

# 生命パスウェイの脆弱部位推定の一考察

北風裕教\* 中司弘樹\*\* 松野浩嗣\*\*\* 宮野悟\*\*\*\*

## Consideration on Prediction of Debacle Points for Robustness of Biological Pathway

Hironori KITAKAZE, Hiroki NAKATSUKA, Hiroshi MATSUNO and Satoru MIYANO

### Abstract

The aim of our study group is developing the software Cell Illustrator, by using hybrid functional Petri net as its basic architecture, which can be used in modeling and simulation of life development. To use this tool to estimate the debacle points of biological pathways, it is necessary to stop the considerable reaction paths by many sequential manual procedures so that it needs a huge amount of time. In order to solve this problem, we have been developing a system to modificate the HFPN into RNN. However, it still produces a low prediction rate. It was because we did not take into account the three kinds of arc of HFPN when modifcating it into RNN. In this paper we report the examination we made after taking into consideration the three arcs which were introduced in the model to reform the RNN.

**Key words:** hybrid functional Petri net, recurrent neural networks, biological pathways

### 1. はじめに

生体は外部異常と内部異常に対して複数の分子機構が互いに補償することで恒常性を保っている。しかし、何か鍵となる分子やシステム異常により恒常性が途端に崩れるという脆弱性も同時に持ち合わせている[1]。我々は、ハイブリッド関数ペトリネット（HFPN）を基本アーキテクチャとして計算機上でシミュレーションできるツール Cell Illustrator(CI)[2,3,4,5]を開発し、脆弱部位の特定に努めてきた。

しかしこのシステムは、生命経路（生命パスウェイ）から脆弱部位の推定を行なう際に、考慮できる経路の反応を順次手作業で停止（ノックアウト）し、その都度実行処理を繰り返す必要があるため膨大な処理時間を要するという問題があり、脆弱部位の自動推定と時間の短縮が望まれている。我々はこれまでに、HFPN で記述された複雑な生命パスウェイを、時間成分を記憶可能なりカレントニューラルネットワーク（RNN）[6]に再形成し、自動で連続的にノックアウトを実現するシステムについて提案してきた[4]。しかしこのシステムは、HFPN に比べて処理時間の向上の面は優れているが、HFPN の結果に対して推定率が約 87%程度であり、実システムとして利用できるまでには至ってはいなかった。その理由として、HFPN では connector の種類がペトリネットで“アーク”を意味する“process”と、“促進アーク”を意味する“association”そして、“抑止アーク”を意味する“inhibitory”の 3 種類あるが全て“process”と仮定して RNN に再形成したため、促進アークと抑止アークについての閾値に対する概念を無視したことによる影響が出ていた可能性があった。RNN の学習により、アークの条件は吸収される可能性もあるため、確認を行う作業が必要となった。

そこで本研究は、RNN の再形成の際に、3 種類のコネクターによる閾値を考慮に入れた改良モデルを提案する。また、サーカディアンモデル[7]を用いて簡易実験を行い、従来法による脆弱部位推定システムを用いた結果との比較を行うことで、3 種類のアークの必要性について考察し、今後の検討を行う。

\*商船学科 \*\*電子・情報システム工学専攻 \*\*\*山口大学  
\*\*\*\*東京大学

## 2. 脆弱経路推定法

### 2.1 伝統的な推定方法 (Cell Illustrator を用いた手作業による手法)

HFPN を用いた伝統的な脆弱部位の推定方法を図 1 に示す。HFPN をコンピュータで実現するために、Cell Illustrator を利用する。まず、対象とする生命パスウェイを、ユーザーの知識に基づき HFPN を用いてモデル化する。次に生物実験でノックアウトを行うように Cell Illustrator のシミュレーション上で connector の連結を削除する。削除した状態でシミュレーション実験を行う。ネットワークが破綻するような脆弱部位であるか、状態が変化しない頑健部位であるかの確認を行う。この作業を全ての経路で繰り返すことで、ネットワーク全体の推定を行う[8]。

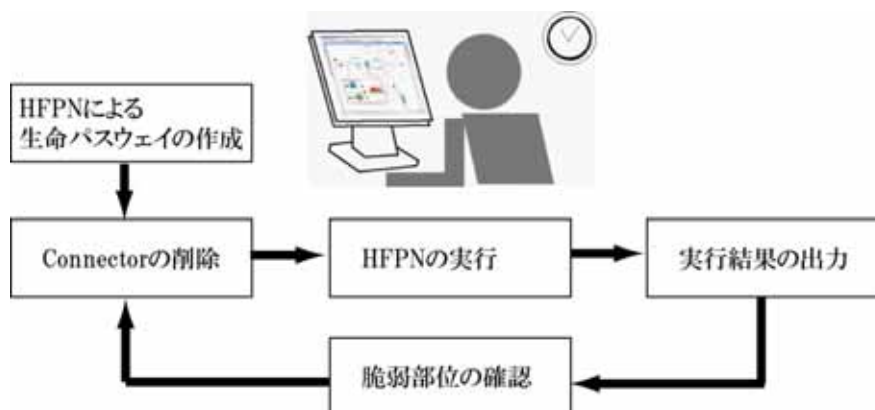


図 1：伝統的な脆弱部位の推定方法

### 2.2 従来のモデル (RNN を用いた自動推定手法)

伝統的な手法はシミュレーション実験を行うことにより、生命の犠牲や費用の問題を解決して脆弱部位の推定が行えるが、膨大になるシミュレーション時間の問題を解決することができなかった。そこで、RNN を用いてシミュレーション時間を短縮する方法を提案してきた(図 2)。この手法は次の手順で行う。まず、HFPN によりモデル化した生命パスウェイを RNN に再形成を行う。再形成には、入力と出力の関係が同じようになるように Back Propagation Through Time (BPTT)法を用いて学習を行う。学習後に HFPN に対応した RNN の結合荷重を連続的にノックアウトすることで脆弱部位を推定する[9][10]。

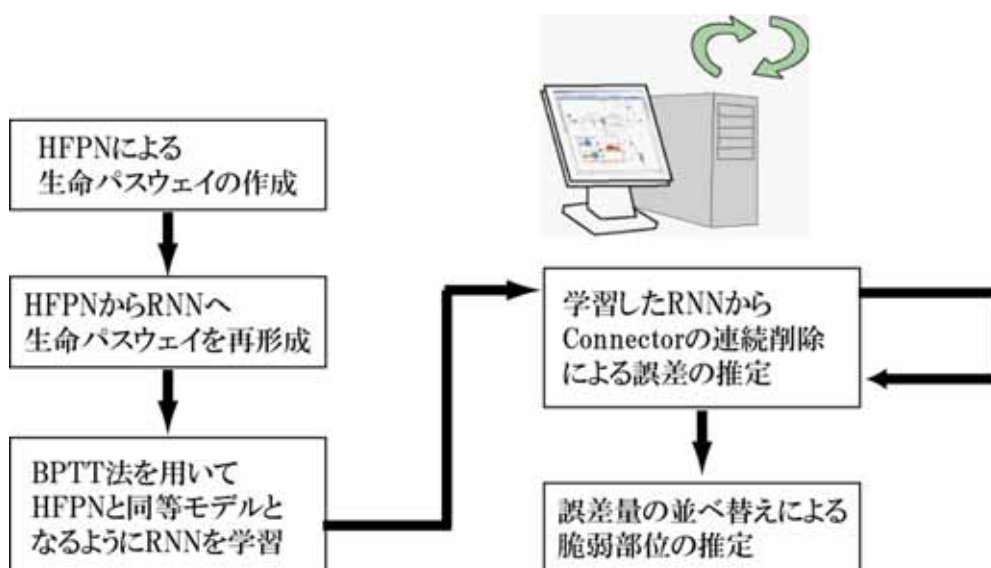


図 2：RNN を用いた自動推定方法

### 2.3 改良モデル (RNN を用いた自動推定手法の改良版)

従来のモデルでは, HFPN の 3 種類ある connector ( アーク:process, 促進アーク:association, 抑止アーク:inhibitory ) について 1 種類のアーク (process) として処理を行った為に, 抑止アークと促進アークが持つ特徴を RNN に反映させることができない問題があり, また, アークの閾値情報についても同様に反映できなかった. 改良モデルでは, 従来のモデルに次の条件を与えコネクタと閾値を設定する. これは HFPN から RNN へ再形成を行った後, BPTT へ時間展開を行う前に行う.

#### 2.3.1 条件 1 : process の場合

- ・HFPN のトークン a の値が閾値 th の値より大きい ( $a > th$ ) 場合, RNN の A C の結合を行う.
- ・HFPN のトークン a の値が閾値 th の値より小さい ( $a < th$ ) 場合, RNN の A C の結合を行わない.

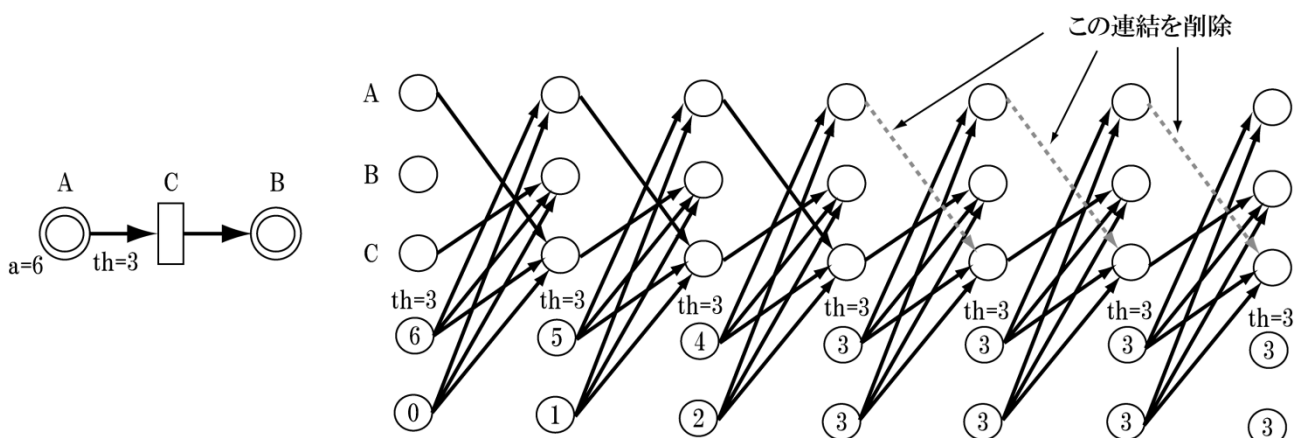


図 3 : process の場合 ( 左 : HFPN , 右 : RNN )

HFPN のアークが”process”で, プレース A のトークンが  $a=6$ , アークの閾値が  $th=3$  の例を考える ( 図 3 左 ). このとき 改良モデルにおける BPTT により時間展開された RNN の結果は図 3 右のようになり, トークン a の値が閾値  $th=3$  より大きい場合は, RNN のユニット A からユニット C は連結させ, a の値が  $th=3$  以下の間は, ユニット A からユニット C の連結は削除される.

#### 2.3.2 条件 2 : association の場合

- ・HFPN のトークン a の値が閾値 th の値より大きい ( $a > th$ ) 場合, RNN の A C の結合を行う.
- ・HFPN のトークン a の値が閾値 th の値より小さい ( $a < th$ ) 場合, RNN の A C の結合を行わない.

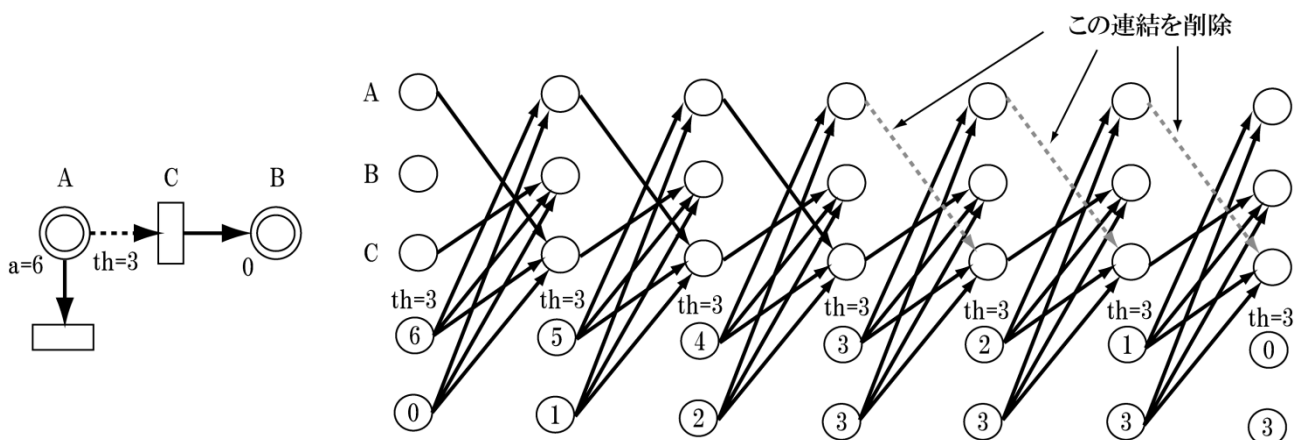


図 4 : association の場合 ( 左 : HFPN , 右 : RNN )

HFPN のアークが”association”で，プレース A のトークンが  $a=6$ ，アークの閾値が  $th=3$  の例を考える（図 4 左）．このとき，改良モデルにおける BPTT により時間展開された RNN の結果は図 4 右のようになり，トークン  $a$  の値が閾値  $th=3$  より大きい場合は，RNN のユニット A からユニット C は連結させ， $a$  の値が  $th=3$  以下の間は，ユニット A からユニット C の連結は削除される．

### 2.3.3 条件 3：inhibitory の場合

- ・HFPN のトークン  $a$  の値が閾値  $th$  の値より大きい ( $a > th$ ) 場合，RNN の A C の結合を行う．
- ・HFPN のトークン  $a$  の値が閾値  $th$  の値より小さい ( $a < th$ ) 場合，RNN の A C の結合を行わない．

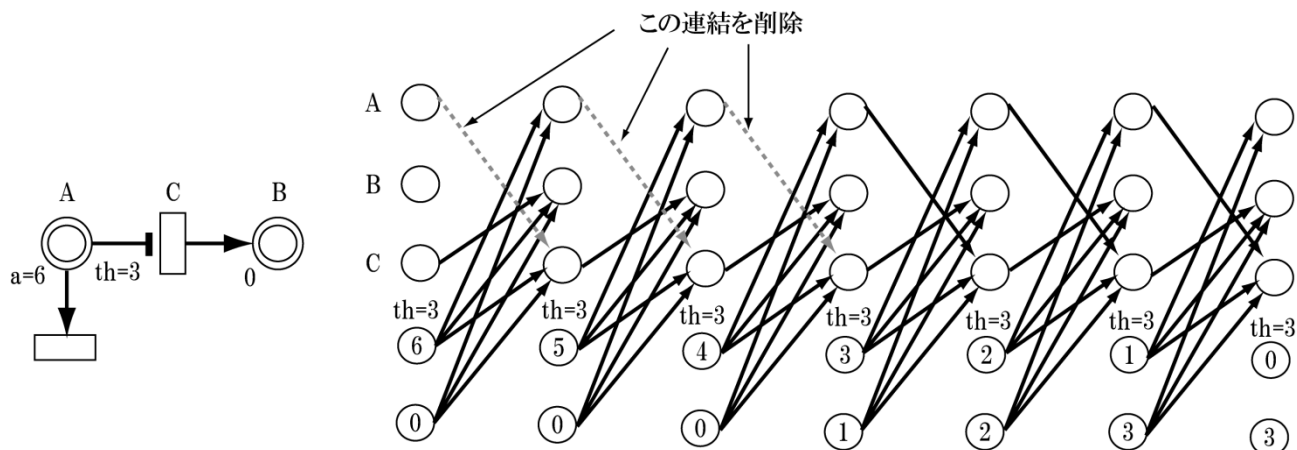


図 5：inhibitory の場合（左：HFPN，右：RNN）

HFPN のアークが”inhibitory”で，プレース A のトークンが  $a=6$ ，アークの閾値が  $th=3$  の例を考える（図 5 左）．このとき，改良モデルにおける BPTT により時間展開された RNN の結果は図 5 右のようになり，トークン  $a$  の値が閾値  $th=3$  より大きい場合は，RNN のユニット A からユニット C の連結は削除させ， $a$  の値が  $th=3$  以下の間は，ユニット A からユニット C は連結される．

## 3. 実験結果と考察

### 3.1 本手法による実験結果

図 6 に今回の実験で利用した HFPN で記述したサーカディアンモデル[2,3]を示す．予め HFPN を用いて従来法によるロックアウト作業で指標を作成した．赤色のアークがロックダウンにより破綻する脆弱部位であり，青色のアークがロックダウンをしても破綻を起こさない頑健な部位である．図 7 左に RNN を用いて脆弱部位を推定した結果を示す．丸で囲まれたアークが誤って推定されたアークである．それ以外のアークに関しては正しく推定することができた．

表 1 上段に指標の結果を示す．脆弱部位が 23 箇所，頑健部位が 18 箇所である．これに対し，提案モデルによる脆弱部位の予測結果を表 1 下段に示す．提案モデルを用いて脆弱部位であると正しく予測できた結果が 87.0%(20 箇所)であり，正しく頑健部位であると予測した結果が 77.8%(14 箇所)であった．HFPN で脆弱部位ではあるが，提案モデルにおいて頑健であると誤った率が 13.0%(3 箇所)あり，HFPN で頑健であるが脆弱部位であると誤った率が 22.2%(4 箇所)となった．

表 1：提案手法による脆弱部位推定結果(1)

		脆弱部位	頑健部位
HFPN を用いて確認した指標		23	18
提案モデルによる脆弱部位予測結果	的中率	約 87.0%(20)	約 77.8%(14)
	誤り率	約 13.0%(3)	約 22.2%(4)

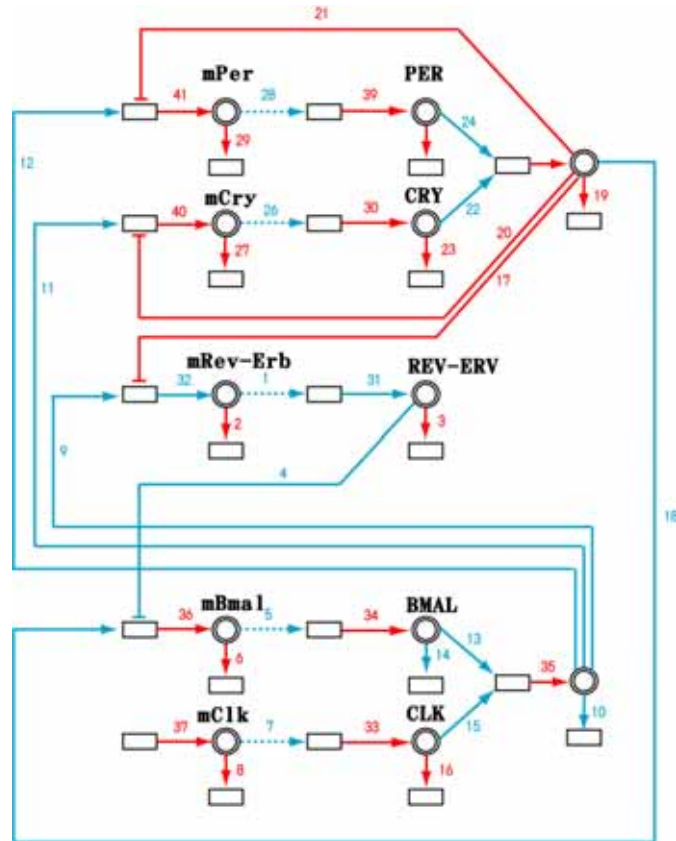


図 6：HFPN を用いて確認したサーカディアンモデルの結果

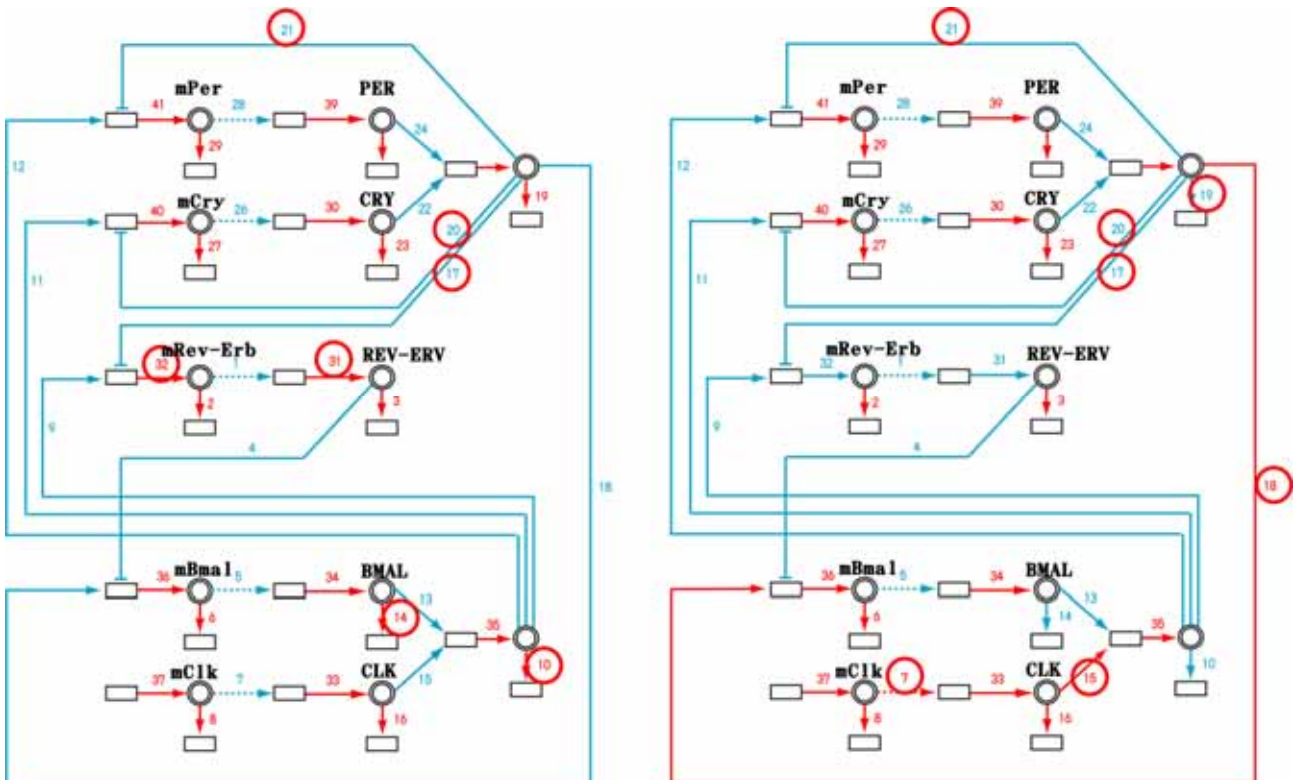


図 7：RNN を用いたサーカディアンモデルの脆弱部位の推定結果

RNN を用いた学習は非線形であるために、ユニット間における結合荷重の初期値によって、最終的な結果が異なる。図 7 右に初期値を変更した場合の結果を示す。初期値はランダム関数を用いて定義した。この結果から分かるように、脆弱部位推定で誤認識をしたアークが図 7 左の結果と異なっていることが言える。

これをまとめた結果を表 2 に示す。しかし、的中率と誤り率を表 1 と表 2 で比較したところ、それほど結果に対して違いは見られなかった。

表 2：提案手法による脆弱部位推定結果(2)

		脆弱部位	頑健部位
HFPN を用いて確認した指標		23	18
提案モデルによる脆弱部位予測結果	的中率	約 82.6%(19)	約 83.3%(15)
	誤り率	約 17.4%(4)	約 16.6%(3)

### 3.2 従来結果との比較

従来方法による学習結果を表 3 に示す[10]。この結果からも分かるように、改良モデルの結果である表 1,2 とほぼ同じ結果となった。これから判断できることは、3 種類のアークの違いを RNN に反映させるために改良モデルを作成したが、改良モデルの RNN の学習段階で吸収されてしまったことになる。これは、改良モデルがうまく動作しなかったわけではない。従来の RNN を用いた提案モデルでトークンの変化（プレース内の濃度量）が時間情報として RNN の入力ユニットに入っていることから、すでに 3 種類のアークの動作の違いについても、BPTT 法で学習した際に再形成出来ていた可能性が高い。言い換えれば改良モデルでは強制的に“process”, “association”, “inhibitory”の条件を与えたが、従来の RNN の学習モデルは 3 種類のアークの違いを教えることなく、学習なしのニューラルネットワークのように、類似度に応じて分類できていた可能性がある。したがって、改良モデルの条件を与えても、それに反応するユニットが存在せず同じ結果になってしまったと言える。

表 3：従来手法による脆弱部位推定結果

		脆弱部位	頑健部位
HFPN を用いて確認した指標		23	18
従来モデルによる脆弱部位予測結果	的中率	約 87.0%(20)	約 77.8%(14)
	誤り率	約 13.0%(3)	約 22.2%(4)

## 4. 結言

HFPN で記述された複雑な生命パスウェイを、時間成分を記憶可能な RNN に再形成し、自動で連続的にロックアウトを実現するシステムの開発を試みた。ここで、3 種類のアーク（“process”, “association”, “inhibitory”）の情報を考慮に入れた改良モデルを提案し、シミュレーション実験を試みることによって、アークの違いによる学習結果の違いについて考察を行った。サーカディアンモデルを用いて簡易実験を行い、従来の手法との推定率の比較を行った結果、従来の手法と改良モデルとでは結果として違いがないことが明らかとなった。これにより、3 種類のアークは従来の RNN の学習で吸収されていると位置づけられた。この結果は、今後の脆弱部位推定システムの開発において、演算処理を減少させるために重要な要素と考えられる。

今後の予定として、アークの種類による脆弱部位の推定率を向上できなかったため、その他の条件を定義することで脆弱部位の推定率を向上させる必要がある。その一つとして HFPN におけるトランジションの設定で連結するアークの速度情報が定義されているが、RNN では速度情報は定義されていない。脆弱部位の推定率を向上させるために、速度情報の重要性を考慮した結合荷重の情報を RNN の学習へ導入することを考えている。HFPN のトークン値をトランジションに連結するアークの速度式に代入し、トークンの変化量を算出することで推定率の向上を目指す。

## 参考文献

- [1] 山下富義, 生命をシステムとしてみる薬剤開発, 月刊薬事 Vol.46, No7, pp.15-17,2004.
- [2] <http://www.genomicobject.net>
- [3] <http://genome.ib.sci.yamaguchi-u.ac.jp/~gon/>
- [4] <http://pathway.sci.yamaguchi-u.ac.jp/>
- [5] Matsuno, H., Tanaka, Y., Aoshima, H., Doi, A., and Miyano, S., Biopathways representation and simulation on hybrid functional Petrinet, In *Silico Biology*, 3(3):389-404, 2003
- [6] Williams, R. and Zipser, D., A learning algorithm for continually running fully recurrent neural networks, *Neural computation*, 1:270-280, 1989.
- [7] 藤井靖, 松野浩嗣, 宮野悟, 井上慎一, ハイブリッド関数ペトリネットによる哺乳類の時計遺伝子機構のモデル化とシミュレーション, *時間生物学* Vol.11.No.1, 2005
- [8] 北風裕教, 松野浩嗣, 宮野悟, リカレントニューラルネットワークを用いた細胞内反応システムにおけるロバストネス瓦解部位予測, 独立行政法人国立高等専門学校機構大島商船高等専門学校紀要 第37号 pp.1-8, 2004
- [9] Kitakaze, H., Matsuno, H., Ikeda, N., Miyano, S., 5th International Workshop Bioinformatics and Systems Biology 22-25 August 2005 Berlin, Germany, *Genome Informatics* 16(1):192-202, 2005
- [10] 北風裕教, 松野浩嗣, 宮野悟, リカレントニューラルネットワークを用いた ハイブリッド関数ペトリネットの脆弱部位の推定, 電気・情報関連学会中国支部第57回連合大会 pp.203-204 (0701-6), 2006

