

平成 30 年 5 月 22 日現在

機関番号：32634

研究種目：基盤研究(C) (一般)

研究期間：2014～2017

課題番号：26420371

研究課題名(和文) 未知情報を含む多端子情報理論の精密化

研究課題名(英文) Explicit Analysis of Multi-Terminal Information Theory with Unknown Parameters

研究代表者

野村 亮 (Nomura, Ryo)

専修大学・ネットワーク情報学部・准教授

研究者番号：90329102

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：情報理論に基づき確率構造が未知のネットワーク、特に代表的な通信の数理モデルである1.情報源符号化問題、2.通信路符号化問題、における高信頼符号化に関する研究を行った。これらの問題において誤り確率を適切に設定すると確率構造が未知の状況が混合情報源、混合通信路により表されることに着目し特に混合情報源、混合通信路を対象とした解析を行った。研究期間を通じては多対一および一対一情報源符号化モデル、一対一通信路符号化モデルに関する結果を得ることができた。

研究成果の概要(英文)：We have analyzed fundamental properties of information theory. In particular, we have considered mixed sources and mixed channels, and derived the fundamental limits on the reliable communication. The situation where we know the class of probabilistic model but we do not know its parameter, can be formulated by the mixed probabilistic models. Hence, our approach is meaningful from the view point of the practical situation. We have determined the optimal first- and second-order achievable rates or rate regions in several problems.

研究分野：通信・ネットワーク工学，情報理論

キーワード：情報理論 情報スペクトル理論 シャノン理論

1. 研究開始当初の背景

情報社会の進展に伴い、ネットワーク上で通信される情報量は増加の一途を辿っている。そうした中、センサネットワークや放送型通信、分散映像符号化等の多対多通信の重要性も高まっている。情報通信の基礎理論の一つである情報理論においては、一対一通信の様々な問題設定において 1) 誤りなくデータを伝送可能な条件、2) それを達成する効率的なアルゴリズムの提案、に関する研究が行われてきた。代表的な問題設定は情報源符号化モデル、通信路符号化モデルである。研究開始当初この条件を従来より精密に評価する研究が登場し、理論的な側面のみならず実用的な面からも注目を浴びていた。これらの成果は有限長解析や2次最適レートの評価等と呼ばれている。

一方、1970年代より一対一通信の問題を複数の送信者から複数の受信者へと拡張した多対多通信問題が扱われるようになった。これらは多端子情報理論と呼ばれ、一対一通信の問題と同様に誤りなく伝送可能な条件や実用的なアルゴリズムが研究されてきた。研究開始当初センサネットワークの情報センシングの過程が多端子情報理論を用いて定式化できることが指摘されており、実用面からも多端子情報理論の発展が望まれていた。従来の情報理論で、二つのモデルが代表的であったように、多端子情報理論でも、複数の情報源の符号化モデル、多入力多出力通信路における符号化モデルの二つが代表的である。

研究開始当初一対一通信の問題設定で行われていた2次最適レートの評価の多くは情報源、通信路の確率構造が既知であることを仮定していた。この問題設定は実際の通信で言えば、それぞれ観測信号の生起確率が既知である設定、ネットワークにおける誤り(ノイズ)の生起確率が既知(平均0、分散 σ^2 の正規分布に従う等)である設定に相当するが、応用上これらの確率構造が未知の状況(例えば分散 σ^2 は未知である状況等)は少なくないと考えられる。実際、現在lzhやgzip等に応用されている一対一の情報源符号化アルゴリズムは、対象の生起確率が未知な状況(例えば多項分布のパラメータが未知である状況等)を対象としている。また最近注目を集めている分散映像符号化問題においても、確率構造が未知である状況への対応が望まれる状況であった。

2. 研究の目的

研究代表者は一対一と多対一の情報源符号化モデルにおいて、混合情報源を仮定した場合の2次最適レートを導出していた。混合情報源はそれ自体重要な情報源であるが、誤り確率等を適切に設定すれば、確率構造が未知の情報源符号化モデルを記述することも判明している。すなわちこれらの結果は確率構造が未知の情報源符号化モデルにおけ

る2次最適レートを評価したと考えることができる。

そこで本研究ではこれらの結果を拡張し、様々な多端子情報理論モデルにおいて情報源、通信路の確率構造が未知である場合に誤りなく伝送可能な条件を精密に評価し、また伝送可能な符号化・復号化アルゴリズムを提案することを目標とした。

3. 研究の方法

本研究では、一対一通信モデルから多対多通信モデルまでを対象に、混合情報源、混合通信路を仮定した場合の、2次最適レートの導出を試みた。特に解析手法として近年有力な手法とされている情報スペクトル的方法を用いた。ただし前提として2次最適レートの導出には1次最適レートの導出が不可欠である。1次最適レートが導出されていない問題に対しては1次最適レートの導出を優先した。また、情報スペクトル的手法ではまず確率構造に特別な仮定をおかない一般情報源や一般通信路を考え、最適レートの一般式を導くことが目標となる。本研究でも一般式が未導出である問題に対しては一般式の導出を行い、その後に混合情報源や混合通信路への適用を試みた。

4. 研究成果

(1) 一対一通信モデル

仮説検定問題における1次および2次最適指数の導出

仮説検定問題は、情報学のみならず社会科学においても重要な問題である。本研究では最も基本的な単純仮説検定問題において第一種の誤り確率が漸近的に以下になるもとでの第二種誤り確率の最適指数の導出を試みた。帰無仮説および対立仮説を定常無記憶な確率分布(情報源)とした場合の1次最適指数はSteinの補題として知られている。本研究ではまず帰無仮説および対立仮説の双方が混合情報源(すなわち定常無記憶分布の混合で表される情報源)である場合の1次最適指数を導出した。この結果はSteinの補題の一般化と言える。また帰無仮説が混合情報源、対立仮説が定常無記憶情報源の場合の2次最適指数も導出した。さらには対立仮説および帰無仮説が一般の確率分布の混合である場合に関して、複合仮説検定との関係についても明らかにした。複合仮説検定とは実際の確率分布が未知の状況を表現したモデルで本研究課題と密接な関係がある。この結果より情報源が未知の仮説検定問題と混合情報源を仮定した仮説検定問題の新たな関係が明らかになった。

微少な誤り確率を許容するコスト付き可変長情報源符号化問題における1次および2次最適レートの導出

現在実用化されているgzipやlzhなどのデータ圧縮手法のベースとなる可変長情報源符号化問題を対象としている。応用上極めて

重要な同問題に対して、本研究では二つの一般化を考えた。一つは符号長を一般化した概念である符号コストに対してその最適符号化レート（平均符号コストの最小値）を考えること、もう一つは誤りを許容しない問題から微少な誤り確率を許容する問題へと問題を拡張すること、である。各符号語に非対称な符号コストが課せられている状況は例えば、モルズ符号におけるドットおよびダッシュの送信時間が異なる状況などを表す。（従来の符号長で考えるとドットおよびダッシュの送信にかかるコストが同じと仮定していることになる。）このように符号長でなく符号コストを考えることは実用上大変重要である。また微少な誤り確率を許容する可変長符号化問題は近年特に注目を集めている問題であり、定常無記憶情報源に対する2次最適レートの振る舞いが誤りを許容しない場合と異なることが知られている。また歪みを許す情報源符号化問題と密接な関係もあり、こちらも応用上重要である。

本研究では、同一一般化問題に対してまず符号コストに、ある制約を課したもとの1次最適符号コストレートの一般式を導出した。また異なる符号コストを考えた場合の同最適レートの関係についても考察を行っている。さらに2次最適符号コストレートの一般式についても導出した。この一般式を用いて定常無記憶情報源に対しても2次最適符号コストレートを導出している。混合情報源に関する適用は今後の課題である。

コスト制約付き混合通信路符号化問題における2次最適レートの導出

情報源符号化問題と対をなす重要な問題の一つが通信路符号化問題であり、情報源符号化問題が情報をいかに効率よく送信するかに着目しているのに対し、通信路符号化問題は情報をいかに正確に送信するかに着目している。本研究では混合通信路の要素通信路のクラスに仮定をおくことにより2次最適レートの導出を行った。本結果において特筆すべきは、コスト制約付き混合通信路というより実用的な通信路クラスを考えたこと、および要素通信路の連続の混合を含む一般的な混合通信路に対して結果を得たこと、の二つである。

微少な誤り確率を許容する可変長情報源符号化問題における1次および2次最適オーバーフローしきい値の導出

可変長情報源符号化問題の評価基準の一つとしてオーバーフロー確率がある。オーバーフロー確率とは符号長があるしきい値を超過する確率を表し、オーバーフロー確率の解析はバッファに制約のある符号器や復号器の設計開発に寄与する。本研究では微少な誤り確率を許容する可変長符号化においてオーバーフロー確率が一定値以下になるためのしきい値の下限（最適オーバーフローしきい値）の一般式導出を試み、実際に1次および2次最適しきい値の一般式導出に成功

した。得られた一般式は一見複雑であるが、別表現を考えることにより混合情報源に対して容易に適用できることが確認できた。この問題は今後符号コストのオーバーフロー確率への発展が期待できる。

微少な歪みおよび符号長超過確率を許容する可変長有歪み情報源符号化問題における達成可能条件の導出

有歪み情報源符号化問題は、符号化前のデータと復号化後のデータ間に一定の歪みを許容する問題であり、画像圧縮や音声圧縮などへの適用が考えられる。有歪み情報源符号化は可変長符号化および固定長符号化に大別できる。本研究では可変長有歪み符号化問題を対象とし、歪み超過と符号長超過の和事象の発生確率が一定値以下になるための必要十分条件を考察した。得られた一般式は一般の情報源を対象とした広いクラスに対する結果である。一方この結果を直接混合情報源に適用することは困難であることが分かったため、今後は混合情報源に対する必要十分条件を導出することが目標となる。

歪み超過確率を許す固定長有歪み符号化における達成可能条件の導出

固定長有歪み符号化問題を対象としている同問題においては従来、歪み超過確率を許容しない場合の符号長の最小値を表すいわゆるレート・歪み関数が考察されていた。近年では、混合情報源を含む一般情報源に対するレート・歪み関数も得られていた。一方で歪み超過確率を許容する場合のレート・歪み関数に対しては定常無記憶情報源という限定的な情報源に対しての結果が得られているのみであった。本研究ではまず歪み超過確率を許容する場合の1次および2次のレート・歪み関数の一般式を導出した。しかしながらこの課題と同様に得られた一般式を直接混合情報源へ適用することが困難であることも分かった。混合情報源に対する結果を得ることが今後の課題となる。

Kullback-Leibler (KL) 情報量を基準とした情報源 Resolvability 問題における1次および2次最適レートの導出

与えられた確率分布を一樣乱数を用いて近似する問題は情報源 Resolvability 問題と呼ばれている。近似の尺度としては変動距離や KL 情報量が一般的である。従来変動距離および正規化された KL 情報量を近似尺度とした場合の最適レートの一般式については導出されていたが、(正規化しない) KL 情報量を近似尺度とした場合の同レートについては未解決であった。本研究ではこの問題設定において1次および2次最適レートの一般式を導出した。得られた結果から本問題と固定長情報源符号化問題の間に密接な関係があることが分かった。この関係を用いると固定長情報源符号化問題の結果を直ちに情報源 Resolvability 問題へと適用可能である。実際に混合情報源に対する1次および2次最適レートを固定長情報源符号化問題に

関する従来結果より得ることができた。

(2)多対一通信モデル

相関のある複数の情報源に対する符号化問題における2次最適レートの導出
多対一の無歪み情報源符号化モデルにおける最も基本的な二対一符号化モデルはSlepian-Wolf(SW)符号化モデルと呼ばれる。研究代表者は以前にSW符号化モデルにおける2次最適レート領域の一般公式を導出し、それを用いて混合情報源における2時最適レート領域を導出していた。本研究では同手法をN対一の符号化モデルへと適用し、まず2次最適レート領域の一般公式を導出した。さらに定常無記憶情報源への適用も行った。混合情報源への適用が今後の課題となる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計4件)

- (1) Te Sun Han and Ryo Nomura, "First- and Second-order hypothesis testing for mixed memoryless sources," *Entropy*, vol. 20, no.3, 2018. 査読有.
DOI: 10.3390/e20030174.
- (2) Hideki Yagi and Ryo Nomura, "Variable-Length Coding with Cost Allowing Non-Vanishing Error Probability," *電子情報通信学会論文誌*, vol. E100-A, no.8, pp. 1683-1692, 2017. 査読有.
DOI: 10.1587/transfun.E100.A.1683.
- (3) Hideki Yagi, Te Sun Han, and Ryo Nomura, "First- and Second-Order Coding Theorems for Mixed Memoryless Channels with General Mixture," *IEEE Trans. Information Theory*, vol. 62, no. 8, pp.4395-4412, 2016. 査読有.
DOI: 10.1109/TIT.2016.2573310.
- (4) Ryo Nomura and Te Sun Han, "Second-Order Slepian-Wolf Coding Theorems for Non-Mixed and Mixed Sources," *IEEE Trans. Information Theory*, vol. 60, no. 9, pp. 5553-5572, 2014. 査読有.
DOI: 10.1109/TIT.2014.2339231.

[学会発表](計17件)

- (1) Ryo Nomura, "Source Resolvability with Kullback-Leibler Divergence," 2018 IEEE International Symposium on Information Theory, 2017.
- (2) Ryo Nomura and Hideki Yagi, "Overflow Probability of Variable-length Codes Allowing Non-Vanishing Error Probability," 2017 IEEE Information Theory Workshop, 2017.
- (3) 野村亮, 八木秀樹, "微少な誤り確率を

許容する可変長符号化における楽観的符号化定理,"電子情報通信学会情報理論研究会, 2017.

- (4) Te Sun Han and Ryo Nomura, "First- and Second-Order Hypothesis Testing for Mixed Memoryless Sources with General Mixture," 2017 IEEE International Symposium on Information Theory, 2017.
- (5) 野村亮, 八木秀樹, "微少な誤り確率を許容する可変長符号化におけるオーバーフロー確率について,"電子情報通信学会情報理論研究会, 2017.
- (6) Hideki Yagi and Ryo Nomura, "Variable-Length Coding with Cost Allowing Non-Vanishing Error Probability," 2016 International Symposium on Information Theory and Its Applications, 2016.
- (7) Ryo Nomura and Hideki Yagi "Variable-Length Lossy Source Coding Allowing Some Probability of Union of Overflow and Excess Distortion," 2016 IEEE International Symposium on Information Theory, 2016.
- (8) Hideki Yagi and Ryo Nomura, "Variable-Length Coding with Cost Allowing Non-Vanishing Error Probability," 第38回情報理論とその応用シンポジウム, 2015.
- (9) 野村亮, "相関のある複数の情報源に対する二次符号化定理,"第9回シャノン理論ワークショップ, 2015.
- (10) Hideki Yagi, Te Sun Han and Ryo Nomura, "First- and Second-Order Coding Theorems for Mixed Memoryless Channels with General Mixture," 2015 IEEE International Symposium on Information Theory, 2015.
- (11) Hideki Yagi and Ryo Nomura, "Variable-Length Coding with Epsilon-Fidelity Criteria for General Sources," 2015 IEEE International Symposium on Information Theory, 2015.
- (12) Ryo Nomura and Hideki Yagi "Information Spectrum Approach to Fixed-Length Source Coding Problem with Some Excess Distortion Probability," 2015 IEEE International Symposium on Information Theory, 2015.
- (13) 八木秀樹, 野村亮, "符号長又は歪みの超過確率を許容する忠実度規範付き楽観的符号化と強逆定理,"第37回情報理論とその応用シンポジウム, 2014.
- (14) 野村亮, 八木秀樹, "符号長超過と歪み超過の和事象に基づく忠実度規範付き可変長符号化定理,"第37回情報理論とその応用シンポジウム, 2014.
- (15) Hideki Yagi and Ryo Nomura, "Channel

Dispersion for Well-Ordered Mixed Channels Decomposed into Memoryless Channels," 2014 International Symposium on Information Theory and Its Applications, 2014.

- (16) 八木秀樹, 野村亮, “微少なオーバーフロー確率を許容する忠実度規範付き符号化,” 電子情報通信学会情報理論研究会, 2014.
- (17) 野村亮, 八木秀樹 “微少な歪み超過確率を許す固定長有歪み符号化定理,” 電子情報通信学会情報理論研究会, 2014.

〔その他〕

ホームページ等

<http://reach.acc.senshu-u.ac.jp/Nornir/search.do;jsessionid=DC9C1237D18D9168806E2C0F190E56A2?type=v01&uid=1208060>

6. 研究組織

(1) 研究代表者

野村 亮 (Nomura, Ryo)

専修大学・ネットワーク情報学部・准教授

研究者番号：90329102