Fighting Fraud in Uncertain Economic Times

A conversation with

Dr. Victor Lee
Head of Product Strategy & Developer Relations
TigerGraph
Moderated by **Mike Meriton**
Co-Founder & COO, EDM Council

- Joined EDM Council full-time 2015 to lead Industry Engagement
- EDM Council Co-Founder & First Chairman (2005-2007)
- Former CEO GoldenSource (2002-2015)
- Former Executive for D&B Software and Oracle
- FinTech Innovation Lab – Executive Mentor (2011 – Present)
Dr. Victor Lee
Head of Product Strategy and Developer Relations

- Joined TigerGraph in 2015
- 20 years in tech industry
- B.S. University of California, Berkeley and M.S. Stanford University in Electrical Engineering
- Ph.D. Kent State University in Computer Science
  Research in Graph Data Mining
- Former Visiting Professor of Computer Science,
  John Carroll University
Fraud is a Huge and Growing Problem

Credit Card Fraud alone reached $30B in 2020.

33 million US cardholders had **blocked sales** amounting to $118B

Only 1 in 5 blocked transactions are actually fraudulent → significant loss of potential sales

The pandemic has led to rises in identity theft, fake charities, phony businesses, insurance fraud, phishing, and other types of fraud.
Stopping Fraud is a Constant Battle

● Technology gives fraudsters new methods, with speed and volume
  ○ Use disposable phones, phone numbers, email addresses, and online accounts
  ○ Send a million phishing emails for pennies
  ○ Coordinate with remote colleagues in real time
  ○ Produce and print authentic-looking documents and websites

⇒ Complex, multi-step patterns designed to hide their tracks

● You need to also handle complex analytics, with speed and volume

● Detect and learn new fraud patterns ⇒ Machine Learning
Why Graph, Why Now?

“Graph analysis is possibly the single most effective competitive differentiator for organizations pursuing data-driven operations and decisions after the design of data capture.” - Gartner
“Finding relationships in combinations of diverse data, using graph techniques at scale, will form the foundation of modern data and analytics.”

- Graph analytics is a set of analytic techniques that allows for the exploration of relationships between entities of interest such as organizations, people and transactions.

- Graph analytics will grow in the next few years due to the need to ask complex questions across complex data, which is not always practical or even possible at scale using SQL queries.

- Through 2022, the application of graph processing and graph databases will grow at 100% annually to accelerate data preparation and integration, and enable more adaptive data science.
Graph-Powered Analytics & Machine Learning

1. Richer Data
   - Relationships are 1st Class Citizens
   - Connects different datasets and silos

2. Deeper Questions
   - Look for semantic patterns of relationship
   - Search far and wide more easily

3. More Analytical Tools
   - Graph algorithms
   - Graph-enhanced machine learning

4. Explainable Results
   - Semantic data model, queries, and answers
   - Visual exploration and results
Advanced Analytics on Connected Data

Deep
- **Analyze** relationships deeper into the data to find hidden patterns

Wide
- **Connect** all datasets to uncover undisclosed relationships

Operational
- **Process** transactions in real-time to provide the next best action

At Scale
Benefiting From Both Semantic Graphs And Property Graphs
Complementary Technologies To Address Different Needs

Semantic (RDF) Knowledge Graphs:
- Logical inference
- Standards-based
- Pattern matching

Property Graphs:
- High performance for queries, transactions, and advanced analytics
- Pattern matching
- Schema-free or Schema-based
- Schema-based allows application-specific tuning

Knowledge

Information

Data

Ontology

Schema

RDF triples

nodes and edges

Semantic context, Principles and rules

Structured, Categorized, Useful

Individual raw facts and measurements

© 2020 TigerGraph. All Rights Reserved
1. Richer Data: Connect the Dots Across Internal & External Data with Graph

Deliver consistently high quality service while controlling costs
- Detect and prevent fraud and money laundering
- Optimize Network & IT Infrastructure to meet the service levels at lower costs

Increase consumer loyalty and usage
- Understand consumer needs & deliver tailored products and services

Improve merchant/bank market share
- Understand merchant/bank needs and deliver tailored products and services
Intuit Implements Real-Time Fraud Investigation with Customer 360

Business Challenge
Intuit realized their homegrown link analysis solution for fraud detection needed an upgrade. Risk investigators were switching between multiple tools causing a manual, time consuming process.

Solution
• Payments with potential fraudulent activity are flagged in real-time
• Over 50 investigators explore each suspected transaction using GraphStudio to understand linkage to customer accounts and previous transactions with fraudulent activity

Business Benefits
Fraud Detection solution with graph has improved the productivity and efficiency of the risk investigators, allowing them to identify and investigate fraud in real-time connecting information across the organization.
2. Deeper Questions
Fraud Detection in Financial Payments

Diagram showing relationships between User, Account, Email, Bank, Device, Payment, Phone number, and other entities with arrows indicating "Sets_Up," "Authenticated_by," "Has," "Associated_with," "Send_payment," "Receive_payment," and "Used_for."
Fraud Detection in Financial Services (Payments)

Regular Analytics vs Advanced Analytics with TigerGraph (Deep)

User 1
- Sets_Up
- Authenticated_by
- Has

Account 1
- Sets_Up
- Authenticated_by
- Has

Payment 1
- Sends_payment
- Receive_payment
- From

Email 1
- From

Credit Card
- From

Bank
- From

User 2
- Sets_Up
- Has

Payment 101
- Has

Account 101
- Sets_Up

Stolen Credit Card
- Has

Device 1 (phone)
- Used_for
- Linked_to

Phone_number 1
- Used_with

Device 101
- Used_for
- Send_payment
- Has

Phone_number 101
- Has

User 3
- Sets_Up

Involved in Prior Fraud Cases

Sign up FREE for TigerGraph Cloud at TGCLOUD.IO to use the starter kit for fraud detection (payments)
OpenCorporates Upgrades Performance and Functionality

Business Challenge
OpenCorporates, world’s largest open database of corporate information experienced challenges in terms of scalability, lack of support for simple queries and speed as it ramped up to production using a first-generation graph technology.

Solution
- Support queries of up to five degrees of separation to help uncover relationships between entities and see which relationships are active vs. dead
- Unlock insight into how relationships and networks have changed and evolved over time (temporal graph search)

Business Benefits
OpenCorporates is scaling up its database with 170 million corporate entities to provide users with deeper analysis of the information, helping them uncover instances of criminal or anti-social activity - such as corruption, money laundering, and organized crime.

“As our work continues and our data grows, we had challenges scaling our data to meet our business needs. TigerGraph’s excellent scalability and performance enables us to achieve things we previously could not do, and to better support ongoing investigative work in the process.”

Additional details in the press release - https://tinyurl.com/y36skysr
Pagantis Delivers Faster Consumer Finance Services with TigerGraph on AWS

Business Challenge
Pagantis must **assess credit worthiness and fraud risk in real-time** for customers to allow them to pay for their purchase in monthly installments. Risk assessment with relational databases was taking too long, delaying the time for loan approvals.

Solution
- Real-time calculation of customer’s credit rating using their current activities as well as all available historical data
- A scalable, high-performance system to deliver insights into complex relationship-based workflows for **credit scoring, fraud detection, recommendation engines and risk analysis**

Business Benefits
Pagantis can now offer a faster and seamless consumer finance solution for the eCommerce merchants throughout Italy, France and Spain.

Know Your Customer
Combine Internal Data With Third-party Sources


© 2020 TigerGraph. All Rights Reserved

Regular Analytics (Narrow) vs Advanced Analytics with TigerGraph (Wide)
3. More Analytic Tools - Graph Algorithms

- Graph algorithms answer precise, meaningful questions about connected data
- Each algorithm is like a tool in an analytics toolkit
- Building blocks for more complex business questions

Specialized functions

Combine to make something better
# Graph Algorithm Types and Uses

<table>
<thead>
<tr>
<th>Path-Finding</th>
<th>Centrality</th>
<th>Community Detection</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is there a close but otherwise hidden relationship between two entities?</td>
<td>Is an account/merchant/bank/etc the focus of unusually high activity, either directly or indirectly</td>
<td>Find a group of closely interconnected individuals --Could signal either good or bad situation, depending on context</td>
<td>One you find a pattern or individual of interest, where are their similar individuals of interest --having the same type of relationships</td>
</tr>
<tr>
<td>• Shortest path</td>
<td>• PageRank</td>
<td>• Connected component</td>
<td>• Cosine similarity</td>
</tr>
<tr>
<td>• Shortest weighted path</td>
<td>• Betweenness Centrality</td>
<td>• Louvain modularity</td>
<td>• Jaccard similarity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Label propagation</td>
<td></td>
</tr>
</tbody>
</table>
7 Key Data Science Capabilities Powered By A Native Parallel Graph

1. **Deep Link Analysis**
   - From a set of entities (e.g. customers, accounts, doctors), show all links or connections.

2. **Multi-dimensional Entity & Pattern Matching**
   - Query Pattern P
   - Match
   - Given a pattern (e.g. a type of suspicious activity), find similar patterns in the graph.

3. **Relational Commonality Discovery and Computation**
   - Given 2 entities (e.g. customers, merchants, devices), follow their relationships to find commonality.

4. **Hub & Community Detection**
   - Find most influential members (customers, doctors, citizens) & detect community around them.

5. **Geospatial Graph Analysis**
   - Analyze changes in entities & relationships with location data.

6. **Temporal (Time-Series) Graph Analysis**
   - Analyze changes in entities & relationships over time.

7. **Machine Learning Feature Generation & Explainable AI**
   - Extract graph-based features to feed as training data for machine learning; Power Explainable AI.
IceKredit Uses a Customer 360 Graph and Machine Learning For Credit Rating & Risk Assessment

Business Challenge
With rapid growth in size and complexity of the interconnected global financial markets making it difficult for banks to process loan applications for home, automobile, etc.

Solution
- Leverage Machine Learning and AI for custom advanced models and analytics to build comprehensive credit views for applicant
- Quantify applicant’s fraud probability and compares it with actual business activity
- Find undisclosed relationships and connections within data; assign and update risk ratings in real time

Business Benefits
IceKredit is empowering lenders by reducing their fraud risks with more accurate, detailed credit ratings for applicants that are not tracked by traditional credit bureaus.

Additional details in the TechTarget article - https://tinyurl.com/y696vkqf
Detecting Phone-Based Fraud by Analyzing Network or Graph Relationship Features at China Mobile

GOOD PHONE

- # of call-SD
- # of call-SD
- Many in-group connections
- 3-step friend relation

BAD PHONE

- Empty stable group
- Short call duration
- No call back
- Many rejected calls
- Average distance > 3

GOOD PHONE FEATURES
1. High call back phone
2. Stable group
3. Long term phone
4. Many in-group connections
5. 3-step friend relation

BAD PHONE FEATURES
1. Short call duration
2. Empty stable group
3. No call back phone
4. Many rejected calls
5. Average distance > 3

3. More Analytical Tools & 4. Explainable Results

Graph Feature Extraction → External Training

TRAIN

Training data → GSQL Queries to Extract Features → Feature vectors → ML Training

DEPLOY

Query → Prediction
## The TigerGraph Difference

<table>
<thead>
<tr>
<th>Feature</th>
<th>Design Difference</th>
<th>Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep-Link Pattern Discovery</td>
<td>• Native Graph, for speed and efficiency</td>
<td>• Uncovers hard-to-find fraud</td>
</tr>
<tr>
<td></td>
<td>• Uncovers hard-to-find fraud</td>
<td>• Operational, real-time fraud detection</td>
</tr>
<tr>
<td></td>
<td>5 to 10+ hops deep</td>
<td></td>
</tr>
<tr>
<td>Handling Massive Scale</td>
<td>• Distributed DB architecture</td>
<td>• Integrates all your data</td>
</tr>
<tr>
<td></td>
<td>• Massively parallel processing</td>
<td>• Automatic partitioning</td>
</tr>
<tr>
<td></td>
<td>• Compressed storage reduces footprint and messaging</td>
<td>• Complete data → Better detection</td>
</tr>
<tr>
<td>In-Database Analytics</td>
<td>• GSQL: High-level yet Turing-complete language</td>
<td>• Avoids transferring data</td>
</tr>
<tr>
<td></td>
<td>• User-extensible graph algorithm library, runs in-DB</td>
<td>• Richer graph context</td>
</tr>
<tr>
<td></td>
<td>• ACID (OLTP) and Accumulators (OLAP)</td>
<td>• Option for in-DB machine learning</td>
</tr>
</tbody>
</table>
TigerGraph: 3rd Generation Graph Database

Real-time Performance
Sub-second response for queries touching tens of millions of entities/relationships

Transactional (Mutable) Graph
Hundreds of thousands of updates per second, billions of transactions per day

Deep Link Multi-Hop Analytics
Queries traverse 10+ hops deep into the graph performing complex calculations

Ease of Development & Deployment
- GraphStudio - visual SDK
- GSQL - Intuitive, Turing complete graph query language for developing complex analytics in days
- User extensible graph algorithms library

Scalability for Massive Datasets
100 B+ entities, 1 Trillion+ relationships

Enterprise Grade Security
- Encryption at-rest and in-transit
- Control access to sensitive data based on user role, dept or organization with MultiGraph
Summary

● Defeating fraud requires **connecting the dots** and **seeing hidden patterns**.

● Graph databases and analytics provide the deep insight and performance needed.

● Graph analytics is in production, delivering results.

● A native parallel, distributed graph provides the best performance for high-performance analytics and machine learning.

Get Started with TigerGraph Cloud today with 25+ starter kits – [https://www.tigergraph.com/starterkits/](https://www.tigergraph.com/starterkits/)
Questions?
FOR MORE INFORMATION:

Victor Lee
Head of Product Strategy
TigerGraph
victor@tigergraph.com
Backup Slides
Graph enables organizations of all sizes to ask their best questions of connected data.

Graph integrates data from multiple internal and increasingly diverse external sources. Larger, richer, and more diverse datasets mean more entities can be analyzed as more connections are understood.

According to Gartner, over the next five years the graph analytics market will see 100% cumulative annual growth.
Top 5 Reasons Why Fighting Fraud is Hard without Graph + Machine Learning

5. **Hidden Dependencies**: So many connections and hidden dependencies with limited ‘human decision time’ to uncover **good and bad** events

4. **Risk Scoring**: *Human processing* of risk scoring is complex & mistakes are costly

3. **Data Volumes**: **Massive** data set, logs, history, system data, application data

2. **Changing Landscape**: Fintech and fraud tech are changing rapidly

1. **Real Time**: **Complexity** amplified by connected data. Nested entities in SQL Systems → Queries are complex, fragile, and cumbersome
Automated Fraud Prevention with Advanced Analytics

Deep
Analyze relationships deeper into the data to find hidden fraud patterns

Wide
Connect all datasets to uncover undisclosed relationships

Operational
Process transactions in real-time to flag possible fraud

At Scale
Automated Fraud Prevention

Problem: Losing $M to fraudsters because of manual and slow investigation processes.

Solution: By adding graph to their fraud systems, payment and credit processors have created a rapid and precise process automation system that significantly reduces the number of false-positives.

Investigators can now focus on investigating high-value fraud cases.
Credit Risk Assessment - Combine Traditional & Non-traditional Data Sources

Regular Analytics (Narrow) vs Advanced Analytics with TigerGraph (Wide)

EDM Webinar Guidelines & Planning References
# Production Timeline & Collaboration

<table>
<thead>
<tr>
<th>Week</th>
<th>Activity</th>
<th>Responsible Parties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>Kick-off call &amp; discuss topic idea; confirm event date</td>
<td>EDMC &amp; Sponsor</td>
</tr>
<tr>
<td>Week 2</td>
<td>Finalize topic and presenter(s)</td>
<td>Sponsor with EDMC input</td>
</tr>
<tr>
<td></td>
<td>Set-up webinar platform and registration form</td>
<td>EDMC</td>
</tr>
<tr>
<td></td>
<td>Draft marketing communications (e.g. emails)</td>
<td>EDMC (Paragon), with sponsor input on topic</td>
</tr>
<tr>
<td>Week 3</td>
<td>Outline content/flow/talking points</td>
<td>Sponsor</td>
</tr>
<tr>
<td></td>
<td>Launch marketing communications &amp; Registration (Weekly email and social posts)</td>
<td>EDMC (Paragon) to lead; Sponsor to also promote directly</td>
</tr>
<tr>
<td>Week 4</td>
<td>Create presentation</td>
<td>Sponsors</td>
</tr>
<tr>
<td>Week 5</td>
<td>Run-throughs &amp; edits</td>
<td>Sponsor &amp; EDMC</td>
</tr>
<tr>
<td>Week 6</td>
<td>Final run-through and tech rehearsal</td>
<td>Sponsor &amp; EDMC</td>
</tr>
<tr>
<td>Week 7</td>
<td>LIVE Webinar</td>
<td>Sponsor &amp; EDMC</td>
</tr>
<tr>
<td></td>
<td>Final marcomms, plus reminders to attendees</td>
<td>EDMC (Paragon)</td>
</tr>
<tr>
<td>Week 7+</td>
<td>Post-comms &amp; promos with recorded link</td>
<td>EDMC (Paragon) to lead; Sponsor to also promote directly</td>
</tr>
</tbody>
</table>
Industry Narrative Format

1. Business Problem / Challenge
2. Traditional ways to solve this problem and limitations of these approaches
3. New / Innovative way to solve these issues
4. Illustration of new approach (case stories/demo as proof) > business value delivered
5. How to prepare for this new approach / lessons learned
6. How to learn more about this opportunity
General Guidelines – Topic / Content

• Industry topic (not company description) leads the beginning and core content for the presentation

• Citing real use cases is the highest value and should be given the most time

• Take a deeper dive on 1-2 points to show depth of thought on key areas
General Guidelines – Slides

• Deck should have 20-25 slides to meet 30-35 minutes of delivery

• This is a thought leadership piece, avoid making it a “sales pitch”

• Defer company credentials to 1-2 slides towards conclusion of webinar to “learn more”

• Close with lessons learned and what to do to prepare for this new capability/approach

• Provide 1-2 open questions for EDMC moderator to kick off the Q&A and 1 concluding question to wrap-up
Speaking Roles

- Administrator: Carole Mahoney
- Moderator: Mike Meriton or John Bottega
- Subject Matter Expert: Sponsor
Tips When **LIVE**

- Join 15-minutes early for a final prep
- We recommend sharing your webcam – adds a personal touch
  - Wear business casual attire in a well-lit room
  - Look directly into the camera (avoid reading from off-screen notes or looking at phone)
- Assign a back-up presenter (in case main presenter faces technical difficulties or connection loss)
- Use GoToWebinar private chat to message other organizers during live event if needed
- Control distractions
  - Turn off phone prior to webinar
  - Mute desktop notifications
  - Make arrangement for pets/children (avoid barking, background noise, etc.)
Live Timeline

- **10:45 AM** - Panelists must arrive 15 minutes early for final prep

- **11:00 AM** - Carole starts, giving 3 minutes for people to join
  - Covers Housekeeping – recording / how to use Q&A panel

- **11:03 AM** – MM sets the stage for the industry challenge and intro main speaker

- **11:05 AM** – Core TLS materials

- **11:40 AM** – Q&A
  - MM to frame 1-2 known questions to start
  - MM to moderate open Q&A
  - MM delivers final known wrap up question

- **11:50 AM** – EDM Webinar concludes
Next Steps & Responsibilities

• **Sponsor** and **EDM Council** to agree on webinar date and time

• **Sponsor** to finalize topic
  - **EDM Council** will need the following information: list of webinar speakers and webinar highlights (why is this webinar relevant, what will you cover, who is the target audience, etc.)

• Based on finalized topic and details provided, **EDM Council** to:
  - Create GoToWebinar registration link
  - Draft promotions (email + social) - distributed to entire database
  - Post on **EDM Council** website calendar & homepage

• **Sponsor** to finalize webinar deck (1-2 weeks prior to webinar)
  - Feel free to send WIP files for **EDM Council** review throughout the process

• **Sponsor** and **EDM Council** to schedule webinar rehearsal (1 week prior to webinar)
Deliverables

• **Sponsor** has perpetual rights to share and host final webinar recording

• **Sponsor** will receive the full registration list (excluding those that chose to opt-out)
  • On average, opt-in rate is over 80%
  • Registration details included:
    • First & Last Name (required field)
    • Email Address (required field)
    • Organization (required field)
    • Job Title (optional field)

• **EDM Council** will send the transcribed Q&A + presentation PDF to registrants post-event

• **EDM Council** will promote the webinar recording post-event

• **EDM Council** will host the webinar recording on the EDMC website
Co-marketing the webinar

• We encourage and recommend Sponsors to also promote the EDM Webinar

• Best practices for promoting a webinar
  • Distribute at least 2 promotional emails
  • Publish at least 3 promotional social posts (LinkedIn + Twitter)

• EDM Council will be happy to help draft promotional copy as well as design custom social graphics as needed
  • Please contact edm@paragonpr.com if you need assistance or if you have any additional questions regarding EDM Webinars
Post Webinar

• The webinar recording will be sent to all registrants upon completion of the webinar

• Q&A transcription sent out within 2-3 days of event – responses drafted by partner and QR by EDMC

• Opt-In Registrants shared with Sponsors for their additional direct follow-up