Neural Message Passing for Multi-Label Classification
Jack Lanchantin, Arshdeep Sekhon, Yanjun Qi
University of Virginia, Department of Computer Science
Crisis on Wall Street as Lehman Totters, Merrill Seeks Buyer, AIG Hunts for Cash

U.S. Opt to Avoid Lehman Rescue, Stirring a Momentous Weekend for American Finance; Traders Brace for a Chaotic Monday

BY CARRIC MOLLENKAMP, SUSANNE CRAIG AND SERENA NG

The American financial system was shaken to its core on Sunday. Lehman Brothers Holdings Inc. faced the prospect of liquidation and Merrill Lynch & Co. was close to a deal to sell itself to Bank of America Corp.

The U.S. government, which bailed out Fannie Mae and Freddie Mac a week ago and orchestrated the sale of Bear Stearns Cos. to J.P. Morgan & Co. earlier this year, drew a line in the sand with Lehman. It refused to provide a financial lifeline to potential buyers.

Without such support, Bar- clays PLC and Bank of America, the two most interested buyers, walked away. On Sunday night, Bank of America was close to striking a deal to buy Merrill Lynch for about $44 billion, or $28 a share. Lehman was working on a possible bankruptcy filing.

As worries spread across Wall Street that Lehman would not survive, brokerage firms, hedge funds and other traders moved to disentangle themselves from trades with Lehman. When hopes of a potential sale dimmed, a quiet Sunday on Wall Street turned into a mad rush. Executives and traders hurried to their offices or worked their phones to unwind outstanding contracts with Lehman and to gauge their overall exposure.

A sense of foreboding gripped the Street as top executives feared collateral damage from a Lehman liquidation. Attention turned to Merrill Lynch, which boasts the largest force of retail brokers, and to American International Group Inc., the insurer giant. Both firms have seen their stocks get hammered, and their managers spent the weekend frantically coming up with plans to reassure the markets.

"Monday will be a day of reckoning for the financial markets," said Carlos Mendez, senior managing director of ICP Capital, a boutique investment bank in New York. On Sunday, he said, "it was like a fire alarm went off and people ran in all directions."

AIG executives spent the weekend trying to raise cash, either from asset sales or a capital infusion from private-equity firms, or both. AIG executives were meeting with regulators to see if they could transfer capital from some of its subsidiaries to the holding company.

Merrill, whose retail brokerage force is largest in the country and is known as the "Thundering herd," quietly engaged in discussions with Bank of America, whose retail bank branches stretch coast to coast. Wall Street executives said the Federal Reserve was involved in orchestrating the sale, figuring that it was "better to save the relatively healthy patient instead of the dying one," said a lawyer involved in the discussions.

"We are in uncharted waters here," said a top executive of a big bank. "If Merrill can pull off a deal this weekend, that would certainly help."

The U.S. dollar, which had strengthened in the past few weeks, fell against all four of its major rivals on Sunday—the euro, the Swiss franc, the U.K. pound and the Japanese yen. Some executives involved in the Lehman discussions held out hope that an 11th-hour reprieve would materialize. Under one scenario aimed at limiting the ripple effects of Lehman's demise, a group of about 15 banks were in discussions Sunday to pool about $100 billion, which would be used to buy assets of the battered securities firm, according to one person familiar with the situation. Details were being finalized Sunday night. One possibility is that the Federal Reserve will support the move by opening its wholesale-borrowing window and relaxing collateral requirements for borrowers.

Lehman, a 158-year-old firm

Ultimatum

By Paulson Sparked Frenzied End

After more than a century and a half as an investment house, Lehman Brothers Holdings Inc. tumbled on the brink Sunday night. But its latest troubles were set in motion days earlier.

In midweek, Treasury Secretary Timothy F. Geithner said the government was formulating plans to rescue Lehman. Geithner, who has been warning of a "cliff" that threatened the stability of the financial system, was expected to outline a rescue plan for the embattled firm.

By Deborah Solomon, Dennis K. Berman, Suzanne Craig and Carrick Mollenkamp

Please turn to page A18
Crisis on Wall Street as Lehman Totters, Merrill Seeks Buyer, AIG Hunts for Cash

U.S. Opt to Avoid Lehman Rescue, Stirring a Momentous Weekend for American Finance; Traders Brace for a Chaotic Monday

BY CARRICEL MOLLENKAMP, SUSANNE CRAIG
AND SERENA NG

The American financial system was shaken to its core on Sunday. Lehman Brothers Holdings Inc. faced the prospect of liquidation and Merrill Lynch & Co. was close to a deal to sell itself to Bank of America Corp. The U.S. government, which bailed out Fannie Mae and Freddie Mac a week ago and orchestrated the sale of Bear Stearns Cos. to J.P. Morgan & Co. earlier this year, drew a line in the sand with Lehman. It refused to provide a financial backstop to potential buyers.

Without such support, Barclays PLC and Bank of America, the two most interested buyers, walked away. On Sunday night, Bank of America was close to striking a deal to buy Merrill Lynch for about $44 billion, or $22 a share. Lehman was working on a possible bankruptcy filing. As worries spread across Wall Street that Lehman wouldn't survive, brokerage firms, hedge funds and other traders moved to disentangle themselves from trades with Lehman. When hopes of a potential sale dimmed, a quiet Sunday on Wall Street turned into a mad rush. Executives and traders hurried to their offices or worked their phones to unwind outstanding contracts with Lehman and to gauge their overall exposure.

A sense of foreboding gripped the Street as top executives feared collateral damage from a Lehman liquidation. Attention turned to Merrill Lynch, which boasts the largest force of retail brokers, and to American International Group Inc., the insurer giant. Both firms have seen their stocks get hammered, and their managements spent the weekend trying to come up with plans to reassure the markets.

"Monday will be a day of reckoning for the financial markets," said Carlos Mendez, senior managing director of ICP Capital, a boutique investment firm in New York. On Sunday, he said, "it was like a fire alarm went off and people ran in all directions."

AIG executives spent the weekend trying to raise cash, either from asset sales or a capital infusion from private-equity firms, or both. AIG executives were meeting with regulators to see if they could transfer capital from some of its subsidiaries to the holding company.

Merrill, whose retail brokerage arm is No. 1 in the country and is known as the "Thundering herd," quietly engaged in discussions with Bank of America, whose retail bank branches stretch coast to coast. Wall Street executives said the Federal Reserve was helping orchestrate the sale, figuring that it was "better to save the relatively healthy patient instead of the dying one," said a lawyer involved in the discussions.

"We are in uncharted waters here," said a top executive of a big bank. "If Merrill can pull off a deal this weekend, that would certainly help." The U.S. dollar, which had strengthened in the past few weeks, fell against all four of its major rivals on Sunday—"the euro, the Swiss franc, the U.K. pound and the Japanese yen."

Some executives involved in the Lehman discussions held out hope that an 11th-hour rescue would materialize. Under one scenario, aimed at limiting the ripple effects of Lehman's demise, a group of about 15 banks were in discussions Sunday to pool about $100 billion, which would be used to buy assets of the bailed-out securities firm, according to one person familiar with the situation. Details were being finalized Sunday night. One possibility is that the Federal Reserve will support the move by opening its wholesale-borrowing window and relaxing collateral requirements for borrowers.

Lehman, a 158-year-old firm

Ultimatum
By Paulson Sparked
Frenzied End

By Deborah Solomon, Dennis K. Berman
Susan Craig and Carrick Mollenkamp

After more than a century and a half as an investment house, Lehman Brothers Holdings Inc. tumbled on the brink Sunday night. Its latest troubles were set in motion days earlier.

In midweek, Treasury Secretary Henry M. Paulson Jr. met with Lehman executives and demanded an immediate plan to keep it afloat. Lehman missed a major loan it had been counting on from Bear Stearns and agreed to cut its exposure to Lehman by $1 billion. Lehman said it had a plan to raise $1.5 billion from a group of investors that included AIG and Credit Suisse Group AG.

Lehman had recently disclosed that it was in talks to sell itself to Bank of America. But after Bank of America walked away Sunday, Lehman executives were scrambling to raise cash. Lehman officials were expected to meet with several of the investors on Monday to discuss the loan, according to a person who was briefed on the situation. Other investors were eager to meet with Lehman officials to discuss other plans for the company.

Lehman's bankruptcy filing Monday will likely be one of the largest in U.S. history. The company has reported $65 billion in assets and $45 billion in debt. Lehman's failure will likely trigger a domino effect, shaking up the global financial system and threatening to cause a recession.

Lehman's collapse is the latest in a string of high-profile failures in the financial industry. The collapse of Lehman Brothers follows a series of botched deals and mismanagement, including the company's decision to issue more than $20 billion in toxic mortgage-backed securities.

Lehman's failure is a blow to the reputation of former Chief Executive Richard Fuld, who led the company through its recent difficulties.

Lehman's collapse is likely to have a ripple effect across the global economy, as banks and investors grapple with the fallout from the financial crisis. The collapse of Lehman may lead to a new round of bailouts and government interventions, as governments try to prevent the crisis from spilling over into other sectors of the economy.

Lehman's collapse is likely to have a significant impact on the financial markets, as investors and banks adjust to the news of the company's failure. The failure of Lehman is likely to lead to a spike in volatility in financial markets, as investors and banks grapple with the news.

Lehman's collapse is likely to have a significant impact on the global economy, as governments try to prevent the crisis from spilling over into other sectors of the economy. The failure of Lehman is likely to lead to a new round of bailouts and government interventions, as governments try to prevent the crisis from spilling over into other sectors of the economy. The failure of Lehman is likely to lead to a spike in volatility in financial markets, as investors and banks grapple with the news.
Crisis on Wall Street, Lehman Totters, Merrill Seeks Buyer, AIG Hunts for Cash

Finance

United States

Economy

Housing
Multi Label Classification (MLC)

- MLC is the task of assigning a set of target labels for a given sample
- Given input $x$, predict the set of labels $\{y_1, y_2, ..., y_L\}$, $y_i \in \{0,1\}$

$x$

$y_1$ Castle
$y_2$ City
$y_3$ Mountains
$y_4$ Car
$y_5$ Road
Multi Label Classification (MLC)

- MLC is the task of assigning a set of target labels for a given sample
- Given input $x$, predict the set of labels $\{y_1, y_2, \ldots, y_L\}$, $y_i \in \{0,1\}$
Binary Relevance Classifiers

\[ p(y|x) \approx \prod_{i=1}^{L} p(y_i|x) \]
There are dependencies between labels!

\[ p(y|x) \approx \prod_{i=1}^{L} p(y_i|x) \]
Probabilistic Chain Classifiers

Read et. al. 2009, Wang et. al 2016, Nam et. al. 2017

\[ p(y|x) = \prod_{i=1}^{L} p(y_i|y_{1:i-1}, x) \]
Deep Structured Prediction Classifiers
Belanger & McCallum 2016, Gygli et. al. 2017

\[ y = \arg \min \ E(\bar{y}; F(x)) \]
\[
\bar{y} \in \{0,1\}^L
\]
Embedding Models
Bhatia et. al. 2015, Wu et. al. 2017

\[ \text{sim}(x, y_1), \text{sim}(x, y_2), \ldots \]
Interpretable vs Fast

- Probabilistic chain classifiers
- Embedding and energy function classifiers
Interpretable vs. Fast

- Probabilistic chain classifiers
- Embedding and energy function classifiers

This Work
Background: Message Passing Neural Networks
Message Passing Neural Networks (MPNNs)

Generalization of Graph Neural Networks
Message Passing Neural Networks (MPNNs)
Message Passing Neural Networks (MPNNs)
Message Passing Neural Networks (MPNNs)
Message Passing Neural Networks (MPNNs)
Message Passing Neural Networks (MPNNs)

- MPNNs can learn useful node representations for classifying nodes by encoding local graph structures and node attributes
- **Main idea**: pass messages between pairs of nodes and update them
Message Passing Neural Networks (MPNNs)

\[ G = (V, E) \]

\[ \mathbf{v}_i^t \in \mathbb{R}^d \]
Message Passing Neural Networks (MPNNs)

\[ G = (V,E) \]
\[ \mathbf{v}_i^t \in \mathbb{R}^d \]

message function \( F_m \)

\[ m_i^t = \sum_{j \in \mathcal{N}(i)} F_m(\mathbf{v}_i^t, \mathbf{v}_j^t), \]

node update function \( F_u \)

\[ \mathbf{v}_i^{t+1} = F_u(m_i^t) \]
Message Passing Neural Networks (MPNNs)

\[ G = (V, E) \]

\[ \mathbf{v}_i^t \in \mathbb{R}^d \]

message function \( F_m \)

\[ m_i^t = \sum_{j \in \mathcal{N}(i)} F_m(\mathbf{v}_i^t, \mathbf{v}_j^t), \]

node update function \( F_u \)

\[ \mathbf{v}_i^{t+1} = F_u(m_i^t) \]
MPNNs are good at modeling **relationships** (joint representation of nodes and edges)
MPNNs for Multi Label Classification
Neural Message Passing for MLC

- Castle
- City
- Mountains
- Car
- Road
Neural Message Passing for MLC
Neural Message Passing for MLC
Neural Message Passing for MLC
This Paper: Label Message Passing (LaMP) Networks

- **Main Idea:** Labels are represented as nodes in a label-interaction graph.
- Given input embedding $x$, the goal of LaMP is to model the conditional dependencies between label embeddings $\{u^t_{1:L}\}$ using Message Passing Neural Networks.
- LaMP uses MPNN modules to update label embeddings at each step $t$ in two parts.
Part A: Input-to-Label Message Passing

The first step passes messages from input $\mathbf{x}$ to each label embedding $\mathbf{u}^t_i$ to produce the intermediate embedding state $\mathbf{u}'^t_i$.

$$m^t_i = F_m(u^t_i, \mathbf{x}),$$

$$u'^t_i = F_u(m^t_i).$$
Part A: Input-to-Label Message Passing

The first step passes messages from input $\mathbf{x}$ to each label embedding $u_i^t$ to produce the intermediate embedding state $u_i^{t'}$.

\[
m_i^t = \sum_{j \in \mathbf{x}} F_m(u_i^t, x_i),
\]

\[
u_i^{t'} = F_u(m_i^t).
\]
Part B: Label-to-Label Message Passing

The second step passes messages between label embeddings to update their states conditioned on $x$ to produce updated embedding $u_i^{t+1}$:

$$m_i^{t'} = \sum_{j \in \mathcal{N}(i)} F_m(u_i^{t'}, u_j^{t'}),$$

$$u_i^{t+1} = F_u(m_i^{t'}).$$
Readout Function

- After $T$ updates to the label embeddings, the last module predicts the probabilities of each label being positive $\{\hat{y}_1, \hat{y}_2, ..., \hat{y}_L\}$
- Readout function $R$ projects each of the $L$ label embeddings $\{\mathbf{u}_{1:L}^T\}$ into a scalar value

$$\hat{y}_i = R(\mathbf{u}_i^T; \mathbf{W}^o) = \text{sigmoid}(\mathbf{W}_i^o \mathbf{u}_i^T)$$
Label Message Passing (LaMP) Networks

for T steps

Input-to-Label MP

Label-to-Label MP

x

u_1^t

u_2^t

u_3^t

u_4^t

u_5^t

\hat{y}_1

\hat{y}_2

\hat{y}_3

\hat{y}_4

\hat{y}_5

R

Neural Message Passing for MLC - Lanchantin, Sekhon, & Qi
Loss Function

Binary cross entropy, however since LaMP iteratively updates the label state for $T$ steps, we can impose a loss at each step $t$ by using the readout function to obtain intermediate predictions $\hat{y}_i^t = R(u_i^t; W^o)$

$$\text{Loss}(y, \hat{y}^t) = \frac{1}{T} \sum_{t=0}^{T} \frac{1}{L} \sum_{i=1}^{L} -(y_i \log(\hat{y}_i^t) + (1 - y_i) \log(1 - \hat{y}_i^t))$$
Label Graph Structure
Prior Label Graph Structure

- For datasets with **known label structure** (hierarchy, protein interaction), use known
- For all other datasets, we place an edge on the adjacency matrix for all labels that co-occur in any sample of the training set
Learned Label Graph Structure Using Attention

- Prior graphs may not be the best assumption since they are not conditioned on inputs
- We propose **using attention to learn the graph end-to-end** while training the classifier
Attention-Based MPNNs

\[ m_i^t = \sum_{j \in \mathcal{N}(i)} F_{m\alpha}(v_i^t, v_j^t), \]

\[ F_{m\alpha}(v_i^t, v_j^t) = \alpha_{ij}^t W^v v_j^t \]
Attention-Based MPNNs

Attention coefficient $e_{ij}^t$ for pair of nodes $(\mathbf{v}_i^t, \mathbf{v}_j^t)$ is computed using some attention function producing a scalar value representing the relationship between nodes:

$$e_{ij}^t = (W^q \mathbf{v}_i^t)^\top (W^u \mathbf{v}_j^t)$$

Attention weight $\alpha_{ij}^t$ produced by normalizing over neighboring nodes

$$\alpha_{ij}^t = \text{softmax}_j(e_{ij}^t) = \frac{\exp(e_{ij}^t)}{\sum_{k \in \mathcal{N}(i)} \exp(e_{ik}^t)}.$$
Attention-Based MPNNs
Experiments and Results
Datasets

- We validate our model on **8 real world MLC datasets**
- These datasets vary in the number of samples, number of labels, input type (sequential, tabular, graph, vector), and output type (unknown label graph, prior label graph)
LaMP Variations

**LaMP\textsubscript{el}** uses an **edgeless label graph** and messages are not passed between labels, assuming no label dependencies.

**LaMP\textsubscript{pr}** uses a **prior label graph** where each label is able to attend to only other labels from the known label graph.

**LaMP\textsubscript{fc}** uses a **fully connected label graph** where each label is able to attend to all other labels.
Example-Based F1 Scores (ebF1)
LaMP vs RNN Speed

Neural Message Passing for MLC - Lanchantin, Sekhon, & Qi
Visualization: Intermediate Predictions
Visualization: Label to Label Attention
Visualization: Label to Input Attention
Conclusion

- LaMP models label interactions for MLC by placing labels as nodes on a graph
- LaMP networks are as accurate, faster, and more interpretable than the previous state-of-the-art MLC classifiers
- **Representation learning**: exploiting the inductive bias of a model automatically discover the representations needed for classification
Thank You

code available at: github.com/QData/LaMP