

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-208188  
(P2019-208188A)

(43) 公開日 令和1年12月5日(2019.12.5)

(51) Int.Cl.	F I	テーマコード (参考)
HO4W 28/02 (2009.01)	HO4W 28/02	5K067
HO4W 24/02 (2009.01)	HO4W 24/02	
HO4W 84/12 (2009.01)	HO4W 84/12	
HO4B 17/373 (2015.01)	HO4B 17/373	

審査請求 未請求 請求項の数 8 O L (全 17 頁)

(21) 出願番号 特願2018-103999 (P2018-103999)  
(22) 出願日 平成30年5月30日 (2018. 5. 30)

(71) 出願人 000004226  
日本電信電話株式会社  
東京都千代田区大手町一丁目5番1号  
(71) 出願人 504132272  
国立大学法人京都大学  
京都府京都市左京区吉田本町36番地1  
(74) 代理人 110001634  
特許業務法人 志賀国際特許事務所  
(72) 発明者 官武 遼  
東京都千代田区大手町一丁目5番1号 日  
本電信電話株式会社内  
(72) 発明者 浅井 裕介  
東京都千代田区大手町一丁目5番1号 日  
本電信電話株式会社内

最終頁に続く

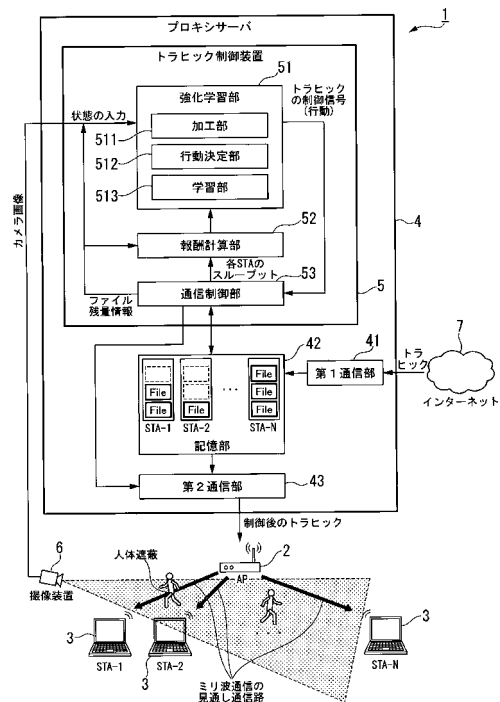
(54) 【発明の名称】 通信システム、トラフィック制御装置及びトラフィック制御方法

(57) 【要約】

【課題】 移動する障害物により無線通信のための見通し通信路に一時的に遮蔽が生じる環境下における合計スループットを増加させる。

【解決手段】 トラフィック制御装置5の行動決定部512は、AP2とSTA3との間の通信環境を撮像した画像データと、プロキシサーバ4の記憶部42が記憶するSTA3宛てのデータのデータ量の情報とを用いて、STA3それぞれのトラフィックの組み合わせとして表される行動の価値を算出する価値関数により複数種類の行動それぞれの価値を算出し、価値に基づいて行動を決定する。通信制御部53は、決定された行動に従ってSTA3宛てのデータが配信されるように通信を制御する。報酬計算部52は、この制御が行われたことによるSTA3の通信状況を取得し、取得した通信状況に基づいて過去の通信状況から向上した程度を表す報酬を計算する。学習部513は、計算された報酬に基づいて価値関数を更新する。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

第 1 通信装置と、前記第 1 通信装置と無線により通信する 1 台以上の第 2 通信装置と、第 1 通信装置から前記第 2 通信装置に送信するデータを取得する第 3 通信装置と、トラヒック制御装置とを有する通信システムであって、

前記トラヒック制御装置は、

前記第 1 通信装置と前記第 2 通信装置との間の通信環境を撮像した画像データと、前記第 3 通信装置が記憶する前記第 2 通信装置宛ての未送信の前記データのデータ量の情報とを用いて、前記第 2 通信装置それぞれのトラヒックの組み合わせにより表される行動の価値を算出する価値関数により複数種類の行動それぞれの価値を算出し、算出した前記価値に基づいて行動を決定する行動決定部と、

10

前記行動決定部が決定した前記行動が表す前記第 2 通信装置それぞれのトラヒックに従って、前記第 2 通信装置宛ての前記データを前記第 1 通信装置に送信するよう前記第 3 通信装置を制御する通信制御部と、

前記通信制御部による制御が行われたことによる前記第 2 通信装置の通信状況を取得し、取得した前記通信状況が過去の通信状況から向上した程度を表す報酬を計算する報酬計算部と、

前記報酬計算部が計算した報酬に基づいて前記価値関数を更新する学習部と、  
を備え、

前記第 1 通信装置は、前記第 3 通信装置から受信した前記第 2 通信装置宛ての前記データを無線により前記第 2 通信装置へ送信する、  
通信システム。

20

**【請求項 2】**

第 1 通信装置と 1 台以上の第 2 通信装置との間の通信環境を撮像した画像データと前記第 2 通信装置宛ての未送信のデータのデータ量の情報とを用いて、前記第 2 通信装置それぞれのトラヒックの組み合わせとして表される行動の価値を算出する価値関数により複数種類の行動それぞれの価値を算出し、算出した前記価値に基づいて行動を決定する行動決定部と、

前記行動決定部が決定した前記行動が表す前記第 2 通信装置それぞれのトラヒックに従って、前記第 1 通信装置から前記第 2 通信装置宛ての前記データが配信されるよう通信を制御する通信制御部と、

30

前記通信制御部による制御が行われたことによる前記第 2 通信装置の通信状況を取得し、取得した前記通信状況が過去の通信状況から向上した程度を表す報酬を計算する報酬計算部と、

前記報酬計算部が計算した報酬に基づいて前記価値関数を更新する学習部と、  
を備えるトラヒック制御装置。

**【請求項 3】**

前記第 2 通信装置の前記通信状況は、前記第 2 通信装置におけるスループット、又は、前記第 2 通信装置宛ての前記データの送信にかかった時間を表す情報である、

請求項 2 に記載のトラヒック制御装置。

40

**【請求項 4】**

前記価値関数は、深層ニューラルネットワークにより近似される、

請求項 2 又は請求項 3 に記載のトラヒック制御装置。

**【請求項 5】**

前記価値関数に用いられる前記画像データは、異なるタイミングにおいて撮影された複数の画像データそれぞれの解像度を低減したのちにピクセル値を正規化したデータである、

請求項 4 に記載のトラヒック制御装置。

**【請求項 6】**

前記価値関数に用いられる未送信の前記第 2 通信装置宛てのデータ量の情報は、複数の

50

前記第2通信装置それぞれ宛ての未送信の前記データ量をOne - Hot表現により表したベクトルを並べた情報である、

請求項4に記載のトラヒック制御装置。

【請求項7】

前記画像データは、深度画像データである、

請求項2から請求項6のいずれか一項に記載のトラヒック制御装置。

【請求項8】

第1通信装置と1台以上の第2通信装置との間の通信環境を撮像した画像データと前記第2通信装置宛ての未送信のデータのデータ量の情報とを用いて、前記第2通信装置それぞれのトラヒックの組み合わせとして表される行動の価値を算出する価値関数により複数種類の行動それぞれの価値を算出し、算出した前記価値に基づいて行動を決定する行動決定ステップと、

前記行動決定ステップにおいて決定された前記行動が表す前記第2通信装置それぞれのトラヒックに従って、前記第1通信装置から前記第2通信装置宛ての前記データが配信されるよう通信を制御する通信制御ステップと、

前記通信制御ステップによる制御が行われたことによる前記第2通信装置の通信状況を取得し、取得した前記通信状況が過去の通信状況から向上した程度を表す報酬を計算する報酬計算ステップと、

前記報酬計算ステップにおいて計算された報酬に基づいて前記価値関数を更新する学習ステップと、

を有するトラヒック制御方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、通信システム、トラヒック制御装置及びトラヒック制御方法に関する。

【背景技術】

【0002】

大容量かつ高速通信を実現できる次世代無線通信技術として、ミリ波通信に期待が集まっている（例えば、非特許文献1参照）。ミリ波通信の利点の一つは利用可能な周波数幅が広帯域な点であり、1 G b i t / s（ギガビット毎秒）を超える高速通信が可能である。その一方で、ミリ波は水分や酸素による減衰が大きく、見通し通信路が人体等で遮蔽されると通信品質が急峻に低下するという欠点がある（例えば、非特許文献2参照）。この遮蔽による急峻な通信品質低下問題に対処するため、遮蔽された通信路の流量やトラヒックの経路を制御する装置が必要となる。具体的には、図9のようにA P（Access Point：アクセスポイント）が複数のS T A（Station；無線局）とミリ波で通信している環境の無線通信システムにおいては、A PとS T Aとの見通し通信路を人体が遮蔽しうる状況であり、このような状況のA Pの無線帯域を有効利用するための制御装置が必要となる。以下では、N台（Nは1以上の整数）のS T Aを、S T A - 1 ~ S T A - Nとも記載する。

【0003】

ミリ波通信における通信制御問題の解決手法として、R G B - Dカメラを用いた人体遮蔽予測に基づくトラヒック制御装置が提案されている（例えば、非特許文献3参照）。従来技術では、R G B - Dカメラから得られた画像・動画データを用いて人体を検知し、その移動先を予測する。その移動先への移動によって人体がA PとS T Aとの見通し通信路を遮蔽する場合、遮蔽が起こる直前にA PとS T A間のトラヒックを停止し、遮蔽されていない通信路のトラヒックを優先して送信する。この制御によって、制御しない場合と比べて、A Pにおける合計スループットを増加できる。つまり、無線帯域を有効利用するためのトラヒック制御が可能となる。また、遮蔽を予測し、遮蔽が起こる直前にプロアクティブに制御をかけるため、スループットが低下してから制御をかける従来のリアクティブな制御方式と比較して、合計スループットを増加できる。

【0004】

10

20

30

40

50

図10は、非特許文献3の技術を適用したトラフィック制御装置の機能ブロック図である。同図では、APとSTA-1～STA-Nとが無線通信する無線通信システムのプロキシサーバに、トラフィック制御装置が搭載されている。トラフィック制御装置は、画像解析部と、遮蔽判定部と、通信制御部とを備える。トラフィック制御装置を稼働させる際には、初期設定として遮蔽判定部に通信路を設定しておく。画像解析部は、RGB-Dカメラから得られた画像を用いて、ミリ波通信における人体（障害物）の位置推定を行う。次に、遮蔽判定部は、推定された人体の位置とその移動速度から、予め設定した見通し通信路が人体によって遮蔽されるか否かを判定し、遮蔽されると判定した場合にはそのタイミングを推定する。

#### 【0005】

通信制御部は、遮蔽判定部が推定した見通し通信路の遮蔽状況に基づいて、遮蔽が起こると推定された時間にそのトラフィックを停止するようにトラフィックの流量を制御する。具体的には、通信制御部は、インターネットから受信した、見通し通信路が遮蔽されるSTA宛のパケットの送信を停止する。また、通信制御部は、遮蔽が解除されると推定された時間に、STA宛のパケットの送信を再開する。このトラフィック制御によって、APは、あるSTAとの通信において人体遮蔽に伴いスループットが低下するときにおいても、別のSTAとの通信にリソースを割り当てることができる。よって、トラフィック制御を行わない場合と比較して、APにおける合計スループットを増加できる。

#### 【先行技術文献】

#### 【非特許文献】

#### 【0006】

【非特許文献1】P. Wang, Y. Li, L. Song, and B. Vucetic, "Multi-gigabit millimeter wave wireless communications for 5G: From fixed access to cellular networks," IEEE Communications Magazine, 2015年1月, vol.53, no.1, p.168 - 178

【非特許文献2】S. Collonge, G. Zaharia, and G.E. Zein, "Influence of the human activity on wide-band characteristics of the 60 GHz indoor radio channel," IEEE Transactions on Wireless Communications, 2004年11月, vol.3, no.6, p.2396 - 2406

【非特許文献3】T. Nishio, R. Arai, K. Yamamoto, and M. Morikura, "Proactive traffic control based on human blockage prediction using RGBD cameras for millimeter-wave communications," Proc. 2015 IEEE Consumer Communications and Networking Conference(CCNC), Las Vegas, Nevada, USA, 2015年1月, p.152 - 153

#### 【発明の概要】

#### 【発明が解決しようとする課題】

#### 【0007】

非特許文献3の技術では、見通し通信路が遮蔽されそうなときにその見通し通信路を使用するSTAとの通信を遮断し、別のSTAとの通信にリソースを割り当てるといったルールベースの制御を行っている。この方式では、環境に合わせてルールを手で作る必要がある。例えば、見通し通信路の遮蔽が通信品質に影響しないような環境（反射による通信路ができるような環境）では、その見通し通信路が遮蔽された場合でも通信を停止する必要がない。しかし、ミリ波通信環境は、ミリ波基地局や家具の配置によって容易に変化するため、その度に設定し直す必要がある。

#### 【0008】

加えて、人手では適切なルールの設計が難しい環境、例えば、遮蔽する歩行者が多数存在し到来にムラがあるような場合や、動画や音声通話などアプリケーションが異なる場合などでは、適切なトラフィック制御方策は変わることが考えられる。しかしながら、適切な制御方策を決めるのは容易ではない。

#### 【0009】

さらには、画像から人体認識、移動予測、見通し通信路遮蔽予測など、様々な処理を行う必要がある。それらの性能は、通信制御の性能に強く影響を与える。

#### 【0010】

10

20

30

40

50

上記事情に鑑み、本発明は、移動する障害物により無線通信のための見通し通信路に一時的に遮蔽が生じる環境下における合計スループットを増加させることができる通信システム、トラフィック制御装置及びトラフィック制御方法を提供することを目的としている。

【課題を解決するための手段】

【0011】

本発明の一態様は、第1通信装置と、前記第1通信装置と無線により通信する1台以上の第2通信装置と、第1通信装置から前記第2通信装置に送信するデータを取得する第3通信装置と、トラフィック制御装置とを有する通信システムであって、前記トラフィック制御装置は、前記第1通信装置と前記第2通信装置との間の通信環境を撮像した画像データと、前記第3通信装置が記憶する前記第2通信装置宛ての未送信の前記データのデータ量の情報とを用いて、前記第2通信装置それぞれのトラフィックの組み合わせにより表される行動の価値を算出する価値関数により複数種類の行動それぞれの価値を算出し、算出した前記価値に基づいて行動を決定する行動決定部と、前記行動決定部が決定した前記行動が表す前記第2通信装置それぞれのトラフィックに従って、前記第2通信装置宛ての前記データを前記第1通信装置に送信するよう前記第3通信装置を制御する通信制御部と、前記通信制御部による制御が行われたことによる前記第2通信装置の通信状況を取得し、取得した前記通信状況が過去の通信状況から向上した程度を表す報酬を計算する報酬計算部と、前記報酬計算部が計算した報酬に基づいて前記価値関数を更新する学習部と、を備え、前記第1通信装置は、前記第3通信装置から受信した前記第2通信装置宛ての前記データを無線により前記第2通信装置へ送信する、通信システムである。

10

20

【0012】

本発明の一態様は、第1通信装置と1台以上の第2通信装置との間の通信環境を撮像した画像データと前記第2通信装置宛ての未送信のデータのデータ量の情報とを用いて、前記第2通信装置それぞれのトラフィックの組み合わせとして表される行動の価値を算出する価値関数により複数種類の行動それぞれの価値を算出し、算出した前記価値に基づいて行動を決定する行動決定部と、前記行動決定部が決定した前記行動が表す前記第2通信装置それぞれのトラフィックに従って、前記第1通信装置から前記第2通信装置宛ての前記データが配信されるよう通信を制御する通信制御部と、前記通信制御部による制御が行われたことによる前記第2通信装置の通信状況を取得し、取得した前記通信状況が過去の通信状況から向上した程度を表す報酬を計算する報酬計算部と、前記報酬計算部が計算した報酬に基づいて前記価値関数を更新する学習部と、を備えるトラフィック制御装置である。

30

【0013】

本発明の一態様は、上述のトラフィック制御装置であって、前記第2通信装置の前記通信状況は、前記第2通信装置におけるスループット、又は、前記第2通信装置宛ての前記データの送信にかかった時間を表す情報である。

【0014】

本発明の一態様は、上述のトラフィック制御装置であって、前記価値関数は、深層ニューラルネットワークにより近似される。

【0015】

本発明の一態様は、上述のトラフィック制御装置であって、前記価値関数に用いられる前記画像データは、異なるタイミングにおいて撮影された複数の画像データそれぞれの解像度を低減したのちにピクセル値を正規化したデータである。

40

【0016】

本発明の一態様は、上述のトラフィック制御装置であって、前記価値関数に用いられる未送信の前記第2通信装置宛てのデータ量の情報は、複数の前記第2通信装置それぞれ宛ての未送信の前記データ量をOne-Hot表現により表したベクトルを並べた情報である。

【0017】

本発明の一態様は、上述のトラフィック制御装置であって、前記画像データは、深度画像データである。

50

## 【0018】

本発明の一態様は、第1通信装置と1台以上の第2通信装置との間の通信環境を撮像した画像データと前記第2通信装置宛ての未送信のデータのデータ量の情報とを用いて、前記第2通信装置それぞれのトラヒックの組み合わせとして表される行動の価値を算出する価値関数により複数種類の行動それぞれの価値を算出し、算出した前記価値に基づいて行動を決定する行動決定ステップと、前記行動決定ステップにおいて決定された前記行動が表示前記第2通信装置それぞれのトラヒックに従って、前記第1通信装置から前記第2通信装置宛ての前記データが配信されるよう通信を制御する通信制御ステップと、前記通信制御ステップによる制御が行われたことによる前記第2通信装置の通信状況を取得し、取得した前記通信状況が過去の通信状況から向上した程度を表す報酬を計算する報酬計算ステップと、前記報酬計算ステップにおいて計算された報酬に基づいて前記価値関数を更新する学習ステップと、を有するトラヒック制御方法である。

10

## 【発明の効果】

## 【0019】

本発明により、移動する障害物により無線通信のための見通し通信路に一時的に遮蔽が生じる環境下における合計スループットを増加させることが可能となる。

## 【図面の簡単な説明】

## 【0020】

【図1】本発明の一実施形態による無線通信システムの構成例を示す図である。

【図2】同実施形態によるトラヒック制御装置の処理の流れを示すフロー図である。

20

【図3】同実施形態によるエピソードを説明するための図である。

【図4】同実施形態によるカメラ画像から入力データへの加工を示す図である。

【図5】同実施形態によるファイル残量情報から入力データへの加工を示す図である。

【図6】同実施形態による行動評価関数の層設計を示す図である。

【図7】同実施形態によるトラヒック制御装置のシミュレーション評価の諸元を示す図である。

【図8】同実施形態によるトラヒック制御装置のシミュレーション評価結果を示す図である。

【図9】制御対象の無線通信システムの構成例を示す図である。

【図10】従来技術によるトラヒック制御装置の機能ブロック図である。

30

## 【発明を実施するための形態】

## 【0021】

以下、図面を参照しながら本発明の実施形態を詳細に説明する。

本実施形態のトラヒック制御装置は、従来の問題点を解決するために、深層強化学習を用いる。本実施形態のトラヒック制御装置は、カメラ画像とトラヒックバッファとを「状態」として用い、その「状態」に適切な制御を試行錯誤により学習的に獲得する。強化学習とは、行動主体であるエージェントが環境に対して試行錯誤をしながら行動し、その行動に対して環境から報酬を与えられることによって、より良い方策を獲得する機械学習の一種である。エージェントは、「状態」から期待される報酬を表す価値関数に従って行動し、得られた報酬によってこの価値関数を更新する。深層強化学習では、この価値関数に畳み込みニューラルネットワーク(CNN; Convolutional Neural Network)などのニューラルネットワークを用いて関数近似をする。これによって、状態数が膨大な問題に適用できることに加え、畳み込み層を用いることで画像を入力とするような問題に対して効果を発揮する。

40

## 【0022】

図1は、本発明の一実施形態による通信システム1を示す図である。通信システム1は、アクセスポイント(AP)2、無線局(STA)3と、プロキシサーバ4と、トラヒック制御装置5と、撮像装置6とを備える。N台(Nは1以上の整数)のSTA3のうちn台目(nは1以上N以下の整数)のSTA3を、STA-nと記載する。また、同図において、トラヒック制御装置5は、プロキシサーバ4に搭載される。同図に示す通信システ

50

ム 1 は、図 10 に示す従来のトラヒック制御装置を、トラヒック制御装置 5 に置き換えた構成である。

【0023】

A P 2 は、1 台以上の S T A 3 と無線通信する。A P 2 は、インターネット 7 を介して接続される通信装置からプロキシサーバ 4 が受信した S T A 3 宛ての packets を無線により送信する。また、A P 2 は、インターネット 7 を介して接続される通信装置宛ての packets を S T A 3 から無線により受信し、プロキシサーバ 4 に送信する。プロキシサーバ 4 は、S T A 3 の代理としてインターネット 7 を介した通信を行う。撮像装置 6 は、例えば、R G B - D カメラである。R G B - D カメラは、R G B 画像（カラー画像）と深度画像とを撮像する。撮像装置 6 は、A P 2 と複数の S T A 3 との間の無線の見通し通信路と及びその周辺を含んだ環境の画像を所定周期で撮像する。撮像装置 6 は、撮像した画像のデータであるカメラ画像をトラヒック制御装置 5 に送信する。

10

【0024】

プロキシサーバ 4 は、第 1 通信部 4 1 と、記憶部 4 2 と、第 2 通信部 4 3 と、トラヒック制御装置 5 とを備える。第 1 通信部 4 1 は、インターネット 7 を介して受信した S T A 3 宛てのファイルの packets を受信し、S T A 3 別に記憶部 4 2 に書き込む。記憶部 4 2 は、複数のファイルバッファを有している。S T A 3 に割り当てられたファイルバッファに、当該 S T A 3 宛てのファイルが記憶される。1 台の S T A 3 に対して複数のファイルバッファを割り当てることができる。1 台の S T A 3 に対して割り当て可能なファイルバッファに上限を設けてもよい。本実施形態では、1 台の S T A 3 に 3 つのファイルバッファを割り当て可能とする。第 2 通信部 4 3 は、トラヒック制御装置 5 の制御に従って、S T A 3 宛てのファイルを記憶部 4 2 から読み出して A P 2 に送信する。

20

【0025】

トラヒック制御装置 5 は、強化学習部 5 1 と、報酬計算部 5 2 と、通信制御部 5 3 とを備える。強化学習部 5 1 は、加工部 5 1 1 と、行動決定部 5 1 2 と、学習部 5 1 3 とを備える。行動決定部 5 1 2 及び学習部 5 1 3 は、深層強化学習アルゴリズムの処理部である。加工部 5 1 1 は、撮像装置 6 から入力されたカメラ画像と、トラヒックバッファ情報とを処理に適したデータ形式に加工し、深層強化学習アルゴリズムの処理部に出力する。行動決定部 5 1 2 は、データ形式が加工されたカメラ画像とトラヒックバッファ情報とを含む「状態」に基づいて、トラヒックの制御信号を「行動」として出力する。トラヒックバッファ情報とは、プロキシサーバ 4 に蓄積されている各 S T A 3 宛ての未送信のデータのデータ量である。本実施形態では、トラヒックバッファ情報として、ファイル残量が用いられる。ファイル残量は、記憶部 4 2 に記憶される未送信の各 S T A 3 宛てのファイルの容量である。学習部 5 1 3 は、出力した「行動」について報酬計算部 5 2 が計算した報酬に基づいて、より良い制御方法を学習する。

30

【0026】

報酬計算部 5 2 は、各 S T A 3 のスループット及びトラヒックバッファ情報、あるいは、それらの一部から、目的に合わせて設計された報酬を出力する。通信制御部 5 3 は、A P 2 と各 S T A 3 とのトラヒックをスケジューリングしながら S T A 3 宛てのファイルを配信するようプロキシサーバ 4 の第 2 通信部を制御する。これは、ミリ波通信においては、その高速通信という利点を活かし、容量の大きいファイルを送信するという実用例が想定されるためである。

40

【0027】

なお、トラヒック制御装置 5 が、プロキシサーバ 4 の第 1 通信部 4 1 と、記憶部 4 2 と、第 2 通信部 4 3 とのうち任意の一以上の機能部を有してもよい。また、第 1 通信部 4 1 と通信制御部 5 3 とが同一の機能部でもよく、第 2 通信部 4 3 と通信制御部 5 3 とが同一の機能部でもよく、第 1 通信部 4 1、第 2 通信部 4 3 及び通信制御部 5 3 が同一の機能部でもよい。また、トラヒック制御装置 5 は、プロキシサーバ 4 と通信ネットワークにより接続される外部の装置であってもよい。また、第 1 通信部 4 1 と、記憶部 4 2 と、第 2 通信部 4 3 と、強化学習部 5 1 と、報酬計算部 5 2 と、通信制御部 5 3 とのうち一以上の任

50

意の機能部を、プロキシサーバ4及びトラヒック制御装置5とが協働して実現してもよい。

【0028】

図2は、トラヒック制御装置5の処理の流れを示すフロー図である。

トラヒック制御装置5が起動すると、撮像装置6は、一定時間間隔で通信環境を撮影してカメラ画像を生成し、強化学習部51へ送信する(ステップS1)。一方で、通信制御部53は、各STA3のファイルバッファ内のファイル残量を取得し、強化学習部51へ送信する(ステップS2)。加工部511は、撮像装置6及び通信制御部53のそれぞれから受信したデータを深層強化学習の設計に合わせて事前処理した後、行動決定部512に入力する(ステップS3)。

10

【0029】

深層強化学習では価値関数にニューラルネットワークを用いるため、加工部511は、カメラ画像とファイル残量情報を、設計されたニューラルネットワークに適した入力データに加工する。この価値関数のニューラルネットワークの例として、全結合層のみの単純なものや、画像認識の分野でよく用いられる畳み込み層を含んだものが挙げられる。例として、価値関数が全結合層のみのニューラルネットワークの場合、加工部511は、カメラ画像のうち深度画像の解像度を低くした後に1次元のデータにして、各深度値を0から1までの値に正規化する。また、加工部511は、各STA3のファイルバッファに残っているファイルの容量を離散化してOne-Hot表現化したファイル残量情報を生成し、入力データとする。One-Hot表現とは、ある要素のみが1であり、それ以外の要素が0となるベクトル表現のことである。ファイル容量を表すベクトルの各要素はそれぞれファイル容量の範囲に対応しており、ファイルバッファに残っているファイル容量に対応した要素に1が設定され、他の要素には0が設定される。

20

【0030】

行動決定部512は、深層強化学習アルゴリズムを用いて、価値関数の出力結果に基づいて各STA3の通信のトラヒック(強化学習の「行動」)を決定する(ステップS4)。具体的には、行動決定部512は、カメラ画像とファイルバッファのファイル残量情報という「状態」において、とりうる「行動」のうち、それら各「行動」によって最も価値が高くなるような状態遷移を起こす「行動」(各STA3のトラフィック)を優先的に採用する。行動決定部512は、決定した各STA3の通信のトラヒック制御情報を通信制御部53に送信する。これを受信した通信制御部53は、そのトラヒック制御情報に従って、ファイルバッファに保持していたファイルをパケットに設定してAP2へ送信するようプロキシサーバ4の第2通信部43を制御する(ステップS5)。

30

【0031】

パケット送信後、通信制御部53は、各STA3宛てのバッファ内のファイル残量とその時点での各STA3のスループットを取得し、報酬計算部52へ送信する(ステップS6)。報酬計算部52は、受信したファイル残量及びスループット情報を用いて報酬を計算する(ステップS7)。報酬は、トラヒック制御の詳細な目的に合わせて設計される。詳細な目的の例としては、AP2の合計スループットの最大化、ファイル送信時間の合計の最小化等が挙げられる。AP2の合計スループットの最大化が目的の場合、報酬計算部52は、行動決定部512が行動を決定し、その決定に基づいて通信制御部53が行動する度に毎回、その時点でのAP2の合計スループットを報酬として与える。ファイル送信時間の合計の最小化が目的の場合、報酬計算部52は、行動決定部512が行動を決定し、その決定に基づいて通信制御部53が行動する度に毎回、ファイルがプロキシサーバ4に到着してからSTA3へファイルの送信を完了するまでの間、負の定数を報酬として与える。つまり、報酬の累積和が、ファイル送信時間の合計に比例した値になる。

40

【0032】

例えば、AP2の合計スループットの最大化が目的の場合、時間ステップ $t$ における報酬 $r_t$ は、以下の式(1)のように算出される。

【0033】

50



【数 1】

$$r_t = T_t - \sum_{i=0}^t c(i)T_i \quad \dots (1)$$

【0034】

$T_t$  は時間ステップ  $t$  における合計スループット、 $c(t)$  は時間パラメータ  $t$  に応じた係数である。 の項はこれまでの合計スループットを時間等のパラメータにより加重平均した値である。例えば、各  $c(i)$  を、式(1)の第2項において時間に応じた荷重平均スループットが得られるように決定してもよい。また、 $c(i) = 1$  ( $i$  は  $t$  以下の整数) とすると、報酬  $r_t$  は、以下の式(2)により算出される。

【0035】

【数 2】

$$r_t = T_t - \frac{1}{t+1} \sum_{i=0}^t T_i \quad \dots (2)$$

【0036】

また、報酬を、式(3)に示すように AP2 全体の平均のスループット  $T_t$  で正規化したスループットとしてもよく、式(4)に示すように、正規化したスループットの差分としてもよい。

【0037】

【数 3】

$$r_t = T_t / \bar{T}_t \quad \dots (3)$$

【0038】

【数 4】

$$r_t = (T_t - \bar{T}_t) / \bar{T}_t \quad \dots (4)$$

【0039】

また、以下の式(5)のように、スループットの平均からの減衰率が一定値 を下回ったときに大きな負の報酬を与えるようにしてもよい。

【0040】

10

20

30

40

【数5】

$$r_t = \begin{cases} 1, & \text{if } T_t/\bar{T}_t > \alpha \\ -100, & \text{otherwise} \end{cases} \quad \dots(5)$$

【0041】

10

また、式(1)～式(5)におけるスループットを、ミリ波通信の物理伝送速度に置き換えてもよい。

【0042】

報酬計算部52は、計算した報酬を強化学習部51に送信する。強化学習部51は、通知された報酬に基づいて、深層強化学習アルゴリズムによって価値関数を更新していくことで学習を進める(ステップS8)。

【0043】

この一連の動作を繰り返すことにより、強化学習部51は、入力された報酬の累積和が最大となるように学習を進めながら各STA3のトラヒックのトラヒックを決定していく。従って、学習が進むに連れてトラヒック制御装置5を設置した環境に適応したトラヒック制御方法を自動的に獲得する。

20

【0044】

トラヒック制御装置5は、複数エピソードを実施した結果に基づいて、上記の処理を行い、行動評価関数を学習する。図3は、エピソードを説明するための図である。エピソードとは、記憶部42におけるファイルバッファ内のファイルが全て送信完了するまでの一連の流れを表す。プロキシサーバ4は、トラヒック制御装置5の通信制御部53の制御に従って、ファイルバッファに記憶されるファイルを、AP2を介して各STA3へ送信していき、ファイルバッファ内のファイルを全て送信し終えた時点で1エピソードの終了とする。1エピソードの途中ではファイルは追加されない。エピソードが進むに連れて、本実施形態のトラヒック制御装置5の学習も進む。なお、学習する上限数をあらかじめ決めておき、エピソードが上限数に達した場合には学習を終了してもよい。

30

【0045】

価値関数として用いられる深層ニューラルネットワーク(CNN)の入力データ及び層設計の例を説明する。

図4は、ステップS3におけるカメラ画像から入力データへの加工を示す図である。強化学習部51は、1秒間における過去5枚分のカメラ画像に含まれる深度画像データをそれぞれ20×20ピクセルの二次元画像データに圧縮する。強化学習部51は、5枚の深度画像データそれぞれを圧縮して得られた5チャンネルの二次元画像をCNNへの入力データとする。

【0046】

40

図5は、ステップS3におけるファイル残量情報から入力データへの加工を示す図である。まず、各ファイルの残量を複数段階に離散化する。ここでは、ファイル容量の最大値が2000Mbit(メガビット)であり、10段階に離散化する場合を例とする。この場合、ファイル残量情報として用いられるOne-Hot表現のベクトルの各要素を、[(0-200Mbit), (200-400Mbit), (400-600Mbit), (600-800Mbit), ..., (1800-2000bit)]と定める。記憶部42から取得したSTA-n(nは1以上N以下の整数)のファイル残量が容量700Mbitである場合、ファイル残量情報はベクトル[0, 0, 0, 1, 0, 0, 0, 0, 0, 0]と表される。強化学習部51、STA-1、STA-2、...、STA-Nについて生成したファイル残量情報を表すベクトルを並べて結合し、入力データとする。

50

## 【0047】

図6は、CNNの層設計を示す図である。なお、「Affine, a - b」は、a次元ベクトルを全結合層に入力し、b次元ベクトルを出力する演算を表す。「k x l 2D Conversion, a - b」は、k x lの二次元フィルタにより、aチャンネルの入力を畳み込み、bチャンネルにして出力する演算を表す。また、「k x l 2D Max Pooling」は、サイズがk x lのグリッドに入力を分割し、各グリッドの最大値を代表値として出力する演算を表す。「ReLU」は、活性化関数ReLU (Rectified Linear Units)に入力する演算を表す。活性化関数ReLUは、マイナスの値を0に変換する。

## 【0048】

入力層では、図3に示した処理により5チャンネルの二次元画像(5 Channels 2D Image)を生成する。さらに、入力層では、図4に示した処理により各STA3のファイル残量をOne - Hot表現のベクトルに変換し、結合して60次元ベクトルを生成する。

## 【0049】

隠れ層には、1a層~8a層と、1b層~2b層と、8a層及び2b層の出力を入力とする9層とがある。

1a層では、5チャンネルの二次元画像(5 Channels 2D Image)を、5 x 5の二次元フィルタにより畳み込み、20チャンネルにして出力する。2a層では、20チャンネルの1a層の出力を活性化関数ReLUに入力し、マイナスの値を取り除く。3a層では、20チャンネルの2a層の出力を2 x 2のグリッドに分割し、各グリッドの最大値を出力する。4a層では、20チャンネルの3a層の出力を、5 x 5の二次元フィルタにより畳み込み、50チャンネルにして出力する。5a層では、50チャンネルの4a層の出力を活性化関数ReLUに入力し、マイナスの値を取り除く。6a層では、50チャンネルの5a層の出力を2 x 2のグリッドに分割し、各グリッドの最大値を出力する。7a層では、6a層の1250次元ベクトルを全結合層に入力し、500次元ベクトルを出力する。8a層では、7a層の出力を活性化関数ReLUに入力し、マイナスの値を取り除く。

## 【0050】

一方、1b層では、各STA3のファイル残量に基づいて得られた60次元ベクトルを全結合層に入力し、100次元ベクトルを出力する。なお、STA3の台数Nと、One - Hot表現のベクトルの要素数との乗算が60であるとする。2b層では、1b層の出力を活性化関数ReLUに入力し、マイナスの値を取り除く。

## 【0051】

9層では、8a層の出力及び2b層の出力を併せた600次元ベクトルを全結合層に入力し、各行動の評価値を得る。出力層は、各行動の評価値を出力する。各行動は、各STA3との通信をONにするかOFFにするかの組み合わせでもよく、N台のSTA3それぞれのトラフィック量の組み合わせでもよい。同図では、2台のSTA3それぞれとの通信をONにするかOFFにするかの組み合わせから、2台ともOFFの組み合わせを除いたものである。つまり、(STA - 1, STA - 2)を(OFF, OFF)、(ON, OFF)、(OFF, ON)とする3種類の行動である。この3種類の行動それぞれの評価値を得るため、9層からは3次元ベクトルが出力される。

## 【0052】

なお、Conversion層については、入力層に近いところにおいては画像から特徴量抽出するフィルタが学習されることが期待され、出力層に近いところでは特徴量から値を予測するフィルタが学習されることが期待される。ReLUは、活性化関数として広く用いられる。ReLUは、他の活性化関数(シグモイド関数など)とくらべて、経験的に学習速度が早く、性能が高くなることが知られている。また、Max Pooling層は、Conversion層を通すことにより増大したパラメータ数を削減することで学習時間を短縮するために使用される。Affine層は、CNNにより抽出された特徴量から値を予測することを期待して使用される。CNNのみで構成するような層設計と比較して、学習時間の短縮が期待できることが経験的に知られている。

10

20

30

40

50

## 【0053】

学習部513は、価値関数として用いられるCNNを更新する。具体的には、学習部513は、報酬計算部52により計算される報酬に基づいて、全結合層における重みを更新する。例えば、行動決定部512において、AP2とSTA-1の通信ON、AP2とSTA-2の通信OFFという結果が得られた場合、通信制御部53は、AP2とSTA-1との通信のみをONにするよう制御を行う。例えば、通信制御部53は、STA-1宛てのファイルをAP2に出力し、STA-2宛てのファイルをAP2に出力しないようにプロキシサーバ4の第2通信部43を制御する。あるいは、プロキシサーバ4の第2通信部43を介して、AP2に対してSTA-1との通信を行い、STA-2との通信を行わないよう制御信号を送信してもよい。しかしながら、このような制御を行っても、AP2とSTA-1間で遮蔽が発生している、マルチパスで反射が発生しているなど、実際はAP2とSTA-1間の伝搬路の状態が悪い場合、通信速度は低くなる。極端な例として、AP2とSTA-1間に金属の壁があり、STA-1にまったく電波が届かない場合は、通信がONの状態でもスループットは0Mbit/sとなる。学習部513は、そのようなことが発生しないように、各STA3のON/OFFを制御するための学習を行うことができる。

10

## 【0054】

本実施形態のトラフィック制御装置5によれば、カメラ画像を入力とした深層強化学習によりトラフィック制御を行い、様々な通信環境に自動的に適応して無線帯域を有効利用することが可能となる。また、通信端末やカメラの設置環境が変化した際にも、変化した環境に適応して自動的にトラフィックを制御することが可能となる。特に、ミリ波通信機能を搭載した無線LAN(Local Area Network)ルータと、複数のミリ波通信端末とが接続された通信システムにおいて、人体遮蔽が起こりうる状況に有用である。また、無線LANルータやミリ波通信端末の設置環境が変化する場合にも対応可能である。

20

## 【0055】

トラフィック制御装置5の実測データを用いたシミュレーション評価について述べる。図7は、シミュレーション評価の諸元を示す図である。このシミュレーション評価では、1台のAP2に、2台のSTA3を接続した場合を想定し、本実施形態のトラフィック制御を行った場合と、ファイル送信完了ごとに交互に送信宛先を切り替えるラウンドロビン方式で制御を行った場合のAPにおける合計スループットを得た。AP2は、ミリ波APである。シミュレーションで用いるミリ波通信の見通し通信時、遮蔽時のスループット及びカメラ画像は実機実験から測定した値を用いた。カメラ画像は、RGB-Dカメラで撮影した画像のデータを用いた。また、AP2及びSTA3も市販のものを用いた。

30

## 【0056】

図8は、シミュレーション評価結果を示す図である。同図は、本実施形態のトラフィック制御を行った場合とラウンドロビン方式で制御を行った場合のエピソード数に対する合計スループットの推移を示す。同図のグラフにおけるAP2の合計スループットとして、各エピソードにおけるAP2の合計スループットの時間平均として表示している。このシミュレーションでは、プロキシサーバ4のファイルバッファには最初、ファイルがランダムなサイズで与えられ、AP2を通して各STA3へファイルを送信していく。ファイルバッファ内のファイルを全て送信し終えた時点で1エピソードが終了する。同図に示す評価結果から、エピソードが進み、トラフィック制御装置5の学習が進むに連れて、ラウンドロビン方式による制御を行った場合のスループットよりも、本実施形態のトラフィック制御を行った場合の合計スループットが上回っていることがわかる。

40

## 【0057】

以上説明した実施形態によれば、通信システムは、第1通信装置と、第1通信装置と無線により通信する1台以上の第2通信装置と、第1通信装置から第2通信装置に送信するデータを取得する第3通信装置と、トラフィック制御装置とを有する。例えば、第1通信装置はAP2であり、第2通信装置はSTA3であり、第3通信装置はプロキシサーバ4である。

50

## 【 0 0 5 8 】

トラフィック制御装置は、行動決定部と、通信制御部と、報酬計算部と、学習部とを有する。行動決定部は、第1通信装置と第2通信装置との間の通信環境を撮像した画像データと、第3通信装置が記憶する第2通信装置宛ての未送信のデータのデータ量の情報とを用いて、第2通信装置それぞれのトラフィックの組み合わせにより表される行動の価値を算出する価値関数により、複数種類の行動それぞれの価値を算出する。価値関数は、深層ニューラルネットワークにより近似されてもよい。この場合、深層ニューラルネットワークに入力される画像データは、異なるタイミングにおいて撮影された複数の画像データそれぞれの解像度を低減したのちにピクセル値を正規化したデータである。また、深層ニューラルネットワークに入力される未送信の第2通信装置宛てのデータ量の情報は、複数の第2通信装置それぞれ宛ての未送信のデータ量を One - Hot 表現により表したベクトルを並べた情報である。行動決定部は、算出した価値に基づいて行動を決定する。

10

## 【 0 0 5 9 】

通信制御部は、行動決定部が決定した行動が表す第2通信装置それぞれのトラフィックに従って、第2通信装置宛てのデータを第1通信装置に送信するよう第3通信装置を制御する。報酬計算部は、通信制御部による制御が行われたことによる第2通信装置の通信状況を取得し、取得した通信状況が過去の通信状況から向上した程度を表す報酬を計算する。第2通信装置の通信状況は、第2通信装置におけるスループット、又は、第2通信装置宛てのデータの送信にかかった時間を表す。学習部は、計算された報酬に基づいて価値関数を更新する。第1通信装置は、第3通信装置から受信した第2通信装置宛てのデータを無線により第2通信装置に送信する。

20

## 【 0 0 6 0 】

上述した実施形態におけるトラフィック制御装置5の機能をコンピュータで実現するようにしてもよい。その場合、トラフィック制御装置5はこの機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することによって実現してもよい。なお、ここでいう「コンピュータシステム」とは、OSや周辺機器等のハードウェアを含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムを送信する場合の通信線のように、短時間の間、動的にプログラムを保持するもの、その場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含んでもよい。また上記プログラムは、前述した機能の一部を実現するためのものであってもよく、さらに前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるものであってもよい。

30

## 【 0 0 6 1 】

以上、この発明の実施形態について図面を参照して詳述してきたが、具体的な構成はこの実施形態に限られるものではなく、この発明の要旨を逸脱しない範囲の設計等も含まれる。

40

## 【 産業上の利用可能性 】

## 【 0 0 6 2 】

無線通信を行う通信システムに利用可能である。

## 【 符号の説明 】

## 【 0 0 6 3 】

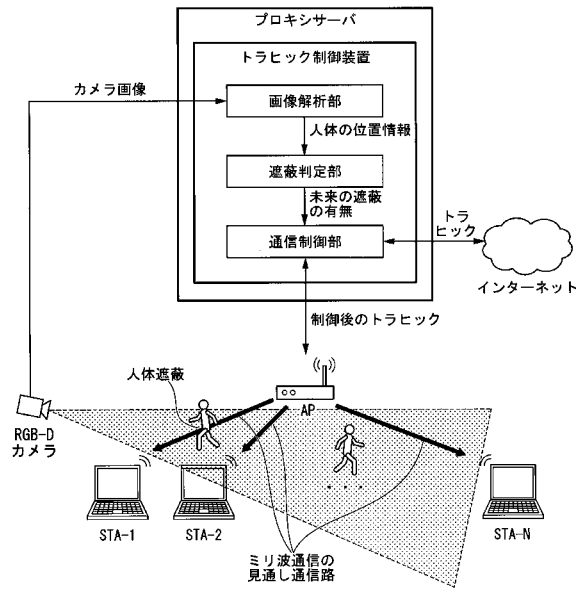
1 ... 通信システム、 2 ... アクセスポイント、 3 ... 無線局、 4 ... プロキシサーバ、  
5 ... トラフィック制御装置、 6 ... 撮像装置、 7 ... インターネット、 4 1 ... 第1通信部、  
4 2 ... 記憶部、 4 3 ... 第2通信部、 5 1 ... 強化学習部、 5 2 ... 報酬計算部、  
5 3 ... 通信制御部、 5 1 1 ... 加工部、 5 1 2 ... 行動決定部、 5 1 3 ... 学習部

50





【図 10】





フロントページの続き

(72)発明者 西尾 理志

京都府京都市左京区吉田本町3番地1 国立大学法人京都大学内

Fターム(参考) 5K067 AA13 EE02 EE10 EE16