Technische Universität Berlin

ü

Institut fÃ¼r Telekommunikationssysteme
Fachgebiet Architektur der Vermittlungsknoten ?

Master Thesis

Hysteretic neural network s ~~and trading~~

Zongxiong Chen

Matriculation Number: 0390198
10.10.2019

Supervised by
~~Prof. Dr. Pavel Gurevich~~ there should be your formal advisor

Assistant Supervisor
Dr. Hannes Stuke
Dr. Nikita Begun

1 Introduction

Hysteresis (see Figures 1.1a to 1.1d) is defined as a *rate independent* process with *memory effect* [Visintin, 2013]. It's ubiquitous in various fields, including microelectronics [Bondurant and Gnadinger, 1989, Jiles and Atherton, 1983], materials [Kaltenbacher and Krejcí, 2014, Krejčí and Sprekels, 2007, Al Janaideh et al., 2009, Ge and Jouaneh, 1995], mechanics [Truesdell and Noll, 2004, Krejčí et al., 2014, Kunze and Marques, 2000], economics [Belke et al., 2013, Göcke, 2002, Belke et al., 2014, Blanchard and Summers, 1986], etc. The nonlinear operators like *non-ideal relay operator* (see Figure 1.1b), *stop operator* [Krejcı, 1996] (see Figure 1.1c), *play operator* [Krejcı, 1996] (see Figure 1.1d) and their generalizations are used as essential blocks to model dynamical systems with hysteresis. Preisach-type model, constructed as a superposition of non-ideal relays, is quite general in applications and it's equivalent to a linear combination of ~~*generalized*~~ nonlinear *plays*~~, namely, a *generalized Prandtl-Ishlinskii operator of play-type*~~ [Visintin, 2013].

or of

check this statement is correct

see details in Sec... of the thesis.

In this thesis, ~~we want~~ is to approximate a wide class of hysteretic processes, Preisach-type model(s), by recurrent neural network. Intuitively, recurrent neural network (RNN) is a optional approach to approximate systems with *memory* since it allows previous output to be used as input while having hidden states. Wang et al. [2018] applied Internal Time-Delay RNN to describe the hysteresis and showed promising performance that their network described both the major and minor hysteresis loops well. However, they preprocessed input by the ground-truth play operator and tried to find relationship between output and processed input. In other words, they didn't approximate *nonsmooth* operator, like play operator, in their model. We also checked Long Short Term Memory (LSTM) networks [Hochreiter and Schmidhuber, 1997], the state-of-the-art architecture of RNN, and found that it didn't perform well enough to reveal the relations between output and original input in hysteretic systems. In order to achieve better performance, **we develop a new neural network architecture, namely hysteretic neural network (HNN) (see Figure 1.2)**. It is a realization of a linear combination of ~~generalized~~ nonlinear plays and hence it's able to approximate any *Preisach operator* [Visintin, 2013]. ~~Restated, herein we approximate Preisach-type model via learning unknown generalized Prandtl-Ishlinskii operator.~~

I don't understand

add theorem

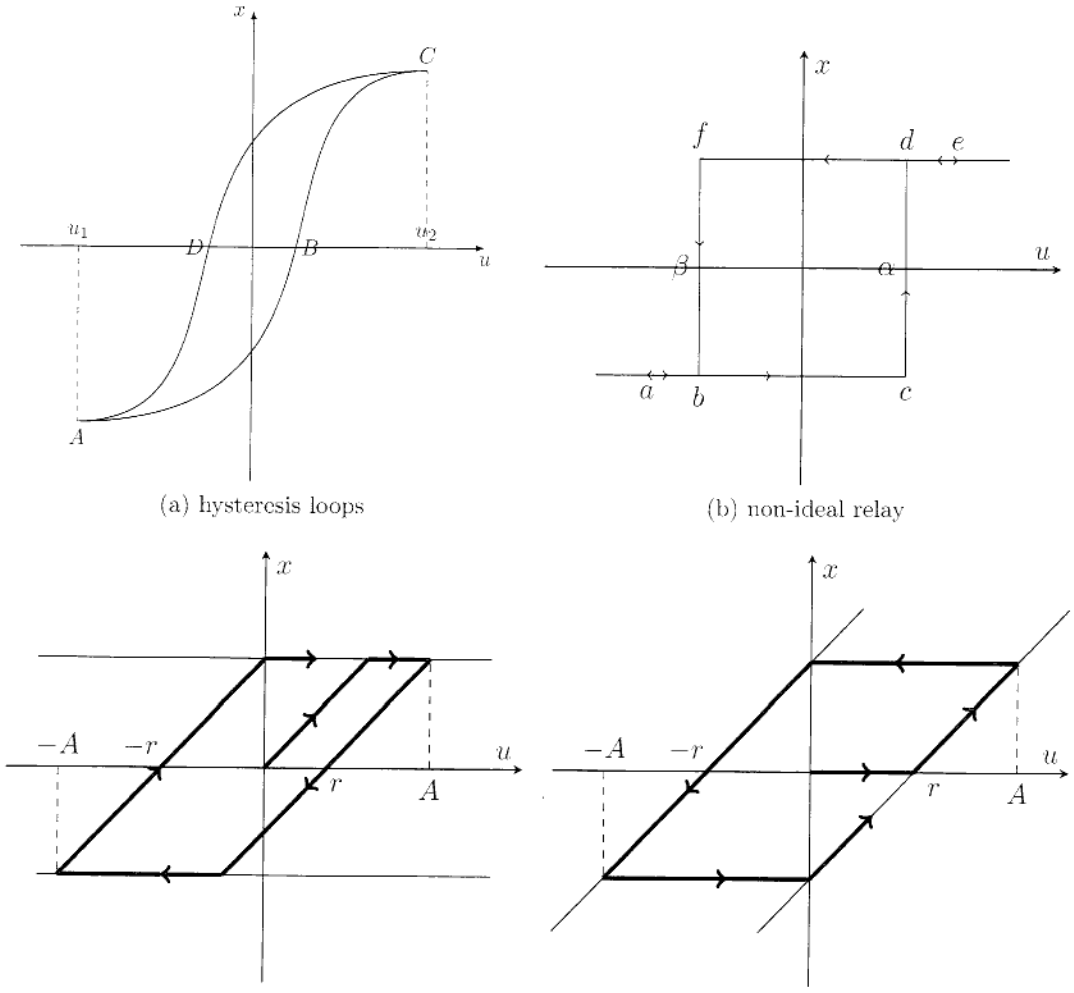
Given an observation (x_n, y_n) $(n = 1, \ldots, N)$ underlying hysteretic input-output relations, both LSTM and HNN minimize mean square error (MSE) between predicted target $\hat{y}_n$ and observed target y_n. **It shows HNN outperforms LSTM by comparing root mean square error (RMSE). In particular, HNN is able to reconstruct *minor hysteresis loops* well whereas LSTM fails.**

should I put some plots here to show that hnn outperforms lstm

yes

Based on the HNN we obtain, we study a particular application, momentum-based trading strategies in financial market, with hysteretic property in economic. Krejčí et al. [2014] proposed a market model and used Prandtl-Ishlinskii operator to model

(a particular case of the Preisach model)



(a) hysteresis loops (b) non-ideal relay

(c) stop (d) play

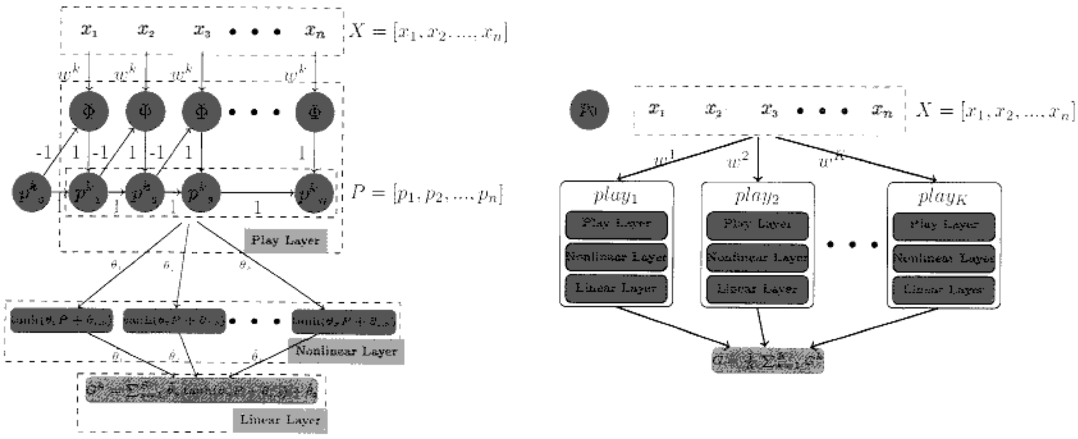
Figure 1.1: Interpretation of simplest *hysteresis loop*, *non-ideal relay*, *stop* and *play*. (Figure 1.1a) If u monotonically increases from u_1 to u_2, then the coordinate (u, x) moves along the path $A \to B \to C$; conversely, if u monotonically decreases from u_2 to u_1, then (u, x) moves along the path $C \to D \to A$. (Figure 1.1b) Hyper-parameters α and β correspond to *on* and *off* switching values of input, respectively. As the input u monotonically increased, the ascending branch $a \to b \to c \to d \to e$ is followed. When the input is monotonically decreased, the descending branch $e \to d \to f \to b \to a$ is traced [Mayergoyz, 1986]. Figure 1.1c and Figure 1.1d Input-output diagram for *stop* and *play* in the case $\dim X = 1, Z = [-r, r], u(t) = A\sin(\omega t)\ for\ A > r > 0$ [Krejcı, 1996]

trading strategies within their market model. ~~generalizing the model by supposing that the market agents also have a network structure and each agent reacts not only to the~~

too detailed for the intro.

~~price but to the states of their network neighbors~~ It provides a promising insight to make play-hysteresis economic models compatible with multi-agent modeling framework [Cross et al., 2007, Cartwright et al., 1999, Lamba and Seaman, 2008]. However, they only considered single trading strategy that agents take reaction to the change of *price trend*, called *agents D* (cf. ??) in this thesis, to simulate market movements. Another trading strategy, which is common in financial trading pattern as well, is also threshold-based where agents in markets are sensitive to the fluctuation of some *fixed price value* instead of *price trend*, called *agents N* (cf. ??) in this thesis. **We generalize Krejčí et al. [2014]'s market model by introducing two different agents, agents D and agents N based on ~~generalized~~ play operator and non-ideal relay operator respectively.**

Again, we learn this financial market model using HNN and LSTM, by ~~using~~ maximizing log-likelihood of price distribution ~~as loss function~~ to train both networks. It reveals HNN models this kind of market model better than LSTM. Even HNN is possible to reconstruct the *unobserved state changes underlying the market*, which is useful to interpret the *avalanche of market movements*, since it inherently contains agents that use different trading strategies in micro level.



(a) generalized play operator

(b) linear combination of multiple generalized play operator to approximate Preisach operator

Figure 1.2

The thesis is organized as follows. In the next chapter, we presents detail discussion about our financial market model and show how to generate synthetic data for training and test. In chapter 3, a detail methodology for HNN is discussed, explaining how we approximate the market model by neural networks and formulating the methods to learn hysteretic systems. In chapter 4, we evaluate performance between LSTM and HNN in different aspects, including the complexity of data sets, network complexity and the accuracy of predicted results. The last chapter contains our conclusions and future works.

change the order of chapters: first lstm, then HNN, then Market then HNN for Market