OPIC
OFFICE DE LA PROPRIÉTÉ INTELLECTUELLE DU CANADA

CIPO
CANADIAN INTELLECTUAL PROPERTY OFFICE

(12) (19) (CA) **Demande-Application**

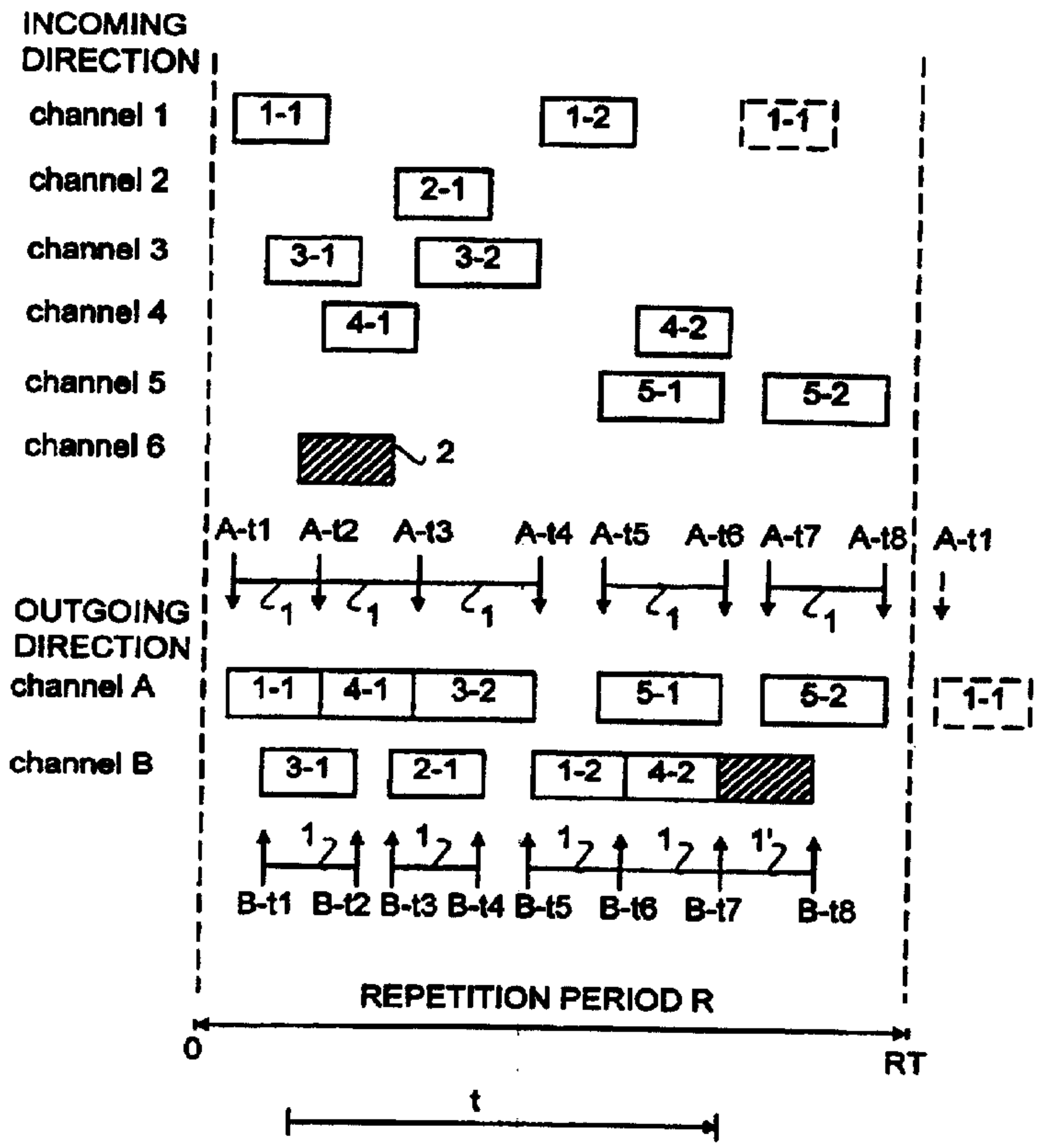
(21) (A1) **2,289,924**
(86) 1998/05/12
(87) 1998/11/19

(72) HIPPELÄINEN, LEO, FI
(71) NOKIA NETWORKS OY, FI

(51) Int.Cl.^{7} H04L 12/56, H04Q 7/22
(30) 1997/05/13 (972039) FI

(54) **PROCEDE DE TRANSMISSION DE DONNEES COMMUTEE PAR PAQUETS**

(54) **A METHOD FOR PACKET SWITCHED DATA TRANSMISSION**



(57) L'invention concerne la transmission d'appels commutée par paquets et notamment la transmission d'appels vocaux commutée par paquets, en temps réel. L'objet de l'invention est de fournir un procédé permettant l'utilisation efficace de la capacité de transmission sans détériorer la qualité des appels en temps réel et consiste à transmettre des informations en mode commuté par paquets, de manière que les retards de transmission restent courts et le retard de transmission d'un et même appel ne varie pas à l'extrémité de réception. Le procédé de l'invention consiste à définir une période de répétition (R) sensiblement plus longue

(57) The invention relates to packet-switched transmission of calls and particularly to real-time packet-switched transmission of speech calls. The object of the invention is to provide a method that allows efficient use of the transmission capacity without impairing the quality of real-time calls: information can here be transmitted by a packet-switched method so that the transmission delays remain short and the transmission delay of one and the same call does not vary at the receiving end. The method of the invention comprises defining a repetition period (R) that is notably longer than the duration of the transmission of a single packet,

OPIC
OFFICE DE LA PROPRIÉTÉ INTELLECTUELLE DU CANADA

CIPO
CANADIAN INTELLECTUAL PROPERTY OFFICE

(21) (A1) **2,289,924**
(86) 1998/05/12
(87) 1998/11/19

que la durée de la transmission d'un seul paquet, à définir le point de départ (0) de la période de répétition, à recevoir un paquet (2), à affecter une tranche de temps (1') de la même durée que celle du paquet à partir de la période de répétition pour la connexion à laquelle le paquet est associé, et à transmettre le paquet dans la même tranche de temps affectée et lorsqu'un nouveau paquet associé à la même connexion est reçu, à transmettre le paquet dans la tranche de temps affectée à la connexion. Un autre objet de l'invention est d'établir un noeud de réseau dans un réseau de transmission de données.

defining the starting point (0) of the repetition period, receiving a packet (2), allocating a time slice (1') of the same duration as the packet from the repetition period for the connection that the packet is associated with, and transmitting the packet in the time slice allocated and when a new packet associated with the same connection is received, transmitting the packet in the time slice allocated for the connection. Another object of the invention is to provide a network node in a data transmission network.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04L 12/56, H04Q 7/22 — A2

(11) International Publication Number: WO 98/52327

(43) International Publication Date: 19 November 1998 (19.11.98)

(21) International Application Number: PCT/FI98/00401

(22) International Filing Date: 12 May 1998 (12.05.98)

(30) Priority Data:
972039 13 May 1997 (13.05.97) FI

(71) Applicant (*for all designated States except US*): NOKIA TELECOMMUNICATIONS OY [FI/FI]; Keilalahdentie 4, FIN–02150 Espoo (FI).

(72) Inventor; and
(75) Inventor/Applicant (*for US only*): HIPPELÄINEN, Leo [FI/FI]; Sysimiehenpolku 8 E, FIN–00670 Helsinki (FI).

(74) Agent: KOLSTER OY AB; Iso Roobertinkatu 23, P.O. Box 148, FIN–00121 Helsinki (FI).

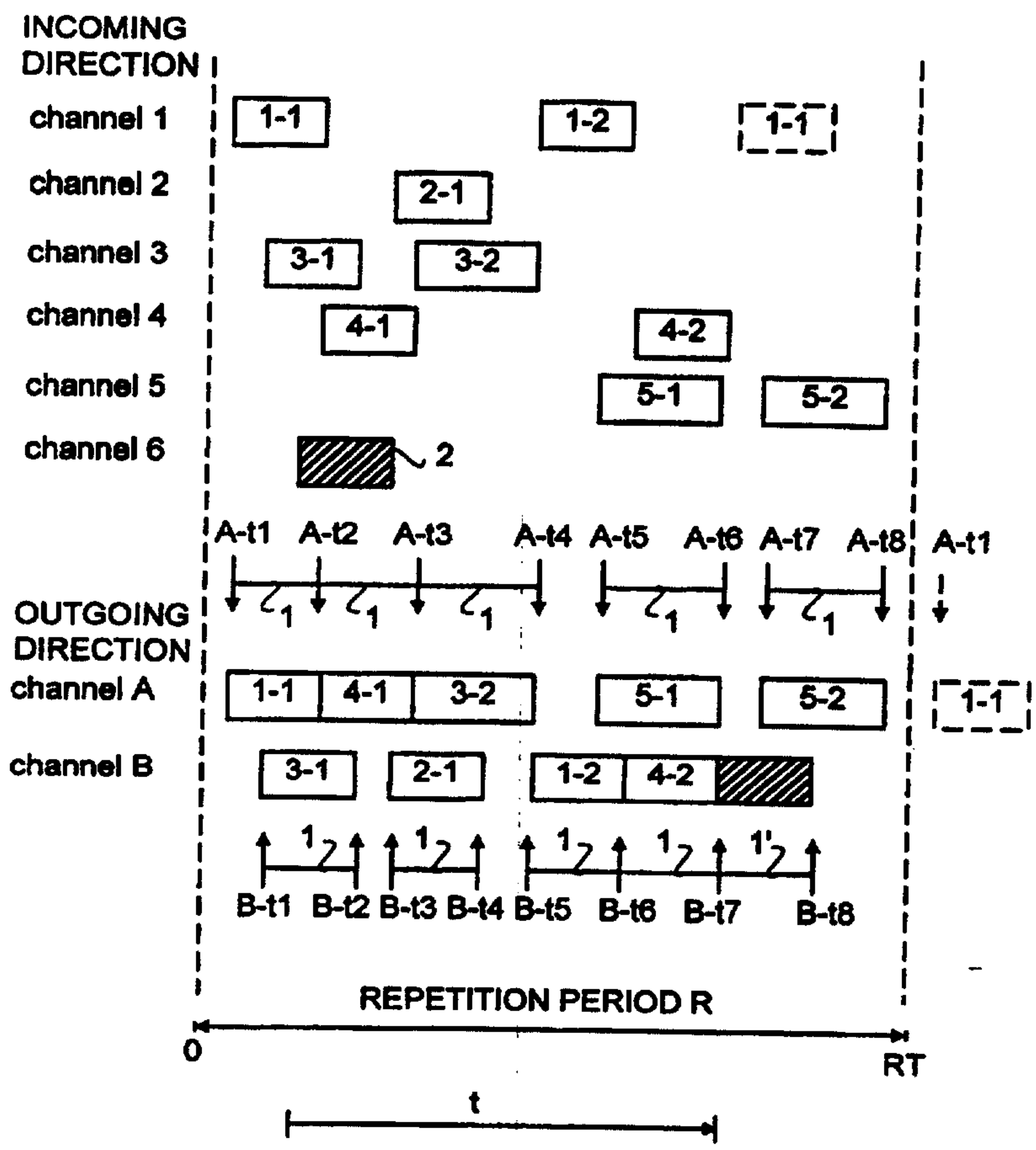
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
In English translation (filed in Finnish).
Without international search report and to be republished upon receipt of that report.

(54) Title: A METHOD FOR PACKET SWITCHED DATA TRANSMISSION

(57) Abstract

The invention relates to packet–switched transmission of calls and particularly to real–time packet–switched transmission of speech calls. The object of the invention is to provide a method that allows efficient use of the transmission capacity without impairing the quality of real–time calls: information can here be transmitted by a packet–switched method so that the transmission delays remain short and the transmission delay of one and the same call does not vary at the receiving end. The method of the invention comprises defining a repetition period (R) that is notably longer than the duration of the transmission of a single packet, defining the starting point (0) of the repetition period, receiving a packet (2), allocating a time slice (1') of the same duration as the packet from the repetition period for the connection that the packet is associated with, and transmitting the packet in the time slice allocated and when a new packet associated with the same connection is received, transmitting the packet in the time slice allocated for the connection. Another object of the invention is to provide a network node in a data transmission network.



A METHOD FOR PACKET SWITCHED DATA TRANSMISSION

BACKGROUND OF INVENTION

The invention relates to packet-switched transmission of calls and particularly to real-time packet-switched transmission of speech calls on the circuits of a telecommunication network of a mobile telephone system.

In most digital mobile telephone systems the carrier wave of the radio path is divided between several users by using, for example, a TDMA (Time Division Multiple Access) or CDMA (Code Division Multiple Access) method. The common feature of the methods is that the call is divided into call packets of a certain length and the packets are transferred in frames on the radio path. On the circuits of a network infrastructure, calls are usually transmitted using a circuit-switched system. If digital speech coding, such as ACELP (Algebraic Code-Excited Linear Predictive), is used on the radio path, the packets received from the radio path in packet form must be disassembled or, correspondingly, assembled when they are to be transmitted onto the radio path. This causes extra work particularly if both parties use radio telephones of the same mobile telephone system.

In a circuit-switched network a different transmission channel, or even a different channel in each direction of transmission, is allocated for each call. In a circuit-switched system a channel remains allocated even if nothing is transmitted, whereby the efficiency is poor. The efficiency of the circuit-switched system can be improved by reducing the transmission rate. A low transmission rate, however, prolongs the transmission delay. In a mobile telephone system where information is transmitted in packets on the radio path, the relevant delay is the delay encountered by the last bit of the packet, since the packet cannot be coded for transmission on the radio path until the whole packet is available. The packet may thus miss the time slot allocated for it, whereby the transmission capacity will not be in efficient use and the quality of the call may be impaired.

Another problem in the circuit-switched network is that as the load increases the capacity will be abruptly finished, whereas in the packet-switched network an increase in the load appears as a reduction in the transmission rate and deterioration of the quality of the connection.

In the packet-switched network the arrival rate of the packets at the receiving end varies, for example, with the load of the network nodes on the way. If the periods of transmission of the packets, i.e. the arrival rate, at the receiving end vary too much, the quality of speech will be impaired where speech calls are concerned. In speech transmission it is important that the maximum delays are short and that the delays remain unchanged throughout the entire speech item. In other words, the transmission delays of the speech call packets must be as identical as possible. In addition, any variation in the transmission times of the data call packets in mobile telephone systems is also undesired, since the call packets must be sent onto the radio path at an even rate in accordance with the radio protocol.

One way of overcoming the above problem, i.e. the variation in the transmission times in the packet-switched network, is to transfer the packets by emulating packet-switching, i.e. by allocating a separate channel or channels for them. The problems in this solution are the same as in the circuit-switched solution, i.e. capacity is allocated even through it may not be needed, and as the load increases the capacity will be abruptly finished.

Another way of levelling the transmission delays of the packets in the packet-switched network is to use a very high-rate transmission channel, whereby the transmission time of the packets is short as compared with the interval between the packets. The problem in high-rate transmission networks is that the networks are expensive and that if the rates on the different channels of the transmission network are different, a speech packet must be received in full before it can be transmitted to a faster channel. This may cause an additional delay as compared with a situation where the entire transmission network operates at the same rate, whereby the beginning of a packet may arrive before the end has even been transmitted.

Yet another alternative is to buffer a sufficient number of packets at the receiving end so as to level any variation in the transmission delays. The problem is that the real-time transmission of speech does not allow large buffering: in a public safety and security call, for example, the delay must be less than 400 milliseconds. In the Internet the speech-transmitting software uses large buffering, for a delay of a couple of seconds is irrelevant there in speech transmission. Since speech must be put through as quickly as possible in a system transmitting real-time speech, the buffer must be very short, and so buffering cannot be relied on to level the transmission delays.

BRIEF DESCRIPTION OF INVENTION

The object of the present invention is to provide a method and a network node enabling a packet-switched transmission network transmitting real-time speech. The object is achieved by a packet-switched data transmission method of the invention, the method being characterized by defining a repetition period that is considerably longer than the time needed for sending a single packet and allocating time slices of a different length from the repetition period for the connections; defining the starting point of the repetition period; maintaining information on the current point in the repetition period; maintaining information on the time slices allocated from the repetition period and on the connections associated with them; receiving a packet; identifying the connection associated with the packet as well as the destination address from the identification data of the packet; checking whether a time slice has been allocated for the connection of the packet from the repetition period; if yes, starting the transmission of the packet when the starting point of the time slice allocated is reached in the repetition period; if no, defining the duration of the packet transmission time; searching the repetition period for a free period that is at least as long as the transmission time; allocating from the free period of the repetition period a time slice which is at least as long as the packet transmission time for the connection of the packet; and starting the transmission of the packet when the starting point of the time slice allocated is reached in the repetition period.

A connection here means either a data transmission, or call, connection, or a connection on which signalling information is transmitted between network elements. In systems where a call is handled as separate speech items, a connection means a single speech item.

The invention also relates to a network node that can be utilized in the method of the invention. A network node of a packet-switched data transmission network according to the invention, to which there leads at least one incoming channel and from which there leads at least one outgoing channel and which comprises transmission means for transmitting packets toward a destination address, is characterized by further comprising setting means for setting the duration of a repetition period to be notably longer than the transmission time of a single packet; at least one time counter, which resets to zero at an interval of one repetition period, to indicate the current point in the repetition period so as to time the transmission of the packets; identification means for identifying the connection associated with the packet received and for detecting whether packets associated with the connection have already been received in the network node before said packet or whether the packet is the first packet associated with the connection; allocation means for searching the repetition period for a free time slice that is at least as long as the packet transmission time and for allocating a time slice that is as long as the packet transmission time from the free time slice for the connection associated with the packet in response to the reception of the first packet associated with the connection, the starting point of the allocated time slice indicated by the time counter defining the starting moment of the transmission of the packets associated with the connection; and control means for transmitting packets toward destination addresses in the time slices allocated for the connections associated with the packets so that the packets associated with the same connection are transmitted from the network node one packet at a time at an interval of one repetition period.

The invention is based on the idea that the transmission delays of the packets are standardized by allocating a transmission slot of the size of a packet for each connection in the network node, whereby the connection packets associated with a single connection are forwarded at certain intervals. The advantage of the invention is thus that the transmission capacity is in efficient use without that the quality of the calls is impaired, since information can be transmitted by a packet-switched method so that the transmission delays remain short and the transmission delay of one and the same call does not vary at the reception end. The invention is the most advantageous in simplex (half duplex) communication, where data can be transmitted in both directions but only in one direction at a time. Packet-switched packets are supplied and resources are allocated only in the direction where information is being transmitted, whereas in the circuit-switched transmission channels must be allocated in both directions and yet only one channel is used at a time. The transmission capacity of the circuits can thus be, at best, doubled with the invention.

Another advantage of the invention is that a separate channel need not be allocated for signalling, but the signalling information is transferred on the same channels as other data. This further improves the efficiency.

Yet another advantage of the invention is that it can be applied to very different systems and that no empty spaces need to be transmitted in the packets in addition to the information, since the invention does not set any limits to the size of the packet. To each connection it is possible to forward packets whose length differs from the lengths of the packets of the other connections.

In a preferred embodiment of the invention the network nodes deallocate a time slice if there have been no packets to transmit to the connection for two times in succession. This has the advantage that time slots are not allocated in vain but the whole capacity is in efficient use. A further advantage is that the amount of signalling is reduced, since control information on the termination of a call need not be separately sent to the network nodes to deallocate resources.

In another preferred embodiment of the invention, control packets containing signalling information are separated from call packets and transmitted between the call packets whenever there is enough time. The call packets contain speech or data. The advantage is that transmission capacity can be used more efficiently when the less time-critical control packets 'give way', giving priority to the more time-critical call packets. This further ensures that the transmission delays of the call packets are short and even.

The preferred embodiments of the claimed method and network node appear from the attached dependent claims 2 to 6 and 8 to 11.

BRIEF DESCRIPTION OF DRAWINGS

In the following the invention will be described in greater detail in connection with preferred embodiments and with reference to the attached drawings, in which

Fig. 1 is a block diagram of a transmission network according to the TETRA system,

Fig. 2 illustrates the relationship between the packets of the radio path and those transmitted in the transmission network,

Fig. 3 illustrates the allocation of time slices according to a first preferred embodiment on the incoming and outgoing channels of a network node,

Fig. 4 is a block diagram of a network node,

Fig. 5 is a flow diagram illustrating the division of the packets into control packets and connection packets, and

Fig. 6 is a flow diagram illustrating the operation of a network node.

DETAILED DESCRIPTION OF INVENTION

The invention will now be described by way of an example in a packet-switched telecommunication network that is designed to act as a transmission network of a digital radio network according to the TETRA standard (Trans-European Trunked Radio) defined by the ETSI (European Telecommunications Standards Institute). The TETRA standard provides standards, for example, for interfaces to other networks, an air interface and an interface to another TETRA network. The internal structure of the transmission network, however, is not standardized in the TETRA standard, so a TETRA network is a good example of a network. In addition, the TETRA standard defines diverse highly time-critical call and network services, since the network defined by the TETRA standard is also designed to be used as a private mobile radio network by the authorities. The invention, however, is not limited to radio networks or other wireless networks, but it will be obvious to a person skilled in the art that the invention can be applied to other data transmission systems both in networks based on wireless data transmission and in fixed networks.

Fig. 1 shows an exemplary structure of a TETRA transmission network. A Mobile station MS (Mobile Subscriber) communicates with a Base Station BS over the radio path Air. The radio interface is defined in the TETRA standard. Each base station BS contains a Node N, which is connected by a circuit to a Digital Exchange DXT for TETRA of the fixed transmission network. The TETRA exchanges DXT are connected to other exchanges DXT and to a Digital Central Exchange DXTc for TETRA by a fixed circuit, the DXTc being an exchange to which are connected other exchanges DXT and/or other central exchanges DXTc so as to provide alternative paths for the traffic. The interface to another TETRA network is here arranged in the central exchange DXTc, but it can also be located in the other exchanges DXT. The external interfaces, defined by the standard, to the Public Switched Telephone Network PSTN, the Integrated Services Digital Network ISDN, the Private Automatic Branch Exchange PABX and the Packet Data Network PDN are here located in one exchange DXT, but they can also be arranged, for example, in every exchange. The TETRA transmission network also comprises other interfaces and peripheral units, which are not shown in the figure. They include, for example, network management systems and dispatcher systems.

In the TETRA system or the like a call is handled as separate speech items, but in the present application a call stands for a speech item and call packets stand for speech item packets.

In a first preferred embodiment the node N of the base station BS is connected to the exchange DXT by a 64 kbit/s circuit. The exchange DXT is connected to the rest of the network at a higher rate, for example, at 2Mbit/s. A high-rate channel is preferably used in the first preferred embodiment as parallel 64 kbit/s connections, whereby 32 packets, multiplexed with respect to each other, can be sent in parallel. The situation is then the same as when the packets arrive from different directions: the packets are in random order. Packets associated with the same connection are, however, in chronological order, for they are sent in succession and travel the same distance at the same rate. A high-rate channel can also be used as such without dividing it into parallel channels. The transmission rates are presented only by way of an example to illustrate the invention.

The TETRA standard defines diverse speech and data services with different priorities. For each user logging in the network, a priority is stored in the database. The priorities and the quality of service affect, for example, the allocation of the resources. One speech service defined in the standard is a group call, which is established, for example, from the site by dialling a group number and announcing that the call is a point-to-multipoint call. When a group call is being established, only the group identity is transmitted to the base stations BS within the area of the group. The identities of the groups that the mobile subscribers belong to are stored in the mobile stations MS. The mobile stations MS check the group identity received and connect to the call if the group identity concerned is found in the memory of the mobile station MS: In the present example it is assumed that the exchange or the node in which the group call may be copied is able to make a sufficient number of copies of the packets of the group call and to forward them as separate packets so that a separate packet can be sent to each base station BS within the area of the group for forwarding to the mobile subscribers of the group. A group call packet can thus be processes in the same way as a conventional call packet in the first preferred embodiment of the invention.

Fig. 2 illustrates the relationship between the packets of the radio path and those to be transferred in the transmission network. When information is transmitted on the radio path, frames are transmitted in succession on a carrier wave. One frame contains carrier-wave-specific channels, or connections, and thereby a packet associated with a call destined to a receiving mobile station on the channel. A base station, in turn, receives successive frames in which a channel is allocated for each connection. In the currently used methods based on dividing a carrier wave into channels, the frames are either TDMA or CDMA frames. The duration of the frame transmission is dependent on the radio system used. For example, in a TETRA system the transmission of one TDMA frame takes about 60 ms, and in narrowband data transmission implemented by the CDMA method the transmission of a frame takes about 20 ms. Fig. 2 shows packets C1, C2, C3 and C4 to be transmitted on the carrier wave in the first preferred embodiment according to the TETRA standard. The radio interface Air of the TETRA standard is based on a four-channel time division multiplexed carrier wave with a bandwidth of 25 kHz. For the sake of clarity, the figure shows only one carrier wave and only in one direction. A base station can have several carrier waves, and there may be carrier waves both in the uplink and in the downlink directions. Usually one channel is allocated for each connection, but the radio system of the TETRA standard allows the allocation of even several channels for one connection. However, for the sake of clarity it is here assumed that one channel is allocated for one connection. A group call is also one connection. With reference to Fig. 2 it is assumed that speech that has not been channel-coded is transmitted on channel 1, whereby the actual information content of the packet, i.e. the payload PL1, is 432 bits. The speech transmitted on channel 3 is channel-coded speech, whereby the payload PL2 of the packet is 274 bits. In the node N at the base station the payload received from the radio path is separated from the other information transferred on the radio path, and identification data ID is added to the payload. The identification data ID1, ID2 indicate the connection that the packet is associated with, i.e. the destination of the packet. On the basis of the data the node identifies the data transmission connection associated with the packet. The identifier indicates the destination address of the packet, i.e. where it is to be transmitted to. The identification data is presented, for example, by 56 bits. Correspondingly, the identification data is deleted from the payload of the packets received from the transmission network at the node N, and the other necessary information is added thereto on the radio path. To or from the actual transmission network are thus transmitted call packets PA whose size is either 330 bits or 488 bits.

The transmission times of the packets on a 64 kbit/s transmission path are about 5.2 ms and 7.6 ms, respectively. About 15 ms are reserved for each channel on the radio path. The network node of the transmission network, i.e. the exchange DXT in the first preferred embodiment of the invention, receives packets PA and forwards them preferably on the basis of the identification data ID.

Fig. 3 illustrates the operation of a node in a packet-switched data transmission network in the first preferred embodiment of the invention. The node is, for example, the exchange DXT of a TETRA transmission network. The node can also be any other network node that contains at least so much intelligence that it will not transmit the packets it has received back in the same direction. The node need not process the packet in any way; all it has to do is to forward the packets. The packets contain a payload and an identifier in accordance with Fig. 2. For the sake of clarity, it is assumed in Fig. 3 that the transmission rates of incoming channels 1, 2, 3, 4, 5 and 6 and those of outgoing channels A and B are identical. The advantage is that the transmission of the packet can be started immediately when the reception of the packet begins. If the transmission rates were different, the transmission could not be started until the whole packet would have been received. In Fig. 3 the different data transmission connections are differentiated from one another by numbering them in the order of incoming channels. The packets associated with the same connection have the same number. The function of the numbering is purely illustrative. The numbered data transmission connections have already been assigned time slices 1 in a repetition period R, which is of a predefined length RT. In accordance with Fig. 3, the repetition period has to be much longer than the duration of the transmission of a single packet, so that there is enough time to send packets of various connections within the period. If the control packets containing signalling information internal to the network are separated from the call packets, the length of the control packets need not be taken into account when the length of the repetition period is defined. The control packets are usually short, but for example when the network is updated it may be worthwhile to use large control packets, whereby the decisive factor in defining their maximum size is that the transmission of the control packet must not exceed the duration of the repetition period. When the invention is applied to radio systems, it is natural to define the length of the repetition period to be equal to the duration of the transmission of a frame on the radio path, for one packet per connection is needed per frame, since the packets associated with the same connection are always sent on the same channel. When the length of the repetition period is the same as the duration of the transmission of the frame on the radio path, the operation of the whole system is synchronized. If the repetition period were shorter than the duration of the frame transmission, the packets would still not reach the receiving party any faster but would instead be buffered at the base station, where they would wait for access onto the radio path. If the repetition period were longer, some packets would not arrive in time for transmission onto the radio path but would miss their transmission slots and thereby impair the quality of the connection. In TETRA networks the length of the repetition period is thus preferably 60 ms. It is about 11.5 times the period needed to transmit a shorter 330 bit call packet and about 8 times the period needed to transmit a longer 488 bit packet at the 64 kbit/s transmission rate used in the example.

In Fig. 3 a packet 2 associated with a connection for which no time slice 1 has been allocated arrives on an incoming channel 6. Packet 2 is not transmitted after packet 1-1 in accordance with the order of arrival, since it contains a time slice allocated for connection 4-1, but the transmission of packet 2 is delayed by a time t until a sufficient amount of free time is found on the outgoing channels. In the example of Fig. 3 the first sufficiently long free period is found on channel B after a time instant B-t7. A time slice 1' is allocated for the connection of the channel 6 from channel B. Even if all the other connections deallocated their time slices, the packets arriving at the connection of channel 6 would be started to be forwarded to channel B at time instant B-t7. The transmission delay can thus be maintained even. The delay of the first packet can preferably be used, for example, for optimal routing of the packet and for defining its length.

In Fig. 3 the network node has received a second packet to data transmission connection 1-1 during the same repetition period. The packet is marked with hatched lines in the figure. The packet, however, is not transmitted immediately when the transmission is possible, i.e. on channel B at time instant B-t8, but it is preferably transmitted only at time instant A-t1 in the next repetition period, i.e. in a time slice allocated for the connection and preferably on the same channel. In this way, the packets can be forwarded at even intervals and the variation in the transmission delays can be eliminated and the problems caused by it minimized. The use of the same channel ensures that the packets travel an equal distance along an equally congested route. The effect of channel variation on the variation in the transmission delays can thus be eliminated. If a higher-rate channel is divided into a plural number of channels, the channels form a unit that can be regarded as a single channel with several layers of time slices. A packet can also be transmitted on any channel whatsoever or to any number of channels whatsoever, as long as the interval between the starting times of the transmission of successive packets destined to one and the same connection is of the same length as the repetition period R.

Although Fig. 3 illustrates the allocation of time slices according to the invention by using the lengths of the call packets of the TETRA system shown in Fig. 2, it is apparent from Fig. 3 that the invention does not limit the length of the packets in any way nor require that they should be of a certain length. The size of a packet can preferably be checked from an identifier located in the packet, if the identifier contains information on the size of the packet as well as information on the connection. The information on the size of the packet can also be derived from a piece of information contained in the identifier. For example, in a TETRA system a sufficient piece of information is that the connection is, for example, a speech connection and that channel-coded speech will be transmitted on it. The size of the packet can also be checked by delaying the first packet so that the packet is received in full before the transmission is started. The size of the packet can also be checked by sending a control packet to the network node during the call set-up, the control packet indicating that packets of a defined length will be transmitted on the connection.

When there is no free capacity, resources can be allocated in many different ways, which are at least partly dependent on the resource allocation algorithms of the data transmission system utilizing the invention. The invention allows, for example, optimization of the transmission capacity by transferring the time slices of a connection or connections to provide the time slice desired. The time slices are preferably transferred so that a time slice moves either to the same time instant in the repetition period on another channel or to an earlier time instant in the repetition period on the same channel. These alternatives, however, are not always possible. As regards the quality of the connection, the essential point is that such transfers are as rare as possible. A once transferred allocation of a time slice is thus not transferred back even if it is possible, but the packets are transmitted in a new time slice. When resources are allocated, it is possible to try to prioritize different connections, for example, so that time slices allocated for a speech connection are transferred only when the problem cannot be solved by transferring time slices allocated for a data connection. On the quality of a data connection the time has a less essential effect.

In a system according to the TETRA standard resources are allocated in the order of priority and time. Each user and group has a priority defined in the database, and the priority is utilized in allocating resources that have become free or in setting down calls. For example, an emergency call can set down all other connections except other emergency calls. If there are no free resources when a call is being set up, the call will be put in a queue. When resources become available, they are allocated for the calls in the order of priority. The resource allocation principles according to the standard can be well combined with the optimization of transmission capacity made possible by the invention so as to yield the resource allocation algorithm that provides the best transmission capacity.

Fig. 4 is a block diagram illustrating a network node of the invention in the first preferred embodiment. For the sake of illustration, the network node is the same node whose operation has been described above in connection with Fig. 3. In accordance with Fig. 3, six incoming channels thus lead to the network node and two outgoing channels lead away from it. The only restriction of the number of outgoing and incoming channels of the network node according to the invention is that there is at least one incoming and one outgoing channel in the node. The network node comprises a clock CLO for the synchronization and the timing of the operation. To the clock is connected a time counter TC that resets to zero at an interval of one repetition period, the time counter indicating the current point in the repetition period so as to time the transmission of the packets. The time counter enables correct timing of the transmission of the packets from the network node. The network node further comprises exchange terminals ET for transmitting and receiving packets from other network nodes or other networks. The network node further comprises a maintenance unit OMU, which connects the network node with the network management system. The maintenance unit OMU sets the duration of the repetition period in the network node in accordance with the commands received from the network management system. The network node also comprises a call control unit CCU, whose functions are call control and resource management. The call control unit CCU identifies a connection associated with the packet received; checks whether the packet received is the first packet received or whether a time slice in which the packet can be sent has already been allocated for the connection; allocates time slices for the connections and controls the transmission of the packets through the exchange terminals ET so that a packet is transmitted in a time slice allocated for the connection such that the packets of one and the same connection are transmitted at an interval of one repetition period. In order for the call control unit CCU to successfully perform the functions, the unit can, for example, maintain a Table T on the connections that it is forwarding. The information can also be stored in some other way. The table T contains information on the connections for which time slices have been allocated, i.e. on calls 3 in the first preferred embodiment, and information on the moments 4 when the transmission of the calls is started. The starting moment is the time instant in the repetition period indicated by the time counter at which the transmission of the packet is started, i.e. the time instant from which onward a time slice in the repetition period is reserved. Further, the length of the reserved time slice must be found out. In the present example the problem is solved by maintaining, in the table T, a column 5 in which the length of the reserved time slice is recorded in bits, from which the duration of the time slice can be easily calculated when the transmission rate is known. The length can also be indicated in seconds. The length of the reserved time slice can also be indicated by indicating the moment when the time slice is terminated. This information is the minimum information needed by the call control unit CCU. The information allows the CCU to allocate time slices, identify the connections for which time slices have been allocated, and time the transmissions correctly.

In addition to the minimum information, the table T can preferably also contain other information, such as information on the packets 6 sent to the connection and, if there are a plural number of outgoing channels in the network node, information on the outgoing channel 7 to which packets are being sent. The table T can contain both types of information, neither type of information, or only one of these types of information. The information on the outgoing channel ensures that the packets are actually transmitted on the same outgoing channel, whereby it is possible to ensure that the transmission delay does not vary because of the channel.

Fig. 4 shows an alternative way of maintaining information on the packets transmitted to the connection. The column 'sent' 6 in the table T indicates, as will be described below in connection with Fig. 6, whether or not packets have been sent in the previous time slice. There are also other alternatives: for example, the column could contain information on all the packets transmitted and on the time instants when the packets have been transmitted. The alternative illustrated in Fig. 4 makes it possible to avoid unnecessary allocation of capacity. Unlike the other alternative, this alternative cannot be used to collect billing information.

In the first preferred embodiment, the call control unit CCU comprises a buffer BUF to buffer packets. The buffer BUF can also be arranged in other units in the network node or it can be a separate module. In some other embodiments a separate buffer is not necessarily needed at all.

Fig. 5 is a flow diagram illustrating the first preferred embodiment of the invention, in which call packets and control packets containing signalling or control information on the data transmission network are separated in the network node. A separate queue is maintained in the buffer of the network node for control packets. In the network node of Fig. 4 the call control unit comprises a buffer. The buffer can also be in some other unit in the network node or it can be a separate module. The call packets may contain either speech or data. When a new packet is received 20 in the network node, the network node detects from the identification data contained in the packet whether the packet is 21 a control packet. If it is 21 a control packet, the received packet is supplied 22 to a control packet queue. If it is a call packet, it is checked whether a time slice has been allocated 23 for the connection associated with it. If a time slice has been allocated for the connection, the starting moment of the packet transmission, i.e. the output instant, is looked for 25. Depending on the starting moment, the packet either waits for the output instant in the buffer or the transmission of the packet is started immediately. If a time slice has not been allocated, the call is a new call, and a time slice is allocated 24 for the call in the manner described above in connection with Fig. 3. When a time slice has been allocated, an output instant is defined 25 for the packet and the packet is either put in the buffer to wait for the output instant or the transmission is started immediately.

If the control packets and the call packets are not separated, only some of the method steps illustrated in Fig. 5 are taken. When a new packet has been received, it is checked whether a time slice has been allocated 23 for the connection associated with the packet. If yes, an output instant is looked for 25 for the packet and the packet is either put in the buffer to wait for the output instant or the transmission is started immediately. If no time slice has been allocated, a time slice is allocated 24 and the starting moment of the transmission is defined 25 in the same way as above.

Fig. 6 illustrates the operation of the transmission side of the network node according to the invention in the first preferred embodiment, in which the network node is the network node shown in Fig. 4. The situation in Fig. 6 is one in which a transmission moment of the next call packet is being waited for 26. When the moment arrives, it is checked whether there is 27 a packet waiting in the buffer. If there is a packet associated with the call in the buffer, then the packet is sent and a P is marked 28 in the 'sent' column of the table maintained in the network node in accordance with the solution described in connection with Fig. 4 so as to indicate that a packet has been sent. After the transmission of the packet the time left before the next call packet is to be sent, i.e. transmission moment of the next call packet, is calculated 29. The queue of control packets is then preferably monitored 30 to find a first control message that is sufficiently short so that there is enough time to send it. If one is found 31, then the control message is sent 32.

After the transmission of the control message, it is checked whether 33 is the transmission moment of the next call packet. If yes, the routine returns to step 27 to check whether there is a packet waiting in the buffer. If the transmission moment of the next call packet is not yet at hand, the routine returns to step 29, where the interval to the transmission moment of the next call packet is calculated. If a sufficiently short control message has not been found 31, the transmission moment of the next call packet is waited for 26. Alternatively, instead of looking for a sufficiently short control message, control messages can also be selected for transmission in the order of arrival, whereby it is not checked whether a sufficiently short control message can be found, as above, but whether there is enough time to send the next control message. If there is not enough time, the message is returned to the queue and the transmission moment of the next call packet is waited for. If the call packets are provided with priorities, they are put in a queue in the order of priorities, and the control packets with the highest priorities can even be interpreted as call packets and a separate time slice can be allocated for them.

If there is 27 no next call packet, it is checked 34 whether there is an E in the 'sent' column of the table according to the solution presented in connection with Fig. 4. If there is an E in the column, it means that the previous time there was no packet, either, and the allocation of the time slice for the call is cancelled as unnecessary from the table of Fig. 4, i.e. the time slice is deallocated 35. After that, the routine proceeds to step 29, where the interval to the transmission moment of the next call packet is calculated. If the check carried out in step 34 shows a P in the 'sent' column of the table presented in Fig. 4, an E is marked 36 in the 'sent' column and the routine proceeds to step 29, where the interval to the transmission moment of the next call packet is calculated. The allocation of the time slice is also cancelled, i.e. the time slice is deallocated, in response to a call set-down message indicating that the call has been terminated. The above-described deallocation of a time slice after two successive 'no packet' indications thus improves the efficiency of the transmission capacity, since the capacity that is not used is here made available. Further, a separate call set-down message need not be sent, and the amount of signalling can thus be reduced. The fact that there are no packets can also be detected in many other ways, for example, by comparing the interval to the previous packet transmitted, as mentioned above in connection with Fig. 4. The number of packets that must be lacking before a time slice will be deallocated can also be other than two, but the number must be determined in advance and the network node must be arranged to function accordingly. If packets associated with the call are received for forwarding after the time slice allocated for the call has been deallocated, the call will be regarded as a new call and a new time slice will be allocated for it.

If control packets and call packets are not separated, only some of the method steps illustrated in Fig. 6 are taken. A call packet here means all packets to be transmitted. The starting situation is one in which the transmission moment of the next call packet is waited for 26. When the moment arrives, it is checked whether there is 27 a packet waiting in the buffer. If there is a packet associated with the connection in the buffer, the packet is transmitted and a P is marked 28 in the 'sent' column of the table maintained in the network node in accordance with the solution presented in connection with Fig. 4 to indicate that a packet has been sent, and the routine then proceeds to step 26 to wait for the transmission moment of the next packet. If there is no next call packet, it is checked 34 whether there is an E in the 'sent' column of the table of the solution presented in connection with Fig. 4. If there is an E in the table, it means that there was no packet the preceding time, either, and the allocation of the time slice for the connection will then be cancelled as unnecessary from the table of Fig. 4, i.e. the time slice will be deallocated 35. After that, the routine proceeds to step 26 to wait for the transmission moment of the next call packet. If the check in step 34 shows a P in the 'sent' column of the table presented in Fig. 4, an E is marked in the 'sent' column, and the routine then proceeds to step 26 to wait for the transmission moment of the next call packet.

With regard to the different handling of the packets in Figs. 5 and 6 and the description of the figures based on the information content of the packets indicated by the identification information, it is pointed out that a similar division could also be conducted by separating the speech-containing packets from the other packets, whereby the speech packets would be handled in the same way as the call packets in connection with Figs. 5 and 6, i.e. time slices would be allocated only for speech packets. Both the control packets and the data-containing call packets would then be handled in the same way as the call packet in connection with the description of Figs. 5 and 6, i.e. they would be transmitted between the speech packets whenever there is enough time. The packets can also be divided into control, speech and data packets, whereby time slices would be allocated for the speech packets and the data packets would be given priority over the control packets. It is also possible to divide the packets generally into control packets and call packets, but in the case of congestion to divide the call packets into speech packets and data packets, so that different types of packets can be handled differently. If real-time image is transferred in the network, the image packets can be compared to speech packets and handled in the same way.

The figures and the description of the figures are only intended to illustrate the present invention and its application to a mobile system. It will be obvious to those skilled in the art that the invention can be varied and modified in many ways without deviating from the scope and spirit of the invention as disclosed in the attached claims.

CLAIMS

1. A method for packet-switched data transmission, in which a packet received is transmitted toward the destination address of the packet, **characterized** by comprising the steps of

defining a repetition period that is considerably longer than the time needed for sending a single packet and allocating time slices of a different length from the repetition period for the connections;

defining the starting point of the repetition period;

maintaining information on the current point in the repetition period;

maintaining information on the time slices allocated from the repetition period and on the connections associated with them;

receiving a packet (20);

identifying the connection associated with the packet as well as the destination address from the identification data of the packet;

checking whether a time slice has been allocated for the connection of the packet from the repetition period; and

if yes, starting the transmission of the packet when the starting point of the time slice allocated is reached in the repetition period;

if no,

- defining the duration of the packet transmission time;

- searching the repetition period for a free period that is at least as long as the transmission time;

- allocating from the free period of the repetition period a time slice (24) which is at least as long as the duration of the packet transmission time for the connection of the packet; and

- starting the transmission of the packet when the starting point of the time slice allocated is reached in the repetition period.

2. A method according to claim 1, **characterized** by further comprising the steps of

deallocating a time slice (35) allocated for the connection if there are two successive repetition periods during which there are no packets associated with the connection to be sent in the time slice, and

if a packet associated with the connection is received after the deallocation of the connection, allocating a new time slice.

3. A method according to claim 1 or 2, **characterized** by further comprising the steps of

defining a repetition period and its starting point separately for each available outgoing channel,

selecting the channel on which a time slice suitable for allocation is found the quickest as an outgoing channel for the new connection,

sending the packets associated with the same connection on the same outgoing channel, and

maintaining information channel-specifically on the time slices allocated from the repetition period and on the connections associated with them.

4. A method according to claim 1, 2 or 3, **characterized** by

dividing the packets on the basis of the identification data contained in them into call packets and control packets that contain signalling information,

allocating time slices only for connections associated with call packets,

sending the call packets in the time slices allocated, and

sending the control packets in such points of the repetition period in which they can be sent in full before the next allocated time slice.

5. A method according to claim 1, 2 or 3, **characterized** by

dividing packets on the basis of the identification data contained in them into speech-containing speech packets and other packets,

allocating time slices only for connections associated with speech packets,

sending the speech packets in the time slices allocated, and

sending the other packets in such points of the repetition period in which they can be sent in full before the next allocated time slice.

6. A method according to any one of the preceding claims used in a transmission network of a mobile telephone system, **characterized** by defining the length of the repetition period as the length of the time needed to transmit a frame on the radio path.

7. A network node of a packet-switched data transmission network to which there leads at least one incoming channel and from which there leads at least one outgoing channel, the node comprising transmission means for transmitting packets toward a destination address, **characterized** by the node further comprising

setting means (OMU) for setting the duration of a repetition period to be notably longer than the transmission time of a single packet;

at least one time counter (TC), which resets to zero at an interval of one repetition period, to indicate the current point in the repetition period so as to time the transmission of the packets;

identification means (CCU) for identifying the connection associated with the packet received and for detecting whether packets associated with the connection have already been received in the network node before said packet or whether the packet is the first packet associated with the connection;

allocation means (CCU) for searching the repetition period for a free time slice that is at least as long as the transmission time of the packet and for allocating a time slice that is as long as the packet transmission time from the free time slice for the connection associated with the packet in response to the reception of the first packet associated with the connection, the starting point of the allocated time slice indicated by the time counter defining the starting moment of the transmission of the packets associated with the connection, and

control means (CCU) for transmitting packets toward destination addresses in the time slices allocated for the connections associated with the packets so that the packets associated with the same connection are transmitted from the network node one packet at a time at an interval of one repetition period.

8. A network node according to claim 7, **characterized** by further comprising recording means (7) for recording, for each connection, the packets transmitted and the time slices with no packets and, in response to two successive time slices of one and the same connection with no packets, for deallocating the time slice.

9. A network node according to claim 7 or 8 and further comprising at least a second outgoing channel, **characterized** in that

each channel has a separate repetition period,

the allocation means (CCU) are arranged to allocate time slices from each channel, and

the control means (CCU) are arranged to transmit the packets associated with the same connection on the same outgoing channel.

10. A network node according to claim 7, 8 or 9, **character-ized** in that

the network node is located in the transmission network of a mobile telephone system,

the repetition period is of the same length as the time needed to transmit a frame on the radio path, and

the lengths of the packets (PA) to be transmitted are defined on the basis of the lengths (PL1, PL2) of the packets transmitted on the radio path in the mobile telephone system.

11. A network node according to claim 10, **characterized** in that

the identification means (CCU) are arranged to separate call packets and control packets that contain signalling information,

the allocation means (CCU) are arranged to allocate time slices only in response of the reception of a first call packet, and

the control means (CCU) are arranged to transmit the control packets in such intervals between the allocated time slices in which they can be sent in full.

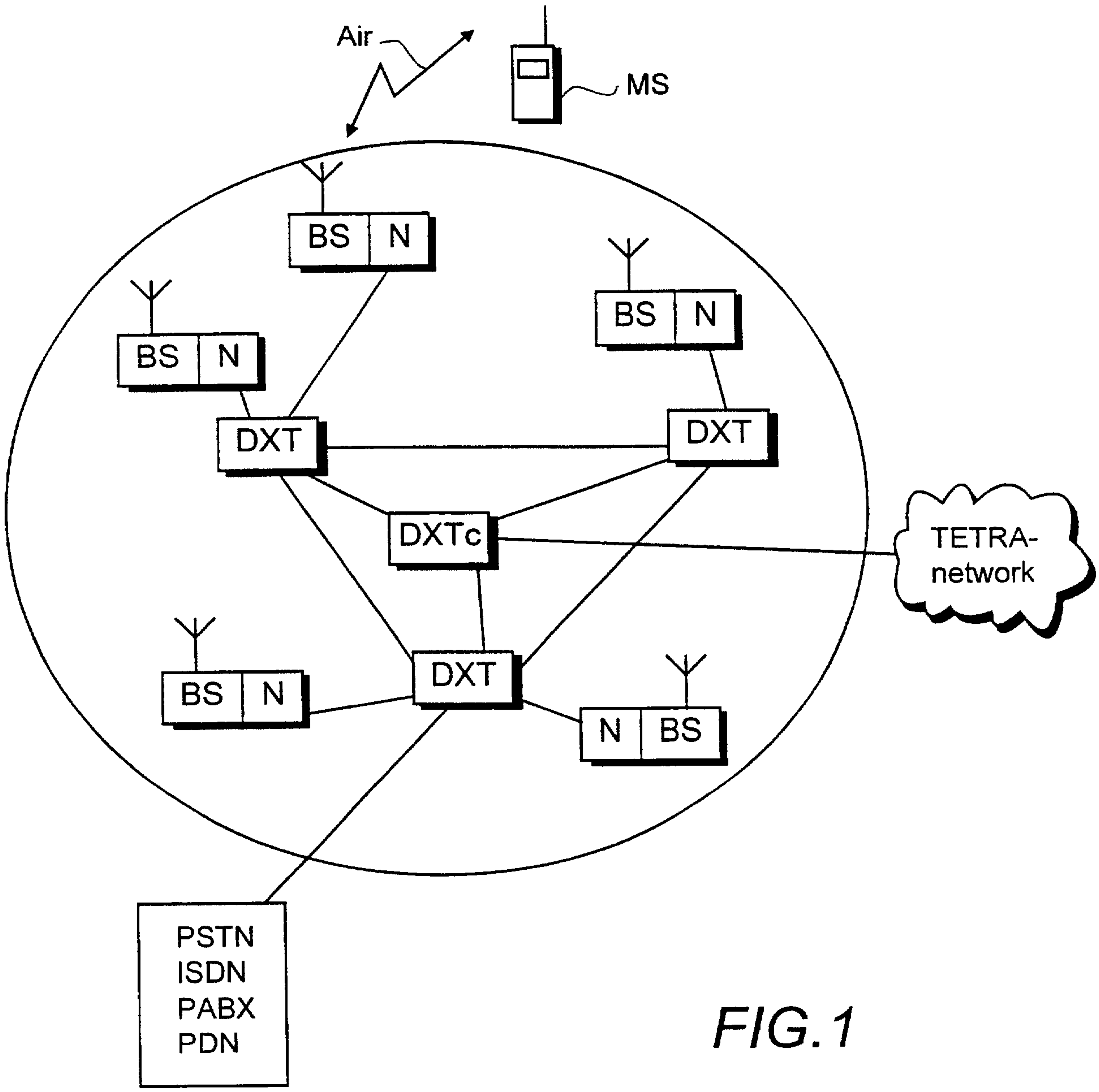

Air
MS
BS
N
BS
N
BS
N
DXT
DXT
DXTc
TETRA-
network
BS
N
DXT
N
BS
PSTN
ISDN
PABX
PDN

FIG.1

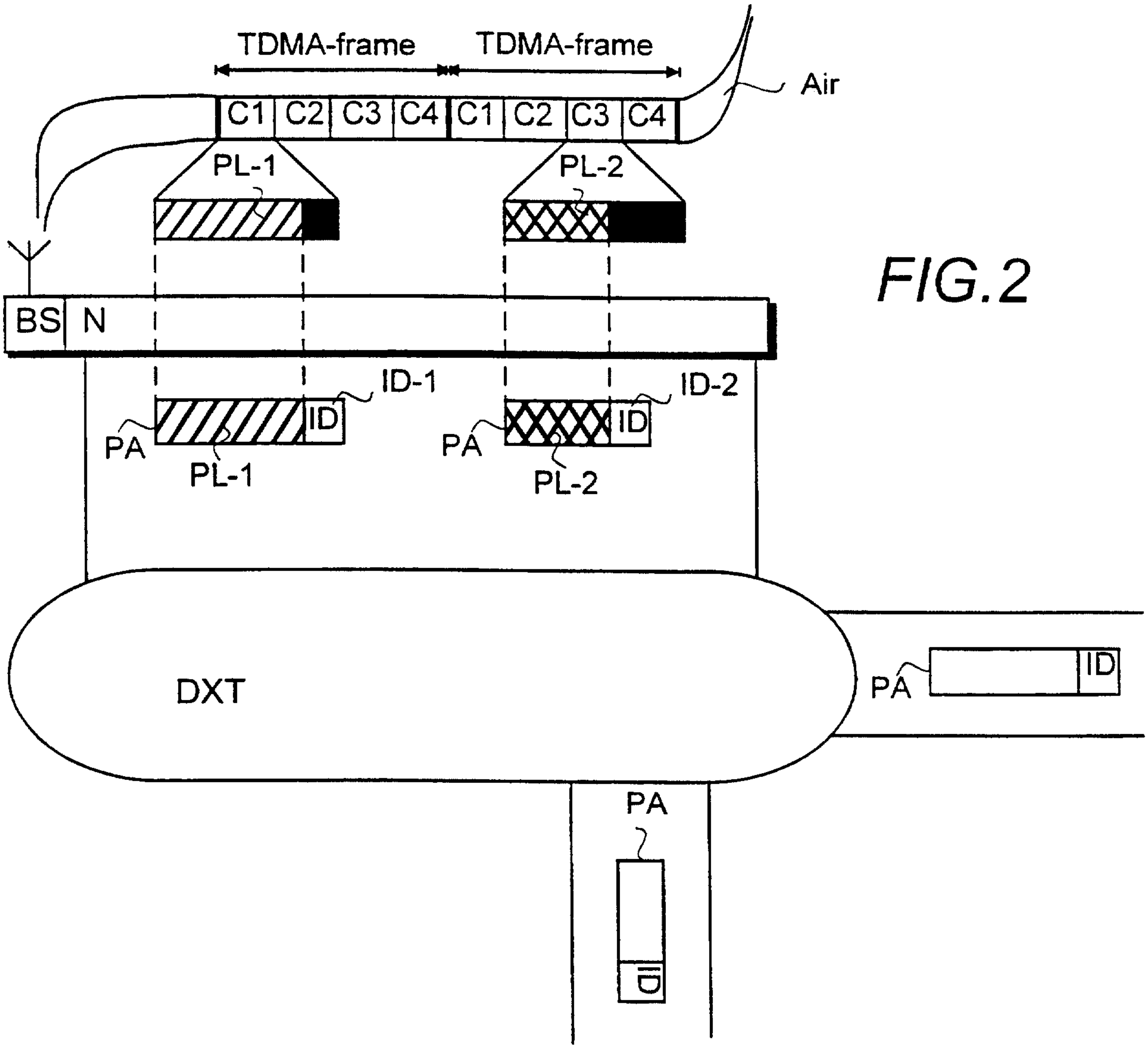

TDMA-frame
TDMA-frame
Air
C1
C2
C3
C4
C1
C2
C3
C4
PL-1
PL-2
BS
N
ID-1
ID-2
ID
ID
PA
PA
PL-1
PL-2
DXT
PA
ID
PA
ID

FIG.2

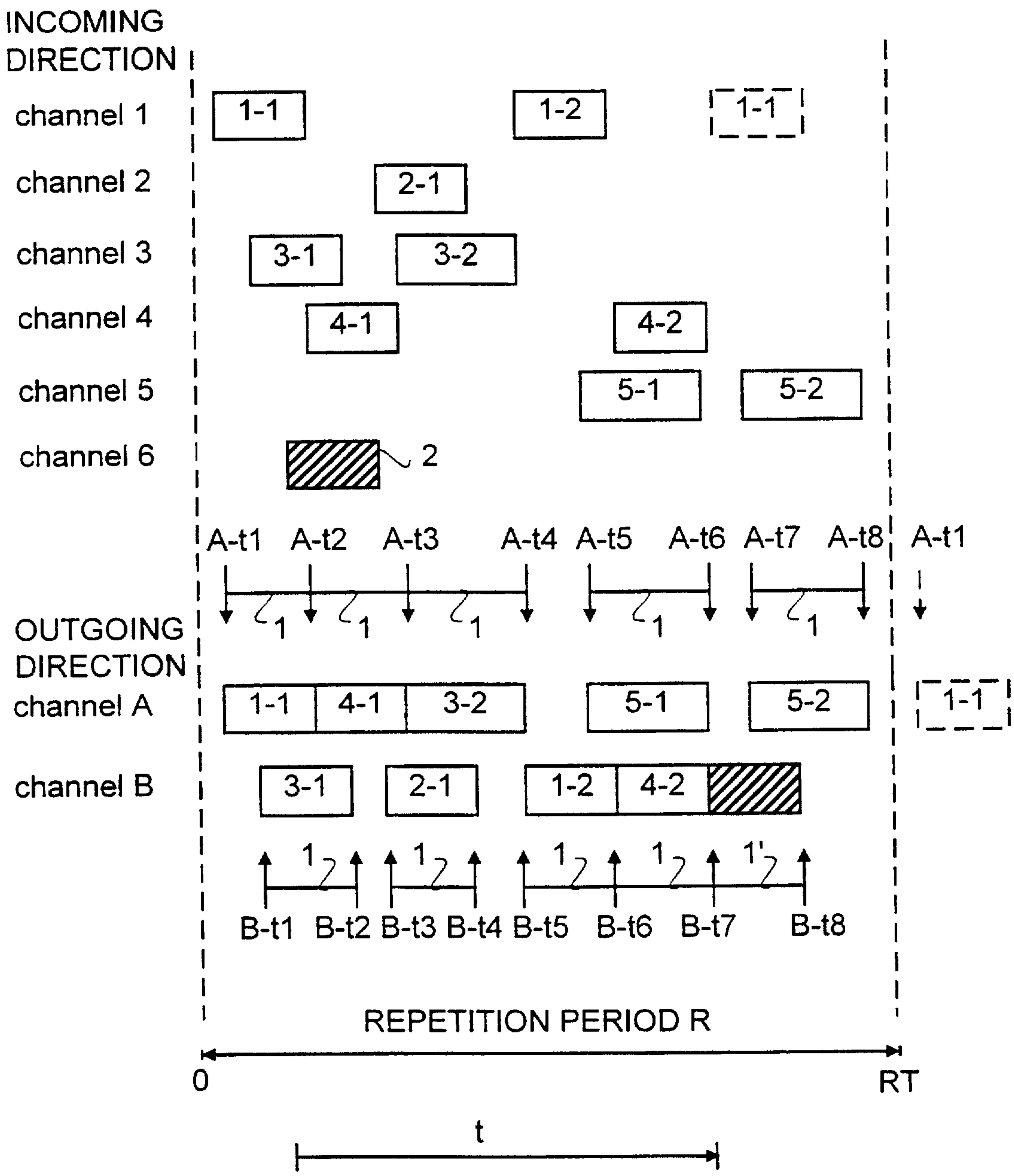
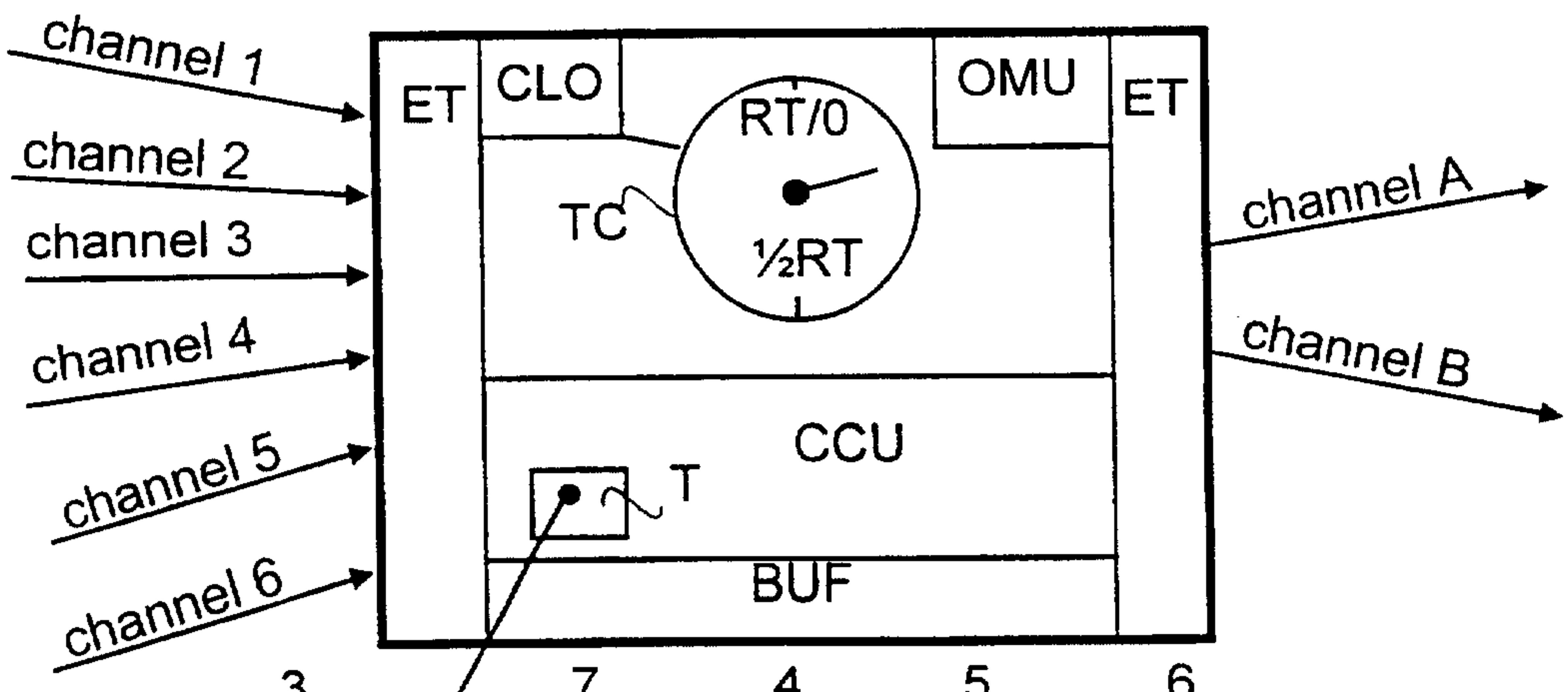

INCOMING DIRECTION
channel 1
1-1
1-2
1-1
channel 2
2-1
channel 3
3-1
3-2
channel 4
4-1
4-2
channel 5
5-1
5-2
channel 6
2
A-t1 A-t2 A-t3 A-t4 A-t5 A-t6 A-t7 A-t8 A-t1
1
OUTGOING DIRECTION
channel A
1-1 4-1 3-2
5-1
5-2
1-1
channel B
3-1
2-1
1-2 4-2
1
1'
B-t1 B-t2 B-t3 B-t4 B-t5 B-t6 B-t7 B-t8
REPETITION PERIOD R
0
RT
t

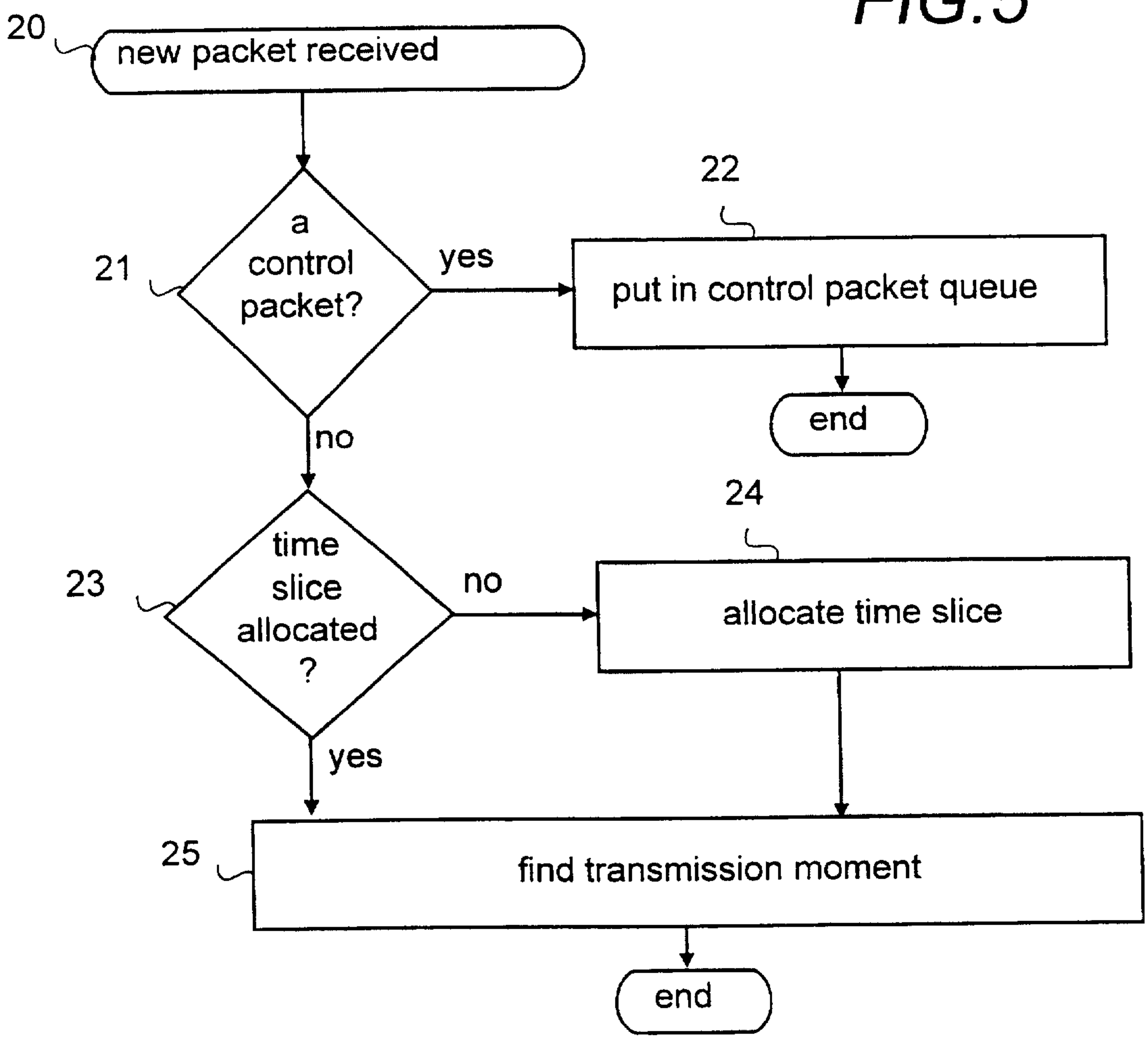
FIG.3



3	7	4	5	6
call	outgoing channel	starting moment	length	sent
1-1	A	A-t1	330	P
3-1	B	B-t1	330	P
4-1	A	A-t2	330	P
2-1	B	B-t3	330	P
3-2	A	A-t3	488	P
1-2	B	B-t5	330	P
5-1	A	A-t5	488	P
4-2	B	B-t6	330	P
5-2	A	A-t7	488	P

FIG.4

FIG.5


20
new packet received
21
a control packet?
yes
22
put in control packet queue
end
no
23
time slice allocated ?
no
24
allocate time slice
yes
25
find transmission moment
end

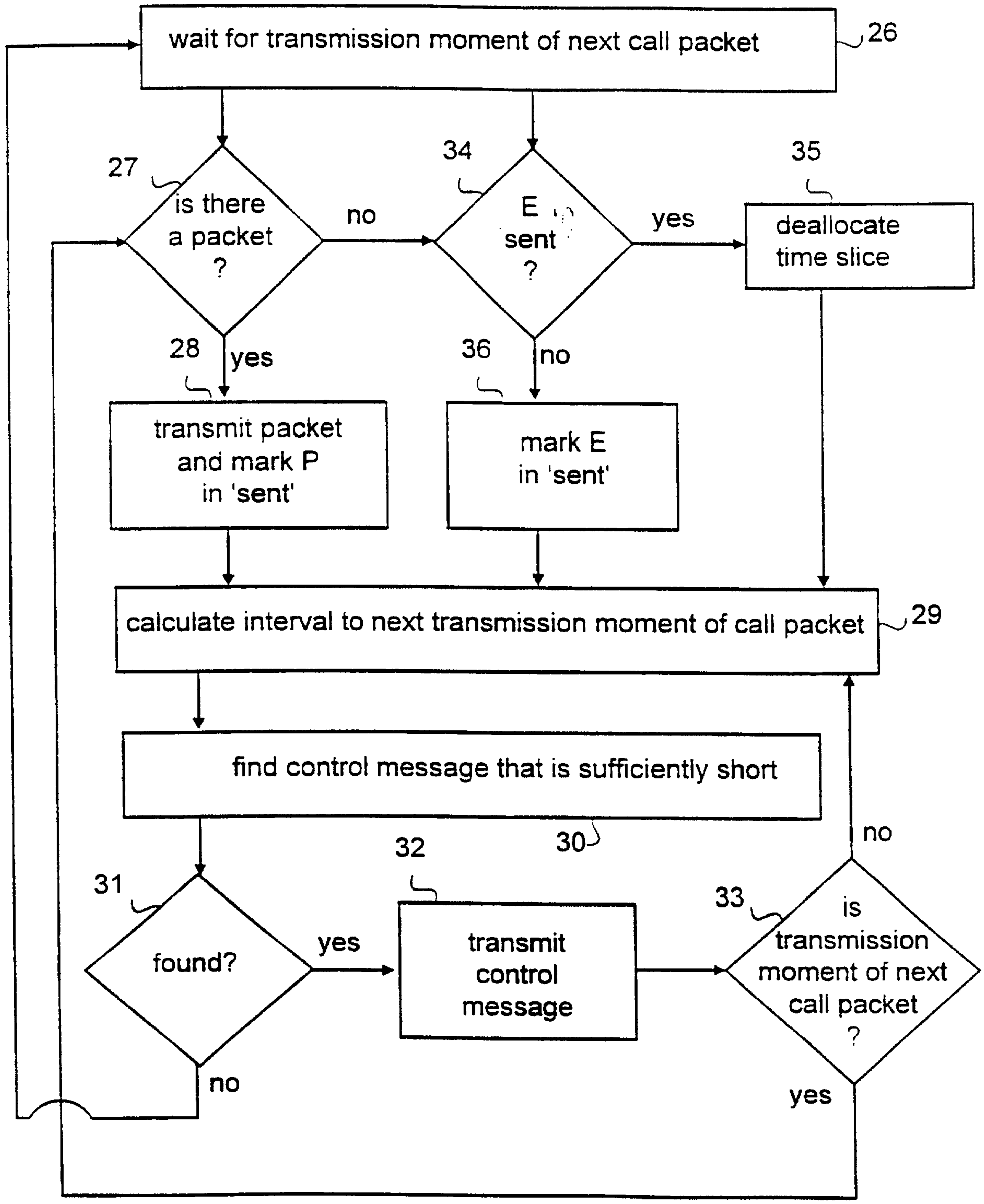

wait for transmission moment of next call packet
26
27
is there a packet ?
no
34
E sent ?
yes
35
deallocate time slice
28
yes
36
no
transmit packet and mark P in 'sent'
mark E in 'sent'
calculate interval to next transmission moment of call packet
29
find control message that is sufficiently short
30
32
31
found?
yes
transmit control message
33
is transmission moment of next call packet ?
no
no
yes

FIG.6