

学位論文題名

統計学とグラフ理論に基づくコンテンツ権利処理技術と その応用に関する研究

学位論文内容の要旨

近年、高速なネットワークの普及や高機能・高性能な端末の普及により、映像を中心としたいわゆるコンテンツ配信サービスが普及期を迎えている。しかし、ネットワークや端末の普及と比較して、市場に供給されるデジタルコンテンツはまだ限定的で、放送、映画などの従来のメディアからみると、質・量の面で大幅に見劣りするのが現状である。コンテンツの市場への供給を妨げる要因としては、法制度などの問題や、不正なコンテンツ利用などが指摘されているが、供給側から見たときに大きな障害となっているのが、コンテンツの権利処理の煩雑さである。

デジタルコンテンツの特性として、消費者側がいくらコンテンツを利用しても無くなるものではないため、コンテンツを消費者に提供するサービス事業者から権利者への適正なフィードバック（利用報告）がなければ、コンテンツの権利者に利益が還元されないという問題がある。現状、このコンテンツの利用報告が大きな問題を抱えている。ひとつは日本で制作されるコンテンツの9割以上を占める放送コンテンツの権利処理に関するものであり、もうひとつは、近年携帯端末、i-Podなどの普及により急激に利用が増えている、音楽配信の権利処理である。

放送コンテンツは、インターネットへのコンテンツ配信のもっとも重要なコンテンツ供給元と考えられており、それに向けた技術開発、法整備などが進められた結果、NHKが本年12月からVODサービスを開始し、民放各局も準備を進めている。しかし、その供給の前提となるコンテンツの権利処理、特に番組中で使用される音楽の権利処理が十分に進んでいないのが現状である。従って、実際に供給可能な放送コンテンツは、放送されたコンテンツ全体から見ると極一部に過ぎない。一方、音楽配信に関しては、当初はCD等のパッケージの売り上げを阻害するとの懸念から十分な供給がなされなかったが、i-Podの急激な普及などにより、急激に配信回数、配信楽曲数が増加している。その結果、上記の放送の利用楽曲にも共通の課題が浮かび上がった。権利者と信託契約を結んでいる著作権管理団体において、権利使用料の分配業務が、急激な利用報告数の増加に対応しきれなくなってきた。このままの状態が続けば、利用者が支払った使用料が権利者に分配できなくなるという危機的な状況となっている。

両者に共通する課題は、(1) 利用状況の把握に際しての「全量」的な処理が、現状の制度、コストなどの観点から困難であること、(2) 報告数の増加に対して、権利使用料の分配が、現状の仕組みでは対応しきれなくなってきたこと、と考えられる。

本研究では、前者に対しては、いわゆるサンプリングの方法で、後者に対しては、既存の権利処理データベースの統合という方法で、これらの課題の解決を目指す。

本論文は、全7章から構成される。各章の概要は以下の通りである。

第1章では、本論文の背景、目的、および構成について述べている。

第2章では、はじめに、コンテンツの利用状況把握に関して現状の状況を分析し、放送における楽曲利用状況の把握に現状大きな問題があることを示す。さらにその問題について、権利管理団体を中心とした今までの取り組みを説明することにより、その本質的な課題が利用状況把握にかかるコストと精度のトレードオフにあることと、それが現状解決されていないことを明らかにする。その課題の解決法として、フィンガープリント技術と呼ばれるコンテンツの特定技術の適用法とサンプリングによる利用状況把握の可能性を示す。

第3章では、上記課題に対して、サンプリングに基づく利用状況把握の具体的な方法を提案する。まず、放送における楽曲の利用率の推定が、視聴率などの一般の母比率推定とは異なる問題であることを指摘し、利用率の推定問題に関するサンプリングのモデル化を行う。次にこのモデルに基づき、サンプリングに伴う誤差を理論的に解析し、サンプリングの期間が十分短いときに、サンプリングの誤差がほぼ一定となるという事実を理論的に明らかにする。さらに、FMラジオ局で用いられた楽曲の、実際の利用状況データを用いたシミュレーションにより、サンプリングに基づく誤差の挙動の分析を行う。最後にこれらの結果を用いて、具体的なサンプリングの際に決めなければならない、サンプリングの割合、およびサンプリングの期間の決定法を示すとともに、利用率の推定に必要なコストと誤差の関係を明らかにする。その結果を用いて、従来、誤差の見積もりが困難であったために、「全量」報告を前提とせざるを得なかった権利の利用許諾に、一定の誤差を前提とした利用許諾という新しいスキームを提案する。

第4章では、適正なコンテンツ使用料の分配の基礎となる、コンテンツ権利情報データベースの課題を分析し、その課題の根本的な原因が、目的の異なる2種類のデータベース(流通用と権利処理用)の存在にあることを示す。また、この根本的な原因の解決に繋がらない、今までのコンテンツ権利情報データベースの整備に関する取り組みとその問題点を分析し、その本質的な課題が、情報量が極めて少ないという楽曲情報の特性と、その特性を考慮していないその表現法にあることを示す。これらの課題に対する解決策として、複数のデータベース統合のための楽曲情報の階層的表現法を提案する。この表現法の特徴は、コンテンツの情報をコンテンツそのもの(オブジェクト)とその属性として表現することにある。統合されたデータベースの運用に必要な新規データの登録・管理法と、複数の楽曲情報データベースの名寄せを効率的に行う方法を提案する。

第5章では、上記の階層的表現法を前提とし、コンテンツ権利情報の登録・管理法について議論する。コンテンツの権利情報の登録時の問題は、新たに登録しようとする情報と、既存の情報との関係を事前に確認することが困難な点にある。本章では、オブジェクトを点、属性を(有向)枝とする有向グラフと考え、さらに有向グラフの一般的な性質として縮退可能という概念を提案する。この概念に基づく点の集合(コンポーネントと呼ぶ)を用いることにより、既存の権利情報の柔軟な表示と検索が可能となり、既存の情報と整合性を保った新規の情報の登録が可能となる。このコンポーネントを効率よく発見するアルゴリズムを提案し、実験により有効性を確認した。

第6章では、第5章と同様に上記の階層的表現法を前提とし、コンテンツ権利情報の統合の問題を、オブジェクトを点、属性をラベル付有向枝と考え、複数のラベル付有向グラフと属性間の関係からなるシステムの最適化問題として定式化する。さらにこの最適化問題に対する効率的なアルゴリズムを提案し、実験により有効性を確認する。従来、個別のラベル付有向グラフの最適化問題は

広く論じられてきたが、互いに (制約) 関係を持つ複数のラベル付有向グラフの最適化は議論されてこなかった。実験の結果、複雑で互いに制約条件を持つラベル付有向グラフに関して、このアルゴリズムの効果が大きいことを示した。

最後に第 7 章で、本論文の成果を総括とともに、実システムへの適用に向けた課題を整理した。

学位論文審査の要旨

主査	教授	水田	正弘
副査	教授	大内	東
副査	教授	赤間	清
副査	准教授	南	弘征

学位論文題名

統計学とグラフ理論に基づくコンテンツ権利処理技術と その応用に関する研究

近年のデジタルコンテンツの普及に関し、権利処理の煩雑さが、おもにコンテンツ供給側での問題として問題となっている。デジタルコンテンツは物理的に消費されないため、サービス事業者から権利者への適正なフィードバック (利用報告) がなければ、権利者に利益が還元されない。代表的なデジタルコンテンツである放送コンテンツにおいても、特に番組中で使用される音楽の権利処理は、大雑把な計測に基づいており、現在もなお精度が不十分である。一方、昨今の音楽配信の急速な普及などをあわせて考えた際、仮にコンテンツの利用報告が完全であっても、権利者と信託契約を結んでいる著作権管理団体の処理能力を越える可能性が指摘されつつあり、使用料が権利者に分配できなくなるという危機的な状況となっている。

本論文では、以上の背景の本質を、(1) 利用状況の「全量」把握とその扱いが、現在きわめて困難であること、(2) 権利使用料の分配も、現状では対応しきれなくなってきたこと、とし、前者に対して統計科学的な技法の導入を、後者に対して既存のデータベースの有機的統合をそれぞれ行い、これら課題の解決を目指している。

本論文は、全7章から構成されている。

第1章では、本論文の背景、目的、および構成が記されている。

第2章では、コンテンツの利用状況把握に関する現状分析と、それを踏まえた問題の具体的な指摘と、それに対する権利管理団体を中心とした従来の施策を述べ、利用状況把握のコストと精度のトレードオフが課題の本質であること、現状なお未解決であることを示している。さらに、その解決法として、フィンガープリント技術と呼ばれるコンテンツ特定技術の適用法とサンプリングによる利用状況把握の可能性を示している。

第3章では、サンプリングに基づく利用状況把握の技法を提案している。放送における楽曲の利用率の推定は、視聴率などの一般的な母比率推定と異なることを示し、利用率の推定問題に関するサンプリングのモデル化を行っている。さらに、このモデルに伴う誤差を解析し、サンプリング期間が十分短い場合、誤差がほぼ一定となることを理論的に示し、FM ラジオで放送された楽曲データを用いて、誤差の挙動を分析している。以上の結果を踏まえ、具体的なサンプリングの際に決め

なければならない、サンプリングの割合および期間の決定法を示すとともに、利用率の推定に必要なコストと誤差の関係を論じている。さらに、従来は「全量」報告を前提とせざるを得なかった権利の利用許諾に、一定の誤差を前提とした利用許諾という新しいスキームを提案している。

第4章では、適正なコンテンツ使用料の分配の基礎となる、コンテンツ権利情報データベースの課題を分析し、根本原因が、目的の異なる2種類のデータベース(流通用と権利処理用)の存在であることを論じている。また、従前のコンテンツ権利情報データベースの整備に関する取り組みがこの原因の解決に寄与せず、問題の本質が、情報量が極めて少ないという楽曲情報の特性と、それを考慮していないデータベースの表現法にあるとし、これらの解決策として、複数のデータベース統合のための楽曲情報の階層的表現法が提案されている。具体的には、コンテンツの情報をコンテンツそのもの(オブジェクト)とその属性として表現し、統合されたデータベースの運用に必要な新規データの登録・管理法と、複数の楽曲情報データベースの名寄せを効率的に行う方法を論じている。

第5章では、前章の階層的表現法を用いた、コンテンツ権利情報の登録・管理法を論じている。コンテンツの権利情報の登録時の問題が、新規登録情報と、既存情報との関係の事前確認が困難な点にあることを指摘し、オブジェクトを点、属性を(有向)枝とする有向グラフと考え、一般的な性質として縮退可能という概念を提案し、これに基づく点の集合(コンポーネント)を用いることで、既存の権利情報の柔軟な表示と検索、整合性を保った新規情報登録が可能となる。このコンポーネントを効率よく発見するアルゴリズムを提案し、実験により有効性を確認している。

第6章では、コンテンツ権利情報の統合の問題を、複数のラベル付有向グラフと属性間の関係からなるシステムの最適化問題として定式化した上で、効率的解法を提案し、実験による有効性を確認している。その結果、複雑で互いに制約条件を持つラベル付有向グラフに関して、提案アルゴリズムの効果が大きいことを示している。

最後に第7章で、本論文の成果を総括し、実システムへの適用に向けた課題を挙げている。

これを要するに、本論文はデジタルコンテンツにおける権利処理に対する技法について新知見を得たものであり、統計科学ならびに情報科学に貢献するところ大なるものがある。

よって、申請者は北海道大学博士(情報科学)の学位を授与される資格あるものと認める。