

令和 5 年 6 月 27 日現在

機関番号：13601

研究種目：基盤研究(C)（一般）

研究期間：2018～2022

課題番号：18K04131

研究課題名（和文）算術符号の精度に関する遅延と符号化レートの解析

研究課題名（英文）Mathematical analysis of delay and coding rate with respect to precision of arithmetic code

研究代表者

西新 幹彦（Nishiara, Mikihiro）

信州大学・学術研究院工学系・准教授

研究者番号：90333492

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：遅延を歪みと解釈したりコストと解釈したりすることによって従来の情報理論の枠組みで主要となる問題を数理工学的に解析することができた。歪みに関する結果として、符号化法における固定長符号化と可変長符号化の違いと歪み制約における平均歪み制約と最大歪み制約の違いによって定式化の違いが生まれるが、そのすべてにおいて雑音を伴う一般情報源に対するレート歪み問題を雑音のない問題へ帰着することができた。コストに関する結果としては、従来のコストの概念を拡張して通信路の出力に掛かるコストに関する最適解を導いたほか、この問題の解法を応用することによって適用範囲の広い確率的コストについても最適解を導くことができた。

研究成果の学術的意義や社会的意義

研究開始当初は遅延の解析を目標に掲げていたが、遅延の解釈を一般化することによって当初想定していたよりも適用範囲の広い結果が得られた。歪みの概念を決定論的なものから確率論的なものへ一般化したことで揺らぎのある歪みを伴う問題に対する理論的体系を示すことができた。一方、コストの概念を決定論的なものから確率論的なものへ一般化したことで間接的な要因によるコスト変動に対する理論的体系を示すことができた。

研究成果の概要（英文）：Interpreting a delay as a distortion or a cost, the mathematical aspect of the main problems in the framework of conventional information theory were analyzed scientifically. As a result about the distortion, for various formulations due to the difference in fixed length encoding and variable-length record encoding in the encoding method and the difference between mean distortion condition and maximum distortion limitation in the distortion limitation, the rate-distortion problems for general sources with noise reduced into the corresponding problems without noise. As a result about the cost, expanding the concept of the conventional cost, the optimal solution about the cost that the output of the channel cost was derived as well as the optimal solution about the stochastic cost.

研究分野：情報理論

キーワード：情報理論 通信の数学的理論 情報源符号化 通信路符号化

### 1. 研究開始当初の背景

高度な情報処理が求められる現在、情報通信に要求される 2 つの大きな指標として、効率と遅延がある。効率とは、1 秒あたりに送信できる情報量のことであり、できるだけ大きいことが望ましい。もう一つの遅延とは、情報が送信されてから受信先で正しく解釈されるまでの時間のことであり、この時間はできるだけ短いことが望ましい。遅延を小さくすることが要求される通信の例として緊急警報信号や電子商取引が挙げられる。通信の遅延を小さくすることは国民の生命財産の保全や社会の発展につながる重要な課題である。一般に通信の効率と遅延の間にはトレードオフが存在する。つまり、通信の効率を良くするための処理を複雑なものにするということは送信前の処理と受信後の処理を複雑にするということを含み、情報が発生してから受信者に届くまでに時間を要することになる。

情報が発生してから受信者に届くまでの時間は広い意味で遅延と呼ばれるが、遅延の主要な原因は 3 つある。1 つ目は伝播遅延と呼ばれるもので、送信機と受信機の間を信号が物理的に伝播するのにかかる時間である。これは送信機と受信機間の物理的な距離に比例する。例えば無線通信における電波の伝播時間がそれである。言うまでもなく電波は光の速度で伝播する。伝播遅延を小さくするには送信機と受信機間の物理的な距離(伝播経路の長さ)を小さくするしか方法が存在しない。伝播遅延は本研究の対象外である。2 つ目は処理遅延と呼ばれ、送信機や受信機の情報処理の時間に起因する遅延である。大抵の送信機や受信機は情報処理のための一種の電子計算機を内蔵している。この内部の計算機の処理能力やアルゴリズムが貧弱であると処理遅延が大きくなる。計算機の能力を上げたり高速なアルゴリズムを採用したりすることで処理遅延を小さくすることができるが、本研究で対象とするのは処理遅延ではない。3 つ目の遅延は符号化遅延と呼ばれる。本研究では符号化遅延を研究の対象とする。符号化遅延とは、符号器と復号器を状態遷移機械(オートマトン)とみなした時の入力系列と出力系列の長さの差のことである。符号器と復号器の内部状態として情報が蓄積されることにより、一般に出力系列は入力系列よりも短くなる。つまり受信者が把握している情報は最新の情報ではないということになる。符号器と復号器の計算能力を大きくすることで状態遷移先を決定する計算速度を早くしても、内部状態が本質的に情報を蓄積していることには変わりがないので、符号化遅延の改善にはならない。符号化遅延を小さくするためには符号器と復号器のオートマトンとしての性質、ひいては写像としての数学的性質を改善する必要がある。

符号器と復号器をオートマトンとして捉える考え方は情報理論では逐次符号と呼ばれている。このような符号として代表的なのは算術符号である。算術符号は四則演算を基本とした符号化法であり、研究開始当初の背景として、符号化確率を演算する際の演算精度が符号化遅延に関係していることが分かりつつあった。そこで算術符号の演算精度と符号化遅延の関係を解明することによって遅延を改善することを本研究の狙いとした。以降では符号化遅延のことを単に遅延と呼ぶ。

### 2. 研究の目的

通信を数学的に考える際には、問題設定は大きく 2 つに分類される。一つは与えられた情報源から出力される系列をなるべく短い符号語で送信しようとする情報源符号化の問題であり、もう一つは与えられた通信路を利用してなるべく多くの情報を送ろうとする通信路符号化の問題である。つまり情報源符号化においては符号化レートが小さいことが効率の良い通信であり、通信路符号化においては符号化レートがなるべく大きいことが効率の良い通信である。

情報源符号化の問題は数学的には情報源系列の長さを大きくすることで符号化レートの上極限を上から評価する問題として定式化される。他に制約がない場合、情報源系列の長さを大きくすると遅延も大きくなる。したがって、遅延に制約を課すことによって符号化レートを小さくすることができず、効率は悪くなるのが定性的に知られている。

一方、通信路符号化の問題を数学的に定式化する際には、符号語長を大きくしながら復号誤り確率を極限でゼロにするという条件のもとで符号化レートの下極限を下から評価する。このとき、通信路符号化の場合と同様に、符号語の長さを大きくすると遅延も大きくなる。したがって、遅延に制約を課すことによって伝えられる情報量に制約がかかり、通信の効率は悪くなるのが定性的に知られている。

本研究の目的は、上記の 2 つの意味で符号化レートと遅延の間のトレードオフを明らかにすることである。

### 3. 研究の方法

上記の通り、遅延と符号化レートには 2 つの意味があることから、遅延の研究は大きく 2 つに分けることができる。ひとつは遅延を特別な歪みとして捉える方法である。歪みとは入力系列と出力系列がどれくらい異なるのかを測る概念であるが、遅延は入力系列と出力系列の長さの違いであることから、遅延を一般化して考えると歪みの概念に一致する。すなわち通信における歪みに関する一般論は遅延に関する考察を含んでいる。この事実に基づき、本研究では歪み

に関する一般論の観点から遅延の問題に取り組んだ。

遅延のもう一つの意味はコストである。コストとは、情報を伝達するにあたって支払うべき費用や消費されるエネルギーなどのことである。情報を伝達する際に経過する時間もコストの概念に含まれる。すなわち通信におけるコストに関する一般論は遅延に関する考察を含んでいる。この事実に基づき、本研究ではコストに関する一般論の観点からも遅延の問題に取り組んだ。

#### 4. 研究成果

主要な成果を以下の通り報告する。

##### (1) 着払いコスト付き通信路符号化

コスト付き情報源符号化問題では符号アルファベットに対してコストが定義される。これは雑音のない通信路に対するコスト付き通信路符号化問題と見することもできる。この場合、通信路の入力アルファベットと出力アルファベットは区別されない。ところが通信路に雑音がある場合では両者は区別されねばならない。すると雑音のある通信路に対するコスト付き通信路符号化問題を考える場合、コストをどちらのアルファベットに対して定義するのか、選択肢が生まれる。従来のコスト付き通信路符号化問題では、通信路の入力アルファベットに対してコストが定義された問題が考えられてきた。本研究では、通信路の出力アルファベットに対してコストを定義し、コスト付き通信路符号化定理を証明した。従来の場合とは符号語を送信する前に正確なコストが分かるのに対し、本研究の設定ではコストの値は受信語を見るまで確定しないという点が拡張になっている。

さらに問題は次のように一般化される。通信路の出力にコストを定義したもとで問題を解くと、出力側でコスト制約を満たすかどうかを入力側で制御する必要がある。言い換えると出力側で測られるコストの値は入力を与えられたもとでの条件付確率に従う。このことから、コストの定義は入力系列に依存する確率変数へと拡張される。本研究では確率的コストを用いたコスト付き通信路符号化定理も証明した。

##### (2) コスト超過を許す通信路符号化

通信路符号化問題では送り手と受け手の間に通信路の存在を仮定し、送り手から送られたメッセージを受け手が正しく受け取るための伝送レートと復号誤り確率の関係について取り扱う。実際のメッセージの送信には時間や電力といった何らかのコストを要する。符号語にかかるコストに制約をおいた問題をコスト制約付き通信路符号化問題という。すでに知られている結果として、各符号語長に対してコスト制約を課す場合の一般通信路に対する通信路容量が明らかになっている。本研究では、コスト制約付き通信路符号化問題の拡張と、コスト制約の定義を緩和したもとで通信路容量を明らかにした。まず、コスト制約の概念を決定論的なものから確率論的なものに拡張し、コスト制約を満たさない確率（コスト超過確率）を評価できるようにした。そのもとでコスト超過確率が漸近的にゼロになる場合やコスト超過確率の上極限が指定した正数以下になることを達成可能な条件とした。問題を解く際には新たな証明手法を用いる必要があった。具体的には、従来定理の証明の際にはコスト制約を満たす符号語を選ぶことによって符号を構成すればよかったが、本研究の問題設定ではコスト制約を満たす符号語とそうでない符号語の2段階に分けて符号を構成するという新しい手法を用いた。このようにして一般通信路の通信路容量を明らかにした。

##### (3) 雑音を伴う一般情報源に対するレート歪み問題

本研究では情報源を観測する際に雑音が混入する場合を考える。これは物理的な対象をセンサーなどで観測する場合の数理的なモデルである。また同様に復号器の出力と宛先の間にも雑音があることを想定する。このような問題はリモートソースの問題と呼ばれる。誤り訂正が不可能な位置に雑音があることからこの問題は必然的に歪みのある符号化問題すなわちレート歪み問題となる。定常無記憶情報源に対するリモートソースの問題は解かれているが、一般情報源に対しては解かれていない。本研究では一般情報源に対するリモートソースの問題を解いた。一般情報源に対するレート歪み問題には、符号化法と歪み制約の取り方によって全部で4通りの組み合わせがある。符号化法に関しては固定長符号化と可変長符号化の2通りがあり、歪み制約の取り方として平均歪み制約と最大歪み制約の2通りがある。

平均歪み制約を用いる場合はすでに知られている定常無記憶情報源と同じ解法で解くことができた。これは雑音の影響を考慮した代替歪み測度を定義する際、期待値を評価することになるので、問題設定は通常のレート歪み問題と同じになることに起因する。このようにして4通りのうち2通りの符号化定理を証明することができた。

最大歪み制約を用いる場合は、雑音の影響を考慮した代替歪み測度を確率変数として扱う必要があるため、通常のレート歪み問題に帰着させることができず、確率的歪みの概念を導入する必要があった。すなわち、代替歪み測度は確率的歪みとなる。この事実に着目することによって最大歪み制約の固定長符号化問題は確率的歪みを用いたレート歪み問題として解決することができた。最大歪み制約の可変長符号化問題も確率的歪みを用いたレート歪み問題として解くことができることは判明したが、帰着先の確率的歪みを用いた可変長符号化問題が未解決のため、この解決が今後の課題として残った。

#### (4) 情報スペクトルを保存する情報源符号化

情報スペクトルが符号化問題において重要な意味を持っていることはよく知られている。可変長符号に対する標準的な評価指標は平均符号語長であるが、もうひとつの評価指標として符号語長のオーバーフロー確率がある。無歪み可変長符号化のオーバーフローと復号誤りに関してはいくつかの従来研究がある。また、レート歪み問題のオーバーフローと歪み超過に関してもいくつかの従来研究がある。これらの問題では情報スペクトルの裾確率が重要な役割を担っている。期待値や裾確率は広がりをもつ確率分布に対する代表値である。これに対し、本研究では代表値を用いずに確率分布の広がりそのものを評価する指標を新しく導入し、この指標の意味で最適な符号は情報源の情報スペクトルを保存することを明らかにした。この指標は従来の代表値による評価の一般化になっており、証明した内容も従来の結果を一般化したものになっている。具体的には歪みのない情報源符号化に対する符号化を基礎として歪みのある情報源符号化に対する符号化についても符号化定理を証明した。歪みのない情報源符号化に対する符号化定理はよく知られているシャノン符号を用いることで容易に証明できた。一方、歪みのある情報源符号化に対する符号化定理を証明するには、新しい証明手法を用いた。情報理論における符号化定理の証明手法は大きくランダム符号と非ランダム符号に分類されるが、本手法は非ランダム符号に相当する。さらに、情報スペクトル的方法において特徴的な「情報スペクトルの切り出し」という手法を適用することで最適な符号を構成することができた。具体的には、情報スペクトルを適切な幅で切り出した上で、小さい情報スペクトルをもつ系列から順に符号語を割り当てていく。通常の固定長符号化における非ランダム符号では符号語の割り当てはいずれ終了するが、本研究では可変長符号化を考えているため、いくらでも長い符号語が生成される可能性があり、一般に符号語の割り当ての終了が存在しないのが特徴である。この場合においても符号の存在は数学的帰納法により保証される。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 NISHIARA Mikihiro	4. 巻 E105.A
2. 論文標題 Channel Coding with Cost Paid on Delivery	5. 発行年 2022年
3. 雑誌名 IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences	6. 最初と最後の頁 345 ~ 352
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transfun.2021TAP0002	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計28件（うち招待講演 1件 / うち国際学会 4件）

1. 発表者名 西新幹彦
2. 発表標題 情報スペクトル指向の情報源可変長符号化定理について
3. 学会等名 第45回情報理論とその応用シンポジウム(SITA2022)
4. 発表年 2022年

1. 発表者名 金子歩夢, 西新幹彦
2. 発表標題 雑音を含んで観測される一般情報源に対するレート歪み理論について
3. 学会等名 第45回情報理論とその応用シンポジウム(SITA2022)
4. 発表年 2022年

1. 発表者名 Masaki Hori, Mikihiro Nishiara
2. 発表標題 Channel Capacity with Cost Constraint Allowing Some Cost Overrun
3. 学会等名 International Symposium on Information Theory and Its Applications (ISITA2022) (国際学会)
4. 発表年 2022年

1. 発表者名 佐藤奏杜, 西新幹彦
2. 発表標題 パケット間隔で情報を送る通信路に対するSum-Product復号法の検証
3. 学会等名 2022年度電子情報通信学会信越支部大会
4. 発表年 2022年

1. 発表者名 金子歩夢, 西新幹彦
2. 発表標題 一般情報源に対する観測値に雑音があるシステムの平均歪み制約レート歪み理論
3. 学会等名 2022年度電子情報通信学会信越支部大会
4. 発表年 2022年

1. 発表者名 細海俊介, 西新幹彦
2. 発表標題 情報鮮度の視点から見たバッファで割り込みのある待ち行列通信システムの実験的考察
3. 学会等名 2022年度電子情報通信学会信越支部大会
4. 発表年 2022年

1. 発表者名 堀真樹, 西新幹彦
2. 発表標題 漸近的成本制約付き通信路符号化
3. 学会等名 第44回情報理論とその応用シンポジウム
4. 発表年 2021年

1. 発表者名 佐藤奏杜, 西新幹彦
2. 発表標題 パケット間隔で情報を送る通信路の記憶を考慮したSum-Product復号法の検証
3. 学会等名 2021年度電子情報通信学会信越支部大会
4. 発表年 2021年

1. 発表者名 大橋輝路, 西新幹彦
2. 発表標題 二元定常無記憶情報源に対するレート・遅延関数
3. 学会等名 電子情報通信学会情報理論研究会
4. 発表年 2020年

1. 発表者名 堀真樹, 西新幹彦
2. 発表標題 復号誤りとコスト超過を許容した漸近的成本制約付き通信路符号化定理
3. 学会等名 電子情報通信学会情報理論研究会
4. 発表年 2020年

1. 発表者名 大橋輝路, 西新幹彦
2. 発表標題 定常無記憶情報源に対するレート・遅延関数の導出について
3. 学会等名 電子情報通信学会信越支部大会
4. 発表年 2020年

1. 発表者名 塚田芳寿, 西新幹彦
2. 発表標題 誤り訂正符号としてのラテン方陣の復号方法について
3. 学会等名 電子情報通信学会信越支部大会
4. 発表年 2020年

1. 発表者名 堀真樹, 西新幹彦
2. 発表標題 微小なコスト超過を許容した一般通信路における漸近的成本制約付通信路符号化定理
3. 学会等名 電子情報通信学会信越支部大会
4. 発表年 2020年

1. 発表者名 千葉直紀, 西新幹彦
2. 発表標題 情報鮮度の観点に基づく通信システムの数理的評価に向けて
3. 学会等名 電子情報通信学会信越支部大会
4. 発表年 2020年

1. 発表者名 Mikihiko Nishiara
2. 発表標題 On Channel Coding with Cost Constraint on Delivery
3. 学会等名 11th Asia-Europe workshop on Concepts in Information Theory (国際学会)
4. 発表年 2019年



1. 発表者名 Naohiro Uchida, Mikihiro Nishiara
2. 発表標題 On Searching for Optimal Non-alphabetic Arithmetic Codes with Low Delay Based on A* Algorithm
3. 学会等名 2019 IEEE International Symposium on Information Theory, Recent Results Session (国際学会)
4. 発表年 2019年

1. 発表者名 大橋輝路, 西新幹彦
2. 発表標題 レート・遅延理論に関する基礎的検討
3. 学会等名 電子情報通信学会情報理論研究会
4. 発表年 2019年

1. 発表者名 川合康太, 西新幹彦
2. 発表標題 離散無記憶通信路に対するコストあたり通信路容量の算出に関する考察
3. 学会等名 第42回情報理論とその応用シンポジウム
4. 発表年 2019年

1. 発表者名 堀真樹, 西新幹彦
2. 発表標題 一般通信路に対する漸近的成本制約付き通信路符号化定理
3. 学会等名 電子情報通信学会信越支部大会
4. 発表年 2019年

1. 発表者名 川合康太, 西新幹彦
2. 発表標題 離散無記憶通信路に対するコストあたり通信路容量の算出に向けて
3. 学会等名 電子情報通信学会信越支部大会
4. 発表年 2019年

1. 発表者名 打田尚大, 西新幹彦
2. 発表標題 順序保存性のない最適な算術符号の探索アルゴリズムの評価
3. 学会等名 電子情報通信学会信越支部大会
4. 発表年 2019年

1. 発表者名 大橋輝路, 西新幹彦
2. 発表標題 レート・遅延理論に関する基礎的検討
3. 学会等名 電子情報通信学会信越支部大会
4. 発表年 2019年

1. 発表者名 Mikihiko Nishiara, Masashi Yamakawa
2. 発表標題 An A* algorithm for constructing optimal AIFV codes
3. 学会等名 Asia Pacific Society for Computing and Information Technology 2018 Annual Meeting (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 打田尚大, 西新幹彦
2. 発表標題 A*アルゴリズムを用いた順序保存性のない最適な算術符号の探索について
3. 学会等名 電子情報通信学会情報理論研究会
4. 発表年 2018年

1. 発表者名 西新幹彦
2. 発表標題 微小なコスト超過を許容する着払いコスト制約付き通信路符号化について
3. 学会等名 第41回情報理論とその応用シンポジウム
4. 発表年 2018年

1. 発表者名 伊藤佑樹, 西新幹彦
2. 発表標題 ランダム符号化を用いない一般情報源に対するレート・歪み理論の順定理
3. 学会等名 第41回情報理論とその応用シンポジウム
4. 発表年 2018年

1. 発表者名 打田尚大, 西新幹彦
2. 発表標題 順序保存性のない最適な算術符号の探索アルゴリズム
3. 学会等名 平成30年度電子情報通信学会信越支部大会
4. 発表年 2018年

1. 発表者名 大橋輝路, 西新幹彦
2. 発表標題 レート歪み理論と遅延の関係について
3. 学会等名 平成30年度電子情報通信学会信越支部大会
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------