Available online www.jsaer.com

***Journal of Scientific and Engineering Research*, 2023, 10(9):124-132**

Research Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Research on Hybrid Automatic Repeat reQuest (HARQ) technology in V2X

Miaohui Yang

School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo Shandong 255000, China

Abstract With the rapid economic development of our country, the number of private cars has increased sharply, which has brought great pressure to the domestic transportation system. At the same time, with the development of communication technology, the amount of data transmitted in 5G communication system has also increased significantly, and users have higher and higher requirements for communication quality. All these have prompted people to put forward new requirements for HARQ re-transmission mechanism, mainly reflected in the pursuit of data transmission rate and system reliability. The traditional HARQ mechanism will bring a lot of redundant re-transmission when the amount of data is large, which will lead to the system performance decline. In order to improve the success rate of information transmission when vehicles pass each other, this paper studies the vehicle-mounted information transmission mechanism. In this case, this paper presents a new hybrid information re-transmission technology, is firstly established by using Matlab simulation software simulation environment to get the data, packet real-time prediction calculation model is established, then based on cloud computing, real-time monitoring of the network environment for predicting to may transmit packets to a transmission or re-transmission consecutive time delay. Taking multiple linear regression and dynamic analytic hierarchy process as the core, the real-time prediction model of data transmission block error rate is constructed, and the framework of decision making and implementation rules of two alternative schemes is set up.

Keywords 5G, hybrid automatic repeat request, vehicle to everything, cloud computing, continuous re-transmission

1. Introduction

With the rise of various high and new technologies, technologies such as big data, 5G, Internet of Things, and artificial intelligence have been very mature. Vehicle network is the embodiment of the Internet of Things in the intelligent transportation system, and it is also one of the key technologies in the intelligent transportation system, which can complete the wireless connection between pedestrians, vehicles and road infrastructure. The ultra-low latency of the 5G network enables the sensors in the intelligent transportation system to provide various information in almost real time, making the driver's grasp of the road surface information more accurately. Under the combined action of the Internet of Vehicles and artificial intelligence, autonomous driving systems are becoming more and more mature, and a large amount of data is generated by sensors inside and outside the vehicle during the driving process.

In the wireless network environment, data transmission errors are inevitable due to external factors, and the reliability of information transmission cannot be guaranteed. If this data is analyzed and fitted in advance, the predictive models needed to support autonomous driving can be obtained so that the system can more accurately predict vehicle behavior and improve and avoid errors.

The communication technology of direct vehicle-to-vehicle connection enables the exchange of information between vehicles, but the reliability of packet transmission cannot be guaranteed in the wireless network environment, especially in the case of autonomous driving sharing information exchange, vehicle cooperative maneuvering, cooperative lane change, and amplification of sensing range.

In order to improve the reliability of information transmission, it is necessary to develop an information retransmission technology. Vehicle-to-vehicle communication (5GNR-V2X) technology realizes the packet loss retransmission mechanism in the media access control layer and physical layer, with the retransmission mechanism of Hybrid Automatic Repeat reQuest (HARQ) combined with semi-persistent scheduling (SPS) [1]. Semi-persistent scheduling schedules the radio resources of the first transmitted packet in a fixed manner, and if the channel quality changes due to the movement of the receiver, changes in the surrounding environment, or interference from other users, the quality of service cannot be ensured.

This topic focuses on information exchange between vehicles in a wireless network environment to ensure reliable packet transmission. In general, the retransmission mechanism will receive a resend request after the packet transmission fails, and then retransmit the corresponding packet. This paper takes a different starting point, predicts the probability of packet transmission error by monitoring the network environment of each packet transmission in real time, and then delays the transmission or retransmission of 2~4 consecutive times for the packet with a large probability of transmission error. It is expected that the use of vehicle-to-vehicle information sharing will allow vehicles with a long distance to have a higher priority when retransmitting data, so as to increase the probability of successful packet transmission, make full use of limited channel resources, and reduce the number of information retransmissions and increase the probability of successful data packet transmission. In turn, it improves the user's perception of transmitting information, reduces the number of traffic accidents, and improves the efficiency of traffic traffic.

2. Related work

The hybrid automatic retransmission request protocol can improve the reliability of the connected car system and reduce the error rate of the system, but each retransmission will consume a certain amount of energy. When researchers study the reliability of the retransmission system, the retransmission energy consumption of the system is often not taken into account, and the channel gain monitoring (CGM) technology is studied in view of the retransmission energy consumption of the retransmission protocol, and the relationship function between the codeword reception probability of the CGM receiver and the average transmission times of the system is obtained, and CGM technology is introduced into the Internet of Vehicles [2]. The receiving end only receives codewords with channel gain greater than the preselected channel gain threshold, which can better meet the performance requirements of receiving packet error interrupts.

Li Y et al. in 2015 studied the throughput of HARQ systems based on incremental redundancy (IR) and fixed transmission rate under the presence of statistical queuing constraints, that is, the presence of buffer overflow probability limits, and compared them with Type-I type HARQ and Chase-based HARQ merger [3]. Xu K et al. proposed a Type-III HARQ protocol for joint network coding and distributed Turbo coding in 2015, which adopts a XOR network encoding combined retransmission strategy when accessing nodes, and a federated network-Turbo decoding strategy at the user endpoint, and the average transmission times are significantly reduced [4].

How to correct errors caused by multipath channels and noise during wireless transmission has always been an important research content. Current error control methods are divided into three categories: forward error control (FEC), automatic retransmission (ARQ), and HARQ mode. Hybrid automatic retransmission technology has attracted widespread attention from researchers since the 70s of the twentieth century. After the 20th century, with the development and maturity of multiple input multiple output (MIMO) and channel coding technology, the use of different types of channel coding methods in the process of HARQ retransmission has received extensive attention and research [5].

With the withdrawal of major operators from 5G packages, 5G officially began the commercial period in China, in order to meet the requirements of bandwidth, delay, traffic rate and higher quality under various conditions of 5G, HARQ technology has made corresponding changes. Yang Xiaohang proposed an adaptive HARQ retransmission scheme for 5G, which remaps the starting position of the retransmission according to the equivalent channel quality feedback of the receiver feedback [6] to ensure that the packets transmitted through the poor channel last time can experience better new quality when retransmitted. The simulation results show that the newly proposed retransmission scheme increases the gain of 0.45~0.5dB over the original standard retransmission scheme on the TDL-A fading channel.

In 2021, Gao Yuehong analyzed the non-negligible waste of resources caused by traditional HARQ technology in the scenario of 5G data volume transmission, and introduced the principle of CBG HARQ technology, which makes full use of limited spectrum resources by reducing redundant retransmission [7].

3. Real-time prediction model design of BLER

Analysis of influencing factors of BLER

Channel Quality Indicator (CQI) as one of the important influencing factors of the decoding error rate (BLER), the information measurement from the terminal is reported to the base station, CQI generally refers to the downstream channel quality [8]. The base station MAC-layer scheduler selects the appropriate scheduling algorithm and downstream block size according to the CQI information feedback from the user equipment (UE) to ensure the best downstream communication performance of UE in different wireless channel environments. In this study, the average CQI (Ave CQI) of each resource block downlink was used.

The worse the channel quality, the more redundant information the channel encoding, the stronger the error correction ability, the lower the order modulation, and the lower the data transmission rate; The better the channel quality, the less redundant information the channel encodes, the weaker the error correction, the higher order modulation is used, and the lower the rate at which data is sent.

Since the downlink scheduling is determined by the base station, and the base station, as the transmitter, does not know the quality of the channel, the quality of the channel is also measured by the UE. The LTE protocol quantifies the quality of the channel into a sequence of 0~15 and defines CQI [9].

In addition to CQI, BLER's influencing factors also take into account the number of user devices (NumUEs) that transmit information at the same time, the distance between user devices, that is, the smart car and the base station (UE-Distance), and the number of resource blocks (NumRBs).

The premise of this study is that each user device is the same distance from the base station, and the BLER in the case of 4, 5, and 6 user devices at the same time is studied. The number of resource blocks is 25, 50, 75, 100, and 125, respectively, and the distance between the user equipment and the base station is set in the range of 100~2500 (m).

Matlab simulation and data acquisition

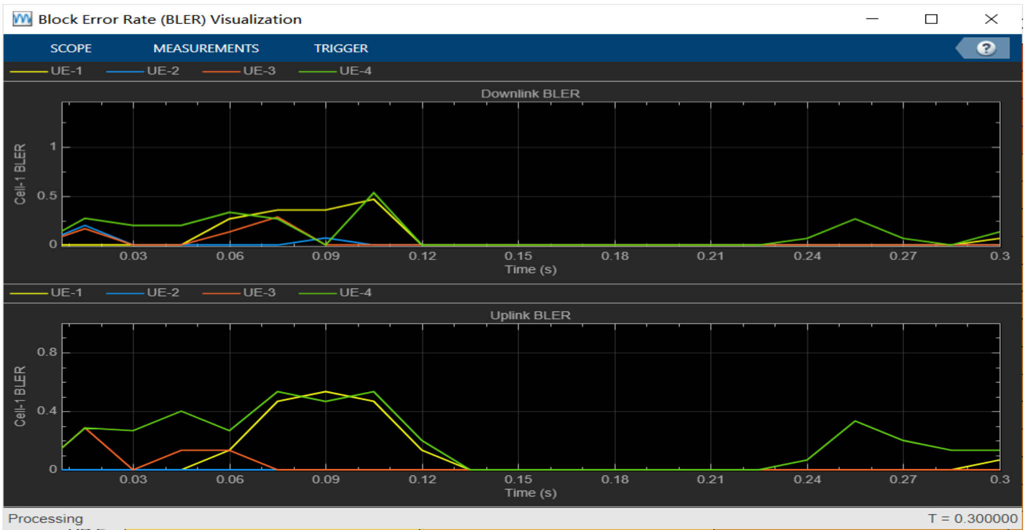

Figure 1: Illustration of BLER running results

First, the BLER module is called up as the simulation environment in the 5GToolBox of Matlab software, and then the corresponding parameters are modified and debugged to obtain the data required for this study. In order to avoid excessive variables and too many data, the following parameters in this study are fixed and set as follows:

• Subcarrier interval SCS fixed at 15 kilohertz (KHz)
• Assume that the entire bandwidth is allocated to the physical uplink/downstream shared channel (PUSCH/PDSCH)
• Assume that the uplink/downlink (UL/DL) carrier has symmetrical channel bandwidth
• 2 decibel milliwatts (dBm) transmit power for all terminals
• 5G base station transmission power of 29 decibel milliwatts (dBm)
• Receiver antenna gain at gNB SINR to CQI mapping table 0.1 BLER
• UL message generation time of 30 milliseconds (ms)
•The size of the UL packet generated by the terminal is 8000 bytes
•The period of DL packets generated by 5G base station terminals is 30 milliseconds (ms)
•The size of DL packets generated by 5G base station terminals is 8500 bytes

Then modify the parameters of NumUEs, UE-Distance, and NumRBs respectively, and run the program every time the parameters are modified, the BLER results obtained are collected and averaged, and the CQI data is also collected and averaged, and the data is recorded in an Excel table. Figure 1 is, the abscissa is the time, a total of 300ms, the ordinate is BLER, and the four different lines represent the BLER of each moment of the four UEs themselves. Using the 2021a version of Matlab, a total of 118 simulations were successfully run and 118 sets of valid data were obtained, due to the large number of data, a part (22 groups) will be shown, see Table 1.

Table 1: Partial UE information data summary table

AveBLER	UE-Distance	AveCQI	NumRBs	NumUEs
0.0025	500	4.0000	25	4
0.0200	100	10.7200	50	4
0.0050	500	3.4400	50	4
0.0100	500	3.1067	75	4
0.0500	700	2.6800	100	4
0.0800	700	2.6160	125	4
0.1750	1200	1.5520	125	4
0.0140	1200	1.7959	50	5
0.1660	1600	1.5733	75	5
0.0400	700	2.8000	100	5
0.0720	100	10.2080	125	5
0.0740	700	2.6160	125	5
0.3240	2000	1.1280	125	5
0.3980	2500	10.4000	125	5
0.0067	500	4.0000	25	6
0.0067	1600	1.6800	25	6
0.0067	2000	1.6800	25	6
0.0150	1200	1.6400	50	6
0.1617	2500	1.0000	50	6
0.1817	1600	1.5926	75	6
0.0283	500	3.4400	100	6
0.1400	1200	1.6160	125	6

These data can be performed on multiple linear regression to obtain a BLER real-time prediction model, so that the packet transmission scheme can be adjusted in real time to improve the success rate of data transmission.

4. Real-time prediction model construction of BLER

Based on the above obtained data, if you want to establish the regression equation between variables (UE-Distance, AveCQI, NumRBs, NumUEs) and AveBLER, the relationship between the dependent variables is linear, so consider using multiple linear regression. Multiple linear regression analysis requires the following four conditions [10]:

• At least 4 independent variables (4 in this study) and independent of each other;
• The dependent variable should be continuous (BLER is continuous);
• The data have the characteristics of homogeneity of variance, no outliers and normal distribution (verified to meet the conditions);
• There is no multicollinearity between the independent variables (verified to meet the criteria).

The verification process is more cumbersome, I will not repeat it here, after verification, the data meet the above four conditions, and multiple regression operations can be performed. First, import the sorted Excel data into SPSS software, click "Analysis", "Linear", "Regression" in turn, and then select the independent variables and response variables respectively for regression analysis.

Table 2 is the model summary table in the analysis results, where R=0.908 is the multiple correlation coefficient, which can be used to determine the linear relationship between the independent variable and the dependent variable, and is also an index for evaluating the fit of the regression model, and can be used as a reference index for model goodness [11]. The R square (0.825) and adjusted R square (0.819) in the table can be regarded as the degree to which the independent variable explains the variation of the dependent variable in regression analysis, and is generally measured by the adjusted R square. Therefore, this regression analysis shows that the four indicators of UE-Distance, AveCQI, NumRBs and NumUEs can explain 81.9% of the change of AveBLER, and the UE information and various indicators of channel communication quality measured by simulation can better explain the variation of block error rate, and UE information and channel traffic quality indicators have a high degree of influence on block error rate.

Table 2: Regression model summary table

Model	R	R^2	Adjustment afterR^2	Standard Error of Estimation	Durbin-Watson
1	.908[a]	.825	.819	.05098772	.857

a. Predictor variable: (constant), NumUEs, AveCQI, NumRBs, UE-Distance

b. dependent variable：AveBLER

Table 3 is the regression model significance test (ANOVA table), F=133.179 in the table is the result of the F test, the significance P<0.001 calculated according to the F value, and P<0.005 indicates that there is a linear correlation between the dependent variable and the independent variable, and indicates that at least one independent variable in the multiple linear regression model has a non-zero coefficient. Statistically significant, it is also shown that the inclusion of these independent variables helps predict the dependent variable, and this regression model is better than the null model [12].

Table 3: Regression model ANOVA table

Model	Sum of Square	Degrees of Freedom	Mean Square	F	Significance
Return	1.385	4	.346	133.179	.000[b]
Residual	.294	113	.003		
Total	1.679	117			

a. dependent variable：AveBLER

b. Predictor variable: (constant), NumUEs, AveCQI, NumRBs, UE-Distance

Table 4 is the table of regression model coefficients, where B is the model coefficient, Beta is the standardized coefficient, significance is the significance test of the variable in the model, the VIF value is used in the previous collinearity test results, the VIF value is less than 10, and the data basically meets the premise of multivariate linear analysis, that is, there is no multicollinearity problem [13]. The predictive model in this study can be assumed as:

$$AveBLER = B_0(\text{常量}) + B_1 \cdot UE\text{-}Disance + B_2 \cdot AveCQI + B_3 \cdot NumRBs + B_4 \cdot NumUEs$$

Table 4: Table of regression coefficients

Model	**Un-standardized Coefficient**		**Standardized coefficient**	**t**	**Significance**	**Correlation**			**Collinearity Statistics**	
	B	Std Error	Beta			0^{th} Order	Partial	Part	Tolerance	VIF
(constant)	-.224	.035		-6.477	.000					
UE-Distance	.00011	.000	.76940	14.756	.000	.636	.811	.581	.570	1.76
AveCQI	.00922	.002	.23434	4.490	.000	-.306	.389	.177	.569	1.76
NumRBs	.00215	.000	.63283	16.046	.000	.644	.834	.631	.996	1.00
NumU	.00014	.006	.00096	.024	.981	-.001	.002	.001	.998	1.00

a. dependent variable ：AveBLER

Then look at the significance test result P value of the respective variables in the model in the regression coefficient table, if P > 0.05, the independent variable is not statistically significant in this model, and such variables in the regression model should be deleted; If P<0.05, this independent variable is statistically significant in this model and should be retained. From Table 3, it can be seen that only the significance test results of AveCQI, NumRBs and UE-Distance are less than 0.05, so the final real-time BLER prediction model can be expressed as:

$$AveBLER = -0.224 + 0.00011 \cdot UE\text{-}Distance + 0.0092 \cdot AveCQI + 0.00215 \cdot NumRBs$$

Alternative transmission scheme decision model

According to the algorithms and regulations of Chapter 3 and Chapter 4, the real-time packet prediction model obtained in Chapter 3, the two optimization scheme decisions in Chapter 4, the delay time and the number of retransmissions are programmed, and the specific delay or continuous retransmission scheme can be obtained by entering the values of the parameters UE-Distance, AveCQI and NumRBs.

Chapter 4 mentions that when the block error rate ≥h, the decision of the optimization scheme is made, and the determination of the h value requires a large number of simulation data verification, due to the current simulation environment and time and other objective reasons, it is temporarily impossible to give a most suitable value. According to the simulation data currently available in Chapter III, a total of 118 sets of data are selected, of which 29 sets of data have an average block error rate of ≥0.2, accounting for about 25% of the total data. Here, h=0.2 is provisionally specified, that is, when the predicted packet block error rate is greater than or equal to 20%, the decision of delay or continuous retransmission of the optimization scheme is made.

Table 5: Partial optimization scenario decision data

UE-Distance	AveCQI	NumRBs	rand	Optimized scheme	Delay time/number of retransmissions
2500	10.4000	125	8.15	Retransmission	4
2500	1.1280	125	1.14	Delayed	200
2300	1.0640	125	7.20	Retransmission	4
2500	1.0800	100	5.69	Delayed	300
2000	1.2960	125	1.28	Delayed	200
2500	1.1481	75	0.88	Delayed	200
2000	1.1280	125	3.80	Delayed	200
2500	1.3600	75	2.51	Delayed	200
2300	1.1600	100	8.61	Retransmission	4
2300	1.0000	100	8.74	Retransmission	4
2300	1.1200	100	1.42	Delayed	200
2300	1.0000	75	8.68	Retransmission	4
1600	1.3600	125	7.64	Retransmission	3
2300	1.1067	75	8.62	Retransmission	4
2000	1.1600	100	7.13	Retransmission	4
1600	1.4880	125	8.41	Retransmission	3
2000	1.4000	100	0.32	Delayed	200
2300	1.2133	75	4.92	Delayed	200

The four factors of the criterion layer in the decision-making of the two optimization schemes are: AveCQI, transmission distance, NumRBs resource block number and information importance, and the information importance value is randomly generated by the system using the rand function in Matlab. According to the algorithms in Chapters 3 and 4, Table 5 shows the simulation data of some of the solution decisions obtained by using Matlab programming and entering the required parameters.

It can be seen that when H is set to 0.2, the transmission distance is basically greater than 1500m, resulting in the continuous retransmission of the system decision is basically 3 or 4 times. In this case, the value of h can be adjusted, the value of h can be reduced, and the transmission distance will be below 1500m, and the result of the system decision may be retransmitted twice in a row; It is also possible to adjust the distance interval used to determine the number of consecutive retransmissions, and increase the distance interval of 2 consecutive retransmissions, and the result of the system's decision will also be 2 consecutive retransmissions.

Similarly, it can be found that the delay time is basically 200ms or 300ms, and the delay of 100ms can be achieved by adjusting the provisions of the CQI interval used to determine the delay duration. Due to time reasons, the algorithm mechanism currently used in this study is single, and the results are relatively simple.

5. Results & Discussion

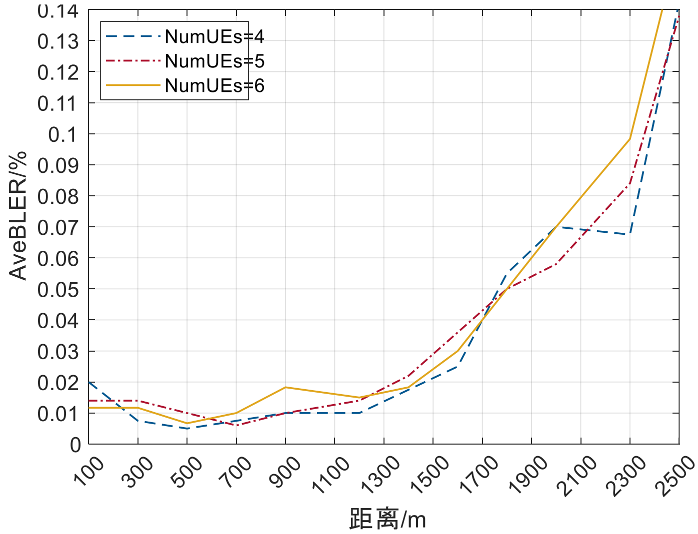

Figure 2: BLER varies with distance for different user situations

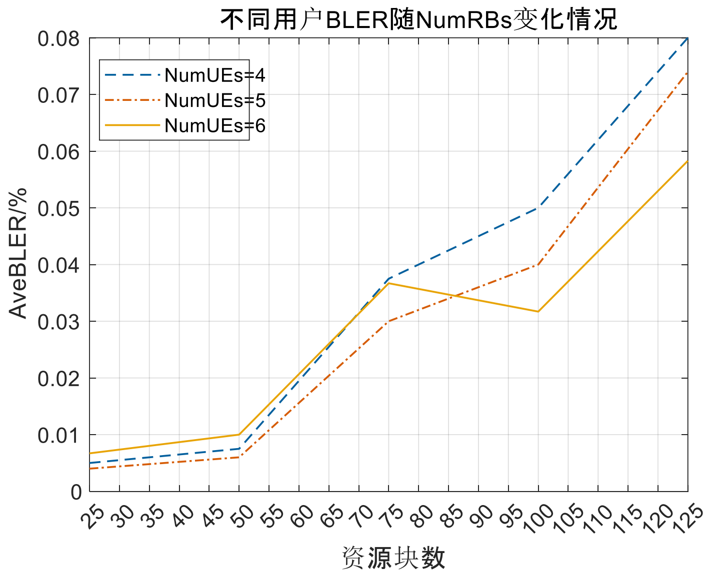

Figure 3: BLER changes with resource blocks for different user situations

Based on the simulation data in Chapter 3, the changes of BLER based on distance UE-Distanc are explained. Due to the large number of parameter data, the premise of this study is that when the resource block NumRBs=50s, based on the change of vehicle distance, the BLER change of block error rate is compared for 4 users, 5 users and 6 users respectively (Figure 2). It can be seen that the overall situation is that with the increase of information transmission distance between vehicles, BLER as a whole has an upward trend, and when the distance reaches more than 1500m, BLER begins to increase sharply with the increase of distance. It can also be seen that the greater the number of users, the more significant the block error rate increases with the increase of distance.

According to the simulation data in Chapter 3, BLER based on the change of resource block can also be explained, due to the large number of parameter data, the premise of this study is that when the information transmission distance between vehicles UE-Distance = 700m, based on the change of vehicle distance, the BLER block error rate BLER change is made for 4 users, 5 users and 6 users respectively Figure 3. It can be seen that the overall situation is that with the increase of the number of information transmission resource blocks NumRBs, BLER as a whole has an upward trend. As can also be seen from the three polylines on the figure, the size of BLER is not directly related to the number of users.

6. Conclusion

This paper aims to improve the success rate of information transmission when vehicles pass each other, study the on-board information transmission mechanism, rely on vehicle network technologies such as vehicle-vehicle interconnection, realize real-time information sharing between vehicles, various on-board sensors can monitor the transmission environment to obtain real-time parameter data, take multiple linear regression and dynamic analytic hierarchy as the core, construct a real-time prediction model of data transmission block error rate, and build a framework for decision-making and implementation rules of two optimization schemes. The provisions therein can be adjusted according to the actual situation to reduce the bit error rate, improve the reliability of information transmission, and make full use of limited information resources. However, there are potential research points that deserve more in-depth study. How to develop a more accurate delay time and continuous retransmission number algorithm. The current algorithm of delay time and continuous retransmission number adopted in this study has a single consideration, and the selection rule is simple, and it is not possible to adopt the delay time or continuous retransmission number that achieves the best efficiency of the entire system for each data packet. Therefore, it is worth studying the algorithm for determining the delay time and the number of consecutive retransmissions.

References

[1]. QUALCOMM Incorporated. (2020). Patent Issued for Techniques And Apparatuses For Ultra Reliable Low Latency Hybrid Automatic Repeat Request Retransmission For Semi-Persistent Scheduling. Network Weekly News, USPTO(10), 797-833.

[2]. Zuoben Yu, Xinhao Lv, Gang Cen, Zhongpeng Wang, Mingwei Wu. (2018). Research on performance of HARQ-CGM system based on effective signal-to-noise ratio in Internet of ehicles. Journal of Zhejiang University of Science and Technology, 30(06):468-473+487.

[3]. Li Y,Gursoy MC,Velipasalar S. (2015). On the throughput of Hybrid-ARQ under statistical queuing constraints. IEEE Transactions on Vehicular Technology, 64(6):2725-2732.

[4]. Vehicle Technology. (2015). Findings from PLA University of Science and Technology Broaden Understanding of Vehicle Technology, NTC-HARQ: Network-Turbo-Coding Based HARQ Protocol for Wireless Broadcasting System, Telecommunications Weekly.

[5]. 박상준. (2020). Selective Symbol-Level Combining Scheme for LDPC Coded MIMO-HARQ Systems. The Journal of Korean Institute of Communications and Information Sciences, 45(4).

[6]. Xiaohang Yang. (2020). Research and implementation of HARQ transmission mechanism under 5G communication system. Beijing University of Posts and Telecommunications.

[7]. Yuehong Gao, Haotian Yang, NingYin, Lu Chen. (2021). Analysis and prospect of CBG HARQ technology in 5G system. Communication Technology, 54(02), 363-368.)

[8]. ÖNEM YILDIZ,RADOSVETA İVANOVA SOKULLU. (2019). Mobility and traffic-aware resource scheduling for downlink transmissions in LTE-A systems. Turkish Journal of Electrical Engineering and Computer Science, (3).

[9]. Wei Zhang. (2014). Cross-layer AMC design and implementation of LTE system. Wuhan University of Technology.

[10]. Jianfei Leng, Xu Gao, Jiaping Zhu. (2016). Application of multiple linear regression statistical prediction model. Statistics and Decision, (07), 82-85.

[11]. Perdana Adityo Cakra,Rachman Rais,Palinggi Monika Datu Mirring. (2021). An analysis of population using Multiple Linear Regression Analysis. IOP Conference Series: Materials Science and Engineering, 1088(1).

[12]. Runyu Zhang. Research on multiple regression and machine learning methods of tropical cyclones. (2020). National Marine Environmental Forecasting Center.

[13]. Jun Li. (2014). Multiple linear regression method based on factor analysis and its application in stock price prediction. Nanjing University.