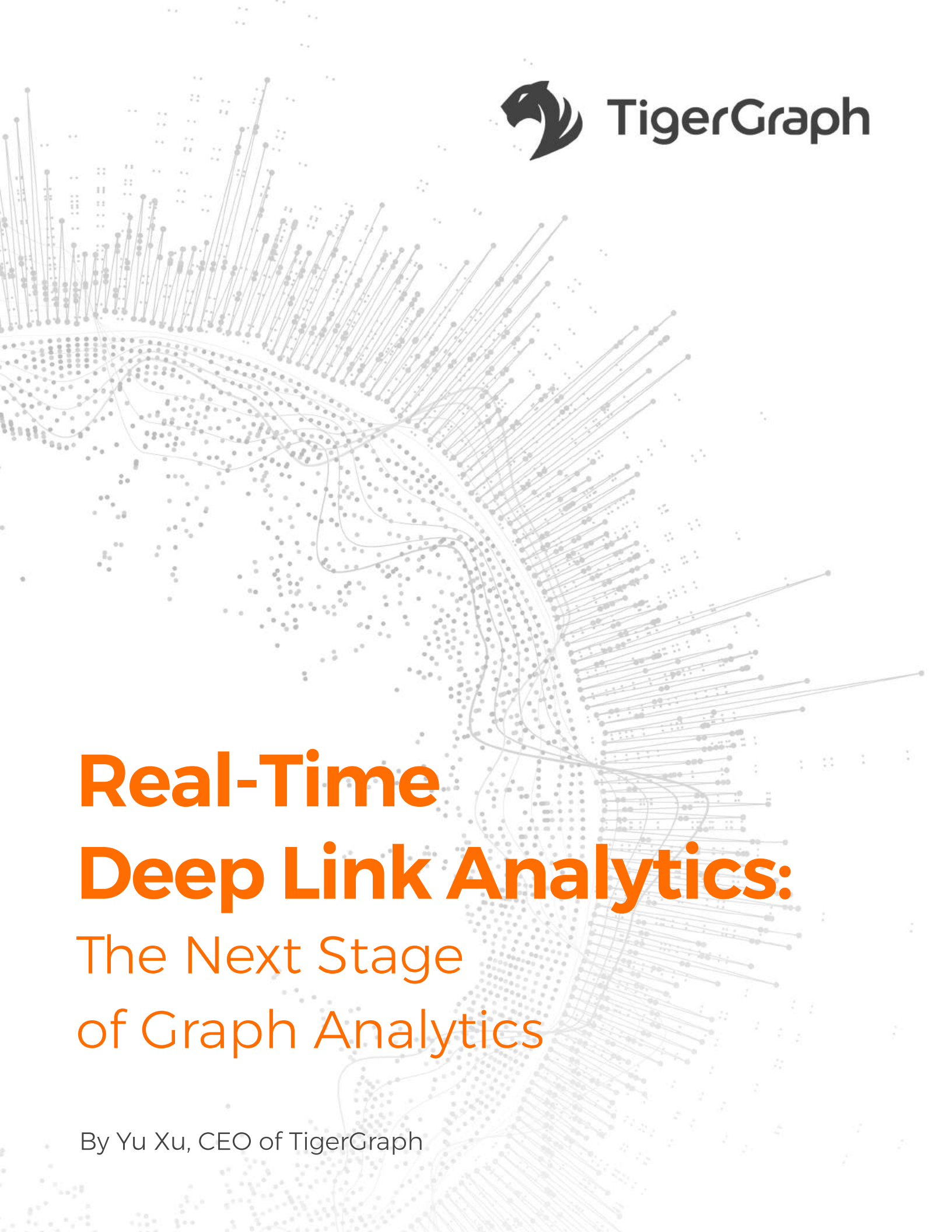


TigerGraph

Real-Time Deep Link Analytics:

The Next Stage of Graph Analytics

By Yu Xu, CEO of TigerGraph

One of the greatest strengths of graph databases is their ability to analyze and answer complex questions about the relationships and connections found in large datasets. Early generation graph technologies have been better at this relationship analysis than traditional RDBMS applications. But they find themselves slow with limited analysis capabilities when connecting to the ever increasing real-time data so prevalent in today's enterprise ecosystems.

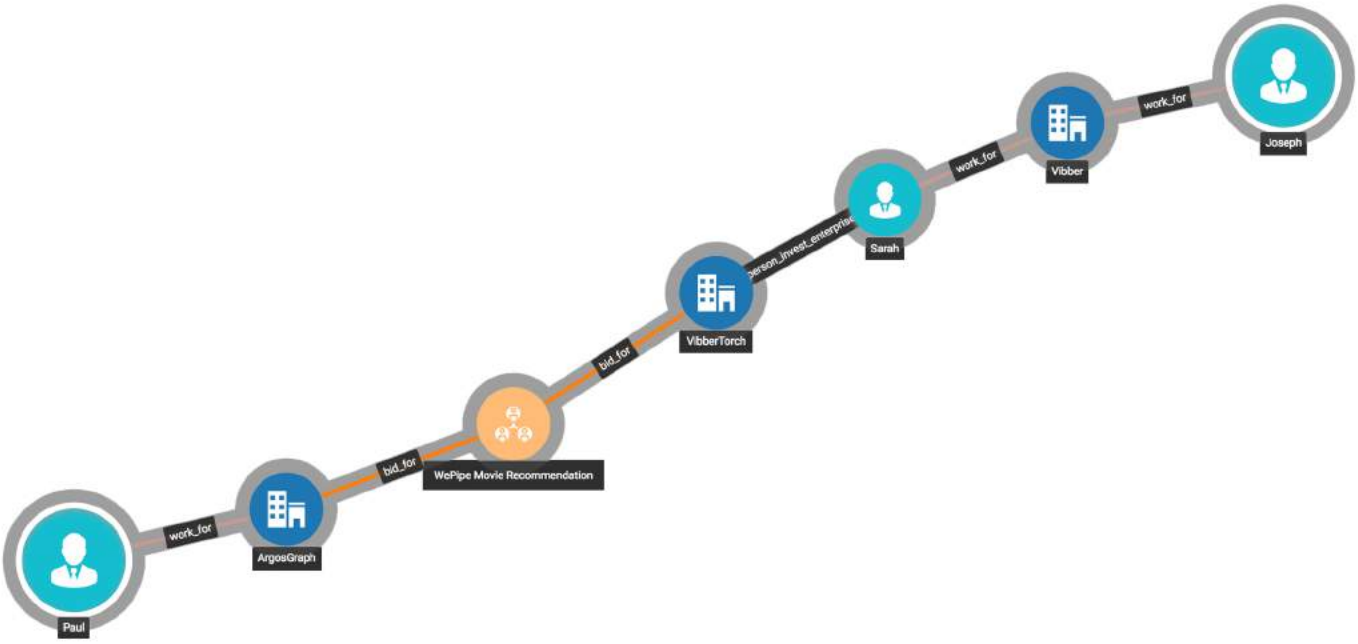
That's because existing graph technologies have trouble loading large quantities of data, cannot deliver fast graph traversal speed, and cannot ingest fast-arriving data in real time. As a result, current graph technologies are typically limited to two hops of traversal (i.e., two degrees of separation) in big graphs.

Let's look at a simple personalized recommendation like "customers who liked what you liked also bought these items." Starting from a person, the query first identifies items you've viewed/liked/bought. Second, it finds other persons who have viewed/liked/bought

those items. Third, it identifies additional items bought by those persons.

```
Person → Product → (other)Persons → (other)Products
```

This query requires 3 hops in real-time, so it is beyond the 2-hop limitation of current-generation graph technologies on larger datasets. Adding in another relationship easily extends the query to 4 or more hops.



Today's enterprises need a real-time graph analytic capabilities that can explore, discover and predict very complex relationships utilizing 3 to 10+ hops. Each additional hop reveals additional connections and hidden insight.

This is **Real-Time Deep Link Analytics**—which has the following performance attributes:

ATTRIBUTE	PERFORMANCE REQUIREMENT
High connectability	3 to 10+ hops
Fast graph traversal speed	100+ M vertex or edge traversals/second per server
Fast data update	100+ K updates/second per server

Real-Time Deep Link Analytics Use Cases

Real-Time Deep Link Analytics empowers the fraud prevention, personalized recommendation, supply-chain optimization, and other analysis capabilities found in today's most critical enterprise applications.

RISK/FRAUD CONTROL

Real-Time Deep Link Analytics combats financial crime by identifying high risk transactions. For example, starting from an incoming credit card transaction, how this transaction is related to other entities can be identified as follows:

```
New Transaction → Credit Card → Cardholder → (other)Credit Cards →
(other)Bad Transactions
```

This query uses 4 hops to find connections only one card away from the incoming transaction. Today's fraudsters try to disguise their activity by having circuitous connections between themselves and known bad activity or bad actors. Any individual connecting path can appear innocent, but if multiple paths from A to B can be found, the likelihood of fraud increases. Thus more hops are needed to find connections two or more transactions away. This traversal pattern applies to many other use cases - where you can simply replace the transaction with a web click event, a phone call record, or a money transfer. With Real-Time Deep Link Analytics, multiple, hidden connections are uncovered and fraud minimized.

MULTIPLE-DIMENSIONAL PERSONALIZED RECOMMENDATION

To find similar customers or to provide personalized recommendation for products, Real-Time Deep Link Analytics is necessary. Consider the following examples of recommendation paths, all of which are longer than 3 hops:

```
Person → Purchases → Products → (other)Purchases →
(other)Persons → (other)Products
```

```
Person → Purchases → Products → Categories → (other)Products
```

```
Person → Clicks → Products → Categories → (other)Products
```

Due to the limitations in current systems, most of the recommendation functions seen today use offline computation. The recommendations are computed in the background and are not able to perform real-time, on-demand analytics using the latest data.

POWER FLOW OPTIMIZATION / SUPPLY-CHAIN LOGISTICS / ROAD TRAFFIC OPTIMIZATION

These applications adjust each entity and/or each connection, until a dynamic equilibrium or optimization is established. For example, in a national or regional power grid, the power output of each generator is adjusted.

```
All Entities
→ relationship to and effect on neighbors
→ Update All Entities
<repeat until a stable state is achieved>.
```

Finding the right values inherently requires iterative computation in a graph (similar to PageRank) until some metric values converge. Moreover, each top-level component (e.g., a

power generator), is the hub for a network of supporting elements, resulting in a multi-tier mesh. A power distribution network could easily have a 6-hop path.

```
Power generator → Transformers → Control units → Lower-level
transformers → Lower-level control units → Meters → Power-consuming
devices
```

Real-Time Deep Link Analytics is absolutely needed for a graph data system to handle this level of computation.

REAL-TIME DEEP LINK ANALYTICS AND ENTERPRISE AI

Real-Time Deep Link Analytics are a key ingredient for Enterprise AI. Due to their technology limitations, earlier data analytics technologies artificially chopped entities and relationships and stored them as rows or key-value pairs scattered across multiple tables/rows/columns. Thus, some of the relationships are lost or are more difficult to process. Real-Time Deep Link Analytics can handle all the connections between data entities enabling users to explore, discover and predict relationships more accurately, more quickly, and with more insight. These capabilities are all essential to enterprise AI applications such as personalized recommendation, fraud prevention, supply-chain logistics optimization, and company knowledge graph construction and querying.

Introducing TigerGraph

TigerGraph is the first and only native parallel real-time graph analytics platform and TigerGraph is the first and only graph database with Real-Time Deep Link Analytics. TigerGraph can traverse hundreds of million of vertices/edges per second per machine, traversing three to ten or more hops, all at orders of magnitude faster that traditional approaches. It also includes fast data update of 100+ K updates per second per server. TigerGraph is designed to combine both native graph storage and computation, supporting real-time graph updates and offering built-in parallel computation. For more on Native Parallel Graphs, see the white paper *Native Parallel Graphs: The Next Stage in the Graph Database Evolution* at www.tigergraph.com.

About TigerGraph

TigerGraph is the world's first Real-Time Graph Analytics Platform powered by Native Parallel Graph (NPG) technology. TigerGraph fulfills the true promise and benefits of the graph platform by supporting real-time deep link analytics for enterprises with complex and colossal amounts of data. TigerGraph's proven technology is used by customers including Alipay, VISA, SoftBank, State Grid Corporation of China, Wish and Elementum.

Founded by Yu Xu, Ph.D. in 2012, TigerGraph is funded by Qiming VC, Baidu, Ant Financial, AME Cloud, Morado Ventures, Zod Nazem, Danhua Capital and DCVC. TigerGraph is based in Redwood City, CA. Learn more at www.tigergraph.com.