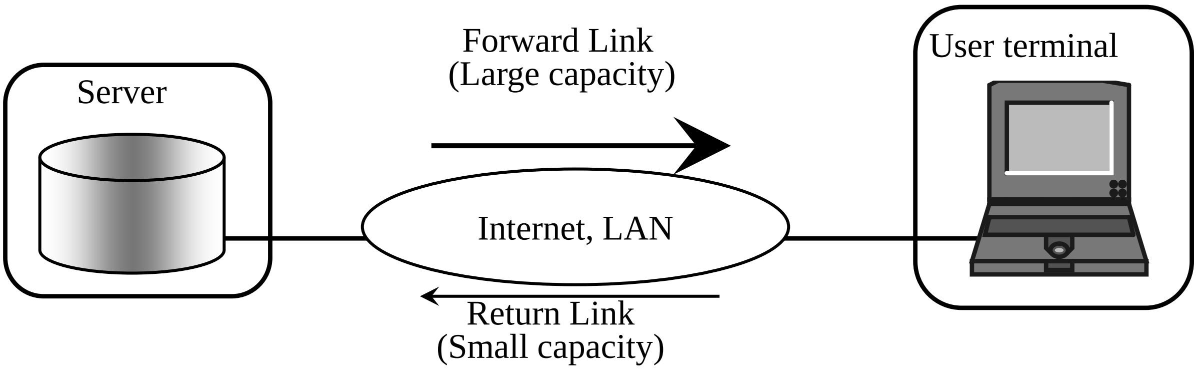


Figure 3-1 Feature of multimedia communications

antenna configuration.

(3) Low cost tracking antenna is made possible by only receiving function.

Figure 3-2 shows the system category that the proposed system falls under in the field of mobile communication systems. The x-axis is the downlink rate and the y-axis is the access link rate.

The satellite system used in the proposed system adopts the radio system developed for DVB to reduce the cost of user equipment. The radio system specifications are shown in Table 3-1. The proposed system receives signals of about 30 Mbit/s as an information rate. The dominant system throughput factors include satellite link latency, window size, and the protocol stack (TCP, UDP etc.) implemented in

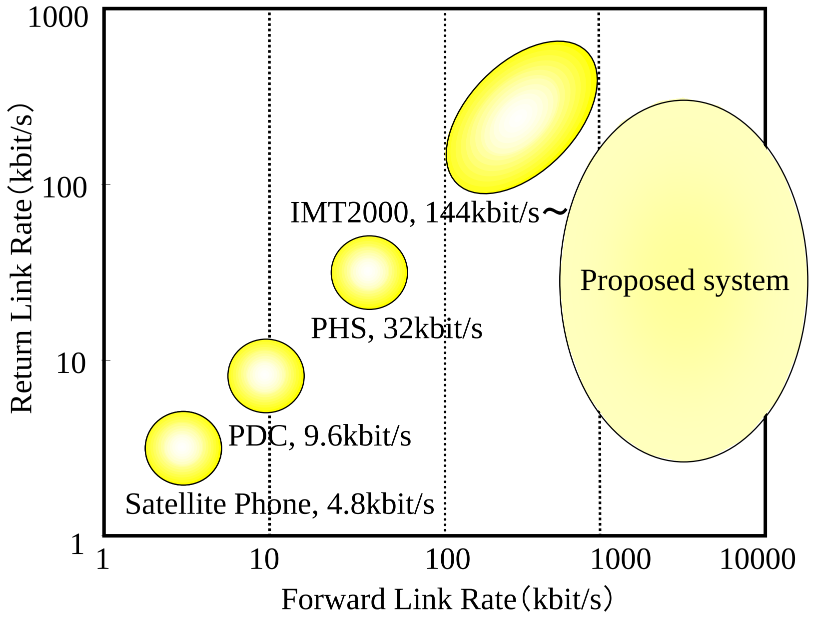


Figure 3-2 System category in mobile communication

Table 3-1 Forward link system specifications.

Item	Specification
Frequency band	Ku band(12.25-12.75GHz)
Modem	QPSK
Transmission rate	42.192Mbit/s
Information rate	29.162Mbit/s
Error correction	Outer coding: Reed Solomon(188/204) Inner coding: Convolutional encoding Viterbi decoding (3/4)
Framing	MPEG2 transport stream
Multiplexing	ATM cell multiplexing

the communication system. The proposed system is expected to achieve downlink throughput from several kbit/s to 30 Mbit/s, although the access link throughput is restricted by the capacity of existing mobile systems.

This system has unique features distinct from those of the present IMT2000 land mobile system and future land mobile systems. In particular, high-speed data transmission and multicast usage are considered to be promising applications for this system.

3.3.2System Configuration

Figure 3-3 shows the configuration of the prototype system constructed for evaluation and demonstration purposes. It comprises a vehicle with a tracking antenna, an Information Server Center (ISC), a Satellite Network Operation Center (NOC), and a Communication Satellite (CS). The NOC and the CS are now in commercial operation [3-4]. The downlink signal stream is indicated by a solid line and the access link stream by a dotted line.

The tracking antenna is for receiving satellite signals, and the PDC terminal is connected to the ISC by terrestrial (mobile and PSTN) networks. User PCs in the vehicle are connected to the satellite router, which receives satellite signals via the tracking antenna and sends access link signals of user PCs to the PDC.

The ISC is connected to the Internet, and the WWW, FTP and VOD servers in the ISC and Internet servers can be accessed from the user PCs to evaluate and demonstrate the system performance. An access server (AS) connects user PCs to each server via a terrestrial network, and the ATM router sends data

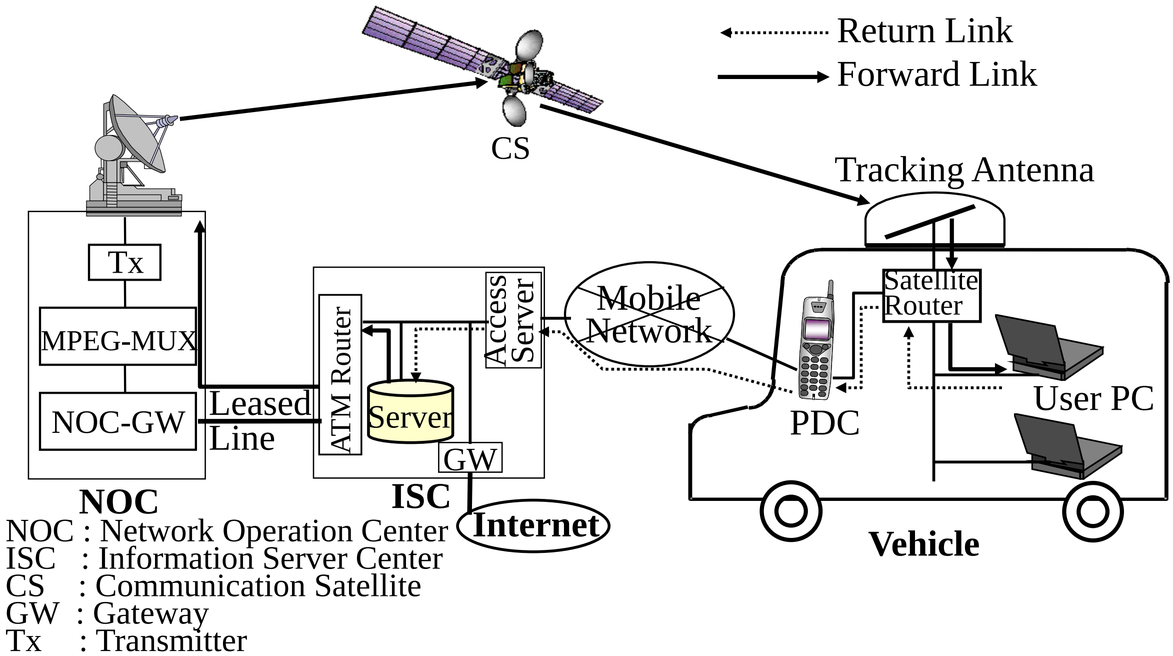


Figure 3-3 System configuration

of the server to the NOC via a leased line.

3.3.3Routing and Protocol

Downlink signals are sent to user PCs via a satellite router. Access link signals, such as requests and acknowledgements, from user PCs are sent to the ISC via the PDC, which is connected to the satellite router. The popular Point-to-Point protocol (PPP) is adopted to achieve IP communication.

Multimedia information in the ISC from each server and the Internet is sent to the NOC over a leased line. This system employs ATM cell multiplexing in the satellite link for data transmission at various speeds and protocols. IP packets are encapsulated into ATM cells at the ATM router and sent to the NOC. Each cell is assigned a Virtual Path Identifier/Virtual Channel Identifier (VPI/VCI) related to user PCs IP address and multicast IP address in the NOC.

This system adopts the radio system developed for the DVB, which adopts the MPEG2 transport stream for its satellite frame. Therefore, the multimedia data formatted into ATM cells are encapsulated into MPEG2 format data in the NOC and then sent to the satellite link [3-5], [3-6].

The signals received by the tracking antenna are sent to the user PCs via the satellite router. The satellite router demodulates the radio signals to MPEG2 format data, and extracts desired ATM cells by checking the VPI/VCI. The selected ATM cells are reassembled into IP packets, and then routed to the destination user PCs[3-7]. Figure 3-4 shows the protocol stack of the proposed system.

3.3.4LAN system in vehicle

Figure 3-5 shows the configuration of the LAN system in the vehicle. It consists of a tracking antenna, a PDC, a satellite router, a PDC modem, and user PCs. There is one subnet under a satellite router. The satellite router multiplexes access link signals from user PCs and sends them to the PDC via an RS232C interface. The satellite router receives the data from the tracking antenna via a coaxial cable

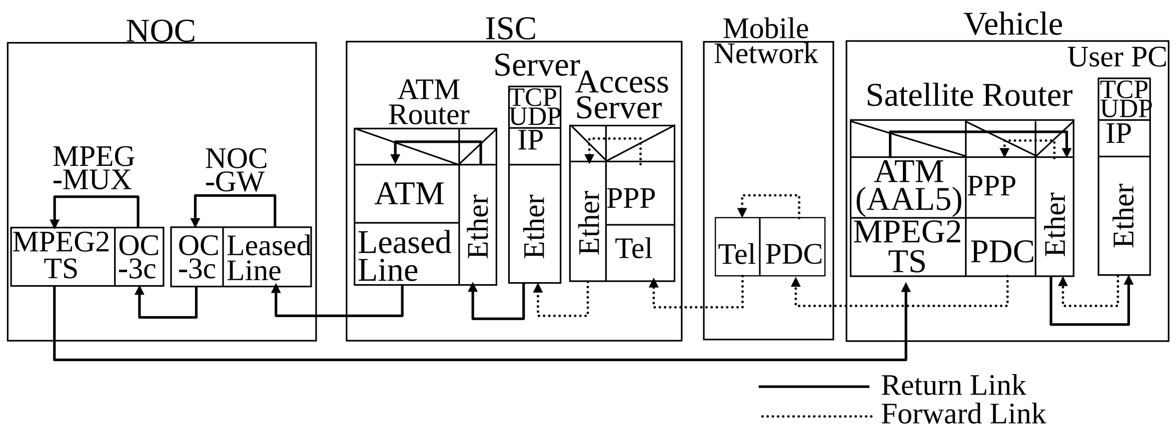


Figure 3-4 Protocol Stack

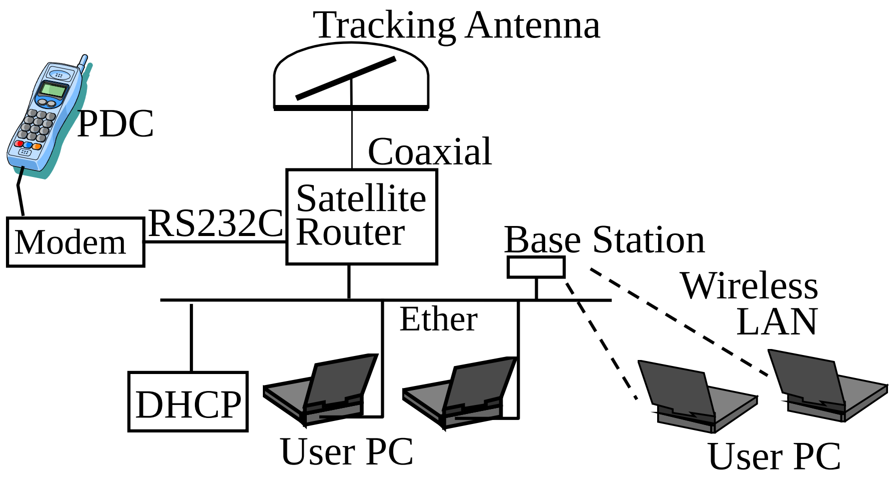


Figure 3-5 Configuration in the vehicle

interface. The satellite router has an Ethernet port as its LAN interface. User PCs are connected to the Ethernet or the base station of a wireless LAN. A wireless LAN system makes it possible to connect the satellite router to user PCs via a radio link and remove constraints on user PCs location. A DHCP server is included in the LAN system to enable automatic assignment of IP addresses of user PCs. This enables each user to connect to the servers and the Internet regardless of the network configuration.

3.4 Tracking Antenna Design

Some types of Ku band tracking antennas have already been developed for vehicle systems; examples include antennas developed for the Satellite News Gathering (SNG) system and for receiving Broadcast Satellite (BS) signals [3-8], [3-9]. The SNG antenna, however, is very expensive and large in size because of its receiving and transmitting functions and because it is a parabola antenna. Furthermore, its usage area is very restricted because a specially modified vehicle is required to house it.

Some BS tracking antennas are already commercially available. Their antenna gain requirement is small because the equivalent isotropically radiated power (EIRP) of BS is twice that of CS, which is the tracking target of the proposed system. To receive high rate DVB signals from CS, an antenna with higher gain than a BS tracking antenna is required. Since the beam width of a CS tracking antenna needs to be narrower than that of a BS antenna in order to achieve high gain, more accurate antenna tracking control is required to suppress the antenna gain reduction caused by tracking errors.

In addition, considering that the antenna is installed on the roof of a vehicle, the size of the antenna, especially its height, must be reduced as much as possible for the sake of appearance and to comply with Japanese traffic regulations.

The major requirements and specifications for the proposed system's tracking antenna are shown in Table 3-2. The radio link design imposes antenna gain of more than 32.0 dBi, since it takes DVB signal

Table 3-2 Main requirements, specification and features of tracking antenna

Requirement	Specification & Feature
High gain antenna	> 32.0 dBi in operation (including radome loss and gain reduction caused by tracking error)
Small size (Height dimension)	< 350 mm (Radome height) Microstrip antenna (32 x 16 elements)
Accurate tracking control	<1 dB gain reduction by tracking error

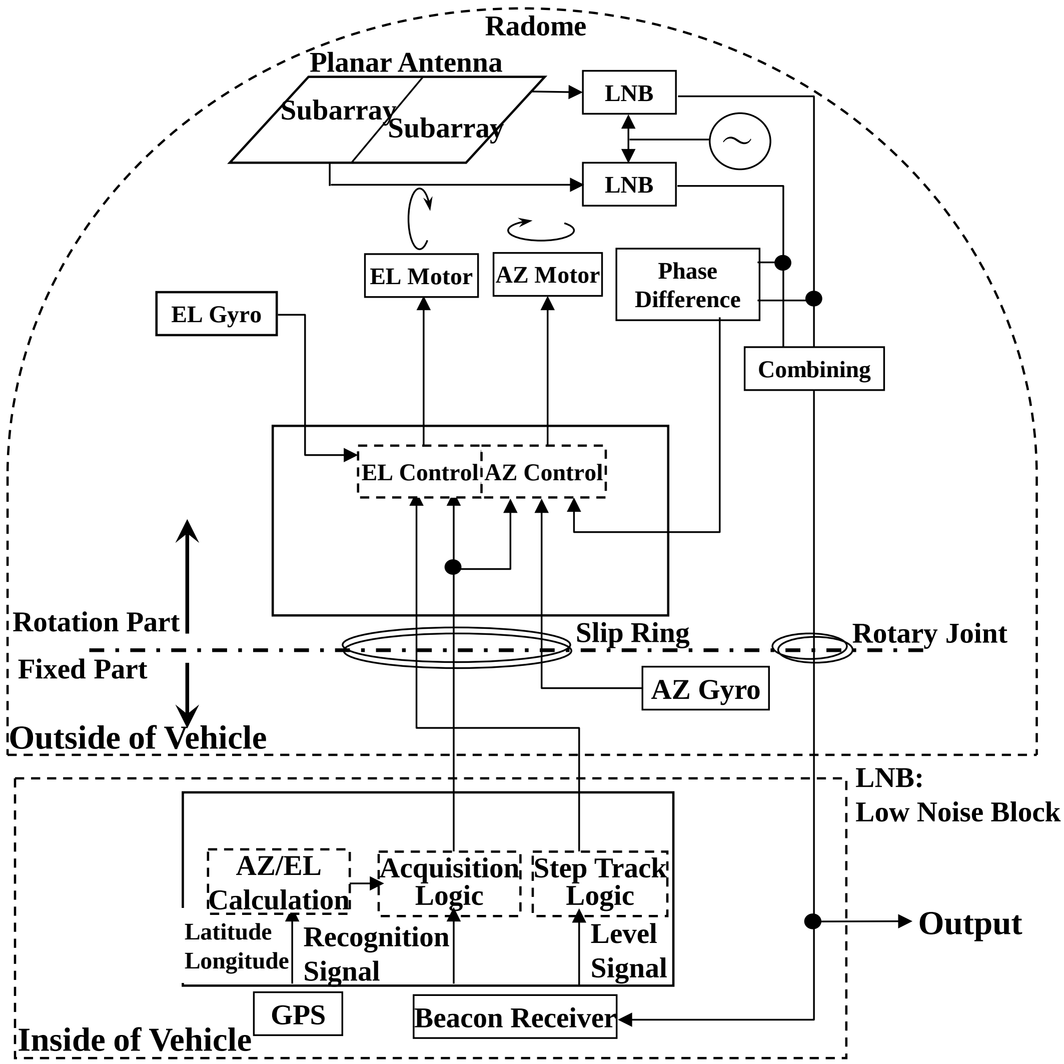


Figure 3-6 Basic tracking antenna configuration

reception into consideration, as well as antenna gain reduction due to tracking errors, radome loss etc. The vehicle height imposes an antenna height of less than 350 mm. Therefore, a planar antenna composed of a microstrip antenna is used in place of a parabola antenna to meet these requirements.

In order to maintain the radio link from the antenna to the satellite, antenna gain reduction due to tracking errors must be maintained at a maximum of 1 dB, and therefore quick, precise tracking control is required for the antenna azimuth angle because of the changes in vehicle direction that occur when it turns corners, for example. The planar antenna is divided into two sub-arrays to enable application of the phase monopulse technique, which detects tracking errors quickly and accurately. The conventional step track (hill climb) method is used for elevation control.

A co-phasing antenna scheme [3-10] is used for the azimuth control. A phase difference between the two arrays occurs when tracking errors occur, and the co-phasing antenna compensates for this difference by adjusting the phase shifter so that the output of the two sub-arrays is combined. This is one type of phase array technique. Use of this technique enables beam widths to be widened equivalently and gain reduction from tracking errors to be suppressed. Consequently, the designed antenna is able to track the targeted satellite both mechanically and electrically.

Since the shadowing environment caused by tunnels, buildings and trees is inevitable in practical use, this antenna system has inertia (gyro) sensors to detect vehicle attitude changes. These sensors control antenna direction in a shadowing environment and enable the targeted satellite to be easily recaptured after the shadowing conditions cease. Figure 3-6 shows the antenna system configuration. Antenna radome size of 1000 mm diameter and 330 mm height is achieved. Antenna radome measures 1000mm in diameter and 400mm in height as shown in Figure 3-7.

3.5 System Evaluation

3.5.1Antenna Performance

Receiving level was measured by rotating the planar antenna direction manually in order to confirm the effects of the co-phasing combining technique. Figure 3-8 shows the receiving level variation, which can be regarded as antenna gain reduction due to tracking errors. The solid line is the result obtained with co-phasing and the dotted line is that obtained without co-phasing, i.e., with no phase adjustment. From the figure, it can be seen that the receiving level deterioration can be suppressed and that equivalent spread of antenna beam width is attained by compensating for the phase difference between the two sub-arrays caused by tracking errors. It can also be seen that the required tracking accuracy is somewhat weakened by the co-phasing scheme.

It is considered that most antenna tracking errors occur when the vehicle is negotiating curves in the road. The number of tracking errors occurring was measured in a driving test that was carried out on a test course. In the course, there were no objects that would impart shadowing conditions to the track-

Figure 3-7 tracking antenna.

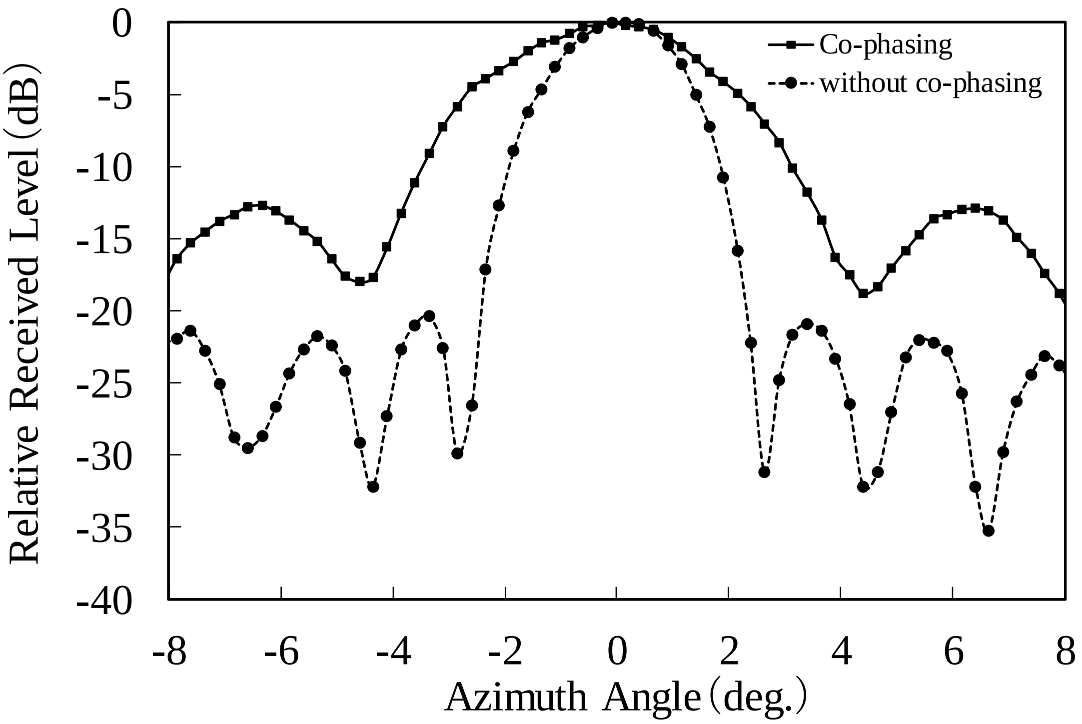


Figure 3-8 Effect of co-phasing combining.

Figure 3-9 Experimented vehicle housing tracking antenna.

ing antenna. In this experiment, a radius of 30 m and velocity of 20 km/h rotation traveling were obtained by assuming one set of practical driving conditions. The test vehicle housing the tracking antenna is shown in Figure 3-9.

Tracking antenna performance tests were carried out by road tests. The measured tracking error probabilities with the laboratory model obtained from the road test are shown in Figure 3-10. The unshaded bars and the shaded bars are tracking error probability of the azimuth direction and the elevation direction, respectively, at every 0.1 degree step. The tracking error was 0.42 degree rms. in the azimuth direction and 0.13 degree rms. in the elevation direction. It was found that these accuracies achieve tracking performance better than the pointing loss requirement, 1dB.

The tracking error of the low-cost model was also measured under emulated vehicle running conditions. The results were 0.80 degree rms. in the azimuth direction and 0.25 degree rms. in the elevation direction. It was confirmed that these accuracies also meet the pointing loss requirement, 1dB

Figure 3-11 shows antenna tracking errors and vehicle direction changes. The tracking errors can be derived from the phase differences between received satellite signals in the two sub-arrays. Vehicle direction changes, that is, azimuth angle changes, are obtained from the output of the inertia (gyro) sensor that is used in the tracking antenna. The vehicle's angular velocity was measured at about 10 deg/sec. The tracking error is 0.4 deg. RMS., and the gain reduction due tracking errors is within 1 dB from the pattern shown in Figure 3-8. The antenna gain of 34.6 dBi was obtained by static antenna tests. These results confirm that the required antenna gain of more than 32.0 dBi can be ensured under vehicle operati-

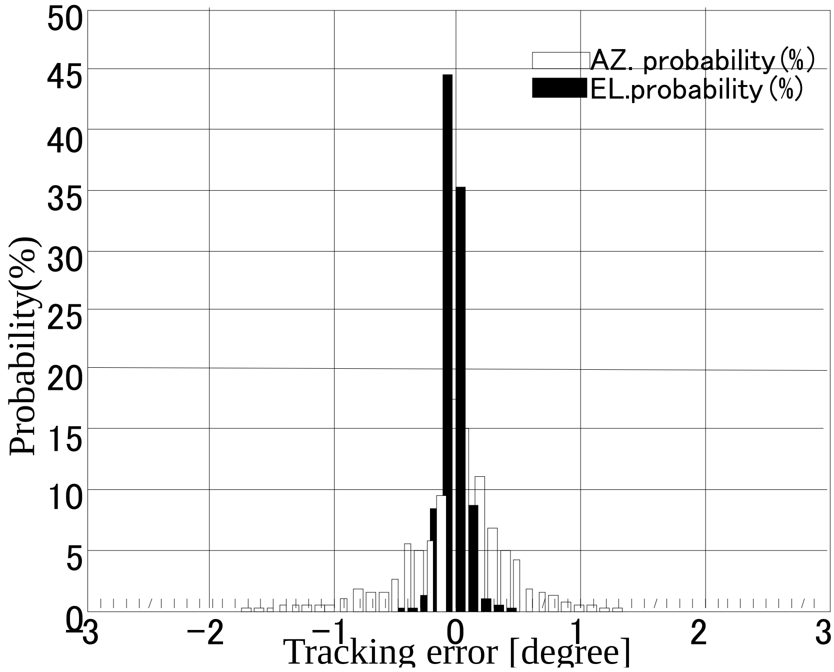


Figure 3-10 Tracking error probability.

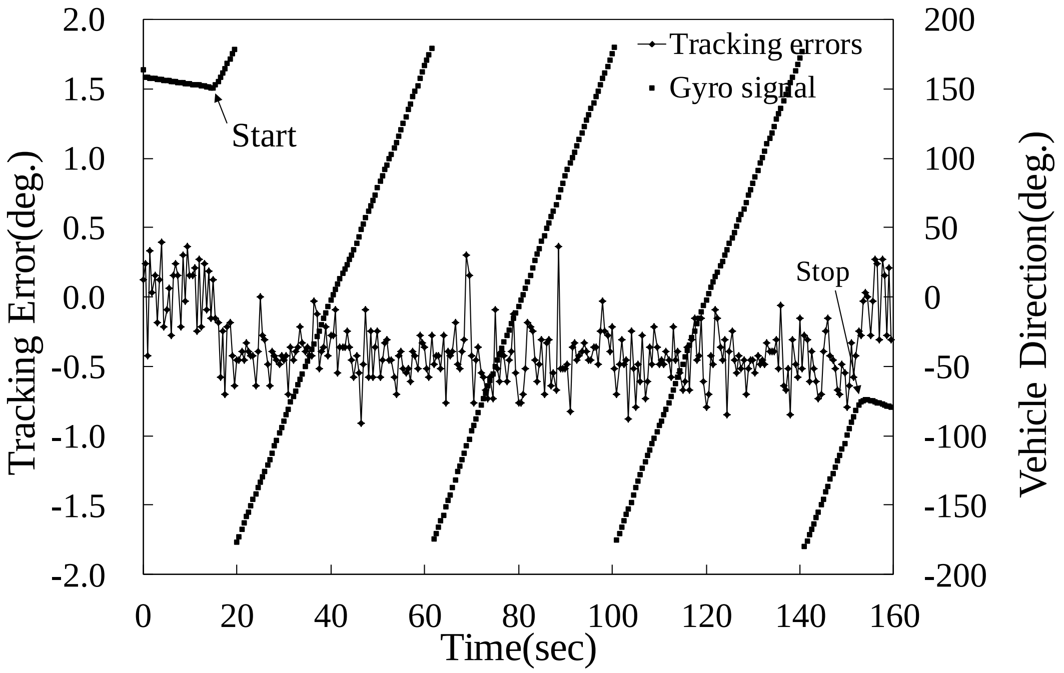


Figure 3-11 Tracking error and vehicle motion.

ng conditions.

On the other hand, the velocity in the EL-direction changes less than in the AZ-direction. It is verified that the tracking antenna can receive the satellite signals without any problems on the roads with slopes.

3.5.2Round Trip Time

To predict TCP throughput performance, ping (ICMP echo) commands are continuously issued from the user PC to the server to measure the round trip time of the system.

Round trip time is the sum of the fixed delay time of the router transactions, satellite link delay time, and transmission delay time by packet size. Long packets increase round trip time. In this system, round trip time is affected by the ping packets size on the low speed return link. Accordingly, ping packet size was chosen that was the same as TCP Ack packet size. Round trip time was measured by 300 pings in stationary environments.

Figure 3-12 shows round trip time variation. The average round trip time by PDC, PDC-Packet, satellite phone, and PSTN (fixed terrestrial networks) for the return link was 504, 434, 936, and 322 msec, respectively. Round trip time varies very little when PSTN is used as the return link. It was found that round trip time with a cellular network for the return link varies more than with PSTN for the return link. Furthermore, it was found that round trip time may become unexpectedly long when using the cellular network for the return link.

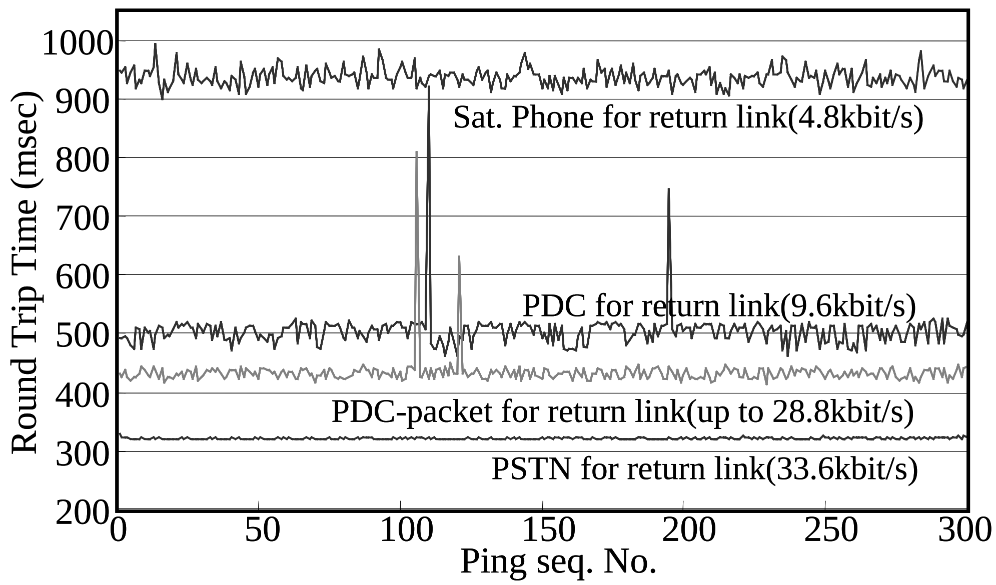


Figure 3-12 Round trip time.

3.5.3TCP performance in stationary environments

To verify the system's TCP performance in stationary environments, we measured TCP throughput by transferring a file in the server to a user PC by FTP.

Figure 3-13 shows TCP throughput in stationary environments. The x-axis is the user PC window size and the y-axis is TCP throughput. The PC window size was 8, 16, 32, and 64 kbytes. Without retransmission, the theoretical TCP throughput P [bit/s] is given by the following equation.

$$P = 8 \cdot \frac{W}{Trt} \qquad (1)$$

where W [byte] is the user PC window size and Trt [s] is the round trip time [3-11].

The measured results with PSTN and PDC for the return link show almost completely coincidence with the theoretical throughputs. The low speed (4.8 kbit/s) the satellite phone used as the return link may degrade a lower measurement value in a large window size. When using the PDC-Packet for the return link, the measurement value was 75 % of the theoretical one. An investigation to determine the reasons for the difference between the theoretical and experimental results remains as a subject for further study.

Figure 3-14 shows TCP throughput variation with a 64kbytes user PC window size. At the beginning of the file transfer, TCP throughput increases by Slow-Start TCP algorithms and is maintained with a little variation. It is shown that the round trip time variation by the return link causes this TCP throughput variation.

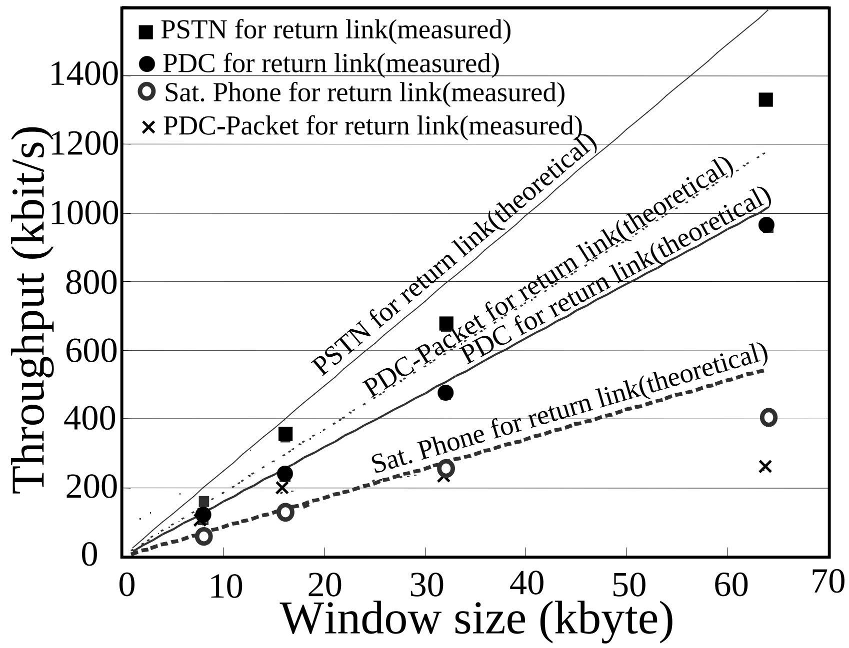


Figure 3-13 TCP throughput.

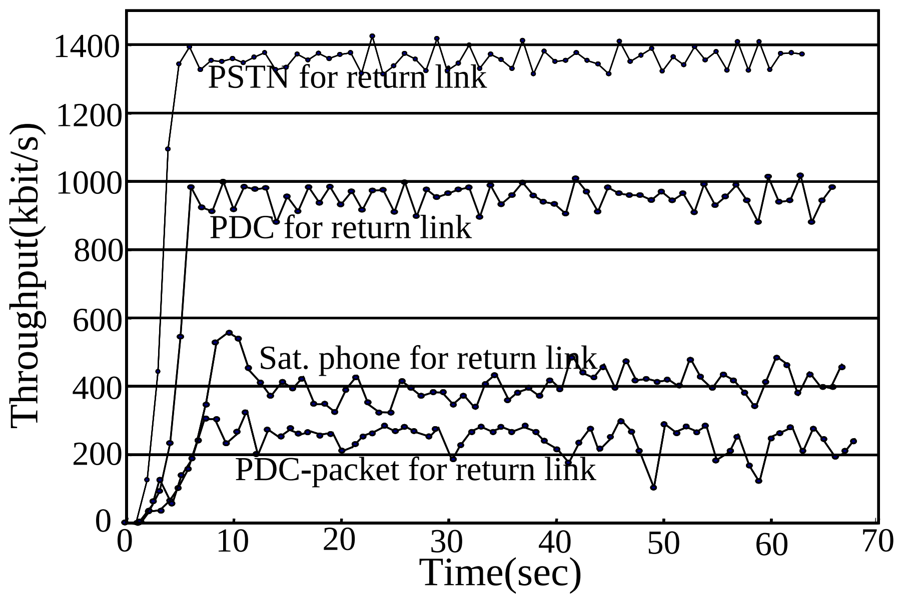


Figure 3-14 TCP throughput variation.

3.5.4XTP performance in stationary environments

XTP (Xpress Transfer Protocol) is designed to boost the throughput of long latency networks such as satellite communication network. In this section, to measure the system's XTP performance in stationary environments, we measured XTP throughput by FTP with XTP protocol gateways. The gateways were inserted into the communication path, with one unit on either side of the ISC and the vehicle. On either end of the gateway, the TCP end-to-end connection is broken and XTP is optimized for satellite links by using mechanisms that use negative acknowledgements, unlimited window sizes, and a streamlined handshake algorithm. XTP also eliminates congestion avoidance mechanisms such as Slow-Start algorithms and has a functionality to match the transmission rate to the available bandwidth. The results of throughput measurement are shown in Table 3-3. The upper-limit throughput in this experiment, which is about 1.7 Mbit/s, depends on the ATM leased-line speed of 2 Mbit/s between the ISC and the NOC. With the measurement, it can be confirmed that the throughput was improved over that of TCP and nearly reached the network limit regardless of kinds of the return link.

Table 3-3 XTP throughput.

Return Link	Throughput
PDC	1680kbit/s
PDC-Packet	1510kbit/s
Satellite Phone	1650kbit/s
PSTN	1690kbit/s

3.5.5TCP performance in mobile environments

To evaluate the system's TCP performance in mobile environments, we carried out FTP while the vehicle was running with PDC for the return link. The tracking antenna mounted on the vehicle can accurately receive the satellite signal in mobile environments [3-12]. Figure 3-15 shows a variation of the satellite signal level received by the tracking antenna. The figure indicates a shadowing environment caused by buildings, trees, tunnels etc., in which the tracking antenna can not receive satellite signals.

Figure 3-16 shows TCP throughput variation with a 64kbyte user PC window size. It was found that throughput variation occurred in this case, and we considered that it was due to the effects of shadowing on the tracking antenna, the PDC link conditions, as well as other factors. We could obtain an average throughput of about 380 kbit/s. The throughput fell to zero when the satellite signals could not be received, and that throughput recovered the same value as in stationary environments when satellite signals could be received again.

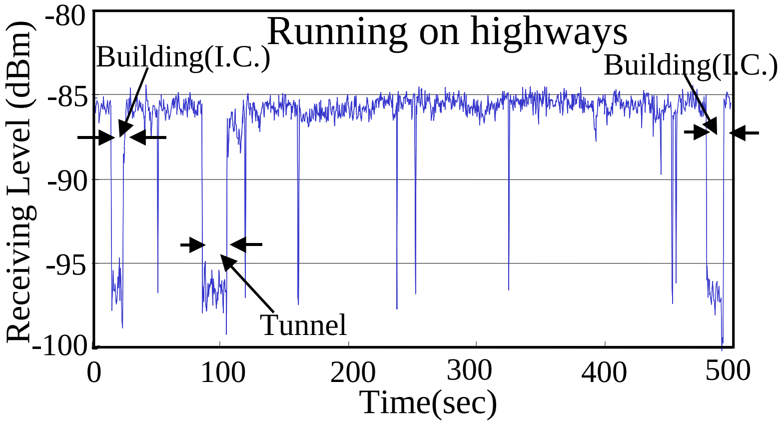


Figure 3-15 Receiving signal level.

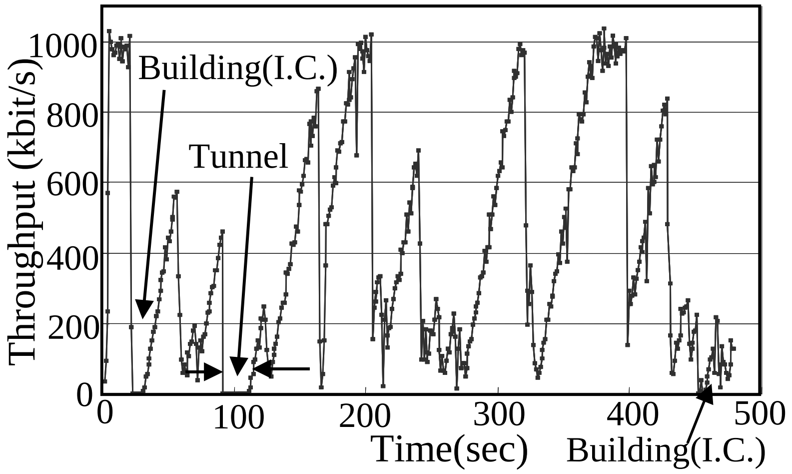


Figure 3-16 TCP throughput variation.

3.5.6XTP throughput in mobile environments

To evaluate the system's XTP performance in mobile environments, we carried out XTP throughput measurement by using FTP in the running vehicle, using PDC for the return link as same conditions as the TCP performance tests.

Figure 3-17 shows XTP throughput variation. We could obtain an average throughput of about 1490 kbit/s. The throughput dropped to zero when the satellite signals could not be received, and that throughput recovered rapidly to the same value as in stationary environments just after satellite signals could be received again. Thus, it was clarified that the XTP is effective even in mobile environments.

3.5.7UDP Performance

In order to verify client receiving rate from server, we evaluated the system's UDP/IP performance. The UDP/IP protocol is widely used for stream applications such as moving picture transmission. With this protocol, UDP/IP packets are sent from the server in the ISC to the user PCs in the vehicle via the NOC by satellite link. In the measurement performed, UDP/IP packets were sent at the rate of 1.5 Mbit/s due to the system restrictions imposed by the leased line rate between the ISC and the NOC.

Figure 3-18 shows the results obtained in the UDP throughput performance test. This test was carried out while the vehicle was traveling on a highway at about 80 km/h. The results obtained made it clear that the time span of throughput deterioration coincides with that of receiving level drop, and that receiving throughput identical to the server transmission rate is obtained except when level deterioration is received. The receiving level drop is due to the shadowing caused by buildings and trees in this experiment. It was confirmed that the tracking antenna can receive all packets in a mobile environment except under shadowing conditions. Thus the validity of the network system design was confirmed in this experiment.

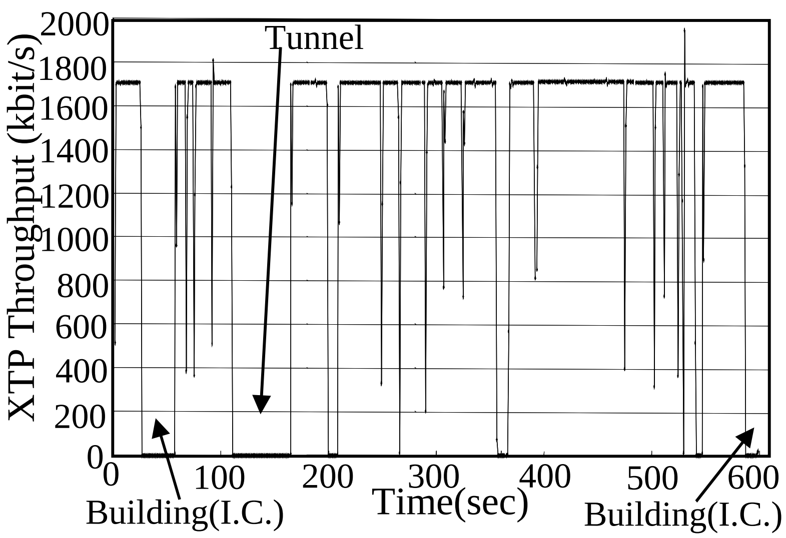


Figure 3-17 XTP throughput variation.

3.5.8FTP Performance

In order to clarify the system's FTP performance, we evaluated the FTP throughput. FTP is used as a TCP application. A Windows NT Server (Version a 4.0) and a PC running Windows98 were used as an FTP server and user terminal respectively. The TCP reception window size of both the server and the user terminal was set at 64 kbytes. Taking the slow start of the TCP procedure into account, it was decided to use a file size of about 2.4 Mbytes for throughput measurement.

About 1 Mbit/s throughput is expected from window size and round trip time that was measured by Internet Control Message Protocol (ICMP) echo. We obtained 842 kbit/s as maximum FTP throughput for both mobile and stationary conditions in our experiments, which were carried out under the same mobile conditions as those in the UDP throughput test. This throughput is affected by the performance of network equipment including the PC and the implemented protocols. It was found that throughput variation occurred in the experiments under mobile conditions. This variation is considered to be due to the effects of shadowing on the tracking antenna, and to the PDC link conditions as well as other factors. It was confirmed that the FTP process was completed without disconnection occurring even under shadowing conditions.

3.6 Conclusion

This paper summarizes a new system for mobile multimedia satellite communications. First, the

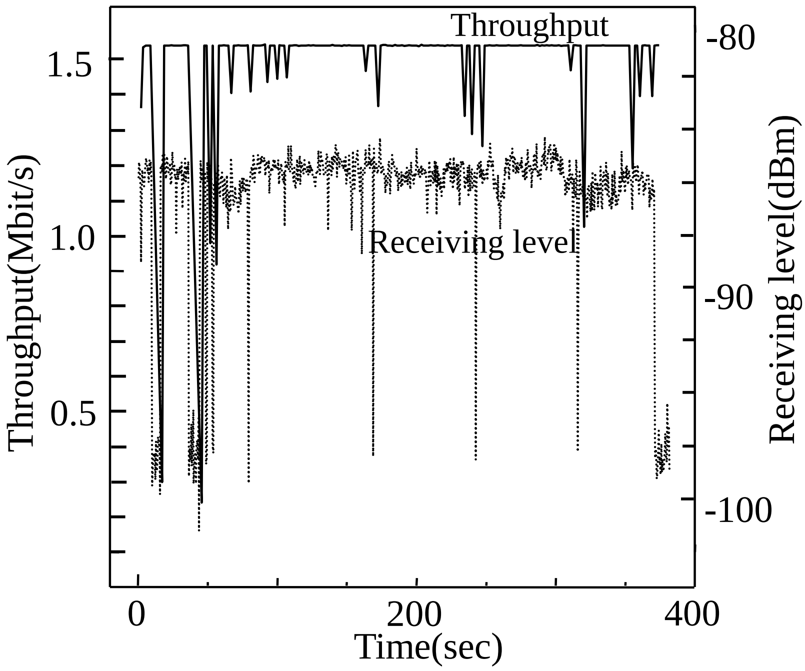


Figure 3-18 UDP/IP performance.

system concept and system configuration comprising a terrestrial mobile network- satellite system combination was described. Also described were the requirements for the tracking antenna installed on the vehicle to receive digital video broadcasting signals (30 Mbit/s) and the antenna configuration.

An experimental system using the newly developed tracking antenna was constructed to evaluate and demonstrate the system. Results obtained in a driving test confirmed that gain reduction caused by tracking errors was suppressed through the use of a co-phasing technique and that the required antenna gain can be achieved in a mobile environment. The validity of the proposed system was verified in a UDP throughput test carried out in a mobile environment. The sending rate (1.5 Mbit/s) from the server coincided with the signal reception rate of the user terminal in this test.

Finally, an FTP throughput test was carried in which throughput of about 850 Kbit/s was obtained under mobile conditions. It was verified that the proposed system can achieve a high communication rate in a mobile environment by making use of the Ku band satellite system. We believe that the proposed system can make a significant contribution to the expansion of the applicable area of satellite communications, i.e., from fixed to mobile usage, and to the ongoing evolution of multimedia communications.

References for chapter 3

[3-1]http://www.tca.or.jp/index.html

[3-2]http://www.mpt.go.jp/pressrelease/japanese/denki/990927j603.html

[3-3]M. Nakayama, M. Nakagawa, Y. Hashimoto, K. Tanaka and H. Nakashima, ’A Satellite Communication System for Multimedia Network,’ IEICE Trans. Commun., VOL. E80-B, NO.1, pp103-108, January 1997.

[3-4]http://www.megawavepro.ne.jp/

[3-5]M. Nakagawa, Y. Hashimoto, and N. Ikeuchi, ‘QoS Evaluation of a Multimedia Interactive Satellite Communications System,’ IEICE Trans. Commun., VOL.E81-B, NO.11, pp2034-2040, November 1998.

[3-6]M. Nakayama, K. Ohata, T. Otsu, H. Kazama, and K. Araki, ‘Portable Interactive Multimedia Satellite Communication System,’ proc. of IAF conference, IAF-98-M.2.07

[3-7]M. Nakayama, S. Yoshino, K. Nidaira, and H. Nakashima, ‘Protocol stack for multimedia interactive satellite communications system,’ Proc. IEICE General Conf., B-266, 1996

[3-8]Y. Minami, S. Aoshima, M. Morita, K. Nishikawa and S. Inamuera, ‘Development of BS Reception System for Small Car,’ Car Technical Journal, Vo.51, No.2, pp.59-63, 1997

[3-9]http://www.datrondbs.com/

[3-10] W. Gabriel, ‘Adaptive Arrays - An Introduction’, Proc. of the IEEE, VOL.64, No.2, pp.239-273, Feb. 1976

[3-11] M. Nakagawa, Y. Hashimoto, and N. Ikeuchi, “QoS Evaluation of a Multimedia Interactive Sat-

ellite Communications System," IEICE Trans. Commun., vol. E81-B, no. 11, pp. 2034-2340, Nov, 1998.

[3-12] F. Nagase, H. Tanaka, M. Nakayama, T. Seki, H. Kazama, and H. Mizuno, "Mobile Multimedia Satellite Communication System at Ku Band," IEICE Trans. Commun., vol. E84-B, no. 4, pp. 903-909, Apr. 2001.

4. Proposal of Layer-3 diversity reception techniques

4.1 Summary

We describe a layer 3 diversity reception scheme that enhances the transmission characteristics of Ku-band mobile satellite communication systems. This scheme can realize high-speed communication for vehicles that experience shadowing caused by terrestrial obstacles such as tunnels, buildings and bridges, especially for trains that frequently experience shadowing from the trolley wire structures. Layer 3 diversity was chosen for long distance diversity to prevent signal shadowing caused by terrestrial obstacles while minimizing the alterations of existing receivers. The technology enables high-speed communication under shadowing conditions in a running train environment.

4.2 Introduction

Mobile communication systems have rapidly come into widespread use, and customers have begun to demand higher speed multimedia communication services in mobile environments, particularly those in commuting and/or moving vehicles. In response to these demands, various experiments for such services have been conducted recently, with respect to high-speed Internet access [4-1], [4-2], [4-3], [4-4]. Mobile satellite communication systems are a powerful way of economically providing high-speed multimedia services to vehicles by virtue of their instantaneous deployment with wide coverage area. The predominant feature of the Internet traffic model is its basic asymmetry. In the return link, the user transmits a small amount of data to the server that includes requests, control signals, acknowledgement signals, and so on. On the other hand, in the forward link, the server transmits large amounts of data such as still images and videos.

In this regard, we have developed a Ku band mobile satellite communication system [4-5], [4-6] that implements the return link on terrestrial mobile networks and the forward link on high-speed Ku-band satellite networks. A large amount of data can be transmitted to vehicles via the Ku-band satellite network link. With this system, the user employs a satellite-tracking antenna, mounted on the vehicle, to receive the satellite signals. Figure 4-1 shows the Ku-band mobile satellite communication system. The system is cost effective because it uses an existing multimedia satellite communication system [4-7], which uses a fixed terrestrial network for the return link, and a Ku-band satellite network with a fixed antenna for the forward link. The system has been shown to be able to provide a high-speed Internet environment for automobiles and vessels [4-8].

Signals, however, may suffer shadowing caused by terrestrial obstacles, such as tunnels, buildings, bridges, trees, and other natural obstacles. Shadowing degrades the network throughput performance. Diversity reception with two or more antennas is effective for offsetting shadowing [4-9]. The initial thrust for developing diversity reception is to implement the Ku band mobile satellite communication system in trains, which experience severe shadowing, particularly periodic and frequent shadowing caused by the trolley wire structures. In section 2, we describe the shadowing caused by the terrestrial obstacles and

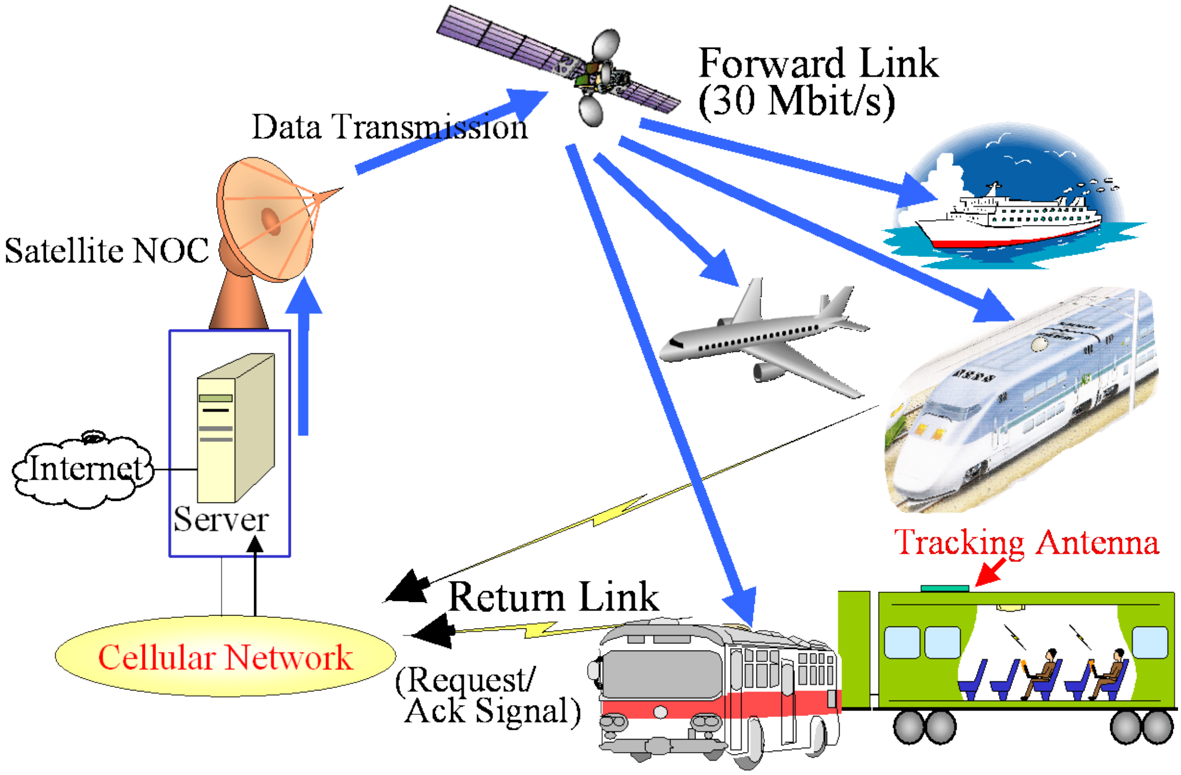


Figure 4-1 The Ku band mobile satellite communication system.

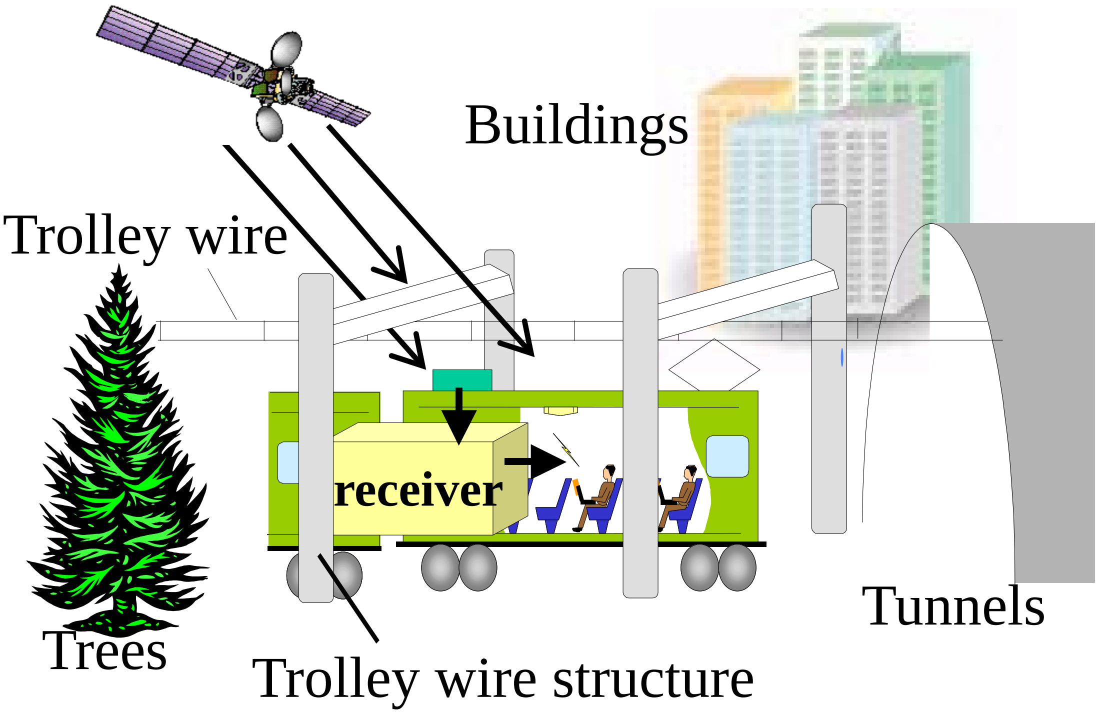


Figure 4-2 Shadowing environment around trains.

throughput on trains. The requirements for the diversity reception system are reviewed in section 3. Layer 3 diversity reception technologies are proposed in section 4. Throughput improvements using the proposed layer 3 technology are presented in section 5.

4.3 Shadowing and throughput on trains

There are many obstacles along railways that cause shadowing during normal train operations, as shown in Figure 4-2. Because the train must receive satellite signals directly, these obstacles greatly degrade network performance. The shadowing caused by tunnels or buildings is mainly countered using a gap-filler system [4-10], [4-11]. Peculiar to the network performance degradation on trains is the structures supporting trolley wires.

In order to clarify the shadowing environment and the influence of the trolley wire structures on network performance, we conducted a field test using the JR lines. The received signal levels were measured at the communication satellite beacon, “N-STAR A” every 5 milliseconds.

Figure 4-3 shows the degradation of the total XTP throughput between stations, and the distance metric caused by the buildings and the trolley wires structures. As can be easily seen, it is impossible to achieve upper limit throughput and, on average, less than half the throughput can be achieved. Figure 4-4 shows the received signal levels, as well as the packet loss caused only by the trolley wire structures. The received signal level was greatly degraded by 7 to 8 dBm, and the communication link was frequently broken. Therefore, a large packet loss is generated and greatly degrades network performance.

To combat the performance degradation caused by this shadowing, we propose a diversity reception technology with multiple satellite tracking antennas on the train roof.

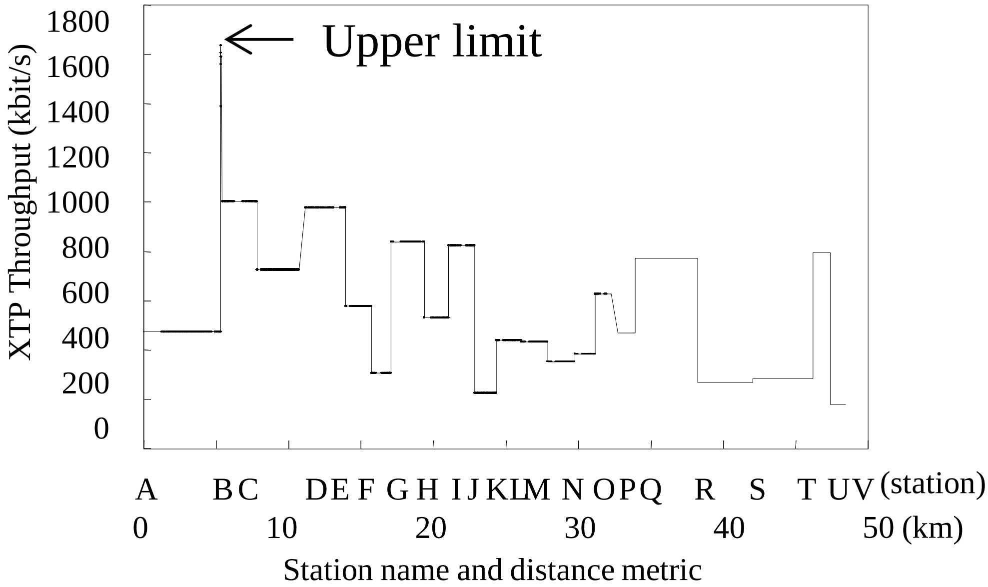


Figure 4-3 Degradation of total XTP throughput.

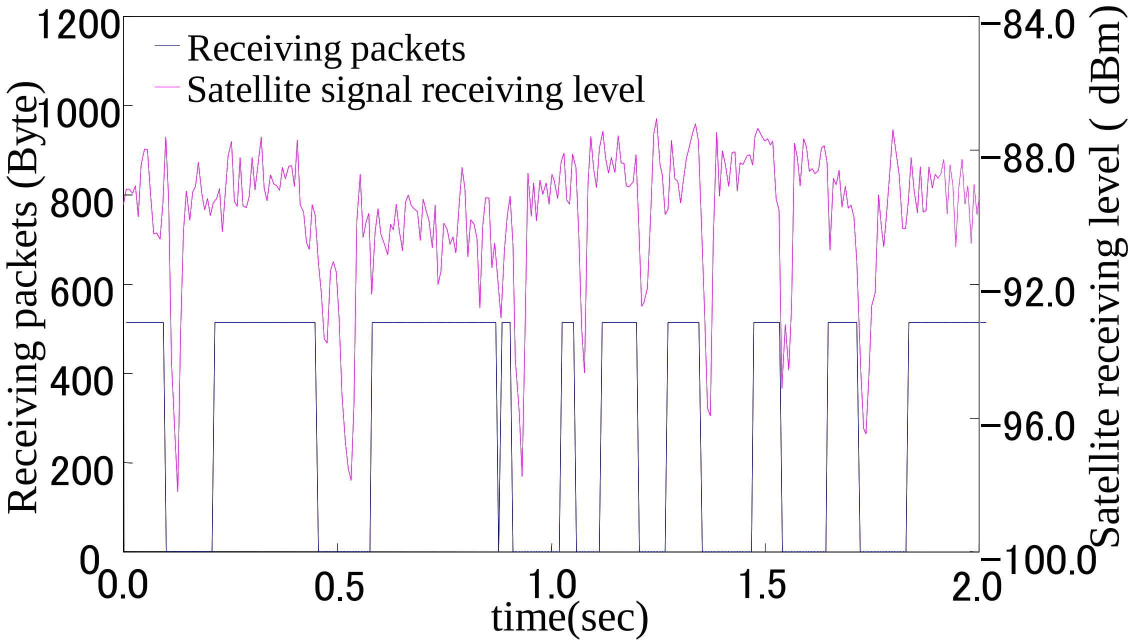


Figure 4-4 Received signal level and receiving packets on a live train.

4.4 Conventional diversity reception and requirements

4.4.1 Correction capability limitations

First, we clarified the correction capability limitations to counter the shadowing caused by the trolley wire structures. Then, for shadowing beyond the correction capability, we clarified the limitations of conventional diversity reception technologies.

Because packets are transmitted on a DVB frame, conforming to Telecommunications Technology Council standards for digital broadcasting: of a 4Mbit/sec band share system, it takes about 5.2 msec to transmit a 1,500-byte packet. Figure 4-5 shows the width of trolley wire structure that can correct packet errors. Packet errors can be recovered within 0.052 msec of shadowing, which corresponds to a bit error rate of 10-2 [4-12]. You can see that because the width of trolley wire structures is about 0.3 to 0.5 meters, packets are lost by shadowing even when the train runs at speeds of 300 km/h.

4.4.2Diversity reception requirements

Requirements for diversity reception can be summarized as follows.

(i) It must adapt to signal symbol timing differences, caused by the diversity distance and the running direction of the train: Because the shadowing influence can be reduced by lengthening the diversity distance, which is effective for avoiding network performance degradation caused by shadowing from terrestrial obstacles. However, this may generate a large symbol difference.

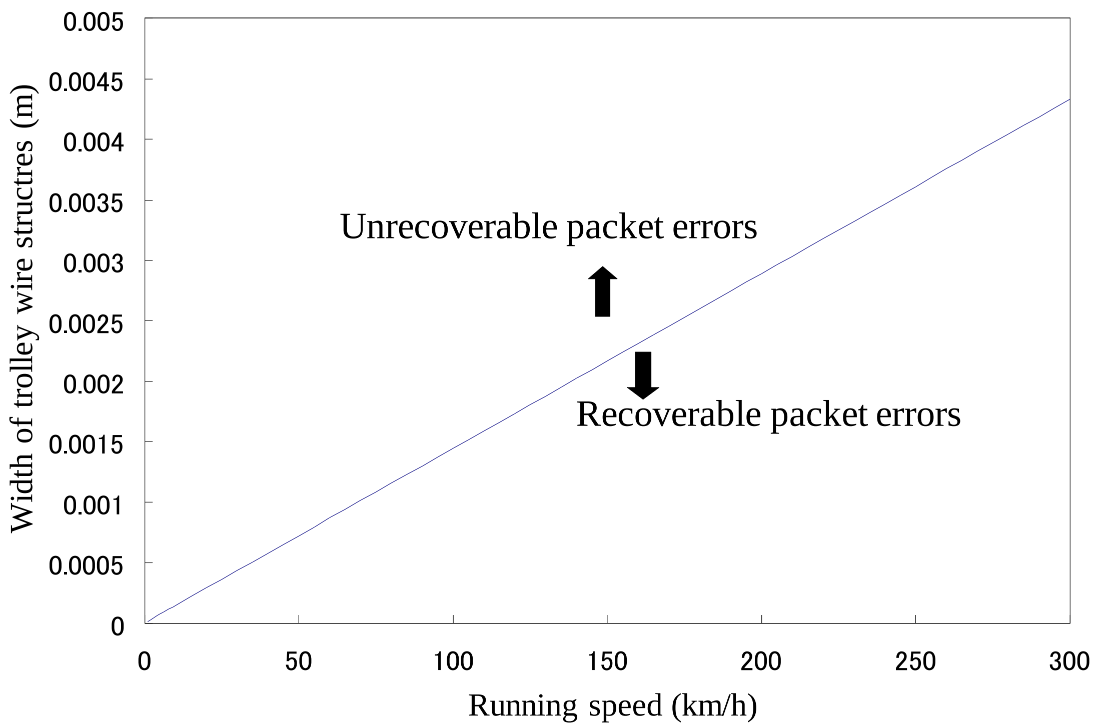


Figure 4-5 Packet error limitations.

(ii) It must furnish seamless connectivity with gap-fillers: Some kind of gap-filler should be used to provide network connectivity in shadowing areas where satellite signals cannot be directly received, such as in tunnels, longer than the train. Ideally, the diversity reception can be applied to both the satellite and gap-filler systems that may exploit different layer 1 protocols.

(iii) It must keep the packet order correct: When the train leaves the tunnel, the packet arrival times on the two branches (satellite and wireless LAN, IEEE802.11b with DCF method9) are quite different. These packet arrival time differences can yield packet order scrambling. Failure to keep the original packet order can degrade the quality of the streaming video.

(iv) A minimum alteration of receivers and transmitting servers: Because we want a technology that can be simply added to existing multimedia satellite communication systems, alterations to existing equipment must be minimized. The requirements mentioned above are summarized in Table 4-1.

4.4.3Conventional diversity reception

Three generic techniques are used for conventional diversity reception; selection-, maximal-ratio-, and equal gain combining. We choose the reception branch with the highest power signal by using a switch for selection combining; the arbitrary diversity branches are first co-phased and then weighted proportionally to their signal levels before summing for

Table 4-1 Requirements of diversity reception systems.

(long) ↑ Shadowing length (short)

Factor	Shadowing length	Measures	Requirements for diversity reception	
Tunnels Buildings	Longer than length of train	Diversity with gap-filler and satellite antenna	・Applicable to different layer 1 protocols	・Minimum alternation of receivers and transmitting station
Bridges Trolley wire structures	Shorter than length of train	Diversity with satellite antennas	・Adaptation to signal symbol timing difference	

Table 4-2 Problems of conventional diversity systems.

Requirements / Diversity system	Adaptation for dynamic symbol difference	Seamless connectivity with gap-fillers	Alteration of receivers and transmitting servers
Selection combining	Need equalizers to compensate for more than one symbol difference	Not able to select or combine different layer 1 signals	Needs switches in receivers
Maximal ratio combining			Needs co-phasing circuits to adapt signal phase differences
Equal gain combining			

the maximal ratio combining; and the diversity branches are summed after co-phasing for equal gain combining.

Conventional diversity techniques need switches or co-phasing circuits in the satellite receiver. Sat ellite receivers need equalizers to compensate for the signal symbol time difference, whose value varies between diversity reception branches due to the diversity distance and the running direction of the train. The signal symbol time difference is described as follows.

$$Dif = \frac{d \cdot \cos\theta}{c} \cdot SR$$

Here, *Dif* is the time difference symbol, d is the diversity distance, θ is the satellite's elevation, c is the light velocity, and *SR* is the symbol rate.

However, it is not easy to adapt the dynamic signal symbol timing difference for equalizers if the diversity distance is long. For example, Dif is about 14 symbols when θ=48 degrees, d=300 meters, and SR=21 Mega symbol/sec. Therefore, conventional diversity usually cannot satisfy requirements (i) and (iii). Furthermore, it is impossible to satisfy requirement (ii) for conventional diversity, because it cannot select or combine signals that exploit different layer 1 protocols. The problems of conventional diversity systems to requirements are summarized in Table 4-2.

Given the weakness of conventional diversity reception schemes, it is effective to use same format signals, which is layer 3 protocol.

4.5 Layer 3 diversity reception

4.5.1 Basic configuration of layer 3 diversity reception

To meet the requirements, we propose a diversity reception technique that can select IP packet, called layer 3 diversity reception. The basic configuration of layer 3 diversity reception is shown in Figure 4-6. It chooses packets using a packet selector from two receivers, each of which decodes the signals to obtain IP packets. The selector has two input ports that receive the IP packets from the two receivers, and one output port. Signals can be received seamlessly, even in tunnels that are longer than the length of the train using gap fillers which exploit a wireless LAN, for example, to satisfy requirement (ii) as shown by grey line in Figure 4-6. In the scheme, the selector must output packets without any duplication and loss. At this time, we developed the selector that has two input ports to confirm primal availability of the technique. The selector naturally can have more than two input ports regarding economy and practicality.

This process is realized by matching a singular packet using content addressable memory (CAM) obtained from the source IP address and the identification data in the IP header. Figure 4-7 shows the function block of the packet selector. Received packets that pass through MAC (Ether controller) are stored in DATA FIFO and entered in CAM, which tries to detect the same packet. The information of a singular packet is stored in CAM HIT FIFO. The transmission processor transmits packets in DATA FIFO, based on information from CAM. Since packets are selected by hardware processing as mentioned above,

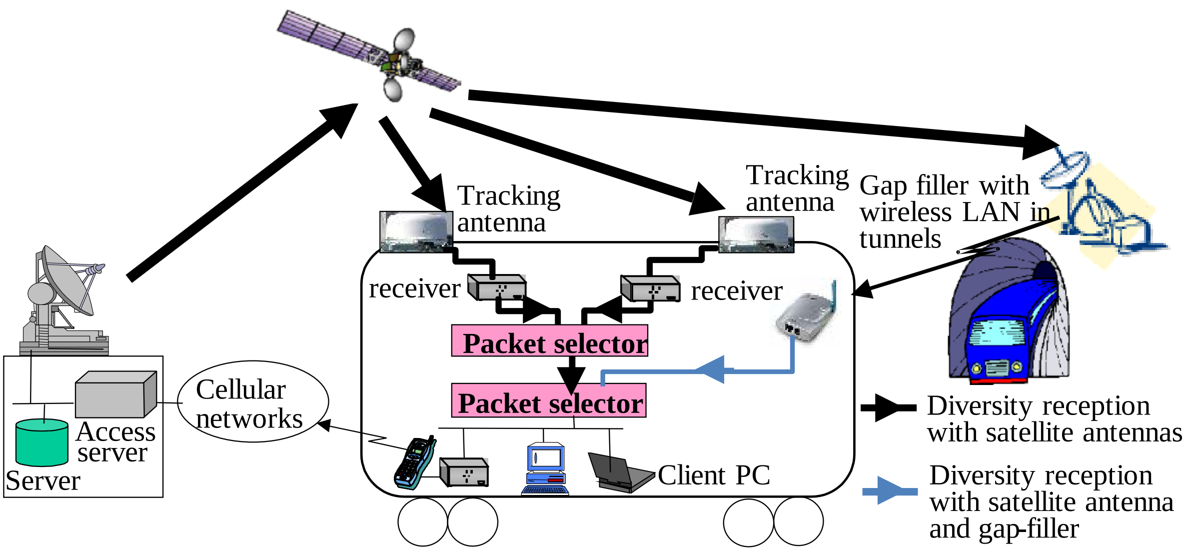


Figure 4-6 Layer 3 diversity reception configuration.

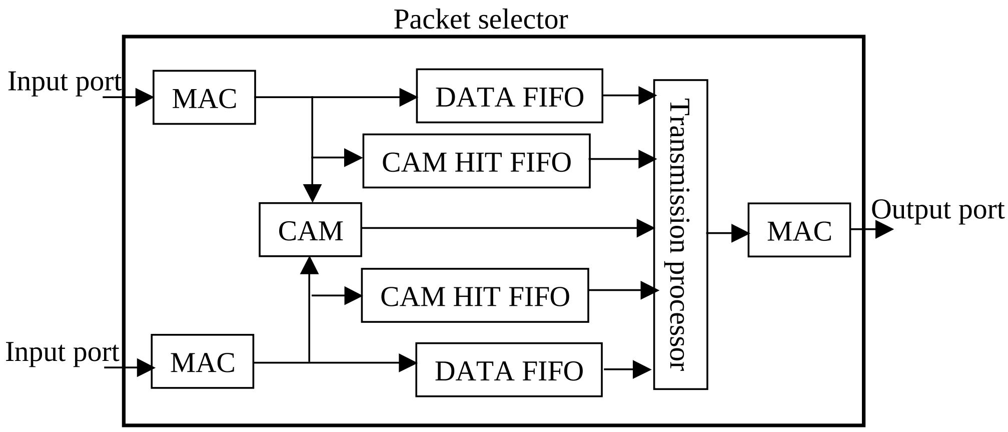


Figure 4-7 Packet selector function block.

the receiving and sending time of Ether frame at the selector is dominant for delay time, which is about 0.2 milliseconds for an IP packet of 1500 Byte. So the delay time at the selector might be too small in the system to influence the system quality.

4.5.2Conventional schemes of packet selector

In the conventional selection scheme, identification numbers are appended to packets at the server side to indicate the order, and the packet selector forwards a packet according to its number to re-order the packet stream. This scheme is illustrated in Figure 4-8. The packet order can be kept as the original stream with this scheme, but the scheme requires additional equipment to append and detach the identification numbers, which may make it too expensive to implement. Also, it doesn't satisfy requirement (iv).

Given the weaknesses of the two conventional schemes, we developed packet-selection logic that can select and forward packets.

4.6 The first proposed packet selection scheme

We developed a packet-selection scheme that can select and forward packets. The proposed scheme places two queues in the selector, one for each receiver. A new packet is forwarded, and one, that has already arrived in the other queue, is discarded. As can be seen in Figure 4-9, packet 1 arrives from receiver B, and packet 2 arrives from receiver A. Because the packets have not arrived, they are output immediately on arrival at the selector. Note that packet 3 from receiver A is output and packet 3 from receiver B is discarded because the same packet has already arrived. Packet 4 from receiver B and packet 5 from receiver A are output just after arriving at the selector because same packets have not arrived. Packet 6 from receiver B is output and packet 6 from receiver A is discarded because the packet from receiver A arrived later than the one from receiver B. Accordingly, packets are transmitted without any duplication and loss. The selector does not require additional information for packet identification, so no significant changes are needed on the server or receiver.

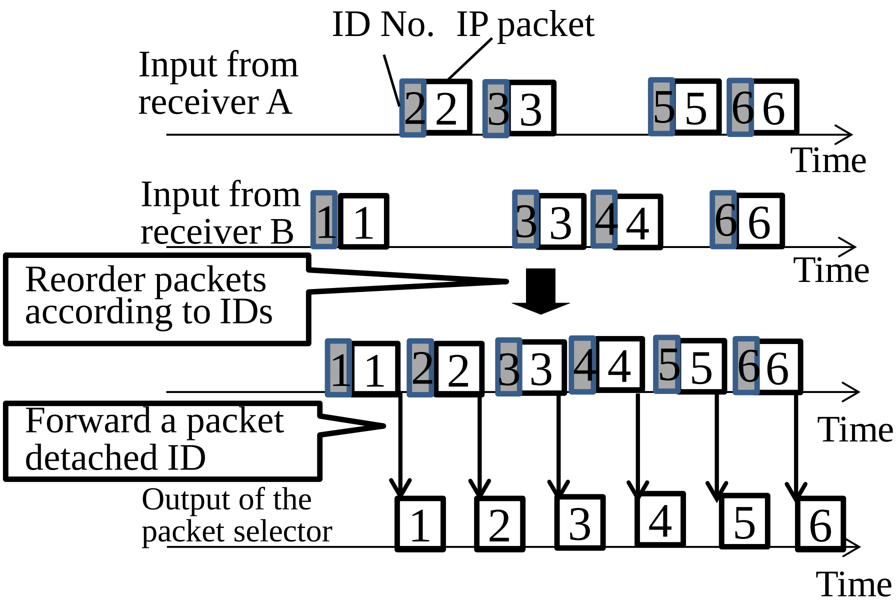


Figure 4-8 The second scheme of packet selection.

4.6.1Layer 3 diversity reception evaluation experiment of the first scheme

We implemented the proposed scheme in the packet selector and evaluated the throughput in an actual train environment.

To determine the effectiveness of layer 3 diversity reception, we carried out a network performance test under the shadowing conditions expected on trains. The characteristic aspects of such shadowing are periodic and frequent caused by the trolley wire structures [4-7]. We emulated this shadowing environment in the laboratory. Structures of 0.3 meters width were assumed to be located every 50 meters along side the railway. The diversity interval of the two antennas was assumed to be 275 meters, which interval is about one train length. Under such conditions, network performance, with and without layer 3 diversity reception, was evaluated by transferring UDP packets from the UDP sender in the server center to a user PC. Figure 4-10 shows the experimental system configuration. Ku-band satellite links were used as the forward links. The received satellite signals were divided into two branches. These shadowing conditions were emulated by switches (SW), which could interrupt the satellite signals by the controller, to recreate the same satellite reception expected in trains.

The experimental system's upper limit throughput was restricted to about 1,800 kbit/s, depending on the ATM's dedicated line speed of 2,000 kbit/s from the server center to the satellite network operations center.

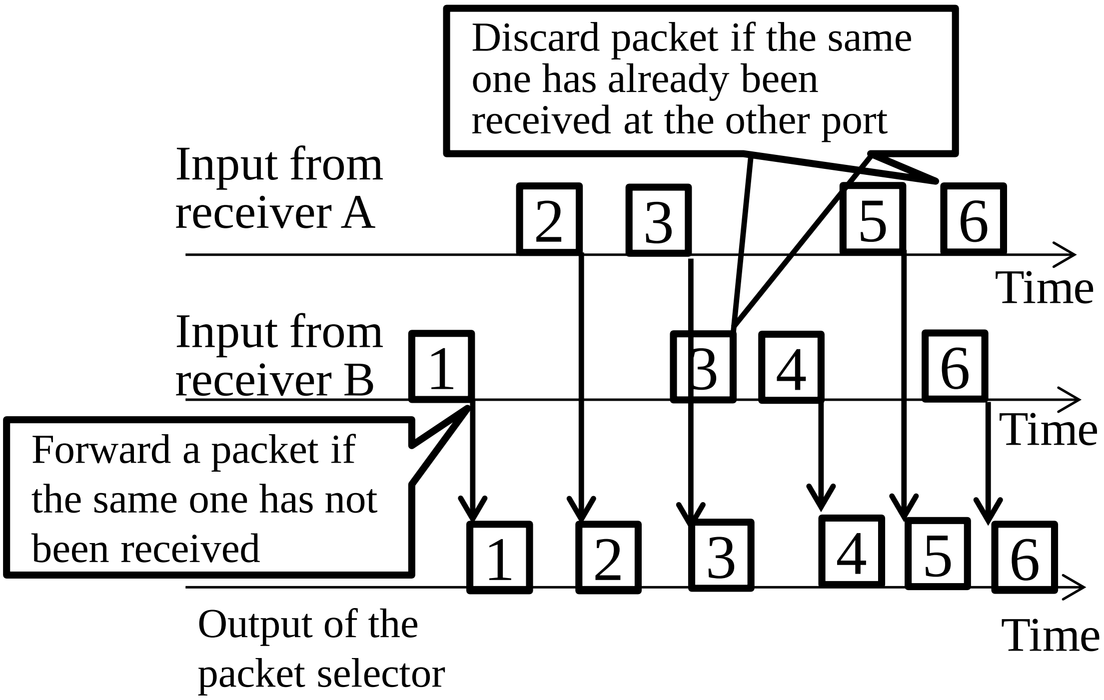


Figure 4-9 The proposed packet selection scheme.

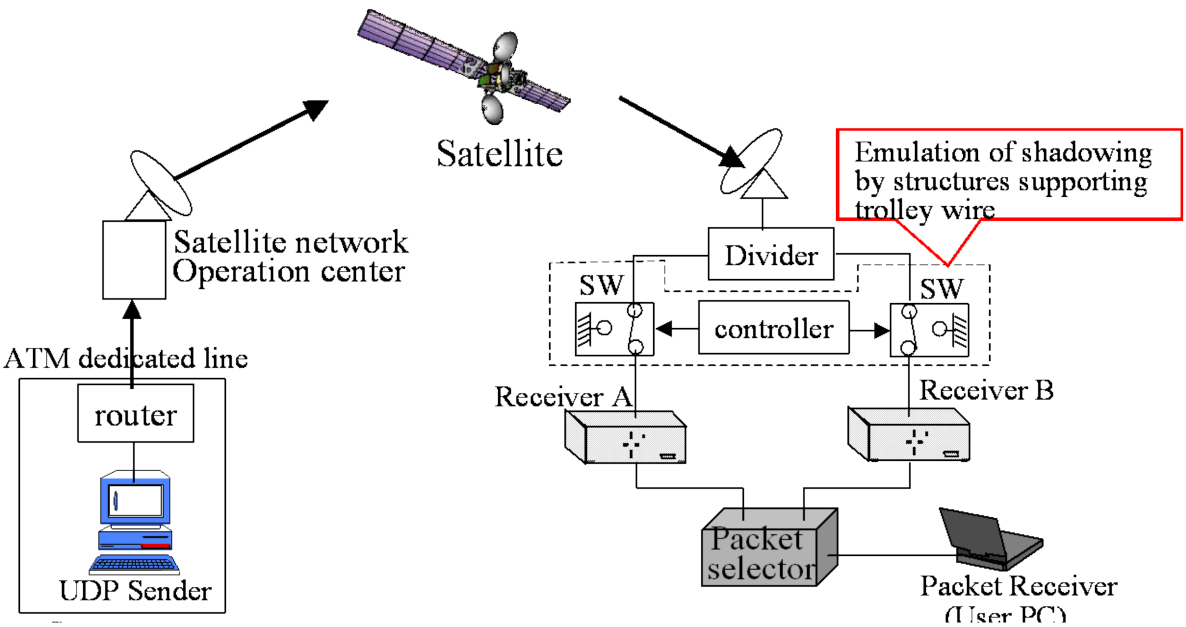


Figure 4-10 Throughput measurement configuration.

The throughput measurement results we obtained from each of the 3 times are shown in Figure 4-11. The results show that throughput without diversity deteriorates as the running speed increases (Appendix I). The throughput with diversity reception was maintained at about 1,800 kbit/s, regardless of the running speed without any packet loss and duplication. We conclude that the layer 3 diversity reception system enables high-speed communication in trains, even with strong shadowing.

4.7 The second proposed packet selection scheme

4.7.1 Issues for the first proposed packet selection scheme

In the first scheme, the selector forwards a packet only if the same one has not arrived at the other port and discards a packet if it has already arrived at the other port. This scheme is illustrated in Figure 4-12. The logic of this method is simple. However, we found that the output packet order may change if one stream is significantly delayed against the other when the early stream recovers from shadowing. We also found that it doesn't satisfy requirement (iii).

4.7.2 The second proposed packet selection scheme

The proposed scheme places two queues in the selector: one for each receiver. A timestamp is appended to each packet when it is placed in the queue. The packet is held in the queue until the same packet is placed in the queue of the other receiver; when this occurs, the selector chooses one of the two packets and releases it, and the other is simply discarded. If no duplicate packet is held for a period exceeding a timeout parameter, or a pair of later duplicate packets is received, the packets which have been held in

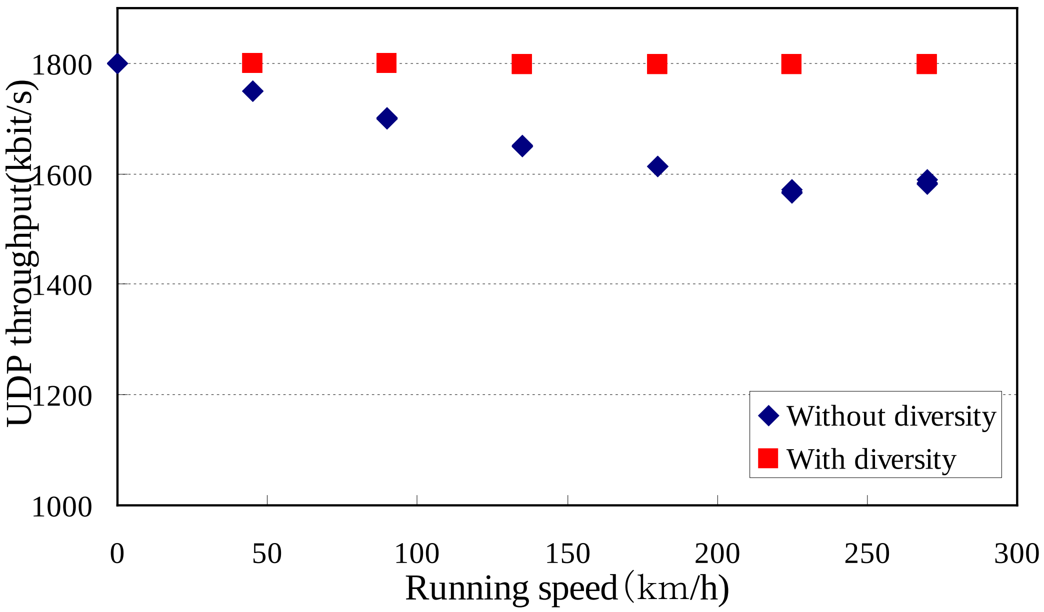


Figure 4-11 UDP throughput, with and without layer 3 diversity reception.

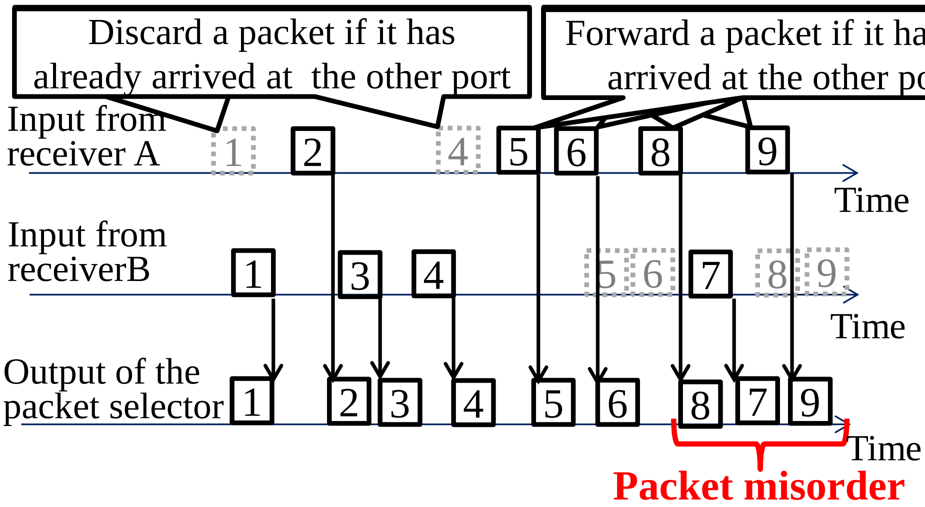


Figure 4-12 Packet misorder with the first scheme of packet selection.

the queues are released by the selector. Packets will occasionally be lost and the timestamps are used to prevent packet misordering as follows. Considering the case in Figure 4-13, packet 4 has been lost on branch B, while branch A has lost packet 3. Note that packet 4 is the output by receiver A ahead of packet 3 from receiver B. In the conventional scheme, this would lead to misordering. The proposed scheme holds packets 4 and 5 in queue A and packet 3 in queue B. The receipt of packet 5 on branch B rewrites the timestamps of the packets in queue A. The offset imposed on the packets in queue A means that packet 4 is now later than that of packet 3. The receipt of the second packet 5 forces all preceding packets held in the queues to be released. Accordingly, packet misordering is prevented. In the scheme, any packets stored in the queue are automatically transmitted when the designated time has elapsed since the packets arrived. This process is realized by matching a singular packet with Content Addressable Memory (CAM) using the source IP address and the identification data in the IP header.

4.7.3Throughput evaluation in a running train environment

We implemented the proposed logic in the packet selector and determined the network performance and packet order recovery.

In order to determine the effectiveness of layer 3 diversity reception, we carried out a network performance test under the shadowing condition expected in a train environment. The characteristic aspects of such shadowing are periodic and frequent shadowing caused by the structures that support the trolley wire7. We emulated this shadowing environment in the laboratory. The structures of 0.3 meters width were assumed to be located every 50 meters along side the railway. The diversity interval of the two antennas was assumed to be 371 meters about one train length. Under such conditions, the network performance with and without layer 3 diversity reception was evaluated by transferring a 5MBytes file in the server (Windows NT4.0 IIS 4.0) to a user PC (Windows 98 SE) by employing FTP with XTP protocol gateways. Figure 4-14 shows the experimental system configuration. The PDC-packet network was used

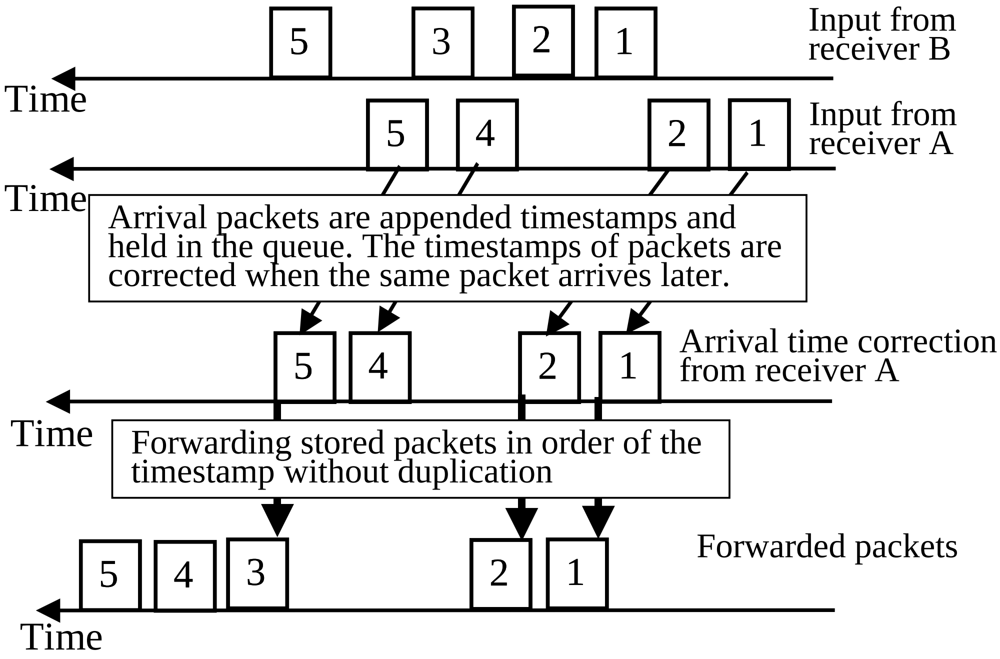


Figure 4-13 The second proposed logic of packet selections.

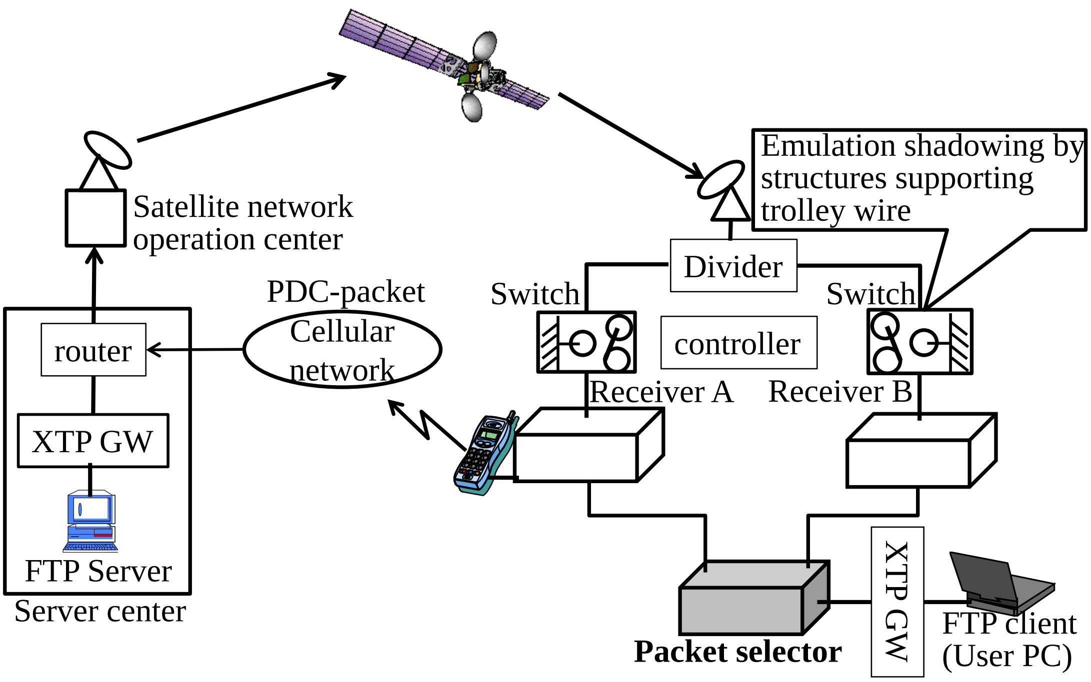


Figure 4-14 Experimental configuration of throughput measurement.

as the return link, and Ku band satellite links were used as the forward link. Received satellite signals were divided into two branches, and the aforementioned shadowing conditions were emulated by switches (SW), which can arbitrarily break the satellite signals that are set by the controller, to recreate the satellite reception expected in running trains.

The dominant system throughput factors include satellite link latency, TCP window size, and the protocol stack (XTP, TCP, UDP, etc.) implemented in the communication system. The experimental system's upper limit throughput is restricted to about 1700 kbit/s, which depends on the ATM dedicated line speed of 2000 kbit/s from the server center to the satellite network operation center.

The throughput measurement results we obtained from each 4 times are shown in Figure 4-15. The results show that the throughput without diversity deteriorates as the running speed increases. The throughput with diversity reception was maintained at about 1500 kbit/s to 1600 kbit/s regardless of the running speed. We conclude that the layer 3 diversity reception system enables high-speed communication in a running train even if the strong shadowing is experienced.

4.7.4Evaluation of packet order recovery

We carried out another experiment to determine that the proposed scheme ensured that the original packet order was maintained. The system examined was assumed to consist of the satellite system plus a gap filler system; the UDP-transmitter sent one 1500 Byte UDP packet every 8 milliseconds. For comparison, we implemented the first conventional scheme in the packet selector. In the scheme, the selector forwards a packet if the same one has not arrived and discards a packet if it has already arrived. As shown in Figure 4-16, the packet order is lost with the first conventional scheme when the train leaves a tunnel if a gap-filler system is used; the packets on the satellite link are much more recent than those on the wireless link. Figure 4-17 shows the configuration of the experimental setup. The average number of misor-

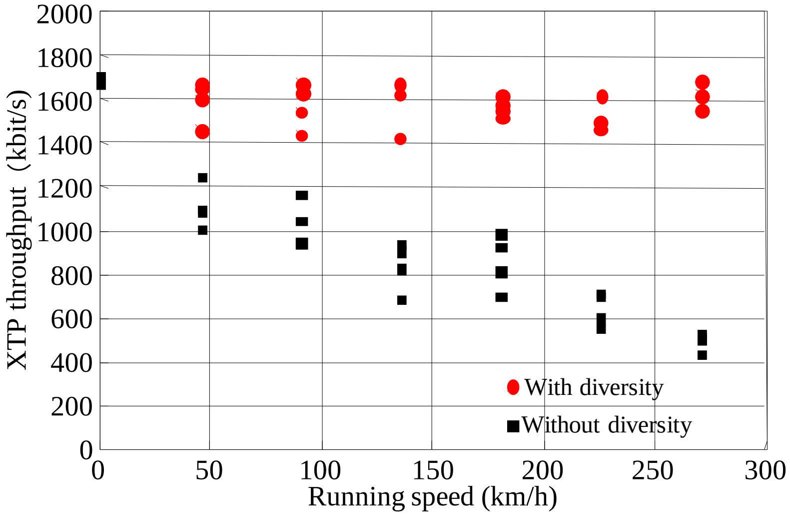


Figure 4-15 XTP throughput with and without layer 3 diversity reception.

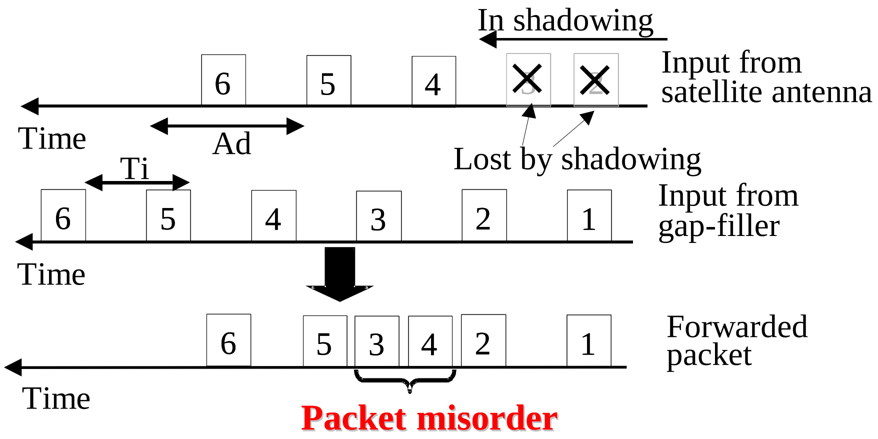


Figure 4-16 Packet misorder.

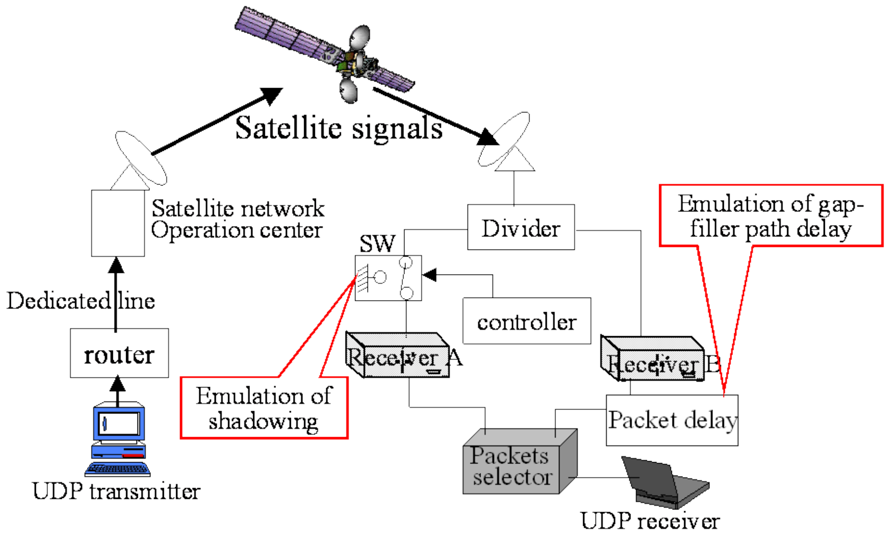


Figure 4-17 Configuration of packet order evaluation.

dered packets was determined from 10 trials under the condition that packets from receiver B were delayed and that packets from receiver A were lost by a switch (SW).

Figure 4-18 shows the results of the experiment. When the conventional scheme is used, the number of misordered packets is given as follows.

$$Ni = 2 \cdot \sum_{i=0}^{\infty} H(Td - i)$$

$$Td = Ad/Ti$$

Here, $H(x) = 0, (x < 0)$
$H(x) = 1, (x \geq 0)$

Ni is the number of misordered packets, Ad is the arrival time difference of the same packets, and Ti is

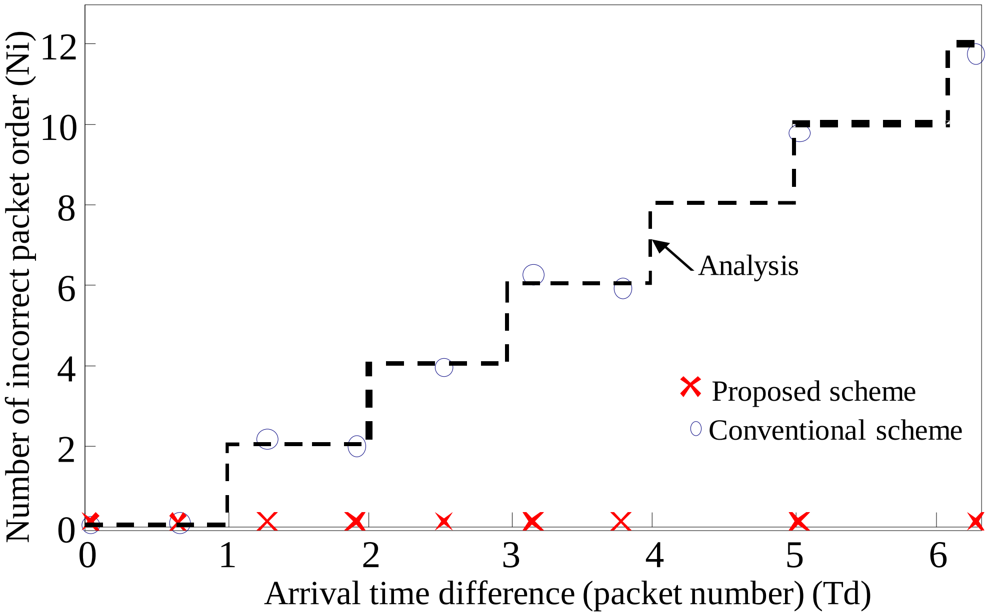


Figure 4-18 Number of misordered packets.

the transmission interval of the packets.

The conventional scheme yields no misordered packets if the arrival time difference is less than the interval of packet arrival; however, the number of misordered packets grows when the arrival time difference exceeds this value. The proposed scheme prevents any packets from being misordered, regardless of the arrival time difference. This verifies that the proposed scheme satisfies the requirement to maintain the packet order.

4.8 Conclusions

We proposed a layer 3 diversity reception system that uses two or more antennas to enhance the network performance of Ku-band mobile satellite communication systems and realize seamless connections with a gap-filler system. The system is realized using a new packet selector, based on IP layer processing. The selector does not require additional numbering for packet identification, so no significant changes are needed on the server or receiver .

We clarified that the proposed technology enables high-speed communications under simulated shadowing conditions on trains.

References for chapter 4

[4-1]Scalise, S., and Schena, V., “Multimedia Service Provision On-Board High Speed Train: Demonstration and Validation of the Satellite-Based FIFTH Solution,” AIAA-paper-2004-3140, ICSSC, 2004.

[4-2]Mori, T., Kagawa, K., and Ishii, J., “Broadband IP Network for Running Trains,” Proc. of Interna-

tional Symposium on Speed-up and Service Technology for Railway and Maglev Systems 2003, JSME No.03-205, pp.94-99.2003.

[4-3]Ohta, G., Kamada,F., Teramura, and N., Hojo, H., “5GHz W-LAN Trials for High Speed Mobile Applications, -Trials on An Express Train and on An Ambulance Car-,” Tech. Report of IEICE, NS2003-66, pp.99-102,2003.

[4-4]Nishimura, T., Mori, T., Ishii, J., Mori, T., Miyatsu, K., Matsuzawa, K., and Tomita, T., “Bluetooth Launch Trial in Hikari Rail Star,” IEE-Japan Tech. Meeting on Transportation & Electric Railway, TER-02-67, PP.11-14, 2002.

[4-5]Nagase, F., Tanaka, H., Nakayama, M., Seki, T., Kazama, H., and Mizuno, H., “Mobile Multimedia Satellite Communication System at Ku band,” IEICE TRANS. COMMUN., Vol. E84-B, No. 4,pp.903-909, April 2001.

[4-6]Mitsugi, J., Nagase, F., and Nakayama, M., “Satellite Communications Technology as an Information Infrastructure for Transportation Vehicles,” The Papers of Tech. Meeting on Intelligent Transport Systems, IEE Japan, ITS-02-31, 2002.

[4-7]Nakayama, M., Nakagawa, M., Hashimoto, Y., Tanaka, K., and Nakashima, H., “A satellite communication system for multimedia network,” IEICE TRANS. COMMUN., Vol. E80-B, No. 1, pp.103-108, Jan. 1997.

[4-8] Nagase, F., Mitsugi, J., Nakayama, M., and Ueba, M., “Ku band Mobile Multimedia Satellite Communications System for Trains,” AIAA-paper-2003-2205, ICSSC, 2003.

[4-9]Clayton, F.M., and McDonald, N.A., “Vehicle antenna for land mobile satellite communications: a diversity system approach,” Vehicular Technology Conference, 1989 IEEE 39th, pp.763-766, vol.2, 1989.

[4-10] Yamaguchi, Y., “S-Band Digital Multimedia Satellite Broadcasting Services for Personal & Mobile Uses in Japan,” AIAA-paper-2004-3213, ICSSC, 2004.

[4-11] Yoshida, T., Kobayashi, K., Ohata, K., and Ueba, M., “Experimental Evaluation for Low Coupling Gap-filler using Null Control,” Technical Report of IEICE, SAT2004-169, pp.17-20, 2005.

[4-12] Araki, K., Kobayashi, M., Nakashima, H., and Kawazoe, K., “ATM Cell/IP Packet Transmission Characteristics Evaluation in Multimedia Interactive Satellite Communications System,” IEICE TRANS. COMMUN., VolJ-81-B-II, No. 5, pp.439-447, May 1998.

Appendix I: Throughput running speed dependency

Unavoidably, Ku-band satellite signals are shadowed by even small obstacles, like trolley wire structures. This slight shadowing degrades throughput at high running speeds. Figure A-1 shows appearance of shadowing time and packet loss time due to signal shadowing caused by trolley wire structures. A receiver can't receive satellite signals during shadowing, which causes packet loss. When a receiver receives satellite signals again, packet reception resumes.

Figure A-2 shows the characteristics of shadowing time and packet loss time. We can see that the packet loss time is constant regardless of shadowing time. This phenomenon occurs due to the demodulation characteristics of the receiver. The satellite signals are shadowed for a specific short time period, during which the receiver demodulates again after receiving a signal from a timer set to a fixed time. For this reason, the throughput depends on the running speed.

UDP throughput (S) is given as follow.

$$S = \frac{Trun - Tplst}{Trun}$$

Here, $Trun$ is total running time of a train, and $Tplst$ is total packet lost time. $Tplst$ is expressed as follow.

$$Tplst = Nst \times Plst$$

Here, Nst is the number of trolley wire structures, and $Plst$ is the packet lost time. Figure A-3 shows UDP throughput by running speeds calculated using the above equations.

The UDP throughput deteriorates at high running speeds, as seen in figure A-3.

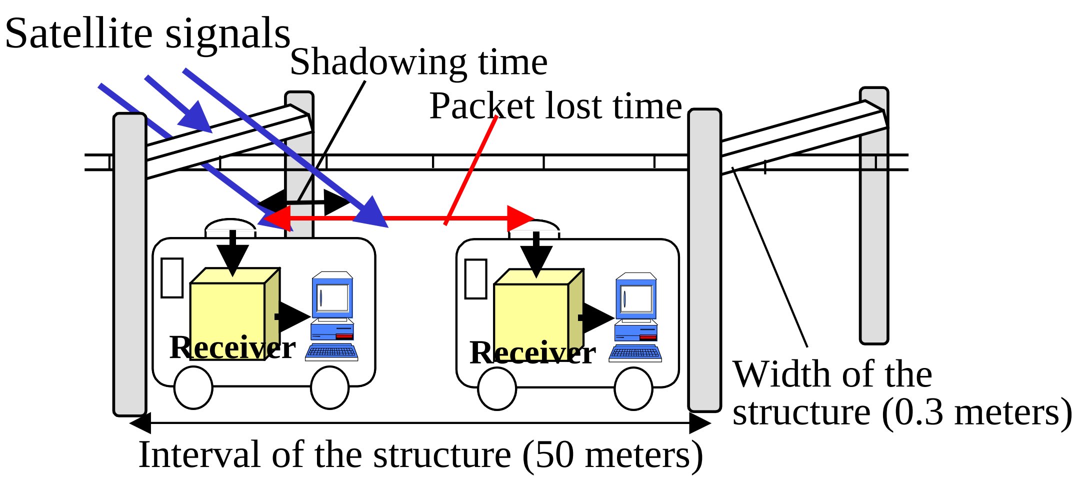


Figure A-1 Appearance of shadowing time and packet loss time caused by shadowing due to structures.

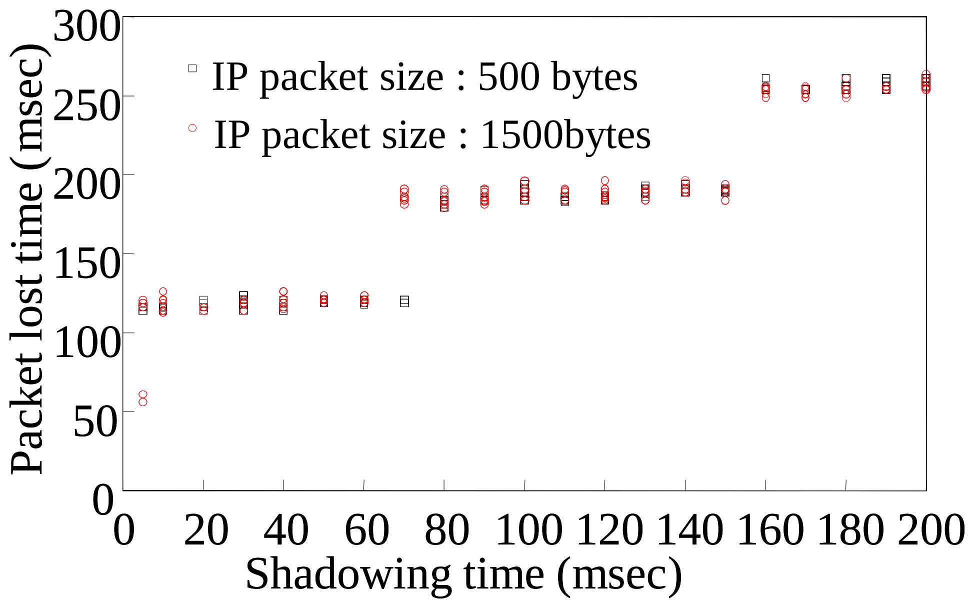


Figure A-2 Shadowing time and packet loss time.

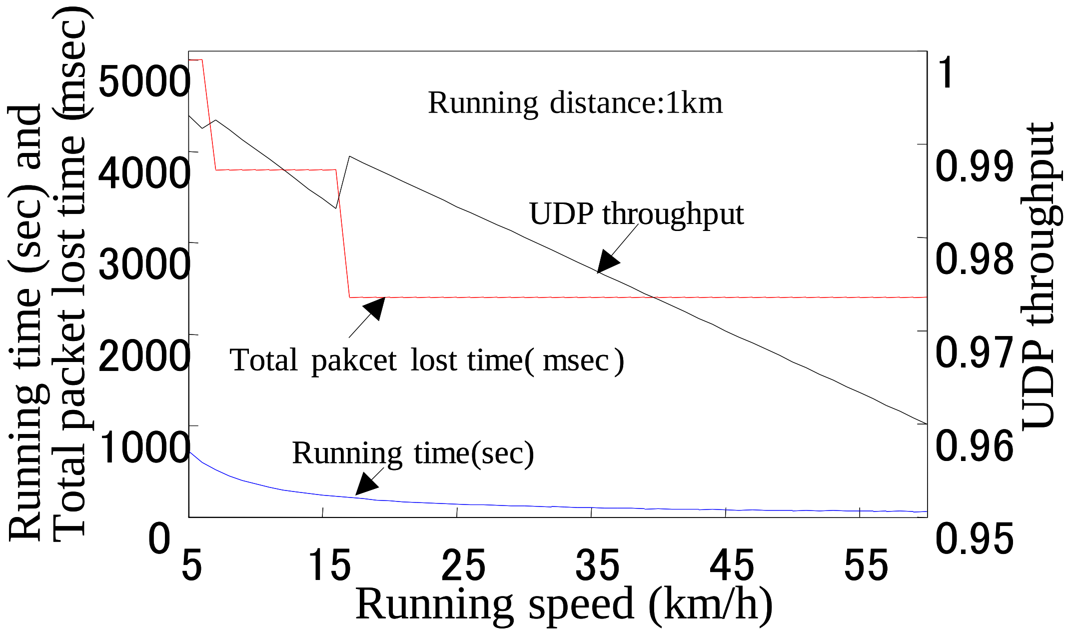


Figure A-3 UDP throughput by running speed.

5. Proposal of WEB prefetching scheme for large latency satellite communication system

5.1 Summary

A scheme for high-speed data transfer via the Internet for Web service in an extremely large delay environment is pro-posed. With the wide-spread use of Internet services in recent years, WLAN Internet service in high-speed trains has commenced. The system for this is composed of a satellite communication system between the train and the ground station, which is characterized by extremely large latency of several hundred milliseconds due to long propagation la-tency. High-speed web access is not available to users in a train in such an extremely large latency network system. Thus, a prefetch scheme for performance acceleration of Web services in this environment is proposed. A test-bed sys-tem that implements the proposed scheme is implemented and is its performance in this test-bed is evaluated. The pro-posed scheme is verified to enable high-speed Web access in the extremely large delay environment compared to con-ventional schemes.

5.2 Introduction

The World Wide Web (WWW), also known as Web ser- vice, is standardized by the World Wide Web Consortium (W3C). WWW is the most popular service among Internet applications. W3C regulates Hyper Text Markup Language (HTML) [5-1] and Extensible Hyper Text Markup Language (XHTML), which are used to achieve Web services. A Web server provides Web contents composed from hypertexts that include multimedia con- tents, such as pictures, moving videos, and music. Web contents are referred to in a Uniform Resource Identifier (URI). Users (clients) send a request for Web contents to the Web server, and the Web server replies to the request. The request and the Web contents are transferred by the Hypertext Transfer Protocol (HTTP) [5-2], [5-3] which is an application protocol.

Web service is the most popular Internet service. Wide-spread use of broadband services can enable high-speed data communication for Web service. Web service with HTTP [5-4], [5-5] often uses the Transmission Control Protocol (TCP) as the transmission protocol. A Web service is composed of several tens or hundreds of web contents, each of which is approximately 10 kBytes. Due to the reliability of TCP, extremely large latency between the clients and the Web servers restricts the throughput of a Web service regardless of the network bandwidth. Therefore, under unusual environments or system condition, the extremely large latency has a major influence on the throughput of a Web service, though it doesn't influence usually. As examples of unusual environment, the con- figuration of high-speed trains is shown in Figure 5-1.

In this system for high-speed trains with satellite communications [5-6], [5-7], since Clients in the trains connects to WLAN (Wireless Local Area Network) [5-8] equipment located at in-vehicle, the latency of between Clients and WLAN will not be almost influenced. However, the latency between the train and the ground station becomes much larger than usual since the propagation distance is long.

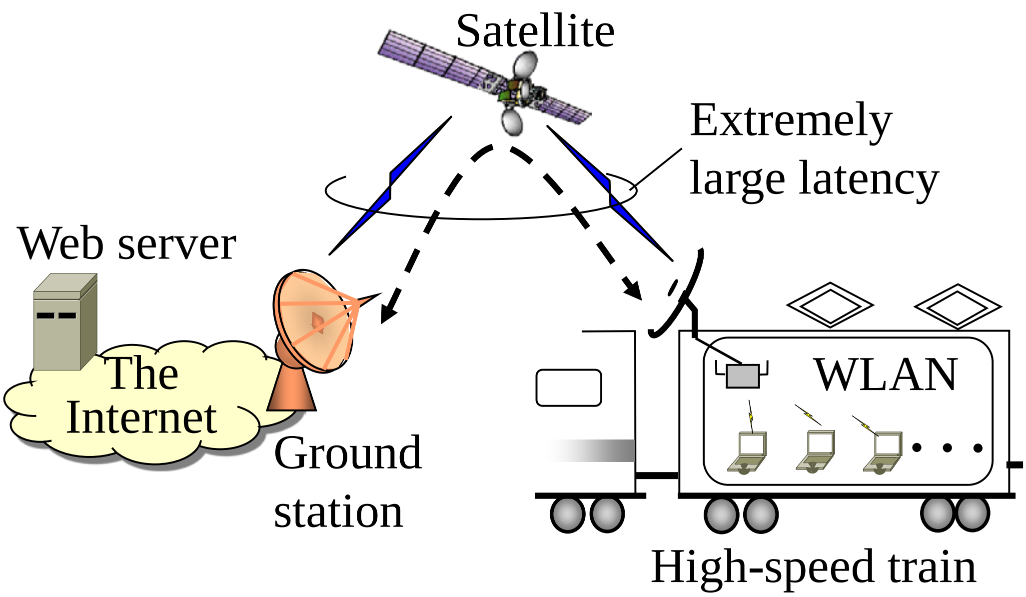


Figure 5-1 Configuration of target WLAN system.

Clients in a high-speed train can access the Internet via the satellite link between the train and the ground station. The long propagation distance via the satellite link causes extremely large latency of about 500 ms between the clients in the train and the Web server. The through- put of a Web service with HTTP is restricted in such an extremely large latency environment [5-9]. To improve these issues, several TCP acceleration schemes and protocol conversion schemes have been studied to boost Web service performance. These schemes may accelerate through- put by enhancement of TCP, which improves congestion control algorithms.

However, since, the Web service transfer a large amount of small-sized data, these schemes will not achieve high-speed Web service. Moreover, the load of the server may degrade its performance, because Web server is installed complex algorithm. Details of these issues is described in Section 2. Thus, a scheme is pro-posed that can enable a high throughput Web service in the extremely large latency environment. The effective-ness of the proposed scheme implemented in a test-bed sys- tem in a laboratory test is shown.

This paper is organized as follows. Section 2 describes conventional schemes and issues, and theoretical values are analyzed and derived. Section 3 presents our pro-posed prefetching scheme for a large latency network, and implementation of the scheme on a test-bed system is described. The evaluation test that was carried out using the test-bed system is described in Section 4. Section 5 concludes this paper.

5.3 Conventional Schemes and Issues

This section shows the signal sequence of HTTP/TCP that is usually used in Web services. A Web service is composed of many Web contents on Web servers (Web Serv.). In a Web service, a client requests Web contents on a Web server, and the server replies to the request by sending an HTML file. The client receiving the file ana- lyzes it and issues the request written in the file. Thus, one content is delivered from the Web server to the client in reply to one request from the client to the Web server. After the content is transferred, the next request for other content is sent. Many contents are transferred repeatedly as described. HTTP/TCP sequence and its performance with normal access is explained hereafter. After that, the conventional prefetching scheme and its issues are de- scribed.

5.3.1HTTP/TCP Sequence

This section introduces the throughput of Web contents with HTTP/TCP. A general overview of the transmission sequence of HTTP is shown in Figure 5-2(a). The HTTP sequence begins after a TCP connection is established through a three-way handshake procedure [5-10]. The TCP connection is established once, and the session continues until the transfer of all contents ceases. Many contents are transferred through the TCP session. One content continues to be sent until an HTTP status-code signal is finally sent.

One content is transferred through the TCP session as shown in Figure 5-2(b). This environment has the latency of round trip time (RTT). In the figure, the content is transferred through a number (k) of data transfers. In data transfer with TCP, window base control and congestion control [5-11], [5-12] are basically used to achieve reliable data transfer. When receiving an acknowledge (ACK) signal, the Web server (Web Serv.) sends data one of k times, the size of which is expressed as Ds(k), with a smaller window base control value (Rwin) and congestion win-dow size at the k-th congestion window size (cwnd(k)), as in the equation denoted below.

$$Ds(k) = Min\{cwnd(k), Rwin\} \tag{1}$$

Here, Rwin is the receive window size advertised from the client to the Web server. The sum of Ds(k) is the size of the content, which is expressed as “Cs” in this paper. The congestion window size at the k-th TCP transmission (cwnd(k)) is expressed as

$$cwnd(k) = 2^{(k-1)} \times cwnd(1) \quad (cwnd(k) < ssthresh)$$
$$cwnd(k) = cwnd(k-1) + MSS \quad (cwnd(k) \geq ssthresh) \tag{2}$$

where ssthresh means the slow start threshold of congestion control by TCP. MSS is the maximum segment size, which is 1460-Bytes in Ethernet, for example.

The congestion window algorithm begins in the exponential growth phase initially with a congestion window size. Since the client sends an ACK, this behavior effectively doubles the window size each round trip of the network. Once the cwnd reaches the ssthresh, TCP goes into congestion avoidance mode, where each ACK in- creases the cwnd by 1 MSS for each ACK received. This results in linear increase of the cwnd. This behavior con- tinues until the cwnd reaches the size of the client’s advertised window (Rwin). The TCP/IP standard allows for Rwin up to 65,535 bytes (= 64 kBytes) in size, which is the maximum value that can be specified in the 16-bit TCP window size field. The maximum value of Rwin

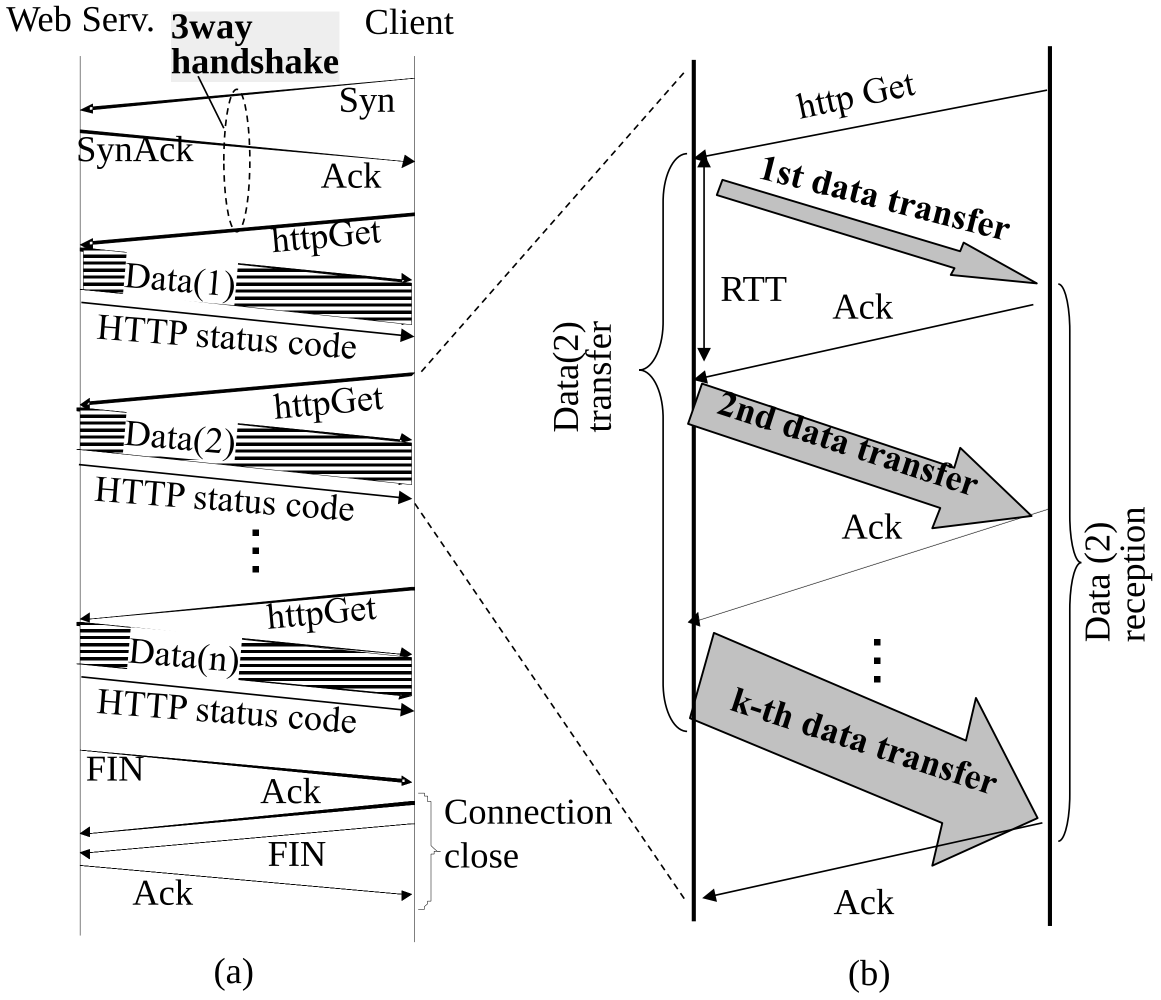


Figure 5-2 Sequence of HTTP (a) and data transfer procedure (b) of TCP.

is accurately expressed as 65,535 = $2^{16} - 1$. The maximum TCP throughput (MaxS) is obtained as follows:

$$MaxS = Rwin / RTT = (64kByte \times 8) / 0.5 = 1024[kbps] \quad (3)$$

Hence, the TCP throughput is restricted to 1024 kbps in a high latency system, regardless of the maximum Rwin value.

5.4 HTTP Performance with Normal Access

This section shows the time for all contents in a Web service to be downloaded (Web load time) and HTTP throughput in a large latency network. Web services are composed of several tens or hundreds of Web contents, which are text, pictures, moving pictures, and so on. The total size of a Web service is generally about several tens of kilobytes to several megabytes. This signifies that the size of one web content is very small, at about 10 kBytes. From the Data(1) transfer to Data(n) transfer, shown in Figure 5-2(a), which expresses each web content as small size. One web service is composed of such a small num- ber n of contents. One web content is transferred by k number of data transfers, shown in Figure 2(b). The client issues the request (HTTP get request) for acquiring the content. Since the data transfer Ds(k) in Equation (1) does not increase to Rwin, the server sends small-sized data (Ds(k)) k times in the sending of one content. Thus, the time for one small-sized content transfer increases to RTT × k, meaning clients cannot obtain high throughput. One content is transferred in the summed time of Ds(k). Since the throughput (S) is the value of the total transfer data size (Cs) divided by the total transfer time (RTT × k), S can be expressed as follows.

$$S = \frac{Cs}{RTT \times k} = \frac{1}{RTT \times k} \sum_{n=1}^{k} Ds(n) \quad (4)$$

The throughput for various transfer data sizes can be derived as shown in Figure 3. RTT is assumed to be 500 ms in Figure 5-3. The throughput for 10-kByte content is restricted to less than 50 kbps regardless of the 64-kByte Rwin. This value is about 20 times smaller than 1024 kbps, which was obtained in Equation (3). This indicates that high throughput cannot be obtained when sending small sized data with the maximum Rwin in a large latency network.

HTTP can transfer contents simultaneously in one TCP session. The time for transferring all contents in the Web service, called Web load time, is expressed as follows.

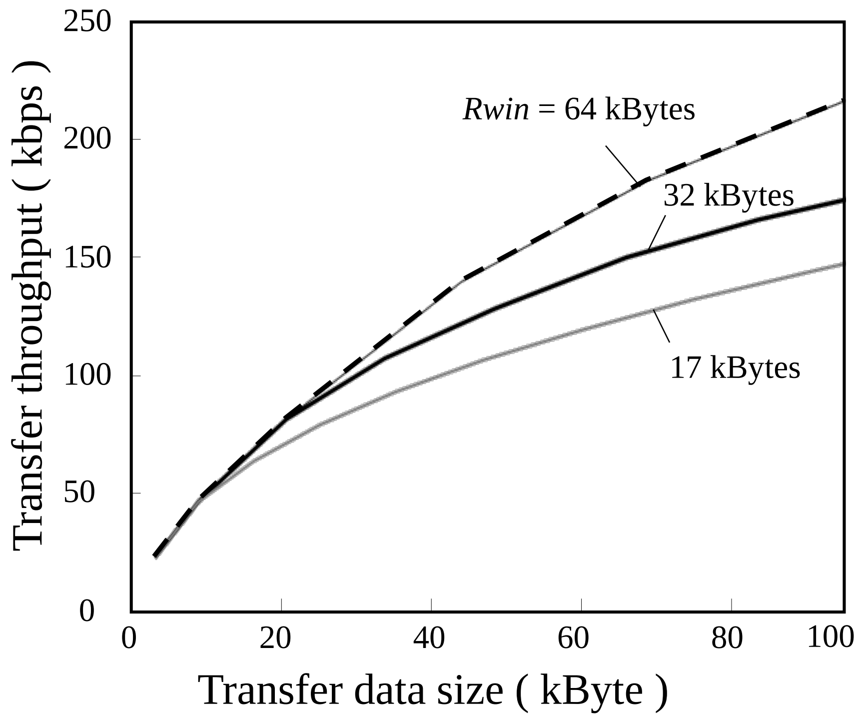


Figure 5-3 Transfer throughput for various data sizes and different Rwin.

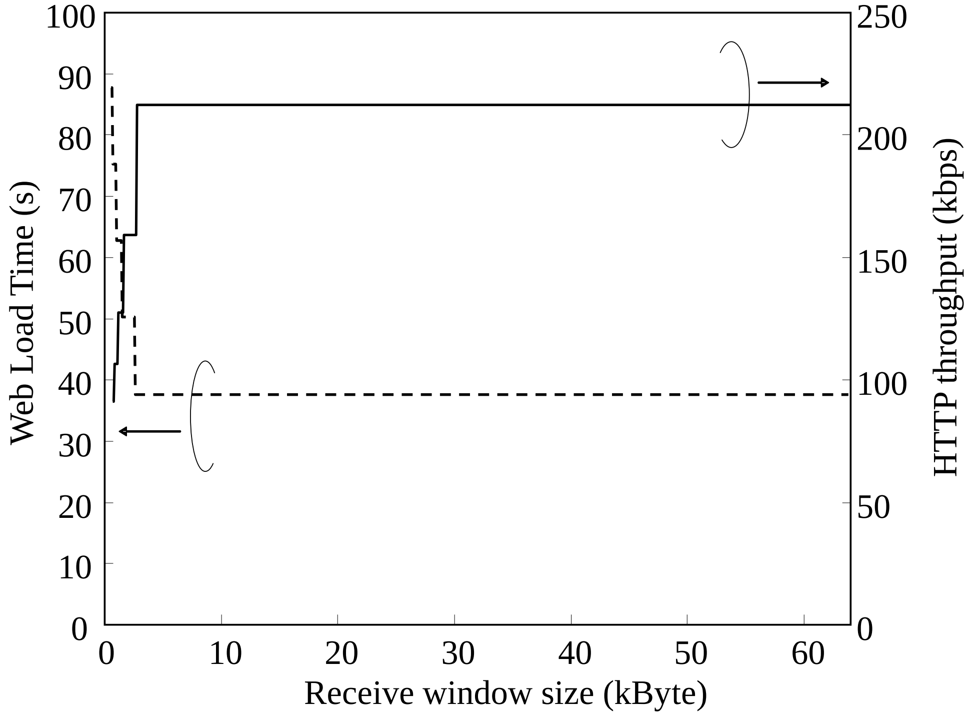


Figure 5-4 HTTP throughput for receive window size.

$$Tl = \frac{\sum_{n=1}^{k} wc(n)}{sm} \tag{5}$$

Here, Tl is the web load time, wc(n) is the transfer time of the n-th content, k is the total number of Web contents in the Web service, and sm is the number of simultaneous connections for transfer. The size of wc(n) is Cs, as de-noted in Equation (4).

Since HTTP throughput (Sh) is derived from the value of the size of all contents divided by the Web load time, Sh is expressed as follows

$$Sh = \frac{\sum_{n=1}^{k} Cs(n)}{Tl} \tag{6}$$

Here, Cs(n) is the size of the n-th content in the Web service. HTTP throughput derived from Equation (6) is shown in Figure 5-4. In Figure 5-4, it is assumed that the content size is 10 kByte, the total number of contents is 100, the number of simultaneous connections is 4, and RTT is 500 ms. As can be seen, HTTP throughput is re-stricted to 212 kbps and the Web load time is larger than 37.6 s when Rwin is only 4 kByte. Since a Web load time within 8 s is commonly known to be demanded for user satisfaction [5-13], [5-14], a Web load time of 30 s by normal access in an extremely large latency environment will not satisfy user demands.

5.5 Conventional Prefetching Scheme

Several TCP acceleration schemes and protocol conversion schemes have been studied to boost Web service performance [5-15]. The schemes can accelerate TCP throughput by expanding Rwin, which improves congestion control algorithms. The Web service characteristic that a large amount of small-sized data is transferred makes achieving high-speed service difficult with such schemes.

Increasing the number of simultaneous connections can increase HTTP throughput, as shown by Equation (5). However, increasing the number of simultaneous con- nections means the Web server must manage a large number of sessions. Increasing the load of the server degrades its performance [5-16]. Hence, it is recommended by RFC2616 so that the number of simultaneous connections is restricted. A pipeline [5-17] scheme has also been studied, but few Web servers support this scheme.

A Web prefetching scheme [5-18], [5-19], [5-20] has been studied to improve HTTP performance. The prefetching server (Pf Serv.) in a wide area network (WAN), shown in Figure 5-5, enables high-speed HTTP throughput in a conventional scheme without expanding Rwin size or increasing the number of simultaneous connections. The prefetching server has two functions: prefetching, and forwarding the stored prefetched information. The sequence of the conventional prefetching scheme is shown in Figure 5-6.

The prefetching server captures the HTML file (data 1), which is the data sent in reply to the request from the client to the server. The prefetching server analyzes the file to issue the prefetching request (Pget 2) and stores the prefetched contents (data 2). After receiving the HTML file, the client may issue a request (get 2) for the same contents acquired by the prefetching server. The prefetching server replies to the request and sends the stored content (data 2). The conventional prefetching schemes effectively doubles the throughput compared with normal access when the prefetching server is located halfway between the Web server and the client. However, the conventional scheme is not as effective for achieving the maximum performance in the target system, which would incorporate the prefetching server in the train or the ground station.

5.6 Proposed scheme

5.6.1 Operation of proposed prefetching scheme

In this subsection, the proposed prefetching scheme is described. The proposed scheme is characterized by its division of the two functions base on the conventional prefetching scheme. Each function is installed so that a WAN (satellite link) may be inserted as shown in Figure 5-7. Therefore, a WAN with the extremely large latency is sandwiched by two servers with different function. The prefetching server, which prefetches for the Web server and forwards the prefetched contents to the information storage server, is installed in the network near the Web server. The information storage server, which stores the prefetched contents and forwards them to the client, is installed in the network near the client. The se-

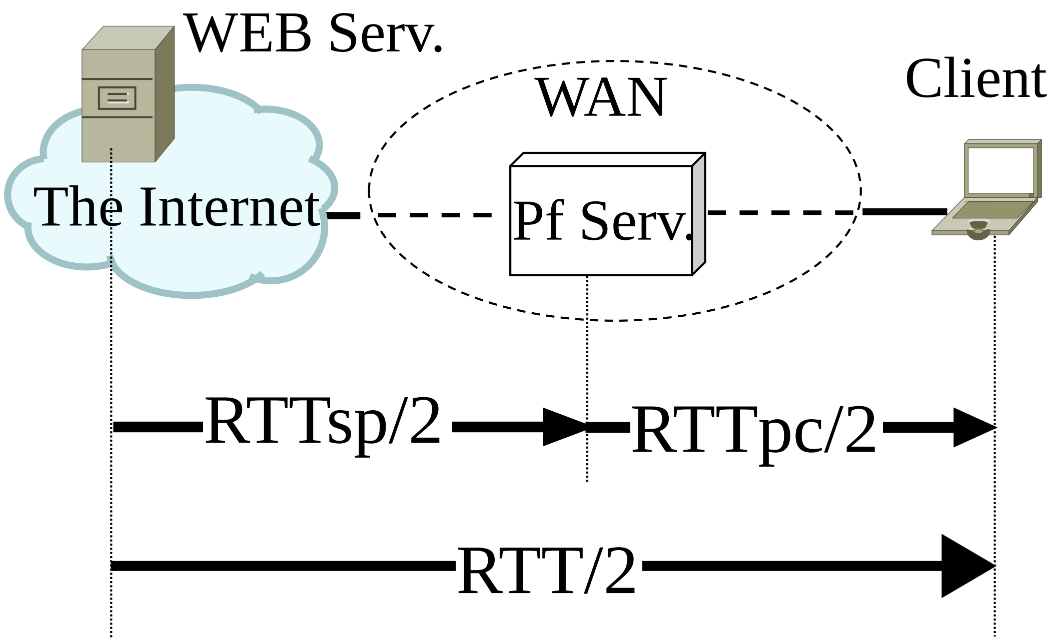


Figure 5-5 Configuration of conventional prefetching scheme.

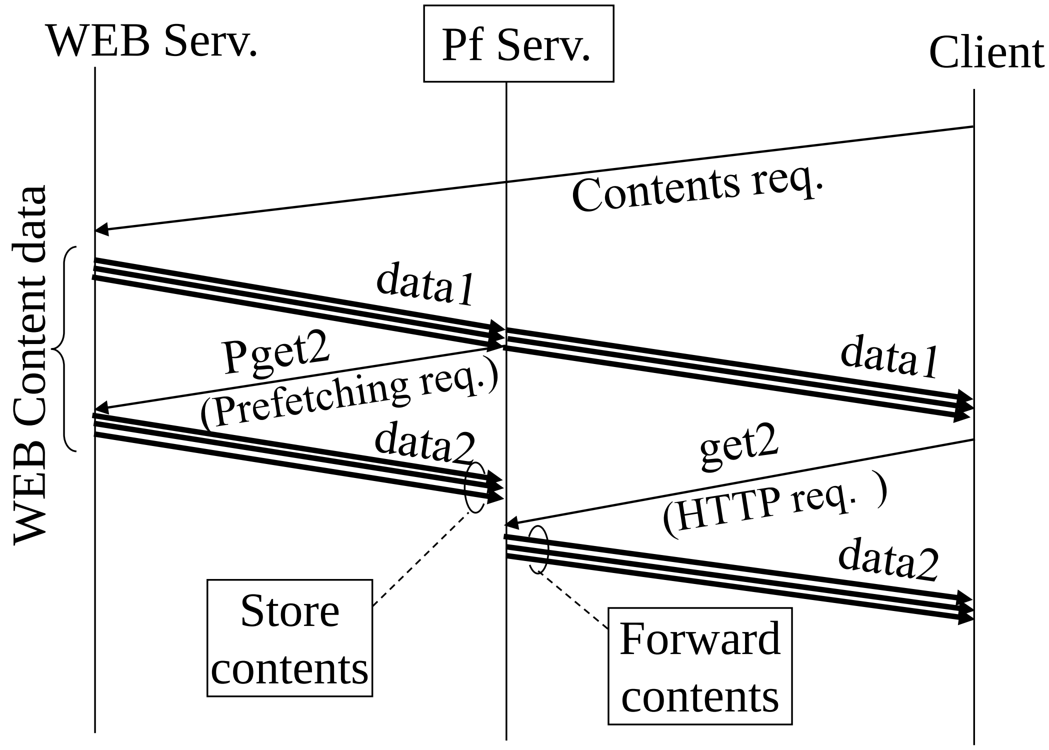


Figure 5-6 Signal sequence of conventional prefetching scheme.

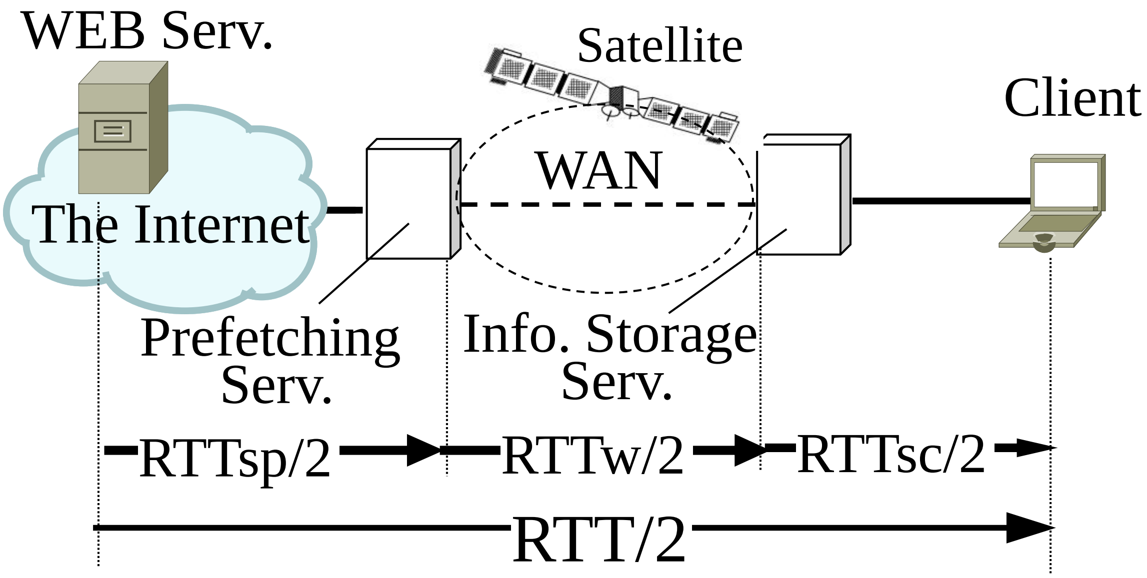


Figure 5-7 Configuration of proposed prefetching scheme.

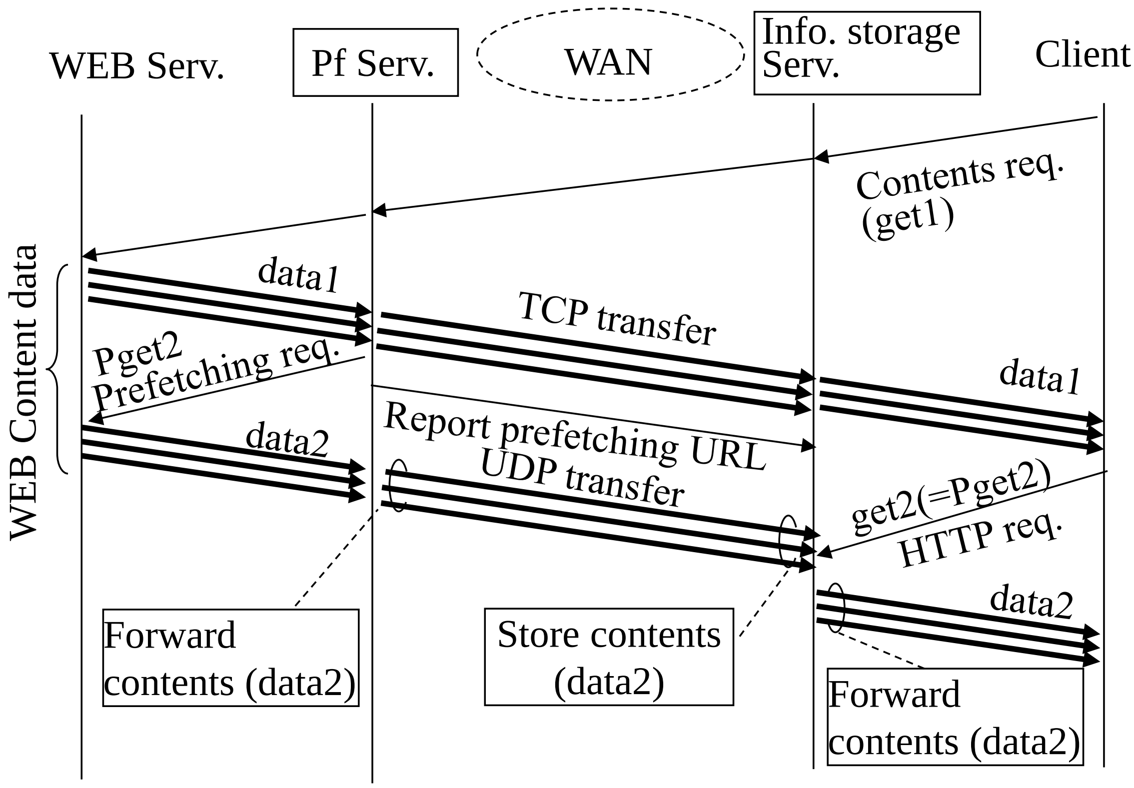


Figure 5-8 Signal sequence of proposed prefetching scheme.

quence of the proposed prefetching scheme is shown in Figure 5-8. The prefetching server captures the HTML file (data 1), analyzes it to issue a prefetching request (Pget 2), and acquires the content (data 2) noted in the HTML file.

Before sending data 2 to the information storage server, the prefetching server reports the prefetch ing URL to the information storage server, which replies with the prefetched content for the request (get 2) from the client.

Reliable communication that can be achieved by TCP with congestion control, retransmission control, and packet reordering may restrict the throughput. The User Datagram Protocol (UDP) is characterized by having no data transfer control such as implicit handshaking dialogues for providing reliability, receiving ordering, or data integrity. Because of this, UDP communication can pro- vide high-speed communication in a large latency network when packet loss is negligible. This transmission characteristic was why we decided to use UDP in the proposed scheme to transfer prefetched contents.

Low cost development of the proposed scheme can be achieved by utilization the functions of the conventional scheme and transferring the prefetching data with high speed UDP.

The time taken to transfer prefetched contents in then only the contents size divided into the transmission speed and the time of propagation latency through WAN, so high-speed performance could be obtained with the pro- posed scheme. The satellite link of long distance accrues packet error by propagation loss. This system environment is significant problem in order to use UDP, since UDP does not have the function for packets error recovery. Scheme to solve this problem will be described in subsection 5.6.4.

5.6.2Theoretical evaluation of Web Load Time

To confirm the effectiveness of the proposed scheme, the Web load time was derived for normal access, i.e. without any prefetching, scheme, access with the conventional prefetching scheme, and access with the proposed prefetching scheme by theoretical analysis. The evaluation parameters are shown in Table 5-1.

Table 5-1 Evaluation parametes of Web load time.

Size of one web content (Sw)	10 kBytes
No. of web contents	N
HTTP version	1.0
No. of simultaneous HTTP contents	1
Rwin	64 kBytes
RTT of client and Web server	510 ms
WAN latency	500 ms

Considering that the TCP uses window base control and congestion control as mentioned in Equations (1) and (2), the first transfer sends 1.46 kBytes (=1 MSS) of the content, the second transfer sends 2.92 kBytes (=2 MSS), and the third transfer sends the remaining 5.62 kbytes (=3.84 MSS, which is within 4 MSS). Since 10-kBytes content can be transferred in 3 transfers, it takes 3 × RTT to load one content. The web load time (Tl) can then be derived as follows.

1) Normal access

Each of the contents takes three RTTs to load. Since the Web service is composed of N number of contents, the Web load time can be expressed as follows.

$$Tl = 3 \times \mathrm{RTT} \times N \quad (7)$$

2) Conventional prefetching scheme

Each of the contents is loaded to the prefetching server in three RTTs between the Web server and the prefetching server (RTTsp). Each of the prefetched contents in the prefetching server is loaded to the client in three RTTs between the prefetching server and the client (RTTpc). Therefore, the web load time is three times the larger value out of the RTTsp and RTTpc as follows.

$$Tl = 3 \times \mathrm{Max}\{\mathrm{RTTsp}, \mathrm{RTCCpc}\} \times N \quad (8)$$

Here, the sum of RTTsp and RTTpc is the RTT of the Web server and the client shown in Figure 5-5.

3) Proposed prefetching scheme

Each of the contents is loaded to the prefetching server in three RTTs between the Web server and the prefetching server (RTTsp). Each of the prefetched contents in the prefetching server is forwarded to the information stor- age server as content size (Sw) divided by the UDP transfer speed (Su). Each of the contents stored in the information storage server is loaded to the client in three RTTs between the information storage server and the client. Therefore, the web load time can be expressed as follows when it is assumed that no packet is lost in the WAN link.

$$Tl = 3 \times \mathrm{Max}\{\mathrm{RTTsp}, \mathrm{RTCCsc}\} \times N + (Sw/Su) \times N \quad (9)$$

The results of the load times derived from Equations (7), (8), and (9) are shown in Figure 5-9. As can be seen, the proposed scheme can significantly shorten the Web load time.

5.6.3Test-Bed system of Proposed Scheme

A test-bed system was created to verify the effectiveness of the proposed scheme. The configuration of the prefetching server and the configuration of the information storage server are shown in Figure 5-10 and Figure 5-11 respectively. Apache Httpd 2.2 and Apache Tomcat 6.0.29 are installed on Red Hat Enterprise Linux 5.4. Each server operates as an HTTP server within a hierarchical structure. The servers perform in a so-called parent and child proxy configuration: the prefetching server is the parent proxy, and the information storage server is the child proxy. As a parent proxy, the prefetching server prefetches contents and forwards them by UDP to the information storage server. As a child proxy, the information storage server stores the prefetched contents and replies to requests from the client.

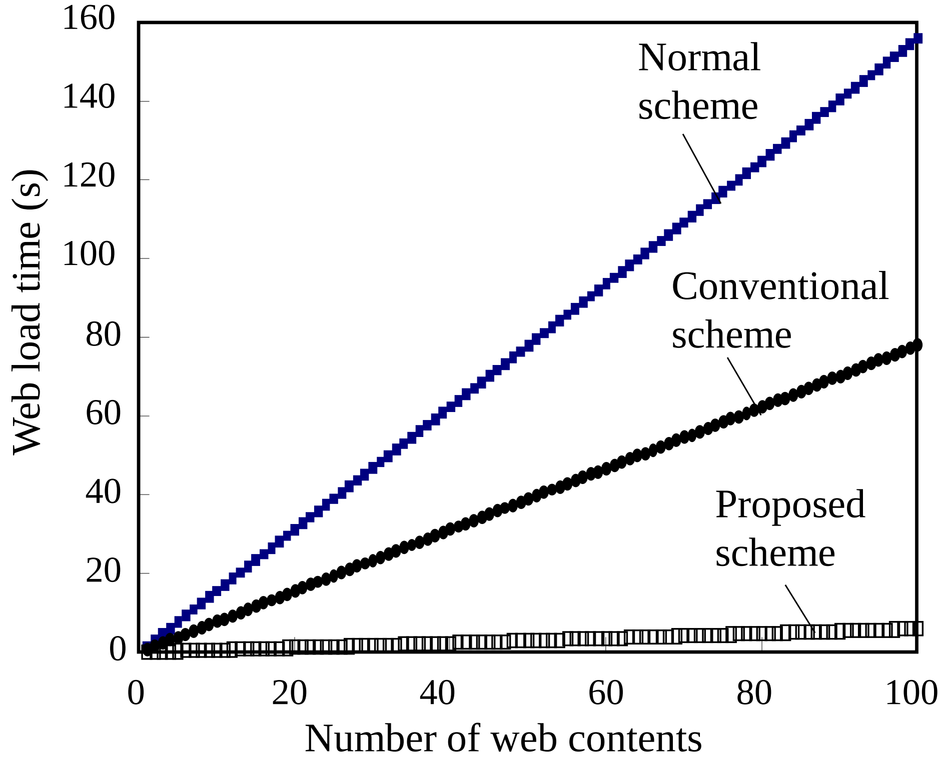


Figure 5-9 Web load time by different types of access for various numbers of web contents.

Considering practical utilization of the developed test-bed system, these servers must perform as an intercepting proxy, also known as a forced proxy or transparent proxy. Therefore, the information storage server is needed as a proxy, which intercepts normal communication, with clients not needing any special configuration to use the proxy. Clients do not need to even be aware of the existence of the proxy. In an intercepting proxy, utilization of the IP tables function installed in Linux detects the re-quest packets from the client to the Web server to for-ward the packets to the port opened by the squid function. The squid function accepts the information storage server as the proxy server system.

Precise prefetching, i.e. acquiring an storing the same content as the client's request, enables high-speed Web service. This section introduces a system for pre- fetching requests from an HTML file. The test-bed sys- tem targets HTML 4.01. The prefetching requests are extracted only when the "Con-tent-Type" header value [5-21] of an HTTP reply includes "text/html", and the system is designed to ex-tract the following elements in the HTML file.

(i) Pictures of src attribution in <img> tags.

(ii)JavaScript of src attribute in <script> tags.

(iii)"Style sheet" of href attribute in <link> tags

(iv)Pictures of href attribute of <link> tags.

5.6.4UDP Packet Loss Recovery Function

Regarding the WAN link in this paper, packet loss may occur when the link is blocked by an obstacle or when packets spill over the network interface card of the client. Moreover, the satellite link of long distance has significant propagation loss compared to a wired link. If the prefetched contents are not transferred, the quality of the Web service utilizing this prefetching scheme may be degraded. Hence, the function to recover the packet loss is required in this system, since UDP does not have the function for packets error recovery.

To compensate for packet loss, a retransmission function (loss recovery function) was imposed in the test-bed system. The fundamental operation for recovering UDP packet loss is shown in Figure 5-12. The prefetching server divides a prefetched content into n number of UDP packets. The size of each

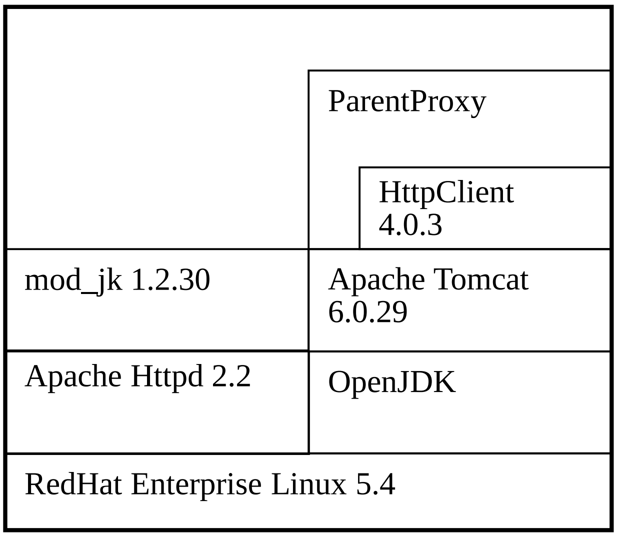


Figure 5-10 Configuration of prefetching server.

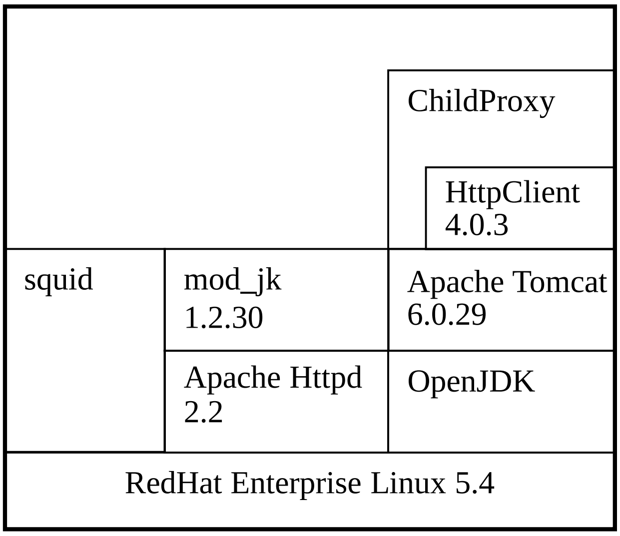


Figure 5-11 Configuration of information storage server.

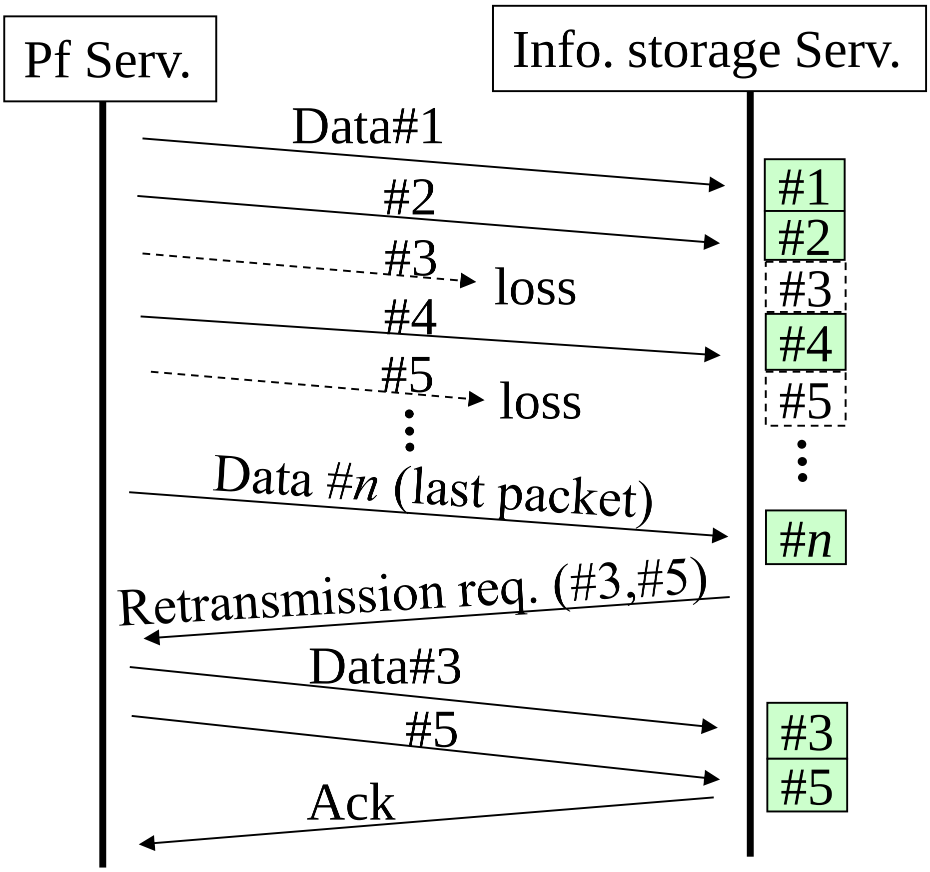


Figure 5-12 UDP packet loss recovery function.

packet is fixed to L Bytes. The packets are added in ascending sequence number from 1 to n, and the n-th packet is indicated as the last packet. They are transferred for every fixed time. The fixed size and the fixed transmission interval can control the transmission speed so that it stays within the link speed of the WAN. The information storage server sends a retransmission request that includes the sequence number of lost packets to the prefetching server when the last packet is received. After receiving the request from the information storage server, the prefetching server retransmits the lost packet. Hence, all of the packets are confirmed to be transferred without any loss in the test-bed system.

5.7 Evaluation of Proposed Scheme

An evaluation test was performed to clarify the effectiveness of the proposed scheme by using the test-bed system. The configuration of the evaluation test is shown in Figure 5-13. The network emulator made a WAN environment that had latency of 250 ms in one-way trip time, i.e., a 500-ms RTT in order to produce the satellite link, and the bandwidth was set to 100 Mbps. The rest of the network was linked with a 100-Mbps Ethernet LAN. The UDP size was fixed to 1414 Bytes of Segment Size, and the transfer interval was fixed to 1 ms. The client accessed the contents in the Web server with HTTP 1.0, and the number of simultaneous connections was 1. The contents, 28 in total, were 1 HTML file, 10 GIF files, 4

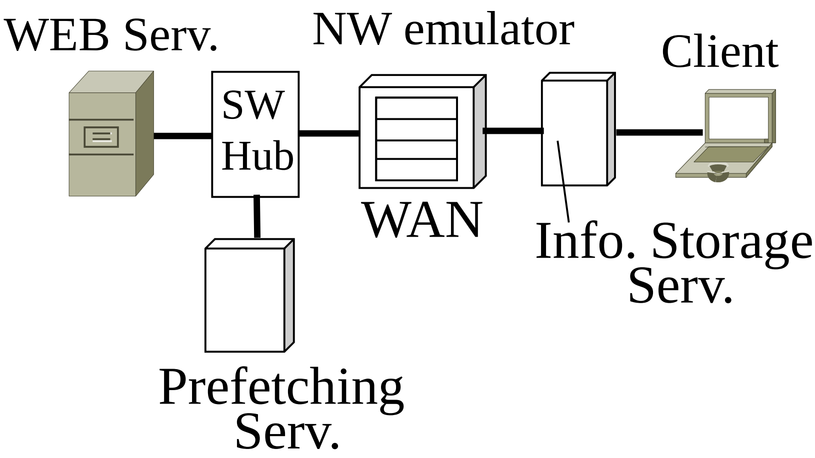


Figure 5-13 Configuration of evaluation test in laboratory.

JPG files, and 13 PNG files: the total size of the contents was 1849.285 kBytes. Each of the contents was transferred in each HTTP session. The content packets were captured in the client, and the final time of each HTTP session was observed. The test results and the theoretical values from Equation (6) are shown in Figure 5-14. Web load times of 47.4 s with normal access and 4.3 s with the proposed prefetching scheme were observed. The derived times using Equation (5) were 89.2 s with normal access and 4.2 s with the proposed prefetched scheme.

An extensive improvement of about 21-times higher-speed web service was obtained, and the values obtained by the test agreed with the theoretical ones.

Next, the throughput variation at every one second with the proposed scheme and normal access are shown in Figure 5-15. The UDP packets are generated 1414 bytes with the interval of 1 ms. In the proposed scheme, a maximum throughput of about 14 Mbps was obtained in the elapsed time of 3 s, where 1.792-MBytes data were transferred from 2.302 to 2.952 s. This shows that high- speed Web service can be achieved with the proposed scheme. The other hand, Normal access was low throughput and long elapsed time.

The UDP packet loss recovery function was finally evaluated. As a methodology for evaluation test, UDP retransmission test was carried out by cutting the network cable off for about 1 s to cause packet loss.

The UDP packets were captured in the information storage server, and the retransmission operation was observed. The cumulative rate of successfully received in the information storage server is shown in Figure 5-16. In this test, the network cable was cut off for about 2 to 3 s of the elapsed time. The results show that 100 % of the UDP packets were perfectly transferred with the scheme.

The information storage server sends a retransmission request when packets are not received after it has waited for a certain time from receiving one packet that is not the last packet, or when packets replying to a retransmission request are not received. In this test, the waiting time was fixed at 5 s, and

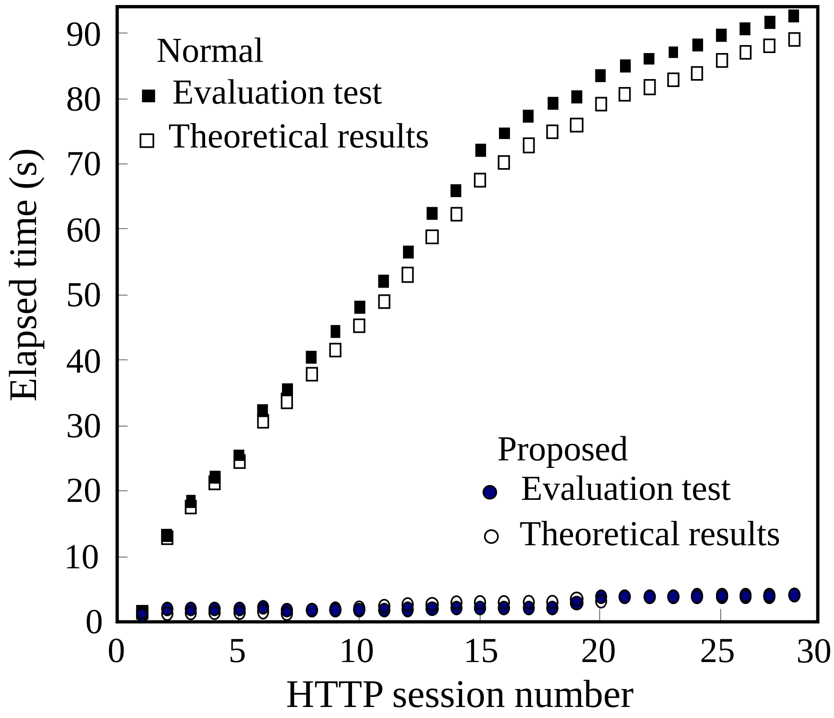


Figure 5-14 Evaluation test of the Web load times.

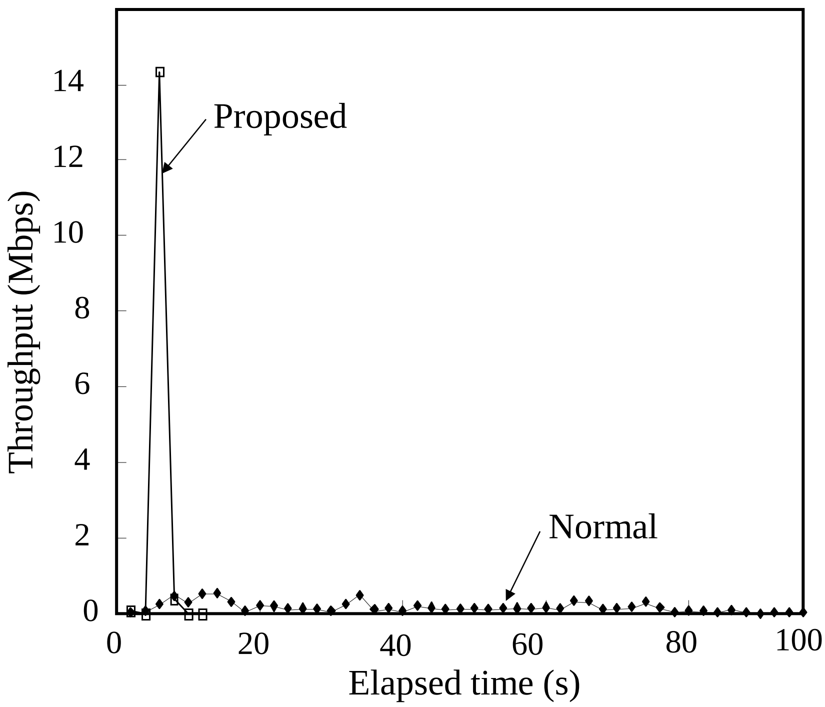


Figure 5-15 Throughput variation with proposed and normal access.

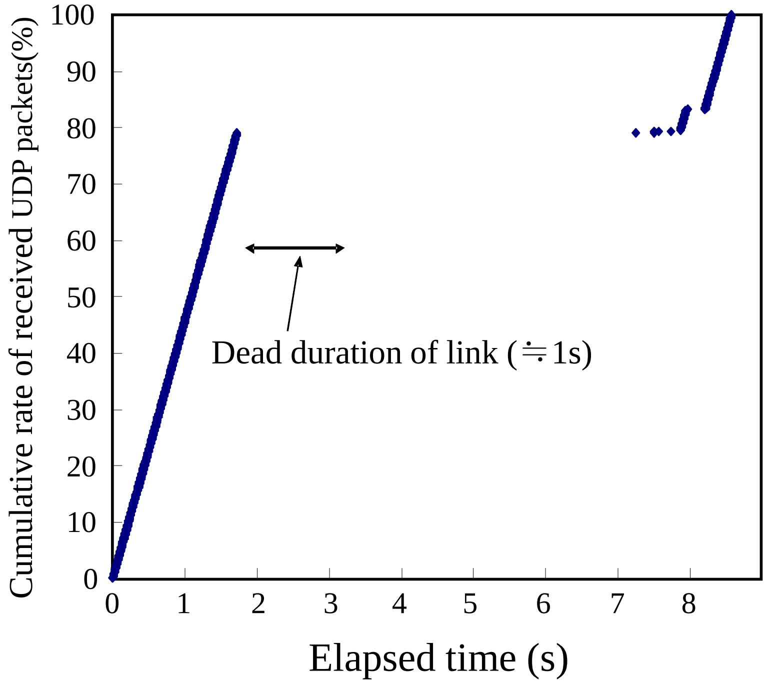


Figure 5-16 Evaluation of recovery function.

retransmission occurred after 7 s had elapsed. It is verified that reliable communication without any packet loss can reach with the installed scheme in the test-bed.

5.8 Conclusion

We investigated boosting web service performance over an extremely large latency network. We derived the theoretical performance and proposed a prefetching scheme that uses a prefetching server and an information storage server. The proposed scheme was implemented on a test-bed system. We carried out an evaluation test using the system and verified that the proposed scheme can enable high-speed Web service over an extremely large latency network. Reliable communication without any packet loss was also verified.

References for chapter 5

[5-1]D. Raggett, A. L. Hors, I. Jacobs, "HTML 4.01 Specifica-tion," W3C Recommendation 24 December 1999.

[5-2]T. Berners-Lee, R. Fielding, and H. Frystyk, "Hypertext Transfer Protocol -- HTTP/1.0.", RFC 1945 MIT/LCS, UC Irvine, May 1996.

[5-3]R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, T. Berners-Lee," Hypertext Transfer Protocol -- HTTP/1.1", Internet draft standard RFC 2616, June 1999.

[5-4]T. Berners-Lee, R. Fielding, and H. Frystyk, "Hypertext Transfer Protocol -- HTTP/1.0.", RFC 1945

MIT/LCS, UC Irvine, May 1996.

[5-5]R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", Internet draft standard RFC 2616, June 1999.

[5-6]X. Liang, F.L.C. Ong, P.M.L. Chan and R.E. Sheriff and P. Conforto, "Mobile Internet Access for High-speed Trains via Heterogeneous Networks," The 14th IEEE 2003 International Symposium on Persona1,lndoor and Mobile Radio Communication Proceedings, pp.177-pp.181, 2003.

[5-7]V. Schena, F. Ceprani, "FIFTH Project solutions demon-strating new satellite broadband communication system for high speed train," Vehicular Technology Conference, 2004. VTC 2004-Spring. 2004 IEEE 59th, pp.2831-pp.2835, 2004.

[5-8]IEEE Std 802.11, Part 11, "Wireless LAN Medium Ac-cess Control (MAC) and Physical Layer (PHY) Specifications," ISO/IEC 8802-11, 2008.

[5-9]H. Obata, K. Ishida, J. Funasaka, and K. Amano, "Data Transfer Time by HTTP1.0/1.1 on Asymmetric Networks Composed of Satellite and Terrestrial Links," IEICE Transactions on Communications, Vol.E85-B, No.12, pp.2895-2903, December 2002.

[5-10] G. Hasegawa, M. Murata and H. Miyahara, "Performance Evaluation of HTTP/TCP on Asymmetric Networks," In-ternational Journal of Communication Systems, Volume 12, Number 4, pp.281-296, July-August 1999.

[5-11] M. Allman, V. Paxson, and W. Stevens, "TCP Congestion Control", RFC 2581, April 1999.

[5-12] W.R. Stevens, "The protocols," in TCP/IP illustrated, vol.1, Addison-Wesley, Reading, Masachusetts, 1994.

[5-13] Z. Reserch, "The Need for Speed II," Zona Market Bulle-tin, Issue 05, April 2001.

[5-14] T. Hyashi, Y. Muto, and I. Nakajima, "Subjective Quality Evaluation on Response Time in Web Service," General Conference of IEICE, B-11-23, 2001.

[5-15] T. Yoshino, Y. Sugawara, K. Inagami, Junji Tamatsukuri, Mary Inaba and Kei Hiraki, "Performance Optimization of TCP/IP over 10 Gigabit Ethernet by Precise Instru-mentation," the Super-Compu-ting2008(SC08).USB-stick.Nov.2008.--Articles snd Conferences, 2008.

[5-16] S. Souders ,"High Performance Web Sites: Essential Knowledge for Frontend Engineers," O'Reilly Media, Sep. 2007.

[5-17] H. F. Nielsen, J. Gettys, A. Baird-Smith, E. Prud'hom-meaux, H. W. Lie and C. Lilley, "Network Performance Effects of HTTP/1.1, CSS1, and PNG", Proc. of ACM SIGCOMM'97 conference on Applications, technologies, architectures, and protocols for computer communication, 1997.

[5-18] K. Chinen, "An interactive prefetching proxy server for improvement of WWW latency," Proc. INET '97 Con-ference, 1997.

[5-19] V. N. Padmanabhan and J. C. Mogul, "Using Predictive Prefetching to Improve World Wide Web Latency," ACM SIGCOMM Computer Communication Review, Vol. 26, No. 3 July 1996.

[5-20] R. Shinkuma, M. Okada, and S. Komaki, "Adaptive Transmission Scheme for Web Prefetching

in Wireless Environment," IEICE Transaction on Electronics, Vol. E85-C, No.3, pp.485-491, March 2002.

[5-21] "MIME Media Types," http://www.iana.org/assignments/media-types/index.html.

6. Proposal of a suitable retransmission control technique for long distance communication system

6.1 Summary

This paper proposes a retransmission scheme in order to achieve high-speed transmission by reducing retransmission delay time caused by packet loss. High-speed transmission can be effectively realized by sending UDP packets continuously over large latency network system such as satellite communication system. Since UDP protocol characterizes connectionless communication, a retransmission scheme over UDP protocol must be used for reliable quality. However, transmission speed will ·drastically degrade by the time of repeatedly retransmission. The proposed scheme suits for long distance transmission of large latency. It is verified that the transmission time can be shorten 2s by theoretical calculation and 20 % of throughput improvement with PER = 0.1 by computation simulation.

6.2 Introduction

Long-distance wireless transmission systems have been widely used in various applications and services. Such conventional transmission systems cannot handle large amounts of data, resulting in long transmission delay times. For example, it is difficult to reasonably provide Web services owing to network delays. Efforts have been expended to overcome this difficulty. User datagram protocol (UDP) packet usage is effective for real-time transmission. Consecutive UDP packet transmission enables efficient data transmission [6-1], [6-2]. However, the UDP cannot compensate for packet loss; hence, a highly reliable scheme is essential for long-distance UDP wireless transmission.

TCP is a highly reliable communication protocol because of its flow control retransmission function [6-3] and [6-4]. This retransmission function uses a slow start procedure to decrease the transmission speed when packet loss occurs. TCP is highly efficient in multi-traffic flow environments, such as the Internet. However, this study assumes wireless relay by peer-to-peer networks, making TCP flow control unsuitable. Therefore, a retransmission function that is suitable for this condition is required. Moreover, same frequencies are assumed for both transmitter and receiver in order to enhance the frequency utilization. In such scenarios, automatic repeat-request (ARQ) [6-5] [6-6], which can reliably transmit data packets, is suitable for long-distance wireless transmission.

In this paper, the suitability of selective-repeat (SR) conventional longdistance ARQ wireless transmission is first described. Problems with SR are also discussed. Next, a novel retransmission scheme based on SR, which will be widely adopted in long-distance wireless transmissions, is proposed. The key advantage of the proposed scheme is that the retransmission waiting time can be dramatically reduced by transmitting data packets in descending order from the last data packet. Finally, it is shown that the proposed scheme outperforms conventional SR from the viewpoints of the average received time and throughput.

The rest of the paper is organized as follows. The conventional scheme and its issues are discussed in Section 2. The proposed scheme is described in Section 3. A computer simulation that verifies that the proposed scheme can better improve the average receiving time and throughput performance than the conventional scheme is presented in Section 4. Finally, conclusions of this paper and future studies are presented in Section 5.

6.3 Conventional retransmission control schemes and their issues

ARQ is classified into three main categories: stop-and-wait (SAW), go-back-N (GBN), and selective-repeat (SR) [6-7]. All these types of ARQs have advantages and disadvantages. SAW is suitable for short-range transmission and environments with high bit errors, such as wireless LAN systems, because of its simplicity. GBN is employed in the TCP/IP employed for the world wide web (WWW). Unlike SAW and GBN, SR is generally used for long-distance wireless communications that have long delays, such as satellite systems, and so on. We adopted SR in this study because of our focus on systems having long delays owing to long-distance transmission.

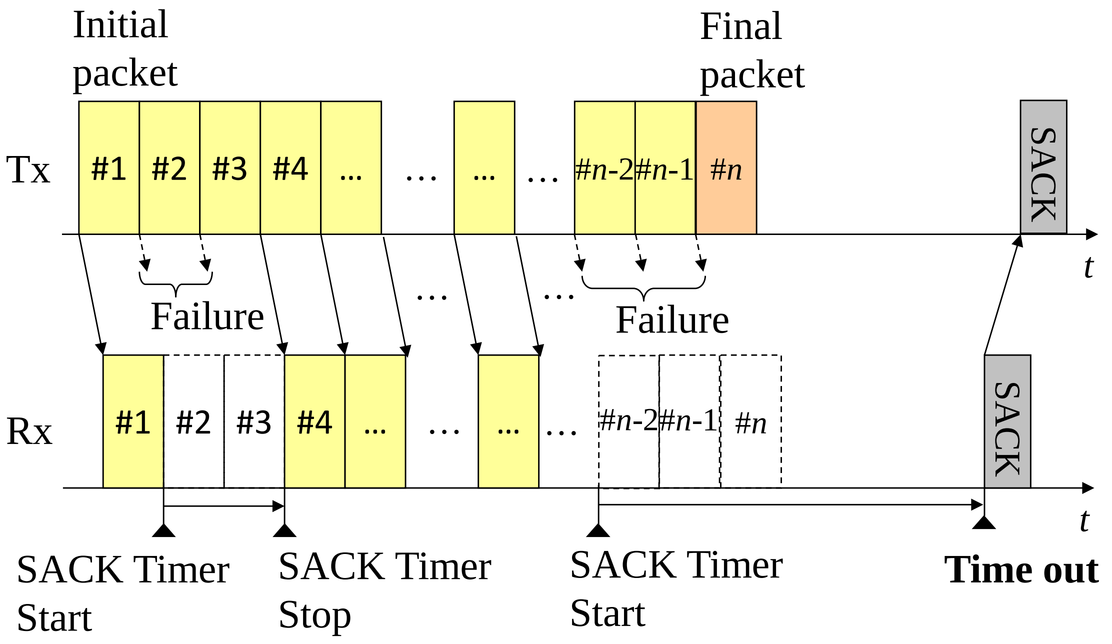
Fig. 6-1(a) shows a conventional SR transmission procedure. Acknowledgment (ACK) and Selective ACK (SACK) messages are used in ARQ to determine whether or not a data packet is successfully received at a receiver. As shown in Fig. 6-1(a), the data packets are successfully received in the conventional SR.

The receiver sends an SACK message, including the successful packet number information, to the transmitter after a certain time lag, called a “Timeout”, when the success/failure of reception of the last data packet is determined, regardless of other packet losses. Therefore, the transmitter will retransmit only failure packets because the successful data packet numbers are transmitted at the transmitter by the SACK message. The receiver waits for a certain time before the SACK, which is retransmitted if appropriate packets are not received. Retransmission is not employed at the transmitter site unless the SACK is received after the transmission of the last data packet.

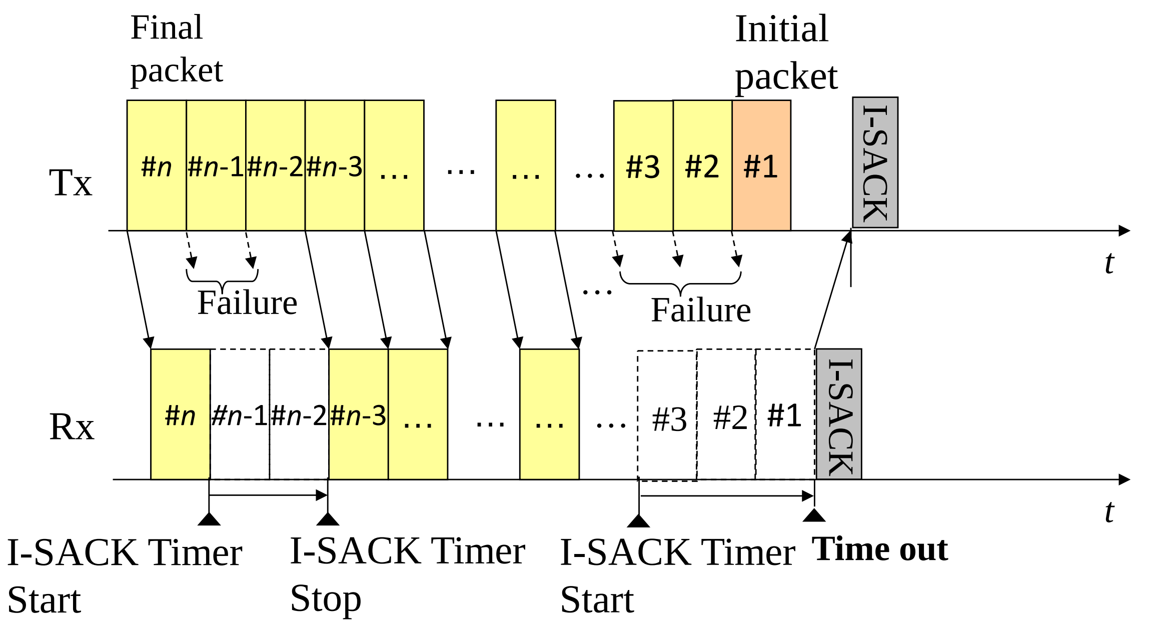
The receiver must wait for a period corresponding to the Timeout when the final data packet cannot be received. Hence, an SACK cannot be instantaneously sent by the receiver site. This problem results in system delays and degradation of frequency utilization. Increased transmission times owing to retransmission control schemes must be avoided in long-distance wireless transmission systems in which real-time TV broadcasts are employed.

6.4 Proposed scheme: instantaneous SACK (I-SACK)

The retransmission flow of the proposed scheme is shown in Fig. 6-1(b). The proposed scheme employs consecutive data packet transmission, similar to the conventional scheme. The key advantage of the proposed scheme is that data packets from the last data packet are transmitted so that the transmitter and receiver share the transmission time at which all the data packets are sent between them in advance. In other words, the data packets from the last data packet are transmitted in descending order in the proposed scheme. SACK can be instantaneously returned by the receiver to the transmitter by employing this



(a) Conventional transmission flow.



(b) Proposed transmission flow.

Figure 6-1 Conventional / proposed selective-repeat transmission procedures.

scheme. We call this scheme instantaneous SACK (I-SACK).

As shown in Fig. 6-1(b), five data packets are not received (#n-1, #n-2, #3, #2, and #1). However, SACK can be instantaneously returned, i.e., I-SACK can be transmitted from the receiver to the transmitter because the receiver can understand the total transmission time per burst in time. The same procedure is repeated when retransmission packet loss occurs. Obviously, the proposed scheme can reduce the total reception time and enhance the throughput as compared to that enhanced by the conventional scheme.

As the scheme of achieving operation equal with the proposed scheme, there is the scheme of using a negotiation function. In this scheme, negotiation between the transmitter and the receiver is employed before consecutive data packets are transmitted and the receiver obtains the information regarding the number of total consecutive data packets from the transmitter. This scheme enables a similar reduction in reception time to the proposed scheme, as the receiver knows the reception time of the final packet because of the negotiation. However, a transmission overhead is required for this negotiation, and the transmitting efficiency will decrease. Moreover, there will be a processing load at the transmitter because the number of total consecutive packets is calculated before transmitting. On the other hand, the proposed scheme is LIFO (last in first out), which simply reverses the consecutive data packets. The objective is achieved because the receiver obtains the sequence number.

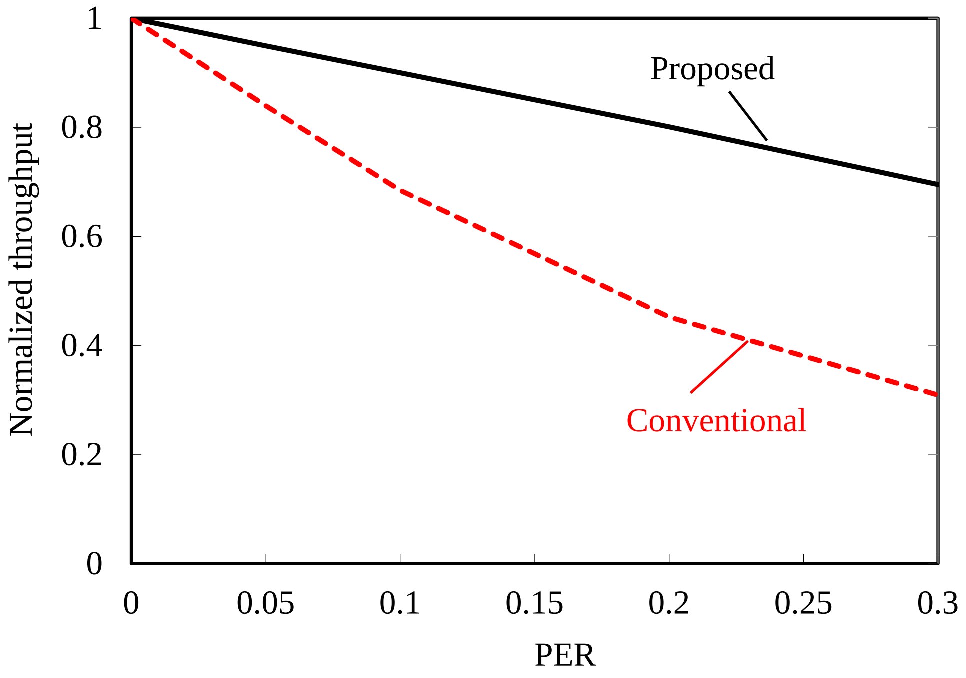
A system that uses the negotiation function is adopted for multiuser satellite communication, etc., on TDMA (time division multiple access) circuit switching, in order to reserve the bandwidth beforehand. However, the target of this study is for systems in which the traffic data volume is dynamically transmitted. If the proposed method is achieved, it will be adopted with a new mobile multimedia satellite communication system, etc., because the conventional system has not yet been achieved.

6.5 Effectiveness of the proposed scheme

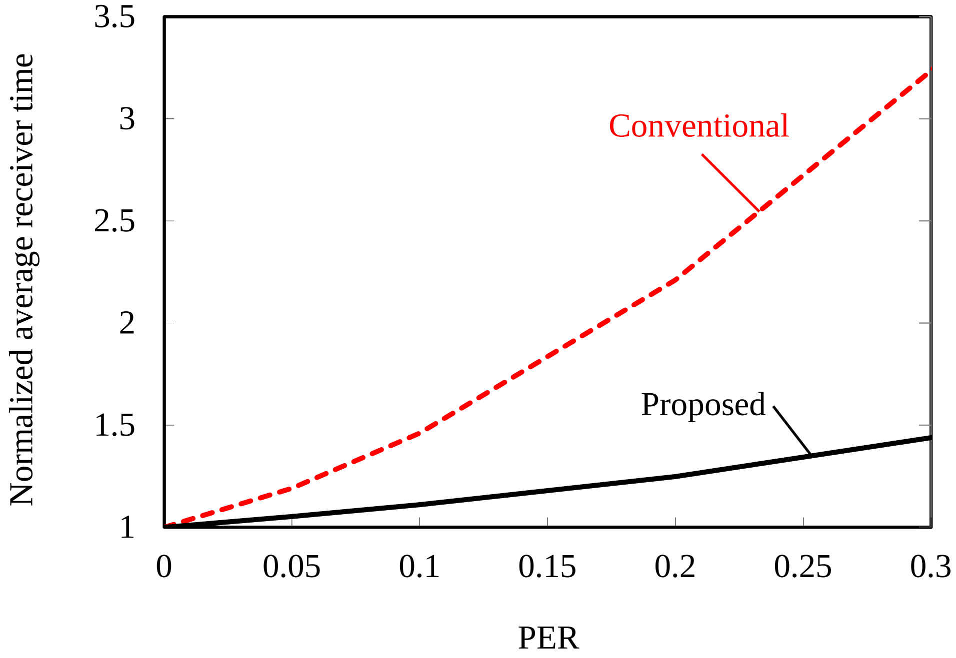
The average receiving time and throughput of the proposed and conventional schemes were determined by a computer simulation and compared in order to verify the effectiveness of the proposed scheme.

Table 6-1 Simulation parameters.

Transmission rate	1 [Mbps]
Number of bursts	1000
Number of data packets per burst	43
Size of data packet	1500 [byte]
ACK size	20 [byte]
Timeout (Conv. Scheme)	1 [s]
Method for generation of error	Random
Packet Error Rate	0 ~ 0.3



(a) Average received time



(b) Throughput

Figure 6-2 Throughput and average received time versus PER.

The proposed scheme assumes that the system does not achieve negotiation functions in a peer-to-peer network; therefore, the proposed scheme was compared with the conventional simple SR scheme using an SACK timer.

Table 6-1 lists the simulation parameters. OPNET 11.5A [6-7] was used for the computer simulation. The data packet size considered in this simulation, i.e., 1,500 bytes, is the maximum Ethernet flame size. In order to evaluate the long-distance communication with high data rate, the transmission rate and delay are set to be 100 Mbps and 100 ms, respectively. The packet error was set in the range 00.3 because the proposed scheme is expected to be more effective than the conventional scheme when many bit errors occur. Transmission delay was not considered in this evaluation, in order to confirm the basic performance of the retransmission control schemes. A transmission distance of 300 km was assumed in this evaluation, and the transmission delay was set as 1 ms.

Figure 6-2 shows the normalized average receiving times and throughput versus the packet error rates (PER) of the proposed and conventional schemes. Normalized averaged received power and throughput values were obtained by considering the results for PER = 0. The average receiving time indicates the time taken for reception to be completed within one burst, including retransmission control, as shown in Figure 6-2, in which the average receiving time of the conventional scheme is 1.8 times that of the proposed scheme for a PER of 0.1. However, the average receiving time for the proposed scheme did not increase much, regardless of the PER. As shown in Figure 6-2, a 20% improvement in throughput was obtained for the proposed scheme relative to the conventional scheme for a PER of 0.1. Therefore, the simulation results confirm that the proposed scheme is effective from the viewpoints of both the average received time and the throughput.

6.6 Conclusion

In this paper, a novel retransmission scheme for long-distance communication based on selective-repeat is proposed. The main advantage of this scheme is that the retransmission waiting time can be dramatically reduced by transmitting data packets in descending order from the last data packet. Computer simulation results confirm that the average receiving time can be approximately halved and a 20% improvement in throughput can be achieved with a PER of 0.1 using the proposed scheme. As a future work, the evaluation when considering concrete parameters in certain systems such as satellite communications is essential.

References for chapter 6

[6-1]J. Nakamura, A. Makishima, D. Umehara, and M. Kawai, ”High-speed data transmission performance on Ka-band GEO-synchronous satellite communication link using TCP”, 2002 Korea-Japan Joint Conference on Satellite Communications, Daejeon, Korea, pp. 33-38, Oct. 2002.

[6-2]K. Ienaga and M. Hirahashi, ”A Study of Fast File Transfer Protocol on PCs”, IEICE Technical Re-

port, NS2002-102, pp.35-40, Sep. 2002.
[6-3]RFC 793: Transmission Control Protocol.
[6-4]RFC 1323: TCP Extensions for High Performance.
[6-5]K. Hasegawa and I. Sasase, ”A Controlling Retransmission Size H-Arq Scheme by BLER based on results of CRC Decoding”, The 18th IEEE Annual International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC 2007), 54, Athens, Greece, Sep.. 2007.
[6-6]S. Lin, D. Costello Jr., and M. Miller, ”Automatic-repeat-request errorcontrol schemes”, IEEE Commun. Mag.,Vol.22, No.12, pp.5-17, Dec. 1984.
[6-7][Online] http://www.opnet.com/index.html

7. Conclusions

High-speed data transmission of broadband communication technology for providing multimedia contents in vehicles such as trains, airplanes, and ships have been carrying out developing and studying by using satellite communications flourishingly in late years. Furthermore, demands of multimedia communication to access the Internet in such vehicles with smartphones are emerged and explosively spread increasing. In the multimedia communication, users often download various kinds of multimedia contents such as texts, still images, animations, and sounds from the Internet through the forward link. On the other hand, data flow volume issuing from users for servers are small size of acknowledgement signals and low traffic through the return link. Since then, traffic feature of multimedia communication characterizes asymmetric among the forward link and the return link, high-speed data transmission in the forward link and small traffic of transmission in the return link. Considering the traffic feature of multimedia communication mentioned above, a network configuration scheme using Ku-band satellite communication system, which can provide high-speed data transmission with its wide band, applying for mobile multimedia satellite communication system for users in vehicles is proposed. "Low-profile and highly precise tracking antenna", and "layer 3 diversity reception technology with packet selection scheme to prevent link quality deterioration" are proposed for vehicles such as buses and railroads. It is clarified that those proposed schemes are substantially effective to realize receiving satellite signals with high link quality through various experiments equipped with on a practical vehicle. In addition, "techniques of http-transmission performance enhancement for WWW multimedia communication services", and "a new suitable retransmission control technique in long-distance communications" are proposed. It is clarified the effectiveness by the experiment with the test-bed system and computer simulations.

Chapter 1 describes the introduction and the background of this study.

Chapter 2 provides techniques and subjects to build a mobile multimedia satellite communication system for high-speed running vehicles economically and quickly expanding over wide area. This chapter also shows related matters with the middleware technology necessary for realization of the system.

Chapter 3 focuses the requirements for the satellite tracking antenna which is the key component of the system to realize the proposed mobile multimedia satellite communication system for vehicles such as buses, trains, and airplanes. Their antenna gain is too small to catch CS signals, which is our target satellite because the equivalent isotropically radiated power (EIRP) of a BS is about 5 times larger (89dBm) than that of a CS (81.7dBm), the developing antennas are required higher-gain and precisely tracking by the beam narrowing of the gain. Gyro tracking scheme, which uses a gyro to detect direction of the satellite, and step truck scheme, which let the receiving signal level drive the highest to turn an antenna, are having utilized to track the satellite. Tracking accuracy is not enough to track the satellite with these schemes when narrow beam band by pointing error. The phase monopulse schemes, dividing an antenna into subarrays and detecting phase difference of receiving signal from each subarrays, are effective for tracking accuracy. In this study, a low-profile tracking antenna that uses a phase monopulse scheme for mounting on a roof of practical-vehicle is proposed. The antenna is composed of four planar antennas by aspect composition and turns mechanically. The developed antenna is mounted onto a practical bus and the mobile multimedia satellite communication system was constructed practically. The tracking accuracy performance of the antenna is demonstrated by running the bus and network performance of the practical system is evaluated. It is confirmed that the tracking accuracy within 1dB can be achieved in the running environment of a bus. It is verified that the satellite signal can be received stably in a Ku band satellite communication system in practical mobile environments of running the bus on roads and highways or practical trains.

Chapter 4 focuses the requirements for receiving stably the satellite signal in large vehicles like trains for high quality communication and proposes packet selection schemes of layer-3 diversity receiving technology by using more than two tracking antennas or receiving systems with gap-filler system. Link quality of the mobile satellite communication system deteriorates in shadowing conditions when satellite signals are blocked by buildings, tunnels and so on for trains and buses. Diversity combining is the technique applied to combine the multiple received signals of a diversity reception device into a single improved signal for overcoming link quality deterioration by shadowing conditions. Selection combining: Of the N received signals, the strongest signal is selected. Therefore, any additional gain diminishes rapidly with the increasing number of channels. Switched combining: The receiver switches to another signal when the currently selected signal drops below a predefined threshold. This is a less efficient technique than selection combining. Equal-gain combining: All the received signals are summed coherently. Maximal-ratio combining is often used in large phased-array systems: The received signals are weighted with respect to their SNR and then summed. Processing to revise the phase of control and the signal which changed a reception system by signal reception level strength is complicated by those methods. Since these schemes are in principle having assumed to use the same wireless system among diversity reception system, it has not been considered that diversity reception among wireless systems of different low layer media where the satellite signals cannot be received. Therefore the possible scheme of diversity reception among different networks of low layer media is proposed where the satellite signals can be reached by the gap filler (relay) systems such as the wireless LAN. The proposed scheme is characterized that the satellite signals, which are received and decoded by two or more receiver systems of the satellite tracking antennas mounted on a vehicle, are sorted out one only packet among decoded equivalent packets from each receiver. Laboratory tests under shadowing condition environment of running train were carried out by using a prototype system equipped with the proposed diversity scheme, and it is verified that equivalent performance of FTP in the condition of without shadowing environment can be obtained. In addition, it is confirmed that the packet order reversal can be prevented by demonstrating UDP packet transmission experiment.

Chapter 5 and chapter 6 describe HTTP performance enhancement schemes and retransmission schemes to provide comfortably the WEB service, which is the most popular application in the Internet, for users of vehicles in large latency network system.

Chapter 5 proposes a web prefetching scheme that can enhance the HTTP throughput of WEB service in large latency network systems. Mathematical evaluation and throughput performance test in the test-bed system with the prototype are demonstrated. The WEB services are characterized that dozens of about 10 kBytes multimedia content are transmitted in each communication sessions. Because a transmission distance of satellite communications network is long about 72,000km, transmission latency is extremely larger than terrestrial communications network. When transferring a multimedia content of such small size under large latency environment, the transferring is completed in slow speed period due to the slow start mechanism that is congestion control of TCP. Such a low-speed transfer over all sessions of each contents is carried out in the WEB services. The transfer speed threshold is restricted by concurrent connection number that can transfer at the same time written in RFC 2616. The WWW transfer time is restricted even if expanding the bandwidth of the satellite link or TCP receive window size, that is one of the performance enhancement technologies in the TCP layer, under such large latency. WEB prefetching technology, that is the application layer of the WEB service, to improve the HTTP throughput has been considered conventionally. Installing a prefetching server of the conventional technology in half points in the delay line between clients can make the greatest effect. That means installing the conventional prefetching server in the satellite, and it is not realistic. Therefore, a new prefetching scheme to enable HTTP performance enhancement in the mobile multimedia satellite communication system is proposed. The two functions in the conventional prefetching function scheme, which are the prefetching function and WEB contents information storage function, were distributing to install prefetching function in the WEB server network and to install WEB contents information storage function in the user's network. The prefetched WEB contents by the prefetching server are transferred to the storage server by a fast transmission protocol. A prototype system equipped with the proposed scheme was developed and a test-bed system equivalent for the environment of the mobile multimedia satellite communication system was constructed in a laboratory. Network performance test was carried out with the test-bed system. It is clarified that the enhancement effect of the HTTP performance.

Chapter 6 describes a new retransmission scheme based on a Selective Repeat (SR) technique in extremely large latency networks for compensate high transmission speed under link conditions of high packet loss ratio. Simulation evaluation of performance effect with the scheme by packet loss ratio is discussed. Large latency via satellite network may cause speed restriction of TCP transmission which is well used on the Internet. Transferring data packet continually by the User Datagram Protocol (UDP) can make high transmission speed over such a large latency network under packet lossless condition. Since UDP is a connectionless protocol, a retransmission control technique to compensate for packet losses to get high quality communication is required. In large latency network, SR scheme is suitable as a retransmission control technique. In the SR scheme, the receiver notifies the transmitter SACK, in which indicates the sequence number of the reception packet as the retransmission request, when the last packet of the data is reached at the receiver or after a certain period (Timeout) of the packets not reached at the receiver. The transmitter receives the SACK and retransmits the lost packets in the receiver, effective and useful transmissions of bandwidth can realize high transmission speed. The receiver waits for certain period of reaching retransmitted packets after notifying SACK, and notifies SACK again at the timeout mentioned above or when the last packet of the data is reached. The transmitter does not retransmit when SACK is not reached after sending the last packet of the data. The receiver waits during for the timeout period when not reaching the last packet which is the trigger of sending SACK. The timeout period may cause the deterioration of the transmission rate in the conventional SR scheme. In this chapter, a new retransmission control scheme suitable for long distance transmission network was proposed. The proposed scheme characterized to transmit in reverse order from the last packet to the first packet. Since then, the receiver can evaluate the time of reaching the last packet (the end time of reception) by receiving a sequence number in decrement order from the last number. When the end time of reception, the receiver notify the transmitter SACK as retransmission request (instantaneous-SACK, called I-SACK below) and the transmitter retransmits according to I-SACK information. The effect of the proposed scheme is clarified by a computer simulation that transmission time was shortened for approximately 0.2 seconds and the throughput was improved by approximately 20% when PER is 0.1.

Dissertation-related primary publications

・**Referred journal publications**

[1] Fumiaki NAGASE, Hiroshi TANAKA, Masayoshi NAKAYAMA, Tomohiro SEKI, Hiroshi KAZAMA and Hideki MIZUNO,” Mobile Multimedia Satellite Communication System at Ku Band,” IEICE TRANS. COMM., VOL. E84-B, NO.4, APRIL 2001.

[2] Fumiaki Nagase, Hiroshi Tanaka, Masayoshi Nakayama, Jin Mitsugi, Masazumi Ueba, and Hideki Mizuno, “Mobile Multimedia Satellite Communication System,” NTT R&D, vol.51, (2002), pp.317-324,2002.

[3] Hironori Ishikawa, Jin Mitsugi, Fumiaki Nagase, Masazumi Ueba, and Mitsunobu Watanabe, “A Feasibility Study on the Program Tracking of the Satellite Tracking Antenna Mounted on Train,” Journal of the Japan Society for Aeronautical and Space Sciences, Vol. 52, No. 607, pp.337-341, 2004.

[4] Fumiaki NAGASE, Jin MITSUGI, Masayoshi NAKAYAMA, Masazumi UEBA, “Layer 3 Diversity Reception Technology for Ku-Band Mobile Satellite Communication System”, IEICE Trans. Comm. Vol.E89-B No.6 pp.1856-1861, (2006).

[5] F.Nagase, T.Hiraguri, A.Kishida, K.Nishimori, and H.Makino, "A new long-distance communication retransmission control scheme, " IEICE CoMEX, Vol.1, No.3 pp.113-118, Aug. 2012.

[6] F.Nagse, T.Hiraguri, K.Nishimori, and H.Makino, "Web Acceleration by Prefetching in Extremely Large Latency Network," American Journal of Operations Research, Vol.2, No.3, pp.339-347, Sept. 2012.

・**Referred international conference publications**

[1] Fumiaki Nagase, Hiroshi Tanaka, and Masazumi Ueba, “IP Network Performance in Mobile Environments over Ku Band Satellite Links”, Proceedings of 2001 Korea-Japan Joint Conference on Satellite Communications (JC-SAT2001), pp.77-pp.82, Kyoto, Japan, Oct. 2001.

[2] Fumiaki Nagase, Hiroshi Tanaka, Jin Mitsugi, and Masazumi Ueba,” IP NETWORK PERFORMANCE OF A Ku BAND MOBILE MULTIMEDIA SATELLTE SYSTEM”,AIAA-paper-2002-2061, 20th AIAA ICSSC, May, 2002.

[3] Fumiaki Nagase、Jin Mitsugi、Masayoshi Nakayama、Masazumi Ueba, "Ku Band Mobile Multimedia Satellite Communication System for Trains", 21st AIAA ICSSC, AIAA-2003-2205, 2003.

[4] Fumiaki Nagase, Jin Mitsugi, Masayoshi Nakayama and Masazumi Ueba,” Shadowing Countermeasure Technology with Layer 3 Diversity Reception for Ku band Mobile Satellite Communication Systems”, 22nd AIAA ICSSC, AIAA-2004-3148,2004.

[5] Yutaka Imaizumi, Satoshi Harada, Fumiaki Nagase, Masayoshi Nakayama, Kohei Ohata and Masazumi Ueba. ”Network performance evaluations for Ku band mobile satellite communications system based on Shinkansen railroad's propagation conditions”, Vehicular Technology Conference, 2005. VTC-2005-Fall. 2005 IEEE 62nd, 2005.

Acknowledgements

I wish to express my gratitude to Professor Hideo Makino and associate professor Kentaro Nishimori for their invaluable advice and for their continuous encouragement. I am most grateful to Professor Hiroyoshi Yamada, Professor Masanobu Yamamoto and Professor Toru Kiryu for giving me the opportunity to accomplish this dissertation and for their invaluable advice. My special thanks also express to associate professor Takefumi Hiraguri and Mr. Akira Kishida for fruitful discussions.

I wish to express my deep gratitude to Professor Hideki Mizuno and Professor Masahiro Umehira who led me to the mobile multimedia satellite communication research and development fields. Their advices always have been encouraging me since joining NTT laboratories. Furthermore, I would like to thank Mr. Hiroshi Hojo, Mr. Hiroshi Kazama and Mr. Toshihiro Manabe for their inspiring and valuable technical suggestions.

I would especially like to express gratitude to Professor Masazumi Ueba, Mr. Katsuhiko Araki, Mr. Kohei Ohata for their continuous encouragement and fruitful suggestions. I am also strongly indebted to thank Professor Hiroshi Tanaka, Professor Jin Mitsugi, Mr. Masayoshi Nakayama, Mr. Tetsu Sakata and Dr. Tomohiro Seki for their valuable technical suggestions, continuous encouragement and helpful guidance over the years. I wish to express my appreciation to Professor Akira Meguro, Dr. Kiyoshi Kobayashi, Mr. Koji Yamamoto, Mr. Mamoru Ogasawara, Mr. Yoshinori Nakasuga, Dr. Katsuya Nakahira, Mr. Yasuyoshi Kojima, Mr. Yutaka Imaizumi, Mr. Yuichi Sagawa, Dr. Fumio Kira, Mr. Hideki Tobe, Dr. Kazuhito Ebina, Mr. Makoto Umeuchi, Mr. Hironori Ishikawa, Dr. Yoshinori Suzuki, Dr. Takana Saitoh, Mr. Hideki Toshinaga, Mr. Kazuhiro Tanabe, Dr. Fumihiro Yamashita, Mr. Tomihiko Yoshida, Mr. Kengo Nagata, Mr. Nobuyuki Kataoka, Mr. Hiroshi Ohashi, Mr. Shinichi Komai, Mr. Shiro Sumitani, Mr. Kenji Akita, Mr. Seiich Hosono and all of my colleagues for many useful discussions and helpful guidance.

I would like to thank my parents for a grant that made it possible to complete this study. Finally, I dedicate this dissertation for my daughter Rina, my son Hitoshi, and my wife Saori of their devoted supports.

Early spring, 2013, Fumiaki Nagase