

卒業論文 2023 年度 (令和 05 年)

リズム予測を用いたネットワーク音楽演奏における遅延の影響の軽減手法
の提案

慶應義塾大学 環境情報学部
松藤舜

リズム予測を用いたネットワーク音楽演奏における遅延の影響の軽減手法の提案

現在、今まで以上にネットワークを介した遠隔で複数の演奏者がリアルタイムで行う音楽演奏への注目が高まっている。これらの演奏をネットワーク音楽演奏と呼ぶ。音楽演奏において遅延は致命的な問題であり、時に演奏を不可能にする。しかし、ネットワークを介したグローバルな演奏を前提とするなら、遅延は避けられない。本研究ではネットワーク音楽演奏における遅延を減らすのではなく、遅延を前提として演奏に影響を与えないシステムを提案する。その過程で、ネットワーク音楽演奏における遅延の問題について考察し、過去の研究を紹介する。そしてそれらの研究をふまえたうえで、本研究では演奏予測を用いたシステムを提案する。また実験の結果、演奏予測を用いたシステムは、遅延がある状況下でも演奏の崩れを抑えることができることを示し、本研究の有効性を示した。

キーワード:

1. ネットワーク音楽演奏, 2. 遅延, 3. 演奏予測, 4. OSC

慶應義塾大学 環境情報学部
松藤舜

The Use of Rhythm Prediction to Reduce the Effects of Latency in Network Music Performance

Recently, there has been an ever increasing growing interest in ' Networked Music Performances (NMP)', in which multiple performers perform music in real time over a network. In NMPs there will always be a delay between the performers, which causes a phase shift in the music performance, causing a breakdown of the performance, which is a fatal problem for NMPs. However, network delay is unavoidable when assuming a global performance over a network. In this research, instead of reducing the latency in NMPs, I propose a system that allows performers to play music with each other regardless of the latency. In doing so, I will discuss the underlying issues of latency in NMPs, introduce past research, and propose a system using music prediction to mitigate the effects of latency. I have also conducted experiments to show that the proposed system can mitigate the effects of latency even in the presence of latency.

Keywords :

1. Network Music Performance, 2. Latency, 3. Music Prediction, 4. OSC

Faculty of Environment and Information Studies, Keio University
Shun Matsufuji

目次

第1章	序論	1
1.1	インターネットを介した音楽演奏への注目	1
1.2	遅延の問題	1
1.3	本論文の構成	2
第2章	背景	3
2.1	ネットワーク音楽演奏	3
2.2	ネットワーク音楽演奏システムの構造	3
2.3	スマートネットワーク音楽演奏	3
2.4	演奏情報の表現	3
2.5	Open Sound Control	3
2.6	Adaptive Metronome	4
2.6.1	相互メトロノームの実験	5
2.7	Tablanet	5
2.8	Alexandraki	5
第3章	提案手法	6
3.1	概要	6
3.2	遅延, ジッター対処としての予測	6
3.3	Adaptive Metronome (既存手法) の課題点	6
3.4	仮説	6
第4章	設計	7
4.1	本システムのアーキテクチャ	7
4.2	拍認識の実装	7
4.2.1	共通テンポの算出	8
4.3	演奏予測の実装	8
4.4	位相修正	9
4.4.1	クロック同期	9
第5章	実験	10
5.1	提案システムの実験	10
5.1.1	手法	10
5.1.2	結果	10

5.1.3 評価	10
第 6 章 結論	11
6.1 本研究のまとめ	11
6.2 本研究の課題	11
6.3 今後の展望	11
謝辞	12

図 目 次

1.1	連成振動のモデル. [1] より引用	2
2.1	OSC メッセージと OSC バンドル	4
2.2	Adaptive Metronome の実験 [2]	4
4.1	本システムのアーキテクチャ	7
4.2	予測システム	8

表 目 次

第1章 序論

本章では本研究の背景，課題及び手法を提示し，本研究の概要を示す．

1.1 インターネットを介した音楽演奏への注目

世界的なインターネット，SNS の普及がもたらしたグローバルなコミュニケーションの発展により，世界中のアーティストをつなぐグローバルな音楽のコミュニティが形成している．私自身も作曲者，ベース演奏者，DJ として様々な方面でこれらのコミュニティに活発に参加している．

従来まではこのようなインターネットを介したコミュニティでは音楽演奏を共有する手段として，事前に音楽演奏を録音し，ファイル共有サービスなどを用いて音声データを送り合うしかなかった．バンド，楽団など複数人の同時に合奏を行う音楽形式ではこのような音声データを集めたのち，一人が演奏音声を合成し一つの楽曲として聞ける形に変換するという手段が主流である．しかし Skype や Zoom などの VoIP アプリケーションが普及し，インターネットを介したリアルタイムの音楽演奏が現実的になってきた．このような音楽演奏を「ネットワーク音楽演奏 (Network Music Performance)[3]」と呼ぶ．

その状況の中，2020 年の新型コロナウイルス感染症が流行し現実で集まった生演奏が難しくなり，このような演奏形式が益々注目されるようになった．この時期に Yamaha は低遅延でネットワーク音楽演奏を行える Syncroom[4][5] を発表し，2023 年 6 月に株式会社ズームはネットワーク音楽演奏が行えるデバイス S6 SessionTrak[6] を発表するなど，直近でもこのネットワーク音楽演奏に対する注目度は高まっていることがわかる．

1.2 遅延の問題

ネットワーク音楽演奏において遅延は致命的な問題である．リアルタイムで相手の演奏に同期して自らの演奏を行う必要があるネットワーク音楽演奏において，このプロセスへの遅延は演奏に多大な影響を与える．この効果は様々な研究で観測され，様々な演奏環境により差があれど遅延時間に差があれど 20～40ms 以上で遅延を感じることができ，80～100ms 以上になると演奏が不可能になるとされている．[7][8]

音楽演奏は一定の「拍子」で行われ，各演奏者は相手の演奏を聴き，拍子を認識し，その拍子に合わせて演奏を行う．単独で演奏を行う奏者は一つのオシレータのように一定の間隔で刻まれる拍に合わせて音を配置していると捉えることができる．この場合遅延がある二人同士のネットワーク音楽演奏は 1.1 のように鍊成振動のモデルで捉えられる．

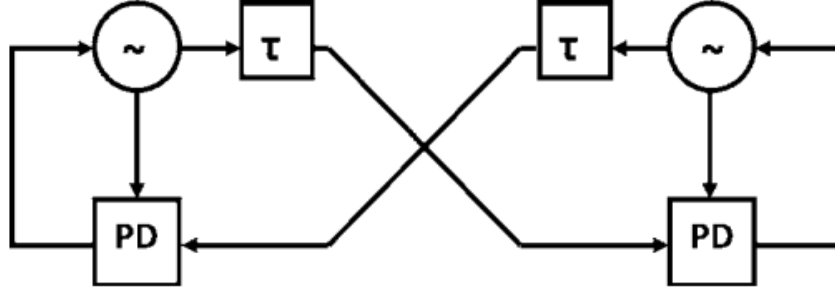


図 1.1: 連成振動のモデル. [1] より引用

ここでPDはPhase Detector, すなわち演奏者の捉えた拍の感覚, τ は遅延時間である. [1] ではこのモデルを踏まえて遅延下の2人の演奏者同士が合意するテンポを以下のように表した.

$$\Omega_{i+1} = \frac{\omega_i}{1 + k\tau}$$

ω は両演奏者の平均テンポ, k は振動の状態更新を表す低数であり, τ は遅延を表す. この通り遅延が大きくなると演奏テンポが楽曲を通して遅くなることがわかり, [8] などの遅延付きのネットワーク音楽演奏の実験でもこの傾向は観察できる.

1.3 本論文の構成

本論文における以降の構成は次の通りである.

2章では, 背景と関連研究を述べる. 3章では, 本研究の提案手法を述べる. 4章では, 3章で述べたシステムの実装について述べる. 5章では, 本システムの実験評価を行い, 考察する. 6章では, 本研究のまとめと今後の課題についてまとめる.

第2章 背景

本章では本研究の背景について述べる.

2.1 ネットワーク音楽演奏

ネットワーク音楽演奏で演奏を届ける手法は, VoIP アプリケーションのように音声の波形情報を処理し低遅延で届ける方法 [4][9][10], 演奏情報を MIDI を使い楽器, 音程, 強度などに抽象化し, 受信した機器が演奏データを元に音を合成するという方法 [11][12] と, 2 種類の手法に分けることができる. 本研究では演奏予測を扱いやすい後者の手法を用いる.

2.2 ネットワーク音楽演奏システムの構造

2.3 スマートネットワーク音楽演奏

[13] [2]

2.4 演奏情報の表現

本研究では演奏を MIDI の形式であらわし, 演奏情報を伝達する. MIDI とは

2.5 Open Sound Control

Open Sound Control (OSC)[14] はソフトやコンピュータ同士でリアルタイムの通信を行うためのプロトコルである. 元々ネットワークを経由した楽器の演奏を前提として開発されているためデータ型として MIDI を使えるうえ, 時刻タグ (OSC Time Tag) をメッセージとまとめて送ることで正確な音の再生のタイミングを送信者から指定することが可能である. そのうえ, OSC は広く使われていてライブラリや実装が豊富であるため本研究のシステムで採用した.

OSC のメッセージは単体で, もしくは時刻を指定して同時に再生する複数のメッセージをバンドルにまとめて送ることができる. 32 ビットで UNIX 時間での秒数を, 32 ビッ

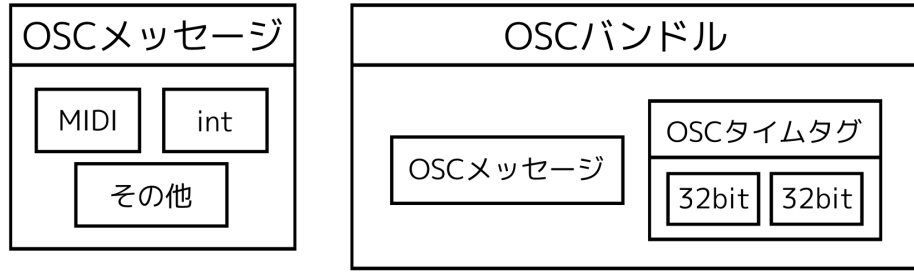


図 2.1: OSC メッセージと OSC バンドル

トで秒数の小数部を表す時刻タグを OSC で使用されるため高解像度で再生の管理が可能であり，時刻タグの精度は本研究では十分とする．

OSC サーバーが OSC メッセージや OSC バンドルを受信するとそれらを引数として OSC メソッドが実行される．OSC メソッドは OSC アドレス空間というツリー構造の中に配置され，OSC メッセージや OSC バンドルに付属される OSC アドレスを宛先として実行される OSC メソッドが選ばれる．

2.6 Adaptive Metronome

Battello らによる Adaptive Metronome[2][15] は，演奏者の演奏をリアルタイムで分析し，演奏者の演奏に合わせてメトロノームのテンポを変化させるシステムである．

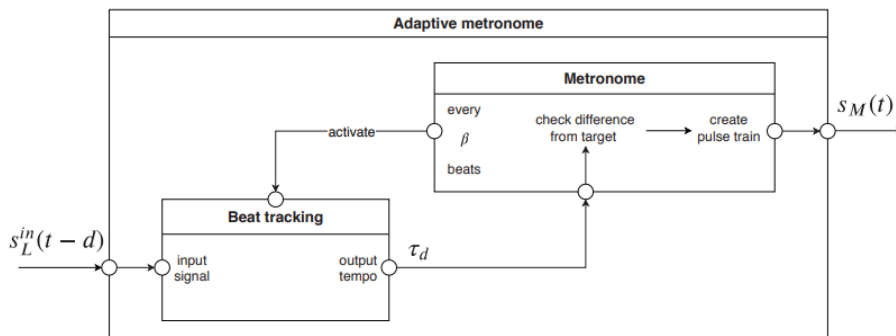


図 2.2: Adaptive Metronome の実験 [2]

Adaptive Metronome は親子構造を持っている．親の演奏者は演奏すると同時にシステムはリアルタイムでその演奏の拍を推定し，子演奏者に音声と拍情報を送信する．子演奏者のシステムはその情報を受取り，親演奏者の演奏に合わせてメトロノームのテンポを変化させる．またこのときメトロノームの音は親から子への遅延を考慮して位相をずれて再生される．

こうすることで小演奏者に聞こえるメトロノームは親演奏者の遅延された演奏に関わらず、常に親演奏者の無遅延の演奏に合わせたテンポで再生される．このシステムを用いた実験では 120ms の遅延下での演奏を行ったうえでも、被験者は抵抗を感じることなく演奏することができたという結果が得られた．[2]

2.6.1 相互メトロノームの実験

当初の Adaptive Metronome の実験では、親演奏者の演奏を子演奏者が聴き、それに合わせて演奏するという形で実験が行われた．後に Battelo らは親子構造を持たず、相互的に Adaptive Metronome を聞きあう実験を行った．実験の結果は現状まだ不十分であるが、このような相互的なシステムでもある程度の効果が得られることが示された．

2.7 Tablanet

Tablanet[16] は、タブラ奏者の演奏をリアルタイムで分析し、演奏者の演奏に合わせてタブラのテンポを変化させるシステムである．

2.8 Alexandraki

Alexandraki らによる研究 [13][17] では、演奏者の演奏をリアルタイムで分析し、事前収録した演奏を実際の演奏に合わせて再生するシステムを提案している．

第3章 提案手法

本章では提案手法について述べる.

3.1 概要

本研究では遅延を減らすのではなく、遅延を前提として快適に演奏を行うことができるシステムを目指す. その手法として演奏相手のリズム予測を行い、本来の演奏と同じタイミングで予測音を再生するシステムを提案する.

3.2 遅延, ジッター対処としての予測

3.3 Adaptive Metronome (既存手法) の課題点

3.4 仮説

遅延のあるネットワーク下でも、演奏相手の演奏を予測しながら演奏を行うことができれば、遅延の量に関係なくまるで同じ部屋にいるかのように演奏できると仮説を立てる.

従来の Adaptive Metronome と違い、メトロノームの音ではなく予測の音を再生することで演奏者はまるで人間の演奏相手と演奏しているかと同等の体験を得られることができると考える. 本システムを用いると音楽における音のニュアンス、表現を保ちつつ、遅延の補償を行うことができる. なお音楽的な表現を保つことが目的であるため、表現を伝達するのに十分な予測精度を得られればよいと考える.

第4章 設計

本章では提案手法の実装について述べる。

4.1 本システムのアーキテクチャ

本システムは従来のネットワーク音楽演奏のシステムに加え，拍認識部分，演奏予測，位相修正の3つの部分を追加している。

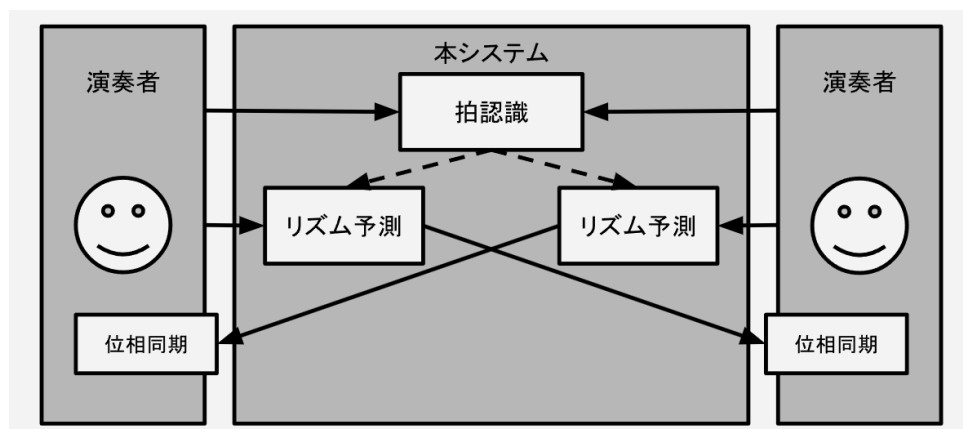


図 4.1: 本システムのアーキテクチャ

4.2 拍認識の実装

演奏予測を行い，相手の演奏を再生するためには2人の演奏のテンポを知る必要がある．人間的な誤差，または表現の一環として楽曲を通して演奏のテンポは変化していくため，リアルタイムで直近数秒間の演奏を参照して現在のテンポを認識するアルゴリズムが必要である．本システムの拍認識では [18] で使われたアルゴリズムを用いる．

この拍認識アルゴリズムでは Dixon のアルゴリズム [19] に従って打鍵間隔 (Inter-onset interval: IOI) のクラスタリングを行う．IOI とは任意の2つの音の開始時点の時間間隔のことを指す．このアルゴリズムの分脈では 0.025 秒から 2.5 秒の間の IOI しか扱わず，これより小さい IOI は「同時に鳴った音」として扱う．ある演奏区間に存在するすべての有効な IOI に対し以下のようにクラスタリングを行う．

Algorithm 1 Dixon's IOI Clustering Algorithm

```

1: for each pair of onset times  $t_i, t_j$  with  $t_i < t_j$  do
2:   if  $0.025 < t_j - t_i < 2.5$  then
3:     Let  $I = t_j - t_i$ 
4:     Find cluster  $C_k$  such that  $|Average(C_k) - I|$  is minimum
5:     if  $k$  exists and  $|Average(C_k) - I| < \Delta$  then
6:        $C_k \leftarrow C_k \cup \{I\}$ 
7:     else
8:       Create new cluster  $C_m := \{I\}$ 
9:     end if
10:  end if
11: end for

```

このアルゴリズムでは全クラスタのうち、平均値が最も近いクラスタに対して、その平均値との差が Δ 以下の場合にのみ IOI をクラスタに追加する。

4.2.1 共通テンポの算出

4.3 演奏予測の実装

演奏予測には様々な手法が存在する。

今回は音楽的な表現の伝達に十分な精度しか必要ないため、/citetablenet で用いられた手法を参考にした簡易的な予測手法を用いる。

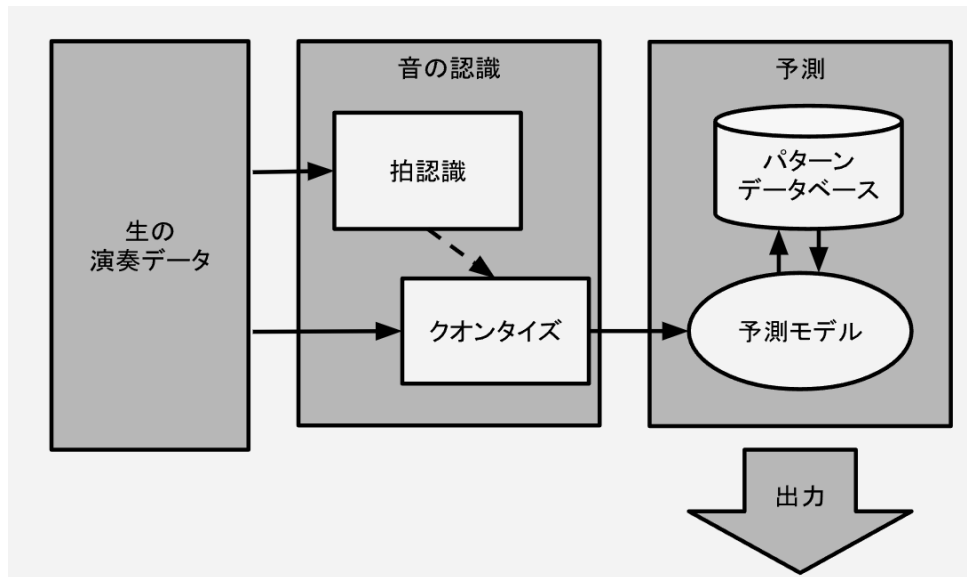


図 4.2: 予測システム

4.4 位相修正

受信部分において演奏相手の演奏予測を受け取った際、遅延を考慮して位相をずらして再生する必要がある。

4.4.1 クロック同期

OSC タイムタグは UNIX 時間を基準にした絶対的な時間で送られるため、各クライアントのわずかな時計のずれで再生時間が変わってしまう。そのため時計のずれを考慮したタイムタグの修正が必要である。本システムでは簡易的であり、十分な精度が得られるという理由で Cristian のアルゴリズム [20] を採用している。

第5章 実験

本章では，4人の被験者に対して行った提案システムの実験について評価を行う．

5.1 提案システムの実験

4人の被験者に対して，提案システム，Adaptive Metronome，従来の遅延付きのネットワーク音楽演奏システムを用いて演奏を行ってもらい，それぞれのシステムにおける演奏のしやすさの評価を行った．被験者はドラム未経験の1人，3年程度のドラム経験の2人，10年以上のドラム経験の1人との計4人である．

5.1.1 手法

5.1.2 結果

5.1.3 評価

第6章 結論

本章では，本研究のまとめと今後の課題を示す．

6.1 本研究のまとめ

6.2 本研究の課題

6.3 今後の展望

謝辞

俺に関わった全てに感謝

参考文献

- [1] Peter Driessen, Thomas Darcie, and Bipin Pillay. The effects of network delay on tempo in musical performance. *Computer Music Journal*, 35:76–89, 03 2011.
- [2] Riccardo Battello, Luca Comanducci, Fabio Antonacci, Augusto Sarti, Stefano Delle Monache, Giovanni Cospito, Enrico Pietrocola, and Filippo Berbenni. An adaptive metronome technique for mitigating the impact of latency in networked music performances. In *2020 27th Conference of Open Innovations Association (FRUCT)*, pages 10–17, 2020.
- [3] John Lazzaro and John Wawrzynek. A case for network musical performance. In *Proceedings of the 11th International Workshop on Network and Operating Systems Support for Digital Audio and Video*, NOSSDAV '01, page 157–166, New York, NY, USA, 2001. Association for Computing Machinery.
- [4] Yamaha. Syncroom について.
- [5] ヤマハ、コロナ禍で遠隔サービス育む 事業化も視野－コロナ禍に挑む. 01 2021.
- [6] 株式会社ズーム. パソコン無しで最大6人のミュージシャンと遠隔セッション、ズーム『s6 sessiontrak』を発売. 06 2023.
- [7] Christopher Bartlette and Mark Bocko. Effect of network latency on interactive musical performance. *Music Perception - MUSIC PERCEPT*, 24:49–62, 09 2006.
- [8] 佑 西堀, 幸生 多田, and 卓朗 曾根. 遅延のある演奏系での遅延の認知に関する実験とその考察. Technical Report 127(2003-MUS-053), ヤマハ株式会社アドバンストシステム開発センター ND プロジェクト, ヤマハ株式会社 アドバンストシステム開発センター ND プロジェクト, ヤマハ株式会社 アドバンストシステム開発センター ND プロジェクト, dec 2003.
- [9] Carlo Drioli, Claudio Allocchio, and Nicola Buso. Networked performances and natural interaction via lola: Low latency high quality a/v streaming system. In Paolo Nesi and Raffaella Santucci, editors, *Information Technologies for Performing Arts, Media Access, and Entertainment*, pages 240–250, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

- [10] Juan-Pablo Cáceres and Chris Chafe. Jacktrip: Under the hood of an engine for network audio. *Journal of New Music Research*, 39(3):183–187, 2010.
- [11] Nicolas Falquet and Dominique Fober. RTP MIDI : Recovery Journal Evaluation and Alternative Proposal. Technical report, GRAME, 2005.
- [12] Robin Renwick. Sourcenode: a network sourced approach to network music performance (nmp). In *ICMC*, 2012.
- [13] Chrisoula Alexandraki and Rolf Bader. Real-time concatenative synthesis for networked musical interactions. *The Journal of the Acoustical Society of America*, 133:3367, 05 2013.
- [14] Wright Matt. Opensoundcontrol specification 1.0, 03 2002.
- [15] Riccardo Battello, Luca Comanducci, Fabio Antonacci, Giovanni Cospito, and Augusto Sarti. Experimenting with adaptive metronomes in networked music performances. 69:737–747, 10 2021.
- [16] Mihir Sarkar and Barry Vercoe. Recognition and prediction in a network music performance system for indian percussion. In *Proceedings of the 7th International Conference on New Interfaces for Musical Expression*, pages 317–320. Association for Computing Machinery, 2007.
- [17] Chrisoula Alexandraki and Rolf Bader. Using computer accompaniment to assist networked music performance. 01 2014.
- [18] Simon Dixon and Emiliós Cambouropoulos. Beat tracking with musical knowledge. *ECAI 2000: Proceedings of the 14th European Conference on Artificial Intelligence*, 06 2000.
- [19] Simon Dixon. Beat induction and rhythm recognition. In Abdul Sattar, editor, *Advanced Topics in Artificial Intelligence*, pages 311–320, Berlin, Heidelberg, 1997. Springer Berlin Heidelberg.
- [20] Flaviu Cristian. Probabilistic clock synchronization. *Distributed Computing*, 3:146–158, 1989.